Palacky University

Faculty of Science

Department of Biophysics

Influence of Ionic Strength on Structure and Energetics of Selected A-RNA Motives

MASTER THESIS

Peter Repiščák

Supervisor: Doc. RNDr. Michal Otyepka, Ph.D.

Olomouc

2011

Declaration of the author

I declare that I have written this thesis by myself using cited sources. Neither the thesis nor any of its part was previously used for obtaining any academic degree.

In Olomouc, 29^{th} April 2011 Peter Repiščák .....

Acknowledgements

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. Especially, I would like to thank my supervisor doc. RNDr. Michal Otyepka, Ph.D. for many valuable discussions and whose stimulating suggestions and guidance helped me during research and writing of this thesis. At the same time, I have to thank to other members of the Department of physical chemistry for their support and friendly atmosphere.

Abstrakt

Ribonukleová kyselina (RNA) je veľmi dôležitá biomolekula, ktorá sa podieľa na mnohých základných biologických procesoch. Štruktúra a energetika RNA hrá nezameniteľnú úlohu pri interakciách RNA s ostatnými molekulami v bunke, ako aj ovplyvňuje funkčnosť mechanizmov, na ktorých sa RNA podieľa. V tejto práci sme využili metódy molekulovej dynamiky (MD) a MD spolu s metódou termodynamickej integrácie (TI) pre štúdium štruktúry a energetiky A-RNA duplexov. Testovali sme vplyv silových polí (*ff*99 and *ff*99bsc0), rôznych typov vodných modelov (SPC/E and TIP3P), pri nízkej koncentrácií iónov (Na^+) ako aj pri vysokej koncentrácií (KCl) na študovanú štruktúru. Ďalej sme skúmali rozdielne vlastnosti K^+ iónov so súborom parametrov navrhnutým Joungom a druhým súborom od Danga. Toto testovanie bolo s cieľom vylepšiť doteraz používaný simulačný protokol. Skúmali sme vplyv rastúcej iónovej sily na geometriu RNA štruktúry. Výskum v tejto oblasti môže poskytnúť veľmi cenné informácie o chovaní RNA v prehistorických podmienkach, kedy sa predpokladá, že prvotné jazierka obsahovali veľmi koncentrované roztoky solí. Posledná časť sa týkala skúmaniu energetiky, hlavne vplyvu mutácie guanínu na inozín. Tu sme testovali TI ako perspektívnu metódu, použiteľnú na získanie dôležitých informácií o energetike nukleových kyselín. V závere potvrdzujeme použitie doterajšieho simulačného protokolu s *ff*99bsc0 a ďalej navrhujeme jeho použitie s vodným typom SPC/E a Joungovými iónovými parametrami, táto kombinácia vykazovala najväčšiu stabilitu spomedzi všetkých simulovaných štruktúr. Pri rastúcej iónovej sile sme pozorovali pokles pre δ a χ torzné uhly a taktiež pokles pre štruktúrne parametre sklz (slide), skrut (twist) and vychýlenie z osi x (x-displacement). Pri najvyššej iónovej sile sa rozdiel v šírke malého a veľkého žliabku vytratil a celá štruktúra sa javila viac otvorená. Predbežne môžeme predpokladať, že pozorované zmeny sú spôsobené vychytávaním K^+ iónov vo veľkom žliabku a taktiež špecifickou väzbov K^+ na ne-Watson-Crickovské páry báz (G/U wobble and AH(+)/C). Na záver sme získali výsledky pre TI, semikvantitatívne porovnateľné s experimentálnymi hodnotami a týmto sme potvrdili potenciál TI metódy pre ďalšie použitie. Avšak, stretli sme sa aj s problémov nedostatočného vzorkovania pre niektoré dihedrálne uhly, hlavne χ dihedrál čo má priamy dopad na získané hodnoty energie.

Keywords: RNA, molekulárna dynamika, termodynamická integrácia, optimálny simulačný protokol, iónová sila, štruktúra a energetika, kvalita vzorkovanie

Abstract

Ribonucleic acid (RNA) is an essential molecule for almost all of the known organism. The structure and energetics of RNA play important roles in the RNA functionality. In order to reveal some of the RNA structural behavior and energetics, we have performed molecular dynamic simulations (MD) and MD with Thermodynamic Integration (TI) of A-RNA duplexes. We have examined an influence of different force fields (*ff*99 and *ff*99bsc0), water type (SPC/E and TIP3P), lower-salt conditions (Na^+) and higher salt conditions (KCl), as well as different set of ion parameters for K^+ ion (Joung and Dang) on simulated RNA structure. The aim of this comparison is to further improve used optimal simulation protocols. Moreover, the dependence of the A-RNA geometry on the increasing ionic strength was investigated, which can help in understanding of RNA behavior under the Early earth conditions. The last aim is to use TI as an effective tool for studies of RNA energetics, especially Guanine (G) to Inosine (I) mutation. The results propose that currently used optimal protocol with *ff*99bsc0 should be complemented with using of the SPC/E water type and the Joung set of parameters. The simulations with SPC/E and Joung parameters seems to be less affected by fluctuations of studied system and therefore yield more stable structures. Simulations with the changing ionic strength show for the highest ionic strength that the difference between major and minor groove widths vanished. The δ and χ torsions together with slide, twist and x-displacement were declining with rising ionic strength. And finally the structure seems to be more opened. This may be caused by tendency of structure to entrap K^+ ions into its major groove as well as specific binding of K^+ by non-canonical Watson Crick base pairs (G/U wobble and AH(+)/C). The TI is capable of providing reasonable results, semiquantitatively comparable with experimental values. Further, we experienced that an insufficient sampling of some torion angles, especially chi torsion, has impact on the obtained energy.

Keywords: RNA, molecular dynamics, thermodynamic integration, optimal protocol, salt conditions, energetics, sampling

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Our life is dependent on simple molecular mechanisms able to sustain its functions, carry basic information from generation to generation, and give organisms the possibility to evolve in time. Ribonucleic acid molecules (RNA) are one of the most important molecules in almost all the organisms. RNA is a more versatile "relative" of DNA, where its small chemical difference results in enormous diversity of final utilization. It is directly involved in catalytic process of proteosynthesis in ribosome and mediates information flow from sequence of bases in DNA to sequence of amino acids in proteins. Therefore, studying RNA molecule, its structure and functionality can bring many helpful information about basic mechanisms in cell, RNA role in the early life evolution as well as it can help in developing new strategies to treat for some diseases.

This work consists of three main parts, in which we try to understand some of RNA structural behavior and energetics. The aim of the first part is to investigate influence of different force fields, water models as well as small salt excess with different ionic type on the RNA structure. Besseova et al.[1] identified that A-RNA structure is more compact and stable under the Parmbsc0 (*ff*99bsc0) force field than under the Parm99 (*ff*99) force field. The Parmbsc0 in comparison to Parm99 force field suppresses excessive occurrence of pathological α/γ flips and introduce some visible structural changes. Moreover, it was suggested that simulations under lower-salt conditions are more realistic and close to X-ray structures, however, higher salt-conditions may stabilize the structure by preventing some deviations of the simulated structure from the X-ray structure. Despite this thorough study, the effects of different ionic types and water models has not been yet distinguished and therefore we try to decouple contributions of water models and ions on the structure and behavior of RNA.

The second part focuses on dependence of the RNA structure on increasing ionic strength. This may have implications in currently very popular study of the Origins of Life. Increased ionic strength in the primordial ponds might have helped the formation of RNA, its catalytic activity and influenced emergence of the early Earth life from RNA.

The aim of last part is to investigate energetics of Guanine (G) to Inosine (I) mutation in RNA. Further, an influence of ionic strength on the energetics and a capability of molecular dynamics to describe energy changes with respect to changing ionic strength is tested. It was observed by Siegfried and Bevilacqua[2] that Guanine to Inosine mutation, which introduces the loss of one hydrogen bond, results in much greater drop in stabilization energy that would correspond to the net energetic loss of one hydrogen bond. It is very interesting to explore other contributions that might have effect on this decrease in stabilization energy. Studying this problem can help to elucidate mechanisms of RNA folding as well as other aspects of RNA energetics and specificity.

In this work theoretical approaches involving computational method of Molecular Dynamics (MD), and Molecular Dynamics with Thermodynamic Integration (TI) are used. This methods possess great advantages over experimental methods as they can provide insight into the structural dynamics of molecular systems at an atomic level without the common problematic background interferences, that can disguise useful information, in more complex experiments. However, MD exhibit also some difficulties and artifacts because of its imperfections and approximate nature[3].

In the second chapter, a brief history of the RNA together with history of the Origins of Life is presented. Further, in this chapter we introduce RNA structure and some basic parameters used for its description. At the end of second chapter, short introduction to the RNA energetics is given. The third chapter presents theory behind MD and TI. Moreover, in this chapter the differences between some water and ion models are explained. The fourth chapter describes analytical tools used to analyze data from MD simulations. In this chapter also simulation protocols and overview of simulations carried out is presented. In the last chapters, results from simulations together with discussion are given. Part of these chapters is also conclusion, summarizing obtained results and proposing the aim of future research.

Chapter 2

RNA

2.1 History

History of nuleic acids can be dated back to 1868 when swiss biologist Friedrich Miescher discovered nuclein, a phosphorus-containing substance. About 30 years later scientists, analysing nuclein chemical structure, realised there are two main types of nucleic acid. First one that contain ribose as sugar was named ribonucleic acid (RNA) and second bears deoxyribose and was named deoxyribonucleic acid (DNA). In 1941 Caspersson[4] discovered relation between activity of the nucleoli and intense synthesis of the cytoplasmic ribonucleic acids and proposed association of RNA with the processes of protein synthesis. In that time scientists still believed that genetic information was stored in proteins. However, it was not until 1944 that the evidence was given by Oswald Avery, Colin MacLeod, and Maclyn McCarty for DNA as the bearer of genetic information[5]. This was further confirmed by Hershey[6] in 1952 that the DNA component of phage T2 carried out genetic specificity. In contrast to known DNA viruses, there are viruses such as plant viruses as well as viruses causing encephalitis and polio that instead of DNA contain RNA. The fact that RNA virus can also carry out function of genetic molecule was proven by Gierer[7] in 1956.

Story of unravelling the DNA double helical structure was accomplished by Watson-Crick[8] in 1953. With help of X-ray diffraction studies done by Rosalind Franklin and Maurice Wilkins they built up a model that follows also for the specific A - T and G - C base equivalences discovered by Chargaff. In the same year they proposed some genetic implications made from DNA structure[9] and year later the relation between DNA and RNA structure was shown by Rich and Crick[10].

In 1958 Crick[11] introduced the statement of central dogma of molecular biology as follows: DNA makes RNA makes protein. Where RNA serves as important link between each step of genetic information transfer. Initially, genetic information in nucleus is transcribed from DNA to messenger RNA (mRNA). The first notion of mRNA is from 1956 made by

Volkin and Astracham[12,13] when they observed, after infection of *Escherichia coli* with bacteriophage T2, an increased level in synthesis of T2-specific RNA. However, mRNA was described by Monod[14] five years later and in the same year an independent proof of its existence was given by Gros[15]. Next stage of information flow covers transfer of created mRNA to ribosomes, which contain ribosomal RNA (rRNA), in cytosol. Here, a genetic code is matched with code on incoming transfer RNA (tRNA) carrying amino acid and peptides are created. The evidence for ribonucleoprotein particles in cytosol, since 1958 known as ribosomes, was given in 1955 by Pallade[16] and they were described by Brenner[17] in 1961. The existence of tRNA was hypothesized by Francis Crick mid-1950s in his note about adaptor hypothesis. Crick anticipated existence of adaptor molecule mediating the translation of the RNA alphabet into the protein alphabet. Structure and properties of tRNA were described by Rich[18] in 1976.

Importance of RNA has grown as its secrets were further revealed over last fifty decades. Its enormous structural richness and functionality has been discovered. However, it was still quite surprising that not only proteins can serve as enzymes but also some RNA ribozymes[19,20] have catalytic activity. Another interesting fact that emerged from studying RNA is that from one gene you can obtain several proteins just by different splicing of pre-mRNA. Further, the central dogma proposed by Crick was completed with discovery of ability of some retroviroses to reverse the usual DNA-to-RNA transcription pathway using reverse transciptase. There exist family of RNA molecules that belong to aptamers[21], a molecules able to bind to a specific target ligand. Finally, new level of RNA significance was introduced by recently found property of small RNA molecules to regulate gene expression by post-transcriptional gene silencing. This effect was first observed in plants[22] and later also in model organism *Caenorhabditis elegans*[23].

2.2 RNA World - Origin of Life

The term "RNA world" describes a hypothetical era in evolution where RNA was the only encoding molecule supporting Darwinian evolution. The properties of RNA such as catalytic activity and ability to store genetic code found in modern organisms make it an ideal candidate for a molecule able to sustain early Earth cellular or pre-cellular life. At the beginning of the 1960s scientists started to speculate about idea of independent RNA life. This idea was first introduced by Rich[24] in 1962, followed by Woese[25] in 1967 and year later by Orgel[26] and Crick[27]. However, it has taken almost thirty years until the title "RNA World" has come to be associated with these speculations[28].

There are two main approaches to evidence about early RNA world. First one is using geological, geochemical and paleontological records together with molecular biology to move back in time from today organisms genome to the last universal common ancestor

(LUCA) and its RNA predecessor. Although some attempts have been successful this model still encounters some problems with decreasing quality of records found on the earth today when it moves far back in time. Hence, inferences made from these records cannot stand for compelling model for the RNA world. Second model focuses on opposite direction, utilizing information from astrochemistry about organic molecules in cosmos or identified in meteorites[29] and assuming their presence on the prebiological Earth, moving forward in time.

In this simulation of how evolution might have happened we need to start with creation of basic organic components, transforming them into RNA components, which are further assembled into polynucleotides. In addition, these polynucleotides must be able to self-replicate with sufficient efficiency and fidelity supporting evolution based on natural selection.

There is the famous Urey—Miller experiment[30,31] describing probable transformation of inorganic precursors to basic organic molecules. Many experiments and theories propose possible synthesis of nucleic acid building blocks from organic molecules. For example, the promising formose reaction suggested for ribose creation either stabilized by phosphate[32] or borate minerals[33]. The formose reaction for sugar synthesis was followed by more successful experiments for synthesis of nucleoside bases; adenine from hydrogen cyanide[34], purines and pyrimidines from cyanoacetylene and cyanate[35,36] and other approaches to synthesis of pyrimidines and purines[37–39]. Various attempts to couple bases with ribose to create nucleosides and nucleotides has been made[40–42]. Prior to the synthesis of polynucleotides, nucleosides must be converted to their activated forms of nucleoside 5'-polyphosphates and finally to nucleotides. There are many proposed ways how this activation could be achieved, for example, from nucleotides and inorganic trimetaphosphate[43] and urea-catalyzed phophorylation reaction[44,45]. The last step in RNA formation is the creation of 3',5'–linked oligonucleotides. If we expect that proteins were not present in prebiotic earth we need to consider a nonenzymatic routes for nucleotide polymerization. Two extensively studied groups of possible catalyst substitutions are adsorption to a specific mineral surface[46–48], providing enough regiospecificity, and metal ions[49,50].

To accomplish feasible model of the RNA World, it must have the capability to replicate synthesized oligonucleotides to support process similar to natural selection. During replication, new complementary oligonucleotide is synthesized by nonenzymatic mechanism under the direction of a preexisting oligonucleotide that serves as template. Nonenzymatic replication could be achieved by template-directed oligomerization of activated monomers[51,52] or by the ligation of short 3',5'–linked oligomers[53,54]. Another possibility is that RNA replication could be accomplished by RNA replicase ribozyme[55] catalyzing the template-directed polymerization of mono- or short oligonucleotides. Its chemistry may resemble chemistry of contemporary group I ribozymes[56–58].

In brief, it is important to mention also other hypothesis that were tested about alternative pre-RNA genetic system ("simple" RNA progenitors) which was later substituted by

RNA ("genetic takeover" hypothesis cite). For instance, possible candidates for RNA progenitors are pyranosyl-RNA (p-RNA)[59], threose nucleic acid (TNA)[60], peptide nucleic acid (PNA)[61] and many more.

The aim of the study of the RNA world is that it can help in undertanding the origins of life as well as elucidate many aspects of the modern RNA function; cellular processes involving RNA, ribozyme enzymology. The part which has not been thoroughly examined is how different ionic strengths influence the RNA structure, its stability, decomposition rate and hydrolytic behavior and how this might helped RNA to emerge from the primordial pond.

2.3 RNA

RNA can be found in important cell structures such as ribozymes, ribosomes, tRNA, mRNA. In addition, RNA plays important role in many processes ranging from transcription regulation[62,63], mRNA degradation[64,65] to catalytic RNA activity and RNA interference.

2.3.1 RNA structure

Ribonucleic acids are biopolymers assembled from four different types of nucleotides into various structural motifs and structures. Nucleotides are basic building blocks of RNA, which comprises a ribose sugar with nucleic acid base connected to carbon C1' and a phosphate group attached through phosphodiester bond to C5' carbon atom. This phosphate is linked to C3' hydroxyl group of adjacent nucleotide and together with ribose forms sugar-phosphate backbone. In the standard Watson-Crick (WC) base pairing, adenine (A) pairs with uracil (U) by forming two hydrogen bonds, and guanine (G) pairs with cytosine (C) by forming three hydrogen bonds. In comparison to deoxyribose sugar in DNA, RNA ribose contains 2'-hydroxyl group at C2' atom, which can act as hydrogen bond donor or acceptor and lead to diverse interactions with its environment. Moreover, this added reactive site together with unique types of base pairing found in RNA support higher diversity and richness of RNA structures. Polyanion character of RNA is mainly caused by a negative charge on each of the phosphate along the RNA backbone. The negative phosphate group is available for an interaction with solvent water molecules and ions in its environment.

Nucleic acid bases interact by stacking or by edge-to-edge interaction. The standard or canonical Watson-Crick base pairs are defined by Watson-Crick edge[a] interaction and with the glycosidic bonds oriented *cis* relative to the axis of the interaction[66]. These canonical Watson-Crick base pairs are characterized by their isostericity, a geometric relationship

[a] Watson-Crick edge comprises for purines A(N6)/G(O6), A or G(N1), A(C2)/G(N2) and for pyrimidines U(O4)/C(N4), U or C(N3), U or C(O2)

between some of the bases, that allows interchange one of the base for another without distorting the three-dimensional structure[67]. In addition, RNA can contain other types of WC base pairs as well as non-WC base pairs. Usually, non-WC base pairs have more significant impact on the specificity and functionality of RNA as they form structures that e.g. mediate long-range RNA-RNA interaction, RNA-protein recognition and create binding sites for small molecule ligands. In some cases protonated adenine paired with cytosine AH(+)/C and G/U wobble pair can substitute for canonical WC/WC pairs. However, these pairs are not self-isosteric that is they are not isosteric to its switched occurrence.

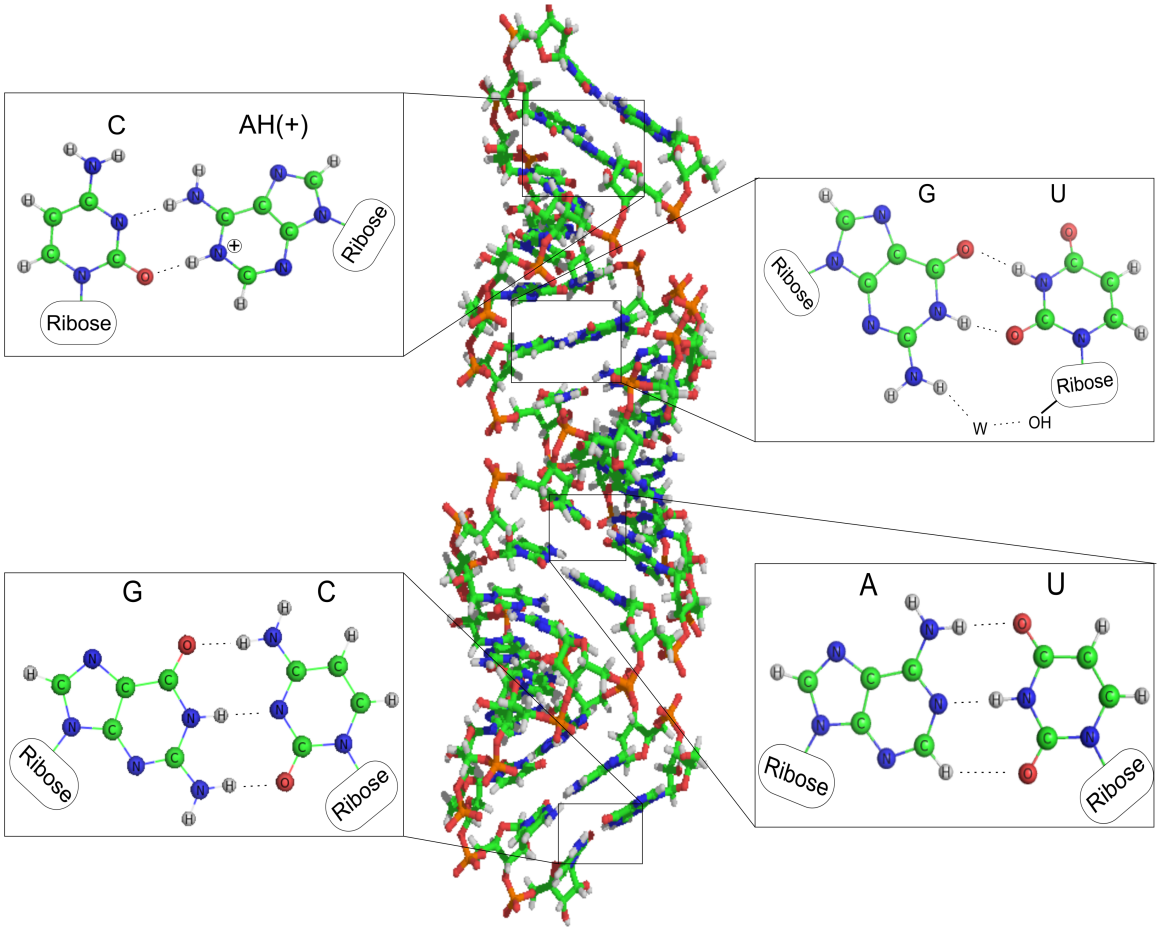

Figure 2.1: Typical A-form RNA helix structure with displayed protonated adenine C/AH(+), G/U wobble pair and canonical A/U and G/C pair. In case of C/AH(+), two hydrogen bonds are formed. In this figure, G/U wobble pair is in *cis* conformation bonded by three hydrogen bonds, where one hydrogen bond is between base and ribose mediated through bridging water molecule.

The linear sequence of covalently bonded nucleotides in a polynucleotide chain form a primary RNA structure. Secondary structure represents relative position of adjacent nucleotides, which is influenced mostly by hydrogen bonding and stacking. Examples of secondary structures are single strand, double strand, hairpin loops, bulge loops and pseu-

doknots. Tertiary structures are three-dimensional structures formed from the secondary structural motives associated together mostly through noncovalent interactions. Double helix is dominant tertiary structure. Helical RNA molecules typically adopt A-form helix structure (Figure 2.1). Highest structural order is quaternary structure, an arrangement of multiple folded RNA-RNA, RNA-DNA and RNA protein structures. Example of quaternary structure is ribosome.

2.3.2 Parameters describing structure

In helical nucleic acid, there are six torsion angles ($\alpha, \beta, \gamma, \delta, \varepsilon, \zeta$) along the sugar-phosphate backbone and one torsion angle (χ) is between the sugar and the base. Ribose sugar is described by pucker type, which is described by the atom or atoms that deviate from plane defined by the remaining furanose ring atoms. For example C2'-endo/C3'-exo states for C3' atom below and C2' atom above the plane defined by the C4', O4' and C1' atoms. In A-form RNA sugar pucker is predominantly in C3'-endo conformation (Figure 2.2). Further, more detailed description of sugar ring is provided with five internal torsional angles, amplitude and phase angle of the sugar ring pseudorotation.

