Molecular dynamics simulation studies of salting effects and ion-mediated molecular interactions

Li, Weifeng

2013

Li, W. (2013). Molecular dynamics simulation studies of salting effects and ion-mediated molecular interactions. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/51118>

<https://doi.org/10.32657/10356/51118>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**MOLECULAR DYNAMICS SIMULATION STUDIES
OF SALTING EFFECTS AND ION-MEDIATED
MOLECULAR INTERACTIONS**

LI WEIFENG

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2013

**MOLECULAR DYNAMICS SIMULATION STUDIES
OF SALTING EFFECTS AND ION-MEDIATED
MOLECULAR INTERACTIONS**

LI WEIFENG

School of Physical and Mathematical Sciences

**A thesis submitted to the Nanyang Technological
University in partial fulfillment of the requirement for the
degree of Doctor of Philosophy**

2013

Molecular Dynamics Simulation Studies of Salting Effects and Ion-Mediated Molecular Interactions

by

LI WEIFENG

Abstract

In this thesis, a series of molecular dynamics simulations have been performed in order to explore some fundamental mechanisms in biological research, such as salting effects and ion-mediated molecular interactions. A general goal of this research is to obtain an atomic level understanding of the structure, dynamics and thermodynamics of co-solutes (including typical ions like Na^+ , Mg^{2+} and complex molecules like urea, guanidinium chloride) and their effects on molecular and electrostatic interactions in protein folding/unfolding and DNA packaging. More interest was paid to the mechanisms that cause protein denaturation in urea solution, the electrostatic and conformational effects that induce, stabilize and regulate DNA condensation in the presence of counterions. These aims are realized by using a combination of molecular dynamics simulation and advanced sampling techniques like umbrella sampling and thermodynamic integration method.

Thesis Supervisor: MU Yuguang

Title: Associate Professor, School of Biological Sciences

Thesis Supervisor: PHAN Anh Tuấn

Title: Associate Professor, School of Physical and Mathematical Sciences

Acknowledgments

First and foremost, I would like to express my deep and sincere gratitude to my supervisor, Assoc Prof. Mu Yuguang, not only for his patient supervision and valuable suggestions to this work, but also the great influence on me by his talent, serious attitude and insistent enthusiasm. High tribute shall be paid to Prof. Lars Nordenskiöld from the School of Biological Sciences, Prof. Zhou Ruhong from IBM Thomas J. Watson Research Center and Assoc Prof. PHAN Anh Tuấn from the School of Physical and Mathematical Sciences, for their instructive advices and useful suggestions on my thesis. Special thanks should go to all the other members in our group, Dr. Xu Weixin, Dr. Lin Xin, Dr. Jiang Ping, Mr. Yang Ye, Miss Han Nanyu, Mr. Zhang Tong and Mr. Wang Yaofeng. I am deeply grateful to their help in the completion of my thesis.

I am also deeply indebted to my parents and my wife, Ms. Yang Yanmei, for their continuous support and encouragement.

Finally, I would like to thank Nanyang Technological University for the financial support, the School of Physical and Mathematical Sciences and the School of Biological Sciences for offering me the academic opportunities during my PhD study.

Contents

1	Introduction	13
1.1	Introduction to Molecular Dynamics Simulation	14
1.1.1	The aims of molecular dynamics	14
1.1.2	Molecular dynamics simulation: the idea	15
1.1.3	Force field in molecular dynamics	16
1.1.4	Statistical mechanics for computer simulations	19
1.1.5	Ensemble average and time average	20
1.1.6	Ergodic hypothesis	21
1.1.7	Advanced sampling technique: thermodynamic integration	22
1.2	Hofmeister Series	23
1.3	Protein Denaturation	25
1.4	DNA Packaging	25
1.5	Thesis Organizations	27
2	Salting Effects in Sodium Chloride Solution	29
2.1	Computational Methods	30
2.1.1	Simulation details	30
2.1.2	Error estimation of ΔG	32
2.1.3	Kirkwood-Buff integral	32
2.2	Results and Discussions	33
2.2.1	Salting effects resolved by hydration free energy calculations	33
2.2.2	The role of ions in salting effects	35

2.2.3	The relation between the salting effects and hydration free energies	37
2.2.4	Interpretation of salting effects by Kirkwood-Buff integral	38
2.2.5	A schematic diagram of salting effects	40
2.3	Conclusion	42
3	Dissociation of Nanoparticles in Aqueous Guanidinium Chloride and Urea So-	
	lutions	48
3.1	Computational Methods	50
3.1.1	Nanoparticle model	50
3.1.2	Simulation details	51
3.1.3	Umbrella sampling method	52
3.2	Results and discussions	52
3.2.1	PMF between two NPs	52
3.2.2	Direct interaction of denaturant with NPs	55
3.2.3	Hardening of NP diffusion	59
3.3	Conclusion	61
4	Salting Effects of Protein Components in Aqueous NaCl and Urea Solutions	62
4.1	Computational Methods	64
4.1.1	Proteinic models	64
4.1.2	Simulation details	64
4.2	Results and Discussions	65
4.2.1	NaCl salts out all 15 molecules	67
4.2.2	Urea salts in all 15 molecules	68
4.2.3	Interpretation of salting effects by Kirkwood-Buff integral	71
4.2.4	Salting-effect is enthalpy driven	72
4.3	Conclusion	73
5	Sequence-Specific Mg^{2+} - DNA Interactions	84
5.1	Computational Methods	85
5.1.1	Simulation details	85

5.1.2	Covariance analysis and conformational entropy	87
5.2	Results and Discussions	88
5.2.1	Solvation structure of free ions	88
5.2.2	Results of isolated DNA model	89
5.2.3	Results of periodic DNA model	98
5.3	Conclusion	102
6	Mg²⁺-mediated DNA Interactions	103
6.1	Computational Methods	105
6.1.1	DNA models	105
6.1.2	Simulation details	105
6.1.3	Umbrella sampling method	106
6.2	Results	107
6.2.1	PMF between two DNA duplexes	107
6.2.2	DNA-DNA Binding pattern	108
6.2.3	Solvent accessible surface of B-DNA and Z-DNA	108
6.2.4	DNA de-hydration through contacting	110
6.2.5	Contact surface between two DNA duplexes	111
6.2.6	A two-stage association of Z-DNA binding	112
6.3	Discussion	113
6.3.1	The structural features of Z-DNA	113
6.3.2	Biological functions of Z-DNA	113
6.4	Conclusion	115
7	Conclusions and Future Plan	117
7.1	Conclusions	117
7.2	Future plan	119
7.2.1	Comparative studies of anions and cations	119
7.2.2	The role of ion's hydration shells	121
A	Publications	122

List of Figures

1-1	Highly simplified description of a molecular dynamics simulation algorithm.	15
1-2	Lennard-Jones pair potential and the r^{-12} and r^{-6} contributions.	17
1-3	Geometry of a simple chain molecule. Illustrations of the definitions of bond length r_{23} , bending angle θ_{234} , and torsion angle ϕ_{1234} are labeled. . .	18
1-4	Illustrations of (a) structure and (b) chemical structure of DNA. (b) is taken from http://en.wikipedia.org under the terms of the GNU Free Documentation License.	26
2-1	(a) Solvation free energy (ΔG) of solute molecules in water; (b) Change of solvation free energy ($\Delta\Delta G$) from pure water to 2M NaCl solution.	35
2-2	Radial distribution function (RDF) of Na^+ and Cl^- around (a) the center solute molecule in region A ($k_q = 1.5$, $k_\sigma = 0.5$) and (b) the center molecule in region B ($k_q = 1.5$, $k_\sigma = 1.1$) in Figure 2-1. Spatial distributions of Na^+ (bluish violet cross, density > 0.1 ions/ $\text{\AA}^3$), Cl^- (yellow cross, density > 0.1 ions/ $\text{\AA}^3$) were shown as inserts. Solute molecule was shown as red and white balls. The solute itself has four strongly bound water molecules (shown in cyan sticks) in a tetrahedron configuration.	36
2-3	The dependence of hydration free energy change ($\Delta\Delta G$) and hydration free energy in pure water (ΔG_{pure}). The dot and dot-dash lines indicate the ΔG of Na^+/Cl^- and water, respectively. Experimental values of methane and ethane (green hollow circles) at 25°C were also shown. The solute molecules in region A and region B in Figure 2-1 are highlighted in blue and red, respectively.	37

2-4	(a) Pearson's correlation between 121ν and $\Delta\Delta G$ with respect to cutoff of integral distance. (b) The convergence of differential Kirkwood-Buff integral (ΔKBI) of three representative solute molecules.	38
2-5	(a) The relationship between salting effect ($\Delta\Delta G$) and differential Kirkwood-Buff integral (ΔKBI) of NaCl (ΔK_{sc}) and water (ΔK_{sw}), (b) 2D map of ΔK_{sc} , and (c) 2D map of ΔK_{sw}	39
2-6	Schematic diagram of salting effects. The origin represents the hydration free energy of a water molecule. The x-axis and y-axis represent the hydration free energy of the ions and the solute, respectively. NS stands for no clear salting effect.	41
3-1	Snapshots of the simulation box. Water and chloride ions are not shown in (b) and (c).	50
3-2	The potential of mean force (PMF) of two (a) no-charge nano-particles (NPs) and (b) charged NPs with respect to their center of mass distance. The direct contact pair (DCP) states in all cases are considered having free energy zero.	52
3-3	Cumulative number of (a) Gdm+ and (b) urea around three kinds of nano-particles (NPs).	55
3-4	Binding orientations of (a) Gdm+ and (b) urea around three kinds of nano-particles (NPs).	56
3-5	(a) Snapshot of two no-charge nano-particles (NPs) in 4M GdmCl. The intruded Gdm+ is red-colored and Gdm+ that form stacked dimer and tetramer are golden-colored. (b) Cumulative charge around NPs in 4M GdmCl. . . .	57
3-6	A snapshot of nearest 65 ureas around positive-charged nano-particle (NP+) and negative-charged nano-particle (NP-). The center of mass distance of two nano-particles is 1.8 nm.	58
3-7	Mean-square displacement of nano-particle dimers.	60

4-1	(a) Solvation free energy changes ($\Delta\Delta G_{NaCl}$) of 15 molecules from pure water to NaCl solution. Three different combinations of force fields were tested: AMBER derived molecule charges with AMBER NaCl (black), RED tool generated molecule charges with AMBER NaCl (red) and RED tool generated molecule charges with new-NaCl model (blue). (b) Solvation free energy changes ($\Delta\Delta G_{urea}$) from pure water to aqueous urea solution. AMBER derived molecule charges (black) and RED tool generated molecule charges (red) were tested.	67
4-2	The linear relationship between the solvation free energy change in aqueous urea solution and the number of heavy atoms of the solute molecule.	68
4-3	Radial distribution function (RDF) and cumulative RDF of urea around (a) NMA and (b) TRP. Cluster size of the urea binding patterns of (c) NMA and (d) TRP.	69
4-4	(a) Binding orientations of urea molecules (within 0.70 nm) around the NMA and TRP molecules. θ is the angle between the vector from the center of mass of NMA or TRP to the center of mass of urea molecule and the normal vector of the urea plane. The representative binding patterns of urea around NMA and TRP are shown in (b) and (c), respectively. The structure snapshots were obtained from the top 25 highest populated clusters, where the clustering method is introduced in the main text.	70
4-5	(a) Correlation of differential Kirkwood-Buff integral with the values of $\Delta\Delta G$. (b) and (c) are radial distribution functions of Na^+ , Cl^- , urea and water around a TRP molecule.	71
4-6	Enthalpic (black) and entropic (red) contributions to solvation free energy changes from pure water to (a) NaCl solution and (b) urea solution. The order of the protein components is $ \Delta\Delta G $ increasing (weak salting-effect to strong salting-effect).	72
4-7	Decomposition of the solute-solvent interaction in urea solvent into the van der Waals (LJ) and the electrostatic terms.	74

5-1	Radial distribution function (solid line) of water oxygen and cumulative number of water (dashed line) around Na^+ (black colored) and Mg^{2+} (red colored) ions. The inserts are (a) snapshot of the spatial configuration of six water molecules bound to Mg^{2+} and (b) distribution of ion-DNA distance.	89
5-2	Time evolution of number of cations on DNA surface: (a) Na^+ ions in Na-DNA system, (b) Na^+ (black line) and Mg^{2+} ions (red line) in Mg-DNA system. (c) Time evolution of root mean squared deviation (RMSD) of DNA in two simulations with respect to a reference structure of a standard B-DNA; (d) The number of Watson-Crick hydrogen bonds (H-bond) in two simulations.	90
5-3	Stereo views of counterions binding map around DNA duplex. Left side represents the Na-DNA system: (a) Na^+ in minor groove, (b) Na^+ in major groove; Right side represents the Mg-DNA system: (c) Mg^{2+} in major groove, (d) Mg^{2+} on phosphate group at A-tract region.	92
5-4	The number of bound ions projected on each base pair. The x axis indicates base pair sequence number (a-d). Base pair atoms used to define bound ions in the major groove and minor groove are shown in the inset of (d).	94
5-5	Width of (a) major groove and (b) minor groove of isolated DNA model. (c) Local bending angles of hex-basepair segments. A monitoring window containing six-basepair slides along basepair 2 to 22 of the isolated DNA duplex, resulting in 16 overlapping hex-basepair segments.	95
5-6	(a) Eigenvalues of first 10 eigenvectors from principal component analysis (PCA) analysis, the inset is the inner-product of the first 5 eigenvectors from Mg-DNA (x-axis) and Na-DNA (y-axis) system; (b) Time evolution of the conformational entropy.	97
5-7	The basepair-averaged number of (a) Na^+ and (b) Mg^{2+} located at major groove, minor groove and backbone of periodic DNA. The overall charge neutralization fraction was also shown accordingly. Width of (c) major groove and (d) minor groove of periodic DNA in Na^+ and Mg^{2+} electrolytes.	99

5-8	Time evolution of conformational entropy of four periodic DNA in Na ⁺ and Mg ²⁺ electrolytes.	101
6-1	Illustration of the cross section of the simulation box and how the two DNAs are connected by an umbrella sampling spring through centers of mass (shown as red balls).	106
6-2	Potentials of mean force (PMF) with respect to inter-DNA distance between two (a) B-DNAs and (b) Z-DNAs. Free energy values at 3.5 nm are set to 0.	107
6-3	Distributions of the relative rotational angles of two parallel DNAs at stable binding state: (a) AT-DNA at 2.4 nm; (b) GC-DNA at 2.4 nm; (c) GC-DNA at 2.1 nm and (d) GC-ZDNA at 1.8 nm.	109
6-4	Snapshots of two DNAs at stable binding states: (a) AT-DNA at 2.4 nm, (b) GC-DNA at 2.4 nm; (c) GC-DNA at 2.1 nm and (d) GC-ZDNA at 1.8 nm. In the upper panel, DNA is depicted by a stick diagram while Mg ²⁺ ions which directly bound to DNA are shown as magentas spheres. In the lower panel, DNA molecules are represented by a surface diagram. The color scheme is: green for hydrophobic (absolute value of atom charge is less than 0.2), red for negative hydrophilic (atom charge is less than -0.2), blue for positive hydrophilic (atom charge is larger than 0.2).	110
6-5	The Potentials of mean force (PMF), number of hydration water in the hydration shell and contact surface area of (a, b, c) AT-DNA, (d, e, f) GC-DNA and (g, h, i) GC-ZDNA with respect to inter-DNA distance.	111
6-6	Radial distribution functions (RDFs) of sugar (C1' – C5') and phosphate (O1P, O2P) atoms of GC-ZDNA, around center of mass of DNA in the xy-plane.	113
7-1	Hofmeister series of ions. The order is given in terms of the ability of ions to stabilize the structure of proteins from most stabilizing to most destabilizing (<i>citrate</i> ³⁻ refers to $C_3H_5O(COO)_3^{3-}$, <i>sulfate</i> ²⁻ refers to SO_4^{2-} and <i>phosphate</i> ²⁻ refers to HPO_4^{2-}).	120

List of Tables

2.1	Force field parameters used in our simulations	31
2.2	Solvation free energies (in kcal/mol) in pure water and NaCl electrolyte of 121 solute molecules	43
3.1	Simulation box contents	51
3.2	Activation energy and energy penalty (kcal/mol) at solvent separated pair (SSP) and 1.8 nm	54
4.1	Free energies (in kcal/mol) in perturbations of van der Waals (LJ) interac- tions using 21 and 53 λ . The absolute difference and average value are shown for comparison	66
4.2	Model molecules used in this study and solvation free energies (in kcal/mol) in pure water in this study and the literatures	66
4.3	List of AMBER derived charges and RED tool generated charges and ab- solute difference	75
5.1	Simulation system content	86
5.2	CHARMM force field parameters of Na^+ and Mg^{2+} used in our simulations	87
6.1	Atoms with surface area $> 0.1 \text{ nm}^2$, the surface area values are shown in the brackets	116

Chapter 1

Introduction

Molecular dynamics (MD) simulation is one of the computational tools to study the time dependent movements/behaviors of a series of atoms and molecules. The idea of the MD method was first introduced by Alder and Wainwright in 1957¹ to study the interactions between hard spheres, and it was first applied to a scientific research by Rahman in 1964² on the thermodynamical properties of liquid argon. Boosted by the rapid development of computation technology in the recent decades, MD is becoming a valuable tool in theoretical studies of many disciplines ranging from materials science to the modeling of biological systems. The advantage of MD simulations is that it allows us to examine and monitor *what is happening* at the atomic level, rather than measuring the *bulk properties* that can only be provided by most experimental studies.

The goal of this thesis is to apply MD simulations to exploring the underlying mechanism of biology-related phenomena, especially to investigating the structure, dynamics and thermodynamics of co-solutes (including typical ions like Na^+ , Mg^{2+} and complex molecules like urea, guanidinium chloride) and their effects on the solvation of biological molecules. The research topics mainly include salting effects and ion-mediated molecular interactions, as well as denaturation of proteins in urea solution and DNA attraction in electrolyte with divalent Mg^{2+} ions.

1.1 Introduction to Molecular Dynamics Simulation

1.1.1 The aims of molecular dynamics

Experiment cannot handle everything. Limited by the current scope of technologies, it is usually difficult to monitor the interactions between molecules at microscopic time and length. An example is the self-diffusion coefficient, $\mathcal{D}$, which governs the evolution of concentration, c (or number density). According to a generalized transport equation:

$$\frac{\partial c}{\partial t} = \mathcal{D} \nabla^2 c \quad (1.1)$$

$\mathcal{D}$ is easy to measure experimentally. Alternatively, we can compute $\mathcal{D}$ through a MD simulation according to the Einstein-Smoluchowsky relation. For a three dimensional system:

$$\frac{\partial \langle r^2 \rangle}{\partial t} = 6\mathcal{D} \quad (1.2)$$

$\langle r^2 \rangle$ is known as mean-squared displacement, MSD. At long time, $\mathcal{D}$ should be independent of time; hence

$$\mathcal{D} = \frac{1}{6} \lim_{t \rightarrow \infty} \frac{d}{dt} (MSD) \quad (1.3)$$

We can calculate the MSD using MD trajectory, and then estimate $\mathcal{D}$ ¹.

Computer simulations act as a bridge between macroscopic world of the experiments and microscopic length and time scales of constituent particles. Through simulations, the hidden details behind bulk properties can be revealed; the “exact” interactions at atomic level can be found. Moreover, we may test a theory by conducting a simulation using the same model, and comparing simulation results with experimental data². We may also carry

¹For a MD trajectory, MSD is the average distance that particles travel in a certain amount of time. The MSD is computed as:

$$MSD = \frac{1}{N} \sum_{i=1}^N \left(r_i(t) - r_i(0) \right)^2 \quad (1.4)$$

Where N is the number of particles, t is time interval, and $r_i(t) - r_i(0)$ is the vector distance traveled by a given particle, i , over the time t . An example is given in Chapter 3

²Such as the Ising model in magnetic theory. We run Monte Carlo simulations implementing 2D Ising

Figure 1-1 Highly simplified description of a molecular dynamics simulation algorithm.

out simulations at extreme conditions (for example, at extremes of temperature or pressure) which are difficult or impossible to be fulfilled in a laboratory.

In the field of computational biology, there are two main families of molecular simulation techniques, molecular dynamics (MD) simulation and Monte Carlo (MC) simulation. In this thesis, all the simulations are MD simulations. Hence in this section we shall only concentrate on the principles of molecular dynamics simulation.

1.1.2 Molecular dynamics simulation: the idea

Molecular dynamics simulation is a computer simulation of physical movements of a series of atoms and molecules. The movements of these objects are predicted by the numerical solution of the Newton's equations of motion:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i \quad (1.5)$$

$$\mathbf{F}_i = -\nabla_i U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (1.6)$$

In an ideal case, we hope to “monitor” the forces, velocities, positions of atoms at all time, which guarantee that the simulation trajectory is a demonstration of the real world. Nevertheless, this is not suitable for a numerical program. In order to be implemented on model to predict the Curie temperature in graphite-based magnetic materials³

a computer with acceptable accuracy and reasonable computing load, the existing methods adopt step-by-step numerical integration. A simple description of the system-evolution algorithm in a molecular dynamics simulation is shown in Figure 1-1. The simulation proceeds by iteratively calculating forces and solving the equations of motion based on the accelerations induced by the new forces. The most commonly used algorithm when updating the position and velocity is the Velocity Verlet algorithm⁴:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 \quad (1.7)$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{\mathbf{a}(t) + \mathbf{a}(t + \Delta t)}{2}\Delta t \quad (1.8)$$

In practise, almost all MD codes use much more complicated versions of the algorithm, including two steps (predictor and corrector) in solving the equations of motion and many additional subroutines for temperature and pressure control, analysis and output. The details are out of the scope of this thesis, and can be easily found elsewhere^{5,6}.

1.1.3 Force field in molecular dynamics

The forces $\mathbf{F}_i$ acting on atoms are derived from the potential energy, U , of the system. In MD simulations, U is computed from the Hamiltonian function which is also named **force field** in the field of MD. U is a function of all the N atoms' coordinates, i.e., $U = U(\mathbf{r}^N)$, where $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$. For simplicity, a force field is an atomic description of interactions. For computer simulations in the area of biological system, the force field is usually formulated and classified into two types:

Type 1: Non-bonded Interactions

In computer simulations, an atom is usually described as a van der Waals (vdW, named after Dutch scientist *Johannes Diderik van der Waals*) ball with a point charge. The properties of a vdW ball are the sum of the attractive and repulsive forces between atoms (or between molecules). VdW interactions are relatively weak compared to normal chemical bonds, but they play a fundamental role because they determine the radius of atoms.

Figure 1-2 Lennard-Jones pair potential and the r^{-12} and r^{-6} contributions.

Meanwhile, the point charge is a simplification of the electrostatic interaction.

These two types of non-bonded interactions are expressed as differentiable pair-potentials. The Lennard-Jones potential (also referred to as the L-J potential, 6-12 potential) is the most commonly used form for vdW interactions:

$$U^{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (1.9)$$

The two parameters: σ is the radius, and ϵ is the attractive well depth. The curves of a L-J potential and its components are illustrated in Figure 1-2. This simple potential was first used, for instance, in the earliest computer simulation studies on the thermodynamic properties of liquid argon^{2,7}.

If electrostatic interactions are present, the Coulomb potential is used

$$U^{Coulomb}(r) = \frac{Q_1 Q_2}{4\pi\epsilon_0 r} \quad (1.10)$$

with three parameters: Q_1 and Q_2 are the atomic charges of two interacting atoms, and ϵ_0 is the permittivity of the free space. The charges of atoms inside a molecule are usually partial. The charge values may be estimated by the electron density distribution of the

Figure 1-3 Geometry of a simple chain molecule. Illustrations of the definitions of bond length r_{23} , bending angle θ_{234} , and torsion angle ϕ_{1234} are labeled.

molecule which can be predicted by a quantum-mechanics (QM) calculation.

Type 2: Bonded Potentials

Besides vdW and Coulomb interactions, for molecules the intramolecular bonding interactions (covalent bonds) need to be considered. A framework of a simple chain-like molecule consisting of 7 atoms is shown in Figure 1-3. Three kinds of interatomic topologies are labeled: the “bond” between adjacent atom 2 and 3 with separation r_{23} ; the “bending angle” between successive bond vectors $\mathbf{r}_{23}$ and $\mathbf{r}_{34}$, θ_{234} ; the “torsion angle” in terms of two planes determined by atom(1,2,3) and atom(2,3,4), ϕ_{1234} .

The first two potentials are usually expressed in harmonic forms with the distance/angular displacements from the equilibrium value:

$$U_{intramolecular}^{bond}(r_{ij}) = \frac{1}{2}k_{ij}^r (r_{ij} - r_{eq})^2 \quad (1.11)$$

$$U_{intramolecular}^{bending-angle}(\theta_{ijk}) = \frac{1}{2}k_{ijk}^{\theta} (\theta_{ijk} - \theta_{eq})^2 \quad (1.12)$$

There are four parameters: r_{eq} and θ_{eq} are the equilibrium values of bond length and bending angle; k_{ij}^r and k_{ijk}^{θ} are the spring constants, respectively. The values of these four parameters are usually determined through QM calculations of the same topology and verified

by experimental results.

Principally, the torsional potential is periodic when two normal vectors, $\hat{n}_{123}$ and $\hat{n}_{234}$, rotate in a circle. Hence, the formulation of the torsional potential involves an expansion in periodic functions:

$$U_{intramolecular}^{torsion-angle}(\phi_{ijkl}) = \frac{1}{2} \sum_m k_{ijkl}^{\phi,m} (1 + \cos(m\phi_{ijkl} - \gamma_m)) \quad (1.13)$$

where m is the periodic order, $m = 1, 2, \dots$ $k_{ijkl}^{\phi,m}$ and γ_m can also be determined through QM calculations.

1.1.4 Statistical mechanics for computer simulations

By conducting a simulation, we obtain a trajectory in which the time-evolution of the instantaneous positions and velocities of all molecules in the system are recorded. However, this kind of information cannot be compared to experimental data, because no experiments can provide us with such detailed information. In order to use computer simulation as the numerical counterpart of experiments, we must analyze the trajectory and calculate *averaged* properties, like self-diffusion coefficient as we have discussed in Section 1.1.1. Here, the word *averaged* is a concept of statistical mechanics, which means that the parameter that we are interested is averaged over a large number of particles and, usually, also averaged over the time of the simulation. In order to explain what kind of computer simulation analysis can be compared with experimental measurements, we have to introduce some basic concepts in statistical mechanics such as ensemble, probability density, partition function, etc. We begin with the definition of ensemble:

Ensemble: infinite collection of identical (independent) copies of the system, each in one of the microscopic states is compatible with the macroscopic state

For a system at temperature T (a canonical ensemble), the *probability density of the ensemble* follows the Boltamann distribution:

$$\rho(\mathbf{p}^N, \mathbf{r}^N) = \frac{1}{Q} \exp \left[\frac{-H(\mathbf{p}^N, \mathbf{r}^N)}{k_B T} \right] \quad (1.14)$$

where $H(\mathbf{p}^N, \mathbf{r}^N)$ is classical Hamiltonian, T is system temperature, k_B is Boltzmann's constant, $\mathbf{p}^N$ and $\mathbf{r}^N$ are momenta and coordinates of the N atoms, respectively. The parameter Q in the denominator is the *Partition function* of the system which can be treated as a constant³.

$$Q = \frac{1}{N!h^{3N}} \iint d\mathbf{p}^N d\mathbf{r}^N \exp \left[\frac{-H(\mathbf{p}^N, \mathbf{r}^N)}{k_B T} \right] \quad (1.15)$$

1.1.5 Ensemble average and time average

To be specific, let us consider certain observable A (like total energy, pressure, density, etc.). From the Boltzmann distribution, we know the probability that a system at temperature T with an energy $H(\mathbf{p}^N, \mathbf{r}^N)$ can be found and we can therefore compute the *ensemble average* (expectation value) of A as

$$\begin{aligned} \langle A \rangle_{ensemble} &= \frac{\iint d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \exp \left[\frac{-H(\mathbf{p}^N, \mathbf{r}^N)}{k_B T} \right]}{\iint d\mathbf{p}^N d\mathbf{r}^N \exp \left[\frac{-H(\mathbf{p}^N, \mathbf{r}^N)}{k_B T} \right]} \\ &= \frac{1}{N!h^{3N}} \iint d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \frac{1}{Q} \exp \left[\frac{-H(\mathbf{p}^N, \mathbf{r}^N)}{k_B T} \right] \\ &= \frac{1}{N!h^{3N}} \iint d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \rho(\mathbf{p}^N, \mathbf{r}^N) \end{aligned} \quad (1.16)$$

where $A(\mathbf{p}^N, \mathbf{r}^N)$ is the observable of interest, and $\rho(\mathbf{p}^N, \mathbf{r}^N)$ is probability density function. We can see that the ensemble average integrates over all of the momentum and position spaces of the system.

Thus far, we have discussed the average behavior of a many body system in a purely statistical mechanics point of view. However, this is not the way that we “measure” some properties from our simulation data. In fact, the idea behind Molecular Dynamics simulations is precisely that we can study the average behavior of a many-particle system simply by computing the time evolution of the system numerically and averaging the quantity of interest over a sufficiently long time. For example, the measurable physical quantity, A , are determined as *time average*

³All the following mathematics are formulated for a gas of N identical classical particles. The reason for the $N!$ factor is to ensure that we do not “over-count” the number of microstates, because the N particles are identical and indistinguishable.

$$\langle A \rangle_{time} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt A(\mathbf{p}^N(t), \mathbf{r}^N(t)) \quad (1.17)$$

where $A(\mathbf{p}^N(t), \mathbf{r}^N(t))$ is the instantaneous value of observable, A , with respect to momenta $\mathbf{p}^N$ and coordinates $\mathbf{r}^N$ at time t .

