# Dynamical Properties of Biomolecules, Ions and Glass-Forming Liquids: A Theoretical and Computational Study

Ailun Wang

A dissertation

submitted to the Faculty of

the department of Chemistry

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

June 2021

Dynamical Properties of Biomolecules, Ions and Glass-Forming Liquids:

A Theoretical and Computational Study

Ailun Wang

Advisor: Professor Udayan Mohanty

**Abstract:**

The conformational dynamics plays a significant role in a wide range of biological systems, from small RNA molecules to the large-scale ribonucleoprotein assemblies, in which ions are found critical and have notable structural and functional impacts. In the glass-forming liquids, the structural dynamics also calls for further investigations and deeper understandings. To this end, using four distinct chapters, this dissertation discusses the ion-related conformational dynamics in various scales of biomolecular systems, as well as the fluctuation effects in the glass-forming liquids.

In chapter 1, we investigate the dynamics of ions and water molecules in the outer solvation sphere of a widely studied 58-nucleotide rRNA fragment. Molecular dynamics (MD) simulations with explicit solvent molecules and atomic details are performed for the RNA fragment in ionic solution. We determine all of the association sites and spatial distributions of residence times for $Mg^{2+}$ $K^{+}$, and water molecules in those sites. In accordance to the analysis of the dynamics of the RNA fragment, we provide insights into how the dynamics of ions and water molecules are intricately linked with the kinetics of the RNA fragment. In addition, the long-lived sites for $Mg^{2+}$ ions identified from the simulation agree with the metal ion locations determined in the X-ray structure. The excess ion atmosphere around the RNA fragment is calculated and compared with the experimental measures. The results from this study indicate that the 58-mer rRNA fragment in ionic solution forms a complex polymer that is encased by a fluctuating network of ions and water.

In chapter 2, the conformational dynamics of a large-scale ribonucleoprotein assembly, ribosome, is studied with molecular dynamics simulations with a newly de-

veloped model that accounts for electrostatic and ionic effects on the biomolecules. In this study, an all-atom structure based model is constructed with explicit representations of non-hydrogen atoms from biomolecules and diffuse ions. Implicit treatment is applied to the solvent molecules with the solvation effect associated with diffuse ions described by effective potentials. Parameters in this model are refined against explicit solvent simulations and experimental measures. This model with refined parameters is able to capture the excess $Mg^{2+}$ ions for prototypical RNA systems and their dependence on the $Mg^{2+}$ concentrations. Motivated by this, we apply the model to a bacterial ribosome and find that the position of the extended L1 stalk region can be controlled by the diffuse ions. This simulation also indicated ion-induced long-range interactions between L1 stalk and tRNA, which provides insights into the impact of ions on the functional rearrangements of ribosome.

In chapter 3, we focus on the dynamics of the glass-forming liquids. In this study, we generalized the Adam–Gibbs model of relaxation in glass-forming liquids and take into account the fluctuations in the number of molecules inside the cooperative rearranging region. We obtain the expressions of configurational fractions at the glass-transition temperature with and without the fluctuation effect in Adam–Gibbs model, and determine the configurational fraction for several glass-forming liquids at glass-transition temperature in the absence of fluctuation effects. A connection between the $\beta$ Kohlrausch–Williams–Watts parameters and the configurational fraction at the glass-transition temperature is also reported in this study.

In chapter 4, we apply the model developed in chapter 2 to a ribosome structure to investigate the effects of diffuse ions on the aminoacyl-tRNA (aa-tRNA) accommodation process. The aa-tRNA accommodation is a critical step in the tRNA selection process which serves the purpose of protein synthesis in the ribosome. Experimental and computational efforts were made to reveal the mechanism and the energetic properties of the accommodation process, while the effects from diffuse ions on this

process remain elusive. To this end, we design and perform MD simulations of ribosome structure with different treatment of electrostatics and diffuse ions in the system. Simulations with various ionic concentrations are also performed to study the concentration effects. The simulation trajectories indicate that diffuse ions can facilitate the aa-tRNA accommodation process and stabilize the accommodated configurations. In addition, we observe that $Mg^{2+}$ ions play critical roles in stabilizing the accommodated configurations and a few millimolar change of $Mg^{2+}$ concentration can alter the tendency of the tRNA configurational change during the accommodation process. This result shed light on the investigations of suitable ionic environment for the tRNA selection in the ribosome. It will be fruitful to extend this strategy into the investigations of other conformational rearrangements in the ribosome, such as tRNA translocation and subunit rotation, which will provide us with deeper understanding about the functional mechanism of the ribosome.

# Table of Contents

# Acknowledgments

This dissertation represents not only my work on my research projects and at the keyboard, it concentrates the knowledge and support from my supervisors, thesis committee members, collaborators, and colleagues. It would not have been possible for me to finish this journey without the support and guidance that I received from many people.

First and foremost, I wish to thank my supervisor, Professor Udayan Mohanty, for his guidance and knowledge that lead me through all my research projects. I also wish to give my thank to Professor Paul C. Whitford, our long-term collaborator and whom I consider as my second supervisor, for his support and guidance in every ribosome project that I was involved in. Second, I wish to express my appreciation to my committee members, Professor Dunwei Wang, Professor Matthias Waegele, and Professor Lucas Bao, and my special appreciation to Professor Frank Tsung, for their support along my journey of research at Boston College. Many thanks also go out to the support and help from my collaborators. I would like to express my thank to Dr. Heiko Lammert, Professor José N. Onuchic for their help in our collaborations in the work presented in Chapter 1. I would like to thank Dr. Mariana Levi for her help and instructions in the early stages of the work presented in Chapter 2. I would like to extend my thanks to Charley Hutchison and Dr. Ajaya Bhattarai for their help in the work presented in Chapter 3. My thanks also go out to the help I received from my colleagues, especially Yang Li, Jingru Luo, Yang Wang, Sandra Byju, Asem

Hassan and Frederico Freitas. Finally, I would like to say a heartfelt thank you to my family and friends, especially my best friend Jia, for their unconditional support and encouragement through not only my journey of research, but also my journey of life.

# List of Abbreviations

$T_{\mathbf{K}}$ Kauzmann temperature.

$T_{\mathbf{g}}$ glass transition temperature.

$\mathbf{K}^{+}$ potassium ion.

$\mathbf{Mg}^{2+}$ magnesium ion.

$\mathbf{Os}^{3+}$ osmium(III) ion.

**Å** Angstrom.

***n*POH** *n*-propanol.

**aa-tRNA** aminoacyl-tRNA.

**BP** base pair.

**BPT** 3-bromopentane.

**cryo-EM** cryogenic electron microscopy.

**DH** Debye-Hückel.

**EF-G** elongation factor G.

**EF-Tu** elongation factor Tu.

**GTP** Guanosine-5'-triphosphate.

**HB** hydrogen bond.

**KWW** Kohlrausch–Williams–Watts.

**MD** molecular dynamics.

**mM** millimolar.

**mNRA** messenger RNA.

**MTHF** 2-methyltetrahydrofuran.

**NLPB** nonlinear Poisson-Boltzmann.

**OTP** *ortho*-terphenyl.

**PC** principle component.

**PCA** principal component analysis.

**PMF** potential of mean force.

**PTC** peptidyl transderase center.

**RDF** radial distribution function.

**RMSF** root mean square fluctuation.

**RNA** $\beta$-D-ribonucleic acid.

**RNP** ribonucleoprotein.

**rRNA** ribosomal RNA.

**SBM** strucsre-based model.

**SDF** spatial distribution function.

**smFRET** single-molecule fluorescence resonance energy transfer.

**tRNA** transfer RNA.

**VFT** Vogel-Fulcher-Tammann.

# Glossary of Terms

**All-atom models.** It is a type of models of molecular dynamics simulations that provides atomic details of the molecules in the system. The coordinates of all atoms (at least all non-hydrogen atoms) are explicitly represented. This model is suitable for simulations that aim to provides details of atomic interactions.

**Aminoacyl-tRNA.** It is a tRNA molecule that delivers amino acid to the ribosome for protein synthesis. In an aminoacyl-tRNA, the cognate amino acid of the tRNA is chemically bond to its 3'-CCA end.

**Coarse-grained model.** It is a type of models that use simplified representations of residues or molecules for the simulations. In this type of models, molecules are not represented by individual atoms. Instead, a group of atoms, such as a nucleic or amino acid residue or a block or residues, is represented by one "pseudo-atom". This model is suitable for simulating the behavior of complex systems.

**Cryogenic electron microscopy (cryo-EM).** It is an electron microscopy technique that is used for measuring the structure of biomolecules at near-atomic resolution. During the measurement, samples are cooled to cryogenic temperature and embedded in an environment of vitreous water. The electron beam then scans the sample and the projection images are collected by the electron detector beneath the sample. The 3D structure of the sample biomolecule can be created from the collected projections.

**Debye-Hückel potential.** It is the potential energy of a screened coulomb

interactions between a pair of charges. Here, the coulomb potential is multiplied with an exponential damping term, with the strength of the damping factor given by the magnitude of Debye length, which corresponds to a given ionic concentration.

**Explicit solvent model.** It is a type of models for molecular dynamics simulations which treats the solvent molecules explicitly (i.e., the coordinates of solvent molecules are explicit and usually at least some of the molecular degrees of freedom are included).

**Force field.** In molecular modelling, force field refers to the functional expression and corresponding parameters that are used to calculate the forces and energetic potentials of a system of interacting atoms or molecules.

**Glass transition.** This refers to a transition from liquid phase to a glassy form in amorphous state that lack the periodicity of crystals but behave mechanically like solids. The most common approach of achieving glass transition is by cooling a viscous liquid fast enough to avoid crystallization (supercooling).

**Glass transition temperature.** The range of temperatures over which the glass transition occurs. The glass transition temperature is always lower than the melting temperature of the crystalline state of the material.

**Implicit solvent model.** It is a type of models that treats solvents as a continuous medium instead of providing individual coordinates for each solvent molecules.

**Molecular dynamics simulation.** A computational simulation methods that allows atoms or molecules to interact for a fixed period of time. The simulation trajectories of atoms or molecules are usually determined by numerically solving Newton's equation of motion for the system that contains interacting atoms/molecules. The forces between atoms/molecules and their potential energy are often calculated with molecular mechanics force fields.

**Preferential interaction coefficient.** This is a quantity that measures the number of excess ions that a polyelectrolyte (e.g., RNA or DNA ) attracts into its

local environment. The excess number of ions in the local environment of DNA or RNA is compared with the number of ions in the bulk solution with the same volume.

**Ribosome.** It is a biomolecular machinery in the living cells that synthesis proteins by translating the sequence of codons on the mRNA into the order of amino acids and add amino acid residues to the peptide.

**Single molecule fluorescence resonance energy transfer (smFRET)** This is a biophysical technique that enables distances between single molecules to be measured at the scale of 1–10 nanometers. In the measurement, a pair of donor and acceptor fluorophores are excited and detected on a single molecule level.

**Structure-base model (SBM).** It is a simplified model that is constructed on the basis of the principal of minimal frustration and the folding funnel concept. It describes an energy landscape where the native interactions are on average more stabilizing than non-native interactions. The potential energy function of SBM is defined by a know structure, usually the native structure.

# Introduction

In this dissertation, I used theoretical and computational methods to address a variety of interdisciplinary questions in the field of biological physics and chemical physics that are significant to both experimentalists and theorists. Some of the questions addressed in this dissertation are: How do ions interact with $\beta$-D-ribonucleic acid (RNA)? What are the dynamics of ions and water molecules in the outer solvation shells of the RNA? Where do the association sites of diffuse ions location in the vicinity of the RNA? What are the timescales of the residence time of diffuse ions in their association sites near the RNA? What are the role of diffuse ions in the conformational dynamics of the large-scale ribonucleoprotein (RNP) assemblies, such as ribosome? How to construct an efficient simplified model to describe the ionic effects on the RNP assemblies? How do diffuse ions regulate the conformational dynamics of ribosome? What are the effects of electrostatics and ionic environment on the functional dynamics of the ribosome and the transfer RNA (tRNA)? Will the diffuse ions alter the the aminoacyl-tRNA (aa-tRNA) accommodation process in the ribosome?

In the following subsections, I summarized the concepts and theory related to the topics investigated in this dissertation, including: (A) Qualitative and quantitative descriptions of ionic interactions with RNA and the dynamics of ions as water in the outer solvation shell of RNA; (B) How do diffuse ions regulate the conformational dynamics of the RNP; (C) Fluctuation effects in the Adam-Gibbs model of coopera-

tive relaxation of glass-forming liquids; (D) The role of diffuse ions in the aa-tRNA accommodation in ribosome.

## 0.1. Dynamics of ions and water in the outer solvation sphere of 58-mer rRNA.

### 0.1.1. Ion-RNA interactions.

RNA are highly negative charged biopolymers, whose conformational dynamics as well as the biological functions are affected by the ionic environments, especially the metal ions with positive charges [1, 2]. Metal ions maintain the compact structure of RNA by reducing the repulsion between negatively charged RNA phosphates, but some ions are much more efficient than others at this task. For instance, millimolar (mM) magnesium ion ($Mg^{2+}$) are able to stabilize RNA structures which are only marginally stable at high monovalent metal ion concentrations[3–5]. The significance of metal ions in the RNA system motivated the investigation of ion-RNA interactions, as well as the competition of ions in terms of the interactions with the same RNA molecule [6, 7]. However, how does metal ion recognize and stabilize RNA structures are not well understood. The ionic effects related to the sequences of RNA, multitude of cellular signals, and the formation of tertiary structures are also elusive.

The ionic environment around the RNA is often partitioned into chelated and diffuse ions according to their distance from the RNA surface atoms and the hydration states of the ions [2]. Diffuse ions are usually fully hydrated with water molecules. The interactions between diffuse ions and the RNA or other ions are dominant by electrostatic interactions. The density of diffuse ions in a local region is dictated by the electrostatic potential of the RNA [8]. In the system of RNA with ions, divalent and monovalent diffuse ions compete for interactions with RNA. When one $Mg^{2+}$

ion associates with RNA, it releases approximately two potassium ion ($K^+$) ions. In contrast to diffuse ions, chelated ions are partially dehydrated to make direct contacts with the molecular surface of RNA [2, 8, 9]. In RNA, those metal ions are usually chelated by negatively charged groups or atoms in RNA, such as phosphate oxygen atoms [10, 11]. Although the chelated ions are held in place by electrostatics forces, the displacement of some water molecules in their first hydration shell is a significant energetic consideration [10].