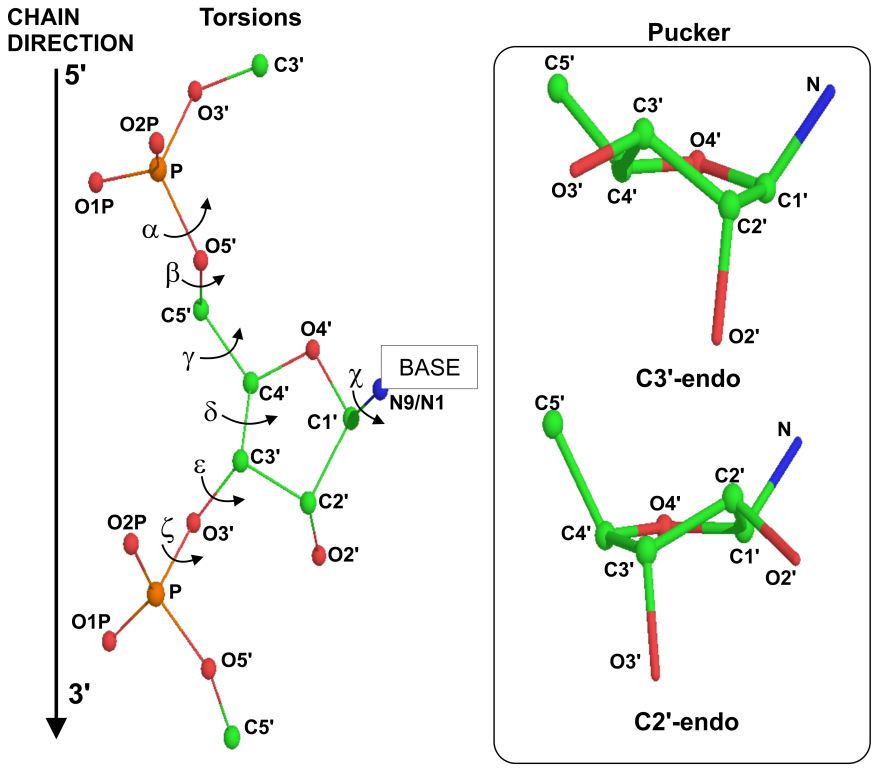

Figure 2.2: On the left side torsion angles are depicted. The right side depict C3'–endo and C2'–endo sugar pucker. Note that hydrogen atoms are not displayed.

Local parameters describing base to base relative position, within base-pair (bp), are

shear, buckle, stretch, propeller twist, stagger and opening. For description of the position and orientation of one bp relative to other two sets of parameters can be used. First one contains three translation parameters (shift, slide and rise) and three rotation parameters (tilt, roll and twist). In second set there are also three parameters for translation (x-displacement, y-displacement and helical rise) and three for rotation description (inclination, tip and helical twist)[68]. These two sets of parameters should be completely dependent, from one you can obtain second and vice versa. In case of A-RNA, there is clear difference between these two sets caused by bps not being perfectly parallel, but for B-form structure, rise and helical rise would be identical, so will be twist and helical twist. Some parameters are shown in Figure 2.3.

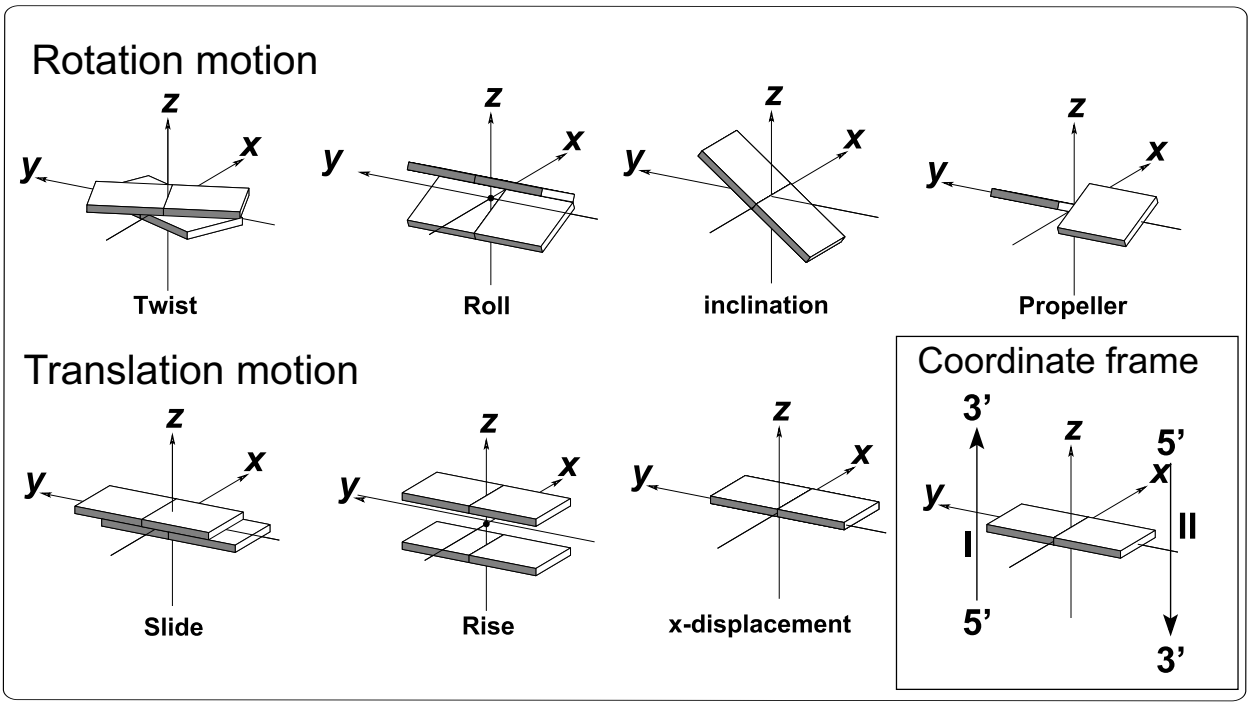

Figure 2.3: Selected parameters describing rotational and translational motion of bases, base pairs (figure adapted from[69]). Note that positive values of selected parameters are depicted.

Another important parameters in nucleic acid structure are widths of major and minor grooves (Figure 2.4). In RNA description these grooves are often denoted as deep (major) and shallow (minor) groove, however, in this work we use standard DNA notation of major and minor groove. Grooves can be described as asymmetric spaces formed by the edges of the base pairs and lie 180° opposite each other in the double helix. Major and minor grooves serve for complex interaction of nucleic acid with solvent molecules and proteins. For example, G/U wobble pair is able to trap water molecule in its minor groove (as can be also seen in Figure 2.1), whereas its major groove does not comprise any hydrophilic donor atom (N7, O6, O4)[70].

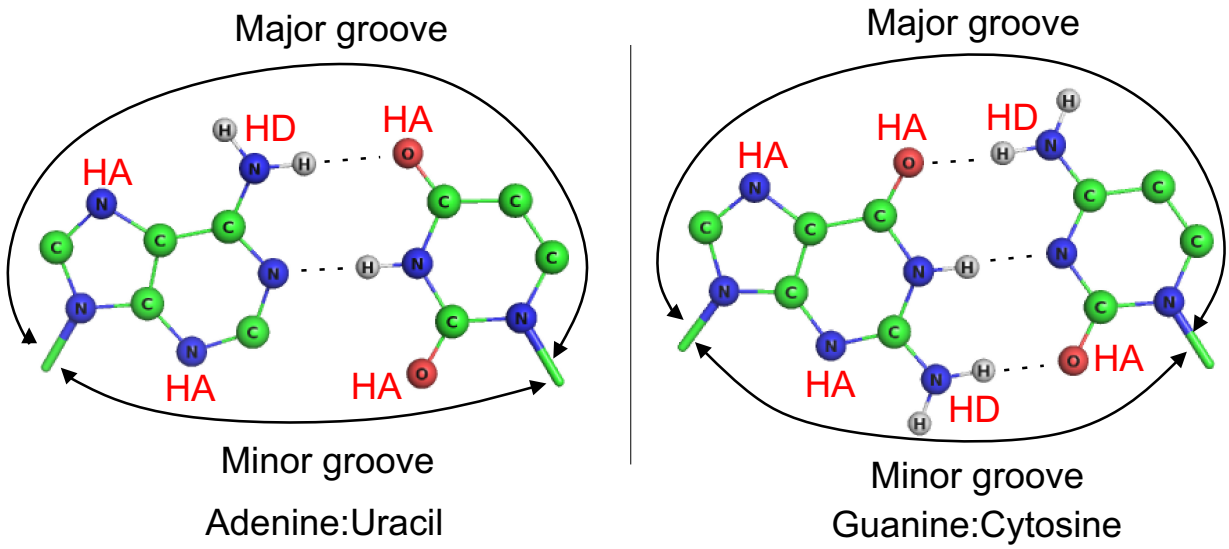

Figure 2.4: Major and minor groove for A:U and G:C base pairs. HA states for hydrogen acceptor and HD hydrogen donor atom.

2.3.3 Nucleic Acid Energetics

It is necessary to know a three dimensional structure of the RNA and DNA molecules in order to fully understand their function. For this purpose, energetics and specificity of interactions between building blocks are essential in complex molecule formation and folding[71]. Mutations of nucleic bases can significantly alter base pairing and thus folding cooperativity and stability of the whole system. For instance, during mutation of G to I, where amino group at C2 atom is replaced by hydrogen atom, one hydrogen bond of GC pair is lost. However, the net energetic loss and contribution to destabilization of duplex is larger that would correspond to loss of one hydrogen bond[72]. G to I mutation does not occur naturally, although it can be found in some RNA molecules, it is very unlikely to be incorporated to DNA due to reparation mechanisms and evolution pressures. In addition, different energetic loss was observed for RNA and DNA[73], i.e. for B-form and A-form duplex.

The challenge is to identify other yet unknown contributions causing these differences in destabilization energy.

Chapter 3

Molecular Dynamics

3.1 Theory

Molecular dynamics (MD) is an advanced computational method based on a rather simple physical model that can provide insight into the structural dynamics and energetic of molecular systems at an atomic level. The MD simulations are carried out from a starting set of atomic coordinates, which is usually obtained from X-ray crystallography, NMR experiments or can be constructed using known model systems for DNA and RNA with, for example, nucleic acid builder (NAB)[74]. The high quality starting structure is of an essential importance for MD sampling the proximity of X-ray potential minimum on a potential energy surface. The reason is that only a small conformational space is sampled around this crystal potential minimum and MD also cannot overcome large energy barriers (>5 kcal/mol)[3,75]. However, for experimental structures with higher resolution (<1.5 Å) MD can further improve parameters such as pairing, stacking patterns or minor backbone distortions[3]. Subsequently, the initial geometry of the studied system, placed in the environment of water and ions, undergoes 1-1000+ ns of dynamics simulated at room temperature and pressure. Immersing the structure into a water box with ions, we are trying to get as close as possible to real biological conditions.

The analysis of the outcome from simulations may provide detailed information about all the aspects of the time evolution (with sub-ps time resolution) of the three-dimensional structure and interactions within studied system[3]. In addition, MD simulation can reveal long-residency ion and water molecules that have influence on structure and function of the entire biomolecule. Despite enormous advance of MD in recent years, the method still encounters two basic limitations. The first one deals with problems of short simulation timescales, which result in limited sampling of conformational space. However, this limitation is slowly waning with faster computers and improvements in codes used in MD. Second, more critical limitation that cannot be easily overcome, is because of the approximate nature of

the force fields used in MD. In comparison to more accurate quantum mechanic calculations MD is able to work with bigger systems with reasonable computational cost.

3.1.1 Force fields

Force fields used in MD are simple analytical atomistic functions relating structure with potential energy[3,75]. In MD set of approximations are introduced to molecular system. Parameters for both bond length and valence angles are presented in form of harmonic springs, supplemented by torsion profiles for dihedral angles[75]. Atoms are approximated as Lennard-Jones spheres with constant point charges localized at the atomic centers. The force fields used are pairwise additive that neglect explicit polarization and charge-transfer effects though these contributions are included indirectly by some parameters. The result is that some important effects are not well described, i.e. inclusion of divalent ions, conformational flexibility of the anionic sugar-phosphate backbone, hydrogen bonding and solute polarization by solvent[3].

Standard AMBER[76] force fields *ff*99[77] and *ff*99bsc0[78] together with recently introduced force field *ff*99bsc0$_{\chi OL}$[79] were used for simulations. In *ff*99bsc0 the change in parametrization results in reduced pathological α/γ quick flips that seems are reversible in RNA, but in case of DNA are irreversible. The *ff*99bsc0$_{\chi OL}$ reparametrization of χ torsion profiles suppresses the high-*anti* χ states that often caused presence of ladder-like structures[80] in simulations of small RNAs, while it also modifies the *anti* vs *syn* balance and *syn* region shape.

3.1.2 Water and ion parameters

Parameters of the planar three-site SPC/E[81,82] and TIP3P[83,84] explicit water models, as used in the AMBER package are listed in Table 3.1 and shown in Figure 3.1.

Table 3.1: Parameters of explicit water models

Model	σ_O [a] [Å]	ε [a] [kJ.mol^{-1}]	r_{OH} [b] [Å]	q_O [c] [e]	q_H [c] [e]	θ_{HOH} [d] [°]
SPC/E	3.1656	0.1554	1.0000	-0.848	+0.424	109.47
TIP3P	3.1506	0.1521	0.9572	-0.834	+0.417	104.52

[a] ε and σ are the well depth and van der Waals radius Lennard-Jones parameters, respectively.
[b] r_{OH} is the oxygen-hydrogen pair distance
[c] q_O and q_H are the partial charges of hydrogen and oxygen, respectively.
[d] θ_{HOH} is the hydrogen-oxygen-hydrogen pair angle

Water models exhibit difference in physicochemical properties (Table 3.2) that affects interaction of solvent with solute. For example, a difference in diffusion coefficients has effect on kinetics of transfer between structural sub-states[85].

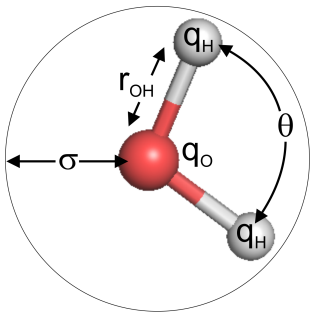

Figure 3.1: Illustrating picture of representation of three-site SPC/E and TIP3P water models

Table 3.2: Physicochemical properties of water models, where density and self-diffusion are under (298 K, 1 atm) (table adapted from[85])

	Models	
	SPC/E	TIP3P
dipole moment [D]	2.35	2.35
dielectric constant	71.0	82.0
density [g.cm^{-3}]	0.999 ± 0.010	0.986 ± 0.010
self-diffusion [10^{-5} cm^2.s^{-1}]	5.5	2.5
density maximum [K]	∼235.15	∼182.15
melting temperature [K]	214.95	145.55

Parameters of Na^+, K^+ and Cl^- ions used during simulations can be seen in Table 3.3. Ions with different parameters exhibit also different physicochemical properties. In the case of wrong usage of combination rules and ionic parameters this may lead to significant simulation artifacts like formation of salt aggregates at low concentrations[86]. Therefore, optimal set of ion parameters for some common water models was proposed by Joung[87],recently.

Table 3.3: Parameters of the ions used in simulations

Ion	Dang/Åqvist[a]		Joung SPC/E[b]		Joung TIP3P[c]	
	σ [Å]	ε [kcal.mol^{-1}]	σ [Å]	ε [kcal.mol^{-1}]	σ [Å]	ε [kcal.mol^{-1}]
Na^+	1.8680	0.0028	1.212	0.3526	1.369	0.0874
K^+	1.8687	0.1000	1.593	0.4297	1.705	0.1937
Cl^-	2.4700	0.1000	2.711	0.0128	2.513	0.0356

[a] Na^+, K^+ and Cl^- used with Aqvist[88] (default AMBER), Dang[89] and Smith[90] parameters, respectively.
[b] Joung[87] Na^+, K^+ and Cl^- parameters optimized for SPC/E waters
[c] Joung[87] Na^+, K^+ and Cl^- parameters optimized for TIP3P waters

3.1.3 Potential energy and motion equations

Propagation of the studied system in time is acquired by calculating the Newtonian motion equation (3.1). Initial velocities are assigned to each of the atoms according to Maxwell–Boltzmann distribution for simulated temperature.

$$\vec{f}_i = m_i \ddot{\vec{r}}_i = -\nabla \mathcal{U} \tag{3.1}$$

where forces $\vec{f}_i$ acting on the atoms are usually derived from a potential energy $\mathcal{U}$ which is given as a sum of individual energy terms for covalent and noncovalent contributions (3.2):

$$\mathcal{U} = \mathcal{U}_{\text{covalent}} + \mathcal{U}_{\text{noncovalent}} \tag{3.2}$$

Continuing to discuss, for simplicity, a system composed of atoms with coordinates $r_1, \cdots, r_N$ and potential energy $\mathcal{U}(r_1, \cdots, r_N)$, we introduce the atomic momenta $p_1, \cdots, p_N$, in terms of which the kinetic energy may be written $\mathcal{K}(p_1, \cdots, p_N) = \sum_{i=1}^{N} |p_i|^2/2m_i$. The covalent part of potential energy consist of bond, angle and torsion angle contributions. Illustrating figure of covalent contibutions are shown in Figure 3.2 and related equations (3.3):

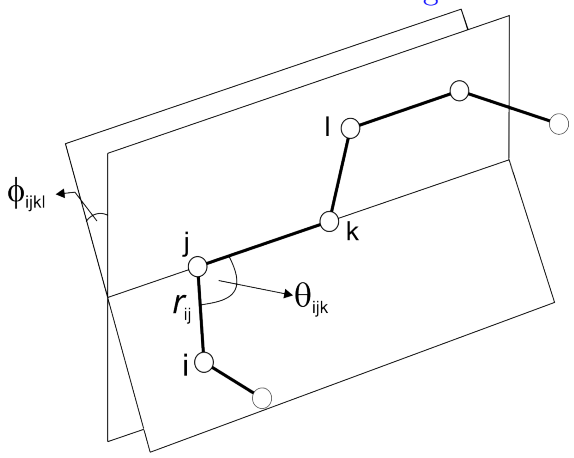

Figure 3.2: Illustrating picture for the definition of bond r_{ij} (interatomic distance), bend angle θ_{ijk} and torsion angle ϕ_{ijkl}.

$$\begin{aligned} \mathcal{U}_{\text{covalent}} &= \frac{1}{2}\sum_{\text{bonds}} k_{ij}(r_{ij}-r_{ij0})^2 \\ &+ \frac{1}{2}\sum_{\substack{\text{bend}\\ \text{angles}}} k_{ijk}(\theta_{ijk}-\theta_{ijk0})^2 \\ &+ \frac{1}{2}\sum_{\substack{\text{torsion}\\ \text{angles}}}\sum_{n} k_{ijkl}(1+\cos(n\phi_{ijkl}-\phi_{ijkl0})) \end{aligned} \tag{3.3}$$

The noncovalent part of potential energy is in the form of pair additive potential where three-body and higher order interactions are neglected. Model of pair potential can be seen in (3.4) and real noncovalent potential used in simulations (3.5) consist of vdW part, which has form of the Lennard-Jones potential, and Coulomb part for electrostatic charges.

$$\mathcal{U}(r_1,\cdots,r_{\text{N}}) = \sum_{i<j} u(r_i,r_j) \tag{3.4}$$

$$\mathcal{U}_{\text{noncovalent}} = \underbrace{4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12}-\left(\frac{\sigma}{r_{ij}}\right)^{6}\right]}_{\text{Lennard-Jones potential}} + \underbrace{\frac{1}{4\pi\varepsilon_0}\frac{q_iq_j}{r_{ij}}}_{\text{Coulomb potential}} \tag{3.5}$$

AMBER uses leapfrog algorithm, which step-by-step numerically integrate coupled differential equations of motion (3.6) and (3.7). In case of leapfrog, positions are defined at times $t_i, t_{i+1}, t_{i+2}, \ldots$, spaced at constant intervals dt, while the velocities are defined at times halfway in between, indicated by $t_{i-1/2}, t_{i+1/2}, t_{i+3/2}, \ldots$.

$$\text{r}(t+\delta t) = \text{r}(t) + \text{v}(t+\frac{1}{2}\delta t)\delta t \tag{3.6}$$

$$\text{v}(t+\frac{1}{2}\delta t) = \text{v}(t-\frac{1}{2}\delta t) + \frac{1}{m}\text{F}(t)\delta t \tag{3.7}$$

3.2 Thermodynamic Integration

The available experimental techniques to study energetics and stability of double stranded nucleic acids can be efficiently complemented by MD combined with method of Thermodynamic Integration (TI). Furthermore, TI is a powerful tool that can provide insight into destabilization at atomic level, decompose obtained interaction energies to particular components and evaluate their contribution to total interaction energy. In this work the thermodynamic parameters of a difference in Gibbs free energy $\Delta\Delta G$, which serves as an indicator

of duplex stability, obtained from the TI are compared with experimental values from the UV melting experiments. There is a significant difference between experimental and theoretical way how to obtain final free energy differences $\Delta\Delta G$ of G to I mutation, which are, however, equivalent due to thermodynamic cycle (Figure 3.3). In experiments, the thermodynamic properties are calculated from data obtained by melting the ordered RNA (DNA) duplex structure comprising G into the disordered, denatured single strands. The same is done for I and $\Delta\Delta G$ is then calculated as in (3.8).

$$\Delta\Delta G = \Delta G_{1,\mathrm{exp}} - \Delta G_{2,\mathrm{exp}} \tag{3.8}$$

However, in theoretical approach the melting procedure is very difficult if not impossible. Instead, G in double strand is directly transformed to I in double strand and same is done for single strands. The comparison of about mentioned experimental and theoretical approach can be seen in Figure 3.3.

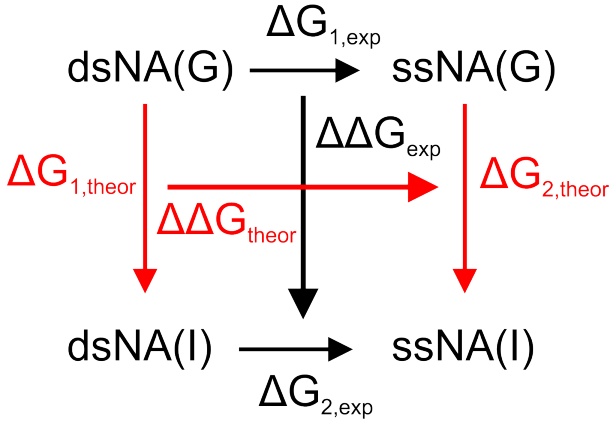

Figure 3.3: Comparison of thermodynamic cycles for experimental and theoretical approach to $\Delta\Delta G$; where black arrows represent experimental and red arrows theoretical approach.

And equivalence between results obtained by theoretical and experimental method is proven in following equation(3.9):

$$\Delta G_{1,\mathrm{exp}} + \Delta G_{2,\mathrm{theor}} - \Delta G_{2,\mathrm{exp}} - \Delta G_{1,\mathrm{theor}} = 0 \tag{3.9}$$

which results in(3.10):

$$\Delta\Delta G_{\mathrm{exp}} = \Delta\Delta G_{\mathrm{theor}} \tag{3.10}$$

In order to obtain $\Delta\Delta G$ of A→B in TI a set of simulations is carried out according to hybrid potential:

$$V^{\mathrm{tot}} = (1-\lambda)V_{\mathrm{A}} + \lambda V_{\mathrm{B}} \tag{3.11}$$

where λ is mixing parameter evolving system from A to B. V_{A} is the potential with the original Hamiltonian, and V_{B} is the potential with the perturbed Hamiltonian.