1.1.6 Ergodic hypothesis

In writing down the equation 1.17, we have assumed that, for sufficiently long simulation time T , the time average does not depend on the initial conditions. This is, in fact, a subtle assumption that is not true in general⁸. However, we shall disregard subtleties and simply assume that, once we have specified N , V and E , time-averages do not depend on the initial coordinates and momenta. If that is true, then we would not change our result for $\langle A \rangle_{time}$ if we average over many different initial conditions, *i.e.* we consider the hypothetical situation where we run a large number of Molecular Dynamics simulations at the same values for N , V and E , but with different initial coordinates and momenta,

$$\langle A \rangle_{time} = \frac{\sum_{\text{initial conditions}} \left(\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt A(\mathbf{p}^N(0), \mathbf{r}^N(0), t) \right)}{\text{number of initial conditions}} \quad (1.18)$$

We now consider the limiting case where we average over all initial conditions compatible with the imposed values of N , V and E . In that case, we can replace the sum over initial conditions by an integral

$$\frac{\sum_{\text{initial conditions}} f(\mathbf{p}^N(0), \mathbf{r}^N(0))}{\text{number of initial conditions}} \rightarrow \frac{\iint_E d\mathbf{p}^N d\mathbf{r}^N f(\mathbf{p}^N, \mathbf{r}^N)}{Q(N, V, E)} \quad (1.19)$$

where Q is partition function. The subscript E on the integrals indicates that the integration is restricted to a space of constant energy E . If we switch the order of the time-averaging and the averaging over initial conditions, we find

$$\langle A \rangle_{time} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau dt \langle A(\mathbf{p}^N(0), \mathbf{r}^N(0), t) \rangle_{ensemble} \quad (1.20)$$

It is clear that the ensemble average $\langle A(\mathbf{p}^N(0), \mathbf{r}^N(0), t) \rangle_{ensemble}$ in the above equation

does not depend on the measuring time t . In other words, averaging over all initial phase-space coordinates is equivalent to averaging over the time-evolved phase-space coordinates:

$$\langle A \rangle_{time} = \langle A \rangle_{ensemble} \quad (1.21)$$

which is also known as *ergodic hypothesis*. The above equation states that, if we wish to compute the average of a function of the coordinates and momenta of a many-particle system, we can either compute that quantity by time-averaging (the ‘Molecular Dynamics’ approach), or by ensemble-averaging (the ‘Monte Carlo’ approach).

1.1.7 Advanced sampling technique: thermodynamic integration

Thermodynamic integration is a method used to compare the difference in free energy between two given states (*e.g.*, A and B) via sampling state configurations in a molecular dynamics or Metropolis Monte Carlo simulations. Because the free energy of a system is not simply a function of the phase space coordinates of the system, but is instead related to the canonical partition function $Q(N, V, T)$, the free energy difference between two states cannot be calculated directly. In the thermodynamic integration approach, the free energy difference is calculated by defining a thermodynamic path between the states and integrating over enthalpy changes along the path. Such paths can either be real chemical processes or alchemical processes. A good example of the alchemical process is the Kirkwood’s coupling parameter method⁹. The derivation of Kirkwood’s coupling parameter method is summarized as follows:

Consider two systems, A and B, with potential energies U_A and U_B . The potential energy in either system can be calculated as an ensemble average over configurations sampled from a molecular dynamics or Monte Carlo simulation with proper Boltzmann weighting. Now consider a new potential energy function defined as:

$$U(\lambda) = U_A + \lambda(U_B - U_A) \quad (1.22)$$

Here, λ is defined as a coupling parameter with a value between 0 and 1, and thus the

potential energy as a function of λ varies from the energy of system A for $\lambda = 0$ and system B for $\lambda = 1$. In the canonical ensemble, the partition function of the system can be written as:

$$Q(N, V, T, \lambda) = \sum_s \exp \left[\frac{-U_s(\lambda)}{k_B T} \right] \quad (1.23)$$

In this notation, $U_s(\lambda)$ is the potential energy of state s in the ensemble with potential energy function $U(\lambda)$ as defined above. The free energy of this system is defined as:

$$F(N, V, T, \lambda) = -k_B T \ln Q(N, V, T, \lambda) \quad (1.24)$$

If we take the derivative of F with respect to λ , we will get that it equals the ensemble average of the derivative of potential energy with respect to λ .

$$\begin{aligned} \Delta F(A \rightarrow B) &= \int_0^1 d\lambda \frac{\partial F(\lambda)}{\partial \lambda} = - \int_0^1 d\lambda \frac{k_B T}{Q} \frac{\partial Q}{\partial \lambda} \\ &= \int_0^1 d\lambda \frac{k_B T}{Q} \sum_s \frac{1}{k_B T} \exp \left[\frac{-U_s(\lambda)}{k_B T} \right] \frac{\partial U(\lambda)}{\partial \lambda} = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \end{aligned} \quad (1.25)$$

The change in free energy between states A and B can thus be computed from the integral of the ensemble average of the change in potential energy with the coupling parameter λ ⁶. In practice, this calculation is performed by first defining a potential energy function $U(\lambda)$, sampling a series of equilibrium configurations for different values of λ , finding the ensemble average of the derivative of the system energy with respect to λ from the configurations in these separate equilibrium runs, then finally numerically computing the integral over all of these derivatives.

1.2 Hofmeister Series

Common cations and anions have effectiveness on the solubility of other molecules in the aqueous solution. These effects are ion-specific, either salting-out (decrease the solubility) or salting-in (increase the solubility) other molecules. This change of solubility is usually called Hofmeister effect because it is first systematically investigated by *Franz Hofmeister*

(Professor of Pharmacology at the Charles University in Prague) and his colleagues during 1880s and 1890s. By a series of milestone experiments, *Prof. Hofmeister* proposed a series of ions that have consistent effects when precipitating proteins from the egg white, which is usually called *Hofmeister series* or *lyotropic series*. Based on the ability to salt out proteins¹⁰, cations are usually ordered as:

while for anions:

Former members of the series decrease the solubility of proteins (“salting out” effect); by contrast, the latter members in the series increase the solubility of proteins (“salting in” effect). Later, it was found that apart from protein solubility, ions have effects on a wide variety of phenomena, like protein folding and crystallization, enzymatic activity and bacterial growth, which also follow the order in Hofmeister series.

To date, the mechanism of the Hofmeister effect is not entirely clear. Several reasons like changes in water structure, surface tension and osmotic pressure have been proposed^{11–20}.

It is found that specific interactions (direct interactions) between ions and proteins may be more important for the latter members (such as guanidinium⁺, I[−], ClO₄[−] and SCN[−]) in Hofmeister series that have a strong “salting in” effect. These ions are also known as protein denaturants^{21,22}. They salt in the peptide group, and thus interact much more strongly with the unfolded form of a protein than with its native form. Consequently, they shift the chemical equilibrium of the folding/unfolding reaction towards the unfolded state, causing **protein denaturation**.

1.3 Protein Denaturation

Proteins are polymers of amino acids. Once a protein is created by RNA ribosomes during gene translation, it begins to fold (sometimes spontaneously and sometimes with enzymatic assistance). It curls up on itself so that hydrophobic amino acids are buried inside the structure and hydrophilic/polar amino acids end up on the surface. The final shape of a protein determines its function and how it interacts with the environment. There are four types of interactions maintaining the shape (secondary and tertiary structure) of a protein, which includes: hydrogen bonding, salt bridges, disulfide bonds, and hydrophobic contacts.

Protein denaturation is a process in which a protein loses its biological functions caused by the changes of external environment, such as applications of some external compounds like a strong acid, a concentrated inorganic salt (e.g., the latter members in the Hofmeister series), an organic solvent (e.g., alcohol and urea), or heat. In such denaturing environments, these four stabilizing interactions will be disrupted, resulting in a loss of the tertiary and secondary structures that are present in the native state.

A classic example of protein denaturation comes from egg whites, which are mostly egg albumins in water. The fresh egg whites are transparent and in a liquid form. Cooking turns them opaque, solid, and insoluble. In such a case, the denaturation results in loss of solubility. On the contrary, adding of ethanol, urea, Guanidinium Chloride, etc., solubilizes proteins in the solution, thereby denatures the proteins. In such cases, the protein denaturation is mainly caused by structure unfolding.

1.4 DNA Packaging

Deoxyribonucleic acid (DNA) is the hereditary material in almost all organisms (with the exception of RNA viruses)²³. Structurally, DNA is made of repeating units called nucleotides as shown in Figure 1-4. The nucleotide unit contains a nucleobase, a phosphate group and a 2-deoxyribose sugar. There are four types of nucleobase, which are adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T). The coded genomic DNA sequence specifies the sequence of the amino acids in proteins. The phosphate and sugar

Figure 1-4 Illustrations of (a) structure and (b) chemical structure of DNA. (b) is taken from <http://en.wikipedia.org> under the terms of the GNU Free Documentation License.

rings constitute the backbone of the DNA strand, which is negative-charged. DNA usually exists in duplex strands. The nucleobase from one strand interacts with another nucleobase from the other strand: A forms hydrogen bonding only to T, and C bonds only to G, which is called *complementary base pairing*. The DNA duplex stores the hereditary information through the 4-letter sequence²⁴.

Except from DNA transition, meters of genomic DNA are packed by nature into compact structures for all living beings in most of time. This compact structure protects DNA from damage induced by oxidizing agents, alkylating agents and also high-energy electromagnetic radiation such as ultraviolet light and X-rays²⁵⁻²⁷.

The DNA-DNA interaction is electrostatically repulsive without counterions. To condense DNA, the attractive force must outweigh the repulsive force. In chromosomes, this is done by specialized proteins. In phage heads and sperm, DNA is packaged by simple counterions^{28,29}.

1.5 Thesis Organizations

The organizations of this thesis are as follows:

Chapter one is the introduction to this thesis, and it provides a summary of the background knowledge that includes the basis of molecular dynamics simulation, Hofmeister series, protein denaturation, and DNA packaging.

Chapter two takes advantage of the thermodynamic integration method to calculate the solvation free energies of a series of model solute molecules in pure water and 2M NaCl electrolyte, respectively. Through the analysis, we proposed that the salting effect depends not only on the ion's position in Hofmeister series, but also on the solute's specifics. The salting-in and salting-out behaviors are a result of delicate interactions among solute, ions and water molecules. Taking the solvation free energies of solute and ion as independent variables, a schematic diagram of salting effects in NaCl is suggested.

Chapter three discusses the dissociation of nano-particles (NPs) in aqueous guanidinium chloride (GdmCl) and urea solutions. The aim of this work is to answer the question why GdmCl is a stronger protein denaturant than urea. The dissociation behavior of two NPs were reflected through the potential of mean force with respect to their center of mass distance, calculated through umbrella sampling simulations. Through such a NP model, GdmCl shows stronger dissociation power than urea. A series of possible reasons are analyzed.

Chapter four also focuses on the mechanism of protein denaturation. The basic idea of this chapter is that, protein unfolding in urea solution is due to a change of better-solubility of amino acids from pure water to urea solution. NaCl electrolyte is taken into our consideration as a control solution. In contrast to Chapter three where NPs were used as model solute, we have employed more realistic molecules: 14 amino acid side chain analogues, and *N*-Methylacetamide (NMA) which mimics the protein backbone. The solvation free energies of the 15 molecules were calculated in pure water, aqueous urea, and NaCl solutions. Our results show that NaCl displays a strong capability to salt out all 15 molecules, while urea facilitates the solvation (salting-in) of all the 15 molecules which can be extrapolated to initiate the protein unfolding process.

Chapter five reports the MD simulation results of ion-DNA interactions. The binding behaviors of Mg^{2+} and Na^+ to isolate DNA were studied. Moreover, the ionic effects on DNA structure and conformational dynamics were compared between Mg^{2+} and Na^+ .

Chapter six provides insights into Mg^{2+} -mediated DNA interactions. A hexagonal simulation cell comprised of two identical DNA duplexes with periodic boundary condition along the DNA helix was used. The simulation setup mimics the lateral interactions in parallel-oriented DNAs. The effects from both DNA sequence and conformation were taken into consideration. Our findings that Z-DNA attraction is much more stronger than those of B-DNA can explain the experimental observations that Z-form DNA segments can induce plasmids condensation. The more hydrophobic nature of Z-form DNA is highlighted and may be important in its biological functions.

Chapter seven suggests possible directions for future research.

Chapter 2

Salting Effects in Sodium Chloride Solution

As the beginning of this thesis, this chapter focuses on the most-fundamental theory of salting-effects, the well known Hofmeister Effect, which was first observed in 1888 when precipitating proteins from egg whites by adding salts^{30,31}. It describes ions' ability to salt out or salt in proteins. The typical order based on the ability of ions to salt out proteins, is called *Hofmeister series* and shown here, for cations: $NH_4^+ > K^+ > Na^+ > Li^+ > Ca^{2+} > Mg^{2+} > guanidinium^+$; and for anions: $F^- \approx SO_4^{2-} > HPO_4^{2-} > Cl^- > Br^- > I^- > ClO_4^- > SCN^-$ ¹⁰. Apart from protein solubility, a wide variety of phenomena, from protein folding and crystallization, enzymatic activity and bacterial growth, has been shown to follow the Hofmeister series^{22,32,33}.

An early explanation of the mechanism of the Hofmeister effect is that ions alter the water structure, like breaking the hydrogen-bonded network of bulk water^{16,18}. However, recent experiments show that ions have no influence on the dynamics of water molecules outside the first solvation shell and the change of bulk water structuring is very small and negligible^{34,35}. Another explanation is the so called “cavity model”¹⁷. When proteins or macromolecules are introduced in water, their existences will make cavity. The ions at the left side of the series (called kosmotropes) can increase surface tension which tends to make the “cavity” smaller and thus salt proteins out. Those at the right side (called chaotropes)

can decrease surface tension which tends to make the “cavity” larger and thus salt proteins in. The cavity model has clear limitations. It does explain the salting-out effect of nonpolar solute molecules which can be approximately treated as “cavities” in water. However it ignores interactions between water and solute molecules and fails in accounting for hydrophilic molecules or charged/polar amino acids due to the strong Coulomb interactions, such as hydrogen bonds formed between solute molecules and water.

A recent molecular dynamics simulation study provided detailed pictures of the salt-induced modification of the attractive interaction between two hydrophobic plates in aqueous salt solutions³⁶. The results showed that the strength of the hydrophobic interaction is amplified, similar to the salting-out phenomenon, by high charge density ions. On the other hand, preferential binding of low charge density ions causes an attenuation of the hydrophobic interaction, likely salting-in³⁶.

Most biologically relevant molecules are, however, not purely hydrophobic. How the solvation properties of these molecules are affected by salts is still not completely clear. In this chapter we have performed computer simulations to study the hydration of a series of solute molecules in pure water and 2M NaCl solution. We find that the change of hydration free energy is dominated by the cosolvent’s preferential binding to the solute molecule. Accumulation of Na^+ and/or Cl^- ions around the solute molecule will cause salting-in, while depletion of ions will cause salting-out. Our results show that for a certain type of salts, the salting-in or salting-out effects depend on the specifics of the solute.

2.1 Computational Methods

2.1.1 Simulation details

MD simulations were performed using GROMACS³⁷. The partial charges q and Lennard-Jones interaction range σ and depth ϵ of Na^+ and Cl^- ions are listed in Table 5.2. The three-site TIP3P³⁸ water model was used as a template for the solute, referred as MWS (modified water as solute). To cover a broader range of solute properties, values of σ and q of MWS’s atoms were rescaled with factor k_σ and k_q from 0.5 to 1.5, in a step of 0.1,

Table 2.1 Force field parameters used in our simulations

	atoms	σ (Å)	ϵ (kJ/mol)	charge
Ions	Na ⁺	3.328	0.003	1
	Cl ⁻	4.401	0.100	-1
TIP3P	OW	3.151	0.152	-0.834
	HW	0.000	0.000	0.417

separately. These modifications generated a total of 121 solute molecules. Solvation in pure water was simulated by embedding one MWS in a cubic water box composed of 907 water molecules with periodic boundary conditions applied. For simulations in 2M NaCl, 35 pairs of Na⁺ and Cl⁻ ions were added. All the simulations were done in the NPT ensemble (300K, 1atm). SHAKE constraints³⁹ were applied to all bonds involving hydrogen. The movement integration time was 2 fs. Van der Waals (vdW) interactions were considered with a distance cutoff of 12 Å and a long range vdW correction was employed at distances greater than the cutoff. Electrostatic interactions were treated using the Particle Mesh Ewald (PME) algorithm^{40–42}. The non-bonded interaction pair-list was updated every 10 fs.

The solvation free energy (ΔG) was calculated using the thermodynamic integration (TI) technique^{43–45} with the soft-core method⁴⁶. The transfer of MWS from solvent into vacuum was broken down into two steps. In the first step, the atomic charges on MWS were turned off and in the second step the vdW interactions of the neutralized MWS were decoupled. Each transformation was performed using 21 equally spaced λ windows ($\lambda = 0.00, 0.05, \dots, 1.00$). For each window, simulations were equilibrated for 200 ps, followed by 2.3 ns data collection¹.

Normal MD simulations at constant temperature (300K) and constant pressure (1 bar) maintained by Berendsen coupling⁴⁷ were performed for 20ns for each MWS. All of the statistical analysis is based on these normal simulations.

¹The values of k_σ , k_q , solvation free energies and error estimations in pure water and NaCl electrolyte of 121 MWSs can be found in Table 2.2 at the end of this chapter.

2.1.2 Error estimation of ΔG

The uncertainty of an observable A is given by⁴

$$\delta A = \sqrt{\langle A^2 \rangle_C} \sqrt{\frac{2\tau_A}{T}} \quad (2.1)$$

where $\langle A^2 \rangle_C$ and τ_A are the standard deviation and correlation time respectively of the observable, A , over the entire time, T . According to this definition, we first calculated $\delta(dH/d\lambda)$ at each λ window. The statistical uncertainty of the predicted ΔG can be calculated from individual $\delta(dH/d\lambda)$ by weighting them with the width of their λ window through Gaussian error propagation

$$\delta(\Delta G) = \sqrt{\sum_i \left((\lambda_{j+1} - \lambda_j) \times \delta\left(\frac{dH}{d\lambda_i}\right) \right)^2} \quad (2.2)$$

We ran 2 separate simulations with the same setup but different initial velocities and compared the individual results to monitor whether the simulations had really converged. The largest error of our results is within 0.05 kcal/mol. The same range of errors was reported in a similar free energy calculation⁴³. For simplicity, the error bars are shown in Figure 3 only.

2.1.3 Kirkwood-Buff integral

Kirkwood-Buff (KB) theory is a statistical mechanical theory of solutions⁴⁸. The principal quantities of interest are the KB integrals

$$K_{ij} = 4\pi \int_0^\infty \left[g_{ij}^{\mu VT}(r) - 1 \right] r^2 dr \quad (2.3)$$

where r is the distance between two species, i and j . $g_{ij}(r)$ is the radial distribution function (RDF) for species j around the center of species i . These integrals can be used to relate the thermodynamic properties of a solution to the molecular distributions in that solution. Here we focus on the chemical potential μ of a solute molecule (s) in solvent

$$\mu_s = \mu_s^{id} + \mu_s^{ex} \quad (2.4)$$

where μ_s^{id} and μ_s^{ex} denote the ideal and excess (also known as hydration free energy, ΔG) contributions to the chemical potential of the solute.

For a ternary system in which the solute is infinitely dilute, the derivative of the excess chemical potential with respect to the concentration of cosolvent is given by⁴⁹

$$-\beta \left(\frac{\partial \mu_s^{ex}}{\partial \ln \rho_c} \right)_{p,T,\rho_s \rightarrow 0} = \frac{\rho_c (K_{sc} - K_{sw})}{1 + \rho_c (K_{cc} - K_{cw})} \quad (2.5)$$

where K_{sc} and K_{sw} are the KB integrals of cosolvent (c) and water (w) around solute. The part $\rho_c (K_{sc} - K_{sw})$ is usually called the preferential binding parameter, v . In our present simulation, we have used the same concentration of 2M NaCl in all the solutions. Thus ρ_c , K_{cc} and K_{cw} are constants in the formula, so that $\Delta \mu_s^{ex} \propto -(\Delta K_{sc} - \Delta K_{sw})$. This gives a direct link of the hydration free energy change ($\Delta \Delta G$) to the differential KB integral ($\Delta KBI = \Delta K_{sc} - \Delta K_{sw}$) from pure water to 2M NaCl.

The KB integral was formulated in the grand canonical ensemble (μVT). Realizing that the values of K_{ij} reflect excesses (positive K_{ij}) or deficits (negative K_{ij}) of species j around the central i , the upper limit (∞) of the KB integrals can be reduced to a finite value, R_c , when the integral is calculated in the NPT ensemble^{50,51}. R_c is an estimate of the scale of the preferential interaction distance over which the local composition is different from the bulk. The choice of R_c will be discussed later. An equivalent relation between the free energy change and the excess (or deficit) number of solutes or water molecules can be derived from the Gibbs-Duhem equation⁵².

2.2 Results and Discussions

2.2.1 Salting effects resolved by hydration free energy calculations

The solvation free energies in pure water, ΔG_{pure} , of MWS are shown in Figure 2-1(a). The MWS with $k_q = 1.0$ and $k_\sigma = 1.0$ is the standard TIP3P water molecule, with $\Delta G = -6.09$

kcal/mol. This value agrees well with another theoretical result of -6.10 kcal/mol⁴⁵ and the experimental data of -6.32⁵³. The solvation behavior of MWSs with $\Delta G = -6.09$ kcal/mol is similar to that of bulk water molecules even though the partial charges and the molecular sizes are different.

The values of the vdW radius and atomic charge determine the hydrophobicity of the atoms: the atoms with small vdW radius and large atomic charge have large values of charge density, thus they are hydrophilic; on the contrary, the atoms with large vdW radius and small atomic charge have small values of charge density, thus are hydrophobic. In the experiment, a number of methods have been developed to determine the hydrophobicity^{54–56}, like partitioning method, accessible surface area method, chromatographic method and physical property methods. The solvation free energies shown in Figure 2-1(a) are reasonable index to define the hydrophobicity of the molecules^{36,57}. Because water is the most commonly used medium as in our present simulations, hence the solvation free energy of TIP3P water is treated as the boundary between hydrophobic and hydrophilic species. The line with $\Delta G = -6.09$ kcal/mol divides solute molecules into two classes: hydrophobic MWSs (with $\Delta G > -6.09$) and hydrophilic MWSs (with negative $\Delta G < -6.09$). ΔG decreases quickly when the partial charges become large and the vdW radii become small. The lowest ΔG value is -174.80 kcal/mol for MWS with $k_q = 1.5$ and $k_\sigma = 0.5$. On the contrary, ΔG increases when the charges become small and the vdW radii become large. The highest ΔG is 4.03 kcal/mol for MWS with $k_q = 0.5$ and $k_\sigma = 1.5$.

We also calculated the solvation free energies in 2M NaCl, ΔG_{NaCl} as shown in Table 2.2. The solubility of certain MWS in NaCl electrolyte and in pure water can be connected via⁵⁸

$$c_{NaCl} = c_{pure} \times \exp \left(- \frac{\Delta G_{NaCl} - \Delta G_{pure}}{RT} \right) = c_{pure} \times \exp \left(- \frac{\Delta \Delta G}{RT} \right) \quad (2.6)$$

where c_{NaCl} and c_{pure} are the concentration of MWS in NaCl solution and in pure water, respectively. $\Delta \Delta G = \Delta G_{NaCl} - \Delta G_{pure}$ reflects the salting effect by NaCl: a negative $\Delta \Delta G$ reflects better solvated (salting-in); a positive $\Delta \Delta G$ reflects worse solvated (salting-out). The contours of $\Delta \Delta G$ is shown in Figure 2-1(b), from which both salting-in and salting-

Figure 2-1 (a) Solvation free energy (ΔG) of solute molecules in water; (b) Change of solvation free energy ($\Delta\Delta G$) from pure water to 2M NaCl solution.

out phenomena are observed depending on the solute. In other words, the salting effect is solute-specific.

Generally, the hydrophobic species will be salted out, while the hydrophilic species will be salted in. However, some intermediate solutes (solutes in area B in Figure 2-1(b)) are also salted in. Although the variation (q and σ) of solutes from area A to area B is smooth, which is reflected by the continuous ΔG in Figure 2-1(a), the $\Delta\Delta G$ is not continuous.

2.2.2 The role of ions in salting effects

The discontinuity of salting effects between regions A and B implies some differences in the ion-MWS binding. The spatial distributions of Na^+ and Cl^- ions around MWS are shown in the insert in Figure 2. The MWS itself has four strongly bound water molecules in a tetrahedral configuration in space (insert of Figure 6-6(b), four water are shown in light blue sticks). For MWS in region A, only Na^+ can penetrate into the interstices between two water molecules near the modified O atom of MWS. The binding of Na^+ is mainly on the equatorial direction. MWS in region B has the same atomic charges as the above mentioned MWS, but due to the increase of the vdW radius, water molecules are not strongly bound. Some Cl^- ions are able to replace water and bind to the modified H of MWS. Na^+ ions can also replace two waters near the modified O of MWS, and the Na^+ distribution changes

Figure 2-2 Radial distribution function (RDF) of Na^+ and Cl^- around (a) the center solute molecule in region A) ($k_q = 1.5$, $k_\sigma = 0.5$) and (b) the center molecule in region B) ($k_q = 1.5$, $k_\sigma = 1.1$) in Figure 2-1. Spatial distributions of Na^+ (bluish violet cross, density > 0.1 ions/ $\text{\AA}^3$), Cl^- (yellow cross, density > 0.1 ions/ $\text{\AA}^3$) were shown as inserts. Solute molecule was shown as red and white balls. The solute itself has four strongly bound water molecules (shown in cyan sticks) in a tetrahedron configuration.

to the axial direction. The hydration change from A to B can be quantitatively compared through the radial distribution functions (RDF) of Na^+ and Cl^- with respect to the MWS. As shown in Figure 6-6, a significant Cl^- peak appeared within 0.3 nm in the case of MWS in region B, while Cl^- is absent in region A.

The ion distributions shown above indicate that the specific ion binding pattern is highly sensitive on the shape, size and charge density of the solute molecule. On the other hand, based on the following analysis it seems that only the relative excess or deficiency of ions around a solute molecule has high correlation with salting effects, regardless of the specific ion distribution pattern. Thus how the detailed 3-dimensional information of ion binding contributes to the salting effect is still an open question.

Figure 2-3 The dependence of hydration free energy change ($\Delta\Delta G$) and hydration free energy in pure water (ΔG_{pure}). The dot and dot-dash lines indicate the ΔG of Na^+/Cl^- and water, respectively. Experimental values of methane and ethane (green hollow circles) at 25 °C were also shown. The solute molecules in region A and region B in Figure 2-1 are highlighted in blue and red, respectively.

2.2.3 The relation between the salting effects and hydration free energies

The relationship between $\Delta\Delta G$ and ΔG_{pure} is shown in Figure 2-3 for all the 121 MWS. Three regions (labeled 1, 2 and 3) with distinct correlations between $\Delta\Delta G$ and ΔG_{pure} can be seen. The two boundaries (dash and dash-dot lines) are ΔG_{pure} of water molecule (-6.09 kcal/mol) and Na^+/Cl^- (-88 kcal/mol)⁵⁹, respectively. In region 1, $\Delta G_{pure} < -88$ kcal/mol. It is a salting-in region with all $\Delta\Delta G$ less than 0. $\Delta\Delta G$ is roughly linearly correlated to ΔG_{pure} in this region. These MWSs are all highly polarized species (large k_q and small k_σ) and are able to have strong Coulomb interactions with salt ions. Region 3, $\Delta G_{pure} > -6.09$ kcal/mol and $\Delta\Delta G > 0$, is a salting-out region. These MWSs are most hydrophobic with small k_q and large k_σ . A linear relation between $\Delta\Delta G$ and ΔG_{pure} is also observed in this region. For MWS with $-88 \text{ kcal/mol} < \Delta G_{pure} < -6.09 \text{ kcal/mol}$ (region 2), the situation is different. The $\Delta\Delta G$ in this region are usually small and do not

Figure 2-4 (a) Pearson's correlation between $121\ v$ and $\Delta\Delta G$ with respect to cutoff of integral distance. (b) The convergence of differential Kirkwood-Buff integral (ΔKBI) of three representative solute molecules.

have an obvious relationship with ΔG_{pure} . Interestingly, it is found that in region B where specific ion binding patterns are existing (Figure 2-1(b)), the $\Delta\Delta G$ are generally negative (shown by the red points in Figure 3).

In our present analysis, we have adopted an artificial solute model through modification of partial charges and atomic radius of a water molecule. The advantage of this choice is that we can obtain hydration changes with respect to smoothly modified solutes. However, the generalization of our finding to real solutes has to be addressed. The experimental values⁶⁰ of the hydration change of methane and ethane from pure water to 2M NaCl solution measured at 25°C are also shown in the Figure 2-3. We found that the hydration change follows the simulation data nicely. We think our findings can be generalized to other typical salts like KCl, NaI, $MgCl_2$, etc. as well.

2.2.4 Interpretation of salting effects by Kirkwood-Buff integral

Implementation of KB theory in predicting the free energy change requires calculation of the preferential binding parameter v . The choice of R_c is a potential source of errors when

Figure 2-5 (a) The relationship between salting effect ($\Delta\Delta G$) and differential Kirkwood-Buff integral (ΔKBI) of NaCl (ΔK_{sc}) and water (ΔK_{sw}), (b) 2D map of ΔK_{sc} , and (c) 2D map of ΔK_{sw} .

calculating K_{sc} and K_{sw} . Theoretically when R_c is large enough, the RDF of water or ions with respect to the solute will converge to a constant value, the bulk average. The values of K_{sc} and K_{sw} , and the preferential binding parameter, ν , will also converge. In Figure 4-5(a) the values of the differential KB integral (ΔKBI) of three representative MWS as a function of radial distance are shown. The convergence behaviors of ΔKBI of the MWSs are different depending on the specifics of the MWS. For the most hydrophobic MWS ($k_q = 0.5$, $k_\sigma = 1.5$) and water ($k_q = 1.0$, $k_\sigma = 1.0$) ΔKBI converged at about 0.8 nm; and the most hydrophilic ($k_q = 1.5$, $k_\sigma = 0.5$) MWS needs a large R_c (> 1.0 nm) to get converged value. This phenomenon is not strange because the latter one is most significantly salted which is related to a big change of its environment. Although the change of the MWS environment is long ranged, a high Pearson's correlation⁶¹ (> 0.9) of the 121 $\Delta\Delta G \sim \Delta KBI$ pairs is well established when $R_c \sim 0.42$ nm (Figure 4-5(b)). This short distance only covers the first 1~2 hydration shells of MWS which can be interpreted as the directly bound ions or water molecules playing a dominant role in the hydration thermodynamics. This is consistent with a previous study that showed that the first hydration shell (so called coordination sphere) contributes dominantly to ion hydration⁶². The solvent change outside the coordination sphere can be treated as a long range correction.