The underlying RNA architecture and the rugged landscape of its molecular surface manifests in wide-range of temporal and spatial dynamics in aqueous solutions [12–22]. Simulation techniques such as implicit solvent simulations [7, 19, 23–35], and all-atom explicit solvent simulations [14, 36–41] have been utilized to investigate structure and dynamics of RNAs and bio-assemblies in aqueous solution. However, the dynamics of water molecules and monovalent or divalent ions in the vicinity of the RNA fragments has not been well illustrated. How the dynamics of ions and solvent molecules cooperate with the dynamics of RNA molecule is not yet clear. To this end, we applied atomic resolution explicit solvent molecular dynamics (MD) simulations to investigate the dynamics of ions and water in the outer solvation sphere of a widely studied 58-nucleotide ribosomal RNA (rRNA) fragment.

### 0.1.2. 58mer ribosomal RNA fragment.

In the ribosome, the most important functional sites are associated with highly conserved domains of rRNA. In the large-subunit of the rRNA, an extraordinary well conserved region is a 58-nucleotide rRNA domain which is associated with the ribosomal protein L11 and plays significant role in the binding sites of two elongation factors (elongation factor Tu (EF-Tu) and elongation factor G (EF-G)) [42]. The crystal structure of this 58mer rRNA fragment has been solved at 2.8 Angstrom (Å) resolution [10], in which several metal ions were found in the rRNA fragment and the

locations of ions were determined together with the configuration of the 58mer. The crystal structure of the 58mer (PDB: 1HC8) contains two asymmetric protein-RNA assemblies in the unit cell. 21 coordinates of $Mg^{2+}$ ions were reported in the two asymmetric assemblies of the 58mer (chain C and chain D in PDB 1HC8). Alignment of the monomers reveals 13 distinct binding positions for $Mg^{2+}$ ions, where 8 are occupied in both RNA molecules. In these $Mg^{2+}$ ions, Misra *et. al.* reported that only one ion is likely to be strongly bound (i.e. chelated) via nonlinear Poisson-Boltzmann (NLPB) calculations, where the calculated binding free energy was -4.8 kcal/mol [43]. In addition to the chelated $Mg^{2+}$, a chelated $K^{+}$ is also reported in the 58mer rRNA with overwhelmingly large binding free energy (-30 kcal/mol) [10]. Moreover, two osmium(III) ion ($Os^{3+}$) were resolved in the crystal structure of 58mer which were not classified as chelated ions.

### 0.1.3. Excess ion atmosphere of RNA.

Due to the highly negatively charged nature of RNA, positively charged metal ions are attracts to the vicinity of RNA to balance the charge from RNA and stabilize its compact structure. Therefore, the local density of metal ions near the RNA is slightly higher than that in the bulk solution. The "excess" number of ions that accumulate around the RNA molecule due to electrostatic interactions is quantified as the preferential interaction coefficient ($\Gamma$) of ions. The value of $\Gamma$ for a specific RNA is found sensitive to the ionic concentrations of the solution [11, 44]. This quantity can be obtained experimentally using fluorescence titration [11], or theoretically using generalized Manning condensation models [45], or computationally with explicit solvent simulations [46]. The preferential interaction coefficient is an exquisite measure of the long-range electrostatic interactions between ions and the RNA atoms. It also provides an excellent metric for comparing theoretical models and simulations with experiment measures.

In addition to the number of excess ions in the vicinity of RNA, the locations of those excess ions relative to the conformation of the RNA also play a significant role in investigating the ion-RNA interactions and revealing the role of ions in the conformational stability and fluctuations of RNAs [46, 47]. Together with the locations of the association sites of metal ions, the residence time of metal ions in each association sites is also important to deepen our understanding of dynamics of ions near RNA molecule and their correlation with the RNA dynamics.

In order to closely investigate the ion-RNA interaction and unravel the dynamics of ions and water molecules near the RNA molecule, we performed explicit solvent MD simulations with the widely studied and well conserved rRNA fragment, a 58-mer, in ionic solutions. The locations of all the association sites of $Mg^{2+}$, $K^{+}$, and water molecules were determined, as well as the spatial distribution of residence times for those sites. The simulation result provides significant insights not only into the population of ions and water molecules in the outer sphere of the RNA, but also illustrated how their fluctuation are intricately correlated with the kinetics of the 58-mer. In addition, a joint principal component analysis (PCA) analysis was performed for the Cartesian coordinates of the RNA phosphorus atoms as well as the occupation counts of the association sites of ions and water molecules, which indicated that 58-mer rRNA is a complex polymer in conjunction with ions and water in the outer solvation sphere. The details of this work is discussed in the Chapter 1.

## 0.2. The role of metal ions in the configurational dynamics of ribonucleoprotein assemblies.

Besides small RNA fragments and molecules, metal ions are also found critical to the structural stability and the dynamics of large ribonucleoprotein(RNP) assemblies[48, 49]. Ribosome is one of the most well-characterized RNP assemblies. The signifi-

cance of counterions in the assembly[50, 51] and conformational transitions between functional states[52, 53] of ribosome through experiments. However, how metal ion regulate the conformational dynamics of the ribosome remains elusive. Meanwhile, the differentiation of monovalent and divalent ions in stabilizing the ribosome structure is less than well understood.

### 0.2.1. The configurations of ribosome.

Ribosome is a macromolecular machine that performs protein synthesis in living cells. It is composed of three large RNA molecules, 100 $\sim$ 2800 residues of each, and approximately 50 proteins. Each ribosome consists of two subunits that cooperate with each other as well as the movement of tRNA through sequential conformational changes (Fig. 0.1) to translate the codons on messenger RNA (mNRA) into the order of amino acids on the synthesized protein. In this process, the small subunit of the ribosome mainly performs the decoding function from the mRNA, while the large subunit mainly performs the catalytic function and link the amino acid together to produce protein. Each amino acid is delivered by one tRNA. Both subunits of ribosome have three binding sites (A: aminoacyl, P: peptidyl, E: exit) for tRNA, and tRNA sequentially traverse through the A, P, E site in each elongation cycle of the ribosome to bring the amino acid to the polypeptide chain. In the process of translating mRNA sequence into the order of amino acid in the peptide, the ribosome and tRNA will undergo the aa-tRNA selection, aa-tRNA accommodation, tRNA translocation and the E-site tRNA release to add one amino acid to the nascent peptide and prepare for the next cycle (Fig. 0.1).

### 0.2.2. The role of metal ions in ribosome.

It has been known for decades that the structure and function of the large-scale RNP assemblies, such as ribosome, are strongly influenced by the presence of metal

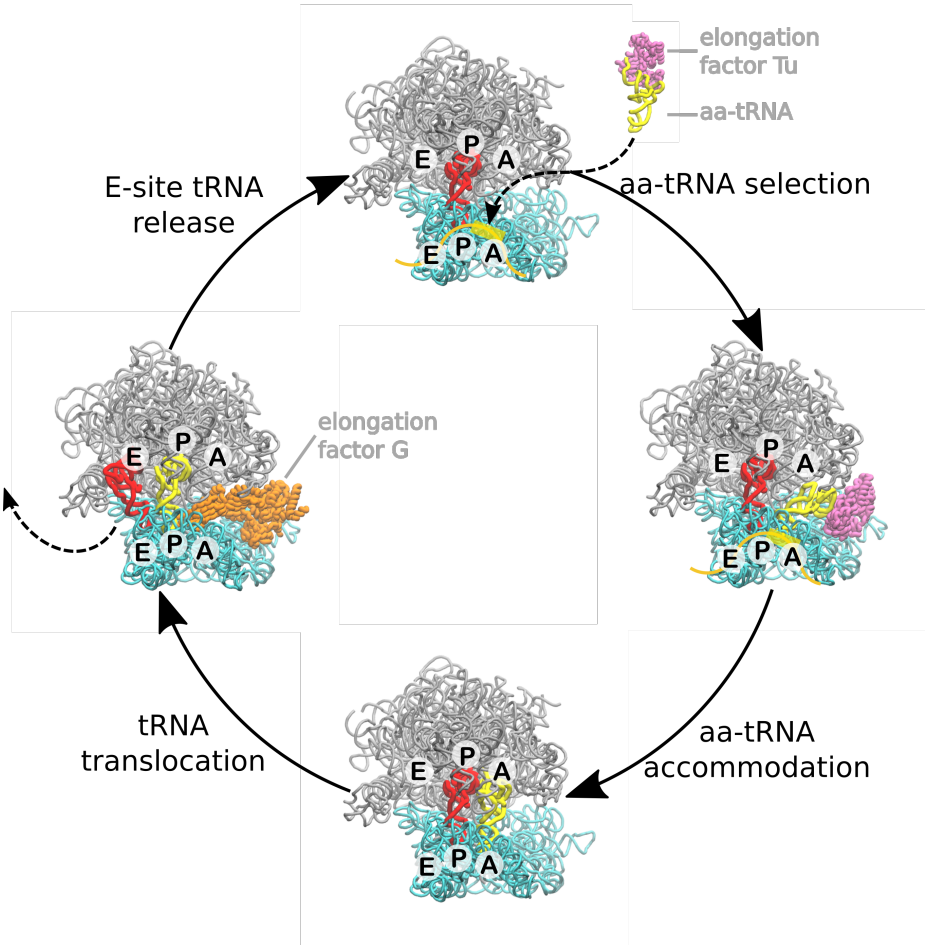


**Figure 0.1: Elongation cycle by the ribosome.** During the process of translating mRNA sequence into polypeptide, the large (grey) and small (cyan) subunits of ribosome and tRNAs (red, yellow) undergoes a series of large-scale conformational rearrangements. At the beginning of each elongation cycle, an aa-tRNA is delivered to the ribosome by EF-Tu (pink) and the anti-codon on the tRNA will pair with the codon on the mRNA through the *tRNA selection* process, so that the tRNA sits at the A-site in the small subunit. Then the aa-tRNA undergoes a $\sim$ 100 Å conformational rearrangement, called *accommodation*, so that the CCA-end of the tRNA binds to the A-site in the large subunit, which corresponds to an A/A-P/P configuration of the ribosome. Then the incoming amino acid is added to the nascent peptide chain. Following the peptide bond formation, both tRNAs will translocate between bindings sites facilitated by the EF-G, where the A-site tRNA moves to P-site and the P-site tRNA moves to E-site along with the displacement of mRNA to expose the next codon. The E-site tRNA will then be released and the ribosome is back to the initial configuration of the next cycle.

ions, such as $Mg^{2+}$ and $K^{+}$ ions [48–53]. For example, in terms of the biological function, *in vitro* studies have shown that how difference in solvent/ionic conditions can dramatically affect the accuracy of the protein synthesis by ribosome [54, 55]. In terms of the ribosome structure, the *in vitro* association of the small and large ribosomal subunits to form intact ribosome was found strongly dependent on the $Mg^{2+}$ concentration [56–58]. Since there are 3 large RNA molecules in the ribosome and the entire ribosome structure carries $\sim 3000$ negative charges, the chelated metal ions were also found significant to the structure stability of the ribosome, which can attenuate the repulsions between the like charged components in the ribosome and maintain its compact configuration. Using X-ray diffraction method, a decent amount of metal ions were recognized in the ribosome structure [50, 59].

To understand the conformational dynamics of ribosome as well as the related structural or environmental factor, much efforts has also been made through computational methods and simulations. However, MD simulations of the entire ribosome still remain challenging for two main reasons: (1) the large size of the ribosome (each ribosome has $\sim 150{,}000$ non-hydrogen atoms) and (2) the long time-scales relevant to the functions of ribosome (e.g. milliseconds or longer simulations might be needed). To this end, the whole arsenal of MD simulation strategies are used and multiple variants are developed on this basis [60]. For example, coarse-grained MD simulations of the entire ribosome were performed to investigate the subunit rotation in the ribosome during the tRNA translocation [61] and the interactions of the $\alpha$-synuclein with the ribosomal surface [62]. Structure-based all-atom MD simulations have been conducted to study the energetic properties of the aa-tRNA accommodation process [63] and the tilting motion in the ribosome during mRNA-tRNA translocation [64]. With the solvent molecules included, explicit solvent simulations of the entire ribosome [65–68] or its reduced/cutout model [68, 69] were performed to investigate the interactions and energetics of the ribosome with higher degree of details. With these existing

models described above, using MD simulations to study the ionic effects of the full ribosome structure is still too expensive for long time-scale dynamics.

In order to study the ionic effects on the full ribosome structure with highly efficient while less expensive models than the explicit solvent simulations, a structure-based model with explicit treatment of diffuse ions while implicit treatment of solvent is introduced in Chapter 2. This model is constructed based on energy landscape principles, which also accounts for explicit Coulomb interactions, implicit solvation and desolvation barriers as well as the interactions between diffuse ions with each other and highly electronegative atoms in RNA and protein. The model parameters are refined using both explicit solvent simulations and experimental measures as a reference. The refined parameters are able to describe the concentration dependent balance between monovalent and divalent ions in the excess ion atmosphere of RNA fragments, which is in agreement with fluorescence titration results. This model was then applied to the 70S bacterial ribosome. The simulation revealed how diffuse ions regulates conformational rearrangement in the extended L1 stalk region and facilitates the interactions between L1 stalk and the E-site tRNA. This model also motivated us to further investigate the role of ions in the functional rearrangements and conformational dynamics in the ribosome structure and other RNP assemblies.

## 0.3. Relaxation properties of Glass-forming liquids

### 0.3.1. Glass-forming liquids.

Glasses are disordered solids, which have liquid-like structure and therefore lack the periodicity of crystals at the molecular level [70]. They can be made through a variety of processes, while the most common route of making a glass is by cooling a liquid sufficiently fast so that it does not have time to crystallize and the supercooled liquid transit into the solid state at the glass transition temperature ($T_g$). The atomic-scale

structures of glasses share the characteristics of the supercooled liquids, even though they behave as solids mechanically.

The glass transition is similar to a second-order phase transition in the Ehrenfest sense with continuity of molar volume and molar enthalpy, but discontinuous changes in their derivatives, such as the thermal expansion coefficient and the specific heat [71]. But the transition is continuous and cooling-rate dependent [72]. The liquid-glass transition is not a transition between equilibrium states. Instead, on cooling of a liquid, the internal degrees of freedom successively fall out of equilibrium. Therefore, the term relaxation time ($\tau$) is used to describe how fast a macroscopic stress relaxes and also determines the the typical time between molecular displacements or reorientations.