From simulations a pseudo force acting on λ ($\left\langle \frac{\partial V^{tot}}{\partial \lambda} \right\rangle$) is obtained, which after integration yields all no pressure-volume work under NpT that is ΔG. AMBER module *Sander* primarily provides the tools to collect the statistics that are further processed numerically to estimate the $\Delta\Delta G$. Figure 3.4 illustrates the integral of the ensemble average of the derivative of the Hamiltonian ($\left\langle \frac{\partial V}{\partial \lambda} \right\rangle$) along a reaction coordinate of G to I transformation.

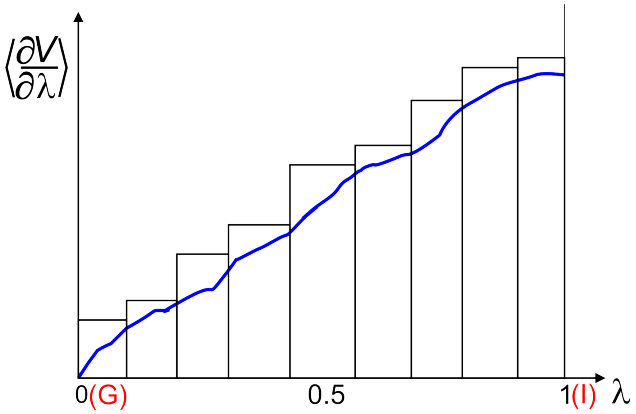

Figure 3.4: Illustrating figure of energy integral for G to I transformation

Integral is over set of nine λ mixing parameters as the system evolves from G to I. The integral is further estimated using numerical nine point Gaussian quadrature (Equation 3.12).

$$\Delta G = \int_0^1 \left\langle \frac{\partial V}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda \approx \sum_{\mathrm{i}} w_{\mathrm{i}} \left\langle \frac{\partial V}{\partial \lambda} \right\rangle_{\mathrm{i}} \tag{3.12}$$

In this equation w_{i} represents weight of selected quadrature point.

Chapter 4

Analytical Tools and Simulation Protocols

4.1 Analytical Tools

Calculated trajectories were analyzed using Ptraj module from AMBER package. Ptraj was also used for calculation of Root Mean Square deviation (RMSd) over backbone atoms. RMSd is used as a measure of the differences in position between selected atoms in starting structure and the same atoms in each time frame. Furthermore, Ptraj protocols were used for production of density maps of ions. Density map represents the histogram of atoms, molecules (ions, waters) on the 3D grid around solute molecule. Trajectories were visualized in VMD[91] and structures were displayed in PyMOL[92]. Moreover, VMD was used for visualization of density maps around RNA. Analysis of torsion angles and structural parameters was done utilizing X3DNA software[93].

4.2 Simulation Protocol

Note that ångström (Å) unit used in this work is commonly used unit for distance description not only in MD, but also in X-ray experiments etc. It can be converted to standard SI unit meter (m) as follows: 1 Å = 10^{-10} m.

Influence of ionic strength, different ionic parameters and force fields were tested on double stranded 18 base pair (bp) A-RNA helix[94]. This A-RNA occurs in form of hairpin in natural environment at dilute concentrations. However, in crystals it forms an 18-mer self-complementary double helix with 19th nucleotide flipped out of the helix, which was cut off (Figure 4.1).

Head of a hairpin is recognized by a zing-finger protein and interestingly its binding region preserve also in the duplex structure. Structure contains four AH(+)/C base pairs,

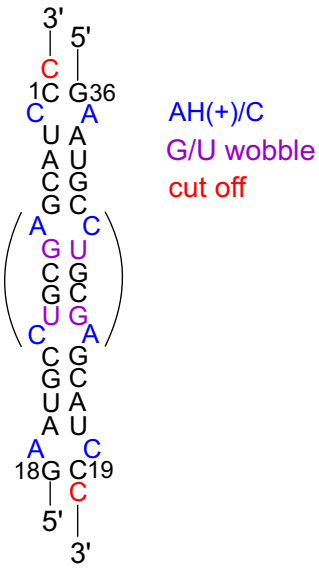

Figure 4.1: RNA simulation structure as used in simulations. The cytosine residue (marked red) was cut off.

two G/U wobble pairs and the rest are canonical WC/WC base pairs. Starting geometry for MD simulations was taken from crystal structure with 1.6 Å resolution from Nucleic Acid Database (deposit number 1KFO). Initially, the X-ray structure was processed in AMBER Leap module. It was immersed in a rectangular SPC/E or TIP3P water box and negative charge on phosphates was neutralized by Na^+ counter ions. In case of salt excess, K^+ and Cl^- ions were used. System prepared in this way was minimized through several minimization steps prior to the simulation. Special minimization step was used for high ionic strength, where solute and solvent were hold with big forces and ions were minimized with decreasing force constants. In next step, system was slowly heated up to final simulation temperature 298.15 K and pressure was set to 1 atm (10^5 Pa). Simulations were performed with *Sander* module of AMBER 10.04 program package using 2-fs integration steps. The length of simulations was 200 ns, though only 50 ns trajectories were available for some simulations in time of analysis.

4.3 Simulation Protocol of TI

First of all , an initial A-RNA structure for TI was built using web based Nucleic Acid Builder[74]. The 12-mer double stranded A-RNA (dsARNA) with nucleic acid sequence taken from stem (see Figure 4.2 below, residue 1-8) of RNA with addition of two GC pairs at both ends to prevent base fraying of terminal base pairs[1]. Secondly, created RNA.pdf file with guanine at RES 6 (G6) was further processed in preparatory Amber program Leap: Na^+,Cl^- ions were added and solvated with a rectangular TIP3P water box. In third step, residue

G6 (in original hairpin denoted as residue 4, but notice that two GC pairs were added) was changed to Inosine (I6) and thus prepared structures were minimized. One set of coordinates is needed for thermodynamic integration calculations. Hence, the coordinates were prepared in following way: After minimization Guanine coordinates were used as a reference frame, C2-H2 distance of minimized Inosine was measured. In order to create I coordinates, the hydrogen atoms of N2 amino group of minimized Guanine structure were deleted and the N2 nitrogen was renamed to H2 hydrogen and its position was shifted toward C2 carbon into a distance of C2-H2 from minimized I. This results in two identical coordinate files one for Guanine and one for Inosine. The nine point Gaussian quadrature was used for the integral evaluation, therefore for each of the lambda mixing parameter one simulation was carried out (together nine multi-sander jobs).

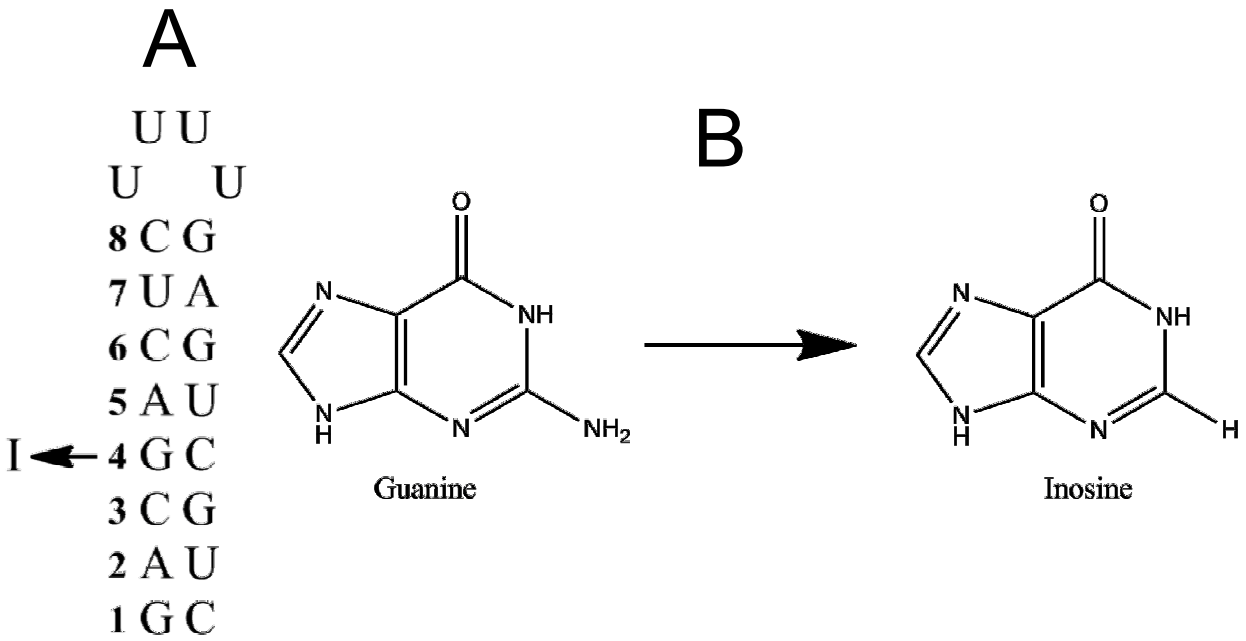

Figure 4.2: (A) Hairpin structure as used in experiments by Siegfried and Bevilacqua. (B) Guanine to Inosine mutation; NH_2 group at C2 atom is substituted by H atom.

Together 27 separate simulations were carried out for each of the ionic strength.

All Thermodynamic free energy calculations were carried out using AMBER 10.04 program package with force field *ff*99bsc0$_{\chi OL}$[78,95]. Two prmtop files were necessary, first for starting state containing Guanine (V_G) and second corresponding to the ending state with Inosine (V_I). However, only one set of coordinates that are propagated in the molecular dynamics algorithm are needed. Softcore potential algorithm was used to handle problem with different number of atoms in Guanine and Inosine. Using softcore potentials an unique atoms, disappearing NH_2 group of G and appearing H atom of I, can be present at the same time. This is provided by modified version of the vdW equation, which ensures that

non-bonded interactions of unique atoms with their common atom neighbors are smoothly switched off. For more detailed description see Amber 10 manual.

A problem for some lambda values occurred with modified vdW interaction for unique atoms as coulombic interaction prevailed. This resulted in strong attraction of adjacent water molecule to disappearing/appearing atoms and crash of the simulation. Hence, a charge neutralizing bypass (G$\rightarrow$G$_0$ $\rightarrow$I$_0$ $\rightarrow$I) was introduced, first neutralizing charges at Guanine NH_2 group then morph G to I (G$_0$ $\rightarrow$I$_0$) including soft-core potentials with zero charge at Inosine H atom and finally reset the charge at H (I$_0$ $\rightarrow$I). The above mentioned "obstacles" are easily solved in one step transformation using soft core electrostatics in Amber 11 (for the future use).

Following table (Table 4.1) lists overview of performed simulations for methodological part, evaluation of influence of ionic strength and thermodynamic integration.

Table 4.1: Overview of simulations performed with different ionic strength, ionic types, water types and force fields. First part of table shows simulations used for comparison of different water models, ionic parameters and force fields. In second part of table simulations mostly used for evaluation of ionic strength influence on structural parameters are listed. Third part lists simulations used in Thermodynamic Integration. RMSD values are calculated over 3 to 16 base pairs, omitting two terminal base pairs at both ends, with respect to the first snapshot.

Simulation	Length [ns]	Ionic strength [mM]	Ion type	Wat box	Force field	RMSD [Å]
Na99SPC	50	355	Na Aqvist	SPC/E	parm99	2.2 ± 0.6
NaB0SPC	50	356	Na Aqvist	SPC/E	parmbsc0	1.8 ± 0.4
Na99T3P	50	350	Na Aqvist	TIP3P	parm99	3.0 ± 0.8
NaB0T3P	50	351	Na Aqvist	TIP3P	parm99	1.8 ± 0.4
KCl99DangSPC	50	708	KCl Dang	SPC/E	parm99	1.8 ± 0.4
KCl99JoungSPC	50	710	KCl Joung	SPC/E	parm99	2.7 ± 0.4
KClB0JoungSPC	50	710	KCl Joung	SPC/E	parmbsc0	2.4 ± 0.3
05MKClB0SPC	200	498	KCl Joung	SPC/E	parmbsc0	1.9 ± 0.3
1MKClB0SPC	200	1005	KCl Joung	SPC/E	parmbsc0	2.5 ± 0.3
2MKClB0SPC	200	1941	KCl Joung	SPC/E	parmbsc0	2.3 ± 0.4
3MKClB0SPC	200	2804	KCl Joung	SPC/E	parmbsc0	1.9 ± 0.3
4MKClB0SPC	200	3590	KCl Joung	SPC/E	parmbsc0	2.2 ± 0.3
5MKClB0SPC	200	4367	KCl Joung	SPC/E	parmbsc0	2.3 ± 0.4
ds/ssTI100	200	113	NaCl Joung	TIP3P	parmbsc0OL	
ds/ssTI200	200	206	NaCl Joung	TIP3P	parmbsc0OL	
ds/ssTI500	200	497	NaCl Joung	TIP3P	parmbsc0OL	
ds/ssTI1000	200	1021	NaCl Joung	TIP3P	parmbsc0OL	

Chapter 5

Experimental Part and Results

5.1 Results and Discussion

5.1.1 Different force fields, water models, salt excess and ionic type

Two different force fields, *ff*99 and *ff*bsc0 are compared. Further, two different water models TIP3P and SPC/E and Na^+ and K^+ ions were used. In case of K^+ also Dang and Joung set of parameters for ions were compared.

Besseova et al. identified dependence of the A-RNA geometry (sequence-dependent) on the used force field as well as they observed that modest salt excess may in some cases prevent deviations of simulated structures from experimental[1]. One of the results was that *ff*99bsc0 force field yield, in comparison with *ff*99, more compact structure by its stabilization. For *ff*99bsc0 the major groove is narrower and they also observed small increase in the helical twist. Moreover, under low salt conditions *ff*99bsc0 yields lower RMSD values than *ff*99. The same drop was observed also for high salt conditions, furthermore RMSD experienced additional reduction in comparison to low salt conditions. In our simulations we observed similar behavior of averaged RMSD for *ff*99bsc0 and *ff*99, however, surprisingly under higher salt conditions these values did not further declined to lower values, but instead, exhibited higher values of averaged RMSD (Figure 5.1 and in previous chapter in Table 4.1 averaged RMSD values are given.).

The difference between averaged values of RMSD for *ff*99bsc0 and *ff*99 with Na^+ and SPC/E is $\Delta_{\mathrm{RMSD}} = 0.4$ Å. In case of similar simulation, but with TIP3P water model, this difference is three times bigger $\Delta_{\mathrm{RMSD}} = 1.2$ Å. This is caused, when we compare RMSD time progress for both simulations Na^++*ff*99+SPC or TIP3P, by fluctuations for the Na^++*ff*99+TIP3P, which are more often and with bigger amplitudes than for the SPC/E. It seems that a small salt excess (twice the lower salt conditions) somehow substitute stabilizing properties of *ff*99bsc0. Na^++*ff*99bsc0+SPC, Na^++*ff*99bsc0+TIP3P and KCl+Dang+*ff*99+SPC yield equal values of averaged RMSD. The change of ion parameters

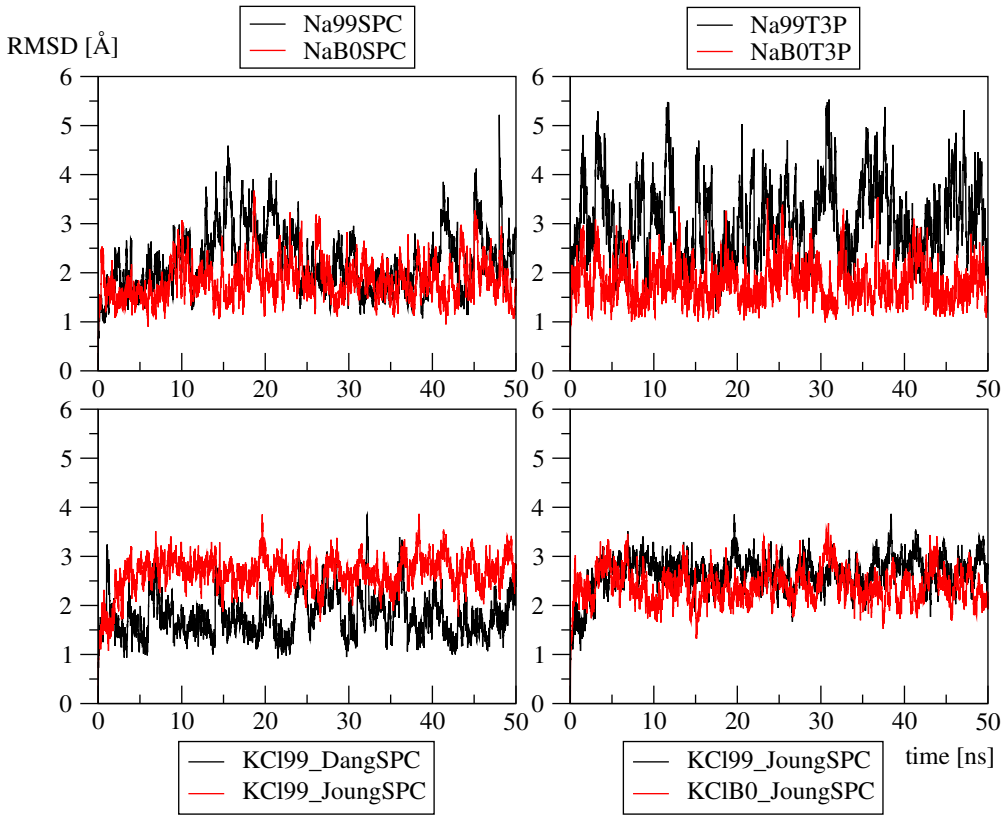

Figure 5.1: Comparison of RMSD time progress for *ff*99, *ff*99bsc0, SPC/E and TIP3P water models and for different ion type. For K^+ also different sets of parameters were used (Dang and Joung).

from Dang to Joung caused rise of RMSD value, but minor reduction in fluctuation amplitude. Interestingly, different combinations of force field, salt conditions, water type and ion parameters result in the same RMSD average values for KCl99DangSPC (salt excess) and NaB0T3P/SPCE (low salt conditions).

Comparing occurrence of the α/γ t/t flips, our results are in agreement with results published by Besseova[1]. The *ff*99bsc0 significantly suppress the α/γ t/t substates as well as their amplitudes. Furthermore, difference for these substates can be observed for TIP3P and SPC/E water types. The TIP3P has faster increase to shifted states and also torsions stay longer in flipped state. Increased salt conditions reduce the amplitude of α/γ flips and together with use of *ff*99bsc0 an additional decrease in their occurrence and magnitude can be seen.

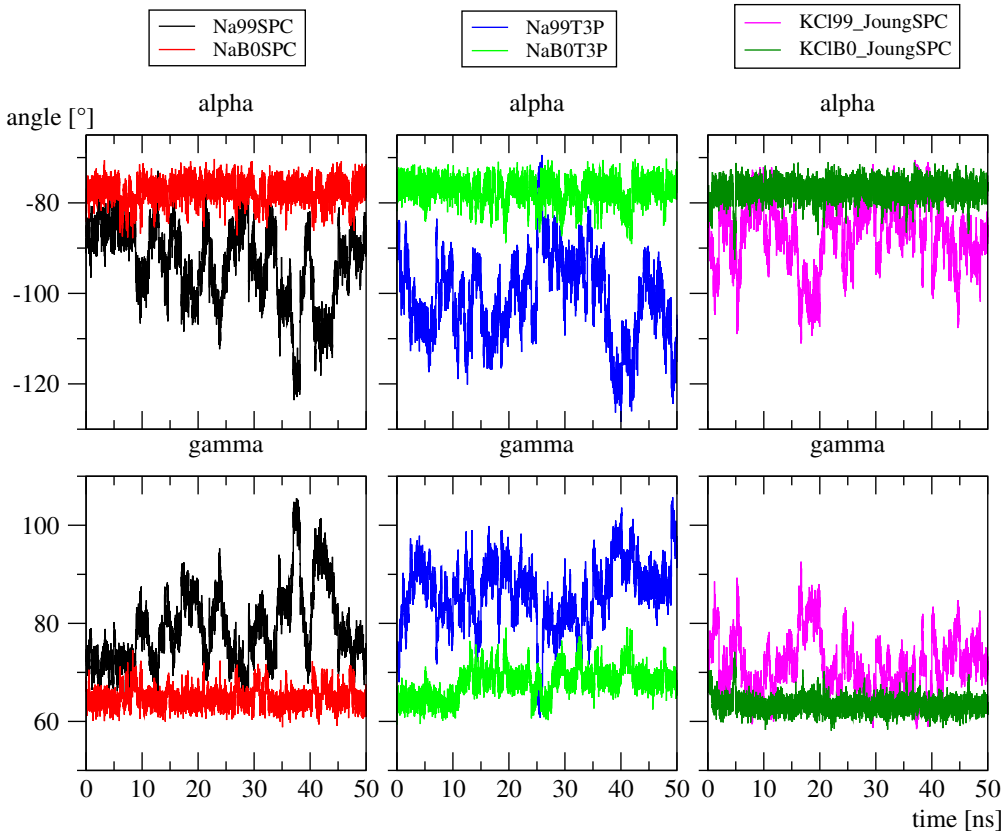

Figure 5.2: Progress in time of α, γ torsion angle for different force fields, water type and low/high salt conditions.

Already, from these RMSD and α/γ torsion results it is possible to distinguish influence of TIP3P on higher kinetics (TIP3P has higher diffusion coefficient than SPC/E). In the TIP3P, deviations has faster start and bigger magnitude in comparison to the similar simulations, but with SPC/E.

A reduction in major groove width and increase in minor groove width was observed

under higher salt conditions and even more reduction when *ff*99bsc0 is used (Figure 5.3). This had opposite effect on minor groove width, which was increasing.

All the simulated structures with *ff*99bsc0 force field, in comparison to *ff*99 force field simulations, result in increase of the base pair roll and slide for each of the ion and water type combinations (Figure 5.4). The same trend was observed for base pair inclination and x-displacement (Figure 5.5). The twist and propeller display fluctuating behavior, where only values for KCl99/B0+JoungSPC are more distinguished from others. Again, we observe additional increase in roll, slide, inclination and x-displacement for high-salt conditions. The narrowing of major groove together with changes toward larger values in base pair roll, helical inclination and twist parameters are described by Besseova to cause a shifting of the simulated A-RNA structures more towards the A-form[1]. This is assumed to stabilize simulation structures. Although, in our simulations we did not observed an increase of the twist parameter, the rest of parameters exhibit reported behavior for *ff*99bsc0 and low/high salt-conditions.

In the end, we agree with Besseova over optimal simulation protocol that could be based on the *ff*99bsc0. In addition, we suggest using SPC/E water type as it obviously in some cases suppress bigger deviations seen for TIP3P. To use Joung parameters for simulations is also more than recommended, as these parameters have special variations optimized for SPC/E and TIP3P. However, it is difficult to propose at this point whether higher salt conditions yielding smaller deviations in simulations or lower salt conditions, which are closer to biological conditions should be used. This issue needs more complex testing together with testing of new force field *ff*99bsc0$_{\chi OL}$[79] that suppresses artificial ladder-like structures[80,95]. Moreover, it is also important to examine how the parameters change in a sequence dependent manner for this A-RNA structure.