The relationships between $\Delta\Delta G$ and ΔK_{sc} , ΔK_{sw} and $\Delta K_{sc} - \Delta K_{sw}$ are shown in Figure 2-5(a). The boundary between salting-in and salting-out, $\Delta\Delta G = 0$, is also the separatrix of the KB integral of zero. The high correlation between $\Delta\Delta G$ and the preferential binding

parameter indicates that a significant hydration change is expected with a large ΔK_{sc} or a large ΔK_{sw} . It is interesting to check the patterns of ΔK_{sc} and ΔK_{sw} of all 121 MWSs in comparison with that of $\Delta\Delta G$ shown in Figure 2-1b. The 2D map of ΔK_{sc} and ΔK_{sw} ($R_c = 0.47$ nm) of all MWSs are shown in Figure 2-5(b) and (c). One can find that the patterns well reproduce the free energy change as shown in Figure 2-1b. In the salting-in region (area A and B in Figure 2-1(b)), Na^+ and Cl^- ions are accumulated around the solute, and accordingly some water molecules are repelled away. It needs emphasizing that the KB integrals were calculated with respect to the COM of the MWS. Thus the detailed molecular structure is ignored. The spatial distributions in Figure 6-6 show that effects from Na^+ and Cl^- , respectively, are not well captured by the 2D map. On the contrary, in the salting-out region, water molecules are accumulated and Na^+ and Cl^- ions are deficient. It is found that a hydrophobic solute in NaCl solution has a “relatively” enhanced first water shell than that in pure water. Such behavior is consistent with the high charge density ions induced salting-out effects on hydrophobic plates in aqueous solution³⁶.

2.2.5 A schematic diagram of salting effects

In this work, the specific property of solute is used as a parameter to explore the salting-in and salting-out effects. In a recent study, the properties of salt ions, mainly the ionic charge density, were modified to study how the association free energy of two nanoscale hydrophobic surfaces is affected by the different types of salts³⁶. It is clear that the salting effects are dependent on the tripartite interaction among water, solute and ions. We summarize such delicate relations into a simple salting schematic diagram shown in Figure 2-6.

The hydration free energy of a water molecule is positioned at the origin. The x-axis and y-axis represent the hydration free energy of the ions and the solute, respectively. We modify the solute properties and keep the salt, NaCl, unchanged. On the schematic diagram, we get samples in the regions 1, 2 and 3 which are directly related the sections of 1, 2, and 3 in Figure 2-3. Within the three regions the hydration free energies of ions (like Na^+ and Cl^-) are less than that of water. In region 1 the hydration free energy of solutes is less than those of ions (highly hydrophilic), the solutes are salted in. In region 3 the hydration

Figure 2-6 Schematic diagram of salting effects. The origin represents the hydration free energy of a water molecule. The x-axis and y-axis represent the hydration free energy of the ions and the solute, respectively. NS stands for no clear salting effect.

free energies of solute are larger than that of water (hydrophobic), the solute will be salted out. This region was also covered in Zangi's study³⁶, where the solute is the hydrophobic plate and the ions have high ionic charge density. When the ionic charge density of ions is lowered, Zangi *et al.* found a transition in the plate solvation behavior from salting-out to salting-in³⁶. In the schematic diagram, it may be related to a translocation from region 3 to region 4. Interestingly region 4 is symmetric with region 1 and both of the regions are salting-in. When the hydration free energy of the solute is in between those of water and ions, like the cases in region 2 on the schematic diagram, we did not find very significant salting-in or salting-out behaviors. Thus we name this region as the region with no clear salting effects (NS). As in the cases of region 1 and region 4 which are symmetric with respect to the origin, both belong to the salting-in space. If such a symmetry is applicable, we can make the prediction that region 5, which is symmetric with region 2, is the NS region and region 6, which is symmetric with region 3, is the salting-out space.

2.3 Conclusion

In summary we have systematically generated a series of solute models from hydrophobic to hydrophilic. The hydration free energy was calculated in pure water and 2 M NaCl solution. Thus a two-dimensional map of hydration free energy change, $\Delta\Delta G$, is obtained. Both salting-in and salting-out effects were observed due to the action of a single kind of salt, NaCl.

The change of hydration free energy can be well reproduced from KB theory. Pearson's correlation analysis shows that the directly bound NaCl ions or water within 1~2 solvation shells play an important role in the hydration thermodynamics. For those molecules being salted-in, excess ions are found in the vicinity of and make strong interaction with them. These molecules are highly polarized with hydration free energy less than that of NaCl. For those molecules being salted-out, ions exhibit preferential exclusion, and the water shell of those solutes is enhanced.

Employing the hydration free energies of solutes and ions as independent variables a schematic diagram of salting effects is suggested. The salting-in and salting-out behaviors are a result of delicate interactions among solutes, ions and water molecules.

Table 2.2 Solvation free energies (in kcal/mol) in pure water and NaCl electrolyte of 121 solute molecules

Force field		Pure water		2M NaCl		Salting effect	
k_q	k_σ	ΔG	error	ΔG	error	$\Delta\Delta G$	error
0.5	0.5	-19.773	0.019	-19.741	0.021	0.031	0.028
0.5	0.6	-9.805	0.017	-9.676	0.017	0.129	0.024
0.5	0.7	-4.290	0.013	-4.222	0.014	0.068	0.019
0.5	0.8	-1.656	0.012	-1.610	0.013	0.047	0.018
0.5	0.9	-0.265	0.013	-0.156	0.015	0.110	0.020
0.5	1.0	0.656	0.015	0.900	0.018	0.243	0.023
0.5	1.1	1.418	0.019	1.678	0.023	0.259	0.030
0.5	1.2	2.133	0.021	2.419	0.026	0.287	0.033
0.5	1.3	2.705	0.025	3.105	0.029	0.399	0.038
0.5	1.4	3.366	0.028	3.753	0.034	0.386	0.044
0.5	1.5	4.026	0.033	4.433	0.038	0.407	0.050
0.6	0.5	-30.448	0.021	-30.330	0.022	0.119	0.030
0.6	0.6	-16.537	0.019	-16.437	0.020	0.100	0.027
0.6	0.7	-8.159	0.016	-8.079	0.017	0.080	0.024
0.6	0.8	-3.708	0.014	-3.650	0.014	0.057	0.020
0.6	0.9	-1.456	0.014	-1.362	0.015	0.095	0.020
0.6	1.0	-0.126	0.015	0.088	0.018	0.214	0.024
0.6	1.1	0.852	0.019	1.090	0.023	0.238	0.030
0.6	1.2	1.695	0.021	1.976	0.026	0.281	0.033
0.6	1.3	2.361	0.025	2.752	0.029	0.391	0.038
0.6	1.4	3.085	0.028	3.459	0.035	0.374	0.044
0.6	1.5	3.789	0.033	4.191	0.038	0.402	0.050
0.7	0.5	-42.327	0.023	-42.230	0.024	0.096	0.033
0.7	0.6	-24.485	0.020	-24.430	0.021	0.055	0.029
0.7	0.7	-13.223	0.018	-13.086	0.019	0.137	0.027

Continued on next page

Table 2.2 – *Continued from previous page*

Force field		Pure water		2M NaCl		Salting effect	
k_q	k_σ	ΔG	error	ΔG	error	$\Delta\Delta G$	error
0.7	0.8	-6.558	0.016	-6.494	0.017	0.064	0.023
0.7	0.9	-3.069	0.015	-3.032	0.016	0.037	0.022
0.7	1.0	-1.146	0.016	-0.941	0.019	0.205	0.024
0.7	1.1	0.135	0.019	0.339	0.023	0.204	0.030
0.7	1.2	1.162	0.021	1.419	0.026	0.257	0.034
0.7	1.3	1.939	0.025	2.313	0.029	0.374	0.038
0.7	1.4	2.741	0.028	3.106	0.035	0.365	0.044
0.7	1.5	3.509	0.033	3.900	0.038	0.392	0.050
0.8	0.5	-55.515	0.024	-55.497	0.025	0.018	0.034
0.8	0.6	-33.257	0.022	-33.237	0.022	0.020	0.031
0.8	0.7	-19.120	0.020	-18.958	0.021	0.162	0.029
0.8	0.8	-10.241	0.018	-10.156	0.019	0.085	0.026
0.8	0.9	-5.199	0.016	-5.192	0.017	0.007	0.024
0.8	1.0	-2.476	0.016	-2.344	0.019	0.132	0.025
0.8	1.1	-0.765	0.019	-0.588	0.023	0.177	0.030
0.8	1.2	0.508	0.021	0.725	0.026	0.217	0.034
0.8	1.3	1.428	0.025	1.788	0.029	0.360	0.038
0.8	1.4	2.339	0.028	2.690	0.035	0.351	0.044
0.8	1.5	3.176	0.033	3.553	0.038	0.376	0.050
0.9	0.5	-69.699	0.025	-69.715	0.026	-0.017	0.036
0.9	0.6	-42.870	0.023	-42.850	0.023	0.020	0.033
0.9	0.7	-25.680	0.021	-25.557	0.022	0.123	0.030
0.9	0.8	-14.657	0.019	-14.564	0.020	0.093	0.028
0.9	0.9	-7.875	0.018	-7.858	0.019	0.017	0.026
0.9	1.0	-4.112	0.017	-4.028	0.020	0.084	0.026
0.9	1.1	-1.860	0.020	-1.701	0.024	0.159	0.031

Continued on next page

Table 2.2 – *Continued from previous page*

Force field		Pure water		2M NaCl		Salting effect	
k_q	k_σ	ΔG	error	ΔG	error	$\Delta\Delta G$	error
0.9	1.2	-0.296	0.022	-0.078	0.026	0.217	0.034
0.9	1.3	0.827	0.025	1.169	0.029	0.342	0.038
0.9	1.4	1.865	0.028	2.213	0.035	0.348	0.044
0.9	1.5	2.791	0.033	3.158	0.038	0.367	0.050
1.0	0.5	-84.925	0.027	-84.876	0.029	0.049	0.040
1.0	0.6	-53.249	0.024	-53.109	0.026	0.140	0.035
1.0	0.7	-32.935	0.022	-32.820	0.023	0.115	0.032
1.0	0.8	-19.634	0.020	-19.533	0.022	0.101	0.030
1.0	0.9	-11.122	0.019	-11.139	0.021	-0.017	0.028
1.0	1.0	-6.088	0.018	-6.053	0.021	0.035	0.028
1.0	1.1	-3.179	0.020	-3.086	0.024	0.093	0.031
1.0	1.2	-1.215	0.022	-1.031	0.026	0.183	0.034
1.0	1.3	0.134	0.025	0.447	0.029	0.313	0.038
1.0	1.4	1.310	0.028	1.631	0.035	0.321	0.045
1.0	1.5	2.349	0.033	2.713	0.038	0.364	0.050
1.1	0.5	-101.105	0.028	-101.291	0.029	-0.186	0.040
1.1	0.6	-64.350	0.025	-64.170	0.027	0.180	0.037
1.1	0.7	-40.767	0.023	-40.634	0.025	0.133	0.034
1.1	0.8	-25.092	0.022	-25.069	0.023	0.023	0.031
1.1	0.9	-14.836	0.020	-14.836	0.022	-0.001	0.030
1.1	1.0	-8.494	0.019	-8.486	0.022	0.008	0.029
1.1	1.1	-4.733	0.021	-4.687	0.025	0.046	0.032
1.1	1.2	-2.305	0.022	-2.208	0.027	0.097	0.035
1.1	1.3	-0.672	0.025	-0.393	0.029	0.279	0.039
1.1	1.4	0.700	0.028	0.987	0.035	0.287	0.045
1.1	1.5	1.852	0.033	2.192	0.038	0.340	0.050

Continued on next page

Table 2.2 – *Continued from previous page*

Force field		Pure water		2M NaCl		Salting effect	
k_q	k_σ	ΔG	error	ΔG	error	$\Delta\Delta G$	error
1.2	0.5	-118.099	0.030	-118.482	0.032	-0.383	0.043
1.2	0.6	-75.942	0.026	-76.003	0.027	-0.061	0.037
1.2	0.7	-49.228	0.025	-49.200	0.026	0.028	0.036
1.2	0.8	-31.089	0.023	-31.068	0.024	0.021	0.033
1.2	0.9	-19.021	0.022	-18.984	0.023	0.037	0.032
1.2	1.0	-11.241	0.021	-11.246	0.023	-0.005	0.031
1.2	1.1	-6.519	0.021	-6.564	0.026	-0.045	0.033
1.2	1.2	-3.565	0.022	-3.473	0.027	0.092	0.035
1.2	1.3	-1.582	0.025	-1.302	0.029	0.280	0.039
1.2	1.4	-0.015	0.028	0.269	0.035	0.284	0.045
1.2	1.5	1.284	0.033	1.617	0.038	0.333	0.050
1.3	0.5	-136.219	0.030	-136.782	0.030	-0.563	0.042
1.3	0.6	-88.304	0.027	-88.480	0.028	-0.176	0.039
1.3	0.7	-58.228	0.026	-58.321	0.027	-0.094	0.037
1.3	0.8	-37.597	0.024	-37.511	0.025	0.086	0.035
1.3	0.9	-23.557	0.023	-23.559	0.024	-0.002	0.033
1.3	1.0	-14.389	0.022	-14.454	0.024	-0.065	0.033
1.3	1.1	-8.624	0.022	-8.679	0.026	-0.055	0.034
1.3	1.2	-4.980	0.023	-4.966	0.028	0.014	0.036
1.3	1.3	-2.641	0.026	-2.443	0.030	0.198	0.039
1.3	1.4	-0.813	0.028	-0.538	0.035	0.276	0.045
1.3	1.5	0.649	0.033	0.959	0.038	0.310	0.050
1.4	0.5	-154.985	0.033	-155.718	0.034	-0.733	0.047
1.4	0.6	-101.280	0.028	-101.663	0.031	-0.383	0.042
1.4	0.7	-67.913	0.027	-67.909	0.028	0.004	0.039
1.4	0.8	-44.538	0.025	-44.516	0.026	0.022	0.037

Continued on next page

Table 2.2 – *Continued from previous page*

Force field		Pure water		2M NaCl		Salting effect	
k_q	k_σ	ΔG	error	ΔG	error	$\Delta\Delta G$	error
1.4	0.9	-28.492	0.024	-28.510	0.025	-0.018	0.035
1.4	1.0	-17.831	0.023	-17.918	0.025	-0.088	0.034
1.4	1.1	-10.989	0.023	-11.102	0.028	-0.113	0.036
1.4	1.2	-6.623	0.023	-6.607	0.028	0.016	0.037
1.4	1.3	-3.804	0.026	-3.602	0.030	0.201	0.040
1.4	1.4	-1.698	0.028	-1.437	0.035	0.261	0.045
1.4	1.5	-0.037	0.033	0.252	0.038	0.289	0.050
1.5	0.5	-174.800	0.033	-175.720	0.033	-0.920	0.047
1.5	0.6	-115.108	0.029	-115.639	0.034	-0.530	0.045
1.5	0.7	-78.104	0.028	-78.167	0.029	-0.063	0.040
1.5	0.8	-51.893	0.026	-51.940	0.028	-0.047	0.038
1.5	0.9	-33.848	0.025	-33.801	0.027	0.047	0.037
1.5	1.0	-21.610	0.024	-21.685	0.026	-0.075	0.035
1.5	1.1	-13.601	0.024	-13.897	0.029	-0.296	0.038
1.5	1.2	-8.432	0.024	-8.515	0.029	-0.083	0.038
1.5	1.3	-5.113	0.026	-4.970	0.030	0.143	0.040
1.5	1.4	-2.665	0.029	-2.413	0.035	0.251	0.045
1.5	1.5	-0.811	0.033	-0.545	0.038	0.266	0.050

Chapter 3

Dissociation of Nanoparticles in Aqueous Guanidinium Chloride and Urea Solutions

For hundreds of years, urea and GdmCl have been the most commonly used chemical agents for protein denaturation. The question of how denaturation happens has been a historical puzzle and has been reviewed from time to time. Earlier experimental researchers provided clues that denaturant molecules can directly bind to protein atoms^{63,64}. In recent years, thanks to the development of computer modeling techniques, many studies have been conducted to rationalize the denaturation process at the atomic level^{19,65–70}. Two opposing mechanisms, “direct interaction mechanism” and “indirect interaction mechanism” have been proposed. In the “direct interaction mechanism”^{67,68,71,72}, direct interactions between proteins and denaturants are supposed to be important. On the contrary, the “indirect interaction mechanism”^{11–15,19,20} emphasizes the effect of changing of water structure through addition of denaturants. It is difficult to distinguish between these two mechanisms because the denaturation may be better treated as a mixture of both. However, the “direct interaction mechanism” is becoming widely accepted while the other one lacks experimental support^{34,35,73–79}.

For urea-denaturation, a 2-stage unfolding of the proteins has been proposed⁶⁸. In the first stage, urea molecules are enriched in the vicinity of the protein surface. These urea molecules replace water to “solvate” the protein surface. In the second stage, urea penetrates into the interior of the protein, forming hydrogen bonds with backbone atoms and breaking the hydrophobic interaction. Thermodynamically, the urea-binding shifts the folding/unfolding equilibrium towards protein unfolding and finally causes denaturation. For simplicity, the central point of the 2-stage denaturation is that urea perturbs important interactions within protein itself or interactions between protein and water. These interactions are the basis of stability of protein structures which can be roughly divided into two types: (1) hydrophobic pair, which mainly consists of hydrophobic residues such as isoleucine, phenylalanine, tryptophan, and tyrosine, etc.; (2) ionized pair, which mainly consists of charged residues, such as arginine(+), lysine(+), aspartic acid(-), glutamic acid(-), and the polar species which potentially form hydrogen bonds with each other or with water.

GdmCl is found experimentally to be about 1.5 to 2.5 times more effective as a protein denaturant than urea.⁸⁰ But to our knowledge, few studies have been conducted to compare the denaturing power between urea and GdmCl^{81,82}. It is still a mystery why GdmCl is more efficient in denaturation of proteins than urea.

In a recent experiment work, the authors monitored the aggregation kinetics of a model system, polystyrene particles, in different salt solutions. They found that the aggregation process of these particles is highly affected by the types of co-existing salts⁸³. Compared with the real protein model in the above mentioned simulation study, the application of such a particle model enables us to interrogate the underlying mechanism which is usually hidden in more complex systems. Inspired by polystyrene particle model, in this study, we have compared the denaturing power of urea and GdmCl in dissociating hydrophobic and charged nano-particles (NPs). By molecular dynamics simulations, we show that 4M GdmCl solution is much more able than 7M urea solution to dissociate both hydrophobic and charged NPs. Both urea and GdmCl affect aggregation of NPs through direct binding to the NP surface. The advantages of GdmCl originate from the net charge of bound guanidinium ions (Gdm⁺) which can generate a local positively charged environment around the hydrophobic and the negatively charged NPs. Such an effective coating introduces

Coulombic repulsion between all the NPs. The Gdm⁺ tends to form dimers and tetramers in stacking configurations which helps to stabilize the separated state. Urea shows certain ability to dissociate the hydrophobic NPs. However, for the charged NPs, urea acts like “glue” that prevents the separation. This is due to the ordered hydrogen bond network formed within urea molecules located between two opposite-charged NPs. Regarding the binding patterns, interestingly, we observe a flipping of the denaturant molecules, more significantly for Gdm⁺, from a parallel configuration (on the surface of the hydrophobic NP) to a perpendicular one (on the surface of the charged NP). Finally, the contribution from “the indirect interaction mechanism” is evaluated through comparing the solute dynamics. The movements of NPs are partially frozen in urea/GdmCl solution compared to that in pure water. However, the dissociation power is not related to the hardening of NPs.

3.1 Computational Methods

Figure 3-1 Snapshots of the simulation box. Water and chloride ions are not shown in (b) and (c).

3.1.1 Nanoparticle model

We used buckyballs (C_{60} , diameter ~ 0.7 nm) as our NP model. The atom type of aromatic carbon (CA) of AMBER(ff03)⁸⁴ force field was assigned to each C_{60} atom. Two NPs with zero partial charge on each atom were used to represent hydrophobic pair-wise interactions. For the charged case, either $+0.1e$ or $-0.1e$ partial charge was assigned to each carbon

Table 3.1 Simulation box contents

Type of NP	Type of solvent	Number of molecules			Volume (nm ³)	Denaturant's concentration (M/L)
		NP	Denaturant	Water		
No-charge NP pair	pure	2	0	2000	62.27	0
	GdmCl	2	170	1660	71.28	3.96
	Urea	2	150	1700	63.18	3.94
	Urea(2)	2	310	1690	73.68	6.99
Charged NP pair	pure	2	0	2000	61.73	0
	GdmCl	2	170	1660	71.03	3.97
	Urea	2	150	1700	62.75	3.97
	Urea(2)	2	310	1690	73.31	7.02

atom of C_{60} . The total charges for two ionized NPs were $+6e$ (designated NP+) and $-6e$ (designated NP-), respectively.

3.1.2 Simulation details

The force field parameters of urea and chloride ion models from the AMBER package were used. For Gdm+, the structural topology was adopted from arginine force field of AMBER(ff03), atomic charges were generated with R.E.D. tools⁸⁵ through quantum mechanics calculations with *Gaussian03*⁸⁶ at the HF/6-31G* level. For water, the TIP3P³⁸ model was used. Snapshots of the simulation box are shown in Figure 3-1. For the solvent, we have considered four different cases: pure water, 4M/7M urea and 4M GdmCl. The details of the eight simulation systems are shown in Table 3.1.

All of the simulations were performed with GROMACS³⁷, using the NPT ensemble at 1 atm and 300 K. SHAKE constraints³⁹ were applied to all bonds involving hydrogen atoms. The MD integration time step was 2 fs. The van der Waals (vdW) interactions were treated with a typical distance cutoff of 12 Å. The electrostatic interactions, on the other hand, were handled with the Particle Mesh Ewald algorithm (with grid spacing of 0.12 nm)^{41,42}. The non-bonded interaction pair-list was updated every 10 fs.

3.1.3 Umbrella sampling method

The potential of mean force (PMF) between two NPs was calculated with the umbrella sampling technique. The distance between the center of mass (COM) of two NPs was restrained at certain distance d_0 with a harmonic force

$$F = k \times (d - d_0)$$

where k is force constant $4000 \text{ kJmol}^{-1}\text{nm}^{-2}$. The sampling distance d_0 was chosen from 0.90 to 1.80 nm with a step of 0.05 nm. At each distance, the system was first equilibrated for 2 ns followed by a 10 ns production simulation.

3.2 Results and discussions

3.2.1 PMF between two NPs

Figure 3-2 The potential of mean force (PMF) of two (a) no-charge nano-particles (NPs) and (b) charged NPs with respect to their center of mass distance. The direct contact pair (DCP) states in all cases are considered having free energy zero.

The PMFs between two hydrophobic NPs are shown in Figure 3-2(a) (black line). The PMFs have two significant minima and one significant peak. The first minimum (at 0.97 nm) corresponds to the direct contact pair (DCP) of two NPs, while the second minimum (roughly at 1.3 nm) corresponds to solvent separated pair (SSP). The two minima are sepa-

rated by a desolvation barrier (roughly at 1.18 nm). For the charged NPs in pure water (Figure 3-2(b), black line), the energy well of the DCP state becomes much deeper implying that the DCP between NP+ and NP- is much more stable than that of the two hydrophobic NPs. The PMF beyond SSP keeps increasing because the electrostatic interaction is long-ranged. Overall, both the hydrophobic and charged NPs tend to aggregate because the DCP state is always the global minimum. Hence in the following discussions, the DCP state is always treated as the reference state (E_{DCP} is set to zero as shown in Figure 3-2).

It is clear that the breaking of the residues' direct contact results in lost of protein activity, which is the central of denaturation. For the NP models, this process corresponds to the transition from DCP to SSP state. Energetically, the rate constant K for transition from DCP to SSP predicted by the collision theory⁸⁷ is

$$K_{DCP \rightarrow SSP} = A \times \exp\left(\frac{-E_{act}}{RT}\right)$$

where E_{act} is the so-called *activation energy*, which is the energy barrier to dissociate a DCP. The probability of finding two NPs separated at distance, d , follows the Boltzmann distribution

$$\frac{\rho(d)}{\rho(DCP)} \propto \exp\left(-\frac{E(d)}{RT}\right)$$

Regarding the dissociation of aggregated NPs, the DCP of NPs is easy to be broken kinetically if E_{act} is lowered or thermodynamically becomes unstable when E_{SSP} is lowered. The free energy at a larger distance (1.8 nm), $E_{1.8nm}$, provides us with the ability for urea/GdmCl to fully disperse the NPs.

The free energy terms are shown in Table 3.2. For no-charge NPs, after addition of 4M urea, E_{act} decreases from 5.0 kcal/mol in pure water to 4.38 kcal/mol, and E_{SSP} decreases from 3.85 to 2.67 kcal/mol. Further addition of urea to 7M results in a stronger dissociation effect: E_{act} decreases to 3.96 kcal/mol, E_{SSP} further decreases to 1.82 kcal/mol. In 4M GdmCl, the free energy changes are more significant compared to those in 7M urea: (1) E_{SSP} decreases to a small value of 0.74 kcal/mol, which indicates a very stable SSP state; (2) $E_{1.8nm}$ is only 1.20 kcal/mol compared to a value of 2.82 kcal/mol in 7M urea.

Table 3.2 Activation energy and energy penalty (kcal/mol) at solvent separated pair (SSP) and 1.8 nm

Type of solvent	No-charge NP pair			Charged NP pair		
	E_{act}	E_{SSP}	$E_{1.8nm}$	E_{act}	E_{SSP}	$E_{1.8nm}$
pure	5.00	3.85	4.05	9.84	8.98	13.06
GdmCl (4M)	4.21	0.74	1.20	6.46	4.04	2.00
UREA (4M)	4.38	2.67	4.17	7.15	4.40	9.33
UREA (7M)	3.96	1.82	2.82	6.80	4.07	9.81

For ionized NPs, the values of E_{act} and E_{SSP} in pure water are very large, 9.84 and 8.98 kcal/mol, respectively. After addition of 4M GdmCl, E_{act} and E_{SSP} significantly decrease to 6.46 and 4.04 kcal/mol, respectively. In 4m/7M urea, the changes of E_{act} and E_{SSP} are roughly similar as those in 4M GdmCl. However, the values of $E_{1.8nm}$ are still very high, 9.33 and 9.81 kcal/mol in 4M and 7M urea, respectively. On the contrary, $E_{1.8nm}$ in 4M GdmCl is effectively lowered to a value of only 2.0 kcal/mol.

Based on the above comparison, both hydrophobic and charged NPs can be well dispersed in 4M GdmCl solution as reflected by small values of $E_{1.8nm}$. On the contrary, for hydrophobic NPs in 4M urea, the $E_{1.8nm}$ value is similar to that in pure water. Although further addition of urea to 7M effectively lowers the $E_{1.8nm}$ to 2.82 kcal/mol, this value is still considerably higher than the value of 1.20 kcal/mol in 4M GdmCl. The weakness of urea is especially significant in the case of the ionized NPs: $E_{1.8nm}$ of 4M/7M urea are still very high. It is clear that although urea can inhibit direct aggregation between NPs, the solvent separated NPs still form “loose aggregates” which are hard to be further separated.

Such loose aggregation can be used to explain the observations of some experiments: urea-denatured proteins have been found to be relatively compact and the native-like topology is still preserved; the secondary structure can be almost maintained^{88–91}. Moreover, there is functional evidence that urea only induces a weak loss of the activity of lysozyme (15% of the maximum) even at very high concentration (10 M)⁹². In the view that enzyme activity is related to structural stability, the weak loss of activity may indicate a moderate change of the conformation of the protein. On the other hand, GdmCl can induce more dramatic changes of protein structure than urea. This is consistent with the experimental

findings that 10 M GdmCl induced a large loss (up to 40%) of of lysozyme activity, while 10 M urea only induced a weak loss (15%) of activity⁹².

3.2.2 Direct interaction of denaturant with NPs

In brief, our results correctly reproduced the stronger denaturing power of GdmCl. GdmCl can not only promote the DCP $\rightarrow$ SSP transition more effectively, but also demonstrates long-ranged dispersive power which is absent in urea. The atomic-resolution picture of denaturant-NP interaction obtained from simulations may help to resolve the underlying denaturation mechanism.

Figure 3-3 Cumulative number of (a) Gdm+ and (b) urea around three kinds of nano-particles (NPs).

Binding shell of denaturant

The cumulative number of denaturant molecules around the three types of NPs is shown in Figure 3-3. These results were obtained from the umbrella sampling simulation trajectories at a separation of 1.80 nm where the two NPs are treated as independent particles. For a no-charge NP, there is an obvious binding shell within distance of ~ 0.85 nm, which consists of 8 Gdm+ ions. The interaction between Gdm+ and a no-charge NP is purely from vdW interaction. The electrostatic attraction between NP- and Gdm+ attracts more Gdm+, where the shell size increases to 17.5. It is not surprising that there is almost no Gdm+ near a NP+ because of electrostatic repulsion. For the case of urea, it seems that the number of bound

urea (within distance ~ 0.85 nm) is less dependent on the charges of the NPs. For no-charge NP, the binding shell consists of 17 urea, while for NP+/NP- the shell size slightly increased to 19. In contrast to the Gdm⁺ ion, urea does not have a net charge so it does not have preference to either NP+ or NP-.

Figure 3-4 Binding orientations of (a) Gdm⁺ and (b) urea around three kinds of nano-particles (NPs).