Glass-forming liquids have some unique characteristics, the major properties of which can be summarized in to three non's [72]:

- The relaxation is almost *nonexponential* [73, 74]. That is, the relaxation function ($M_p$) of property $p$, defined by

$$M_p \equiv \frac{p(t,T) - p(\infty,T)}{p(0,T) - p(\infty,T)}, \tag{0.1}$$

  cannot be accurately represented by an equation of the form

$$M_p = \exp(-t/\tau_p), \tag{0.2}$$

  where $t$ is the time following a jump to temperature $T$. Experimentally, the relaxation function can be expressed by the stretched exponential (Kohlrausch–Williams–Watts (KWW) function [75])

$$M_p = \exp[-(t/\tau_p)^\beta], \tag{0.3}$$

where the exponent $\beta$ has a value between 0 and 1.

- The relaxation time $\tau$ or the viscosity $\eta$ show the *non-Arrhenius* temperature dependence. The non-Arrhenius temperature dependence of the relaxation time is often described by the Vogel-Fulcher-Tammann (VFT) relation [76–78]

$$\tau = \tau_0 \exp\left(\frac{A}{T - T_0}\right) \tag{0.4}$$

where temperature $T_0$ is a characteristics temperature greater than zero.

- The third *non* is the *nonlinearity* of the relaxation state [79–81]. This refers to the finding that near and below $T_g$, relaxation can be studied on systems which are nonergodic and are evolving toward the equilibrium structure on very long time scales.

### 0.3.2. Adam–Gibbs model.

For the glass-forming liquids, Adam and Gibbs proposed a molecular-kinetic theory that explains the temperature dependence of the relaxation behavior in terms of the temperature variation of the size of the cooperatively rearranging region (CRR) [82]. In the Adam–Gibbs model, it is assumed that relaxation can occur if at least a minimum number, $z^*$ of molecules is assembles in CRR. The critical size $z^*$ of the CRR can be expressed in terms of the configurational entropy,

$$z^* = s^* N_A / S_c, \tag{0.5}$$

where $N_A$ is the Avogadro constant, $S_c$ is the configurational entropy of the supercooled liquids, and $s^*$ is the critical configurational entropy corresponding to the lower limit of the size of the representative CRR. Since the relaxation time is reciprocally

related to the transition probability:

$$\tau(T) \propto 1/\bar{W}(T), \tag{0.6}$$

and the average transition probability ($\bar{W}$) can be expressed with

$$\bar{W}(T) = \bar{A}\exp(-z^*\Delta\mu/kT), \tag{0.7}$$

where $\bar{A}$ is a frequency factor, which is approximately independent to temperature, the Adam–Gibbs model predicts a linear relationship between $\log(\tau)$ and $(TS_c)^{-1}$. There have been various generalization of Adam–Gibbs model [83–86].

On the basis of the Adam–Gibbs model, Chapter 3 introduces the theoretical investigation on the fluctuation effects in the Adam–Gibbs model of cooperative relaxation. This work studied the configurational fraction ($f(T)$) at the glass-transition temperature ($T_g$) with and without fluctuation effects in the Adam–Gibbs model, and a connection between the $\beta$ parameter in KWW function and $f(T_g)$ was observed.

## 0.4. The effects of diffuse ions on the aa-tRNA accommodation in the ribosome.

In the elongation cycle of the ribosome (Fig. 0.1), the delivery of the aminoacyl-tRNA(aa-tRNA) is noted as tRNA selection process, which includes two steps: initial selection and proofreading. During the initial section, an aa-tRNA is delivered to the ribosome by EF-Tu. In this step, the anti-codon region in the tRNA pairs with the codon in the mRNA, which positioned in the A-site of the small subunit of the ribosome. After the initial selection, since the tRNA interacts with the ribosomal A site and EF-Tu simultaneously, the aa-tRNA is referred to as an A/T configuration. The formation of the codon-anticodon interactions triggers GTPase activation and

finally Guanosine-5'-triphosphate (GTP) hydrolysis [87]. Upon the GTP hydrolysis, the conformation of EF-Tu changes drastically and dissociates from the aa-tRNA and the ribosome complex. Then the aa-tRNA will go through a large-scale ($\sim$ 100 Å) conformational rearrangement to bind to the A-site of the large subunit of the ribosome. This process is referred as aa-tRNA accommodation (the proofreading step). After the accommodation process, the aa-tRNA ends up with A/A configuration, since the tRNA binds to A-sites in both small and large units.

The accommodation process of aa-tRNA was reported as a multistate mechanism through computational investigation [88] using MD simulations, cryogenic electron microscopy (cryo-EM) reconstruction of the ribosome structure [89], and the single-molecule fluorescence resonance energy transfer (smFRET) characterization of the transition between different states. The multistate mechanism (Fig. 0.2) of the aa-tRNA accommodation can be summarized in to the transition between the (1) A/T ensemble (defined above), (2) elbow accommodated (EA) ensemble, when the elbow region of the tRNA is in an accommodated position while the 3'-CCA end is displaced from the A-site in the large subunit, (3) the arm accommodated state, when both the elbow and the acceptor arm are in accommodated position while the 3'-CCA end does not reach the peptidyl transderase center (PTC), and (4) the fully accommodated (A/A) state. In Chapter 4 of this thesis, the discussion is focused on the transition between the A/T and EA states.

For the elbow accommodation process of the aa-tRNA in the ribosome, numerous computational studies [63, 88, 90, 91] have made careful investigation about the dynamics of the aa-tRNA and the EF-Tu during the elbow accommodation, as well as the role of EF-Tu in the proofreading of the aa-tRNA by the ribosome. Experimental studies [54, 55] found that the concentration of $Mg^{2+}$ ions was able to affect the speed of GTP hydrolysis as well as the speed and accuracy of the tRNA selection. However, the mechanism of how $Mg^{2+}$ ions affect the tRNA accommodation is unclear and

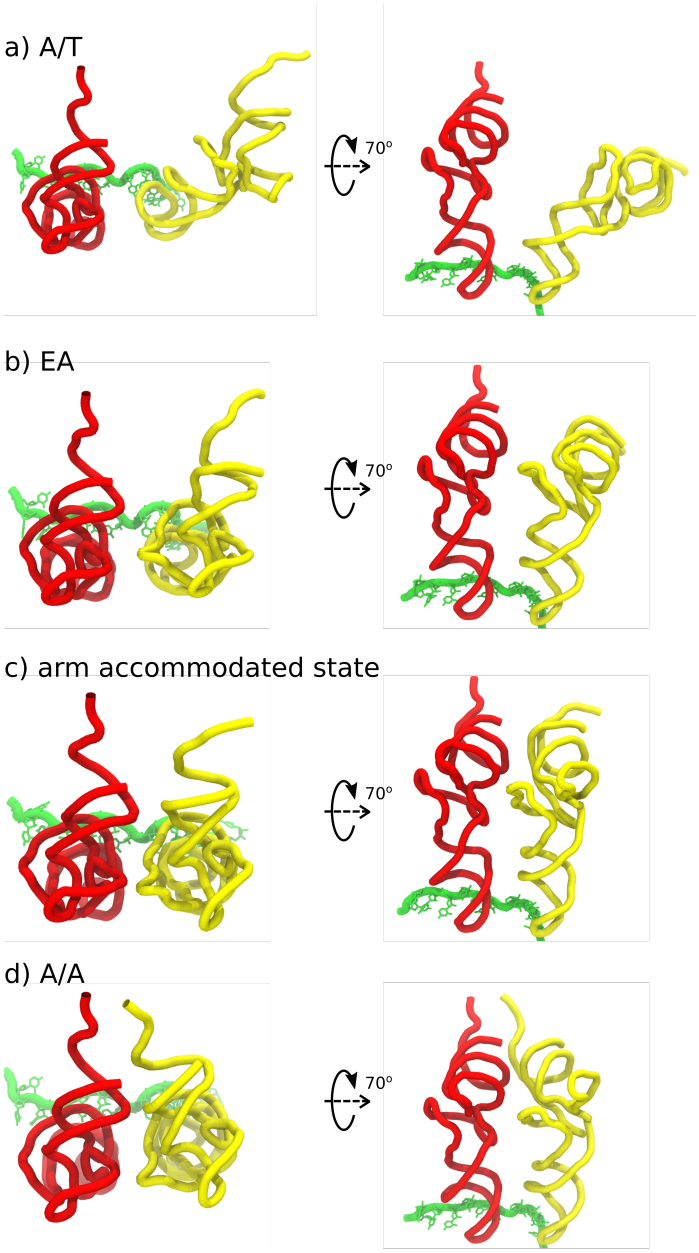


**Figure 0.2: Configurational states during the aa-tRNA accommodation process.** See text for detailed definitions of each state.

require further investigation from both experimental and computational perspective. In Chapter 4, the newly designed model introduced in Chapter 2 will be applied on the full ribosome structure to investigate the role of metal ions in the configurational transition between A/T and EA states during the aa-tRNA accommodation. These simulations also shed light on the differentiation of monovalent and divalent metal ions and their contributions to the elbow accommodation process.

# Chapter 1.

# RNA as a Complex Polymer with Coupled Dynamics of Ions and Water in the Outer Solvation Sphere

*Reproduced with permission from Lammert, H., Wang, A., Mohanty, U. and Onuchic, J.N., 2018. "RNA as a complex polymer with coupled dynamics of ions and water in the outer solvation sphere." *The Journal of Physical Chemistry B*, 122(49), pp.11218-11227.

## 1.1. Introduction

Recent discoveries about gene regulation by flexible mRNAs and the importance of noncoding RNAs have renewed interest in the structure and dynamics of RNA. RNA as almost the same chemical composition as DNA. Due to the presence of hydroxyl group in the ribose ring, the structure and dynamics of RNA is more complex than DNA. The multifunctional nature of RNA polymers allows it to be involved in diverse aspects of cellular metabolism from transcription to pre-mRNA splicing, translation, RNA-protein transport, and RNA decay [1, 2, 8, 92–94]. Most functional RNA molecules are tightly folded into a unique conformations adopted by biopolymer ina cell and in different cell states are crucial for its function [94].

RNA is a highly flexible charged polymer with a rugged landscape, for which formation of tertiary structures is dependent on its sequence, metal ion identity and concentration, solvent, the presence of proteins, and various cellular signals [1, 2, 8, 12–20, 88, 92–97]. Analysis of lifetimes of motional modes underscores the RNA's ability to readily access a heterogeneous ensemble of conformations in response to external perturbations [98]. These motions occur on time scales on the order of a few nanoseconds. Because of the strong interactions between counterions and co-ions and the RNA, the plasticity of RNA, i.e., RNA's ability to rapidly change its heterogeneous ensemble of conformations, interprets into a wide range of spatial and temporal dynamics of the metal ions and solvating water that reflect the underlying architecture of the molecular surface [12–22, 88, 98]. By stabilizing structures that are inherently compact, RNA exhibits sluggish dynamics with an increase of $Mg^{2+}$ concentration. $Mg^{2+}$ can displace other larger metal cations in the diffuse ion atmosphere around RNA. The basis of how divalent and monovalent metal ions affect RNA stability, dynamics, and therefore function is less well understood.

In this study, we investigate the dynamics of water and divalent and monovalent metal ions in the vicinity of a folded 58-nucleotide ribosomal rRNA fragment [10,

11] through atomic resolution explicit solvent MD simulations with five independent runs, each over 180 ns. Due to its small size, the rRNA fragment si ideally suited for MD simulations to address basic questions on spatial and temporal dynamics of the metal ions and solvating water. We determine the excess ion atmosphere around the rRNA and compare it with experimental data and predictions on the basis of a strucsre-based model (SBM) generalization of Manning counterion condensation in which $Mg^{2+}$ ions are treated explicitly while $K^{+}$ ions are described implicitly. We measure the residence time of ions and water molecules in association sites in the outer sphere. We undertake a joint PCA for the coordinates fo the RNA phosphorus atoms and for the occupation counts of the various association sites in the outer sphere to investigate the coupling of metal ion and water dynamics to motions of the rRNA. We determine stabilizing interactions and coupled dynamics for all solvent species, which augment the rRNA's ability for rapid conformational changes. We demonstrate that the 58-mer rRNA in solution forms a complex polymer composed of RNA that is sheathed by a fluctuating network of associated counterions, co-ions, and discrete water molecules.

## 1.2. Methods

**Evaluation of RNA Structure and Magnesium Binding.** The ribosomal 58-mer RNA studied her natively folds into a compact knob on the surface of the 23s subunit, near the binding site of elongation factor EF-Tu and EF-G, with a structure that is further stabilized by binding to ribosomal protein L11 [10]. In solution in the presence of counterions, the 58-mer can adopt a native like compact structure, shown in Fig. 1.1A. (A diagram of the secondary structure and a contact map are available as Appendix Figure A.1 and A.2) Compact RNA structures depend for their stability in solution on associated counterions. Cations become enriched near the

RNA over their bulk concentrations, while anions are locally depleted. The effect is quantified for ions of species $i$ by preferential interaction coefficients $\Gamma_i$, which depend both on the identity of the RNA and on the bulk concentrations of the ions. The spatial distribution of ions near the RNA is heterogeneous and dynamic. Modeling of this local ion atmosphere is the primary goal of our simulations. For comparison, experimental information about the total number of associated ions is available from titration experiments [42], which, however, can not distinguish the different possible regimes of association, including direct inner-sphere contacts between ions and RNA, outer-sphere contacts mediated by a single hydration layer, and diffuse association at longer distances [2, 47]. This information matters, because the different modes of association especially for magnesium ions are not all equally well described by available simulation force fields[99].

Inner-sphere and outer-sphere contacts can best be identified from X-ray structures. The crystal structure of the 58-mer, PDB 1HC8, contains two copies of the RNA, labeled as chains C and D. Apart from the chelated $Mg^{2+}$ and $K^+$ there are 14 distinct positions for metal ions in the two structures, which are shown in Fig. 1.1B. Of these, 12 are occupied by magnesium ions and two by osmium ions, which substitute hexahydrated magnesium. Both the osmium positions and 7 of the 12 magnesium positions are occupied in both copies of the RNA, while 2 magnesium ions are only present in chain C and 3 more only in chain D. (Ion positions that are unique to one copy are labeled with an asterisk below.)