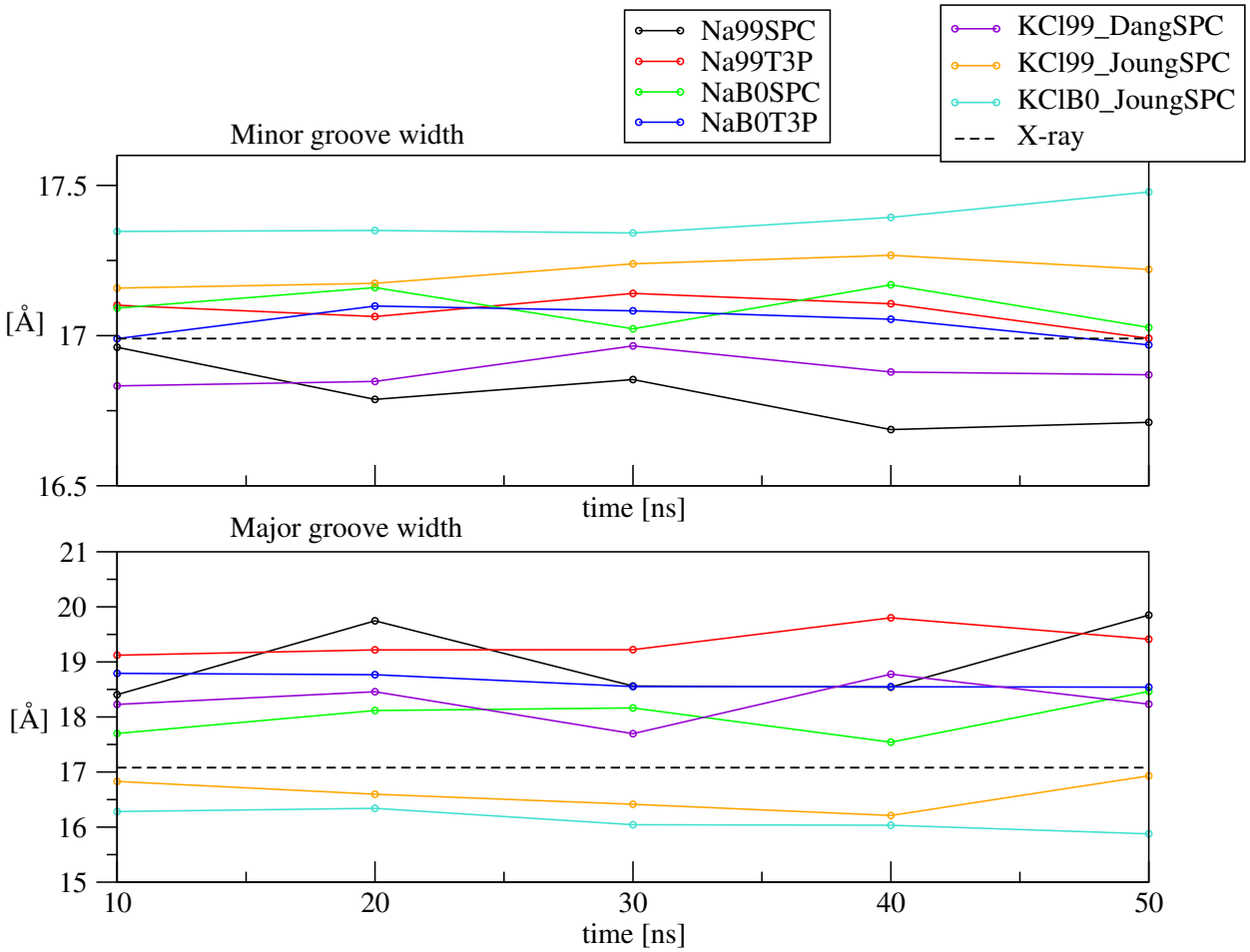

Figure 5.3: Progress in time of averaged values (averaged over base pairs 3 to 16, two terminal base pairs for both ends were omitted, and over 10 ns time intervals) for the major and minor groove width.

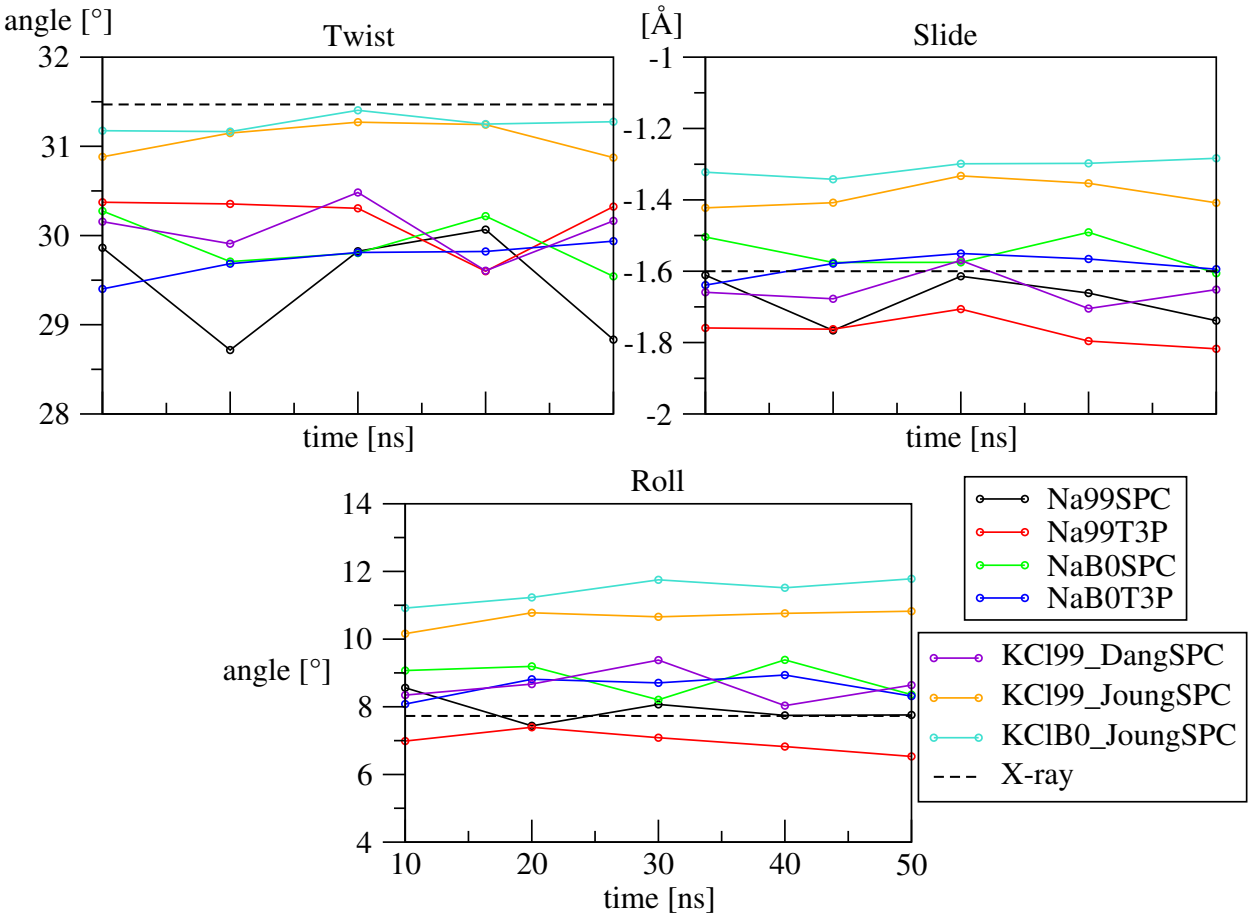

Figure 5.4: Progress in time of averaged parameters (averaged for 10 ns intervals) for twist, slide and roll.

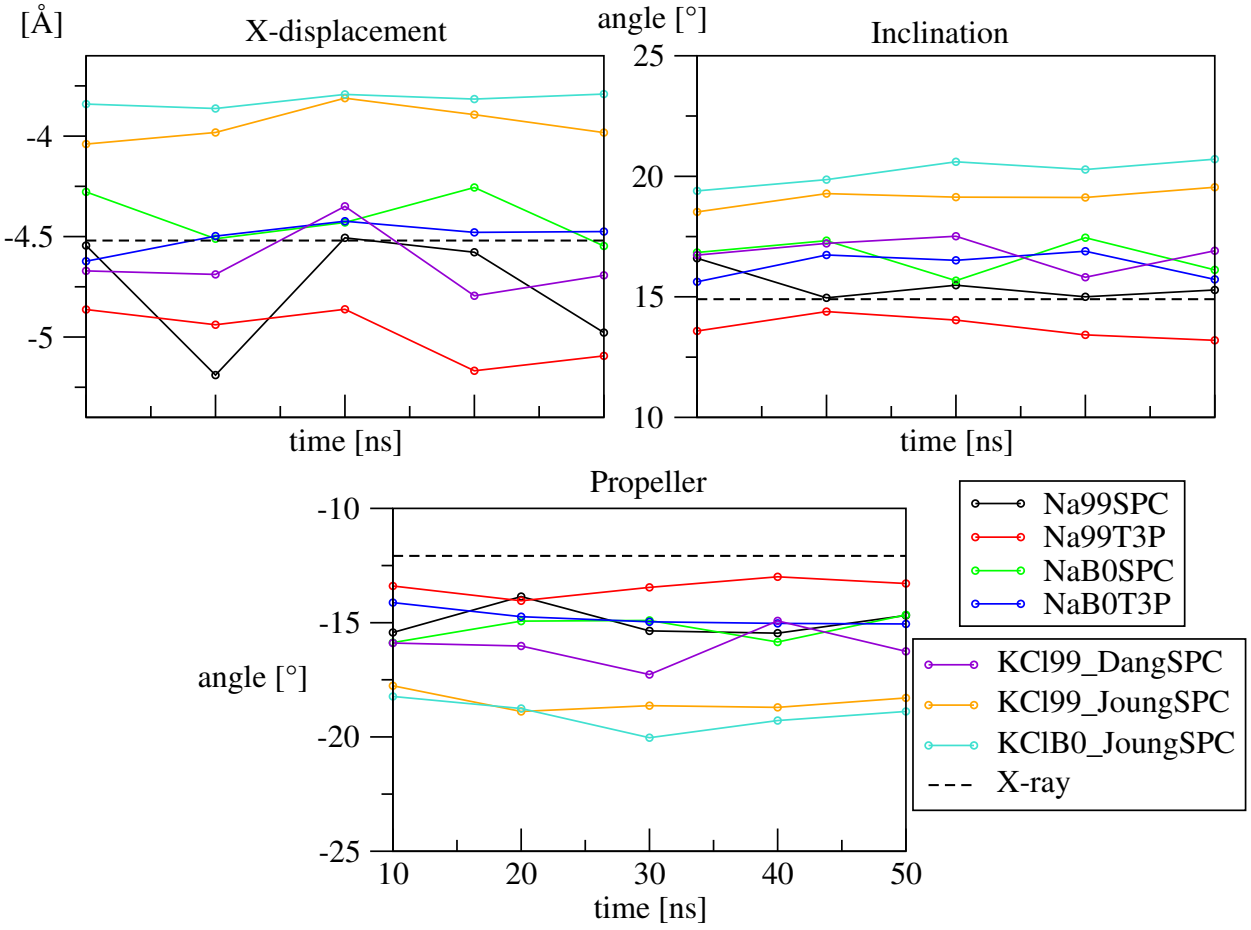

Figure 5.5: Progress in time of averaged parameters (averaged for 10 ns intervals) for x-displacement, inclination and propeller.

5.1.2 Influence of Ionic Strength

Dependence of a DNA on ionic strength has been extensively studied in literature. However, it is very difficult to find any thorough researches concerning an influence of increasing ionic strength on RNA. In this work we carried out six simulations for ionic strength ranging from 498 mM to 4367 mM. Averaged values (calculated and averaged over base pairs 3 to 16, two terminal base pairs were omitted) of RMSD are listed in Table 4.1 at the end of previous chapter. These RMSD averages do not exhibit any trend dependent on increasing ionic strength. Figure 5.6 shows time progress of RMSD values (averaged over base pairs 3 to 16 for each ns time step). In this figure there are significant flips in RMSD values following changes in structure, therefore it is possible to trace down some structural dynamics in selected times. Moreover, amplitude and occurrence of these flips have effect on averaged RMSD.

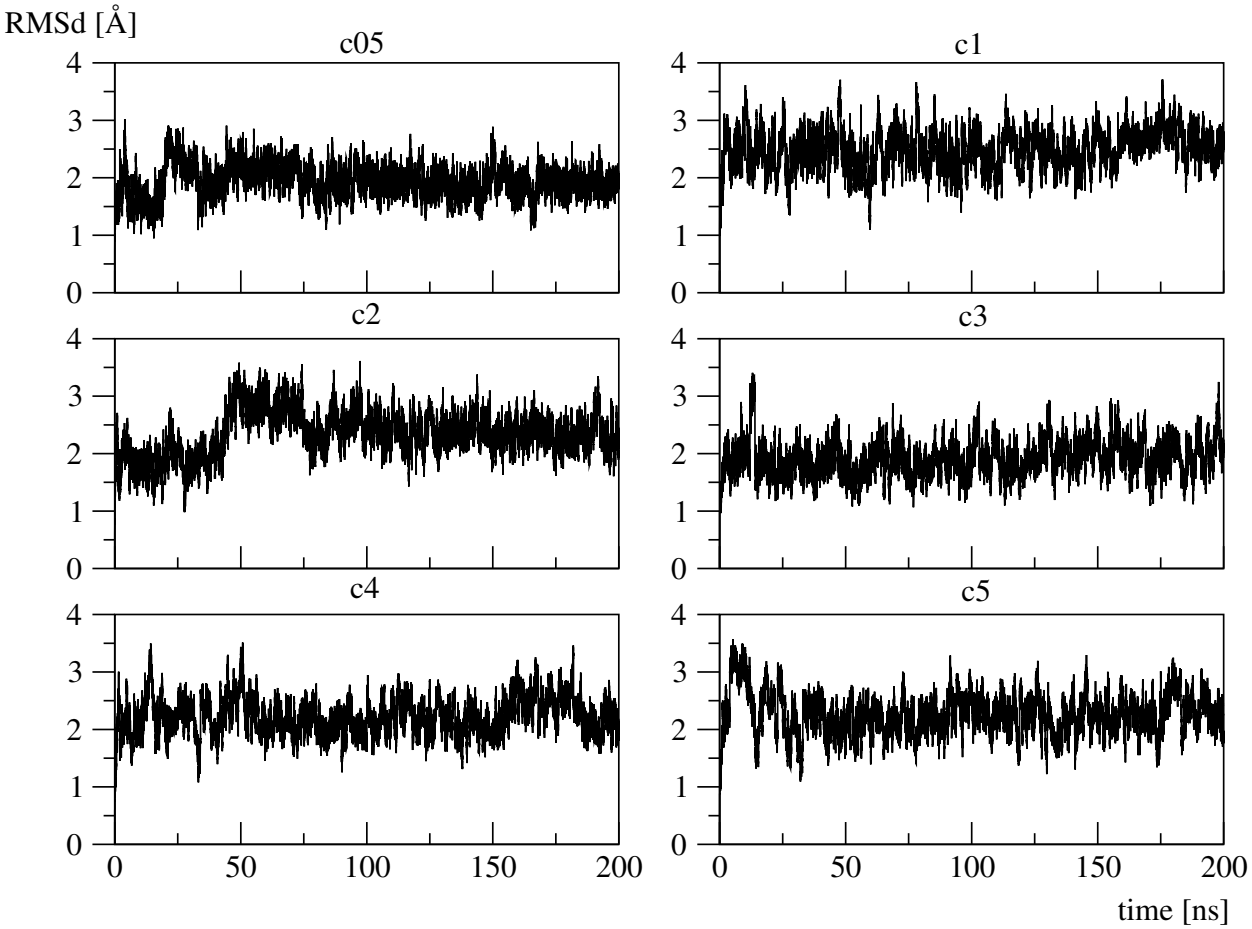

Figure 5.6: RMSD dependence on increasing ionic strength (RMSD is calculated from structures omitting two terminal base pairs on both ends of the helices with respect to the first snapshot).

The minor groove width exhibit fluctuation around ∼17.3 Å with increasing ionic strength

(Figure 5.7). The similar fluctuation pattern will be also seen for torsion angles α, γ, ε and structural parameters inclination, twist and propeller, however, it is important to note that these parameters are interdependent. The major groove width remained stable around value 16 Å from ∼0.5 M to ∼3.5 M ionic strength and suddenly increased to 17.34 Å at ∼4.3 M ionic strength. At the highest ionic strength (4367 mM) the difference between widths of major and minor groove disappeared and it converged to the same value around 17.3 Å.

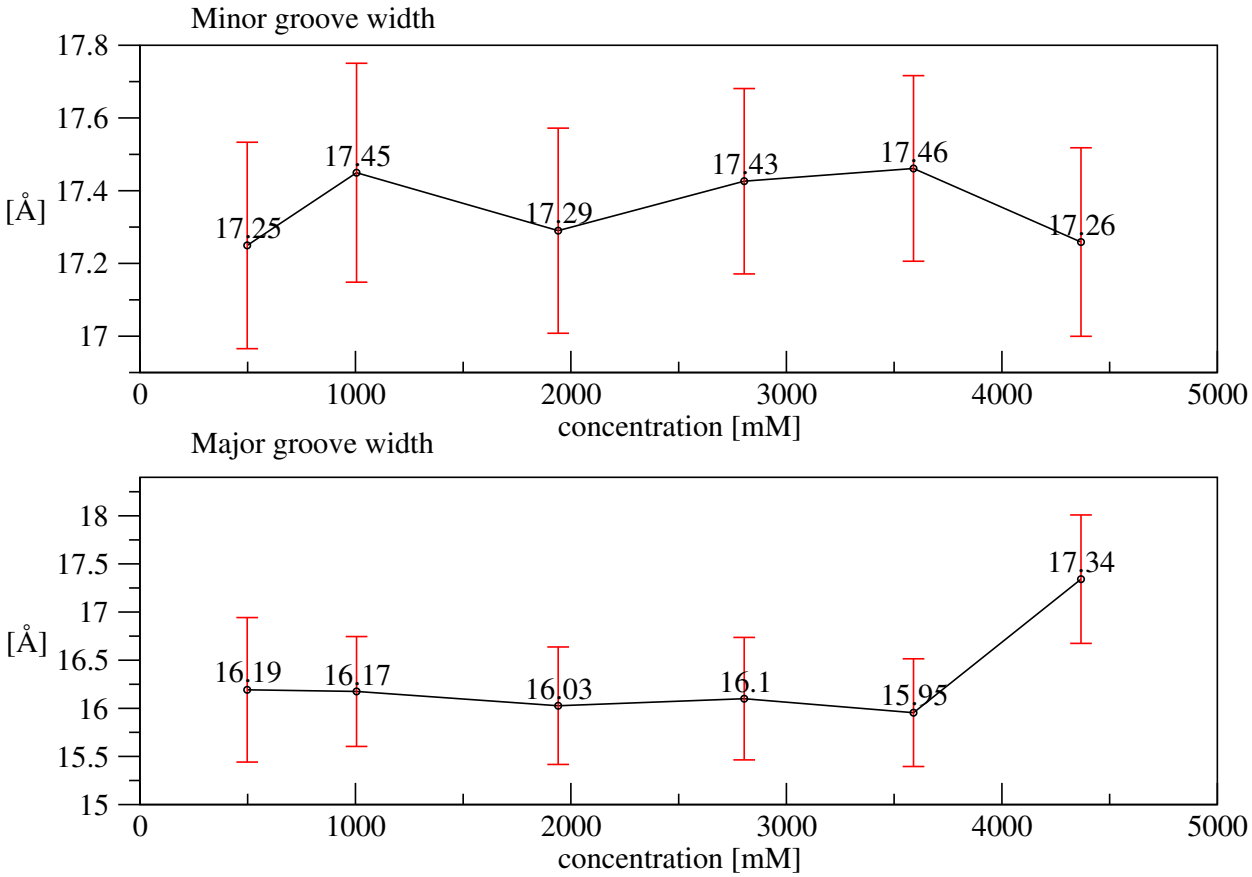

Figure 5.7: Major and minor groove width dependence on increasing ionic strength

Behavior of sugar-phosphate backbone can be monitored using six torion angles. Sugar to base position along glycosidic bond is described by the χ torsion. Shift in the α torsion usually results in the γ torsion shift as these torsions are interdependent (Figure 5.8). At the same figure, the β torsion, except for small fluctuations, does not exhibit dependence on increasing ionic strength and maintain at the same level. The δ torsion slowly decline with rising ionic strength to value 79.1° from original value of 80.9 °.

The ε angle fluctuate around -158.6 ° with changing ionic strength (Figure 5.9). In the case of ζ torsion, though there is small change between angle values at lowest and highest ionic strength, it stays constant for the rest of ionic strengths. The gradually decreasing trend is seen for χ torsion, starting at -153.9° and ending at value of -155.4°. From the last

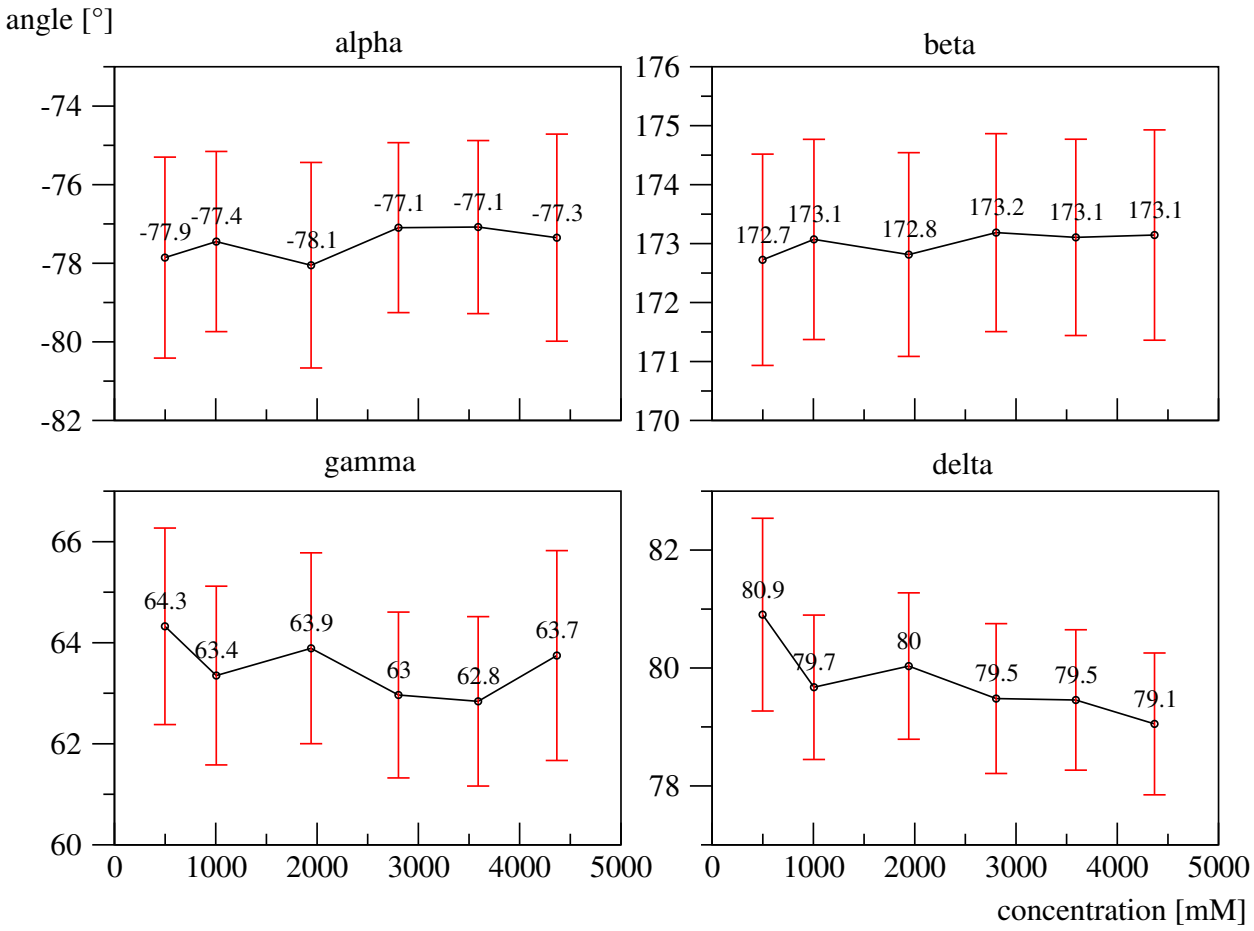

Figure 5.8: Dependence of alpha, beta, gamma and delta torsion angles on increasing ionic strength.

subfigure in Figure 5.9 for pucker phase angle, we can deduce that pucker conformation is not changing with increasing ionic strength and stays in C3'-endo conformation (C3'-endo conformation has distribution peak around 18°).