The number of Gdm⁺ around a NP- is twice as many as that around a no-charge NP. This large difference prompts us to expect a dramatic change of Gdm⁺'s binding pattern from the surface of a no-charge NP to a NP-. Both Gdm⁺ and urea are nearly planar in structure, while the NP also has a smooth surface. The angle θ between the normal vectors of NP surface and bound Gdm⁺/urea can be used to describe the orientation of Gdm⁺/urea relative to the NP surface. The distributions of $\cos(\theta)$ are shown in Figure 3-4. For a no-charge NP, over 86% of the bound Gdm⁺ has the $\cos(\theta)$ value of 0.8-1.0 which indicates that around 7 out of 8 Gdm⁺ ions are parallel to the surface of a no-charge NP as shown in the right-side insert of Figure 3-4(a). When the NP surface is negatively charged, over 66% of the 17.5 Gdm⁺ has the $\cos(\theta)$ value of 0.0-0.2, indicating that Gdm⁺ flips 90 degrees and binds perpendicularly to the NP- (left side insert of Figure 3-4(a)). This binding mode is favored by the close contact of negatively charged NP- atoms and positively charged amide hydrogen atoms and concomitant repulsion of the negatively charged amide nitrogen atoms away from the surface.

It is obvious that the perpendicular binding pattern leaves more space for Gdm⁺ ions

to accumulate nearby. Such a binding mode also applies for urea molecules near the NP-surface due to the fact that both urea and Gdm⁺ have the same amide group. On a no-charge NP, about 42% of the bound urea have the $\cos(\theta)$ value of 0.8-1.0, i.e., only 3-4 out of 8 Gdm⁺ ions are oriented parallel to the surface. The ratio is significantly lower than that of Gdm⁺ ions on the no-charge surface. The difference in ordering may be due to the fact that the symmetry of Gdm⁺ molecules is higher than that of urea molecules resulting in a stronger van der Waals interaction between the symmetric NP surface and Gdm⁺ molecules.

Figure 3-5 (a) Snapshot of two no-charge nano-particles (NPs) in 4M GdmCl. The intruded Gdm⁺ is red-colored and Gdm⁺ that form stacked dimer and tetramer are golden-colored. (b) Cumulative charge around NPs in 4M GdmCl.

Gdm⁺ - NP interaction

The simulation trajectory shows that for the SSP state ($d = 1.3\text{nm}$) of two no-charge NPs, there is often (up to 86.5% of the simulation time) one Gdm⁺ located at the inter-facial site between two NPs (red colored Gdm⁺ in Figure 3-5(a)). Moreover, for two no-charge NPs separated by 1.65nm (corresponds to the third minimum of the blue line in Figure 3-2(a)), there is a stacked Gdm⁺ dimer between the two NPs. The second energy barrier (at 1.45nm) of the PMF of the no-charge NPs corresponds to the transition from the one Gdm⁺ separated state to the Gdm⁺ dimer separated state. Gdm⁺ is structurally planar and has a negative charge center (C atom) and three positive charge centers ($-\text{NH}_2$). Thus Gdm⁺

can form stacked dimers and even tetramers through certain displacements in the lateral direction⁹³. Simulations show that these Gdm⁺ dimers and tetramers can bind firmly to the surface of a no-charge NP, like a growing "ladder" (golden colored Gdm⁺ in Figure 3-5(a)). These Gdm⁺ ions behave like space-blockers and can effectively diminish hydrophobic attraction.

Recalling that a Gdm⁺ ion has a net charge of $+1e$, the binding of Gdm⁺ on NP surface changes the local electrostatic environment. Hence we calculated the cumulative charge (sum of atomic charges on NP, Gdm⁺ and Cl⁻) around the three kinds of NPs. The values are shown in Figure 3-5(b). With the coating of Gdm⁺, there exists a effective positive-charge sphere (0.7-0.95 nm) around a no-charge NP. For a NP⁻, due to the excess number of Gdm⁺ in the binding shell (up to 17.5 Gdm⁺ ions), the positive-charge sphere further spreads from 0.7 to 1.12 nm. For a NP⁺, Cl⁻ ions tend to accumulate nearby which results in a relatively smaller positive-charge sphere. The long-range electrostatic repulsion between three kinds of coated NPs guarantees that the dispersion effect is long-ranged. Hence it is not surprising, as shown in Figure 3-2(b), that the values of $E_{1.8nm}$ of both no-charge and charged NPs are greatly lowered in GdmCl solution. From the above analysis the intrinsic net charge of Gdm⁺ contributes to its strong denaturation power.

Figure 3-6 A snapshot of nearest 65 ureas around positive-charged nano-particle (NP+) and negative-charged nano-particle (NP-). The center of mass distance of two nano-particles is 1.8 nm.

Urea - NP interaction

It is interesting to examine the patterns of urea molecules on charged NPs, because the PMFs of two charged NPs in 4M/7M urea keep increasing at large distance (Figure 3-2(b)). A snapshot of the nearest 65 urea molecules around NP+ and NP- (COM-COM distance is 1.80 nm) is shown in Figure 3-6. Urea molecules associate at the inter-facial site between two NPs through hydrogen bonds with each other. Their dipoles are ordered, pointing from NP+ to NP-. Obviously, these ordered-ureas act like “glue” between NPs which will prevent further dissociation of two NPs. If the charges of the NPs are removed, the ordering of urea will disappear. A random orientation of urea is found in the case of no-charge NPs.

3.2.3 Hardening of NP diffusion

In the previous discussions, we have analyzed the effects of urea and GdmCl on NP association in the framework of “direct interaction mechanism”. Although the direct binding of urea and Gdm+ dissociates the NP dimer in the solutions, this finding could not exclude the “indirect interaction mechanism” totally from the salting-effect. Thus it is valuable to justify the contributions from the indirect interaction to the PMFs shown in Figure 3-2. In the “indirect interaction mechanism”, changing of solvent structure is thought to be the dominant factor in the salting-effect. One reason is that the changes in solvent properties always happen along with the protein denaturation when adding urea^{15,94,95} or GdmCl^{93,96}. Obviously the solvent properties will affect the solute dynamics. In the last part, we would like to analyze the change of NP’s diffusion in different solvents. The diffusion coefficient of NPs is analyzed through calculating its mean square displacement (MSD). The simulation trajectories of ionized NP dimer with $d = 0.90nm$ were used, because the NP dimer is so stable that it can be treated as “one whole molecule”.

As shown in Figure 3-7, the movements of NP dimer in three kinds of denaturant solutions were greatly reduced compared to that in pure water. Numerically, the MSD constant in pure water is $0.57 \times 10^{-5}cm^2/s$. In 4M urea, the MSD constant is reduced to $0.21 \times 10^{-5}cm^2/s$. 4M GdmCl shows more significant effect, where MSD constant is further reduced to $0.14 \times 10^{-5}cm^2/s$. These phenomena are similar to experimental findings

Figure 3-7 Mean-square displacement of nano-particle dimers.

that protein dynamics is strongly hardened imposed by solvent during denaturation⁹². However, the dependence between the denaturing power and the hardening of NP dynamics is not so evident. Because in the case of 7M urea, the MSD constant is greatly reduced to $0.13 \times 10^{-5} \text{ cm}^2/\text{s}$. The PMF changes in three solutions do not linearly depend on NP dynamics. These results show that the change of solute dynamics may be only a consequence after adding denaturants, but not a cause to the denaturation.

The denaturation process is multifaceted. It may need further justifications that how much does the “indirect interaction model” contribute to denaturation. But generally, the “indirect interaction model” should be excluded when comparing the denaturant effects on hydrocarbons¹¹. For large hydrophobic polymers, urea and GdmCl act like *denaturant* which destabilizes the folded states^{97,98}. On the contrary, for small hydrocarbon like methane, urea and GdmCl can enhance the hydrophobic attraction, acting as *renaturant*^{67,98,99}. The dissociation of hydrophobic NPs and unfolding of hydrophobic polymers tend to increase their cavity. However, GdmCl and urea uniformly increase the surface tension of water, this contrasts to easier cavity formation in GdmCl/urea solvent¹⁰⁰. These phenomena hardly support the “indirect interaction model” of denaturation. Therefore, GdmCl and urea actualize their effect through direct interaction with the solute molecules. In other words, denaturant’s effect is solute-dependent.

3.3 Conclusion

To conclude, our results have captured the difference between urea and GdmCl in the dissociation of nano-particles. We found that GdmCl can dissociate both hydrophobic and hydrophilic solutes. Urea has weaker effects on these two kinds of solutes, especially for ionized species. Thus our results correctly reproduced the stronger denaturing power of GdmCl.

The stronger effects of GdmCl over urea can be summarized as following:

1. Gdm⁺ ion has a net charge; the binding of Gdm⁺ to a solute generates a local positively charged environment and introduces electrostatic repulsion between the coated solutes.
2. Gdm⁺ can form stacked dimer and tetramer, which form ladders on the surface of the no-charge particles which can effectively block particle aggregation.
3. Urea will self-aggregate through inter-urea hydrogen bonds. The urea molecules between solute molecules act like “glue” and prevent further separation of the solutes, especially for the ionized species.

Finally, by checking the change of PMF with respect to the solute dynamics, we believe that the effects from the “indirect interaction model” can be hardly linked to the denaturation power. The interesting observation that Gdm⁺/urea undergoes a parallel to perpendicular flipping from no-charged surface to the negative charged surface may be tested through experiments.

Chapter 4

Salting Effects of Protein Components in Aqueous NaCl and Urea Solutions

From the studies in chapter 3, urea and GdmCl show the ability to denature the associated states of NP dimers. The weakened attraction between NPs can be extrapolated as becoming better-solvated of the NPs from pure water to GdmCl and urea solutions. This kind of mechanistic proposal would be much more convincing if we could prove it with more realistic protein-based models for protein denaturation.

For protein denaturation in urea solution, a two-stage unfolding of the proteins has been proposed⁶⁸. In the first stage, urea molecules accumulate in the vicinity of the protein surface. These urea molecules replace water to “solvate” the protein surface and initiate protein unfolding. In the second stage, urea molecules penetrate into the interior of the protein, forming hydrogen bonds with backbone atoms and breaking the hydrophobic interaction. Thermodynamically, the urea-binding biases the folded/unfolded equilibrium towards unfolding and finally causes denaturation.

From another point of view, urea can be treated as a co-solutes and compared with simple ions. Ions’ position in the Hofmeister series^{30,31} reflects their salting effects on pro-

tein stability and solubility. For example, the order of anions is usually given as: $\text{SO}_4^{2-} > \text{HPO}_4^{2-} > \text{Cl}^- > \text{Br}^- > \text{I}^- > \text{ClO}_4^- > \text{SCN}^-$. Anions on the left decrease the solubility of hydrophobic molecules and precipitate proteins from solution (“salting-out”). By contrast, anions on the right increase the solubility of hydrophobic molecules (“salting-in”) and promote the denaturation of proteins²². One approach to explore the mechanisms of co-solutes effects on protein structure and stability is to determine the effects of co-solutes on models of protein components^{101,102}. The basic idea is that we assume protein unfolding in urea solution corresponds to a change of becoming better-solvated upon moving from pure water to urea solution. One key physical quantity relating to the solubility is the transfer/solvation free energy (ΔG). The solvation free energy cannot only be calculated with high accuracy in MD simulation combined with the thermodynamic integration method, but also can be decomposed into physical components, which can provide unambiguous evidence on the underlying mechanism.

Moreover, the decomposition of ΔG can help us to find the dominant forces between urea and protein which have been quite controversial. Because urea has a carbonyl group and two amide groups, it potentially can form hydrogen bonds and other electrostatic interactions when encountering side chain and backbone atoms^{67,71,103–105}. It is therefore not surprising that the action of urea denaturation was attributed to the direct hydrogen bonding⁷¹. And there are numerous experimental works showing that urea does form hydrogen bond with protein^{103,106}. However the direct hydrogen bonding was found to be not necessary to induce denaturation¹⁰⁷. Apart from the electrostatic interaction, van der Waals (vdW) or dispersion force is another driving force that mediates urea-protein interaction^{22,97}. In one recent study, neutral residues, rather than the charged ones, attract more urea in their surroundings¹⁰⁸. Also it is reported that mutation of charged residues to neutral ones can accelerate urea denaturation through computer simulations⁷⁰. Thus it is still an open question whether the denaturation is mainly induced by the electrostatic interaction/hydrogen bonding or by the vdW interaction.

Motivated by the above considerations, we explored protein denaturation through a comparative study of the salting effects from aqueous urea and NaCl solutions. NaCl electrolyte is taken into our consideration as a control system. Fifteen model molecules

were used to mimic protein backbone and side chains. The solvation free energies in these two electrolytes and those in pure water were used to evaluate the salting effect from co-solvents. Our study has successfully reproduced the salting-out effect of NaCl and salting-in effect of urea on these protein compounds. By careful analysis of the thermodynamic properties of the salting effects, the mechanism of urea-induced denaturation is revealed in an unambiguous and unique way.

4.1 Computational Methods

4.1.1 Proteinic models

The 14 neutral amino acid side chain analogs considered were shown in Table 4.2. Other than these, *N*-Methylacetamide (NMA) which is a secondary amide that mimics the protein backbone, was also considered. Topologies of the solute molecules were adopted from AMBER(ff03)⁸⁴. Both AMBER-derived atomic partial charges⁴³ and RED-generated partial charges (using R.E.D. tools⁸⁵ through *Gaussian03*⁸⁶ at the HF/6-31G* level) were used¹. The AMBER supplied urea model was used. For Na⁺ and Cl⁻ ions, it is well known that AMBER force fields usually induce salt crystallization well below their solubility limit. New van der Waals parameters for the Na⁺ and Cl⁻ (new-NaCl) ions developed by *In Suk Joun*⁵⁹ have also been used in this study. Our results show that the inaccuracies caused by the force fields do not impact the salting effects qualitatively.

4.1.2 Simulation details

All of the molecular dynamics (MD) simulations were performed with GROMACS³⁷, using the *NPT* ensemble at 1 atm and 300 K. SHAKE constraints³⁹ were applied to all bonds involving hydrogen atoms. The MD integration time step was 2 fs. The van der Waals (vdW) interactions were treated with a typical distance cutoff of 12 Å and a long range vdW correction was employed at distances greater than the cutoff. The electrostatic interactions, on the other hand, were handled with the Particle Mesh Ewald algorithm^{41,42}. The non-

¹The values of atomic charges can be found in Table 4.3 at the end of this chapter.

bonded interaction pair-list was updated every 10 *fs*. The periodic water box contained 1 solute molecule and 900 TIP3P³⁸ water in the pure water simulation. For the salting simulation, additional 35 NaCl or 70 urea molecules were added corresponding to molar concentrations of 2.05 M NaCl and 3.56 M urea, respectively.

Thermodynamic integration (TI)^{43–45} was used to compute the solvation free energy through computing the ensemble average of $dH/d\lambda$ at a number of values of λ (λ -windows). The electrostatic interactions were first decoupled linearly with 21 λ_C windows from 0.0 to 1.0 with a step length of 0.05. When turning off the Lennard-Jones (LJ) interactions, the soft-core potentials⁴⁶ were used to avoid poor convergence. In addition to the same 21 λ windows in the first step, 32 more λ_{LJ} windows have been added². The average accuracy of the calculated free energies in perturbations of LJ interactions is within 0.02 kcal/mol as shown in Table 4.1. At each λ value, a short (0.4 *ns*) simulation was first conducted for the system to pre-equilibrate, and followed by a 2 *ns* production run. Other than the free energy calculations, normal molecular dynamics simulations were also conducted for 200 *ns* for dynamics analysis. The statistical uncertainties of all the observables were calculated following the methods reported in section 2.1.2.

4.2 Results and Discussions

The solvation free energies in pure water (ΔG_{water}) were shown in Table 4.2. For most molecules, the values are within a range of $-10 \sim 3$ kcal/mol. The RED charges and AMBER charges give similar ΔG_{water} for most of the molecules examined, where RED charges show better consistency with the experiment values of PHE, HIE, ASN, TRP and TYR. Overall, the values are consistent with the experimental and previous theoretical results^{43,109,110}. When co-solutes were added, the solvation free energy change was defined as follows (in the case of NaCl)

$$\Delta\Delta G_{NaCl} = \Delta G_{NaCl} - \Delta G_{water}$$

²These λ_{LJ} are: 0.025, 0.075, 0.125, 0.175, 0.225, 0.275, 0.325, 0.375, 0.425, 0.475, 0.525, 0.575, 0.625, 0.675, 0.71, 0.72, 0.73, 0.74, 0.76, 0.77, 0.78, 0.79, 0.81, 0.82, 0.83, 0.84, 0.86, 0.87, 0.88, 0.89, 0.925, 0.975

Table 4.1 Free energies (in kcal/mol) in perturbations of van der Waals (LJ) interactions using 21 and 53 λ . The absolute difference and average value are shown for comparison

	21 λ	53 λ	difference
NMA	1.30	1.33	0.03
ALA	2.47	2.46	0.01
CYS	2.00	2.03	0.03
PHE	2.11	2.11	0.00
HIE	0.80	0.77	0.03
ILE	2.50	2.52	0.02
LEU	2.47	2.45	0.02
MET	2.15	2.16	0.01
ASN	1.21	1.21	0.00
GLN	1.28	1.33	0.05
SER	1.70	1.68	0.02
THR	1.82	1.80	0.03
VAL	2.46	2.46	0.00
TRP	1.12	1.13	0.01
TYR	1.54	1.61	0.07
		ave	0.02

Table 4.2 Model molecules used in this study and solvation free energies (in kcal/mol) in pure water in this study and the literatures

3-Letter code	1-Letter code	Model compound	AMBER charges	RED charges	Shirts ⁴³	EXP1 ¹⁰⁹	EXP2 ¹¹⁰
NMA	B	N-Methylacetamide	-8.20	-7.80	-	-	-
ALA	A	Methane	2.45	2.45	2.57	1.94	2.00
CYS	C	Methanethiol	0.06	-0.37	0.11	-1.24	-1.24
PHE	F	Toluene	0.54	-0.35	0.10	-0.76	-0.89
HIE	H	4-methylimidazole	-5.31	-7.46	-8.98	-10.27	-
ILE	I	Butane	2.47	2.52	2.84	2.15	2.08
LEU	L	Isobutane	2.41	2.41	2.72	2.28	2.32
MET	M	Methyl-ethylsulfide	0.22	0.28	0.91	-1.48	-
ASN	N	Acetamide	-6.97	-8.49	-7.80	-9.68	-9.71
GLN	Q	Propionamide	-10.75	-7.72	-7.69	-9.38	-
SER	S	Methanol	-4.61	-4.45	-4.37	-5.06	-5.11
THR	T	Ethanol	-3.40	-3.84	-3.83	-4.88	-5.01
VAL	V	Propane	2.33	2.46	2.69	1.99	1.96
TRP	W	Methylindole	-3.60	-4.76	-4.88	-5.88	-
TYR	Y	p-Cresol	-1.92	-4.90	-4.23	-6.11	-6.14

Figure 4-1 (a) Solvation free energy changes ($\Delta\Delta G_{NaCl}$) of 15 molecules from pure water to NaCl solution. Three different combinations of force fields were tested: AMBER derived molecule charges with AMBER NaCl (black), RED tool generated molecule charges with AMBER NaCl (red) and RED tool generated molecule charges with new-NaCl model (blue). (b) Solvation free energy changes ($\Delta\Delta G_{urea}$) from pure water to aqueous urea solution. AMBER derived molecule charges (black) and RED tool generated molecule charges (red) were tested.

According to this definition, a positive value of $\Delta\Delta G$ means worse solvated (salting-out effect), and a negative value means better solvated (salting-in effect).

4.2.1 NaCl salts out all 15 molecules

In order to check the sensitivity of the force fields used, we have performed three sets of simulations with force field combinations of: (1) AMBER partial charges of solute molecules with AMBER NaCl model, (2) RED partial charges of solute molecules with AMBER NaCl model, and (3) RED partial charges of solute molecules with new NaCl model. The results were shown in Figure 4-1(a). All the 15 molecules are uniformly salted-out by NaCl, which is independent of the force fields used. It is found that two sets of partial charges principally do not affect the salting effect (this phenomenon will be discussed later) when the same AMBER NaCl model is used. It is obvious that new-NaCl model results in stronger salting-out effect for all the molecules. Our data show that small and polar molecules like SER and ASN are relatively weakly salted-out, while large and non-polar molecules like LEU, PHE and TRP are most significantly salted-out. This is consistent with our previous studies of the salting-pattern of NaCl electrolyte presented in Chapter 2: the strength of the salting effect is correlated to the hydrophobicity of the solute molecules.

Figure 4-2 The linear relationship between the solvation free energy change in aqueous urea solution and the number of heavy atoms of the solute molecule.

4.2.2 Urea salts in all 15 molecules

In the case of urea, we have performed two sets of simulations with the force field combination: (1) AMBER partial charges of solute molecules with urea, (2) RED partial charges of solute molecules with urea. As shown in Figure 4-1(b), all the molecules are uniformly salted-in. In contrast to the case of NaCl, the strength of the salting-effect depends on the size of solutes rather than the chemical polarity. For example, ALA, SER and CYS are the smallest molecules and are weakly salted-in; while TYR and TRP are the largest molecules and are strongly salted-in. The dependence of $\Delta\Delta G_{urea}$ with respect to the number of heavy atoms of the solute molecules is shown in Figure 4-2. A linear correlation was clearly observed with fitting curve $y = 0.076 - 0.144x$. The linear correlation provides a hint that the salting effects are not mainly of electrostatic or hydrogen bonding nature, as will be elaborated more below. It needs emphasizing that the fitting curve does not go across zero when $x = 0$. This implies that a positive $\Delta\Delta G$ is expected for even tiny molecules. In fact, methane (an ALA analogue) is an exception and was found to be salted-out by urea in an early study¹¹. However, our calculation predicts a weak salting-in effect (-0.04 ± 0.02 kcal/mol) for methane. We realized that ΔG of methane in pure water is 2.45 ± 0.01 kcal/mol in our calculation, which is larger than the experimental value of ~ 2.0 kcal/mol^{109,110}. The methane model used here overestimates the positive solvation free energy. Further improve-

Figure 4-3 Radial distribution function (RDF) and cumulative RDF of urea around (a) NMA and (b) TRP. Cluster size of the urea binding patterns of (c) NMA and (d) TRP.

ment of the methane model is needed.

As shown in Figure 4-3(a) and (b), the binding shells of urea around NMA and TRP have first peak within distance 0.7 nm which contains about 3~5 urea molecules. Hence, the urea molecules within 0.7 nm to the solute molecules were treated as direct bound ureas. The orientation of an urea relative to the solute molecule can be described by the angle θ between the vector from the center of mass of NMA or TRP to the center of mass of an urea molecule and the normal vector of the urea plane. The distributions of $\cos(\theta)$ of the direct bound urea are shown in Figure 4-4(a). It is obvious that the orientations of the urea are generally random, with slightly preference of parallel oriented ($\cos(\theta) = 1$) binding. We have also generated the representative binding structures according to the following methods (taking NMA for example):

1. The whole trajectory (200 ns) was split into three trajectories. Each trajectory con-

Figure 4-4 (a) Binding orientations of urea molecules (within 0.70 nm) around the NMA and TRP molecules. θ is the angle between the vector from the center of mass of NMA or TRP to the center of mass of urea molecule and the normal vector of the urea plane. The representative binding patterns of urea around NMA and TRP are shown in (b) and (c), respectively. The structure snapshots were obtained from the top 25 highest populated clusters, where the clustering method is introduced in the main text.

tains one NMA and one urea molecule.

2. Three trajectories were combined together, which result in a 600 ns trajectory.
3. The binding pattern of the NMA-urea pair was clustered using the 600 ns trajectory.
4. As shown in Figure 4-3(c) and (d), the first 25 largest clusters cover $> 50\%$ of all the clusters found, so they are regarded as representative binding patterns.

In Figure 4-4(b) and (c), the highest populated binding patterns of urea around NMA and TRP were illustrated. One finds that the orientations of urea relative to the solute molecule are quite random, which is consistent with the angle distribution analysis as shown in Figure 4-4(a). Moreover, the positions of urea around solutes are also randomly distributed, which indicates that there is no favorable hydrogen-bond interaction. The preference of parallel-stacked binding pattern as suggested in Figure 4-4(a) is similar to the case of Gdm⁺ binding on hydrophobic nano-particles. The top 4 most salted-in molecules are TRP, TYR, PHE and HIE analogues, which are all aromatic residues and have planar structures. Thus a parallel bound mode can form larger number of contacts between urea and solute, compared to a perpendicular bound case.

Figure 4-5 (a) Correlation of differential Kirkwood-Buff integral with the values of $\Delta\Delta G$. (b) and (c) are radial distribution functions of Na^+ , Cl^- , urea and water around a TRP molecule.

4.2.3 Interpretation of salting effects by Kirkwood-Buff integral

The results of $\Delta\Delta G$ can be interpreted through the KB integral which describes the salting-effect purely through a statistical mechanics point of view^{48,49}. From our previous study in chapter 2, a direct relationship, $\Delta\Delta G \propto \Delta K_{sc} - \Delta K_{sw}$, among 121 model solutes in NaCl solution, was established¹¹¹, where ΔK_{sc} and ΔK_{sw} are changes of KB integral of co-solute (C) and water (W) around solute (S) from pure water to the NaCl solution, respectively.

KB integrals were calculated based on MD simulation trajectories (200 ns) using RED partial charges and the new-NaCl model. Pearson's correlation⁶¹ between $\Delta\Delta G$ and $\Delta K_{sc} - \Delta K_{sw}$ is shown in Figure 4-5(a), where a high correlation is observed at distances larger than 0.35 nm (NaCl solvent) and 0.70 nm (urea solvent), respectively. In the framework of "direct interaction mechanism" of salting-effect and from the salting pattern studies of NaCl in Chapter 2, the salting-out effect could be described as preferential binding of co-solute (NaCl in the current study) to water molecules; while the salting-in effect is caused by the preferential binding of co-solute (urea in the current study) to solute molecules. The radial distribution functions (RDF) of Na^+ , Cl^- , urea and water around TRP are shown in Figures 4-5(b) and (c). It is clear that around a TRP, there is a discernable water shell while Na^+ and Cl^- ions are repelled away. On the contrary, for a TRP in urea solution, urea accumulates nearby and water is repelled away. The correlation value does not change at larger distances, which implies that the modulation of water structures does not contribute to

Figure 4-6 Enthalpic (black) and entropic (red) contributions to solvation free energy changes from pure water to (a) NaCl solution and (b) urea solution. The order of the protein components is $|\Delta\Delta G|$ increasing (weak salting-effect to strong salting-effect).

the salting-effect. This phenomenon supports the “direct interaction mechanism” in which the preferential binding of urea results in denaturation.

4.2.4 Salting-effect is enthalpy driven

In the frame work of the “direct interaction mechanism” there are still controversies regarding the main driving force for the denaturation action of urea^{68,105}. Therefore it is valuable to decompose the free energy changes into enthalpic and entropic contributions. For a solvation process, the enthalpy change (ΔH) is defined as the change of potential energy from the de-coupled state (isolated solute with solvent, $\lambda_{LJ} = 0$, $\lambda_C = 0$) to the coupled state (solvated complex, $\lambda_{LJ} = 1$, $\lambda_C = 1$). The potential energies were computed by averaging over the 200 ns trajectories of the respective states. The entropy change of the solvation process, ΔS , is derived from the relation $\Delta G = \Delta H - T\Delta S$.

The relative enthalpy change from pure water to NaCl solvent is defined as $\Delta\Delta H_{NaCl} = \Delta H_{NaCl} - \Delta H_{water}$ (similar formulation for entropy). The values of $\Delta\Delta H$ and $T\Delta\Delta S$ are shown in Figure 4-6. For the case of NaCl, all the $\Delta\Delta H$ are positive which indicates that the solvation is enthalpically unfavorable. On the other hand, all $T\Delta\Delta S$ values are positive, thus playing a counteracting role on the free energy. The absolute values of enthalpy increases overwhelm those of entropy increases. Hence for NaCl, the salting-out effect is clearly enthalpy-driven¹¹².

In the case of urea, the $\Delta\Delta H$ of all molecules, except ALA, are negative which indicates that the urea solution stabilizes the solvation state energetically. Most of the entropy decreases (negative $T\Delta\Delta S$), however, with the absolute values less than that of the enthalpy. Overall, the salting-effect of urea is also enthalpy-driven⁹⁷, similar to that of NaCl. It is interesting to find that most $T\Delta\Delta S$ have the same sign with $\Delta\Delta H$, which demonstrates the well known entropy-enthalpy compensation^{113–115}. This compensation is commonly observed in the thermodynamic analysis of proteins, nucleic acids - ligands interaction.

The entropy-enthalpy compensation in solvation was reported in a remarkable study by Yu and Karplus in 1988¹¹⁶, where the free energy, energy, and entropy of solvation were analyzed. The authors found that the solvent-solvent interaction term, which contributes the so-called *cavity* term to the energy, is cancelled exactly by a corresponding term in the entropy. Thus only averages over the solute-solvent interaction energy contribute to the solvation free energy. Thus for the case of urea denaturation, it is valuable to examine the dominant force that drives the urea-solute binding. The the solute-solvent interaction in urea solution is further decomposed into parts from the van der Waals and electrostatic interactions as shown in Figure 4-7. It is clear that urea salts in the model solutes mainly through the van der Waals interactions, and the electrostatic interactions only play a very minor role, which is consistent with previous studies on the protein lysozyme in 8M urea⁶⁸ and carbon nanotubes in 8M urea¹¹⁷. This phenomenon challenges a hydrogen-bonding driven mechanism in urea induced protein denaturation. It is worth emphasizing that although urea did not engage in the hydrogen-bond network within protein components, the hydrogen bond in peptide backbones as in alpha-helix and beta-strand, may also be perturbed because NMA was found to be better solvated.