Only a few of the magnesium ions in the X-ray structure make contacts with RNA atoms that are within 2.3 Å, close enough for inner-sphere binding [10]. Apart from the chelated $Mg^{2+}$ **C1167** and the ion **C1165***, which is artificially coordinated by both chains, magnesium **C1163** has one close oxygen neighbor, which it shares with the chelated $K^+$. Single inner-sphere oxygens also occur for magnesium ion **C1161**, for **C1166** near the triphosphate, and for **D1365***. Only **C1161/D1361**, however,

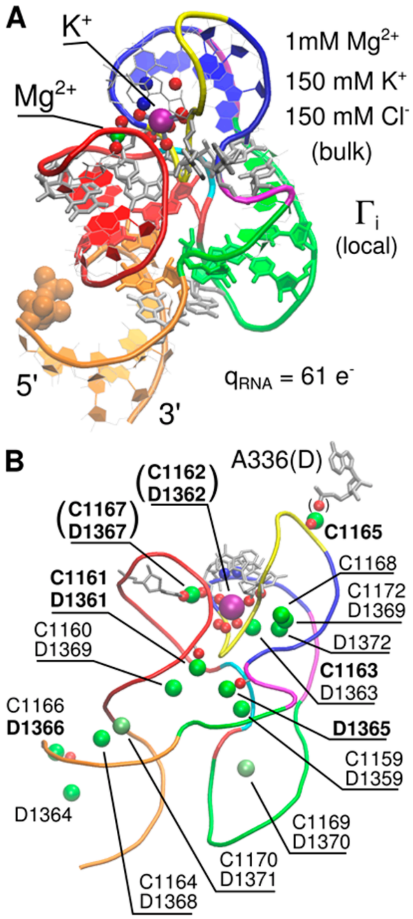


**Figure 1.1: 58-mer RNA system.** (A) X-ray structure of the 58-mer rRNA. Helices A, B , and C are drawn in blue, orange, and red, connecting elements in magenta, yellow, and cyan. Bases forming tertiary contacts are shown in gray. One chelated $Mg^{2+}$ and $K^+$ each are shown in green and purple, together with coordinating RNA oxygen and nitrogen atoms, colored read and blue. The triphosphate group at the 5' end is also highlighted. (B) Metal ion positions from the two copies of the 58-mer in the X-ray structure (PDB:1HC8). Ions are labeled with the chain identifier and residue number from the PDB. The two chelated $Mg^{2+}$ and $K^+$ (purple) ions are in brackets. Bold labels highlight ions with inner-sphere RNA contacts. Coordinating RNA oxygen and nitrogen atoms forming inner-sphere or chelating contacts are shown in red and blue. Also shown is base A336 from chain S, forming an outer-sphere crystal contact with ion C1165. All Structures in all figures were drawn with VMD [100].

forms an inner-sphere contact in both copies of the RNA in the crystal. The ion **D1366** forms an inner-sphere contact only in chain D, and the position of **D1365*** is not occupied at all in chain C. Most of the magnesium ions in the crystal only form outer-sphere contacts. Given that neither all of the inner-sphere contacts nor all of the participating ions are present in both copies of the RNA, the observed inner-sphere contacts do not appear to confer significant additional stability. This view is further supported by thermodynamic calculations of magnesium binding energies in the 58-mer by Misra and Draper [43]: While direct contacts between magnesium and phosphate oxygens are very favorable, this stabilization is generally offset by the high cost of dehydration. Therefore, inner-sphere binding of magnesium to RNA is not generally dominant but occurs only in particularly favorable locations. In the 58-mer, only the chelated ion that is present in the X-ray structure as **C1167** is unequivocally stable. The chelated $K^+$ outcompetes $Mg^{2+}$ in its site for geometrical reasons [10]. Magnesium **C1163**, which is buried on the other side of the chelated $K^+$, has similar but already marginal stability compared to the chelated $Mg^{2+}$. For all other magnesium sites, the calculated stability is lower by an order of magnitude, and persistent inner-sphere binding is not predicted. Any intermittent inner-sphere contacts are not expected to change the character of the interactions between RNA and associated ions [43].

Interactions for $Mg^{2+}$ are more complicated to treat in simulations than for other cations like $Na^+$ or $K^+$. Due to the small size and high charge of $Mg^{2+}$ polarization and charge transfer make important contributions beyond simple electrostatics to the interactions with inner-sphere ligands [99]. Polarizable force fields may well be necessary to predict reversible site-binding of $Mg^{2+}$ to RNA in changing configurations, e.g., during folding or other large-scale dynamics. Meanwhile, point charge force fields for $Mg^{2+}$ are also under active development. Recent parametrizations have aimed to improve the time scale for hydration exchange, which is 1.5 $\mu$s experimentally [101]

but is overestimated by 2 orders of magnitude in a common older force field [36, 102, 103]. However, comparative studies have shown that old and new point charge force fields are equally adequate to describe diffuse association and outer-sphere interactions of hexa-hydrated $Mg^{2+}$ with RNA [103, 104], which are in fact dominated by electrostatics. Since we are simulating an RNA in the native folded state, where the X-ray structure indicates that outer-sphere binding dominates the behavior of the associated $Mg^{2+}$ a point charge model is sufficient for our purpose. The time scale of several $\mu$s, required to equilibrate hydration exchange by $Mg^{2+}$ is beyond our simulation time, and we are neglecting the small contribution from intermittent inner-sphere contacts formed by the associated $Mg^{2+}$. For consistency with earlier work, we are using the $Mg^{2+}$ parameters by Åqvist [36], which overestimate the time scale of hydration exchange. Thus,hydration of $Mg^{2+}$ is effectively permanent in our simulations. The single chelated $Mg^{2+}$ and $K^{+}$, observed in the crystal structure, are included in our simulations, and they remain stable in their native positions without special treatment.

**Simulation Protocol.** System preparation and equilibration follow a well-established protocol, described in refs [47, 105], and [106]. Simulations were run using the GROMACS5 software [107] and the AMBER99 force field [108] with extensions parmbsc0 [109] and chiOL3 [110]. Parameters for the triphosphate group at the 5' end of the RNA were obtained from ref [111]. SPC/E water [112] was used and modified $K^{+}$ parameters [113] in order to avoid crystallization. We retain the $Mg^{2+}$ parameters [36] used in ref [47] for consistency. As discussed above, it is suitable for the treatment of the hydrated $Mg^{2+}$, expected in the system. The RNA portion of the updated force field is well tested, and while quantitative difference might arise, we expect that the major results should also hold with other recent force field.

Molecular dynamics simulations were propagated with a time step of 2 fs, using particle-mesh Ewald electrostatics [114] with an Ewald radius of 15 Å and also a van

der Waals cutoff of 15 Å. Temperature and pressure were maintained with a Berendsen thermostat and barostat, and bond lengths for hydrogen atoms were constrained using LINCS. A cubic box with an initial size of 93 Å was used. The full system contained 76,600 atoms. Ions were included at target concentrations of 150 mM KCl and 0.5 mM $MgCl_2$.

The RNA's starting configuration was taken from the crystal structure (PDB 1HC8) [10] and was initially kept frozen. One $Mg^{2+}$ and one $K^+$ were placed in their respective chelation sites [10, 43, 115] as part of the RNA structure. These ions remained in their chelation sites without special treatment. Additional ions were added randomly into the simulation box, using increased ionic radii to ensure a spacing between ions and RNA that is compatible with hydration. Initial ion positions were pre-equilibrated before addition of explicit water over 10 ns of stochastic dynamics simulations with a 5 fs time step and a dielectric constant of 80. Only then was explicit water added to the box, and equilibrated with frozen RNA and ion positions for 1.2 ns, while increasing the temperature to 300 K over the first 200 ps. Next the ions were freed and equilibrated with the water at constant volume for another 2 ns. Finally, the RNA was gradually released, applying position restraints of decreasing strength (1000, 100, 10, 1 kcal/mol/$nm^2$) in steps of 2 ns at constant pressure. Another 22 ns of unrestrained equilibration was added, before production simulations were continued for data collection over 180 ns. Five independent runs were performed according to this protocol.

**Preferential Interaction Coefficients.** Preferential interaction coefficients in our explicit solvent simulations were computed using the method described in ref [99]. For this purpose, the portion of the box at more than 20 Å from any RNA atom is considered to represent the bulk solution. The time averaged ratio between the number of bulk ions, e.g. $B_{Mg}$, and the number $B_{wat}$ of bulk water molecules in this region is multiplied with the molarity of pure water to yield the bulk concentration

of ions, assuming that the volume effect of the solute ions is negligible.

$$[\mathrm{Mg}]^* = 55.51\mathrm{M}\langle B_{Mg}/B_{wat}\rangle \tag{1.1}$$

A correction is applied to these raw concentrations in order to account for the remaining deviation from charge neutrality in the solution at this distance from the RNA, assuming a linear effect on ion densities of the weak remaining potential.

$$[i] = [i]^*(1 - q_i \sum_j q_j[j]^* / \sum_j q_j^2[j]^*) \tag{1.2}$$

On the basis of the corrected, electroneutral bulk ion concentrations, preferential interaction coefficients are identified as the excess number of ions present in the system, from the numbers $N_i$ and $N_{wat}$ of ions and water near the RNA.

$$\Gamma_i = N_i - N_{wat}[i]/55.51\mathrm{M} \tag{1.3}$$

**Structure-Based Model Simulations for Preferential Interaction Coefficients.** Additional molecular dynamic simulations of ion association were performed with an augmented structure-based model (SBM), developed by Hayes *et al* [45]. In this model, $Mg^{2+}$ ions are treated explicitly, while potassium ions are described implicitly by a generalization of Manning counterion condensation theory. The SBM simulations were performed using Gromacs-4.6.7. A fixed number of 250 $Mg^{2+}$ ions were explicitly included in the system. We have adjusted the effective concentration of the $Mg^{2+}$ ions by changing the size of the cubic simulation boxes in the range from 69 to 161 nm. Each $Mg^{2+}$ ion was initially placed randomly in the box but not too closely to the RNA. In addition, we chose a KCl concentration of 150 mM for direct comparison with previous experimental results by Draper [11]. Besides the $Mg^{2+}$ and $K^+$ ions in the environment, there are one $Mg^{2+}$ and one $K^+$ ion chelated in the 58-mer RNA fragment. The chelated ions were added via the SMOG2 [116] suite

of programs. The chelated $Mg^{2+}$ was harmonically bound to two oxygen atoms in the chelation pocket, while the chelated $K^+$ ion was bound to six neighboring oxygen atoms. Each production simulation was run for $2 \times 108$ time steps of 0.002 ps. Excess ions were calculated by the difference between the actual number of $Mg^{2+}$ ions within 4.5 nm of the RNA and the expected number, which can be calculated by multiplying the bulk concentration with the volume of this region. Bulk concentration was here taken to be the average $Mg^{2+}$ concentration in the region more than 4.5 nm away from the RNA.

**Coordination and Hydrogen Bonds.** Interactions between RNA and solvent were investigated both for magnesium and potassium cations and for the solvent water itself. Coordination bonds between RNA and solvent cations were defined by distance, using cutoff values of $d_{\mathrm{K}} = 3.65$ Å and $d_{\mathrm{Mg}} = 5.0$ Å which encompass the first peaks in the K-RNA and Mg-RNA pair distribution functions, shown below. For water, the analysis was focused on hydrogen bond (HB) formed with the RNA. Hydrogen bonds between water and RNA were identified by geometrical criteria, requiring a distance between the donor (D) and acceptor (A) heavy atoms of $d_{\mathrm{HB}} < 3.5$ Å and an angle of $\theta_{\mathrm{HB}} < 30°$ between the D−H and H−A vectors [117, 118]. It was assumed that every water could form at most four hydrogen bonds, two as donor and two as acceptor. For each of the possible donor bonds, the bond was assigned to that RNA acceptor atom, satisfying the criteria, with the smallest angle $\theta_{\mathrm{HB}}$ and within 1 Å of the closest one. For each of the two possible acceptor bonds, the set of potential RNA donors was also restricted to atoms which were placed on one side of the HOH plane and on the side of the water oxygen facing away from the hydrogens. Water−water HBs were not explicitly considered; i.e., no global optimization of the HB network was carried out. However, in practice, the identified water−RNA HBs sterically exclude competing water−water HBs for the same water.

**Association Sites.** Ion−RNA coordination bonds and water−RNA HBs were

analyzed over the course of the simulation runs, and the total number of bonds formed by every solvent particle at any given time was also tracked. Locations where individual ions or waters formed multiple simultaneous bonds with the RNA were identified, and average positions of the coordinated solvent particle were obtained for each set of ligated RNA atoms. All positions that were populated during at least 1% of the total simulation time were retained for further analysis. These positions trace a network of favorable locations for ions and water molecules, coordinating the RNA. Maximum linkage clustering with a 2 Å cutoff was used to define distinct association sites on the basis of the raw positions of coordinated water and magnesium and potassium ions.

**Residence Times.** Residence times in association sites were determined using an indicator function, $o_{s,i}(t)$, similarly to the positional correlations defined for the ions alone. Here, the reference is the position of the site, and $o_{s,i}(t)$ is defined to equal 1 if ion $i$ is within 2 Å of site $s$ at time $t$ and 0 otherwise. Correlations for $o_{s,i}(t)$ were calculated again without subtracting the mean and were averaged over all times and all ions for each site $s$. The correlation time, identified with the residence time in site $s$, is determined as the time where the correlation function equals 1/e. This approach has been used to define residence times also in ref [47], where a cutoff of 4 Å was chosen. Here we have used 2 Å to match the size of our association sites, as determined by the clustering algorithm used in their identification.

Residence times defined in this way also permit excursions of the particle from the site, as long as it stays close enough to return with increased probability. For comparison, dwell times were also determined, which are defined here as the average duration of the interval that a particle spends strictly within 2 Å of the site. These dwell times are correlated with the residence times but much shorter, especially for sites with long residence times.