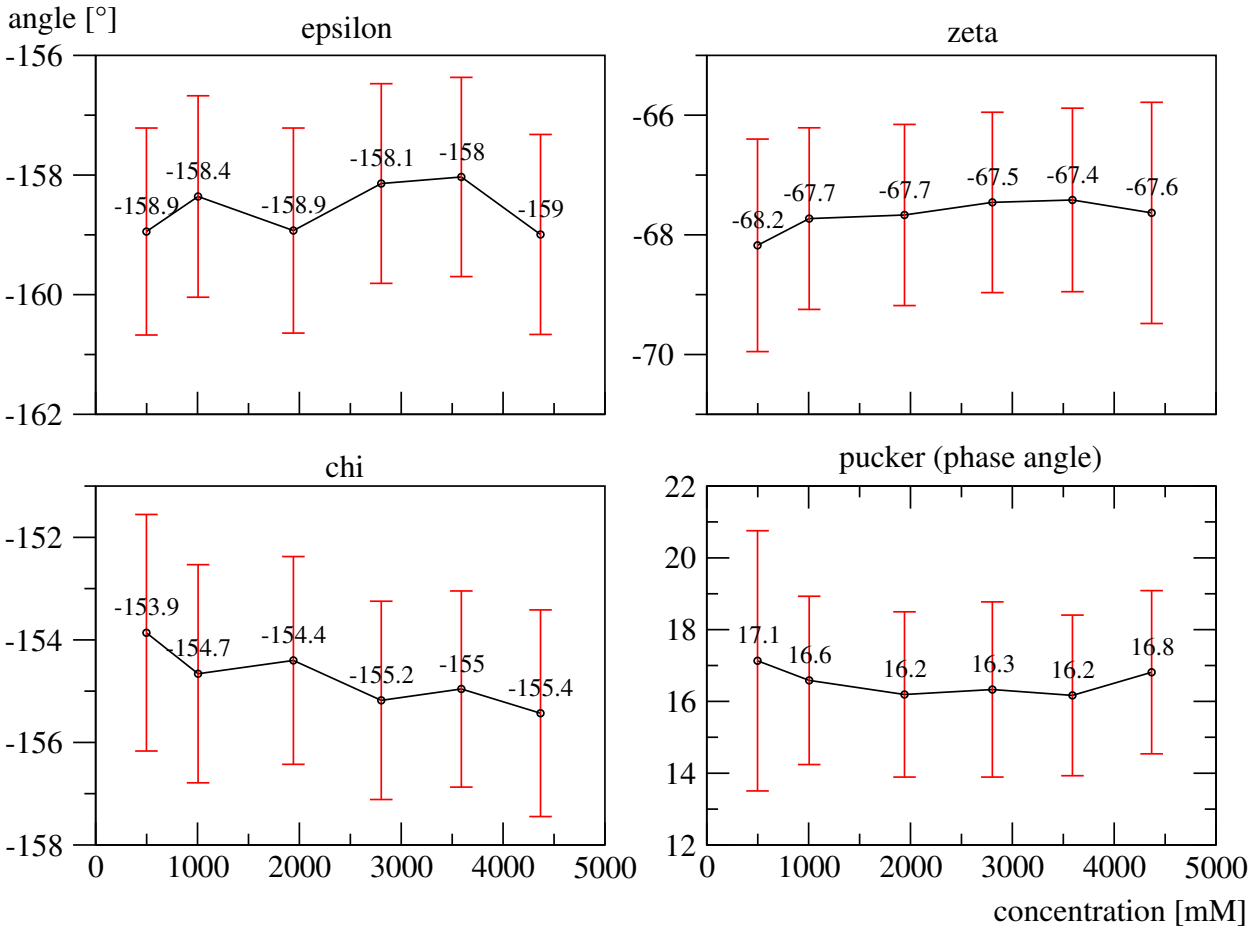

Figure 5.9: Dependence of epsilon, zeta and chi torsion angles and pucker phase angle on increasing ionic strength.

In the following table (Table 5.1) average values and their standard deviations for the torsion angles and pucker phase angle are listed. There are higher values of ζ and χ torsion angles in comparison to X-ray average value, other torsions have lower values than X-ray. Further, α, ε, ζ, χ and phase angle exhibit more significant difference between X-ray average and average from simulated structures. The sugar pucker for X-ray structure was as well as in simulated structures in C3'-endo conformation.

There is a strong dependence between torsion angles and structural parameters. Hence, some changes in torsion angles influence structural parameters and vice versa. Figure 5.10 and Figure 5.11 demonstrate decline of slide and X-displacement toward lower values with increasing ionic strength. Roll and inclination are fluctuating (with similar fluctuation pattern) around $\sim$ 11.4° and $\sim$20.3°, respectively. Twist angle remains stable with increasing

Table 5.1: Average values of torsions and pucker phase angle for X-ray structure. For different ionic strength averaged values and standard deviations of torsion angles and pucker phase angle (averaged over 3 to 16 base pairs over 200ns simulation time).

Simulation	Torsion angles [°]							phase angle [°]
	alpha	beta	gamma	delta	epsilon	zeta	chi	pucker
X-ray	-71.6 ± 34.4	174.6 ± 9.1	66.7 ± 41.8	81.4 ± 2.2	-151.7 ± 6.0	-72.2 ± 4.1	-161.7 ± 8.7	13.7 ± 3.3
c05	-77.9 ± 2.6	172.7 ± 4.7	64.3 ± 1.9	80.9 ± 1.6	-158.9 ± 1.7	-68.2 ± 1.8	-153.9 ± 2.3	17.1 ± 3.6
c1	-77.4 ± 2.3	173.1 ± 1.7	63.4 ± 1.8	79.7 ± 1.2	-158.4 ± 1.7	-67.7 ± 1.5	-154.7 ± 2.1	16.6 ± 2.3
c2	-78.1 ± 2.6	172.8 ± 1.7	63.9 ± 1.9	80.0 ± 1.2	-158.9 ± 1.7	-67.7 ± 1.5	-154.4 ± 2.0	16.2 ± 2.3
c3	-77.1 ± 2.2	173.2 ± 1.7	63.0 ± 1.6	79.5 ± 1.3	-158.1 ± 1.7	-67.5 ± 1.5	-155.2 ± 1.9	16.3 ± 2.4
c4	-77.1 ± 2.2	173.1 ± 1.7	62.8 ± 1.7	79.5 ± 1.2	-158.0 ± 1.7	-67.4 ± 1.5	-155.0 ± 1.9	16.2 ± 2.2
c5	-77.3 ± 2.6	173.1 ± 4.7	63.7 ± 2.1	79.1 ± 1.2	-159.0 ± 1.7	-67.6 ± 1.8	-155.4 ± 2.0	16.8 ± 2.3

ionic strength, except for moderate drop at ∼4.3 M ionic strength. Difference in propeller angle for lowest and highest ionic strength is 0.6° and its values fluctuated with changes in ionic strength.

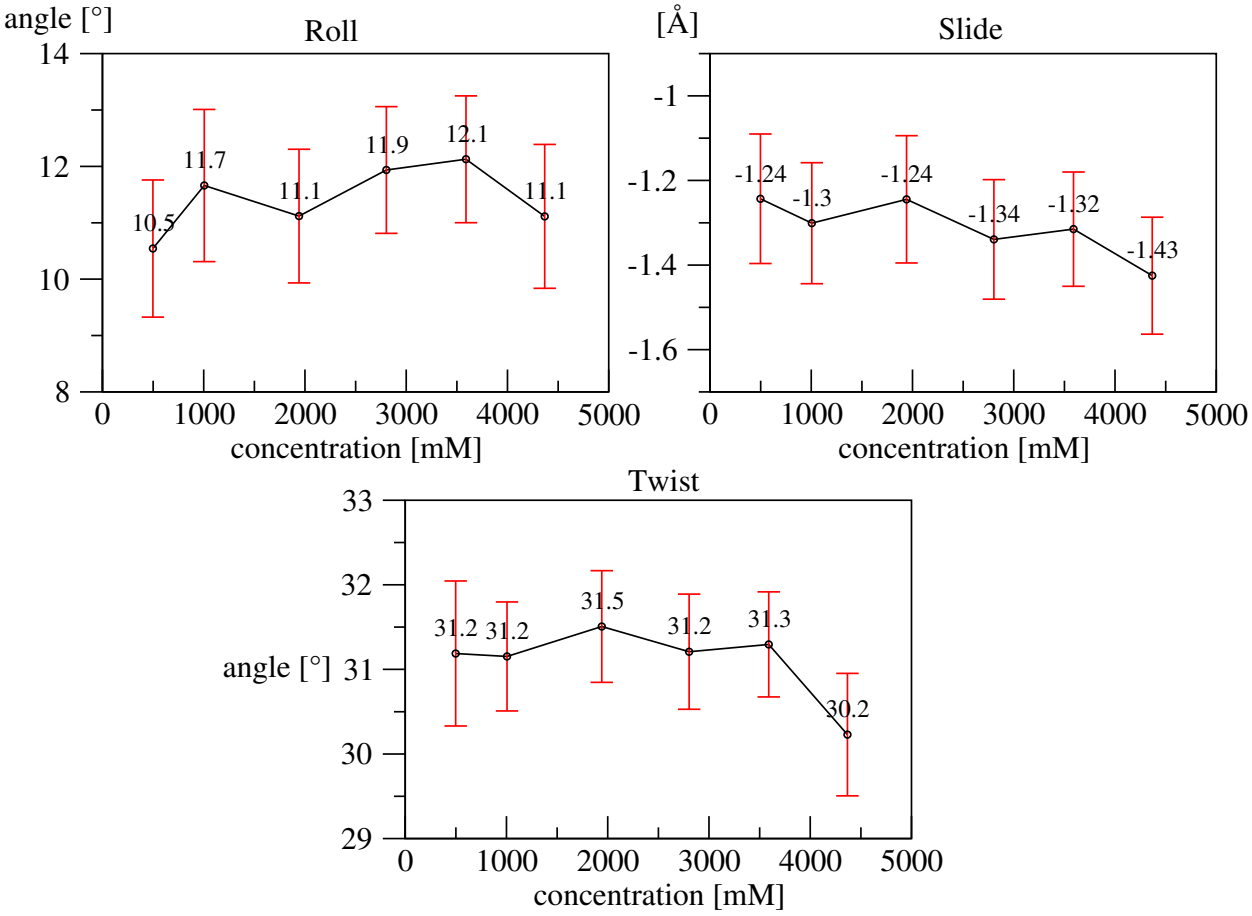

Figure 5.10: Dependence of roll, slide and twist parameters on changing ionic strength.

The value of helical twist is slightly increasing towards 35° at ∼3.5 M and dropped to 33.9° at ∼4.3 M (Figure 5.12). With increasing ionic strength helical rise exhibit fluctuations around ∼2.5 Å.

Table 5.2 and Table 5.3 list parameters used for studied A-RNA structure description. All the listed values for simulations are averaged over 3 to 16 base pairs (omitting two terminal base pairs at the both ends) and over 200 ns simulation time. The twist angle values were close to X-ray average value apart from 1° drop in the twist angle that occurred at the highest ionic strength. For the roll, inclination and propeller simulation structures show higher values in comparison to X-ray. In the X-ray structure minor and major groove widths values were close, the same was observed in the simulated structures for the highest ionic strength.

Structural changes (increase of the major groove width, slight decrease of the helical twist) together with visual examination of trajectories in VMD suggest that with increasing

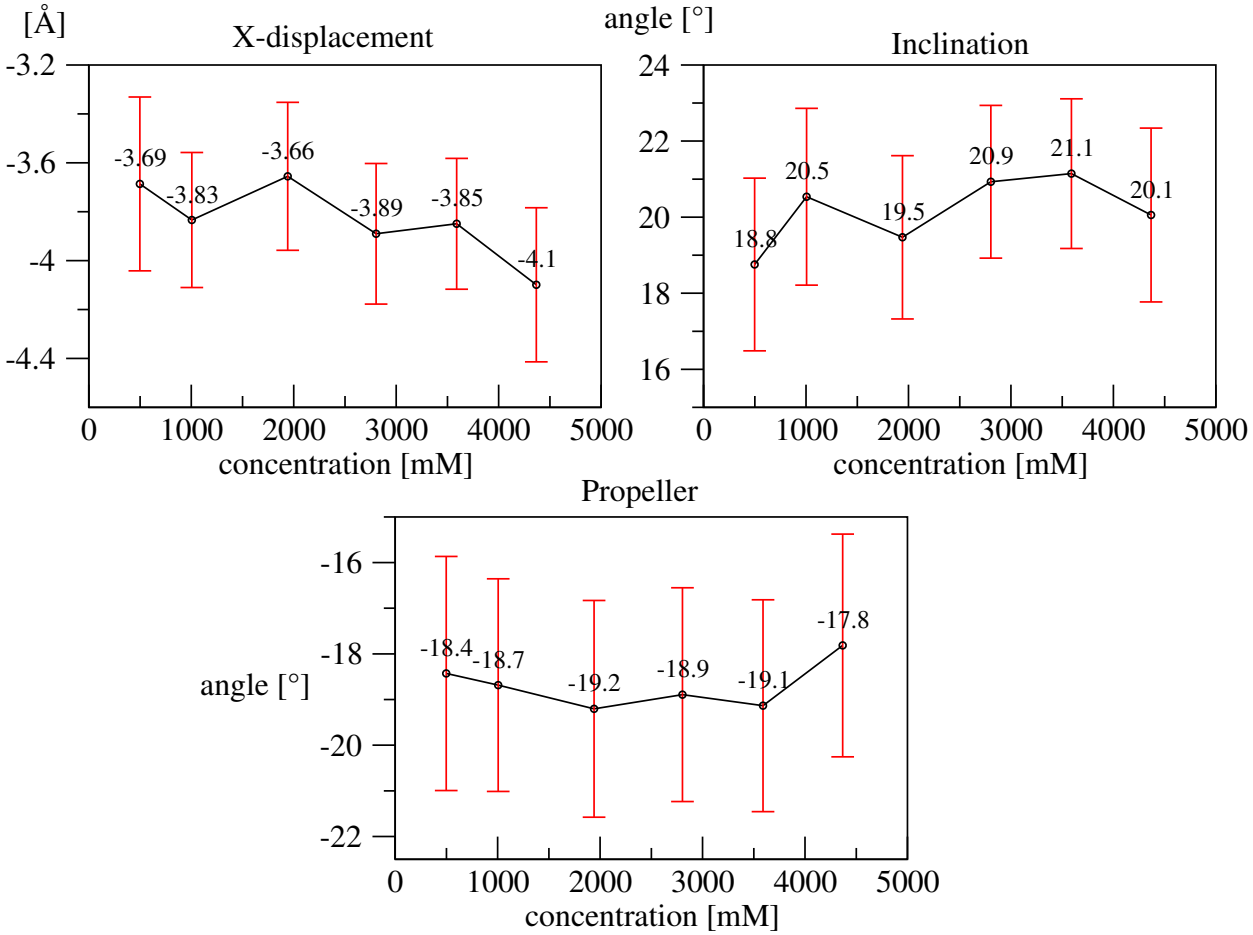

Figure 5.11: Dependence of x-displacement, inclination and propeller on changing ionic strength.

Table 5.2: Average values of structural parameters twist, slide, roll, x-displacement and inclination.

Simulation	twist [°]	slide [Å]	roll [°]	xdisp [Å]	incl [°]
X-ray	31.4 ± 6.4	-1.6 ± 0.6	7.7 ± 2.1	-4.5 ± 2.7	14.9 ± 6.9
c05	31.2 ± 0.9	-1.2 ± 0.2	10.5 ± 1.2	-3.7 ± 0.4	18.8 ± 2.3
c1	31.2 ± 0.6	-1.3 ± 0.1	11.7 ± 1.3	-3.8 ± 0.3	20.5 ± 2.3
c2	31.5 ± 0.7	-1.2 ± 0.2	11.1 ± 1.2	-3.7 ± 0.3	19.5 ± 2.1
c3	31.2 ± 0.7	-1.3 ± 0.1	11.9 ± 1.1	-3.9 ± 0.3	20.9 ± 2.0
c4	31.3 ± 0.6	-1.3 ± 0.1	12.1 ± 1.1	-3.8 ± 0.3	21.1 ± 2.0
c5	30.2 ± 0.7	-1.4 ± 0.1	11.1 ± 1.3	-4.1 ± 0.3	20.1 ± 2.3

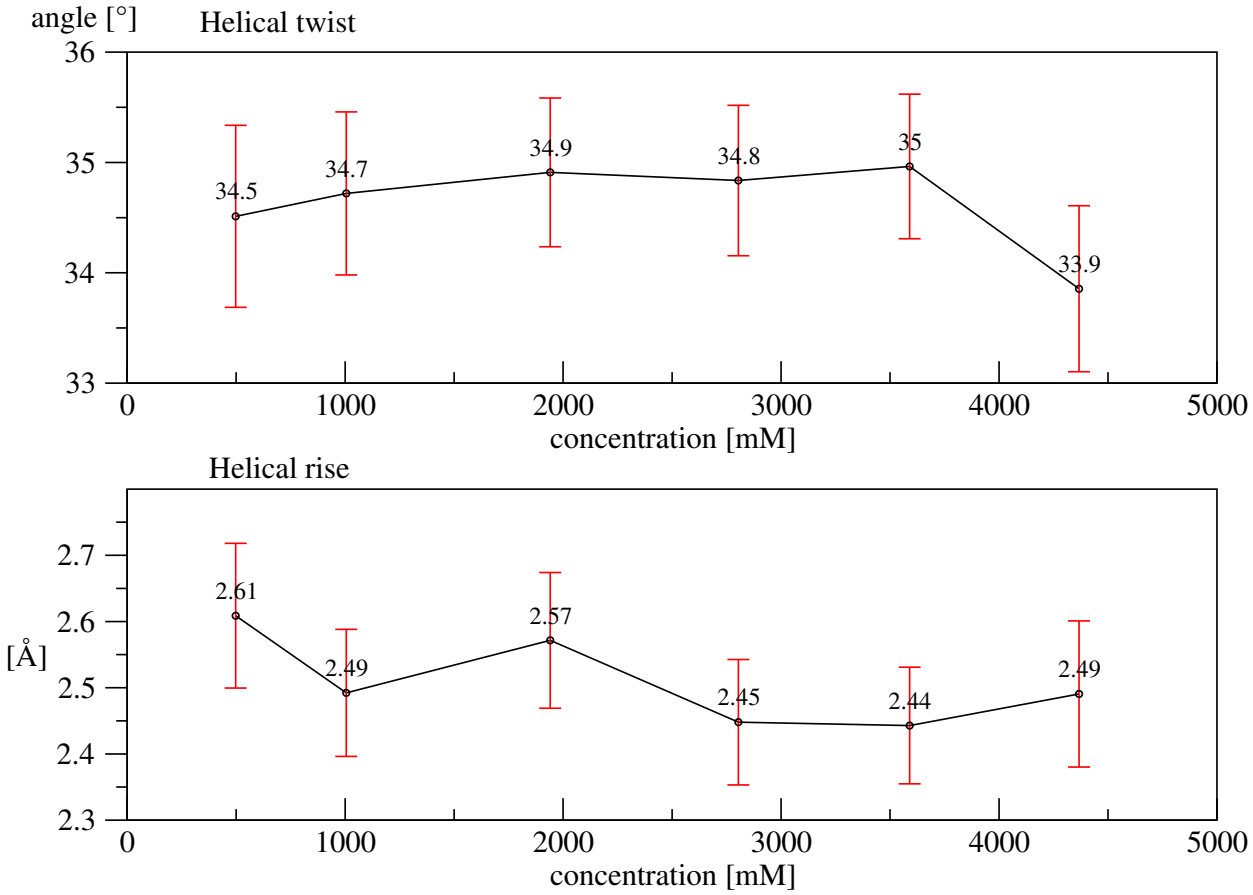

Figure 5.12: Dependence of helical parameters; helical twist and helical rise on changing ionic strength.

ionic strength A-RNA structure is more opened. However, observed structural changes are more complex and need more thorough examination.

Table 5.3: Average values of structural parameters htwist, hrise, propeller, minor and major groove width.

Simulation	htwist [°]	hrise [Å]	propeller [°]	minor [Å]	major [Å]
X-ray	32.7 ± 5.8	2.6 ± 0.6	-12.1 ± 3.5	17.0 ± 0.3	17.1 ± 1.8
c05	34.5 ± 0.8	2.6 ± 0.1	-18.4 ± 2.6	17.2 ± 0.3	16.2 ± 0.8
c1	34.7 ± 0.7	2.5 ± 0.1	-18.7 ± 2.3	17.4 ± 0.3	16.2 ± 0.6
c2	34.9 ± 0.7	2.6 ± 0.1	-19.2 ± 2.4	17.3 ± 0.3	16.0 ± 0.6
c3	34.8 ± 0.7	2.4 ± 0.1	-18.9 ± 2.3	17.4 ± 0.3	16.1 ± 0.6
c4	35.0 ± 0.7	2.4 ± 0.1	-19.1 ± 2.3	17.5 ± 0.3	16.0 ± 0.6
c5	33.9 ± 0.8	2.5 ± 0.1	-17.8 ± 2.4	17.3 ± 0.3	17.3 ± 0.7

One of the explanations for wider major groove in the highest ionic strength (∼4.3 M) might be higher occupation of K^+ in major groove (Figure 5.13). This may also have an influence on other structural parameters and torsions. Moreover, higher occupation of K^+ was observed close to AH(+)/C and G/U wobble base pairs and it might have effects on a structure. There is high probability that these base pairs bind K^+ ion in selected place, however, it is quite surprising as one would expect lower K^+ occupation in proximity of positively charged adenine. Although, all these assumptions seems reasonable, additional more detailed examinations of ion interactions with base pairs and sugar-phosphate backbone are necessary.

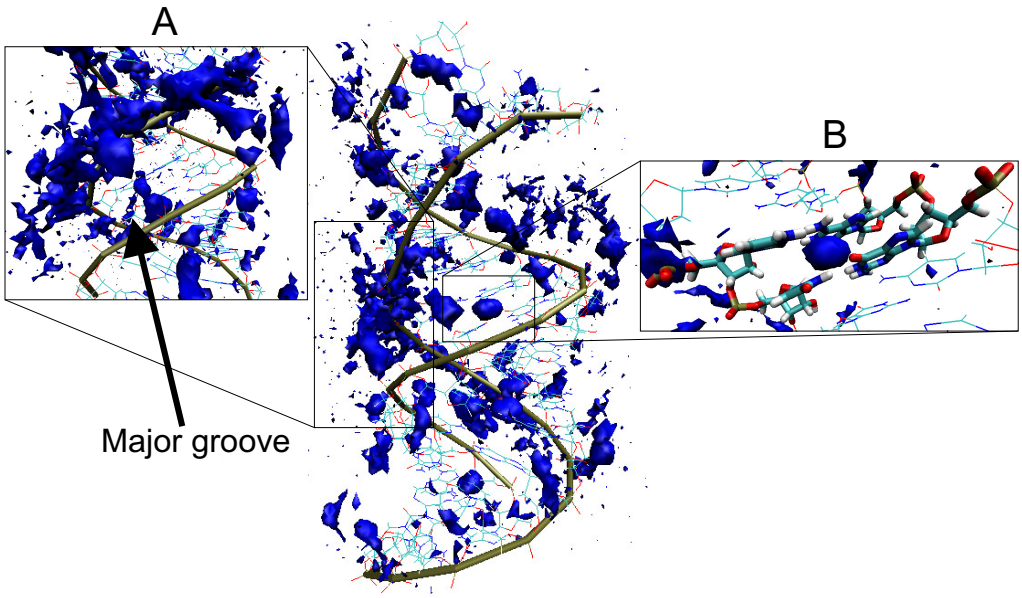

Figure 5.13: Density map of K^+ ions for highest ionic strength. (A) K^+ ions occurrence along sugar-phosphate backbone as well as K^+ presence in major groove. (B) depicts possible K^+ ion trapped inside the structure, where top is AH(+)/C and bottom G/U wobble base pair.