4.3 Conclusion

To conclude, we have successfully reproduced the salting-out effect of NaCl and salting-in effect of urea on protein compounds. The capability of urea to better solvate protein backbone and side chain analogues can be extrapolated to explain a protein's denaturation in aqueous urea solution. By careful analysis of the thermodynamic properties of the salt-

Figure 4-7 Decomposition of the solute-solvent interaction in urea solvent into the van der Waals (LJ) and the electrostatic terms.

ing effects, the mechanism of urea induced denaturation is revealed in an unambiguous and unique way. For NaCl, polar molecules are less affected while large and hydrophobic molecules are significantly salted-out. For urea, the strength of salting-in effect is linearly dependent on the number of heavy atoms, and the salting-out effect is expected for smaller molecules like methane. For both NaCl and urea solutions, the salting-effects are found to be enthalpy driven. The enthalpy change of solvation can be decomposed into two terms: an energy-favorable solute-solvent interaction term and an energy-unfavorable solvent-solvent interaction term. The advantage of urea in denaturing protein is that the decreasing of solute-solvent interaction energy term counterbalances the increasing of solvent-solvent interaction energy term, which finally facilitates the solvation. The van der Waals interactions between urea molecules and protein compounds are found to be attractive forces, which challenges a hydrogen-bonding driven mechanism for the urea denaturation proposed previously.

Table 4.3 List of AMBER derived charges and RED tool generated charges and absolute difference

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
NMA	HH31	0.07601	0.08090	0.00489
	CH3	-0.19026	-0.29430	0.10404
	HH32	0.07601	0.08090	0.00489
	HH33	0.07601	0.08090	0.00489
	C	0.51240	0.69810	0.18570
	O	-0.55017	-0.58500	0.03483
	N	-0.42389	-0.50310	0.07921
	H	0.29011	0.31990	0.02979
	CH3	-0.05429	-0.28000	0.22571
	HH31	0.06269	0.13390	0.07121
	HH32	0.06269	0.13390	0.07121
	HH33	0.06269	0.13390	0.07121
ALA	HA	0.07743	0.07300	0.00443
	CB	-0.30972	-0.29200	0.01772
	HB1	0.07743	0.07300	0.00443
	HB2	0.07743	0.07300	0.00443
	HB3	0.07743	0.07300	0.00443
CYS	HA	0.14654	0.09460	0.05194
	CB	-0.34371	-0.13890	0.20481
	HB2	0.14654	0.09460	0.05194
	HB3	0.14654	0.09460	0.05194
	SG	-0.28518	-0.35490	0.06972
	HG	0.18927	0.21000	0.02073
PHE	HA	0.06099	0.06310	0.00211
	CB	-0.18338	-0.18980	0.00642
	HB2	0.06099	0.06310	0.00211

Continued on next page

Table 4.3 – Continued from previous page

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	HB3	0.06099	0.06310	0.00211
	CG	0.02131	0.09970	0.07839
	CD1	-0.08311	-0.15600	0.07289
	HD1	0.09847	0.13140	0.03293
	CE1	-0.15697	-0.18100	0.02403
	HE1	0.12373	0.14480	0.02107
	CZ	-0.09982	-0.10530	0.00548
	HZ	0.11468	0.12800	0.01332
	CE2	-0.15697	-0.19080	0.03383
	HE2	0.12373	0.14760	0.02387
	CD2	-0.08311	-0.14720	0.06409
	HD2	0.09847	0.12930	0.03083
HIE	HA	0.05461	0.11970	0.06509
	CB	-0.20344	-0.43260	0.22916
	HB2	0.05461	0.11970	0.06509
	HB3	0.05461	0.11970	0.06509
	CG	0.27841	0.43060	0.15219
	ND1	-0.42333	-0.56050	0.13717
	CE1	0.02596	0.16360	0.13764
	HE1	0.12683	0.13270	0.00587
	NE2	-0.09798	-0.28620	0.18822
	HE2	0.26687	0.33070	0.06383
	CD2	-0.29756	-0.34100	0.04344
	HD2	0.16041	0.20360	0.04319
ILE	HA	0.06166	-0.00770	0.06936
	CB	-0.09947	0.03580	0.13527
	HB	0.06166	-0.00770	0.06936

Continued on next page

Table 4.3 – *Continued from previous page*

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	CG2	-0.12999	-0.05670	0.07329
	HG21	0.03023	0.01210	0.01813
	HG22	0.03023	0.01210	0.01813
	HG23	0.03023	0.01210	0.01813
	CG1	0.02223	0.03670	0.01447
	HG12	0.01155	-0.00830	0.01985
	HG13	0.01155	-0.00830	0.01985
	CD1	-0.10125	-0.05280	0.04845
	HD11	0.02379	0.01090	0.01289
	HD12	0.02379	0.01090	0.01289
	HD13	0.02379	0.01090	0.01289
LEU	HA	0.05253	0.05860	0.00607
	CB	-0.23859	-0.27760	0.03901
	HB2	0.05253	0.05860	0.00607
	HB3	0.05253	0.05860	0.00607
	CG	0.19198	0.34610	0.15412
	HG	0.00082	-0.04290	0.04372
	CD1	-0.12304	-0.28000	0.15696
	HD11	0.02238	0.05930	0.03692
	HD12	0.02238	0.05930	0.03692
	HD13	0.02238	0.05930	0.03692
	CD2	-0.12304	-0.26610	0.14306
	HD21	0.02238	0.05560	0.03322
	HD22	0.02238	0.05560	0.03322
	HD23	0.02238	0.05560	0.03322
MET	HA	0.04884	0.01150	0.03734
	CB	-0.07533	-0.01120	0.06413

Continued on next page

Table 4.3 – Continued from previous page

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	HB2	0.04884	0.01150	0.03734
	HB3	0.04884	0.01150	0.03734
	CG	-0.20817	0.03540	0.24357
	HG2	0.12423	0.05130	0.07293
	HG3	0.12423	0.05130	0.07293
	SD	-0.21196	-0.29760	0.08564
	CE	-0.28469	-0.08780	0.19689
	HE1	0.12839	0.07470	0.05369
	HE2	0.12839	0.07470	0.05369
	HE3	0.12839	0.07470	0.05369
ASN	HA	0.04332	0.10030	0.05698
	CB	-0.11509	-0.35420	0.23911
	HB2	0.04332	0.10030	0.05698
	HB3	0.04332	0.10030	0.05698
	CG	0.58352	0.83370	0.25018
	OD1	-0.52681	-0.61100	0.08419
	ND2	-0.78174	-0.99580	0.21406
	HD21	0.35508	0.41320	0.05812
	HD22	0.35508	0.41320	0.05812
GLN	HA	0.03100	0.02310	0.00790
	CB	-0.10686	-0.07150	0.03536
	HB2	0.03100	0.02310	0.00790
	HB3	0.03100	0.02310	0.00790
	CG	-0.02026	-0.02500	0.00474
	HG2	0.03079	0.01500	0.01579
	HG3	0.03079	0.01500	0.01579
	CD	0.66781	0.77790	0.11009

Continued on next page

Table 4.3 – *Continued from previous page*

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	OE1	-0.62848	-0.59450	0.03398
	NE2	-0.88341	-1.01900	0.13559
	HE21	0.40831	0.41640	0.00809
	HE22	0.40831	0.41640	0.00809
SER	HA	0.04008	0.02470	0.01538
	CB	0.07381	0.16450	0.09069
	HB2	0.04008	0.02470	0.01538
	HB3	0.04008	0.02470	0.01538
	OG	-0.64031	-0.66580	0.02549
	HG	0.44626	0.42720	0.01906
THR	HA	0.04469	-0.05350	0.09819
	CB	0.10350	0.38980	0.28630
	HB	0.04469	-0.05350	0.09819
	CG2	-0.17649	-0.08650	0.08999
	HG21	0.06002	0.02590	0.03412
	HG22	0.06002	0.02590	0.03412
	HG23	0.06002	0.02590	0.03412
	OG1	-0.60193	-0.68900	0.08707
	HG1	0.40548	0.41500	0.00952
VAL	HA	-0.11567	-0.01560	0.10007
	CB	0.46548	0.08770	0.37778
	HB	-0.11567	-0.01560	0.10007
	CG1	-0.09013	-0.06220	0.02793
	HG11	-0.00898	0.01130	0.02028
	HG12	-0.00898	0.01130	0.02028
	HG13	-0.00898	0.01130	0.02028
	CG2	-0.09013	-0.06120	0.02893

Continued on next page

Table 4.3 – *Continued from previous page*

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	HG21	-0.00898	0.01100	0.01998
	HG22	-0.00898	0.01100	0.01998
	HG23	-0.00898	0.01100	0.01998
TRP	HA	0.06542	0.07010	0.00468
	CB	-0.17370	-0.20630	0.03260
	HB2	0.06542	0.07010	0.00468
	HB3	0.06542	0.07010	0.00468
	CG	-0.09980	-0.01060	0.08920
	CD1	-0.17405	-0.22200	0.04795
	HD1	0.17063	0.20390	0.03327
	NE1	-0.29843	-0.37640	0.07797
	HE1	0.32238	0.35900	0.03662
	CE2	0.14152	0.11720	0.02432
	CZ2	-0.21070	-0.22160	0.01090
	HZ2	0.12551	0.14590	0.02039
	CH2	-0.13302	-0.13410	0.00108
	HH2	0.11947	0.13830	0.01883
	CZ3	-0.16405	-0.20600	0.04195
	HZ3	0.11925	0.14290	0.02365
	CE3	-0.15399	-0.18600	0.03201
	HE3	0.12308	0.15050	0.02742
	CD2	0.08964	0.09500	0.00536
TYR	HA	0.01915	0.05630	0.03715
	CB	-0.09424	-0.16880	0.07456
	HB2	0.01915	0.05630	0.03715
	HB3	0.01915	0.05630	0.03715
	CG	0.11260	0.09630	0.01630

Continued on next page

Table 4.3 – *Continued from previous page*

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	CD1	-0.18346	-0.19120	0.00774
	HD1	0.13271	0.16430	0.03159
	CE1	-0.18182	-0.32340	0.14158
	HE1	0.13730	0.16230	0.02500
	CZ	0.20628	0.36940	0.16312
	OH	-0.42124	-0.61950	0.19826
	HH	0.32969	0.43980	0.11011
	CE2	-0.18182	-0.21910	0.03728
	HE2	0.13730	0.17080	0.03350
	CD2	-0.18346	-0.20200	0.01854
	HD2	0.13271	0.15220	0.01949
LYS	HA	0.04521	0.04220	0.00301
	CB	-0.08538	-0.08600	0.00062
	HB2	0.04521	0.04220	0.00301
	HB3	0.04521	0.04220	0.00301
	CG	0.03334	-0.00020	0.03354
	HG2	0.01021	0.03360	0.02339
	HG3	0.01021	0.03360	0.02339
	CD	-0.04784	-0.05850	0.01066
	HD2	0.07071	0.04540	0.02531
	HD3	0.07071	0.04540	0.02531
	CE	-0.06997	0.03880	0.10877
	HE2	0.11952	0.09210	0.02742
	HE3	0.11952	0.09210	0.02742
	NZ	-0.25036	-0.28060	0.03024
	HZ1	0.29458	0.30570	0.01112
	HZ2	0.29456	0.30600	0.01144

Continued on next page

Table 4.3 – Continued from previous page

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	HZ3	0.29456	0.30600	0.01144
ARG	HA	0.02801	0.03590	0.00789
	CB	0.01537	-0.09220	0.10757
	HB2	0.02801	0.03590	0.00789
	HB3	0.02801	0.03590	0.00789
	CG	0.01245	-0.01510	0.02755
	HG2	0.00296	0.06300	0.06004
	HG3	0.00296	0.06300	0.06004
	CD	0.12633	-0.03550	0.16183
	HD2	0.06815	0.10170	0.03355
	HD3	0.06815	0.10170	0.03355
	NE	-0.46489	-0.48980	0.02491
	HE	0.32627	0.36970	0.04343
	CZ	0.56554	0.75430	0.18876
	NH1	-0.68575	-0.80590	0.12015
	HH11	0.39105	0.43450	0.04345
	HH12	0.39105	0.43450	0.04345
	NH2	-0.68577	-0.89420	0.20843
	HH21	0.39105	0.45130	0.06025
	HH22	0.39105	0.45130	0.06025
ASP	HA	-0.01484	0.00420	0.01904
	CB	-0.23989	-0.21580	0.02409
	HB2	-0.01484	0.00420	0.01904
	HB3	-0.01484	0.00420	0.01904
	CG	0.74519	0.88570	0.14051
	OD1	-0.73039	-0.83690	0.10651

Continued on next page

Table 4.3 – *Continued from previous page*

3-Letter code	Atom name	AMBER-Q	RED-Q	diff
	OD2	-0.73039	-0.84560	0.11521
GLU	HA	-0.00353	-0.01870	0.01517
	CB	-0.06433	-0.06190	0.00243
	HB2	-0.00353	-0.01870	0.01517
	HB3	-0.00353	-0.01870	0.01517
	CG	-0.03391	0.10830	0.14221
	HG2	-0.00413	-0.07560	0.07147
	HG3	-0.00413	-0.07560	0.07147
	CD	0.76519	0.82150	0.05631
	OE1	-0.82406	-0.84650	0.02244
	OE2	-0.82404	-0.81410	0.00994

Chapter 5

Sequence-Specific Mg^{2+} - DNA Interactions

Due to the polyelectrolyte nature of nucleic acids, their structures and functions are sensitively correlated with the “counterion atmosphere”^{118–122}. Since the DNA backbone is negatively charged, the close contacts between DNA duplexes are strongly repulsive. The counterion atmosphere around DNA plays a role in the electrostatic screening to reduce the repulsion. The direct evidence of intimate interactions between cations and nucleic acids came from crystal structures where dehydrated sodium ions (Na^+) were found in the minor groove of DNA^{123,124}. Compared with monovalent ions, divalent ions (like Mg^{2+}) have stronger interactions with nucleic acids^{125,126} and are critical in genomic DNA folding¹²⁷. DNA structure and dynamics are fundamental in maintaining genome stability¹²⁸.

Ion-DNA interactions are highly dynamic which makes it difficult for experiments to provide a microscopic tempo-spatial description. On the other hand, computer molecular dynamics (MD) simulations based on potential energy functions are effective in deciphering the interaction nature of ions and water with DNA^{129–133}. The advantage of MD simulations is that the simulations can describe not only how ions screen the electrostatic interaction, i.e., the “passive” screening effects; but also the “active” effects of ions on DNA dynamics at the atomic details. So far most studies were targeted on monovalent ions, such as Na^+ and K^+ ions^{130,134–137}. For divalent ions, such as Mg^{2+} ions, the detailed interactions with

DNA obtained from MD simulations are scarcely documented^{28,130}.

One reason for the limited MD studies of DNA in the presence of Mg^{2+} , is that Mg^{2+} ions could form polar covalent bonds with oxygen atoms on the phosphate group of DNA. The chemical polar bond formation process is beyond the scope of classic MD simulations. For most Mg^{2+} ions there is no direct evidence of polar bond formation with the DNA duplex¹³⁸. As a result, directly polar bonded, or “chelated” Mg^{2+} ions only constitute a small number of observed ions¹³⁹. In fact, Mg^{2+} ions have a closed shell electron structure and a stable water coordination shell in solvent. They interact with DNA primarily through electrostatic and van der Waals dispersion forces. These fully hydrated mobile ions contribute a dominant role in stabilization of RNA structure resolved by thermodynamic analysis¹⁴⁰.

Based on the above considerations, we performed classical MD simulations in the presence of explicit water and counterions, to explore the atomic interaction picture of Mg^{2+} with the DNA duplex and the effects of Mg^{2+} ions on the DNA structure and dynamics, in comparison with the DNA system in the presence of Na^+ ions.

5.1 Computational Methods

5.1.1 Simulation details

Two different DNA models have been considered:

In the first model, one isolated double-stranded 23-mer DNA (GGCGGCGGCGCG-GCGTTTTTTGG)¹⁴¹ was solvated in a $10 \times 10 \times 10 \text{ nm}^3$ water box. Two simulations in different electrolytes have been performed: (1) one simulation contained 67 Na^+ , 23 Cl^- and 32007 water (referred to as Na-DNA); (2) the other simulation contained 29 Na^+ , 19 Mg^{2+} , 23 Cl^- and 31980 water (referred to as Mg-DNA).

In the second model, periodic DNA double strands with different sequences were studied. The system contained one double-stranded 10-mer DNA solvated in a $7 \times 7 \times \sin 60^\circ \times 3.4 \text{ nm}^3$ hexagonal water box. DNA duplex is infinite as each strand is covalently bonded to its image over the periodic boundary in the z direction. In order to compare the sequence-specific ions-DNA attraction, four periodic-DNAs (abbreviated pDNA in the lat-

Table 5.1 Simulation system content

System name	Simulation box content					Cation's concentration [M/L]
	DNA	Na ⁺	Mg ²⁺	Cl ⁻	H ₂ O	
Na-DNA	23-mer DNA	67	0	23	32007	0.083
Na-pGG	dG ₁₀ ·dC ₁₀	20	0	0	4314	0.165
Na-pGC	d(GC) ₅ ·d(CG) ₅	20	0	0	4323	0.177
Na-pAA	dA ₁₀ ·dT ₁₀	20	0	0	4309	0.195
Na-pAT	d(AT) ₅ ·d(TA) ₅	20	0	0	4317	0.166
Mg-DNA	23-mer DNA	29	19	23	31980	0.041(Na) / 0.017(Mg)
Mg-pGG	dG ₁₀ ·dC ₁₀	0	20	20	4314	0.185
Mg-pGC	d(GC) ₅ ·d(CG) ₅	0	20	20	4323	0.194
Mg-pAA	dA ₁₀ ·dT ₁₀	0	20	20	4309	0.155
Mg-pAT	d(AT) ₅ ·d(TA) ₅	0	20	20	4317	0.173

er discussions) with different sequences were simulated: (1) polyG·polyC, named as pGG (periodic polyG), (2) polyA·polyT, named as pAA, (3) poly[dGdC]·poly[dCdG], named as pGC (periodic poly[dGdC]) and (4) poly[dAdT]·poly[dTdA], named as pAT. Other than DNA, the simulation box contains 4000 water molecules, 20 Mg²⁺ and 10 Cl⁻ (in the Mg-pDNA case) or 20 Na⁺ (in the Na-pDNA case).

The molar concentration of the ions was evaluated from the solution far from the DNA atoms¹⁴². The ion concentration, together with the contents of all the simulation systems is listed in Table 5.1.

The recently modified AMBER(parmbsc0) force-field¹⁴³ was used to model the DNA molecules. The initial coordinates of DNA models were B-form generated with nucleic acid builder (*nab*)¹⁴⁴. For Na⁺, Cl⁻ and Mg²⁺, the CHARMM parameters optimized by Roux and coworkers¹⁴⁵ were used and are listed in Table 5.2. For water, we used the TIP3P³⁸ water model. All simulations were carried out using the GROMACS³⁷. The systems were first minimized for 1000 steps followed by a pre-equilibration of 100 ps with the atoms of DNA fixed. Data production simulations were performed for 100ns for all our models. All of the simulations were performed within the NPT ensemble (300K, 1atm). SHAKE constraints³⁹ were applied to all bonds involving hydrogen atoms. Van der Waals interactions were considered with a distance cutoff of 10 Å. Electrostatic interactions were

Table 5.2 CHARMM force field parameters of Na^+ and Mg^{2+} used in our simulations

ions	mass	σ (nm)	ε (kJ/mol)
Na^+	22.98977	2.42992×10^{-1}	1.962296×10^{-1}
Mg^{2+}	24.30000	2.11143×10^{-1}	6.276000×10^{-2}

treated using the Particle Mesh Ewald (PME) algorithm^{41,42}. The movement integration time was 2 fs. The non-bonded interaction pair-list was updated every 10 fs.

5.1.2 Covariance analysis and conformational entropy

Covariance analysis, also called principal component analysis (PCA) or essential dynamics¹⁴⁶, can find correlated motions in order to represent the conformational dynamics of a system in terms of a few “principal” modes. The basic idea is that the internal motions are correlated and represented by the covariance matrix C with elements

$$C_{ij} = \left\langle M_{ii}^{\frac{1}{2}} (x_i - \bar{x}_i) M_{jj}^{\frac{1}{2}} (x_j - \bar{x}_j) \right\rangle \quad (5.1)$$

where M is a diagonal matrix containing the masses of the atoms. C is a symmetric $3N \times 3N$ matrix, which can be diagonalized with an orthonormal transformation matrix R :

$$R^T C R = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{3N}) \quad \text{where } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{3N} \quad (5.2)$$

The columns of R are the eigenvectors, also called principal or essential modes. The eigenvectors and eigenvalues describe the modes of collective motions and their amplitudes. The trajectory can be projected on the principal modes to give the principal components $p_i(t)$:

$$p(t) = R^T M^{\frac{1}{2}} (x(t) - \bar{x}) \quad (5.3)$$

The eigenvalue λ_i is the mean squared fluctuation of principal component i . The first few principal modes often describe collective, global motions in the system.

The conformational entropy is calculated by the method introduced by *J. Schlitter*¹⁴⁷.

It is based on the approximation that each degree of freedom can be represented by a quasi-harmonic oscillator with vibration frequency, ω . The overall translational and rotational motions do not contribute to the conformational entropy, thus degrees of freedom of these two motions are removed, corresponding to 6 eigenvalues that are very close to zero. Only the $3N - 6$ nonzero values are substituted into the formula for the conformational entropy¹⁴⁸:

$$S = k \sum_i^{3N-6} \frac{\frac{\hbar\omega_i}{kT}}{e^{\frac{\hbar\omega_i}{kT}} - 1} - \ln(1 - e^{-\frac{\hbar\omega_i}{kT}}) \quad (5.4)$$

where T is the temperature and k is the Boltzmann constant, ω is the quasi-harmonic frequency which can be calculated from the eigenvalues λ_i obtained from the PCA analysis:

$$\omega_i = \sqrt{kT/\lambda_i} \quad (5.5)$$

Under this definition, the conformational entropy can be used to describe the global dynamics of DNA, where lower conformational entropy corresponds to more rigid DNA having smaller amplitudes of motion.

5.2 Results and Discussions

5.2.1 Solvation structure of free ions

First, we investigate the solvation patterns of Mg^{2+} and Na^+ . The radial distribution function (RDF) of water oxygen atoms (OW) around the two cations are shown in Figure 5-1(a) (solid line). Clearly, there is a distinct first solvation shell (FSS) at a distance around 2 Å for Mg^{2+} (red solid line) which is consistent with the findings of X-ray diffraction studies¹⁴⁹. The second solvation shell (SSS) near 4.3 Å is weak. The dashed line in this figure gives the cumulative number of water molecules inside the first solvation shell. It is clear that six water molecules constitute the FSS of Mg^{2+} . One configuration of Mg^{2+} with bound water molecules is shown in the inset: six water molecules form an octahedral shape with all hydrogen atoms pointing outward, so that the $\text{Mg}^{2+} + 6\text{H}_2\text{O}$ complex can function

Figure 5-1 Radial distribution function (solid line) of water oxygen and cumulative number of water (dashed line) around Na⁺ (black colored) and Mg²⁺ (red colored) ions. The inserts are (a) snapshot of the spatial configuration of six water molecules bound to Mg²⁺ and (b) distribution of ion-DNA distance.

as a hydrogen bond donor. The six surrounding water molecules are found to be very stable and did not exchange with bulk water during the 100 ns simulations. The nuclear magnetic resonance study showed that the mean residence time of water molecules bound to Mg²⁺ is $1.5\mu s$ ¹⁵⁰.

For Na⁺, the FSS (black solid line) is around 2.3 Å which is close to the value obtained from a study using recently optimized Na⁺ parameters⁵⁹. Although there are also six water constituting the first solvation shell of Na⁺, the height of the peak is left-shifted and reduced to one third of that of Mg²⁺. This means that the coordination sphere of Na⁺ is relatively loose and less stable than that of Mg²⁺, because of the weaker electrostatic attractions from monovalent Na⁺. This coordination difference between Na⁺ and Mg²⁺ will affect their interactions with DNA, which will be discussed later.

5.2.2 Results of isolated DNA model

Ion-DNA binding pattern

The distinct FSS of Mg²⁺ and Na⁺ ions will affect their binding patterns when encountering the DNA surface. The cation's binding behavior was monitored through calculating the minimal distance between each Na⁺/Mg²⁺ ion and the atoms of DNA (ion-DNA distance). The distributions of the ion-DNA distance are shown in Figure 5-1(b). For Na⁺ ions, there

Figure 5-2 Time evolution of number of cations on DNA surface: (a) Na^+ ions in Na-DNA system, (b) Na^+ (black line) and Mg^{2+} ions (red line) in Mg-DNA system. (c) Time evolution of root mean squared deviation (RMSD) of DNA in two simulations with respect to a reference structure of a standard B-DNA; (d) The number of Watson-Crick hydrogen bonds (H-bond) in two simulations.

are two favorable binding positions: one is around $2.2 \text{ \AA}$ and the other is $4 \text{ \AA}$ from the DNA surface. The small distance of the former position means that Na^+ ions form direct contacts with DNA by discarding some water in its FSS. For the latter position, a Na^+ ion interacts with DNA through its FSS. For Mg^{2+} ions, there is only one obvious bound distance at $4 \text{ \AA}$. Thus, Mg^{2+} can only interact with DNA through its FSS. This is not surprising because the FSS of Mg^{2+} is more stable than that of Na^+ . When the distance is larger than $5 \text{ \AA}$, the ion distribution curve is featureless. This indicates that beyond $5 \text{ \AA}$ the ions are almost free and there is no obvious binding behavior.

Charge-neutralization of the DNA

Cations will accumulate in the vicinity of DNA because of the Coulombic attraction. We have monitored the time evolution of bound ions on DNA as shown in Figure 5-2(a) and

(b). The definition of “bound ions” is made solely from structural point of view. Based on the distribution function in Figure 5-1(b), any ion is supposed to be bound if the ion-DNA distance is less than 5 Å. For Na-DNA system (Figure 5-2(a)), in the beginning of the simulation, the positions of the ions were assigned randomly in the water box, there were only ~ 5 Na^+ bound to DNA. As the simulation went on, more Na^+ ions approached the DNA surface. At $t=10$ ns, the number reached a plateau of around 20. During the rest of the simulation the number of Na^+ was nearly constant, albeit with considerable fluctuations. For the Mg-DNA system the interaction picture is shown in Figure 5-2(b). After a 10 ns simulation, the number of bound Mg^{2+} increased to a plateau around 12. At the same time, the number of Na^+ was reduced to a value around 5. Evidently, Mg^{2+} ions are stronger competitors for the DNA surface than Na^+ ions.

An important role of the counterions is to neutralize the negative charge of the phosphate groups. The close contacts of the bound ions with DNA render their neutralizing effects to full strength. The neutralization-fraction of DNA charge can be defined as:

$$r = \frac{\sum_i Z_i \times N_i^b}{Q_{DNA}} \quad (5.6)$$

where Z_i , N_i^b , and Q_{DNA} are the valence, the number of bound ions of species i , and the charge of DNA respectively. The average r value for the DNA duplex (discarding the two terminal base pairs) is calculated to be 50% (Na-DNA) and 71% (Mg-DNA), respectively.

DNA double helix stability

The structural stability of the DNA double helix was checked through calculating the time evolution of the root mean squared deviation (RMSD) with respect to a standard B-form DNA (Figure 5-2(c)) and the number of Watson-Crick hydrogen bonds (H-bond) (Figure 5-2(d)). During both simulations of Na-DNA and Mg-DNA, the DNA structures are stable. The RMSD of the Mg-DNA case is slightly smaller than that of the Na-DNA case. The H-bonds are very well maintained and less affected by the electrolyte environment.

Figure 5-3 Stereo views of counterions binding map around DNA duplex. Left side represents the Na-DNA system: (a) Na^+ in minor groove, (b) Na^+ in major groove; Right side represents the Mg-DNA system: (c) Mg^{2+} in major groove, (d) Mg^{2+} on phosphate group at A-tract region.

Counterion clouds around DNA

Inspection of the ions binding sites around the DNA gives us a detailed global view of ion-DNA interaction. In order to compute the spatial density map of ions around DNA, the simulation box was divided into three-dimensional grids with grid length of 1 Å and the occupancy of ions in each grid cell was calculated. The yellow and green balls as shown in Figure 5-3 represent grid ion density > 10 M for Na^+ in the Na-DNA case and > 5 M for Mg^{2+} in the Mg-DNA case, respectively. For the Na-DNA system, Na^+ ions can occupy the minor groove of DNA in a non-specific and diffuse way. As shown in Figure 5-3(a), one Na^+ with 5 coordinated water molecules goes deep into the minor groove. The minor groove of the G·C base has three H-bond acceptors: oxygen atoms ($\text{O4}'$) of the 2-deoxyribose sugar, and O2/N3 atoms of the base pairs. In the major groove of the G·C base, N7 and O6 of the guanine base serve as hydrogen bond acceptors. We found that Na^+ can also lose several coordinated water molecules and form direct electrostatic interactions with these negatively charged atoms. As shown in Figure 5-3(b), the distance between Na^+ and the N7 atom is 2.82 Å, which is a direct contact.

For the Mg-DNA system, one distinct feature is that the hydrated Mg^{2+} cannot penetrate into the minor groove because the $\text{Mg}^{2+}+6\text{H}_2\text{O}$ complex is too big for the minor groove to accommodate it. Mg^{2+} only appears near the major groove of the GC-rich segment (Figure 5-3(c)). Most of the bound Mg^{2+} ions exist near the phosphate group because the phosphate groups are highly negatively charged. We also observed that at the A-tract segment (base 16-21) of Mg-DNA, some Mg^{2+} ions are bridging two phosphate groups from the opposite strands. The binding pattern is shown in Figure 5-3(d): three coordinated water molecules of Mg^{2+} form hydrogen bonds with O1P and O2P of phosphate groups from two strands, which facilitates a very stable binding. This implies that the distance between the phosphate backbone from the two strands are smaller than that of GC-rich segments, meaning that the A-tract region should have a much narrower minor groove width (discussed below).