**Principal Component Analysis.** Collective dynamics of RNA and ions and

their time scales were analyzed using principal component analysis (PCA) [119–121]. Starting from trajectory data for a set of coordinates $a_i$, which are, e.g., the Cartesian coordinates for atoms (of selected types) in the system, PCA identifies an orthogonal basis formed by linear combinations of the underlying coordinates $a_i$. Coordinates in the new basis are eigenvectors $v^{(n)}$ of the covariance matrix, $C_{ij} = \langle \sqrt{m_i}(a_i(t) - \langle a_i \rangle)\sqrt{m_j}(a_j(t) - \langle a_j \rangle)\rangle$. The corresponding eigenvalues $\lambda(n)$ give the mean variance exhibited by the dynamics along each eigenvector. As PCA maximizes the fraction of the total remaining variance that is captured by each additional eigenvector, a small number of the principle component (PC) with the highest eigenvalues gives a good overview over the collective dynamics of the system.

We have used PCA of the RNA coordinates to determine time scales of global RNA motions, which were extracted as autocorrelation times after projection of the dynamics onto the first several eigenvectors. The correlation times were identified with the time when the normalized correlation function decreases to a value of 1/e.

In order to study the coupling of ion dynamics to motions of the RNA, we have performed a joint PCA, simultaneously for the Cartesian coordinates $x^P$ of the RNA phosphorus atoms and for the occupation counts $n_i^{\mathrm{Mg}}$, $n_i^{\mathrm{K}}$, and $n_i^{\mathrm{W}}$ of the association sites, identified near the RNA for magnesium, potassium, and water. Prior to this analysis, each of the four types of variables was separately rescaled, such that the mean variance in each type of variable becomes unity, e.g., $\tilde{n}_i^{\mathrm{K}}(t) = n_i^{\mathrm{K}}(t)/V^{\mathrm{K}} = n_i^{\mathrm{K}}(t)/\langle \mathrm{VAR}(n_i^{\mathrm{K}}(t))\rangle_i$. Note that the relationship between different variables $n_i^{\mathrm{K}}$, $n_j^{\mathrm{K}}$ in the same category is not affected by this proportional scaling of the entire category. The scaling ensures, however, that all types of variables contribute equally to the total variance of the joint coordinates that are subjected to the PCA. When the results are interpreted, the actual changes implied by the changes in the scaled variables are calculated, on the basis of the known scaling factors. In order to maximize sampling, all five available runs were analyzed together by forming an aggregate covariance

matrix.

## 1.3. Results and Discussion

**Preferential Interaction Coefficients.** Experimental values of preferential interaction coefficients for $Mg^{2+}$ with the 58-mer RNA were taken from ref [11]. For a system with 150 mM $K^+$ and 1 mM $Mg^{2+}$ the reported $\Gamma_{2+}$ is 10.5, and it is 8.2 at 0.5 mM $Mg^{2+}$. Experimental values for $\Gamma_{2+}$ over a range of $Mg^{2+}$ concentrations from 0.3 to 1.0 mM are replotted in Figure 1.2 together with values from our simulations. The results from the augmented SBM closely match experimental values at the high end of the range. At lower [$Mg^{2+}$], the SBM overestimates $Mg^{2+}$ association with the RNA. For the explicit solvent simulations, the effective concentration of $Mg^{2+}$ is 0.6 $\pm$ 0.2 mM. Concentration and preferential interaction coefficients from the MD simulations are listed in Appendix table A.1. The considerable statistical error in the bulk concentration of $Mg^{2+}$ is due to the small number of ions found in the bulk of even the largest accessible simulation box. The excess number of $Mg^{2+}$, associated with the 58-mer RNA in the simulations, is 10.8 including the chelated $Mg^{2+}$ or 9.8 excluding it. The observed degree of agreement between simulations and experiment indicates that the models are suitable to treat ion association with RNA. The SBM achieves quantitative agreement for excess ion counts, using simple interactions for $Mg^{2+}$ without a dispersive term [122], but with parameters for the repulsive term that were directly optimized to reproduce association numbers. We rely on the more detailed explicit solvent simulations, which produce nearly equivalent association numbers, to study the spatial distributions of associated $Mg^{2+}$ and $K^+$, and also water.

**Ion Distributions.** Pair distribution functions, shown in Figure 1.3, summarize the average spatial relationships between ions and RNA. Enrichment of cations and

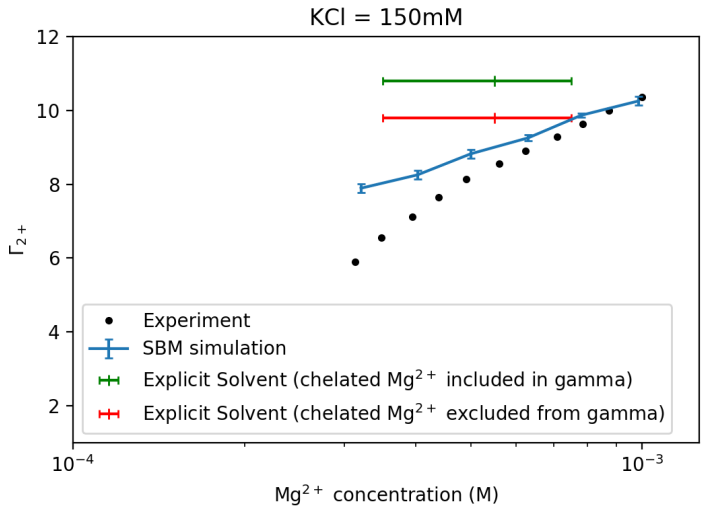


**Figure 1.2:** Preferential interaction coefficients. Experimental data for $\Gamma_{2+}$ are compared to values from the explicit solvent simulations and from simulations with an augmented structure-based model, including explicit $Mg^{2+}$.

depletion of chloride anions near the RNA are apparent in Figure 1.3A. A layer of coordinated water molecules is also visible in the distribution of RNA−oxygen distances. The distribution of $Mg^{2+}$ near the RNA further reflects the different regimes of association, starting with a sharp peak near 2 Å due to inner-sphere contacts, formed by chelated $Mg^{2+}$. The first main peak, corresponding to outer-sphere contacts formed by hexa-hydrated $Mg^{2+}$ via the first hydration layer, ends at 5 Å, where a broad shoulder reflects the enrichment of diffuse $Mg^{2+}$ ions. The first peak for $K^+$ occurs at shorter distances up to 3.65 Å indicating a first layer of $K^+$ that frequently forms inner-sphere contacts with the RNA. As stated above, we use the distances extracted from these distributions as cutoffs to determine contacts between $Mg^{2+}$ or $K^+$ and the RNA. The ion distributions for the unchelated potassium and magnesium ions in Figure 1.3B illustrate cation correlations, although only the K−K pair distribution is well resolved, with the first maximum at 4.5 Å. Both the Mg−Mg and Mg−K distributions are irregular due to the small number of $Mg^{2+}$ in the system. The Mg distribution does, however, show a clear first peak around 6.7 Å. While the distribu-

tion of $Mg^{2+}$ themselves reveals no clear pattern, association with the RNA follows regular preferences. Both $Mg^{2+}$ and $K^{+}$ form regular distributions around RNA P atoms, shown in Figure 1.3C, for nonchelated ions. The split first peak for $Mg^{2+}$ suggests, however, that other partner atoms steer cation association more directly. The distributions of hydrated $Mg^{2+}$ around its preferred immediate coordinating partners are shown in Figure 1.3D. Both phosphate oxygens and electronegative atoms provided by RNA bases coordinate hydrated $Mg^{2+}$ at distances around 4 Å.

**Correlated Dynamics of Ions and Water.** Without initially looking for specific binding sites, the coupling between RNA and solution environment should also be reflected in a slowing down of the dynamics of coordinating ions and water molecules and in correlations between their motions and those of the RNA.

Positional correlation functions, shown in Figure 1.4, quantify how long an unchelated solvent particle remains in a region within 4 Å from its initial position, depending on its initial proximity to the RNA. For $Mg^{2+}$ ions within 3.5 or 4.0 Å of the RNA, correlation times reach 2.4 ns. Some $K^{+}$ approach the RNA more closely than the hexa-hydrated $Mg^{2+}$. The correlation time for $K^{+}$ within 3.0 Å is 1.4 ns, decreasing to 0.8 ns at 3.5 Å and to 0.2 ns at 4.0 Å. Water shows a similar slowing down near the RNA, albeit with shorter correlation times. 0.2 ns is the maximum for initial distances of 3.0 Å, decreasing to 0.1 ns at 3.5 Å and 0.06 ns at 4.0 Å. The temporal scales spanned by water molecules and ions exhibit a broad spectrum due to inherent charge imbalance on the solvent accessible RNA surface [21]. To reveal dynamical features of water coordinating RNA, we have calculated, using a geometric criterion, the average number of hydrogen bonds formed with the 58-mer rRNA. We have determined the lifetimes of water and cation association with the RNA by identifying the hydrogen bonds between water and RNA and the coordination of $K^{+}$ and $Mg^{2+}$ by RNA oxygen atoms (Appendix A).

**Coordination Bonds.** The overall average binding times of metal ions and wa-

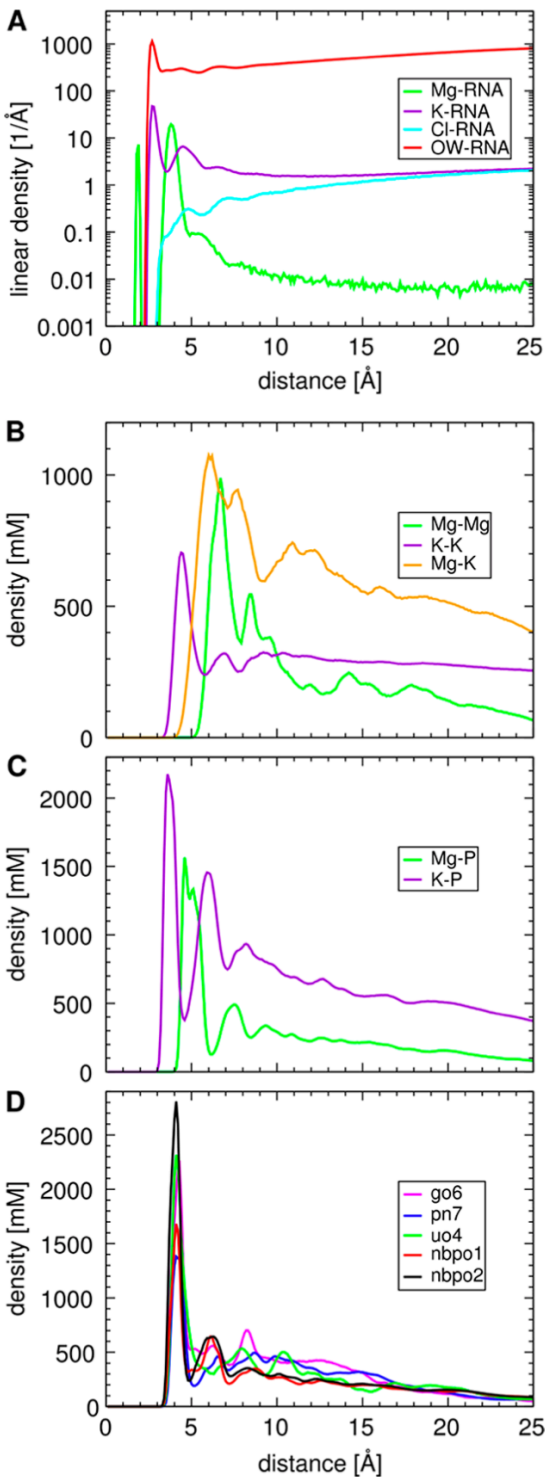


**Figure 1.3:** Pair distribution functions. (A) Distribution of ions and water oxygens near the RNA surface, measured from the closest heavy RNA atom. (B) Mutual pair distribution functions of $Mg^{2+}$ and $K^{+}$ cations. (C) Distributions of $Mg^{2+}$ and $K^{+}$ around RNA P atoms. (D) Distributions of $Mg^{2+}$ around electronegative RNA atoms: guanine O6, purine N7, uridine O4, and phosphate oxygens.

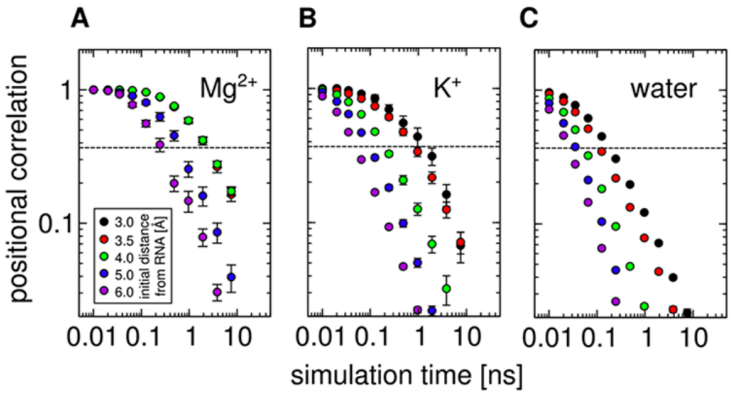


**Figure 1.4:** (A) Positional correlation functions for unchelated $Mg^{2+}$ ions near the RNA, for different initial RNA-Mg distances of 3.5-6.0 Å. (B) Positional correlations for $K^+$, for initial distances from 3.0 to 6.0 Å. (C) Positional correlations for solvent water, for initial distances from 3.0 to 6.0 Å.

ter with RNA are ordered (Figure 1.5A; Figure A.3 and Table A.2 in the Appendix A): $t(Mg^{2+}) > t(K^+) > t(H_2O)$. Lifetimes of individual bonds show considerable variation, also listed in Table A.2. Experiments indicate that RNA stability is governed by charge density of divalent metal cations [123]. The pattern of preferences for specific RNA species is analogous between $Mg^{2+}$ and water but differs for $K^+$. $Mg^{2+}$ and water reach the longest lifetimes at O2P, 311 ps for $Mg^{2+}$ and 51 ps for water. Water also reaches a similar lifetime with 5' triphosphate oxygens (49 ps). Water lifetimes with both of these most favorable partners exceed the shortest residence time for $Mg^{2+}$, which is 28 ps at sugar oxygen O4', where the lifetime for water is also only 12 ps. O4' is also the only species where the residence time for $K^+$ (47 ps) exceeds that of $Mg^{2+}$ even though $K^+$ reaches a lifetime of 101 ps near base atoms.