5.1.3 Results from Thermodynamic Integration

The consequence of Guanine to Inosine mutation in RNA is lost of one hydrogen bond. This should results in a difference of free energy equivalent to net hydrogen bonding value -0.25 kcal.mol^{-1}. However, much greater drop of free energy difference ($\Delta\Delta G$) of 3.11 kcal.mol^{-1} (for 14mM ionic strength) and 3.44 kcal.mol^{-1} (for 114 mM ionic strength) per base pair was observed[72,73]. Hence, some other contributions must play a dominant role in this decrease. Molecular dynamics together with TI is an ideal method beside the experimental techniques as it is able to not only provide necessary statistics for calculation of free energies but also decompose obtained free energies to its particular contributions. In addition, energy changes can be related to changes in structural parameters. Mutation of G6 to I6 was carried out for A-RNA structure and obtained results were compared to experimental values. In the first step we demonstrate capability of TI to describe energetics of G to I transformation. Table 5.4 shows that TI is able to yields reasonable results, semiquantitatively comparable with experimental values.

Table 5.4: Table of experimental and TI results of ΔΔG for RNA structure

	Ionic strength [mM]	ΔΔG [kcal/mol][a]
Experiments	14	3.44 ± 0.06
	114	3.11 ± 0.09
TI	113	2.14 ± 0.52
	206	2.51 ± 0.50
	497	2.81 ± 0.49
	1021	2.44 ± 0.46

[a]calorie (cal) is commonly used unit in thermodynamics. It can be converted to standard Joule (J) as follows: 1 cal ≈ 4.18 J

Experimental values demonstrate decline in $\Delta\Delta G$ for higher ionic strength. However, this was not observed for TI values Figure 5.14.

Instead, $\Delta\Delta G$ followed different trend, yielding lower values for the lowest and highest ionic strength and with peak at ∼500 mM ionic strength. When we look carefully at Figure 5.15, the contribution coming from double stranded A-RNA is independent of ionic strength in comparison to more dynamic single stranded A-RNA that exhibits changes in ΔG for different ionic strength.

Hence, above mentioned different trend for TI is caused by fluctuations in ΔG of single stranded A-RNA. After thorough examination of what may be the cause we propose that these ΔG differences are affected by short timescales of simulations and therefore insufficient sampling for some torsion angles, especially χ torsion. The transition of χ torsion

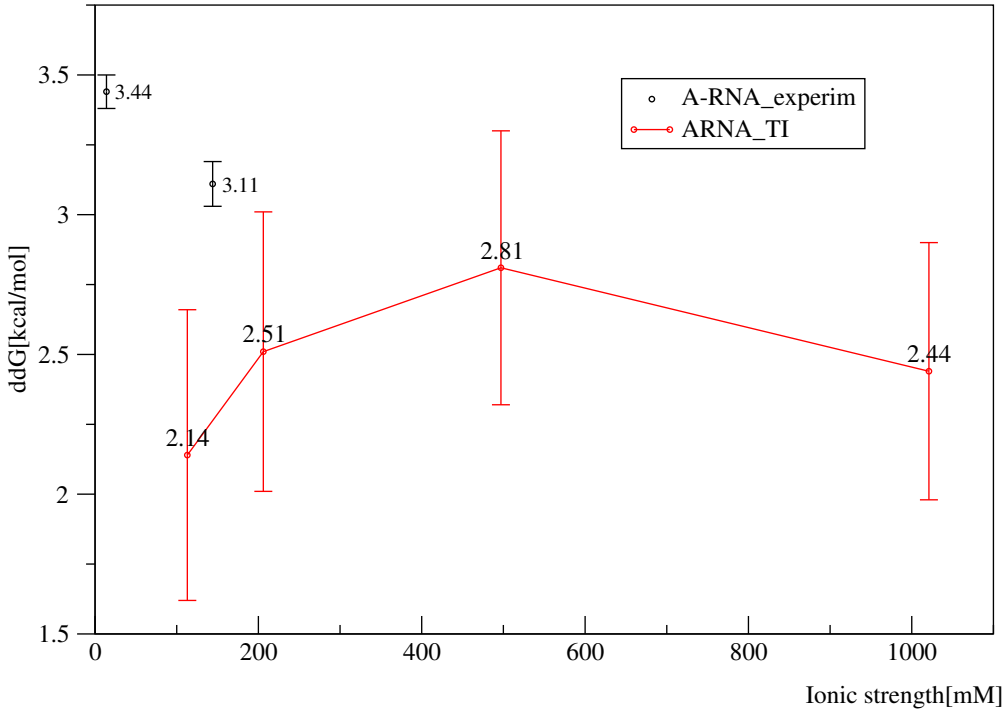

Figure 5.14: Comparison of $\Delta\Delta G$ values from TI with experimental values.

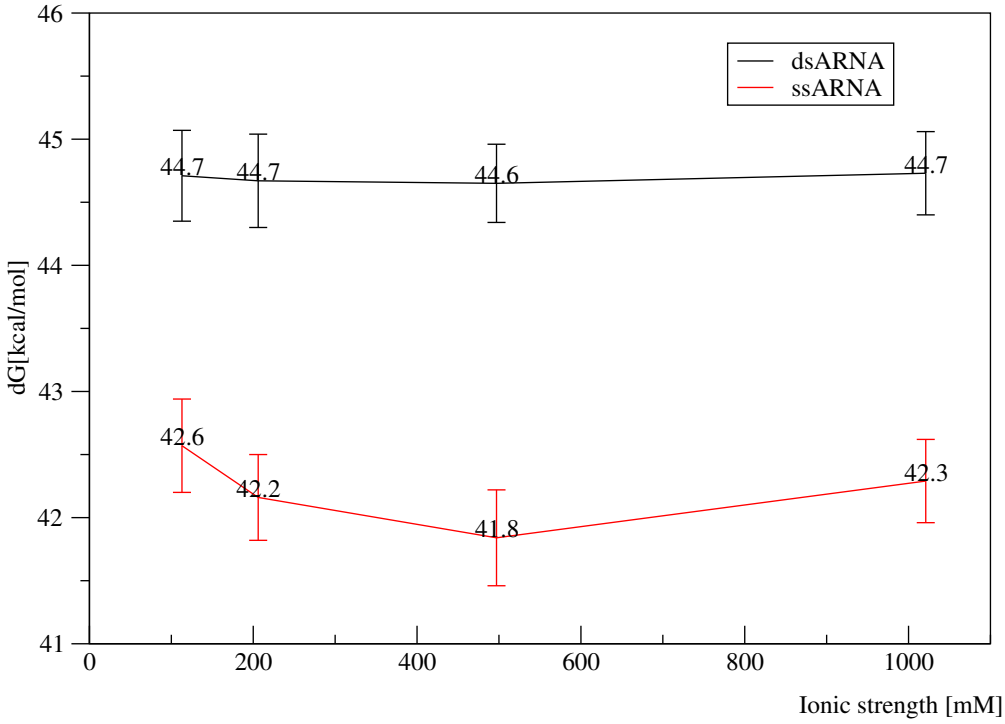

Figure 5.15: ΔG dependence on ionic strength for double and single stranded RNA.

toward high-*anti* value were recently identified to produce the artificial formation of ladder-like structure of A-RNA duplexes[80,95]. Two sub-states of χ torsion low-*anti* and high-*anti*

corresponding to different energy were identified. The relative population of these two sub-states, which is artificial under lower sampling conditions, influence the final ΔG, which is average of energies for these two sub-states. For illustration Figure 5.16 depicts low-*anti* and high-*anti* populations of χ torsion of mutated nucleobase, for different ionic strengths and for one value of mixing parameter $\lambda = 0.5$ of G→G_0.

It seems that for 1000 mM ionic strength low-*anti* is highly populated and in case of the lowest ionic strength high-*anti* state is more frequently populated. However, it is premature to speculate, until additional evidence is given, though we can just guess that longer sampling of χ torsion will yield more accurate information about ΔG and thus result in TI values in agreement with experiments. Nevertheless, it should be noted that different population of high-*anti* and low-*anti* conformations of glycosidic χ torsion might be caused by different ionic strength, but we assume this possibility as less likely.

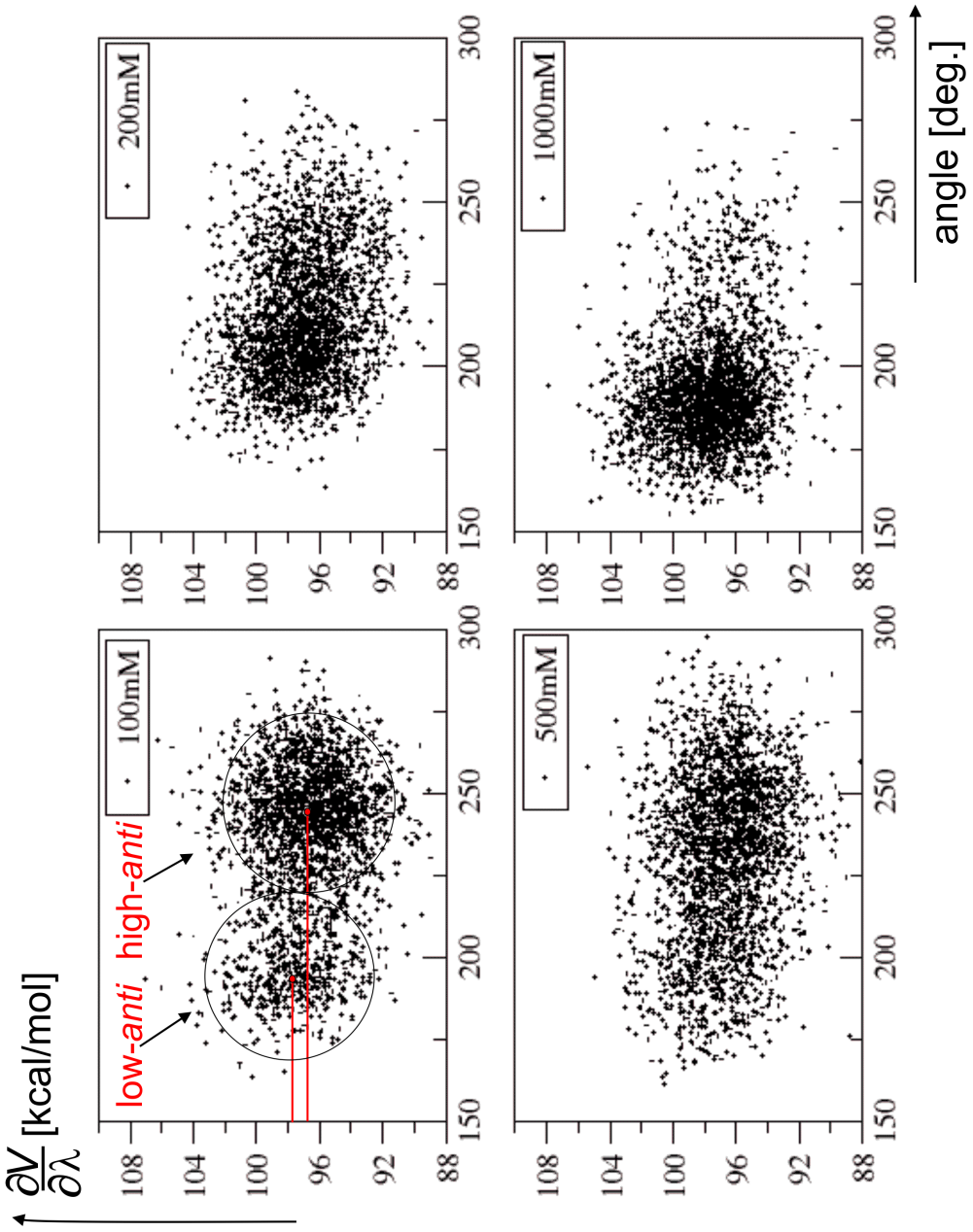

Figure 5.16: Low-*anti* and high-*anti* sub-states of χ torsion angle.

Chapter 6

Conclusions

RNA is very important molecule in most of the organisms. Therefore, studying its structure and energetics under different conditions may reveal invaluable information and elucidate mechanisms of many RNA functions. In this work we used computational method of Molecular Dynamics and Thermodynamic integration that are able to provide insight into structure and structural changes at atomic level. The work itself is divided into three separate parts. The first part tested use of the different force fields (*ff*99 and *ff*bsc0)), water types (SPC/E and TIP3P). Furher, influence of small salt excess (low salt conditions Na^+ ion, high salt excess K^+), together with different ions (Na^+ ion, K^+) and same ions with different parameters (Dang, Joung). Results from the first part suggest using *ff*bsc0 force field that reduce occurrence of pathological α/γ flips, which are irreversible for DNA, but reversible for RNA. Further, this force field should be used together with SPC/E water model. The SPC/E water model has different physico-chemical properties than TIP3P. For example, in case of the TIP3P water type higher diffusion coeficient lead to bigger deviations for RMSD values as well as TIP3P influence the kinetics of α/γ flips. Moreover, Joung parameters, which have been specifically adjusted for each SPC/E and TIP3P water type, were suggested as they exhibited more stabilizing effect in comparison to Dang parameters. Nevertheless, more complex studies for bigger variety of structure, because of sequence dependent parameters, is important in order to improve simulation protocol.

The second part is observing the influence of increasing ionic strength on torsion angles and structural parameters. The main results from this part are that major and minor groove widths ended up at the same values. The δ and χ torsions together with slide, twist and x-displacement were declining with rising ionic strength. And the whole structure seems to be more opened. This may be caused by higher K^+ ion occupancy within major groove, stretching its width. Another contribution might be specific ion binding of AH(+)/C and G/U wobble adjacent base pairs. However, these assumptions need to be further tested and additional analysis of obtained simulations is required.

In this work we proved that Thermodynamic Integration yields reasonable results when

compared with experimental values. It seems that Molecular Dynamics is not able to properly describe $\Delta\Delta G$ dependence on ionic strength. However, some other testing is required. Further, a problem with insufficient sampling of some torsion angles occurred for ssRNA during analysis. The most problematic seems to be the χ torsion with its two main energy levels corresponding to low-*anti* and high-*anti*. Hence, longer sampling times are necessary in order to acquire more accurate results.

References

[1] I. Besseova, M. Otyepka, K. Reblova, and J. Sponer. Dependence of A-RNA simulations on the choice of the force field and salt strength. *Phys Chem Chem Phys*, 11(45):10701–11, 2009. 1, 4.3, 5.1.1, 5.1.1, 5.1.1

[2] N. A. Siegfried and P. C. Bevilacqua. Thinking inside the box: designing, implementing, and interpreting thermodynamic cycles to dissect cooperativity in RNA and DNA folding. *Methods Enzymol*, 455:365–93, 2009. 1

[3] M. A. Ditzler, M. Otyepka, J. Sponer, and N. G. Walter. Molecular dynamics and quantum mechanics of RNA: conformational and chemical change we can believe in. *Acc Chem Res*, 43(1):40–7, 2010. 1, 3.1, 3.1.1

[4] T. Caspersson and J. Schultz. Ribonucleic Acids in Both Nucleus and Cytoplasm, and the Function of the Nucleolus. *Proc Natl Acad Sci U S A*, 26(8):507–15, 1940. 2.1

[5] O. T. Avery, C. M. Macleod, and M. McCarty. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J. Exp. Med.*, 79(2):137–58, 1944. 2.1

[6] A. D. Hershey and M. Chase. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.*, 36(1):39–56, 1952. 2.1

[7] A. Gierer and G. Schramm. Infectivity of ribonucleic acid from tobacco mosaic virus. *Nature*, 177(4511):702–3, 1956. 2.1

[8] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–8, 1953. 2.1

[9] J. D. Watson and F. H. Crick. Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171(4361):964–7, 1953. 2.1

[10] A. Rich and J. D. Watson. Some Relations between DNA and RNA. *Proc Natl Acad Sci U S A*, 40(8):759–64, 1954. 2.1

[11] F. H. Crick. On protein synthesis. *Symp Soc Exp Biol*, 12:138–63, 1958. 2.1

[12] Elliot Volkin and L. Astrachan. Phosphorus incorporation in Escherichia coli ribonucleic acid after infection with bacteriophage T2. *Virology*, 2(2):149–161, April 1956. 2.1

[13] Elliot Volkin and L. Astrachan. Intracellular distribution of labeled ribonucleic acid after phage infection of Escherichia coli. *Virology*, 2(4):433–437, August 1956. 2.1

[14] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, 3:318–56, 1961. 2.1

[15] F. Gros, W. Gilbert, H. H. Hiatt, G. Attardi, P. F. Spahr, and J. D. Watson. Molecular and biological characterization of messenger RNA. *Cold Spring Harb Symp Quant Biol*, 26:111–32, 1961. 2.1

[16] G. E. Palade. A small particulate component of the cytoplasm. *J Biophys Biochem Cytol*, 1(1):59–68, 1955. 2.1

[17] S. Brenner. RNA, ribosomes, and protein synthesis. *Cold Spring Harb Symp Quant Biol*, 26:101–10, 1961. 2.1

[18] A Rich and U. L. RajBhandary. Transfer RNA: Molecular Structure, Sequence, and Properties. *Annu. Rev. Biochem.*, 45(1):805–860, June 1976. 2.1

[19] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, 31(1):147–57, 1982. 2.1

[20] C. Guerrier Takada, K. Gardiner, T. Marsh, N. Pace, and S. Altman. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35(3 Pt 2):849–57, 1983. 2.1

[21] A. D. Ellington and J. W. Szostak. In vitro selection of RNA molecules that bind specific ligands. *Nature*, 346(6287):818–22, 1990. 2.1

[22] J. R. Ecker and R. W. Davis. Inhibition of gene expression in plant cells by expression of antisense RNA. *Proc Natl Acad Sci U S A*, 83(15):5372–6, 1986. 2.1

[23] Andrew Fire, SiQun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver, and Craig C. Mello. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, 391(6669):806–811, February 1998. 2.1

[24] A. Rich. On the problems of evolution and biochemical information transfer. In *Horizons in biochemistry*, pages 103–126. Academic Press, New York., 1962. 2.2

[25] C. Woese. In *The genetic code*, pages 179–195. Harper & Row, New York., 1967. 2.2

[26] L.E. Orgel. Evolution of the genetic apparatus. *J. Mol. Biol.*, 38:381–393, 1968. 2.2

[27] F.H.C. Crick. The origin of the genetic code. *J. Mol. Biol.*, 38:367–379, 1968. 2.2

[28] W. Gilbert. The RNA world. *Nature*, 319:818, 1986. 2.2

[29] S. Pizzarelo. Chemical evolution and meteorites: An update. *Origins Life Evol. Biosph.*, 34:25–34, 2004. 2.2

[30] S. L. Miller. Production of Amino Acids Under Possible Primitive Earth Conditions. *Science*, 117:528, 1953. 2.2

[31] S. L. Miller and H. C. Urey. Organic Compound Synthesis on the Primitive Earth. *Science*, 130:245, 1959. 2.2

[32] D. Muller, S. Pitsch, A. Kittaka, E. Wagner, C.E. Wintner, and A. Eschenmoser. Chemistry of alpha-aminonitriles. Aldomerisation of glycolaldehyde phosphate to rac-hexose 2,4,6-triphosphates and (in presence of formaldehyde) rac-pentose 2,4-diphosphates: rac-allose 2,4,6-triphosphate and rac-ribose 2,4-diphosphate are the main reaction-products. *Helv. Chim. Acta*, 73:1410–1468, 1990. 2.2

[33] A. Ricardo, M. A. Carrigan, A. N. Olcott, and S. A. Benner. Borate minerals stabilize ribose. *Science*, 303:196, 2004. 2.2

[34] J. Oro. Mechanism of synthesis of adenine from hydrogen cyanide under possible primitive earth conditions. *Nature*, 191:1193–1194, 1961. 2.2

[35] R. A. Sanchez, J. P. Ferris, and L. E. Orgel. Studies in prebiotic synthesis. II. Synthesis of purine precursors and amino acids from cyanoacetylene and cyanate. *J. Mol. Biol.*, 30:223–253, 1967. 2.2

[36] J. P. Ferris, R. A. Sanchez, and L. E. Orgel. Studies in prebiotic synthesis. III. Synthesis of pyrimidines from cyanoacetylene and cyanate. *J. Mol. Biol.*, 33:693–704, 1968. 2.2

[37] M.P. Robertson and S. L. Miller. An efficient prebiotic synthesis of cytosine and uracil. *Nature*, 375:772–774, 1995. 2.2

[38] L. E. Orgel. Prebiotic adenine revisited: Eutectics and photochemistry. *Orig. Life Evol. Biosph.*, 34:361–369, 2004.

[39] R. Saladino, C. Crestini, G. Constanzo, and E. DiMauro. Advances in the prebiotic synthesis of nucleic acids bases: Implications for the origin of life. *Curr. Org. Chem.*, 8:1425–1443, 2004. 2.2

[40] W. D. Fuller, R. A. Sanchez, and L. E. Orgel. Studies in prebiotic synthesis. VI. Synthesis of purine nucleosides. *J. Mol. Biol.*, 67:25–33, 1972. 2.2

[41] C. M. Tapiero and J. Nagyvary. Prebiotic formation of cytidine nucleotides. *Nature*, 231:42–43, 1971.

[42] A.-A. Ingar, R. W. A. Luke, B. R. Hayter, and J. D. Sutherland. Synthesis of a cytidine ribonucleotide by stepwise assembly of the heterocycle on a sugar phosphate. *Chembiochem*, 4:504–507, 2003. 2.2

[43] R. Lohrmann. Formation of nucleoside 5'–polyphosphates from nucleotides and trimetaphosphate. *J. Mol. Evol.*, 6:237–252, 1975. 2.2

[44] G. J. Handschuh, R. Lohrmann, and L. E. Orgel. The effect of Mg^{2+} and Ca^{2+} on urea-catalyzed phosphorylation reactions. *J. Mol. Evol.*, 2:251–262, 1973. 2.2

[45] R. Osterberg, L. E. Orgel, and R. Lohrmann. Further studies of urea-catalyzed phophorylation reactions. *J. Mol. Evol.*, 2:231–234, 1973. 2.2

[46] S. Miyakawa and J. P. Ferris. Sequence- and regioselectivity in the montmorillonite-catalyzed synthesis of RNA. *J. Am. Chem. Soc.*, 125:8202–8208, 2003. 2.2

[47] J. P. Ferris, P. C. Joshi, K.-J. Wang, S. Miyakawa, and W. Huang. Catalysis in prebiotic chemistry: Application to the synthesis of RNA oligomers. *Adv. Space Res.*, 33:100–105, 2004.