From the 3D density map, one can find that the distribution of ions is dependent on the DNA sequence. The ion distribution can be quantitatively described in detail by the occupancy of bound ions at each base pair (ions within a distance of 5 Å to the atoms

Figure 5-4 The number of bound ions projected on each base pair. The x axis indicates base pair sequence number (a-d). Base pair atoms used to define bound ions in the major groove and minor groove are shown in the inset of (d).

Figure 5-5 Width of (a) major groove and (b) minor groove of isolated DNA model. (c) Local bending angles of hex-basepair segments. A monitoring window containing six-basepair slides along basepair 2 to 22 of the isolated DNA duplex, resulting in 16 overlapping hex-basepair segments.

labeled in the inset of Figure 5-4(d) are treated as major groove-ions or minor groove-ions). Generally, Na^+ ions do not have distinctly favored base positions. The occupancy percentages of the major groove (Figure 5-4(a)) are slightly larger than those of the minor groove (Figure 5-4(b)) due to the fact that the major groove is more exposed to solvent. At the A-tract region, base pairs 16-19, Na^+ shows lower occupancy, especially for the minor groove. On the contrary, the situation is different for Mg^{2+} ions. As shown in Figure 5-4(c) and (d), Mg^{2+} ions selectively bind to the major groove of GC base pairs, which is consistent with the patterns found in X-ray crystal structures of divalent ion in B-DNA¹⁵¹. At the major groove of the A-tract region and the minor groove of the whole DNA, they are almost absent. Meanwhile, only a few Na^+ ions in the Mg-DNA system can bind to DNA because Mg^{2+} ions act as strong competitors. It is interesting that in the Mg-DNA system, we also found a very low occupancy of Na^+ at the A-tract region even though Mg^{2+} ions are not present. Based on this observation, a more detailed description of ion interaction with the AT-rich DNA sequence is warranted (see discussion below). It is worth noting that the above discussions are valid without considering the terminal nucleic acids. The occupancy values of these bases may not accurately reflect their intrinsic properties because the hydrogen-bonds between the terminal residues usually break, leading to high mobility of these residues.

Groove width of DNA

The groove widths of DNA duplexes are important structural characteristics which may account for the sequence-specific binding of ions. The groove width from base pair 4 to 19 was calculated with the Curves program¹⁵² and is shown in Figure 5-5(a) and (b). The first 3 and last 4 base pairs were ignored because of their high mobility. We observe that the major groove of the GC-rich segment becomes narrower while the minor groove becomes wider comparing NaCl electrolyte and MgCl₂ electrolytes. On the contrary, the A-tract region undergoes the opposite change: the major groove becomes wider while the minor groove becomes narrower. This difference is more obvious in the periodic-DNA model and will be discussed later.

Comparing these two segments, we find that the A-tract segment usually has a broader major groove than that of the CG-rich segment in the presence of Mg²⁺. Thus, the absence of Mg²⁺ at the major groove of the A-tract region is not due to the steric hindrance but is an electrostatic effect. Hydrated Mg²⁺ ions interact with DNA mainly through hydrogen-bond interactions which are sensitive to the environment^{151,153}. As shown in Figure 5-4(d), guanine bases can hold the hydrated Mg²⁺ ions more stably because adjacent N7 and O6 atoms can function as hydrogen bond acceptors and form hydrogen bonds simultaneously with one Mg²⁺+6H₂O complex. On the contrary, the AT base pair does not provide a good electrostatic environment for hydrated Mg²⁺: two hydrogen acceptors, N7 atom of adenine and O4 atom of thymine are separated by the hydrogen bond donor -NH₂ group on the adenine. This special configuration inhibits a Mg²⁺+6H₂O complex forming stable hydrogen bonds.

As for the minor groove (Figure 5-5(b)), it is clear that the A-tract segment has a narrower minor groove than the GC-rich region. This results in low occupancy of Na⁺ (Figure 5-4(b)) and Mg²⁺ (Figure 5-4(d)) in the minor groove at base pair 16-19. Experimental data also indicated that there is no penetration of divalent ions into the spine of hydration of the minor groove¹⁵⁴⁻¹⁵⁶. In the simulations of the periodic-DNA model, we also observed the absence of Na⁺ in the minor groove of the pAA DNA.

Figure 5-6 (a) Eigenvalues of first 10 eigenvectors from principal component analysis (PCA) analysis, the inset is the inner-product of the first 5 eigenvectors from Mg-DNA (x-axis) and Na-DNA (y-axis) system; (b) Time evolution of the conformational entropy.

Structural plasticity of DNA

Counterions are not only “passively” pulled to the vicinity of DNA and screen the electrostatic field around the polyelectrolyte; they can also “actively” modulate DNA dynamics. Analysis of the conformational entropy of DNA can be used to address this issue. First, PCA analysis was conducted in Cartesian coordinate space using the simulation trajectories. Figure 5-6(a) shows magnitudes of the 10 largest eigenvalues for the Na-DNA and Mg-DNA system. The inner-products of the first 5 eigenvectors from the Mg-DNA and Na-DNA system are also shown in the insert. The large values of the inner-products of the first 2 eigenvectors from the two systems indicate that the global motions of DNA molecules in these two systems are similar. Eigenvectors with large eigenvalues depict the global motion of DNA and the magnitude of the eigenvalue reflects the motion amplitude of certain degree of freedom. We can see that the largest eigenvalue for Mg-DNA is smaller than that of Na-DNA, which indicates that the global motion of Mg-DNA is reduced, compared with that of Na-DNA.

The configuration entropy was calculated at a time intervals of 5 ns and shown in Figure 5-6(b). It is found that the entropy of Mg-DNA is always $\sim 800 \text{ J}/(\text{mol} \cdot \text{K})$ smaller than that of Na-DNA system. After checking the virtual trajectory of the first eigenvector, we found that this corresponds to the global bending of the whole DNA molecule. Clearly the binding of Mg^{2+} on DNA makes it more rigid compared to that of Na^+ .

The rigidity of DNA can also be represented by the local bending angles calculated using Madbend¹⁵⁷. To get the local bending angles, a six-basepair window was used to scan the whole DNA strands which resulting in 16 six-basepair overlapping segments (ignoring the first and last base pairs). The average bending angles and error estimations are shown in Figure 5-5(c). It was found that the local bending is reduced in the presence of Mg^{2+} ions compared to Na^+ , except for the first two segments. This reflects that the bending motion of the DNA duplex is partly frozen in the presence of Mg^{2+} .

5.2.3 Results of periodic DNA model

Our simulation results based on an isolated 23mer DNA with an A-tract sequence have elucidated the specific Mg^{2+} ions binding and the related sequence-dependent structural and dynamical properties of DNA. In the following section we aim to confirm the above findings by employing a model of 10mer periodic DNAs with specific sequences. The periodic DNAs (pDNA) do not have dangling ending residues so the ten nucleic acids can be treated uniformly during the data analysis.

Ions around periodic DNA

The average numbers of ions located nearby the DNA major/minor grooves and phosphate groups are shown in Figure 5-7(a) and (b). For the Na-pDNA system, there were around 0.2~0.3 Na^+ at the major groove of the four pDNAs. Nearby the minor groove of pGG, pGC and pAT, there were around 0.2 Na^+ ions. The most significant observation is that there was almost no Na^+ (0.028 per base) at the minor groove of pAA. This phenomenon is consistent with values in Figure 5-4(a) and (b) where the A-tract region has low occupancy of Na^+ at the major and especially the minor groove. However, there is around 0.1 more Na^+ at the backbone of pAA compared to other three types of pDNA. This implies that the minor groove of pAA is very narrow. Thus Na^+ ions accumulate at the backbone without penetrating into the minor groove.

For Mg^{2+} , in the major groove, the occupancy order is: pGC > pGG > pAT $\approx$ pAA. The basepair averaged value of 0.1 indicates only one Mg^{2+} exists at the major groove of

Figure 5-7 The basepair-averaged number of (a) Na⁺ and (b) Mg²⁺ located at major groove, minor groove and backbone of periodic DNA. The overall charge neutralization fraction was also shown accordingly. Width of (c) major groove and (d) minor groove of periodic DNA in Na⁺ and Mg²⁺ electrolytes.

pAA and pAT decamers. For pGC, the base pair-averaged value is above 0.3, which means that there is on the average one Mg^{2+} shared by three base pairs in the major groove. In contrast, the occupancy percentage at the backbone of pGC was lower than the other three periodic DNA models. Moreover, there were almost no Mg^{2+} ions at the minor groove of four periodic DNAs, which is consistent with the isolated DNA model.

As shown in Figure 5-7(a), although the Na^+ ions concentration increased from 0.08 M in the isolated DNA model to as high as > 0.15 M in the four periodic DNA models, the neutralization fraction is still around 50% for all periodic DNAs. Experimental evidence shows that DNA condensation occurs when about 90% of its charge is neutralized by counterions¹²². Structurally, it seems that one DNA base pair can attract one Na^+ at most, so that it is impossible for DNA condensation in Na^+ electrolyte even at very high concentration. On the contrary, Mg^{2+} at 0.15~0.20 M successfully neutralized 90% of the negative charges of DNA in the four cases as shown in Fig. 7b, which may induce DNA condensation. It is interesting to see that pGG and pGC have higher neutralization fraction values of 96% and 94% compared to those of pAA and pAT which are 88%. This indicates that more Mg^{2+} accumulate around G·C than A·T nucleic acids which is consistent with experimental measurements on the dehydration of DNA in Mg^{2+} electrolyte, where a lower dehydration effect is observed when the A·T content becomes higher¹⁵³.

Groove width of pDNA

As shown in Figure 5-7(c) and (d), generally, pGG and pGC have narrower major grooves and broader minor grooves than AT-rich DNAs in the presence of Na^+ or Mg^{2+} . Among these four kinds of pDNA models, pGC shows distinctive structural change in the two types of electrolytes: the major groove of pGC in Na^+ is broader than that in Mg^{2+} , while the minor groove of pGC in Na^+ is narrower than that in Mg^{2+} . It may also be noted that pAA has the narrowest minor groove (~ 4 Å), which facilitates the very stable Mg^{2+} binding on the phosphate as shown in Figure 5-3(d). This also explains the reason that there are almost no ions at the minor groove of pAA.

Figure 5-8 Time evolution of conformational entropy of four periodic DNA in Na^+ and Mg^{2+} electrolytes.

Conformational entropy of pDNA

Finally we studied the conformational entropy in periodic DNA system under the influence of the counterions. As shown in Figure 5-8, we can see that Mg^{2+} ions exert “freezing” effects on DNA dynamics compared with Na^+ ions for pGC, pAA and pAT DNAs. For pGG DNA, the freezing effects are not evident. The conformational entropies in Mg^{2+} and Na^+ do not differ too much (the largest difference is for pAA, around $0.3 \text{ kJ}/(\text{mol} \cdot \text{K})$ per 10 base pairs). The main reason is that in the periodic systems the DNA decamer is covalent bonded to its image, the overall bending in this DNA model is totally inhibited. The conformational entropy of the periodic DNAs may mainly capture local motions like base pair tilt, roll and twist, etc.

5.3 Conclusion

In brief, molecular dynamics simulations of DNA duplex in the presence of Mg^{2+} and Na^+ ions were performed. Na^+ ions have less-stable hydration shells and interact with DNA non-specifically. On the contrary, Mg^{2+} ions have stable hydration shells and interact with DNA mainly through hydrogen-bonds. Thus, Mg^{2+} ion binding is sensitive to the local environment and is sequence-specific. Although A·T nucleic acids are found to have broader major grooves than G·C nucleic acids, Mg^{2+} ions were only found in the major groove of the G·C base pairs. This binding selectivity has an electrostatic nature and is not of steric origin. The adjacent N7 and O6 atoms at guanine base create a negative potential environment and act as hydrogen-bond acceptors for hydrated Mg^{2+} ions, while the positively charged H atoms on N6 amino group of adenine base repel the hydrated cations. Moreover, the binding of Mg^{2+} ions makes the DNA duplex more rigid. The overall conformational entropy of Mg^{2+} -bound DNA is significantly less than that of Na^+ -bound DNA and the local bending amplitudes are also reduced in the presence of Mg^{2+} ions. We expect the sequence-specific binding of Mg^{2+} will lead to distinct DNA condensation patterns in the context of different DNA sequences.

Chapter 6

Mg²⁺-mediated DNA Interactions

As a continuation of the study reported in chapter 5, this chapter also focuses on DNA dynamics in MgCl₂ electrolyte. The electrostatic screening effects and inter-DNA interactions that induce, stabilize and regulate the attraction between DNA duplexes in the context of DNA condensation will be discussed. The central question of this chapter is: how is DNA condensation affected by cations, DNA sequence and conformation?

DNA duplex has two negatively charged phosphate strands spiraling around nucleotide base pairs. The DNA-DNA interaction is electrostatically repulsive without counterions. However, it is interesting that meters of genomic DNA are packed by nature into compact structures for all living beings. To condense DNA, the attractive forces must outweigh the repulsive force. In chromosomes, this is done by specialized proteins. In phage heads and sperm, DNA is packaged by simple counterions. Ion-mediated DNA-DNA interaction has been studied extensively. Monovalent alkali cations (Li⁺, Na⁺, K⁺, Rb⁺, Cs⁺) hardly induce attraction between DNAs, and normal divalent cations (Ca²⁺, Mg²⁺) can induce DNA-DNA attraction only at high salt concentration¹⁵⁸. Cations of charge 3+ or higher are usually required to condense DNA in *in vitro* experiments, such as hexammine cobalt and spermine^{159–161}. Another class of DNA-condensation agent is transition metal ions, like Mn²⁺, Co²⁺, Ni²⁺ and Zn²⁺^{162,163}. Some analytical theories and simulations have been conducted focusing on the importance of cation binding patterns around DNA and their electrostatic screening effect^{28,164–169}. Generally speaking, DNA-DNA interaction is

cation-dependent.

However, there is mounting evidence indicating that sequence-dependent conformational changes may play an important role in DNA condensation^{162,163,170,171}. For example, in sub-millimolar concentrations of Ni^{2+} electrolyte, repeating guanine-cytosine sequence adopts left-hand conformation, Z-DNA, and is easily condensed. On the contrary, AT rich DNA remains B-form and displays very minor condensation even at much higher Ni^{2+} concentrations (10 times higher than that used to condense GC rich DNA)¹⁶².

To date, compared to the extensive experimental and simulation studies of B-DNA, there have been few reports of Z-DNA interaction studies so far. The main reason is that the Z-form is a high energy conformation of the double helix. It is occasionally induced by biological activity and then quickly disappears. The capture of a Z-DNA is, especially for experiments, a challenge^{172–174}. The questions of what are the biological functions of Z-DNA and how does this special configuration interact with other biological molecules like proteins, have not received much attention. This knowledge gap is becoming a hurdle for us to understand the functions of Z-DNA in DNA transactions such as transcription, chromatin remodeling and recombination.

In this chapter, a series of MD simulations are performed to study the lateral interactions between two parallel DNA duplexes in the presence of Mg^{2+} ions, where effects from both DNA sequence and conformation are taken into our consideration. In general, we find that for both B-DNA and Z-DNA, the attractive forces induced by counterions are usually very weak, with a attractive energy well less than -2 kcal/mol. Compared to AT rich B-DNA, GC rich B-DNA double helices can approach each other a little more closely, but with weaker attractive force. For Z-DNA, on the contrary, a significant strong attraction is found. When DNA segments adopt Z-form, the hydrophobic sugar rings are exposed to the solvent while phosphate atoms of one strand are buried inside. The strong attraction of Z-DNA is due to the de-hydration of the sugar rings. Based on contacting surface analysis, a two-stage Z-DNA association process is outlined. During gene transcription, some GC rich DNA segments adopt the Z-form at the transcription starting point. The functions of these Z-form segments are not well known yet. We propose that the rich content of hydrophobic surface of Z-DNA, which is found to be important in DNA interactions, may be relevant in

mediating DNA-protein interaction in the gene transcription process.

6.1 Computational Methods

6.1.1 DNA models

In all of the simulations, we have considered two identical DNA duplexes submerged in a simulation cell with periodic boundary conditions imposed in all directions. The xy -cross section of the simulation cell is hexagonal (as shown in Figure 6-1). For B-DNA, one turn of helix contains 10 base pairs with pitch length of 3.32 nm. Four sequences have been considered: $d(5'-AAAAAAAAAA) \times d(5'-TTTTTTTTTT)$, $d(5'-ATATATATAT) \times d(5'-ATATATATAT)$, $d(5'-GGGGGGGGGG) \times d(5'-CCCCCCCCC)$ and $d(5'-GCGCGCGCGC) \times d(5'-GCGCGCGCGC)$. The “sticky” ends means that the 3' end of each DNA strand is connected to the periodic image of 5' end over the boundary (the cell length in the z direction is equal to the pitch length). This setup mimics two infinite long DNA duplexes where only lateral interaction is considered. These four systems are abbreviated as AA-DNA, AT-DNA, GG-DNA and GC-DNA. For Z-DNA, one turn of helix contains 12 base pairs with pitch length of 4.56 nm. As resolved by experiments, Z-DNA usually forms from repeating guanine-cytosine sequences^{161,162,172–174}. Hence we have only considered one Z-DNA model of sequence: $d(5'-GCGCGCGCGCGC) \times d(5'-GCGCGCGCGCGC)$ (abbreviated GC-ZDNA).

6.1.2 Simulation details

Other than DNAs, the simulation boxes contain 4300 (B-DNA system) and 6200 (Z-DNA system) TIP3P³⁸ water molecules. 25 Mg^{2+} and 10 Cl^- were added in B-DNA system, while for the Z-DNA system there are 35 Mg^{2+} and 22 Cl^- . The choice of Mg^{2+} as cation is driven by the experimental findings that Z-DNA can be found in both 0.7 M $MgCl_2$ ¹⁶¹ and mM concentration of $NiCl_2$ ¹⁶². The force field for Mg^{2+} is well parameterized^{145,175}, while models for transition metal ions like Ni^{2+} are challenging. The molar concentrations of Mg^{2+} were evaluated from the solution far from DNA atoms which are ~ 140 mM in B-DNA system and ~ 100 mM in Z-DNA system¹⁷⁵.

Figure 6-1 Illustration of the cross section of the simulation box and how the two DNAs are connected by an umbrella sampling spring through centers of mass (shown as red balls).

The recent modified AMBER(parmbsc0) force-field¹⁴³ was used to model the DNA molecule, while for counterions, CHARMM parameters optimized by *Roux* and coworkers¹⁴⁵ were used. All simulations were carried out using the GROMACS³⁷.

6.1.3 Umbrella sampling method

In normal MD simulations without any constraints, two freely diffusing DNAs easily make contacts and hardly separate, which limits the sampling efficiency. Thus in our simulations, the potentials of mean force (PMF) between two DNAs were calculated with the umbrella sampling method¹⁷⁶. Statistical uncertainties were estimated using the technique of bootstrap analysis¹⁷⁷. Two DNAs were connected through center of mass (COM) with a harmonic spring as illustrated in Figure 6-1. Two DNAs were fully free except that the inter-DNA (COM-COM) distance in *xy* plane was restricted to certain values in the umbrella sampling. The inter-DNA distance was sampled from 1.90 nm (for B-DNA) and 1.60 nm (for Z-DNA) to 3.60 nm with a resolution of 0.05 nm. At each distance, the system was first equilibrated for 10 ns followed by a 100 ns production simulation.

Figure 6-2 Potentials of mean force (PMF) with respect to inter-DNA distance between two (a) B-DNAs and (b) Z-DNAs. Free energy values at 3.5 nm are set to 0.

6.2 Results

6.2.1 PMF between two DNA duplexes

First we consider the interactions between B-DNAs as shown in Figure 6-2(a). For AA-DNA and AT-DNA, there is only one energy minimum at distance ~ 2.4 nm; the depth of the energy well is ~ -2.0 kcal/mol, close to the experimental estimation¹⁷⁸ and the simulation prediction²⁸. The attractive energy well is quite deep (2 kcal/mol being around $4k_bT$, $T=300$ K) which seems capable of inducing DNA condensation. The interaction landscape of GC rich DNAs is different. The PMFs of GG-DNA and GC-DNA have two local minima located at 2.1 nm (close bound state) and 2.4 nm (loose bound state) with energy well of only ~ -1.3 kcal/mol; the free energy barrier between these two states is ~ 0.5 kcal/mol.

When we take Z-form DNAs into our consideration, the most striking observation is the significant deep free energy well between two Z-DNAs as shown in Figure 6-2(b). The bound state is -8.5 kcal/mol lower than the separated state, which is about fourfold of those of B-DNAs. At the bound state, two Z-DNAs can approach each other very closely with inter-DNA distance of only 1.8 nm. This is not strange because Z-DNA is elongated by about 37% compared with a B-DNA, so that the radius becomes smaller.

6.2.2 DNA-DNA Binding pattern

For two identical parallel DNAs, the relative rotation about their helical axes can be characterized by

$$\Delta\theta = \Delta\theta_Z - 2\pi \frac{\Delta L}{L} \quad (6.1)$$

where $\Delta\theta_Z$ is the difference of the rotational angles of two DNAs about their helical axes, ΔL is the relative displacement of the DNA along the z axis, and L (~ 3.32 nm for B-DNA and ~ 4.56 nm for Z-DNA) is the length of one helical turn. For two B-DNAs at bound states, the distributions of the relative rotational angles are shown in Figure 6-3(a-c). The most sampled bound angles are: 120° for AT rich DNAs at 2.4 nm²⁸, 270° (2.4 nm, loose bound state) and 90° (2.1 nm, close bound state) for GC rich DNAs. For GC rich DNAs, the two DNAs has to rotate relatively by 180° when they jump from the state with 2.4 nm separation to the state with 2.1 nm separation.

Genomic DNA has more complicated sequences, and hence the bound patterns and relative orientations between genomic DNA strands may not be as uniform as the AT and GC only DNA models. Hence it is safe to propose that for genomic DNA of B-form, the attraction may be weaker than predicted by the four PMFs in Figure 6-2(a). Thus DNA condensation in bulk solution by Mg^{2+} is hardly observed in experiments¹⁷⁹.

6.2.3 Solvent accessible surface of B-DNA and Z-DNA

B-DNA and Z-DNA have different surfaces. The specificity of DNA surfaces would affect Mg^{2+} ions binding and, accordingly, DNA-DNA interactions. In order to describe DNA surfaces quantitatively, we first calculated the surface area projected on each atom. The atoms with exposed surface area value larger than 0.1 nm^2 are listed in Table 6.1 (see page 116). It indicates that the transition from B-form to Z-form leaves more sugar ring and base atoms of guanine (G) exposed to the solvent. O6 and N7 atoms in the major groove of a B-DNA are also found to be exposed to the solvent in Z-form DNA. This is consistent with the previous findings that transition metal ions bind specifically to the N7 atom of guanine, inducing the B- to Z- transition, where the N7 atom of guanine becomes more

Figure 6-3 Distributions of the relative rotational angles of two parallel DNAs at stable binding state: (a) AT-DNA at 2.4 nm; (b) GC-DNA at 2.4 nm; (c) GC-DNA at 2.1 nm and (d) GC-ZDNA at 1.8 nm.

accessible^{180–182}.

In the following, the DNA surface is classified into three types according to the atomic charge, Q : (1) hydrophobic, $|Q| \leq 0.2$; (2) negative hydrophilic, $Q < -0.2$; (3) positive hydrophilic, $Q > 0.2$. The representative surfaces of B-DNAs and Z-DNA are shown in the second panel of Figure 6-4. It is clear that the outermost surfaces of B-DNAs are mostly negative hydrophilic and composed of phosphate oxygen atoms. In contrast, for Z-DNA, apart from the negative hydrophilic surface, hydrophobic sugar rings are also exposed outwards. The outermost surface of Z-DNA has an alternative negative-hydrophobic pattern.

A more direct measure of the solvent accessibility of DNA is the size of hydration shell. The hydration shell is defined as the group of water molecules located within 0.5 nm of the DNA surface. These water molecules have been demonstrated to play a dominant role in the hydration thermodynamics^{111,183}. When two DNAs are separated at 3.5 nm, they were treated as non-interacting. The numbers of hydration water (averaged to each base pair)

Figure 6-4 Snapshots of two DNAs at stable binding states: (a) AT-DNA at 2.4 nm, (b) GC-DNA at 2.4 nm; (c) GC-DNA at 2.1 nm and (d) GC-ZDNA at 1.8 nm. In the upper panel, DNA is depicted by a stick diagram while Mg^{2+} ions which directly bound to DNA are shown as magentas spheres. In the lower panel, DNA molecules are represented by a surface diagram. The color scheme is: green for hydrophobic (absolute value of atom charge is less than 0.2), red for negative hydrophilic (atom charge is less than -0.2), blue for positive hydrophilic (atom charge is larger than 0.2).

are: 42.6 for AT-DNA, 43.5 for GC-DNA and 38.7 for GC-ZDNA. These values clearly reflect that the Z-DNA molecule is more hydrophobic than B-DNA.

6.2.4 DNA de-hydration through contacting

When two DNAs approach each other, the water accessible space between DNAs decreases, corresponding to a decrease of the size of the hydration shell (called *de-hydration*). The dependence of the size of the hydration shell (also averaged to each base pair) with respect to inter-DNA distance is shown in Figure 6-5(b, e, h). When the inter-DNA distance changes from 3.5 nm to 2.4 nm (global minimum of PMF), the hydration water decreases from 42.6 to 41.2 for AT-DNA, and from 43.5 to 42.2 for GC-DNA.

In contrast, for two Z-DNAs at 1.8 nm, number of hydration water decreases significant-

Figure 6-5 The Potentials of mean force (PMF), number of hydration water in the hydration shell and contact surface area of (a, b, c) AT-DNA, (d, e, f) GC-DNA and (g, h, i) GC-ZDNA with respect to inter-DNA distance.

ly from 38.7 at 3.5 nm to 34.0 at 1.78 nm. The de-hydrated water number is 4.7, compared to the values of 1.4 and 1.3 for AT-DNA and GC-DNA, respectively. From this comparison, it is clear that the strong attraction between Z-DNA is reinforced by the de-hydration of hydrophobic sugar rings.

6.2.5 Contact surface between two DNA duplexes

It is also interesting to check the composition of the contact surface. The values of contact surface area (normalized to each base pair) are shown in Figure 6-5(c, f, i). For B-DNAs, the bound state at 2.4 nm for GC-DNA and AT-DNA, have negligible contact area, which are 0.035 and 0.045 nm², respectively. Meanwhile most of the contacting surface is hydrophilic. For the closer bound state of GC-DNA at an inter-DNA distance of 2.1 nm, the

contact surface increases to 0.18 nm², where hydrophilic surface is still dominant (67%). The small values of contact area reflect that these bound states are ion-mediated, rather than direct contact as demonstrated in Figure 6-4(a, b). Numbers of Mg²⁺ ions accumulate at the interfacial and exert an effective attractive force between two DNAs^{28,184}, acting like a “bridge” between two DNAs. Thus the attractive force between DNAs depends on the binding affinity of Mg²⁺ to DNA atoms. Monovalent ions bind weakly to the phosphate oxygen thus DNAs repulse each other. Divalent ions, like Mg²⁺ and Ca²⁺ bind more stably to the phosphate oxygen which results in effective inter-DNA attractions.

For the case of Z-DNA at the bound state (inter-DNA distance of 1.80 nm), two Z-DNAs have significantly large contact surfaces of 0.45 nm². This large value indicates that two DNAs form more intimate contacts as shown in Figure 6-4(d). The contents of the contact surface are also reversed compared with those of B-DNAs, which are 59% hydrophobic and 41% hydrophilic. This is consistent with the findings from de-hydration analysis.

6.2.6 A two-stage association of Z-DNA binding

Combining the characteristics of the PMF profile and the features of contact surface in Figure 6-5, a two-stage Z-DNA association process is outlined:

(1) In the first stage (inter-DNA distance larger than 2.0 nm), two Z-DNAs approach each other because counterions have effectively dismissed the repulsive forces. This stage corresponds to a decrease in free energy of -1.80 kcal/mol at 2.0 nm. This value is comparable with those of B-DNAs at the PMF global minimum, which reflects that the fact that charge screening effect is unable to attract two DNA more effectively. Thus the attractive force between DNAs depends on the binding affinity of Mg²⁺ to DNA atoms, and the charge screening effect is almost conformation independent.

(2) In the second stage (inter-DNA distance smaller than 2.0 nm), the PMF dramatically decreases to a value of -5.8 kcal/mol at 1.85 nm and continues to drop to -8.5 kcal/mol at 1.78 nm. Concomitantly, there is a significant increase of the contact area, especially the hydrophobic component. The attraction in this stage is mainly driven by de-hydration of hydrophobic sugar rings which is missing in B-DNA association.

Figure 6-6 Radial distribution functions (RDFs) of sugar (C1' – C5') and phosphate (O1P, O2P) atoms of GC-ZDNA, around center of mass of DNA in the xy -plane.

6.3 Discussion

6.3.1 The structural features of Z-DNA

The participation of hydrophobic interaction in Z-DNAs is related to their structure. The distributions of the sugar (hydrophobic) and phosphate (hydrophilic) atoms with respect to the center of mass of xy -plane are shown in Figure 6-6. For B-DNA, the outermost surfaces are composed of phosphate atoms, while sugar atoms are located ~ 0.25 nm more inside. When a Z-DNA is formed, the phosphate-sugar backbone has a repeated feature in every 2 base pairs: the DG is C3'-endo, while DC is C2'-endo. This dissimilarity results in distinct differences in the spatial distributions of the sugar and phosphate atoms between DG and DC as shown in Figure 6-6b. Only the phosphate group of DC is exposed to the environment, followed by the sugar group of DG. The phosphate group of DG is more hidden inside which will definitely weaken the repulsion to another Z-DNA.

6.3.2 Biological functions of Z-DNA

To date the role of Z-DNA is commonly believed to trap negative supercoiling when transcription occurs^{172,185}. However, the findings of the Z-DNA binding motif, Z α domain

of human RNA adenosine deaminase ($hZ\alpha_{ADAR1}$) by *Alan Herbert*^{186–190} indicate a biological function. The function of $hZ\alpha_{ADAR1}$ may localize this enzyme that modifies the sequence of the newly formed RNA to sites of active transcription¹⁹¹. The participation of Z-DNA in RNA processing is consistent with the observation that potential Z-DNA forming sequences accumulate near the transcriptional start sites (TSS) of human genes, especially at upstream (5') region to the TSS^{192–195}. The possible role of the Z-DNA may be regulating the transcription: $hZ\alpha_{ADAR1}$ has to first bind to Z-DNA segment so that it can find the TSS. However, a key question is how the $hZ\alpha_{ADAR1}$ finds the Z-DNA?