The variation of lifetimes for water HBs with nucleotide groups indicates substantial heterogeneity of water dynamics on the RNA surface. The average relaxation time of water HBs between the nucleotides, each with its unique local environment, is between $\sim 2.5-9$ times larger than the lifetime of bulk HBs in water at room temperature. These results suggest slow dynamics of water near the RNA surface [21].

The lifetime of coordination bonds, formed by water or ions, increases on average

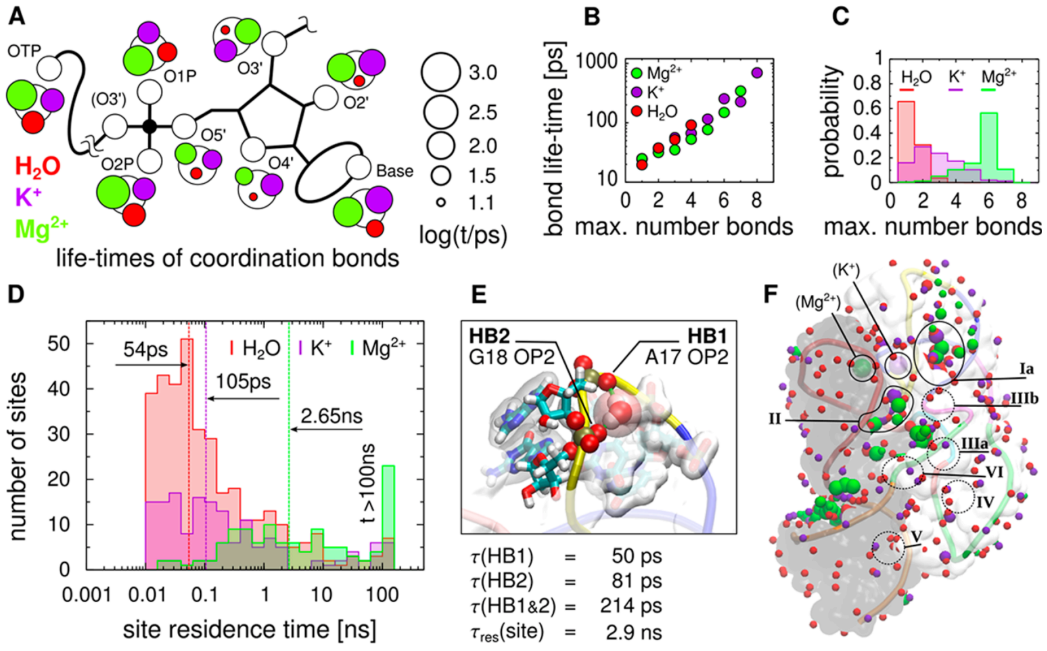


**Figure 1.5:** Coordination bonds and association sites. (A) Summary of lifetimes of coordination bonds between RNA and water or ions. (B) Average coordination bond lifetime as a function of the maximum total number of simultaneous bonds between RNA and the same water or ion. (C) Histogram for the total number of simultaneous coordination bonds formed with the RNA by water, $Mg^{2+}$, and $K^{+}$. (D) Distribution of residence times for $Mg^{2+}$, $K^{+}$, and water in specific association sites near the RNA, determined on the basis of the formation of multiple simultaneous coordination bonds. (E) Illustration of a water molecule in an association site, positioned to form simultaneous hydrogen bonds with two bases. Average lifetimes of single and paired bonds and residence time of water given. (F) Association sites near the RNA surface. Sites for $Mg^{2+}$ (green), $K^{+}$ (purple), and water (red) are distinguished by average residence times: large ($t > 10$ ns), medium (1 ns $< t <$ 10 ns), and small spheres ($t <$ 1 ns). The chelated ions are labeled. Roman numerals indicate locations discussed in the text.

with the total number of such bonds formed simultaneously by the solvent particle involved (Figure 1.5B). This number of simultaneous bonds likely acts as a proxy for the degree of burial, which is known to control residence times. Still, it is surprising how similar the dependence of the average bond lifetime on the number of simultaneous coordination bonds is between both types of ions and water (Figure 1.5B; Figure A.4, Appendix A). The distributions of lifetimes for water HBs with RNA and for metal counterion coordination bonds with RNA indicate that Mg has a significantly more pronounced tail at large bond lifetimes. The differences in average binding times are attributed to the difference in the most probable number of simultaneous binding partners between the species (Figure 1.5C).

**Determination of Association Sites.** On the basis of the observed coordination, we have determined the association sites for ions and water near the 58-mer rRNA. For other recent studies of $Mg^{2+}$ association, see refs [124, 125]. The association sites, similar to observations in more complex systems like riboswitches [47], are all in the outer sphere [47]. The distribution of residence times in the association sites is shown in Figure 1.5D. Examples for median residence times are annotated in the figure. Typical residences last for several ns for $Mg^{2+}$, while potassium residences are shorter, and typical water residences are shorter again. Distributions are, however, broad for all three species and extend from tens of ps to tens of ns. The longest individual residences persist for the length of the simulation run. Residence times beyond 100 ns are collected in the last bin of the histogram, forming a clear spike in the case of $Mg^{2+}$, corresponding to 22 individual long residences.

The sites are identified as locations where water or ions form simultaneous coordination bonds, linking multiple RNA atoms (Figure 1.5E). To quantify time scales of association by water or ions with the RNA, we have examined a positional correlation function between solvent particles and sites. Site residence times were measured where the correlation function fell below 1/e. We have also determined the average

times spent by ligands within 0.2 nm of the association site without excursions. These "dwell times" correlate well with the residence times (Figure A.5 in the Appendix A)).

The association sites for $Mg^{2+}$, $K^+$, and water are shown in Figure 1.5F (and Figure A.6–A.8). We find 120 $Mg^{2+}$ association sites, 161 $K^+$ sites, and 383 water sites in the outer sphere. Among these, 42, 19, and 18 sites, respectively, have residence times greater than 10 ns. The median residence time for Mg is $\sim$ 2.65 ns. All Mg sites with one exception are occupied by hexa-hydrated Mg, some with residence times in excess of 100 ns. In one instance, a $Mg^{2+}$ spontaneously formed two inner sphere contacts with the RNA and remained chelated in X-ray site **C1163** for the rest of the run, in the site that is marginally stable for chelation according to ref [10]. The long-lived outer-sphere association sites, occupied by the hexa-hydrated magnesium, anchor the local ionic RNA environments for long times against fluctuations. The time scales of interactions of the discrete water molecules with RNA nucleotides stretch from several hundred picoseconds to few nanoseconds.

**Placement of Association Sites.** The most stable sites are generally placed in surface crevices, where multiple interactions between RNA and solvent particles are favored. Many are placed in the region around the chelated potassium. Others are located at the packing interface between helix A and helices B and C. One large group of association sites is placed near the chelated $K^+$, however, not at the packing interface between helices A and C, which is bridged by the chelated $Mg^{2+}$, but on the opposite side of the chelated $K^+$. Here, a group of several water, K, and Mg sites form a channel [126] through the structure via loops at the end of helix A, formed by bases 11 to 20 (**Ia**) and near bases 21, 25, 26, and 38 (**Ib**). Note that bases 19−24 bind the chelated $K^+$, and A38 is also involved in a tertiary base pair (BP) with U10. (Roman numerals refer to labels in Figure 1.5F. Additional views, using the same labels, are provided in Figure A.6 in Appendix A.)

Several stable association sites are located near the tertiary base pairs, stabilizing

the main packing interface. One set of three Mg sites is placed in a cleft at the helix packing interface, near the tertiary BPs G21−G41 and C22−C42 (**II**). Matching K sites are observed in the same positions, and several water sites are adjacent, near BP G43−A48 and very close to the chelated potassium. Matching Mg and K sites are also found near the tertiary interface BP U10−A38 (**IIIa**) with water association sites on the other side of the U10−A38 BP (**IIIb**). Another group of Mg, K, and water sites is in a surface cleft near the interface of helices A and B, and also near the strand junction in helix A, formed by BPs U31−A7 and U32−A36 (**IV**). Here the K and water sites are placed deeper inside the pocket than the competing Mg site.

The time scales of interactions of the discrete water molecules with RNA nucleotides stretch from several hundred picoseconds to a few nanoseconds. A tertiary BP, G5−A35, is associated exclusively with a water site (**V**), where a water molecule is able to interpose itself between the binding bases for tens of ns. Outside the interface region, a string of stable Mg sites is found in the groove between helices B and C, starting near the termini and ending in a large pocket at the packing interface, including the BP U10−A38 (**VI, IIIa**) Association sites with shorter residence times are located in the same regions. The groove of helix B contains several more Mg sites with ns residence times and also some K and water sites. Another cluster of water sites is placed more deeply in the groove in helix B, linking U51, C52, and A53 to A4, G5, and G6 via waters with ns residence times, which are either buried or placed in a deep cleft in the structure. Shorter-lived sites are similarly detected near the other stable sites, as well as in other less favored locations.

A string of Mg sites with residence times less than 1 ns follows helix A on the surface of the molecule, clearly visible in Figure A.8A. Another association site is placed further away in helix A, mediating a contact between C29 O2P and A38 O1P. This water site is exposed on the surface of the molecule with a residence time of 180 ps. Exposed K and water sites with residence times below 100 ps trace the entire RNA

chain. Many additional water sites with similar residence times are found elsewhere on the surface, still preferentially in recessed positions. The contact map for interactions mediated by long-lived water bridges is shown in Figure A.9 in Appendix A.

Many of the long-lived magnesium sites correspond to the positions of metal ions that are present in the X-ray structure. Of the 14 magnesium locations found in the two copies of the RNA in the crystal, 9 correspond to long-lived magnesium sites in our simulations (see Figure A.8 in Appendix A)). Two of these, residues **C1164** and **C1170** in the PDB, are placed in the groove of helix B and C. Four more crystal sites are placed on the surface near the end of the groove. Of these, **C1161** and **C1159** are in the locations labeled **II** and **VI** and **C1160** and **D1365**$^*$ correspond to two other nearby sites. (Ions marked with a star are only present in one chain in the crystal structure.) **C1169** corresponds to our site **IV** in helix A. Four more crystal ion positions are located in the channel through the RNA structure, near the chelated ions. **C1172** and **C1163** correspond to our locations **Ia** and **Ib**. **D1372**$^*$ is placed in between, and **C1168**$^*$ is further outside. These two have no corresponding magnesium sites in the simulations. Two more sites that are not occupied in the simulations are **C1166** and **D1364**, near the triphosphate. Possibly the large motions of the terminal in the simulations are responsible. The last unoccupied site is **C1165**$^*$, placed at the other end of the RNA and coordinated by a phosphate oxygen from U116. In the crystal, this ion is also in contact with another phosphate oxygen from A136 in the nearby chain D, which provides an artificial stabilization to this site in the crystal. Hence, this last site is clearly an artifact of the crystal environment. The terminal of the RNA, where two more crystal sites are located, is highly mobile in the simulated solution environment. This may affect the actual stability of the sites, because the binding environment becomes impermanent, but it may also affect the ability of our analysis to identify specific sites in fixed locations relative to the RNA. Finally, the two extra sites in the channel region likely represent a valid alternative pattern of

ion occupation, since they are both occupied in only one of the two copies of the RNA in the crystal. As one of the undetected sites is more distant from the RNA than those observed in the simulations, the configuration that was not sampled in the simulations may be the energetically less stable case.

Overall, the sites identified from the simulations agree with the locations of ions in the X-ray structure, indicating that the force field correctly identifies the most favorable positions for ion association with the RNA. The total number of ions in the crystal is slightly larger than the number of associated ions under the simulated conditions, both according to experiment and in the simulations themselves. One of the extra crystal sites is apparently an artifact. The termini, where another site is not identified by the simulations, are very mobile. In the channel near the chelated ions, the two sites that are occupied by $Mg^{2+}$ in the simulations are present in both crystal structures, while the two unoccupied ones are each only present in one of the copies in the crystal, suggestive of alternative occupation. In any case, ions in the simulations can dynamically visit multiple sites. Simulations, including explicit ions, offer the possibility to investigate correlations between the dynamics of individual ions and dependencies of ion dynamics on the changing conformation of the RNA.

**Collective Dynamics of RNA, Ions and Water.** The collective behavior of RNA, ions, and water that is foundational for the complex polymer system is accessed most directly by a combined analysis of their dynamics. Beyond using principal component analysis (PCA) [120] of the RNA coordinates to determine time scales of global RNA motions (see Figure A.10 and A.11 in Appendix A), we have performed a joint PCA, simultaneously for the coordinates of the RNA phosphorus atoms and for the occupation counts of the magnesium, potassium, and water sites, to reveal the coupling of ion and water dynamics to RNA motions (see the Figure A.12 in Appendix A). Figure 1.6A−C shows the changes in occupation of Mg, K, and water sites, associated with the motions of the P atoms corresponding to the first joint PC,

shown in all three panels. The root-mean-square displacement of the RNA P atoms due to the first mode is also plotted in Figure 1.6D, where some bases that undergo large motions are annotated. The largest RNA displacements, involved in the first joint PCA mode, occur at the termini and near the junction in helix A, around base A35. Base A7, which is placed at the other side of the helix junction, forms a smaller local maximum in the root mean square fluctuation (RMSF), attributed to the first mode. Another local maximum occurs at base C29. The largest motions, near A35 and at the 5' end, are accompanied by the loss of several associated water molecules. Additional water is instead introduced into the core of the RNA. Associated ions are also displaced or exchanged in several affected locations. (1) At the termini, positive displacement along the first mode acts to narrow the groove between helices B and C, while negative displacement acts to widen the groove as the 5' terminus moves outward. As the groove narrows, potassium is lost from the position of X-ray residue **C1170** in the groove, while the magnesium count increases at the expense of water in the adjacent position **C1164**, and in another position closer toward the termini. (2) In the core of the RNA, several sites, placed in the cleft at the end of the groove, are affected. Base A7 is located near the sites in location **VI**, where magnesium is lost in response to a positive displacement along the first mode. Magnesium is also lost from X-ray site **C1161**, while the adjacent position **D1365** gains magnesium occupation. In the more deeply buried site C1159, in location **IIIa**, magnesium replaces potassium. (3) Positive displacement moves base C29 closer in toward a group of magnesium sites on the surface, in location **IV**. In response, magnesium here is shifted from the closer sites away into the more distant sites in the group. (4) Displacements of the RNA are relatively small closer to the chelated ions, and around the nearby channel through the structure, but the occupation of the sites in the channel is still strongly affected. In response to a positive displacement of the RNA, magnesium, found in X-ray sites **C1163** and **C1172**, is replaced by potassium,

which instead occupies site **D1372** placed in between, and positions near **C1172**. The different locations inside the channel that are taken up by magnesium and potassium indicate a correlated response to the changing geometry of the RNA. Overall, larger changes are concentrated in a few regions of the structure. In the groove of helices B and C, changes in ion occupation are directly tied to large displacements of the RNA. Since $Mg^{2+}$ is lost as the groove broadens, the process is not driven by steric repulsion but rather by distortion of the environment, provided by the RNA for the ions. The shift of $Mg^{2+}$ occupation in location **IV**, induced by the movement of base C29, is instead compatible with a steric displacement. No large displacement of the nearby RNA corresponds with the changes in occupation of the sites in locations **IIIa** and **VI** and nearby. The redistribution of ions over several adjacent sites is apparently correlated, and possibly the correlation extends to the other affected sites in adjacent regions, in the groove and near base C29. The replacement of $Mg^{2+}$ by $K^{+}$ in the channel region is also not accompanied by large changes of the RNA structure, but the untypical preference for $K^{+}$ in the nearby chelation site itself shows that subtle structural effects can determine ion occupation. On the other hand, the large displacement of the RNA around base A35 is not directly linked to changes in ion occupation, since only water sites are found in the vicinity. However, the motion itself is linked to the smaller motions of the nearby C29, which visibly affects ion occupations. Meanwhile, A35 and the adjacent A38 are both involved in tertiary BP interactions, stabilizing the RNA structure near the central four-helix junction. The large motion of A35, alongside the introduction of additional water into the RNA as part of the first mode, indicates a reversible disruption of this BP, linked with changes in the structure and composition of the ion atmosphere. (Additional data for the first five PCs are shown in Figure A.13-A.22 in Appendix A) Future work beyond the scope of the current paper, using the combined analysis of RNA motions and site occupations, promises to reveal more insight into the influence of the ion atmosphere

on RNA dynamics.