[48] G. Ertem. Monthmorillonite, oligonucleotides, RNA and origin of life. *Orig. Life Evol. Biosph.*, 34:549–570, 2004. 2.2

[49] H.L. Sleeper and L. E. Orgel. The catalysis of nucleotide polymerization by compounds of divalen lead. *J. Mol. Evol.*, 12:357–364, 1979. 2.2

[50] H. Sawai, K. Kuroda, and T. Hojo. Efficient oligoadenylate synthesis catalyzed by uranyl ion complex in aqueous solution. *Nucleic Acids Symp. Ser.*, 19:5–7, 1988. 2.2

[51] G. F. Joyce. Non-enzymatic template-directed synthesis of informational macromolecules. *Cold Spring Harbor Symp. Quant. Biol.*, 52:41–51, 1987. 2.2

[52] L. E. Orgel. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.*, 39:99–123, 2004. 2.2

[53] R. Rohatgi, D. P. Bartel, and J. W. Szostak. Kinetic and mechanistic analysis of nonenzymatic, template-directed oligoribonucleotide ligation. *J. Am. Chem. Soc.*, 118:3332–3339, 1996. 2.2

[54] R. Rohatgi, D. P. Bartel, and J. W. Szostak. Nonenzymatic, template-directed ligation of oligoribonucleotides is highly regioselective for the formation of 3'–5' phosphodiester bonds. *J. Am. Chem. Soc.*, 118:3340–3344, 1996. 2.2

[55] K.D. James and A. D. Ellington. The fidelity of template-directed oligonucleotide ligation and the inevitability of polymerase funcion. *Orig. Life Evol. Biosph.*, 29:375–390, 1999. 2.2

[56] T. R. Cech. A model for the RNA-catalyzed replication of RNA. *Proc. Natl. Acad. Sci.*, 83:4360–4363, 1986. 2.2

[57] M. D. Been and T. R. Cech. RNA as an RNA polymerase: Net elongation of an RNA primer catalyzed by the *Tetrahymena* ribozyme. *Science*, 239:1412–1416, 1988.

[58] J. A. Doudna and J. W. Szostak. RNA-catalysed synthesis of complementary-strand RNA. *Nature*, 339:519–522, 1989. 2.2

[59] A. Eschenmoser. Towards a chemical etiology of nucleic acid structure. *Orig. Life Evol. Biosph.*, 27:535–553, 1997. 2.2

[60] K. Schoning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser. Chemical etiology of nucleic acid structure: The alpha-threofuranosyl-(3'->2') oligonucleotide system. *Science*, 290:1347–1351, 2000. 2.2

[61] P.E. Nielsen. Peptide nucleic acid (PNA): A model structure for the primordial genetic material? *Orig. Life Evol. Biosph.*, 23:323–327, 1993. 2.2

[62] P. Babitzke, J. Schaak, A. V. Yakhnin, and P. C. Bevilacqua. Role of RNA structure in transcription attenuation in Bacillus subtilis: the trpEDCFBA operon as a model system. *Methods Enzymol*, 371:392–404, 2003. 2.3

[63] A. P. McGraw, P. C. Bevilacqua, and P. Babitzke. TRAP-5' stem loop interaction increases the efficiency of transcription termination in the Bacillus subtilis trpEDCFBA operon leader region. *RNA*, 13(11):2020–33, 2007. 2.3

[64] G. Klug and S. N. Cohen. Combined actions of multiple hairpin loop structures and sites of rate-limiting endonucleolytic cleavage determine differential degradation rates of individual segments within polycistronic puf operon mRNA. *J Bacteriol*, 172(9):5140–6, 1990. 2.3

[65] I. M. Khan and J. M. Coulson. A novel method to stabilise antisense oligonucleotides against exonuclease degradation. *Nucleic Acids Res*, 21(18):4433, 1993. 2.3

[66] Neocles B. Leontis, Jesse Stombaugh, and Eric Westhof. The non–Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Research*, 30(16):3497–3531, August 2002. 2.3.1

[67] N. B. Leontis and E. Westhof. Geometric nomenclature and classification of RNA base pairs. *RNA*, 7(4):499–512, 2001. 2.3.1

[68] R.E. Dickerson. Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Research*, 17(5):1797–1803, March 1989. 2.3.2

[69] X. J. Lu and W. K. Olson. Illustration of base-pair parameters. `http://rutchem.rutgers.edu/~xiangjun/3DNA/examples.html`, April 2011. (document), 2.3

[70] Westhof E. Auffinger P. Hydration of RNA base pairs. *J. Biomol. Struct. Dyn.*, 16:693–707, 1998. 2.3.2

[71] E. M. Moody and P. C. Bevilacqua. Folding of a stable DNA motif involves a highly cooperative network of interactions. *J Am Chem Soc*, 125(52):16285–93, 2003. 2.3.3

[72] N. A. Siegfried, R. Kierzek, and P. C. Bevilacqua. Role of unsatisfied hydrogen bond acceptors in RNA energetics and specificity. *J Am Chem Soc*, 132(15):5342–4, 2010. 2.3.3, 5.1.3

[73] N. A. Siegfried, S. L. Metzger, and P. C. Bevilacqua. Folding cooperativity in RNA and DNA is dependent on position in the helix. *Biochemistry*, 46(1):172–81, 2007. 2.3.3, 5.1.3

[74] T. J. Macke and D. A. Case. Modeling Unusual Nucleic Acid Structures. In Leontis Neocles B. and SantaLucia John, editors, *Molecular Modeling of Nucleic Acids*, pages 379–393. American Chemical Society, 1998. 3.1, 4.3

[75] P. Banas, P. Jurecka, N. G. Walter, J. Sponer, and M. Otyepka. Theoretical studies of RNA catalysis: hybrid QM/MM methods and their comparison with MD and QM. *Methods*, 49(2):202–16, 2009. 3.1, 3.1.1

[76] T.E. Cheatham III C.L. Simmerling J. Wang R.E. Duke R. Luo M. Crowley R.C.Walker W. Zhang K.M. Merz B.Wang S. Hayik A. Roitberg G. Seabra I. KolossvÃˇry K.F.Wong F. Paesani J. Vanicek X.Wu S.R. Brozell T. Steinbrecher H. Gohlke L. Yang C. Tan J. Mongan V. Hornak G. Cui D.H. Mathews M.G. Seetin C. Sagui V. Babin D.A. Case, T.A. Darden and P.A. Kollman. AMBER 10, 2008. 3.1.1

[77] J. M. Wang, P. Cieplak, and P. A. Kollman. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry*, 21(12):1049–1074, 2000. 3.1.1

[78] A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophys J*, 92(11):3817–3829, June 2007. 3.1.1, 4.3

[79] Zgarbova et al. manuscript in preparation. 3.1.1, 5.1.1

[80] V. Mlynsky, P. Banas, D. Hollas, K. Reblova, N. G. Walter, J. Sponer, and M. Otyepka. Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H+ forms are most consistent with crystal structures of hairpin ribozyme. *J Phys Chem B*, 114(19):6642–52, 2010. 3.1.1, 5.1.1, 5.1.3

[81] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, 91:6269–6271, 1987. 3.1.2

[82] P. G. Kusalik and I. M. Svishchev. The spatial structure in liquid water. *Science*, 265(5176):1219–1221, 1994. 3.1.2

[83] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926–935, 1983. 3.1.2

[84] M. W. Mahoney and W. L. Jorgensen. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.*, 112:8910–8922, 2000. 3.1.2

[85] P. Florova, P. Sklenovsky, P. Banas, and M. Otyepka. Explicit Water Models Affect the Specific Solvation and Dynamics of Unfolded Peptides While the Conformational Behavior and Flexibility of Folded Peptides Remain Intact. *J. Chem. Theor. Comp.*, 6(11):3569–3579, 2010. (document), 3.1.2, 3.2

[86] Pascal Auffinger, Thomas E. Cheatham, and Andrea C. Vaiana. Spontaneous Formation of KCl Aggregates in Biomolecular Simulations: A Force Field Issue? *Journal of Chemical Theory and Computation*, 3(5):1851–1859, September 2007. 3.1.2

[87] I. S. Joung and 3rd Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B*, 112(30):9020–41, 2008. 3.1.2, *b*, *c*

[88] J. Aqvist. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.*, 94:8021, 1990. *a*

[89] L. X. Dang. Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: A molecular dynamics study. *J. Am. Chem. Soc.*, 117:6954–6960, 1995. *a*

[90] D. E. Smith and L. X. Dang. Computer simulations of NaCl association in polarizable water. *J. Chem. Phys.*, 100:3757–3765, 1994. *a*

[91] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38, February 1996. 4.1

[92] Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. August 2010. 4.1

[93] Xiang-Jun Lu and Wilma K Olson. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protocols*, 3(7):1213–1227, July 2008. 4.1

[94] S. Lima, J. Hildenbrand, A. Korostelev, S. Hattman, and H. Li. Crystal structure of an RNA helix recognized by a zinc-finger protein: an 18-bp duplex at 1.6 A resolution. *RNA*, 8(7):924–32, 2002. 4.2

[95] Pavel Banas, Daniel Hollas, Marie Zgarbova, Petr Jurecka, Modesto Orozco, Thomas E. Cheatham, Jiri Sponer, and Michal Otyepka. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *Journal of Chemical Theory and Computation*, 6(12):3836–3849, 2010. 4.3, 5.1.1, 5.1.3

Appendix A

Appendix A

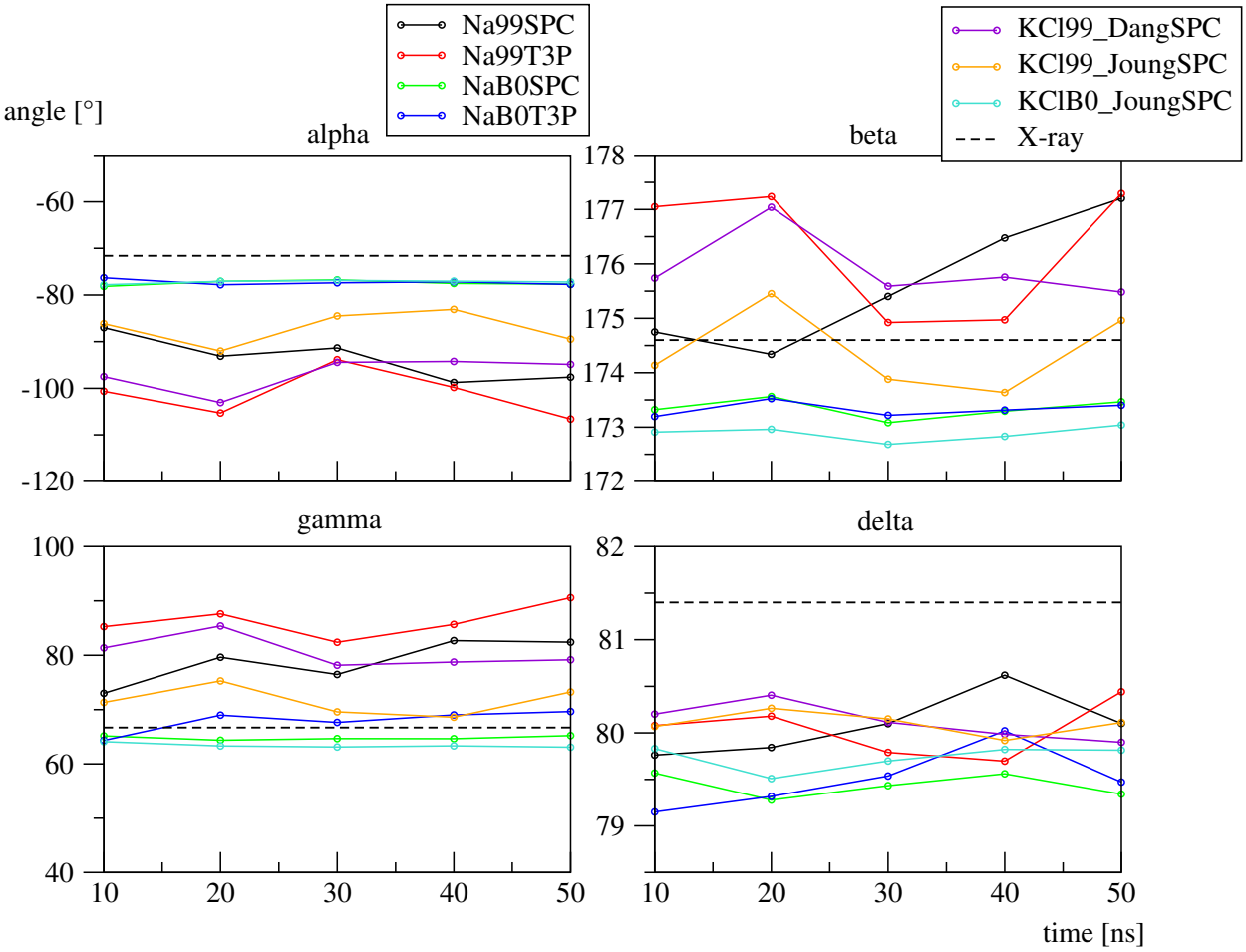

Figure A.1: Progress in time of averaged values (averaged for 10 ns intervals) for alpha, beta, gamma and delta torsion angles.

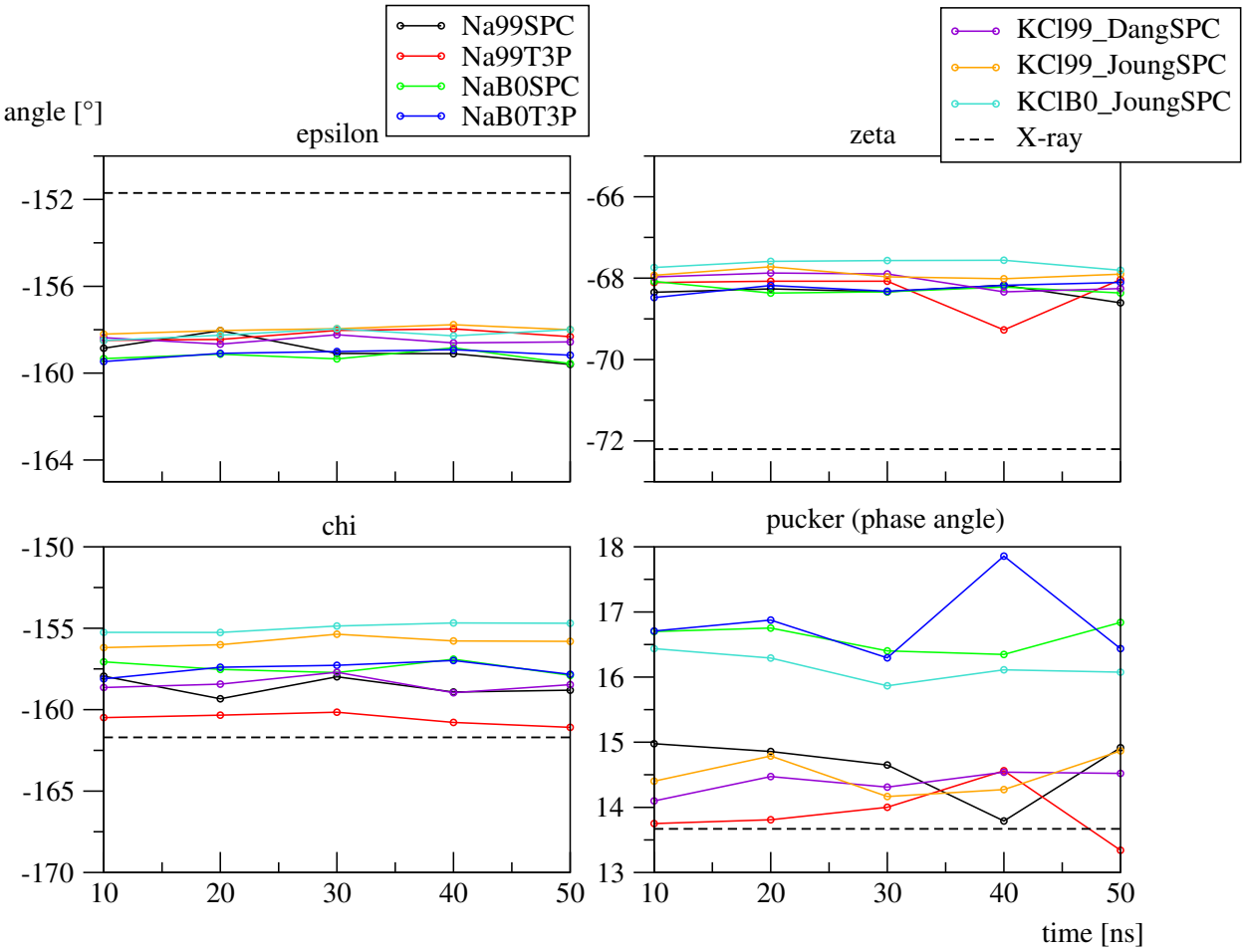

Figure A.2: Progress in time of averaged values (averaged for 10 ns intervals) for epsilon, zeta, chi torsion angles and pucker phase angle.

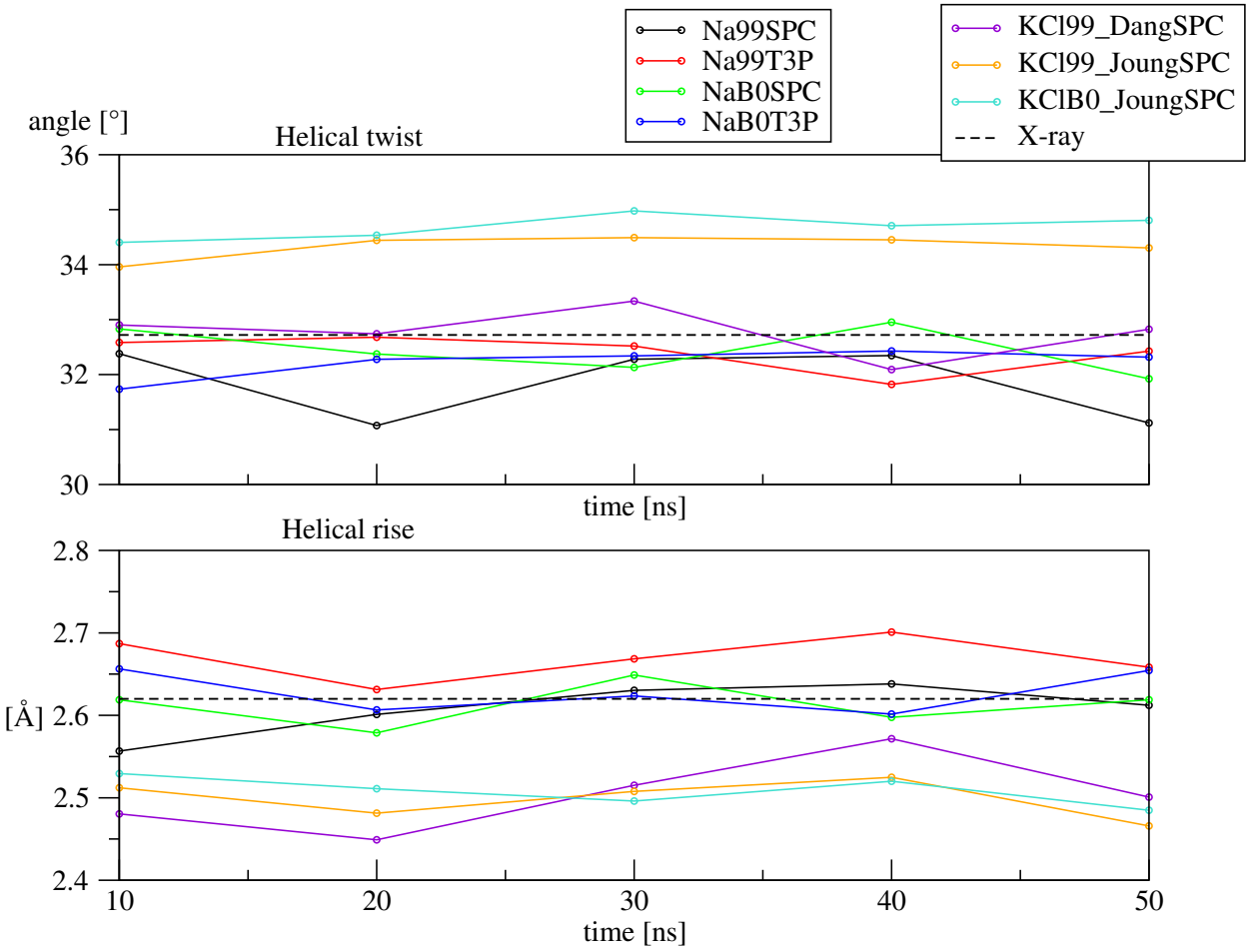

Figure A.3: Progress in time of averaged parameters (averaged for 10 ns intervals) for helical twist and helical rise.

Appendix B

Appendix B

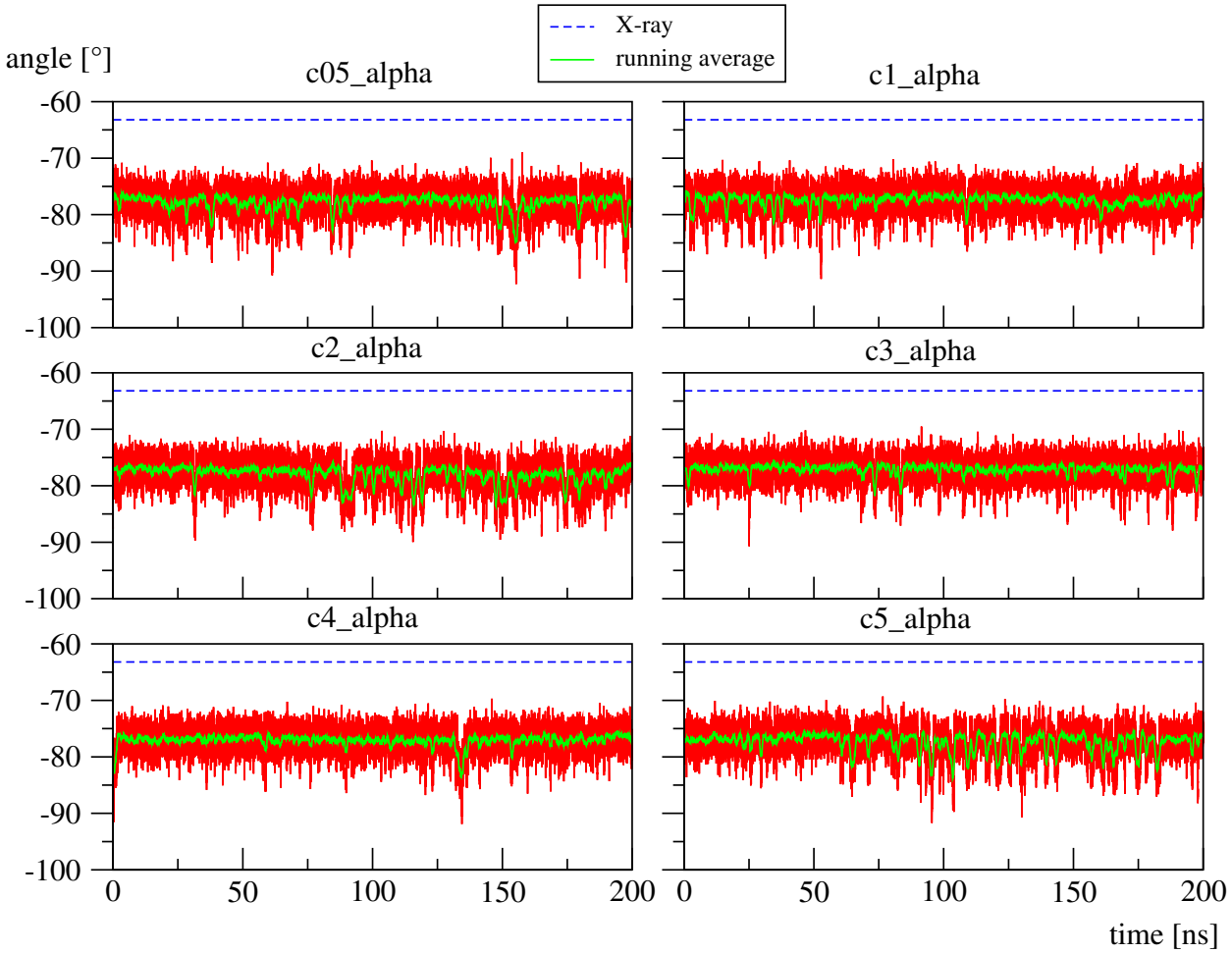

Figure B.1: Progress in time of alpha torsion angle for different ionic strength. Running averages are over set of 100 values.