The observation of hydrophobic contacts in Z-DNA interaction may be relevant in understanding the specific and strong $Z\alpha$ – Z-DNA binding, if we can generalize it to other kinds of interactions involving Z-DNA. A justification is interactions between B-DNA and Z-DNA. It is known that the introduction of Z-DNA has a significant effect on DNA condensation. By inserting 12-bp and 20-bp Z-form d(CG)n sequences into pUC18 plasmids, the plasmids were found to undergo significantly enhanced condensation¹⁶¹. The content of Z-DNA in plasmid is below 1%, such that plasmid condensation should be induced by B-DNA – Z-DNA binding. The Z-DNA – Z-DNA binding is almost negligible.

An exciting experimental clue that points to the significance of hydrophobic interaction is that $Z\alpha$ binds to Z-DNA with limited base-specificity^{187–189}. It can bind equally to GC-ZDNA as well as AT-ZDNA prepared in *in vitro* studies, in other words, the binding is only conformation-specific¹⁸⁷. The absence of sequence-specificity leaves the hydrophobic surface of Z-DNA as the dominant factor which is the “lure” to the $hZ\alpha_{ADAR1}$. In addition, a solution structure shows that $hZ\alpha_{ADAR1}$ has already a pre-positioned binding surface to bind to Z-DNA¹⁸⁹. If $hZ\alpha_{ADAR1}$ binding actively starts the B-to-Z transition, the interaction follows the “induced fit” model¹⁹⁶, where the hydrophobic interaction is passively involved. However, recent single-molecule FRET experiments revealed that intrinsic Z-DNAs are dynamically formed; $hZ\alpha_{ADAR1}$ passively trap preformed Z-DNAs rather than actively induce B-to-Z transition¹⁹⁷. This can explain the rapid association kinetics of $Z\alpha$ – Z-DNA binding¹⁸⁶. Thus, $hZ\alpha_{ADAR1}$ and Z-DNA segment binding is also driven by de-hydration as discussed in the present study.

6.4 Conclusion

From MD simulations, the roles of charge screening effects as well as de-hydration effects in DNA interactions are quantitatively compared. For both B-form and Z-form DNA, the charge screening effect induced attraction is usually weak, corresponding to an attractive energy well depths less than -2kcal/mol, which is insufficient to induce DNA condensation. When DNA adopts Z-form, the sugar rings are exposed at the surface, making Z-DNA hydrophobic. The de-hydration effect from sugar ring contacts can significantly enhance DNA attraction. The strength is about fourfold compared to those of B-DNA. The mechanism may be general and extended beyond inter-Z-DNA interactions, for example, the binding of RNA adenosine deaminase to Z-DNA during transcription.

The revelation of the critical role of the hydrophobic surface on Z-DNA may lead to a broader understanding of the function of Z-DNA in biological environment.

Table 6.1 Atoms with surface area $> 0.1 \text{ nm}^2$, the surface area values are shown in the brackets

	Phosphate	Sugar	Base
AT-DNA	A O1P(0.31), O2P(0.29)	C4'(0.14), C5'(0.23)	–
	T O1P(0.32), O2P(0.30)	C4'(0.13), C5'(0.23)	C7(0.40)
GC-DNA	G O1P(0.32), O2P(0.30)	C2'(0.11), C4'(0.13), C5'(0.23)	C8(0.11)
	C O1P(0.30), O2P(0.31)	C4'(0.14), C5'(0.23)	C5(0.16)
GC-ZDNA	G O1P(0.30), O2P(0.31)	C1'(0.13), C2'(0.16), C4'(0.13), C5'(0.14)	O6(0.13), N7(0.13), C8(0.31)
	C O1P(0.25), O2P(0.30)	C5'(0.12)	–

Chapter 7

Conclusions and Future Plan

7.1 Conclusions

The salting effect is complicated and relies on the sensitive balance of the tripartite interactions among solutes, ions and water molecules. Taking the hydration free energies of solutes and ions in pure water (water is always treated as the medium) as independent variables, a schematic diagram of salting effects is suggested from the studies in Chapter 2. Generally speaking, the salting effect is mainly a thermodynamic equilibrium: strongly-hydrated ions salt in strongly-hydrated solutes, but salt out weakly-hydrated solutes; weakly-hydrated ions salt in weakly-hydrated solutes. The underlying mechanism is related to the preferential binding of ions and water to solutes. In the biological sciences, “strongly-hydrated” is usually named as “hydrophilic”; while “weakly-hydrated” is usually named as “hydrophobic”. Under such kinds of definitions, a more fundamental/obvious explanation of such a diagram is that *Hydrophilic likes hydrophilic, hydrophobic likes hydrophobic*, if we ignore the terms such as “solute” and “ion”, and treat them as *equal components* in water. These conclusions have also highlighted the importance of the specifics of solutes: the salting effect not only depends on the salt’s position in Hofmeister series, but also on the solutes’ specifics.

Urea and guanidinium are two of the most weakly-hydrated solutes characterized by almost non-recognizable hydration shells¹⁹⁸. From our calculations in Chapter 4, most of the

15 molecules (components of protein) are more hydrophobic than water (-6.32 kcal/mol). Hence it is found that urea and guanidinium are protein denaturants, unfolding proteins by salting-in effect. On the contrary, Na^+ and Cl^- are strongly-hydrated ions with hydration free energies of ~ -88 kcal/mol, and hence will salt-out proteinic molecules which are relatively more hydrophobic than them. Structurally, Na^+ and Cl^- prefer binding to water molecules, resulting in a decrease of the number of “free” water. The effective hydration environment for proteins becomes crowded and proteins are inclined to precipitate. The comparative studies of NaCl and urea’s effects on protein solubility also follow the schematic diagram of salting effects proposed in Chapter 2.

For protein denaturation, we have adopted proteinic molecules (amino acid sidechain analogues and N-Methylacetamide which mimics the backbone). The salting of 15 model molecules in NaCl and urea are investigated. The capability of urea to better solvate protein backbone and side chain analogues can be extrapolated to explain a protein’s denaturation in aqueous urea solution. For proteins at folded state, urea molecules accumulate nearby and “solvate” the amino acids on protein surface into solvent. The disruption of the protein surface leaves the protein interior exposed to the urea solvent. Consequently, urea “salts in” the interior amino acids, which biases the folded/unfolded equilibrium towards unfolding and finally causes denaturation. By energy-decomposition, the salting effect is found to be largely enthalpy-driven for both NaCl and urea. Urea salts in molecules through direct binding to the solute surface and the strength is linearly dependent on the number of heavy atoms of solute molecules. The van der Waals interactions are found to be the dominant force. This finding challenges a hydrogen-bonding driven mechanism for protein denaturation proposed previously.

The ion-DNA interactions are dominated by electrostatic forces, rather than a thermodynamic equilibrium in the salting effect discussed in Chapters 2-4. This is because the DNA duplex has two highly negative-charged phosphate strands spiraling around the surface. Ions like Mg^{2+} and Na^+ act like counterions that can electrostatically screen out the repulsive force between DNA duplexes. Mg^{2+} has a stable hydration shell consisting of six water molecules. Thus Mg^{2+} ’s interaction with DNA is mediated by this water shell. Water molecules in the shell may form hydrogen-bonds with the phosphate oxygens of DNA.

Such direct interactions are sensitive to the local environment. Mg^{2+} ions were found at the phosphate backbone and selectively in the major groove of G-C bases. This sequence-dependent specificity is of electrostatic nature and the steric constraint has only a minor effect. The adjacent N7 and O6 atoms at the guanine base create a negative potential environment and act as hydrogen-bond acceptor for hydrated Mg^{2+} ions, while the positively charged H atoms on N6 amino group of adenine base repel the hydrated Mg^{2+} . For DNA dynamics, the binding of Mg^{2+} makes the DNA duplex more rigid compared to Na^+ , which is demonstrated by reduced conformational entropy and restricted local bending motion.

DNA-DNA interactions in MgCl_2 electrolytes is more complicated. From the studies in Chapter 6, the DNA-DNA interaction is found to be sequence-dependent, as well as conformation-dependent. B-DNA attraction is mainly induced by Mg^{2+} 's binding. As counterions, Mg^{2+} ions accumulate at the interfacial phosphate oxygen atoms and exert an effective attractive force between two DNAs. GC rich and AT rich B-DNAs demonstrate distinct attractive energy profiles. Inter-Z-DNA attraction is fourfold stronger than those of B-DNAs which is enforced by the strong de-hydration of Z-DNA that is almost absent in the case of B-DNAs. It is interesting that the ion mediated DNA-DNA attraction also follows the principles of the salting effect that we summarized before. As demonstrated in our analysis, the Z-DNA is more hydrophobic than B-DNA. Thus in the presence of Mg^{2+} ions, Z-DNA would be more strongly salted-out than B-DNA, which is reflected as stronger inter-Z-DNA attraction.

Moreover, combining our simulation results with experimental data, we propose that the hydrophobic nature of Z-form DNA may be relevant in understanding the biological functions of Z-DNA in gene transcription.

7.2 Future plan

7.2.1 Comparative studies of anions and cations

The whole Hofmeister series are shown in the following. The orders of ions are given in terms of their ability to stabilize the structure of proteins.

most stabilizing	most destabilizing
strongly hydrated anions	weakly hydrated anions
<i>citrate</i> ³⁻ > <i>sulfate</i> ²⁻ > <i>phosphate</i> ²⁻ > <i>F</i> ⁻ > <i>Cl</i> ⁻ > <i>Br</i> ⁻ > <i>I</i> ⁻ > <i>NO</i> ₃ ⁻ > <i>ClO</i> ₄ ⁻	
<i>N(CH</i> ₃ <i>)</i> ₄ ⁺ > <i>NH</i> ₄ ⁺ > <i>Cs</i> ⁺ > <i>Rb</i> ⁺ > <i>K</i> ⁺ > <i>Na</i> ⁺ > <i>H</i> ⁺ > <i>Ca</i> ²⁺ > <i>Mg</i> ²⁺ > <i>Al</i> ³⁺	
weakly hydrated cations	strongly hydrated cations

Figure 7-1 Hofmeister series of ions. The order is given in terms of the ability of ions to stabilize the structure of proteins from most stabilizing to most destabilizing (*citrate*³⁻ refers to *C*₃*H*₅*O(COO)*₃³⁻, *sulfate*²⁻ refers to *SO*₄²⁻ and *phosphate*²⁻ refers to *HPO*₄²⁻).

As shown in Figure 7-1, from protein stabilizing species to destabilizing species, it is notable that the intrinsic properties of the ions are reversed between anions and cations. For anions, the most stabilizing anions are strongly hydrated, while the most destabilizing anions are weakly hydrated. For cations, the trends are reversed: the most stabilizing cations are weakly hydrated, while the most destabilizing cations are strongly hydrated. The hydrophobicities of the ions are determined by the molecule charge density of them. We can find that the early members of anions are highly charged with relatively smaller volume, and the later anions in the series are less charged with large volume. For cations the trend is reversed and more obvious: the early members in the series are less charged with larger volume, while the later cations are highly charged with smaller volume.

Although the Hofmeister effect has been studied for more than a century, the underlying mechanism is not entirely clear. Many of the published works are always concerned with how ions affect protein/water properties, like changes in general water structure and surface tension, and specific interactions between proteins and ions/water molecules^{34–36,68,103,106}. However, comprehensive studies which focus on the difference between cations and anions are still missing.

A comparative study of different pairs of anions and cations on the promotion of hydrophobic associations should be a promising and fruitful plan. The hydrocarbon chains model^{97,199} or the nano-particle dimer model in Chapter 2 can be used. The effects from different kinds of ions can be realized by changing types of anions or cations. For example, the effects from different anions can be analyzed by *NaF* vs. *NaClO*₄, while the effect from

cations can be compared by $N(CH_3)_4Cl$ and $NaCl$. The use of identical monovalent anion-s/cations in the examples is for the consideration of easy-control of salt concentrations. In principle, we should always compare the ions effects at same concentrations.

7.2.2 The role of ion's hydration shells

Another important aspect in the Hofmeister series is the role of ion's "hydration shells", especially for that of cations. As discussed in Chapter 5, the hydration shell of Mg^{2+} contains 6 water molecules which are very stable and seldom exchange with the free water in the solvent. On the contrary, the hydration shell of Na^+ is less stable thus a Na^+ can bind directly to the DNA duplex. As a consequence, while Na^+ merely compensates the DNA charge in a diffuse manner, Mg^{2+} binds strongly to the DNA through its hydration shell. The re-arrangement of water around ions is not restricted to the first hydration layer^{200–202} (for example, the RDF of Mg^{2+} has a observable second peak at 4.1 nm in Figure 5-1(a)). Thus in the framework of "direct interaction mechanism" for salting-effect, the effect of the ions may be realized through their "first hydration shell" or even the "second hydration shell". A rough justification to this hypothesis is the similar destabilizing abilities of Al^{3+} and *guanidinium*⁺, if we list both of them in the Hofmeister series^{203–205}: $NH_4^+ > K^+ > Na^+ > Li^+ > Ca^{2+} > Mg^{2+} > Al^{3+} \approx \textit{guanidinium}^+$. As we have discussed in Chapter 4, *guanidinium*⁺ is one of the most weakly-hydrated solutes characterized by almost non-recognizable hydration shells¹⁹⁸. On the contrary, a Al^{3+} has two well defined hydration shells, which contains 6 (the first hydration shell) and 12 water (the second hydration shell), respectively²⁰¹. Generally speaking, the role of a $Al^{3+}(H_2O)_{18}$ complex is similar to a "bare" *guanidinium*⁺. These observations highlight the influence of cation size and charge on the local structure of hydrated ions. A study of the local order of water molecules around ions (especially the cations) and their effect in the salting-effect is desirable and would shed new light on our knowledge of the Hofmeister effect.

Appendix A

Publications

1. **Li, W.**, Zhao, M., Xia, Y., Zhang, R. and Mu, Y. (2009) Covalent-adsorption induced magnetism in graphene. *J. Mater. Chem.*, 19, 9274-9282.
2. **Li, W.**, Zhao, M., Zhao, X., Xia, Y. and Mu, Y. (2010) Hydrogen saturation stabilizes vacancy-induced ferromagnetic ordering in graphene. *Phys. Chem. Chem. Phys.*, 12, 13699-13706.
3. **Li, W.** and Mu, Y. (2011) Hydration patterns and salting effects in sodium chloride solution. *J. Chem. Phys.*, 135, 134502-134506.
4. **Li, W.**, Nordenskiöld, L. and Mu, Y. (2011) Sequence-Specific Mg^{2+} -DNA Interactions: A Molecular Dynamics Simulation Study. *J. Phys. Chem. B*, 115, 14713-14720.
5. **Li, W.** and Mu, Y. (2012) Dissociation of hydrophobic and charged nano particles in aqueous guanidinium chloride and urea solutions: A molecular dynamics study. *Nanoscale*, 4, 1154-1159.
6. **Li, W.**, Zhou, R. and Mu, Y. (2012) Salting Effects on Protein Components in Aqueous NaCl and Urea Solutions: Toward Understanding of Urea-Induced Protein Denaturation. *J. Phys. Chem. B*, 116, 1446-1451.

7. Xu, W., Ping, J., **Li, W.** and Mu, Y. (2009) Assembly dynamics of two-beta sheets revealed by molecular dynamics simulations. *J. Chem. Phys.*, 130, 164709-164708.
8. Yan, L., **Li, W.**, Fan, X.F., Li, W., Chen, Y., Kuo, J.L., Li, L.J., Kwak, S.K., Mu, Y. and Chan-Park, M.B. (2009) Enrichment of (8,4) Single-Walled Carbon Nanotubes Through Coextraction with Heparin. *Small*, 6, 110-118.
9. Jiang, P., **Li, W.**, Shea, J.E. and Mu, Y. (2011) Resveratrol Inhibits the Formation of Multiple-Layered β -Sheet Oligomers of the Human Islet Amyloid Polypeptide Segment 22-27. *Biophys. J.*, 100, 1550-1558.
10. Li, P., Poon, Y.F., **Li, W.**, Zhu, H.Y., Yeap, S.H., Cao, Y., Qi, X., Zhou, C., Lamrani, M., Beuerman, R.W. et al. (2011) A polycationic antimicrobial and biocompatible hydrogel with microbe membrane suctioning ability. *Nat. Mater.*, 10, 149-156.
11. Yan, L.Y., **Li, W.**, Mesgari, S., Leong, S.S.J., Chen, Y., Loo, L.S., Mu, Y. and Chan-Park, M.B. (2011) Use of a Chondroitin Sulfate Isomer as an Effective and Removable Dispersant of Single-Walled Carbon Nanotubes. *Small*, 7, 2758-2768.
12. Wang, W.Z., **Li, W.**, Pan, X.Y., Li, C.M., Li, L.-J., Mu, Y.G., Rogers, J.A. and Chan-Park, M.B. (2011) Degradable Conjugated Polymers: Synthesis and Applications in Enrichment of Semiconducting Single-Walled Carbon Nanotubes. *Adv. Funct. Mater.*, 21, 1643-1651.
13. Yuan, W., **Li, W.**, Mu, Y. and Chan-Park, M.B. (2011) Effect of Side-Chain Structure of Rigid Polyimide Dispersant on Mechanical Properties of Single-Walled Carbon Nanotube/Cyanate Ester Composite. *ACS Appl. Mater. Interfaces*, 3, 1702-1712.

Bibliography

- [1] Alder, B. J. and Wainwright, T. E. (1959) Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics*, **31**(2), 459–466.
- [2] Rahman, A. (1964) Correlations in the Motion of Atoms in Liquid Argon. *Physical Review*, **136**, 405–411.
- [3] Li, W., Zhao, M., Zhao, X., Xia, Y., and Mu, Y. (2010) Hydrogen saturation stabilizes vacancy-induced ferromagnetic ordering in graphene. *Physical Chemistry Chemical Physics*, **12**, 13699–13706.
- [4] Swope, W. C., Andersen, H. C., Berens, P. H., and Wilson, K. R. (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, **76**(1), 637–649.
- [5] Leach, A. R. (2001) *Molecular Modelling: Principles and Applications*, Second edition, Pearson Education, .
- [6] Frenkel, D. and Smit, B. (2007) *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, .
- [7] Verlet, L. (Jul, 1967) Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review*, **159**, 98–103.
- [8] Reichl, L. E. (1980) *A modern course in Statistical Physics*, University of Texas Press, Austin, .
- [9] Kirkwood, J. G. (1935) Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics*, **3**(5), 300–313.
- [10] Cacace, M. G., Landau, E. M., and Ramsden, J. J. (1997) The Hofmeister series: salt and solvent effects on interfacial phenomena. *Quarterly Reviews of Biophysics*, **30**(03), 241–277.
- [11] Watlafer, D. B., Malik, S. K., Stoller, L., and Coffin, R. L. (1964) Nonpolar Group Participation in the Denaturation of Proteins by Urea and Guanidinium Salts. Model Compound Studies. *Journal of the American Chemical Society*, **86**(3), 508–514.

- [12] Hammes, G. G. and Schimmel, P. R. (1967) An Investigation of Water-Urea and Water-Urea-Polyethylene Glycol Interactions1. *Journal of the American Chemical Society*, **89**(2), 442–446.
- [13] Frank, H. S. and Franks, F. (1968) Structural Approach to the Solvent Power of Water for Hydrocarbons; Urea as a Structure Breaker. *The Journal of Chemical Physics*, **48**(10), 4746–4757.
- [14] Barone, G., Rizzo, E., and Vitagliano, V. (1970) Opposite effect of urea and some of its derivatives on water structure. *The Journal of Physical Chemistry*, **74**(10), 2230–2232.
- [15] Finer, E. G., Franks, F., and Tait, M. J. (1972) Nuclear magnetic resonance studies of aqueous urea solutions. *Journal of the American Chemical Society*, **94**(13), 4424–4429.
- [16] Leberman, R. and Soper, A. K. (1995) Effect of high salt concentrations on water structure. *Nature*, **378**(6555), 364–366.
- [17] Baldwin, R. L. (1996) How Hofmeister ion interactions affect protein stability. *Biophysical Journal*, **71**(4), 2056–2063.
- [18] Hribar, B., Southall, N. T., Vlachy, V., and Dill, K. A. (2002) How Ions Affect the Structure of Water. *Journal of the American Chemical Society*, **124**(41), 12302–12311.
- [19] Bennion, B. J. and Daggett, V. (2003) The molecular basis for the chemical denaturation of proteins by urea. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(9), 5142–5147.
- [20] Mountain, R. D. and Thirumalai, D. (2004) Alterations in Water Structure Induced by Guanidinium and Sodium Ions. *The Journal of Physical Chemistry B*, **108**(51), 19711–19716.
- [21] Castellino, F. J. and Barker, R. (1968) The denaturing effectiveness of guanidinium, carbamoylguanidinium, and guanylguanidinium salts. *Biochemistry*, **7**(11), 4135–4138.
- [22] Zhang, Y. and Cremer, P. S. (2010) Chemistry of Hofmeister Anions and Osmolytes. *Annual Review of Physical Chemistry*, **61**(1), 63–83.
- [23] Wolfram, S. (1984) Principles of Nucleic Acid Structure, New York: Springer, .
- [24] Bruce Alberts, Alexander Johnson, J. L. M. R. K. R. and Walter, P. (2002) Molecular Biology of the Cell, 4th edition, New York: Garland Science, .
- [25] Beckman, K. B. and Ames, B. N. (1997) Oxidative Decay of DNA. *Journal of Biological Chemistry*, **272**(32), 19633–19636.

- [26] Cadet, J., Delatour, T., Douki, T., Gasparutto, D., Pouget, J.-P., Ravanat, J.-L., and Sauvaigo, S. (1999) Hydroxyl radicals and DNA base damage. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **424**, 9–21.
- [27] Douki, T., Reynaud-Angelin, A., Cadet, J., and Sage, E. (2003) Bipyrimidine Photoproducts Rather than Oxidative Lesions Are the Main Type of DNA Damage Involved in the Genotoxic Effect of Solar UVA Radiation?. *Biochemistry*, **42**(30), 9221–9226.
- [28] Luan, B. and Aksimentiev, A. (2008) DNA Attraction in Monovalent and Divalent Electrolytes. *Journal of the American Chemical Society*, **130**(47), 15754–15755.
- [29] Mindich, L. (2012) *Viral Molecular Machines: Packaging in dsRNA Viruses*, Springer US, .
- [30] Hofmeister, F. (1888) Zur Lehre von der Wirkung der Salze [Title translation: About the science of the effect of salts]. *Archiv für rimentelle Pathologie und Pharmakologie*, **24**, 247–260.
- [31] Kunz, W., Henle, J., and Ninham, B. W. (2004) Zur Lehre von der Wirkung der Salze (about the science of the effect of salts): Franz Hofmeister’s historical papers. *Current Opinion in Colloid & Interface Science*, **9**(1-2), 19–37.
- [32] Zhang, Y. and Cremer, P. S. (2006) Interactions between macromolecules and ions: the Hofmeister series. *Current Opinion in Chemical Biology*, **10**(6), 658–663.
- [33] Kunz, W. (2010) Specific ion effects in colloidal and biological systems. *Current Opinion in Colloid & Interface Science*, **15**(1-2), 34–39.
- [34] Omta, A. W., Kropman, M. F., Woutersen, S., and Bakker, H. J. (2003) Influence of ions on the hydrogen-bond structure in liquid water. *The Journal of Chemical Physics*, **119**(23), 12457–12461.
- [35] Batchelor, J. D., Olteanu, A., Tripathy, A., and Pielak, G. J. (2004) Impact of Protein Denaturants and Stabilizers on Water Structure. *Journal of the American Chemical Society*, **126**(7), 1958–1961.
- [36] Zangi, R., Hagen, M., and Berne, B. J. (2007) Effect of Ions on the Hydrophobic Interaction between Two Plates. *Journal of the American Chemical Society*, **129**(15), 4678–4686.
- [37] David Van Der, S., Erik, L., Berk, H., Gerrit, G., Alan, E. M., and Herman, J. C. B. (2005) GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, **26**(16), 1701–1718.
- [38] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, **79**(2), 926–935.

- [39] Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, **23**(3), 327–341.
- [40] Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N [center-dot] log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, **98**(12), 10089–10092.
- [41] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *Journal of Chemical Physics*, **103**(19), 8577–8593.
- [42] Darden, T., Perera, L., Li, L., and Pedersen, L. (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure*, **7**(3), 55–60.
- [43] Shirts, M. R., Pitera, J. W., Swope, W. C., and Pande, V. S. (2003) Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *Journal of Chemical Physics*, **119**(11), 5740–5761.
- [44] MacCallum, J. L. and Tieleman, D. P. (2003) Calculation of the water-cyclohexane transfer free energies of neutral amino acid side-chain analogs using the OPLS all-atom force field. *Journal of Computational Chemistry*, **24**(15), 1930–1935.
- [45] Michael, R. S. and Vijay, S. P. (2005) Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *Journal of Chemical Physics*, **122**(13), 134508–134520.
- [46] Steinbrecher, T., Mobley, D. L., and Case, D. A. (2007) Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *Journal of Chemical Physics*, **127**(21), 214108–214120.
- [47] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, **81**(8), 3684–3690.
- [48] Kirkwood, J. G. and Buff, F. P. (1951) The Statistical Mechanical Theory of Solutions. I. *The Journal of Chemical Physics*, **19**(6), 774–777.
- [49] Ben-Naim, A. (1992) *Statistical Thermodynamics for Chemists and Biochemists*. Plenum Press, New York,.
- [50] Lockwood, D. M. and Rossky, P. J. (1999) Evaluation of Functional Group Contributions to Excess Volumetric Properties of Solvated Molecules. *The Journal of Physical Chemistry B*, **103**(11), 1982–1990.

- [51] Sangwai, A. V. and Ashbaugh, H. S. (2008) Aqueous Partial Molar Volumes from Simulation and Individual Group Contributions. *Industrial & Engineering Chemistry Research*, **47**(15), 5169–5174.
- [52] Parsegian, V. A., Rand, R. P., and Rau, D. C. (2000) Osmotic stress, crowding, preferential hydration, and binding: A comparison of perspectives. *Proceedings of the National Academy of Sciences*, **97**(8), 3987–3992.
- [53] Ben-Naim, A. and Marcus, Y. (1984) Solvation thermodynamics of nonionic solutes. *Journal of Chemical Physics*, **81**(4), 2016–2027.
- [54] Kyte, J. and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, **157**(1), 105 – 132.
- [55] Eisenberg, D. (1984) Three-Dimensional Structure of Membrane and Surface Proteins. *Annual Review of Biochemistry*, **53**(1), 595–623.
- [56] (2003) Evaluation of methods for measuring amino acid hydrophobicities and interactions. *Journal of Chromatography A*, **1000**(1-2), 637–655.
- [57] Eisenberg, D. and McLachlan, A. D. (1986) Solvation energy in protein folding and binding. *Nature*, **319**(16), 199–203.
- [58] Chebil, L., Chipot, C., Archambault, F., Humeau, C., Engasser, J. M., Ghoul, M., and Dehez, F. (2010) Solubilities Inferred from the Combination of Experiment and Simulation. Case Study of Quercetin in a Variety of Solvents. *Journal of Physical Chemistry B*, **114**(38), 12308–12313.
- [59] Joung, I. S. and Cheatham, T. E. (2008) Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *The Journal of Physical Chemistry B*, **112**(30), 9020–9041.
- [60] Ben-Naim, A. and Yaacobi, M. (1974) Effects of Solutes on the Strength of Hydrophobic Interaction and Its Temperature Dependence. *The Journal of Physical Chemistry*, **78**(2), 170–175.
- [61] Rodgers, J. L. and Nicewander, W. A. (1988) Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, **42**(1), 59–66.
- [62] Merchant, S. and Asthagiri, D. (2009) Thermodynamically dominant hydration structures of aqueous ions. *The Journal of Chemical Physics*, **130**(19), 195102–195111.
- [63] Lee, J. C. and Timasheff, S. N. (1974) Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride. *Biochemistry*, **13**(2), 257–265.

- [64] Prakash, V., Loucheux, C., Scheufele, S., Gorbunoff, M. J., and Timasheff, S. N. (1981) Interactions of proteins with solvent components in 8 urea. *Archives of Biochemistry and Biophysics*, **210**(2), 455–464.
- [65] Tirado-Rives, J., Orozco, M., and Jorgensen, W. L. (1997) Molecular Dynamics Simulations of the Unfolding of Barnase in Water and 8 M Aqueous Urea. *Biochemistry*, **36**(24), 7313–7329.
- [66] Caffisch, A. and Karplus, M. (1999) Structural details of urea binding to barnase: a molecular dynamics analysis. *Structure*, **7**(5), 477–488.
- [67] Wallqvist, A., Covell, D. G., and Thirumalai, D. (1998) Hydrophobic Interactions in Aqueous Urea Solutions with Implications for the Mechanism of Protein Denaturation. *Journal of the American Chemical Society*, **120**(2), 427–428.
- [68] Hua, L., Zhou, R., Thirumalai, D., and Berne, B. J. (2008) Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proceedings of the National Academy of Sciences*, **105**(44), 16928–16933.
- [69] Das, A. and Mukhopadhyay, C. (2009) Urea-Mediated Protein Denaturation: A Consensus View. *The Journal of Physical Chemistry B*, **113**(38), 12816–12824.
- [70] Wei, H., Yang, L., and Gao, Y. Q. (2010) Mutation of Charged Residues to Neutral Ones Accelerates Urea Denaturation of HP-35. *The Journal of Physical Chemistry B*, **114**(36), 11820–11826.
- [71] Robinson, D. R. and Jencks, W. P. (1965) The Effect of Compounds of the Urea-Guanidinium Class on the Activity Coefficient of Acetyltetraglycine Ethyl Ester and Related Compounds¹. *Journal of the American Chemical Society*, **87**(11), 2462–2470.
- [72] Mountain, R. D. and Thirumalai, D. (2003) Molecular Dynamics Simulations of End-to-End Contact Formation in Hydrocarbon Chains in Water and Aqueous Urea Solution. *Journal of the American Chemical Society*, **125**(7), 1950–1957.
- [73] Omta, A. W., Kropman, M. F., Woutersen, S., and Bakker, H. J. (2003) Negligible Effect of Ions on the Hydrogen-Bond Structure in Liquid Water. *Science*, **301**(5631), 347–349.
- [74] Pegram, L. M., Wendorff, T., Erdmann, R., Shkel, I., Bellissimo, D., Felitsky, D. J., and Record, M. T. (2010) Why Hofmeister effects of many salts favor protein folding but not DNA helix formation. *Proceedings of the National Academy of Sciences*, **107**(17), 7716–7721.
- [75] Gibb, C. L. D. and Gibb, B. C. (2011) Anion Binding to Hydrophobic Concavity Is Central to the Salting-in Effects of Hofmeister Chaotropes. *Journal of the American Chemical Society*, **133**(19), 7344–7347.