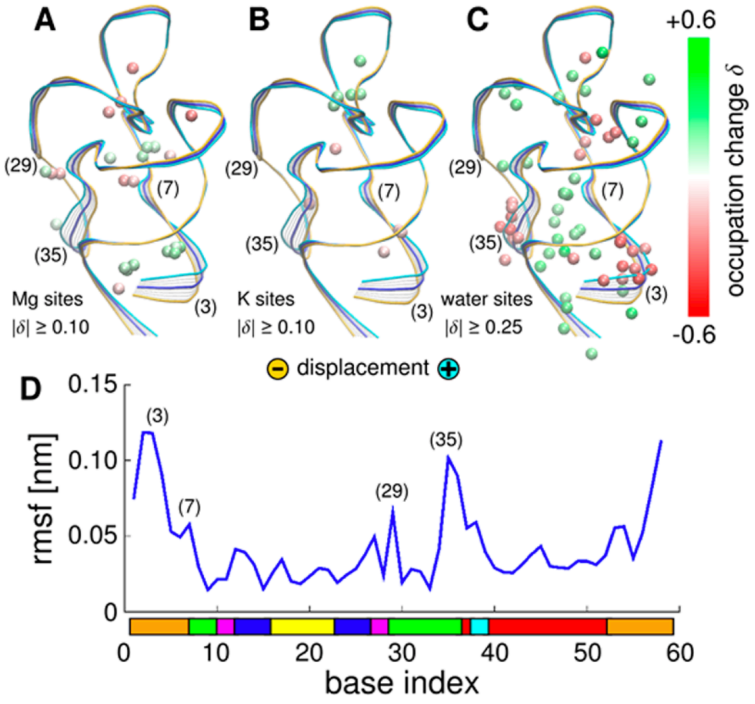


**Figure 1.6:** Joint principle component analysis of RNA dynamics and of Mg, K, and water site occupations. Displacements of RNA P atoms around the average structure (dark blue) due to the first mode are shown, scaled 2×. (A) Color coded changes in Mg occupation counts, associated with the RNA motion, for sites with changes $|\delta| \geq 0.1$. (B) Changes in K counts, for sites with $|\delta| \geq 0.1$. (C) Changes in water counts, for sites with $|\delta| \geq 0.25$. (D) Root-mean-square displacement of RNA P atoms due to the first mode.

## 1.4. Conclusions

We have demonstrated that the 58-mer rRNA fragment in solution forms a complex polymer in conjunction with ions and solvent water, with a wide range of spatial and temporal heterogeneity of the underlying dynamics of its molecular architecture: (1) $Mg^{2+}$ and $K^{+}$ ions near the RNA atoms maintain correlated fluctuations over several ns. (2) The average relaxation time of water HBs between the nucleotides is between $\sim 2.5$–$9$ times larger than the lifetime of bulk HBs in water at room temperature. (3)

The time scales of interactions of water molecules with RNA nucleotides stretch from several hundred picoseconds to a few nanoseconds. (4) Residences in some association sites have lifetimes in excess of 100 ns. The association sites anchor the local ionic environments around the RNA for long times against fluctuations. (5) Joint PCA of the rRNA phosphorus atoms and the occupation numbers in the association sites suggests a microscopic link between ion association and the stability of a key tertiary BP, stabilizing the RNA fold.

# Chapter 2.

# Diffuse ions can coordinate dynamics in a ribonucleoprotein assembly

*Reproduced from Wang, A; Levi, M; Mohanty, U; Whitford, P. "Diffuse ions can coordinate dynamics in a ribonucleoprotein assembly". *Manuscript submitted.*

## 2.1. Introduction

Positively charged cations are required for RNA [1, 2, 8, 92, 94, 127], DNA [128, 129] and ribonucleoprotein (RNP) assemblies [48, 49] to undergo elaborate conformational rearrangements associated with biological function. One of the most well studied RNP assemblies is the ribosome, for which numerous studies have illustrated how appropriate counterion concentrations are required for assembly [50, 51] and rapid transitions between functional states [52, 53]. In addition to altering biochemical kinetics, differences in *in vitro* solution conditions have also been found to dramatically alter the accuracy of protein synthesis [54, 55]. Even though the broad influence of ions on biomolecular kinetics has drawn significant interest, it has been challenging to obtain precise physical-chemical insights into the relationship between the ionic environment and dynamics of RNP assemblies.

When describing RNA, localized ions may be generally categorize as chelated, or diffuse [2, 8]. Chelated ions are characterized as being partially dehydrated, which allows for strong direct contacts to be formed with RNA [2, 8, 9]. Accordingly, chelated ions can remain bound for long (milliseconds to seconds) timescales [130–132]. In contrast, diffuse ions remain fully hydrated (e.g. $Mg(H_2O)_6^{2+}$) and associate less strongly with RNA (Figure 2.1a). The hydration exchange rate of diffuse ion is on the order of microseconds according to experimental investigation [101]. While each diffuse ion only weakly interacts with the RNA, the collective effect of the diffuse ionic environment can be significant, where changes in ion concentrations have been shown to control the balance between folded and unfolded conformations [2]. An additional layer of complexity when describing ionic effects on RNA dynamics is that diffuse monovalent and divalent ions compete for interactions with RNA [6, 7]. Thus, in addition to the transient and weak nature of the interactions, differential charge compensation results in a delicate balance of entropic and enthalpic factors, which together have non-linear effects on biomolecular stability and dynamics.

The essential role of ions on biomolecular dynamics has motivated their study over a wide range of temporal and spatial scales [12–21, 133]. In terms of theoretical/computational techniques, many methods are available to study ionic effects, such as Brownian dynamics simulations [134], implicit-solvent models [19, 35, 135] and all-atom explicit-solvent MD simulations [15, 46, 47]. With regards to implicit-solvent models, there have been many applications of non-linear Poisson Boltzmann (NLPB) theory [23–25], counterions condensation models [6, 7, 26–28, 136] and Debye-Hückel (DH) treatments [30, 31], among other method [32, 137]. However, a general limitation of these approaches is that they do not typically account for ion-ion correlations, or the differential effects of various ionic species (e.g. monovalent vs. divalent) [33, 34]. In the study of larger-scale assemblies, studies that employ coarse-grained models often employ DH representations of the monovalent ions [35]. While DH models are suitable for exploring interactions between opposing charges (e.g. positively charged proteins interacting with DNA [138, 139]), they are unable to capture ion-induced attraction between polyanionic molecules [140]. Motivated by this, there have been efforts to parameterize coarse-grained models with explicit-ion representations, which have been effective in the analysis of DNA-DNA attraction [141, 142]. There is also a long history of ion force field parameterization for use with all-atom explicit-solvent simulations [36, 37], which can circumvent some of the limitations inherent in implicit-solvent models. However, the increased computational requirements of explicit-solvent models often limit simulations of large assemblies to relatively short timescales (i.e. microseconds [66, 67]), making it difficult/impossible to directly describe the relationship between ionic effects and large-scale processes.

To enable the study of diffuse ion dynamics during large-scale conformational events, we present an all-atom (non-Hydrogen atom) model that employs simplified biomolecular energetics, along with a transferrable potential for explicit monovalent ($K^+$,$Cl^-$) and divalent ($Mg^{2+}$) ions. Specifically, an all-atom structure-based

(SMOG) model[116, 143] is used to define intramolecular interactions, while ionic interactions are assigned non-specific effective potentials and coulomb electrostatics. Parameters for ion-ion, ion-RNA and ion-protein interactions were refined against explicit-solvent simulations of small model systems (prototypical RNA fragments (Figure 2.1b,c) and protein S6 (Figure B.3)). A subset of the ion-RNA parameters were then further refined through comparison with an experimental measure of the excess ion atmosphere of a prototypical rRNA fragment. While the model parameters (collectively referred to as the SMOG-ion model) were established using experimental measurements at a single ionic concentration as a reference, we find the model accurately describes the concentration-dependent properties of the diffuse ionic atmosphere for multiple small RNA molecules. With these benchmarks in place, we use the model to simulate a bacterial ribosome in the presence of monovalent and divalent ions. This analysis reveals how the conformation of the ribosome, specifically the L1 stalk, depends on the diffuse ionic environment. We further demonstrate how diffuse ions mediate recruitment and binding of transfer RNA (tRNA) molecules during the elongation cycle of the ribosome. Together, these calculations reveal the intricate relationship between the diffuse ionic environment and the structural dynamics of a large-scale biomolecular assembly.

## 2.2. Methods

### 2.2.1. Structure-based "SMOG" model with explicit ions

The model is constructed by combining the energetics of an all-atom structure-based "SMOG" model [116, 143] with an explicit treatment of diffuse ions ($K^+$, $Cl^-$ and $Mg^{2+}$), where Coulomb electrostatics and effective potentials define the ionic inter-

actions. The potential energy may be generally described in terms of two parts:

$$V = V_{\mathrm{SMOG}} + V_{\mathrm{E}}\cdot, \tag{2.1}$$

where $V_{\mathrm{SMOG}}$ refers to the all-atom structure-based potential energy and $V_{\mathrm{E}}$ describes the potential energy of all electrostatic interactions.

In the all-atom structure-based SMOG model ($V_{\mathrm{SMOG}}$), all non-hydrogen atoms are explicitly represented, and an experimentally-identified configuration is defined as a potential energy minimum. The functional form of the SMOG model is given by:

$$\begin{aligned} V_{\mathrm{SMOG}} = & \sum_{ij\in\mathrm{bonds}} \frac{\epsilon_{\mathrm{b}}}{2}\left(r_{ij}-r_{ij}^{0}\right)^{2} + \sum_{ijk\in\mathrm{angles}} \frac{\epsilon_{\theta}}{2}\left(\theta_{ijk}-\theta_{ijk}^{0}\right)^{2} \\ & + \sum_{i\in\mathrm{improper}} \frac{\epsilon_{\chi\mathrm{imp}}}{2}\left(\chi_i-\chi_i^{0}\right)^{2} + \sum_{i\in\mathrm{planar}} \frac{\epsilon_{\chi\mathrm{planar}}}{2}\left(\chi_i-\chi_i^{0}\right)^{2} \\ & + \sum_{i\in\mathrm{backbone}} \epsilon_{\mathrm{BB}} F_D(\phi_i) + \sum_{i\in\mathrm{sidechains}} \epsilon_{\mathrm{SC}} F_D(\phi_i) \\ & + \sum_{ij\in\mathrm{contacts}} \epsilon_{\mathrm{C}} \left[\left(\frac{r_{ij}^{0}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{0}}{r_{ij}}\right)^{6}\right] + \sum_{ij\notin\mathrm{contacts}} \left(\frac{C_{18}}{r_{ij}^{18}} - \frac{C_{12}}{r_{ij}^{12}}\right), \end{aligned} \tag{2.2}$$

where

$$F_{\mathrm{D}}(\phi_i) = [1-\cos(\phi_i-\phi_i^{0})] + \frac{1}{2}[1-\cos(3(\phi_i-\phi_i^{0}))].$$

$\chi_i^0$ and $\phi_i^0$ were given the values found in a pre-assigned configuration. While $r_{ij}^0$ and $\theta_{ijk}^0$ were assigned the corresponding values found in the Amber99sb-ildn force field [144], as employed in a previous SMOG-AMBER model [145]. Interaction weights were assigned default values (Consistent with previous studies (see ref [116])). Contacts were defined using the Shadow Contact Map algorithm [146] with a 6Å cutoff and a 1Å shadowing radius. In the current study, we introduced a modification, relative to earlier applications of the model, where a 12-18 potential was included for atoms

that are not in contact in the experimentally-defined structure. This was introduced in order to define an excluded volume potential that mimics the AMBER force field, without also including a deep attractive well (Fig. B.1). With this approach, the steric representation provided by the model can mimic that of more highly-detailed models, without introducing a large-degree of non-specific energetic roughness. Here, the coefficient $C_{18}$ and $C_{12}$ were calculated for each type of interaction based on fits to the corresponding 6-12 parameters in the Amber99sb-ildn force field [144] (Fig. B.1).