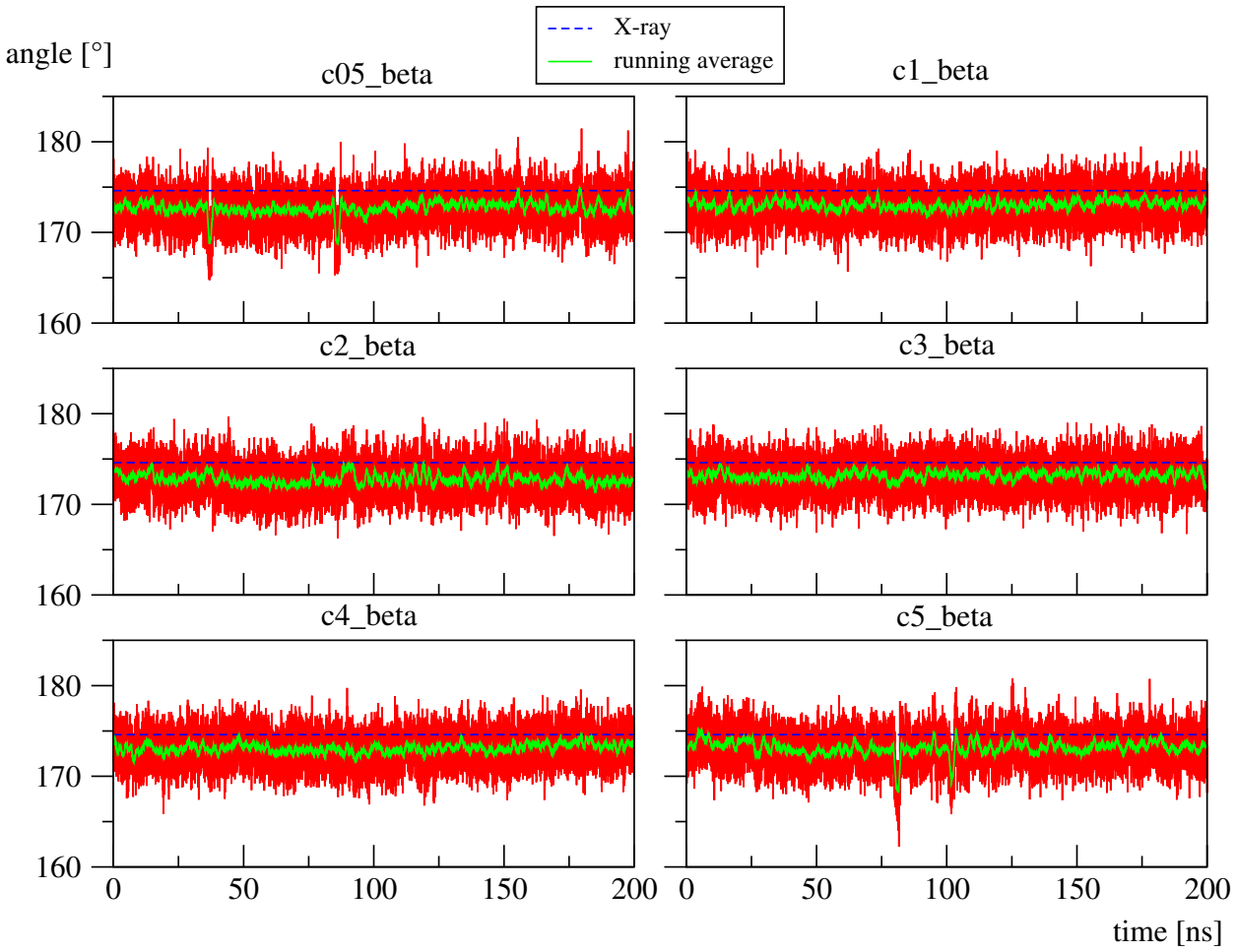

Figure B.2: Progress in time of beta torsion angle for different ionic strength. Running averages are over set of 100 values.

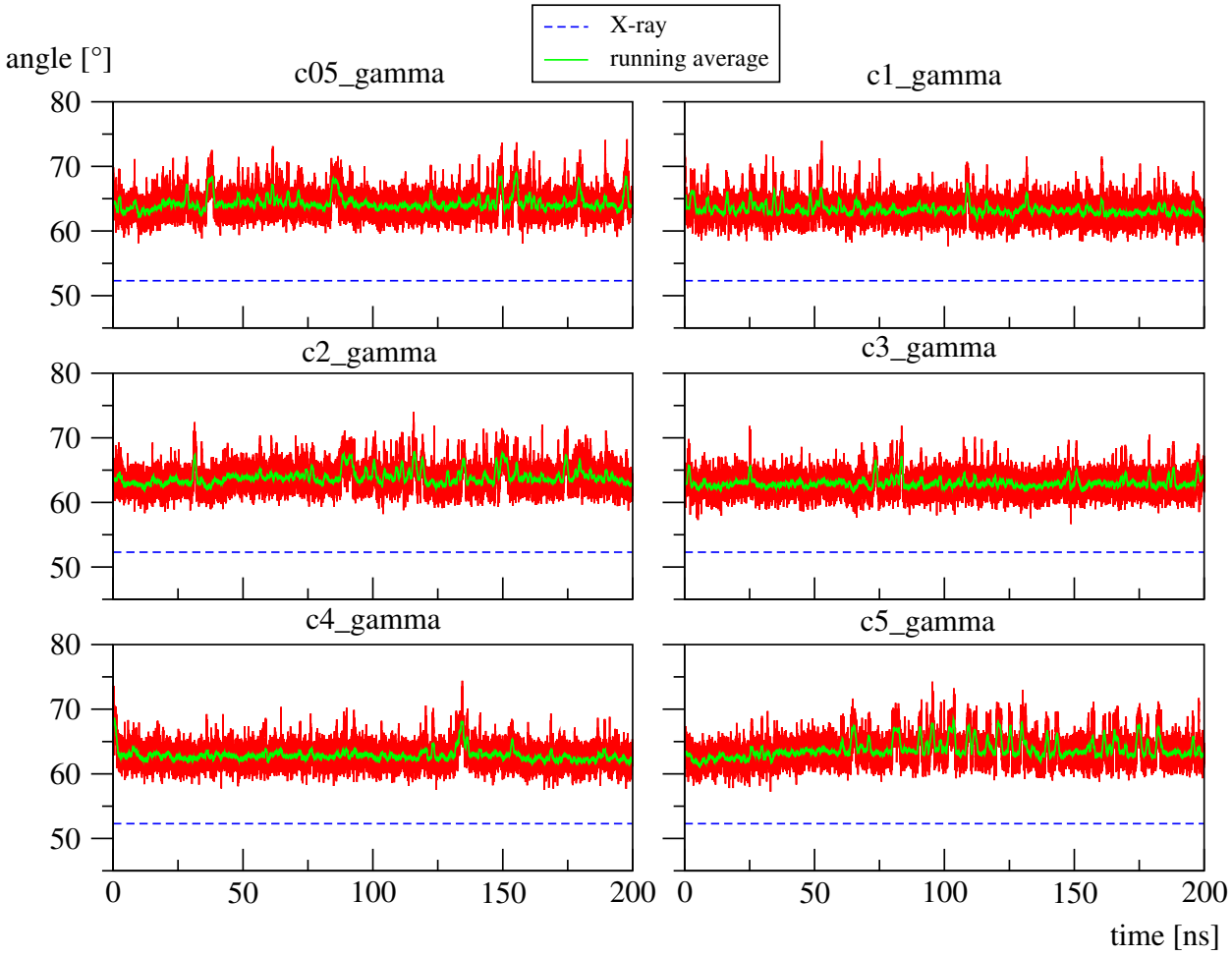

Figure B.3: Progress in time of gamma torsion angle for different ionic strength. Running averages are over set of 100 values.

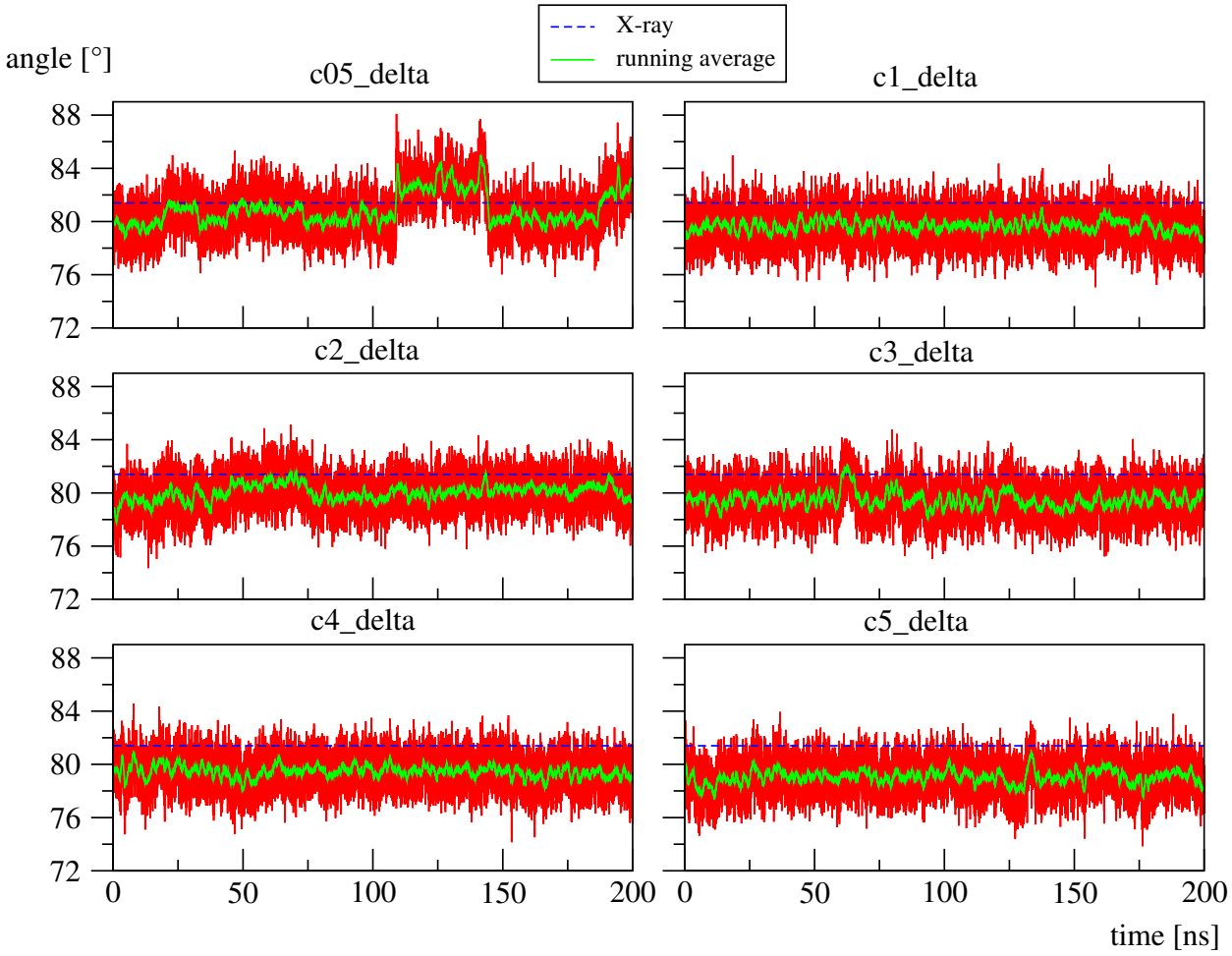

Figure B.4: Progress in time of delta torsion angle for different ionic strength. Running averages are over set of 100 values.

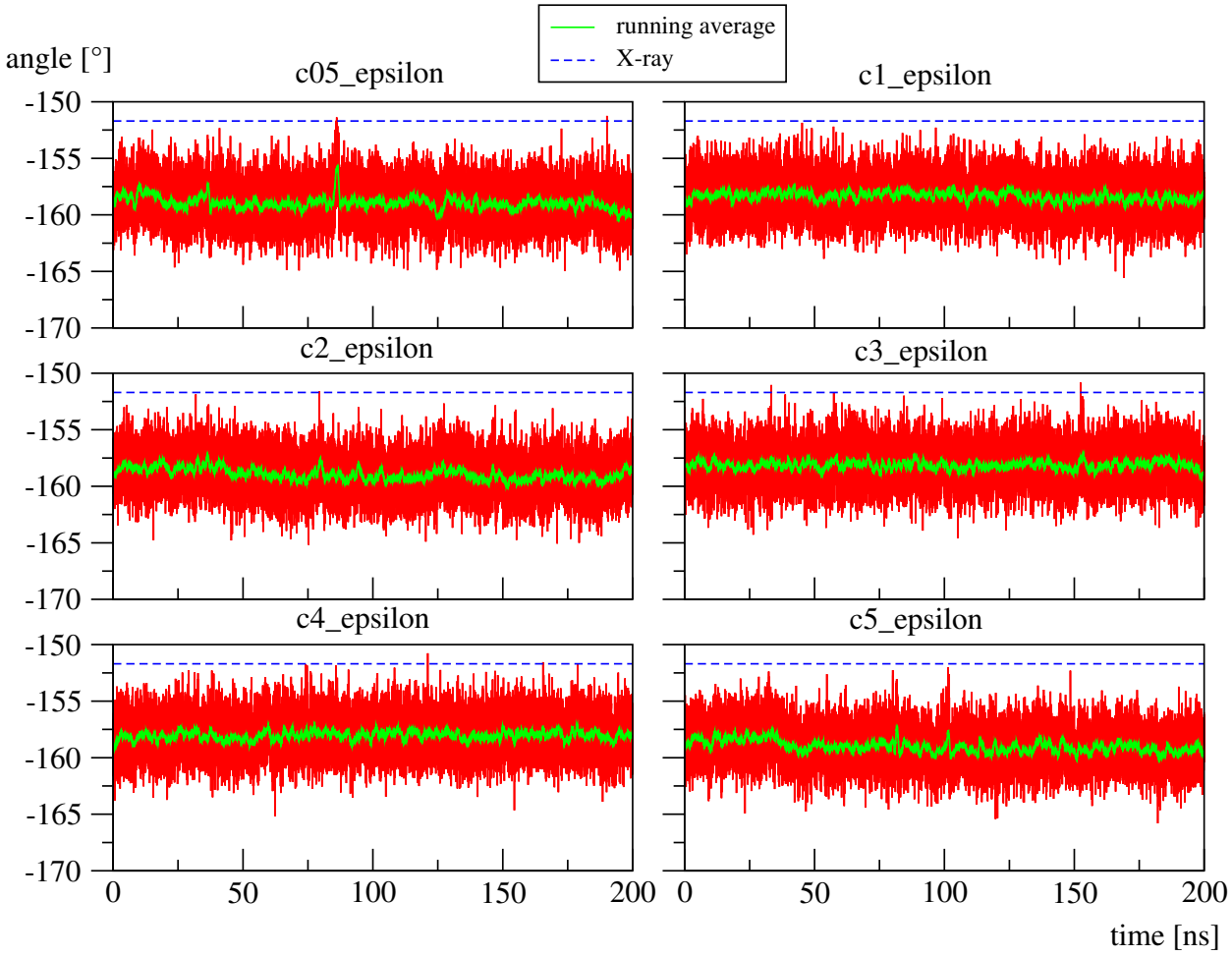

Figure B.5: Progress in time of epsilon torsion angle for different ionic strength. Running averages are over set of 100 values.

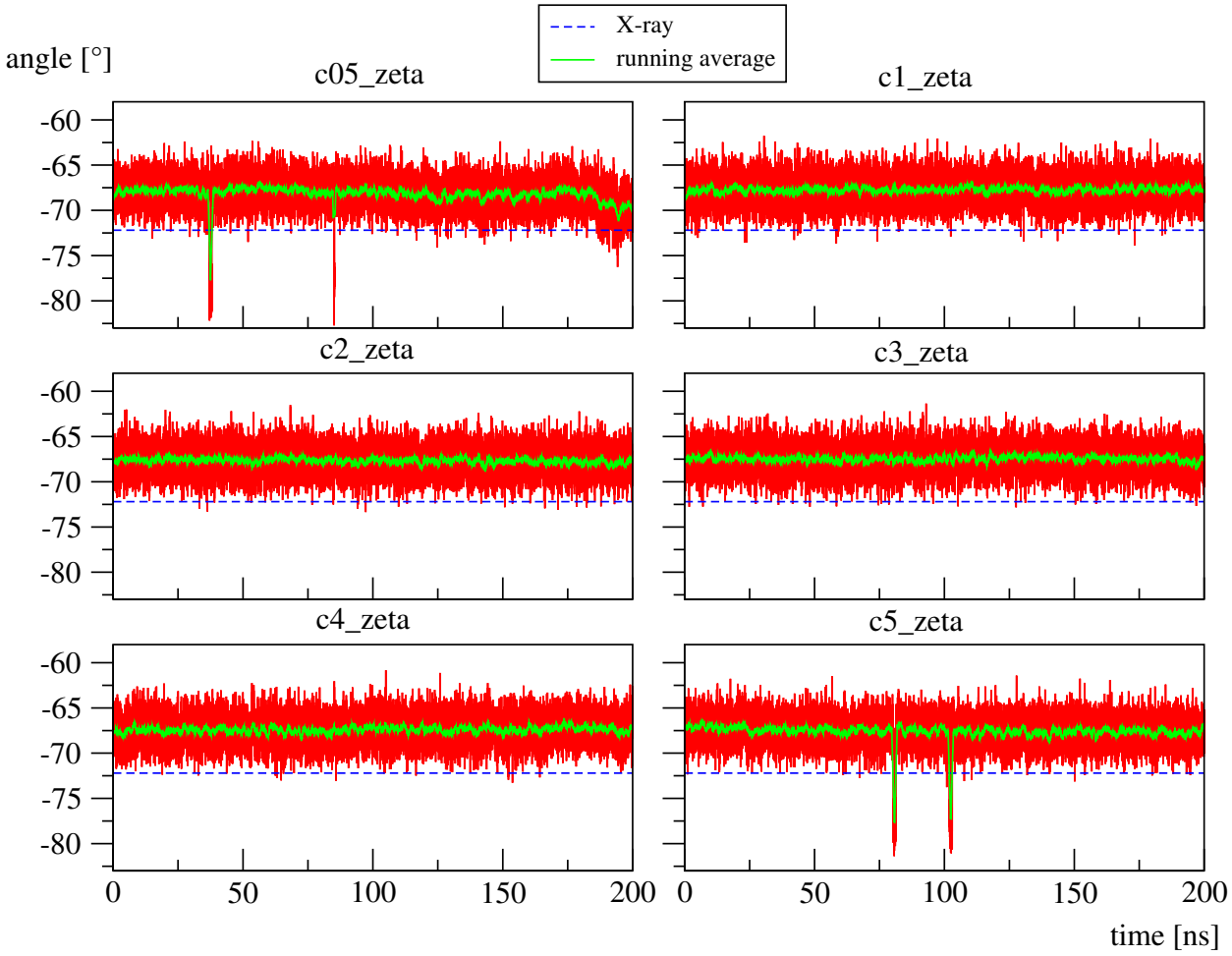

Figure B.6: Progress in time of zeta torsion angle for different ionic strength. Running averages are over set of 100 values.

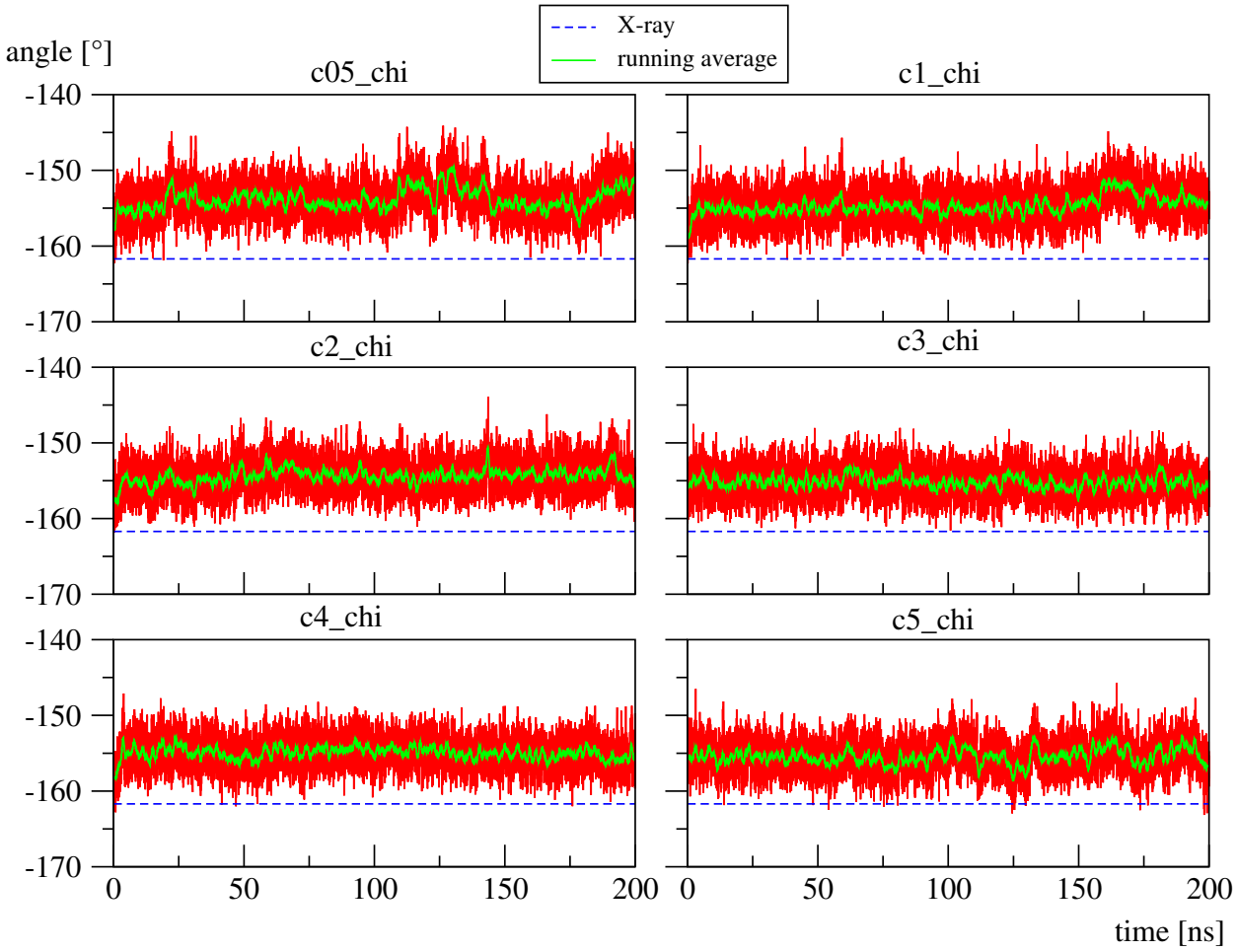

Figure B.7: Progress in time of chi torsion angle for different ionic strength. Running averages are over set of 100 values.