- [76] Tielrooij, K. J., Garcia-Araez, N., Bonn, M., and Bakker, H. J. (2010) Cooperativity in Ion Hydration. *Science*, **328**(5981), 1006–1009.
- [77] Moilanen, D. E., Wong, D., Rosenfeld, D. E., Fenn, E. E., and Fayer, M. D. (2009) Ionwater hydrogen-bond switching observed with 2D IR vibrational echo chemical exchange spectroscopy. *Proceedings of the National Academy of Sciences*, **106**(2), 375–380.
- [78] Turton, D. A., Hunger, J., Hefter, G., Buchner, R., and Wynne, K. (2008) Glasslike behavior in aqueous electrolyte solutions. *The Journal of Chemical Physics*, **128**(16), 161102–161105.
- [79] Rossky, P. J. (2008) Protein denaturation by urea: Slash and bond. *Proceedings of the National Academy of Sciences*, **105**(44), 16825–16826.
- [80] Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology*, Academic Press, **131**, 266–280.
- [81] Camilloni, C., Guerini Rocco, A., Eberini, I., Gianazza, E., Broglia, R. A., and Tiana, G. (2008) Urea and Guanidinium Chloride Denature Protein L in Different Ways in Molecular Dynamics Simulations. *Biophysical Journal*, **94**(12), 4654–4661.
- [82] Heyda, J., Kozisek, M., Bednarova, L., Thompson, G., Konvalinka, J., Vondrasek, J., and Jungwirth, P. (2011) Urea and Guanidinium Induced Denaturation of a Trp-Cage Miniprotein. *The Journal of Physical Chemistry B*, **115**(28), 8910–8924.
- [83] Lopez-Leon, T., Lopez-Lopez, J. M., Odriozola, G., Bastos-Gonzalez, D., and Ortega-Vinuesa, J. L. (2010) Ion-induced reversibility in the aggregation of hydrophobic colloids. *Soft Matter*, **6**, 1114–1116.
- [84] Case, D., Darden, T., Cheatham, T., III, Simmerling, C., Wang, J., Duke, R., Luo, R., Walker, R., Zhang, W., Merz, K., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Götz, A., Kolossváry, I., Wong, K., Paesani, F., Vanicek, J., Wolf, R., Liu, J., Wu, X., Brozell, S., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D., Mathews, D., Seetin, M., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. (2008) AMBER 10. *University of California, San Francisco*.
- [85] Dupradeau, F. Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., Lelong, D., Rosanski, W., and Cieplak, P. (2010) The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Physical Chemistry Chemical Physics*, **12**(28), 7821–7839.
- [86] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Rob, M. A., Cheeseman, J. R., Jr., J. A. M., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda,

- R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., K-
lene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C.,
Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R.,
Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P.,
Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C.,
Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Or-
tiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu,
G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T.,
Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W.,
Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., , and Pople, J. A. Gaussian 03.
Gaussian Inc. Wallingford CT 2003.
- [87] Das (1916) Gesetz der Reaktionsgeschwindigkeit und der Gleichgewichte in Gasen.
Bestätigung der Additivität von $C_v-3/2R$. Neue Bestimmung der Integrationskon-
stanten und der Moleküldurchmesser. *Zeitschrift für Anorganische und Allgemeine
Chemie*, **96**(1), 1–28.
- [88] Plaxco, K. W. and Gross, M. (2001) Unfolded, yes, but random? Never!. *Nature
Structural and Molecular Biology*, **8**(8), 659–660 10.1038/90349.
- [89] Shortle, D. and Ackerman, M. S. (2001) Persistence of Native-Like Topology in a
Denatured Protein in 8 M Urea. *Science*, **293**(5529), 487–489.
- [90] Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ue-
da, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-Range
Interactions Within a Nonnative Protein. *Science*, **295**(5560), 1719–1722.
- [91] Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) NMR determination of
residual structure in a urea-denatured protein, the 434-repressor. *Science*, **257**(5076),
1559–1563.
- [92] Hedoux, A., Krenzlin, S., Paccou, L., Guinet, Y., Flament, M.-P., and Siepmann,
J. (2010) Influence of urea and guanidine hydrochloride on lysozyme stability and
thermal denaturation; a correlation between activity, protein dynamics and confor-
mational changes. *Physical Chemistry Chemical Physics*, **12**(40), 13189–13196.
- [93] Mason, P. E., Neilson, G. W., Enderby, J. E., Sabounji, M.-L., Dempsey, C. E.,
MacKerell, A. D., and Brady, J. W. (2004) The Structure of Aqueous Guanidini-
um Chloride Solutions. *Journal of the American Chemical Society*, **126**(37), 11462–
11470.
- [94] Smith, L. J., Berendsen, H. J. C., and van Gunsteren, W. F. (2003) Computer Simula-
tion of Urea-Water Mixtures: A Test of Force Field Parameters for Use in Biomolec-
ular Simulation. *The Journal of Physical Chemistry B*, **108**(3), 1065–1071.
- [95] Stumpe, M. C. and Grubmüller, H. (2007) Aqueous Urea Solutions: Structure, Ener-
getics, and Urea Aggregation. *The Journal of Physical Chemistry B*, **111**(22), 6220–
6228.

- [96] Mason, P. E., Neilson, G. W., Dempsey, C. E., Barnes, A. C., and Cruickshank, J. M. (2003) The hydration structure of guanidinium and thiocyanate ions: Implications for protein stability in aqueous solution. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(8), 4557–4561.
- [97] Zangi, R., Zhou, R., and Berne, B. J. (2009) Urea’s Action on Hydrophobic Interactions. *Journal of the American Chemical Society*, **131**(4), 1535–1541.
- [98] Godawat, R., Jamadagni, S. N., and Garde, S. (2010) Unfolding of Hydrophobic Polymers in Guanidinium Chloride Solutions. *The Journal of Physical Chemistry B*, **114**(6), 2246–2254.
- [99] Fujita, T., Watanabe, H., and Tanaka, S. (2007) Effects of salt addition on strength and dynamics of hydrophobic interactions. *Chemical Physics Letters*, **434**(1-3), 42–48.
- [100] Breslow, R. and Guo, T. (1990) Surface tension measurements show that chaotropic salting-in denaturants are not just water-structure breakers. *Proceedings of the National Academy of Sciences*, **87**(1), 167–169.
- [101] Nandi, P. K. and Robinson, D. R. (1972) Effects of salts on the free energy of the peptide group. *Journal of the American Chemical Society*, **94**(4), 1299–1308.
- [102] Nandi, P. K. and Robinson, D. R. (1972) Effects of salts on the free energies of nonpolar groups in model peptides. *Journal of the American Chemical Society*, **94**(4), 1308–1315.
- [103] Lim, W. K., Rosgen, J., and Englander, S. W. (2009) Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proceedings of the National Academy of Sciences*, **106**(8), 2595–2600.
- [104] O’Brien, E. P., Dima, R. I., Brooks, B., and Thirumalai, D. (2007) Interactions between Hydrophobic and Ionic Solutes in Aqueous Guanidinium Chloride and Urea Solutions: Lessons for Protein Denaturation Mechanism. *Journal of the American Chemical Society*, **129**(23), 7346–7353.
- [105] Auton, M., Holthauzen, L. M. F., and Bolen, D. W. (2007) Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proceedings of the National Academy of Sciences*, **104**(39), 15317–15322.
- [106] Holthauzen, L. M. F., Rosgen, J., and Bolen, D. W. (2010) Hydrogen Bonding Progressively Strengthens upon Transfer of the Protein Urea-Denatured State to Water and Protecting Osmolytes. *Biochemistry*, **49**(6), 1310–1318.
- [107] Sagle, L. B., Zhang, Y., Litosh, V. A., Chen, X., Cho, Y., and Cremer, P. S. (2009) Investigating the Hydrogen-Bonding Model of Urea Denaturation. *Journal of the American Chemical Society*, **131**(26), 9304–9310.

- [108] Gao, M., She, Z., and Zhou, R. (2010) Key Residues that Play a Critical Role in Urea-Induced Lysozyme Unfolding. *The Journal of Physical Chemistry B*, **114**(47), 15687–15693.
- [109] Wolfenden, R., Andersson, L., Cullis, P. M., and Southgate, C. C. B. (1981) Affinities of amino acid side chains for solvent water. *Biochemistry*, **20**(4), 849–855.
- [110] Cabani, S., Gianni, P., Mollica, V., and Lepori, L. (1981) Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution. *Journal of Solution Chemistry*, **10**(8), 563–595.
- [111] Li, W. and Mu, Y. (2011) Hydration patterns and salting effects in sodium chloride solution. *The Journal of Chemical Physics*, **135**(13), 134502.
- [112] Noubigh, A., Abderrabba, M., and Provost, E. (2007) Temperature and salt addition effects on the solubility behaviour of some phenolic compounds in water. *The Journal of Chemical Thermodynamics*, **39**(2), 297–303.
- [113] Boots, H. M. J. and De Bokx, P. K. (1989) Theory of enthalpy-entropy compensation. *The Journal of Physical Chemistry*, **93**(25), 8240–8243.
- [114] Sharp, K. (2001) Entropy-enthalpy compensation: Fact or artifact?. *Protein Science*, **10**(3), 661–667.
- [115] Starikov, E. B. and Bengt, N. (2007) Enthalpy-Entropy Compensation: A Phantom or Something Useful?. *The Journal of Physical Chemistry B*, **111**(51), 14431–14435.
- [116] Yu, H.-A. and Karplus, M. (1988) A thermodynamic analysis of solvation. *The Journal of Chemical Physics*, **89**(4), 2366–2379.
- [117] Xiu, P., Yang, Z., Zhou, B., Das, P., Fang, H., and Zhou, R. (2011) Urea-Induced Drying of Hydrophobic Nanotubes: Comparison of Different Urea Models. *The Journal of Physical Chemistry B*, **115**(12), 2988–2994.
- [118] Saenger, W. (1984) Principles of Nucleic Acid Structure. *Springer, New York*,.
- [119] Sharp, K. A. and Honig, B. (1995) Salt effects on nucleic acids. *Current Opinion in Structural Biology*, **5**(3), 323–328.
- [120] Draper, D. E. (2004) A guide to ions and RNA structure. *Rna-a Publication of the Rna Society*, **10**(3), 335–343.
- [121] Woodson, S. A. (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Current Opinion in Chemical Biology*, **9**(2), 104–109.
- [122] Victor, A. B. (1997) DNA condensation by multivalent cations. *Biopolymers*, **44**(3), 269–282.

- [123] McFail-Isom, L., Hu, G. G., and Williams, L. D. (1998) The B-DNA Dodecamer at High Resolution Reveals a Spine of Water on Sodium. *Biochemistry*, **37**(23), 8341–8355.
- [124] Sines, C. C., McFail-Isom, L., VanDerveer, D., and Williams, L. D. (1998) Structure of the Potassium Form of CGCGAATTCGCG: DNA Deformation by Electrostatic Collapse around Inorganic Cations. *Biochemistry*, **37**(48), 16877–16887.
- [125] Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. *RNA*, **8**(3), 307–323.
- [126] Kankia, B. I. (2003) Binding of Mg^{2+} to single-stranded polynucleotides: hydration and optical studies. *Biophysical Chemistry*, **104**(3), 643–654.
- [127] Sosnick, T. R. and Pan, T. (2003) RNA folding: models and perspectives. *Current Opinion in Structural Biology*, **13**(3), 309–316.
- [128] Woodcock, C. L. and Dimitrov, S. (2001) Higher-order structure of chromatin and chromosomes. *Current Opinion in Genetics & Development*, **11**(2), 130–135.
- [129] Dixit, S. B., Beveridge, D. L., Case, D. A., Cheatham, T. E., Giudice, E., Lankas, F., Lavery, R., Maddocks, J. H., Osman, R., Sklenar, H., Thayer, K. M., and Varnai, P. (2005) Molecular Dynamics Simulations of the 136 Unique Tetranucleotide Sequences of DNA Oligonucleotides. II: Sequence Context Effects on the Dynamical Structures of the 10 Unique Dinucleotide Steps. *Biophysical Journal*, **89**(6), 3721–3740.
- [130] Varnai, P. and Zakrzewska, K. (2004) DNA and its counterions: a molecular dynamics study. *Nucleic Acids Research*, **32**(14), 4269–4280.
- [131] Auffinger, P. and Hashem, Y. (2007) Nucleic acid solvation: from outside to insight. *Current Opinion in Structural Biology*, **17**(3), 325–333.
- [132] Ponomarev, S. Y., Thayer, K. M., and Beveridge, D. L. (2004) Ion motions in molecular dynamics simulations on DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(41), 14771–14775.
- [133] Varnai, P. and Timsit, Y. (2010) Differential stability of DNA crossovers in solution mediated by divalent cations. *Nucleic Acids Research*, **38**(12), 4163–4172.
- [134] Denisov, V. P. and Halle, B. (2000) Sequence-specific binding of counterions to B-DNA. *Proceedings of the National Academy of Sciences*, **97**(2), 629–633.
- [135] E, T. and III, C. (2004) Simulation and modeling of nucleic acid structure, dynamics and interactions. *Current Opinion in Structural Biology*, **14**(3), 360 – 367.

- [136] Rueda, M., Cubero, E., Laughton, C. A., and Orozco, M. (2004) Exploring the Counterion Atmosphere around DNA: What Can Be Learned from Molecular Dynamics Simulations?. *Biophysical Journal*, **87**(2), 800 – 811.
- [137] Hamelberg, D., Williams, L. D., and Wilson, W. D. (2001) Influence of the Dynamic Positions of Cations on the Structure of the DNA Minor Groove: Sequence-Dependent Effects. *Journal of the American Chemical Society*, **123**(32), 7745–7755.
- [138] Duguid, J., Bloomfield, V. A., Benevides, J., and Thomas, G. J. (1993) Raman spectroscopy of DNA-metal complexes. I. Interactions and conformational effects of the divalent cations: Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd. *Biophysical Journal*, **65**(5), 1916–1928.
- [139] Colmenarejo, G. and Tinoco Jr, I. (1999) Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme. *Journal of Molecular Biology*, **290**(1), 119–135.
- [140] Misra, V. K. and Draper, D. E. (2001) A thermodynamic framework for Mg²⁺ binding to RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(22), 12456–12461.
- [141] Strauss, J. and Maher, L. (1994) DNA bending by asymmetric phosphate neutralization. *Science*, **266**(5192), 1829–1834.
- [142] Prabhu, N. V., Panda, M., Yang, Q., and Sharp, K. A. (2008) Explicit ion, implicit water solvation for molecular dynamics of nucleic acids and highly charged molecules. *Journal of Computational Chemistry*, **29**(7), 1113–1130.
- [143] Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T. E., Laughton, C. A., and Orozco, M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophysical Journal*, **92**(11), 3817–3829.
- [144] Macke Thomas, J. and Case David, A. Modeling Unusual Nucleic Acid Structures Vol. 682 of ACS Symposium Series, chapter 24, pp. 379–393 American Chemical Society (1997).
- [145] MacKerell, A. D., Bashford, D., Bellott, Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., et, and al. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B*, **102**(18), 3586–3616.
- [146] Toshiko, I. and Martin, K. (1991) Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Structure, Function, and Genetics*, **11**(3), 205–217.
- [147] Schlitter, J. (1993) Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chemical Physics Letters*, **215**(6), 617–621.

- [148] Andricioaei, I. and Karplus, M. (2001) On the calculation of entropy from covariance matrices of the atomic fluctuations. *Journal of Chemical Physics*, **115**(14), 6289–6292.
- [149] Subirana, J. A. and Soler-Lpez, M. (2003) CATIONS AS HYDROGEN BOND DONORS: A View of Electrostatic Interactions in DNA. *Annual Review of Biophysics and Biomolecular Structure*, **32**(1), 27–45.
- [150] Bleuzen, A., Pittet, P.-A., Helm, L., and Merbach, A. E. (1997) Water exchange on magnesium(II) in aqueous solution: a variable temperature and pressure 17O NMR study. *Magnetic Resonance in Chemistry*, **35**(11), 765–773.
- [151] Chiu, T. K. and Dickerson, R. E. (2000) 1 Å crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium. *Journal of Molecular Biology*, **301**(4), 915 – 945.
- [152] Stofer, E. and Lavery, R. (1994) Measuring the geometry of DNA grooves. *Biopolymers*, **34**(3), 337–346.
- [153] Buckin, V. A., Kankiya, B. I., Rentzeperis, D., and Marky, L. A. (1994) Mg²⁺ recognizes the sequence of DNA through its hydration shell. *Journal of the American Chemical Society*, **116**(21), 9423–9429.
- [154] Haran, T. E. and Mohanty, U. (2009) The unique structure of A-tracts and intrinsic DNA bending. *Quarterly Reviews of Biophysics*, **42**(01), 41–81.
- [155] Minasov, G., Tereshko, V., and Egli, M. (1999) Atomic-resolution crystal structures of B-DNA reveal specific influences of divalent metal ions on conformation and packing. *Journal of Molecular Biology*, **291**(1), 83 – 99.
- [156] Soler-López, M., Malinina, L., Liu, J., Huynh-Dinh, T., and Subirana, J. A. (1999) Water and Ions in a High Resolution Structure of B-DNA. *Journal of Biological Chemistry*, **274**(34), 23683–23686.
- [157] Strahs, D. and Schlick, T. (2000) A-tract bending: insights into experimental structures by computational models. *Journal of Molecular Biology*, **301**(3), 643–663.
- [158] Qiu, X., Andresen, K., Kwok, L. W., Lamb, J. S., Park, H. Y., and Pollack, L. (2007) Inter-DNA Attraction Mediated by Divalent Counterions. *Physical Review Letters*, **99**(3), 038104–038107.
- [159] Wilson, R. W. and Bloomfield, V. A. (1979) Counterion-induced condensation of deoxyribonucleic acid. A light-scattering study. *Biochemistry*, **18**(11), 2192–2196.
- [160] Gosule, L. C. and Schellman, J. A. (1976) Compact form of DNA induced by spermidine. *Nature*, **259**(5541), 333–335.
- [161] Ma, C., Sun, L., and Bloomfield, V. A. (1995) Condensation of Plasmids Enhanced by Z-DNA Conformation of d(CG)_n Inserts. *Biochemistry*, **34**(11), 3521–3528.

- [162] Sitko, J. C., Mateescu, E. M., and Hansma, H. G. (2003) Sequence-Dependent DNA Condensation and the Electrostatic Zipper. *Biophysical Journal*, **84**(1), 419–431.
- [163] Choi, J. K., Sargsyan, G., Shabbir-Hussain, M., Holmes, A. E., and Balaz, M. (2011) Chiroptical Detection of Condensed Nickel(II)-Z-DNA in the Presence of the B-DNA Via Porphyrin Exciton Coupled Circular Dichroism. *The Journal of Physical Chemistry B*, **115**(33), 10182–10188.
- [164] Grønbech-Jensen, N., Mashl, R. J., Bruinsma, R. F., and Gelbart, W. M. (1997) Counterion-Induced Attraction between Rigid Polyelectrolytes. *Physical Review Letters*, **78**(12), 2477–2480.
- [165] Shklovskii, B. I. (1999) Wigner Crystal Model of Counterion Induced Bundle Formation of Rodlike Polyelectrolytes. *Physical Review Letters*, **82**(16), 3268–3271.
- [166] Allahyarov, E., Gompper, G., and Lowen, H. (2004) Attraction between DNA molecules mediated by multivalent ions. *Physical Review E*, **69**(4), 041904.
- [167] Tan, Z.-J. and Chen, S.-J. (2006) Ion-Mediated Nucleic Acid Helix-Helix Interactions. *Biophysical Journal*, **91**(2), 518–536.
- [168] Rouzina, I. and Bloomfield, V. A. (1996) Macroion Attraction Due to Electrostatic Correlation between Screening Counterions. 1. Mobile Surface-Adsorbed Ions and Diffuse Ion Cloud. *The Journal of Physical Chemistry*, **100**(23), 9977–9989.
- [169] Kornyshev, A. A. and Leikin, S. (1997) Theory of interaction between helical molecules. *The Journal of Chemical Physics*, **107**(9), 3656–3674.
- [170] Behe, M. and Felsenfeld, G. (1981) Effects of methylation on a synthetic polynucleotide: the B–Z transition in poly(dG-m⁵dC).poly(dG-m⁵dC). *Proceedings of the National Academy of Sciences*, **78**(3), 1619–1623.
- [171] Thomas, T. J. and Bloomfield, V. A. (1985) Toroidal condensation of Z DNA and identification of an intermediate in the B to Z transition of poly(dG-m⁵dC).poly(dG-m⁵dC). *Biochemistry*, **24**(3), 713–719.
- [172] Ha, S. C., Lowenhaupt, K., Rich, A., Kim, Y.-G., and Kim, K. K. (2005) Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature*, **437**(7062), 1183–1186.
- [173] Li, H., Xiao, J., Li, J., Lu, L., Feng, S., and Dröge, P. (2009) Human genomic Z-DNA segments probed by the Z α domain of ADAR1. *Nucleic Acids Research*, **37**(8), 2737–2746.
- [174] Brzezinski, K., Brzuszkiewicz, A., Dauter, M., Kubicki, M., Jaskolski, M., and Dauter, Z. (2011) High regularity of Z-DNA revealed by ultra high-resolution crystal structure at 0.55 Å. *Nucleic Acids Research*, **39**(14), 6238–6248.

- [175] Li, W., Nordenskiöld, L., and Mu, Y. (2011) Sequence-Specific Mg^{2+} -DNA Interactions: A Molecular Dynamics Simulation Study. *The Journal of Physical Chemistry B*, **115**(49), 14713–14720.
- [176] Torrie, G. M. and Valleau, J. P. (1974) Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters*, **28**(4), 578 – 581.
- [177] Hub, J. S., de Groot, B. L., and van der Spoel, D. (2010) g_wham – A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *Journal of Chemical Theory and Computation*, **6**(12), 3713–3720.
- [178] Rau, D. C. and Parsegian, V. A. (1992) Direct measurement of temperature-dependent solvation forces between DNA double helices. *Biophysical Journal*, **61**(1), 260–271.
- [179] Koltover, I., Wagner, K., and Safinya, C. R. (2000) DNA condensation in two dimensions. *Proceedings of the National Academy of Sciences*, **97**(26), 14046–14051.
- [180] Schoenknecht, T. and Diebler, H. (1993) Spectrophotometric and kinetic studies of the binding of Ni^{2+} , Co^{2+} , and Mg^{2+} to poly(dG-dC)·poly(dG-dC). Determination of the stoichiometry of the Ni^{2+} -induced B→Z transition. *Journal of Inorganic Biochemistry*, **50**(4), 283–298.
- [181] Guéron, M., Demaret, J. P., and Filoche, M. (2000) A Unified Theory of the B-Z Transition of DNA in High and Low Concentrations of Multivalent Ions. *Biophysical Journal*, **78**(2), 1070–1083.
- [182] Gao, Y.-G., Sriram, M., and Wang, A. H.-J. (1993) Crystallographic studies of metal ion - DNA interactions: different binding modes of cobalt(II), copper(II) and barium(II) to N7 of guanines in Z-DNA and a drug-DNA complex. *Nucleic Acids Research*, **21**(17), 4093–4101.
- [183] Li, W., Zhou, R., and Mu, Y. (2012) Salting Effects on Protein Components in Aqueous NaCl and Urea Solutions: Towards Understanding of Urea Induced Protein Denaturation. *The Journal of Physical Chemistry B*, **116**(4), 1446–1451.
- [184] Dai, L., Mu, Y., Nordenskiöld, L., and van der Maarel, J. R. C. (2008) Molecular Dynamics Simulation of Multivalent-Ion Mediated Attraction between DNA Molecules. *Physical Review Letters*, **100**(11), 118301.
- [185] Rich, A. and Zhang, S. (2003) Z-DNA: the long road to biological function. *Nature Reviews Genetics*, **4**(7), 566–572.
- [186] Herbert, A., Alfken, J., Kim, Y.-G., Mian, I. S., Nishikura, K., and Rich, A. (1997) A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proceedings of the National Academy of Sciences*, **94**(16), 8421–8426.

- [187] Herbert, A., Schade, M., Lowenhaupt, K., Alfken, J., Schwartz, T., Shlyakhtenko, L. S., Lyubchenko, Y. L., and Rich, A. (1998) The Z α domain from human ADAR1 binds to the Z-DNA conformer of many different sequences. *Nucleic Acids Research*, **26**(15), 3486–3493.
- [188] Schwartz, T., Rould, M. A., Lowenhaupt, K., Herbert, A., and Rich, A. (1999) Crystal Structure of the Z α Domain of the Human Editing Enzyme ADAR1 Bound to Left-Handed Z-DNA. *Science*, **284**(5421), 1841–1845.
- [189] Schade, M., Turner, C. J., Kühne, R., Schmieder, P., Lowenhaupt, K., Herbert, A., Rich, A., and Oschkinat, H. (1999) The solution structure of the Z α domain of the human RNA editing enzyme ADAR1 reveals a prepositioned binding surface for Z-DNA. *Proceedings of the National Academy of Sciences*, **96**(22), 12465–12470.
- [190] Kim, Y.-G., Lowenhaupt, K., Maas, S., Herbert, A., Schwartz, T., and Rich, A. (2000) The Zab Domain of the Human RNA Editing Enzyme ADAR1 Recognizes Z-DNA When Surrounded by B-DNA. *Journal of Biological Chemistry*, **275**(35), 26828–26833.
- [191] Herbert, A. and Rich, A. (2001) The role of binding domains for dsRNA and Z-DNA in the in vivo editing of minimal substrates by ADAR1. *Proceedings of the National Academy of Sciences*, **98**(21), 12132–12137.
- [192] Rahmouni, A. and Wells, R. (1989) Stabilization of Z DNA in vivo by localized supercoiling. *Science*, **246**(4928), 358–363.
- [193] Lukomski, S. and Wells, R. D. (1994) Left-handed Z-DNA and in vivo supercoil density in the Escherichia coli chromosome. *Proceedings of the National Academy of Sciences*, **91**(21), 9980–9984.
- [194] Schroth, G. P., Chou, P. J., and Ho, P. S. (1992) Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. *Journal of Biological Chemistry*, **267**(17), 11846–11855.
- [195] Champ, P. C., Maurice, S., Vargason, J. M., Camp, T., and Ho, P. S. (2004) Distributions of Z-DNA and nuclear factor I in human chromosome 22: a model for coupled transcriptional regulation. *Nucleic Acids Research*, **32**(22), 6501–6510.
- [196] Koshland, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences*, **44**(2), 98–104.
- [197] Bae, S., Kim, D., Kim, K. K., Kim, Y.-G., and Hohng, S. (2010) Intrinsic Z-DNA Is Stabilized by the Conformational Selection Mechanism of Z-DNA-Binding Proteins. *Journal of the American Chemical Society*, **133**(4), 668–671.
- [198] Weiss, A. K. H., Hofer, T. S., Randolph, B. R., and Rode, B. M. (2012) Guanidinium in aqueous solution studied by quantum mechanical charge field-molecular dynamics (QMCF-MD). *Physical Chemistry Chemical Physics*, **14**, 7012–7027.

- [199] Mountain, R. D. and Thirumalai, D. (2003) Molecular Dynamics Simulations of End-to-End Contact Formation in Hydrocarbon Chains in Water and Aqueous Urea Solution. *Journal of the American Chemical Society*, **125**(7), 1950–1957.
- [200] N., P. O., P., K. Y., and K., K. M. (2004) X-ray phase analysis of structure of water salt NaCl-H₂O and KCl-H₂O. *Russian Journal of Applied Chemistry*, **77**, 1066–1069.
- [201] Bock, C. W., Markham, G. D., Katz, A. K., and Glusker, J. P. (2006) The Arrangement of First- and Second-shell Water Molecules Around Metal Ions: Effects of Charge and Size. *Theoretical Chemistry Accounts*, **115**, 100–112.
- [202] Schmidt, D. A., Birer, O., Funkner, S., Born, B. P., Gnanasekaran, R., Schwaab, G. W., Leitner, D. M., and Havenith, M. (2009) Rattling in the Cage: Ions as Probes of Sub-picosecond Water Network Dynamics. *Journal of the American Chemical Society*, **131**(51), 18512–18517.
- [203] Pegram, L. M. and Record, M. T. (2007) Hofmeister Salt Effects on Surface Tension Arise from Partitioning of Anions and Cations between Bulk Water and the Air-Water Interface. *The Journal of Physical Chemistry B*, **111**(19), 5411–5417.
- [204] Record, M. T., Guinn, E., Pegram, L., and Capp, M. (2013) Introductory Lecture: Interpreting and predicting Hofmeister salt ion and solute effects on biopolymer and model processes using the solute partitioning model. *Faraday Discuss.*, DOI: 10.1039/C2FD20128C.
- [205] Lo Nostro, P. and Ninham, B. W. (2012) Hofmeister Phenomena: An Update on Ion Specificity in Biology. *Chemical Reviews*, **112**(4), 2286–2322.