In the presented model, the electrostatic representation ($V_{\mathrm{E}}$) included direct Coulomb interactions ($V_{\mathrm{coulomb}}$), effective excluded volume potentials for diffuse ions ($V_{\mathrm{ion-excl}}$) and effective potentials that describe ionic solvation effects ($V_{\mathrm{sol}}$):

$$\begin{aligned} V_E &= V_{\mathrm{coulomb}} + V_{\mathrm{ion-excl}} + V_{\mathrm{sol}} \\ &= \sum_{ij} \frac{q_i q_j}{4\pi\varepsilon\varepsilon_0 r_{ij}} + \sum_{ij} \frac{A}{r_{ij}^{12}} + \sum_{ij} \left( \sum_{k=1}^{5} B^{(k)} e^{-C^{(k)}[r_{ij} - R^{(k)}]^2} \right) \end{aligned} \tag{2.3}$$

$V_{\mathrm{coulomb}}$ represents the direct Coulomb interactions between a pair of charges $q_i$ and $q_j$ with interatomic distance $r_{ij}$, while $\varepsilon$ is the dielectric constant for water (80) and $\varepsilon_0$ is the permittivity of free space. In our model, the partial charges of atoms are obtained from Amber99sb-ildn forcefield [144]. However, since only non-hydrogen atoms are explicitly represented, the partial charge of each hydrogen atom was added to the corresponding non-H atom.

The effects of the excluded volumes of diffuse ions is account for by $V_{\mathrm{ion-excl}}$. Consistent with previous efforts to model explicit ions [141], pairwise potentials of the form $\frac{A}{r_{ij}^{12}}$ were used to denote the effect of excluded volume of ion $i$ with ion/atom $j$. The parameter $A$ is different for each type of interaction in the model (e.g. Cl-Cl, Cl-K, K-O, etc). However, for simplicity it notation, subscripts are not shown. Following the protocol of Savelyev and Papoian [141], the values of these parameters

$\{A\}$ were obtained through refinement based on comparison with explicit-solvent simulations (See Appendix B for details).

The last $V_{\text{sol}}$ describes solvent-mediated ionic interactions, which manifest in the form of ionic shells. The functional form is the same as used previously [141], where there is a sum of Gaussians that describe ionic shells (negative weights) and intervening barriers (positive weights). For each type of interaction considered, up to five Gaussians were included to describe up to three ionic shells. In each Gaussian, the parameter $B^{(k)}$ denotes the amplitude, and the alternative signs of $B^{(k)}$ denote the energy barriers and wells. The location ($R^{(k)}$), width ($C^{(k)}$) and amplitude ($B^{(k)}$) of each Gaussian were refined based on comparison with explicit-solvent simulations and experimental measurements (details in SI). Consistent with the assignment of excluded volume parameters, $B^{(k)}$, $R^{(k)}$ and $C^{(k)}$, different values were used for each type of modeled interaction.

The SMOG-ion model is freely available through the SMOG 2 Force Field Repository (https://smog-server.org/smog2 - Force Field ID: `AA_ions_Wang21`).

### 2.2.2. Calculating preferential interaction coefficient

We calculated the preferential interaction coefficient of $Mg^{2+}$ ions ($\Gamma_{2+}$) from simulations and used it as a metric to compare with experimental measurements. $\Gamma_{2+}$ describes the "excess" number of $Mg^{2+}$ ions that accumulate around an RNA molecule [147] due to electrostatic interactions [11, 44]. Here, it is calculated by taking the difference of the total number of simulated ions ($N_{Mg^{2+}}$) in the system and the expected value based on the bulk density. Specifically, the expected value is the product of the bulk density ($\rho_{Mg^{2+}}$) and the volume of the simulated box ($V_{\text{box}}$), such that $\Gamma_{2+} = N_{Mg^{2+}} - \rho_{Mg^{2+}} \times V_{\text{box}}$. To calculate the bulk density, the RNA fragment was first recentered in the box for each frame of the trajectory. Then the box was then partitioned into five equal-width ($\sim 140$ Å) slabs. The average density of $Mg^{2+}$ was

then calculated by excluding the central (RNA-containing) slab, where the reported uncertainty is the standard deviation over 20 replicas of simulations with the same concentration. To reduce the uncertainty in the calculated bulk density, the average was calculated using multiple simulated replicas for each system. However, despite these efforts, due to the dimensions of the simulated systems the uncertainty in of 1 in $\rho_{Mg^{2+}} \times V_{box}$, and therefore $\Gamma_{2+}$, is due to an uncertainty of $\sim$0.005 mM in the calculated bulk value of [$MgCl_2$].

When evaluating $\Gamma_{2+}$, it is necessary to consider the influence of chelated ions, as well as the diffuse environment. Since chelated ions can be bound to an RNA molecule, they may be treated as effectively partially neutralizing the charge of the system. In some cases, these ions are bound to the RNA and they can be unambiguously identified through X-ray crystallography or NMR spectroscopy [148]. Since the SMOG-ion was only parameterized to describe diffuse ions, strongly bound chelated ions need to be assigned *a priori*. In the Ade riboswitch, there are five chelated $Mg^{2+}$ ions (Figure B.2(b)) that have been identified in the x-ray structure (PDB: 1Y26) [44, 149]. However, the binding site of Mg3 (Figure B.2(b)) is formed by the crystallographic packing interactions and no divalent cation binding could be detected in the vicinity of this position [148]. Accordingly, this ion was not pre-defined to be chelated in current simulations. However, since the remaining four $Mg^{2+}$ ions are bound deep within the grooves of the RNA, they were defined to be harmonically restrained their chelation pockets. Further, these chelated ions were included in the calculation of $\Gamma_{2+}$.

## 2.3. Results

To enable the study of diffuse ions during large-scale biomolecular rearrangements, we developed an all-atom model with simplified energetics and explicit ions ($K^+$, $Cl^-$,

$Mg^{2+}$). Here, an all-atom structure-based (SMOG) model [116, 143] defines the intramolecular interactions, while Coulomb electrostatics and effective potentials ($V_E$, Eq. 2.3) describe ionic interactions. The "structure-based" interactions explicitly stabilize a pre-defined biomolecular structure, while the effective potentials ensure that the competition of ionic species is consistent with in vitro measurements. As described below, we first parameterized and benchmarked the model against experimental measurements for small prototypical molecular systems. We then applied the model to determine how the diffuse ionic environment can influence the conformation of the ribosome. In the current study, we focus on single-basin structure-based models, where individual experimental conformations are defined as stable. In doing so, these calculations provide a basis for integrating ionic effects in multi-basin structure-based models, which have been used extensively to simulate conformational dynamics of large-scale biomolecular assemblies [150, 151].

### 2.3.1. Simplified model reproduces in vitro ionic distribution

To establish parameters for the explicit-ion model, we first iteratively refined the interaction strengths based on comparisons with explicit-solvent simulations of multiple systems (Fig. 2.1) for a specific ionic composition[1]. Using explicit-solvent simulations with the AMBER f99-ildn [144] force field and modified $Mg^{2+}$ parameters described by Åqvist [36] as well as parameters for monovalent ions ($K^+$ and $Cl^-$) reported by Joung and Cheatham [37] as a benchmark, we refined the ion-ion, ion-RNA and ion-protein interactions in our model. Here, we employed the procedure of Savelyev and Papoian [141], where linear parameters in the Hamiltonian are iteratively updated through use of a first-order expansion of the partition function (Eq. B.2b). In the following discussion, we will refer to each stage of refinement in terms of the produced parameter set, rN (i.e. round N). The initial parameters (r0) were estimated

[1] $[MgCl_2] \approx 10$ mM, $[KCl] \approx 100$ mM

based on inspection of radial distribution function (RDF) calculated from explicit-solvent simulations (Figure B.4). As expected, the positions and widths of the peaks in each RDF were consistent between the r0 parameter set and the explicit-solvent model (Fig. 2.1e), though there were significant differences in the relative heights of each peak. However, after separately refining the ion-ion, ion-RNA and ion-protein interactions (i.e. r1 parameter set; details in Appendix B), we obtained excellent agreement (Figure B.5) in the RDFs of $Mg^{2+}$, $K^+$ and $Cl^-$ ions (Figs. 2.1e and B.12). This initial stage of refinement ensures that the positions and widths of the ionic shells are comparable to those predicted by an explicit-solvent model.

In the second stage of refinement, a subset of the ion parameters was adjusted based on comparison with experimental measures of diffuse ion distributions. For this, the preferential interaction coefficient of $Mg^{2+}$ ions ($\Gamma_{2+}$) was calculated from simulations of a 58-nucleotide rRNA fragment (Fig. 2.1(b)). For consistency with the experimental conditions [11, 44], we simulated the 58mer with $[MgCl_2] = 1$ mM and $[KCl] = 150$ mM. Surprisingly, even though the model (r1 parameter set) recapitulates the ionic distributions predicted by the explicit-solvent model (Figure 2.1(e), Figure B.4), $\Gamma_{2+}$ was significantly underestimated for the 58-mer. Specifically, the predicted value of $\Gamma_{2+}$ was $2.2 \pm 0.64$, whereas the experimental value was 10.4 [11].

Since the underestimation of $\Gamma_{2+}$ indicates an imbalance in $K^+$ and $Mg^{2+}$ association with RNA [2], we introduced minor changes to the $Mg^{2+}$ and $K^+$ interactions with highly-electronegative RNA atoms. With regards to $K^+$, explicit-solvent simulations of RNA have found that 40-50% of the excess $K^+$ ions are partially dehydrated, which allows for overstabilization of $K^+$-RNA interactions [47, 122]. This is at odds with NMR experiments that indicate most $K^+$ ions typically remain fully hydrated [9]. Since our parameters we based on comparison with an explicit-solvent model, we removed the effective potential that stabilizes interactions between first-shell $K^+$ ions and electronegative ($q < -0.5$) O and N atoms (r2 parameter set). While this

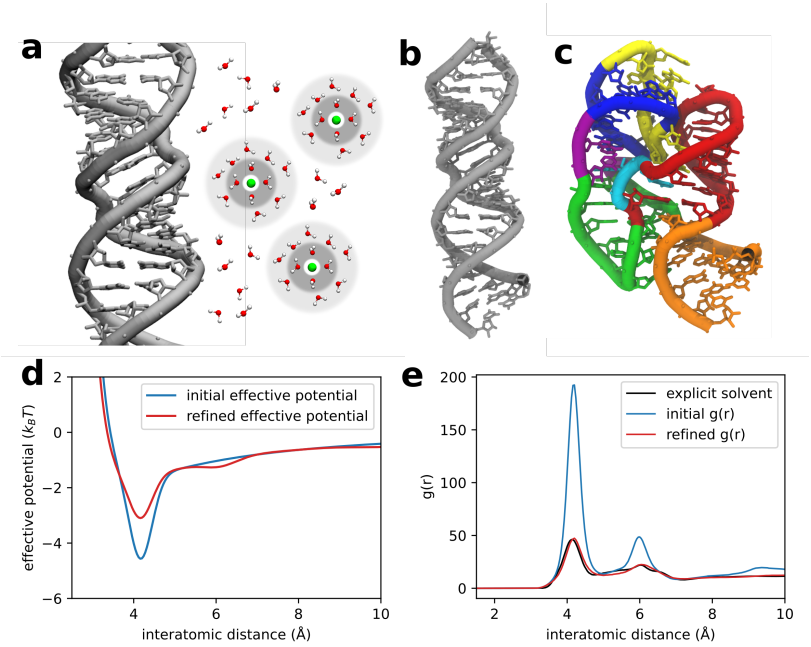


**Figure 2.1: Describing the dynamics of diffuse ions: The SMOG-ion model** a) While chelated ions form strong interactions with biomolecules, diffuse ions (green beads) maintain hydration shells (gray rings) that prevent tight binding. Even though diffuse ions associate weakly, due to their large numbers in solution, they can have substantial collective effects on biomolecular energetics [2]. To study their influence on large molecular assemblies, we developed an all-atom model with simplified energetics (SMOG-ion) in which monovalent and divalent ions are explicitly represented. In this model, intramolecular interactions are defined by a structure-based model [116, 143], partial charges are assigned to each atom, and effective potentials are introduced to account for the effects of ion-ion correlations and hydration. As benchmarks during parameter refinement, we used explicit-solvent simulations of an rRNA helix (h44; panel b) and protein S6 (Fig. B.3), as well and previous experimental measures for the an rRNA 58-mer (panel c). d) Representative interatomic effective potential. Potential for $Mg^{2+}$ interactions with highly-charge RNA oxygen atoms is shown before (s0 parameters, blue) and after (s1 parameters, red) refinement against explicit-solvent simulations of h44. After refinement, the corresponding radial distribution function (panel e) agrees well with that obtained using the explicit-solvent model, which ensures the ionic shells are consistently described by the models. For comparison of all modeled interactions, see Fig. B.5. Minor adjustments to $Mg^{2+}$ and $K^{+}$ interactions were subsequently introduced based on comparison with experiments for the 58-mer [11].

increased $\Gamma_{2+}$ from 2.2 (r1) to 6.4 (r2), the persistent underestimation of $\Gamma_{2+}$ suggested the effective potential for $Mg^{2+}$ was also insufficiently stabilizing. Since $\Gamma_{2+}$ is strongly influenced by the first-shell ionic distributions of $Mg^{2+}$ ions around highly electronegative atoms [47], we modestly increased the stability of these interactions in our model. Based on preliminary free-energy perturbation analysis (not shown), we increased in the strength of the first-shell $Mg^{2+}$ interactions by about 0.15 kcal/mol (r3 parameter set). These parameters predicted a $\Gamma_{2+}$ value of $9.5 \pm 0.68$, which is comparable to the experimental value of 10.4. When considering the uncertainty in calculated $\Gamma_{2+}$ values (see Methods), we decided to terminate the refinement process upon reaching this level of agreement. For all subsequence analysis, we used the r3 parameter set, which will simply be referred to as the SMOG-ion model.

### 2.3.2. Capturing spatially-heterogeneous condensation of diffuse ions

To obtain a qualitative/semi-quantitive assessment of the SMOG-ion model, we compared the predicted ionic spatial distribution function (SDF) to observations from crystallography and explicit-solvent simulations of the 58-mer RNA fragment (Fig. 2.1c). This comparison serves as a preliminary test of the transferability of the model, since the parameters were determined primarily through comparisons with another isolated rRNA helix (Fig. 2.1b).

We find that the SMOG-ion model predicts regions of high ionic density that are corroborated by previous crystallographic analysis. There are two asymmetric assemblies in each unit cell of the crystal structure of the ribosomal protein-RNA complex that contains 58mer (PDB ID: 1HC8), which indicated 8 robust positions of $Mg^{2+}$ ions (Figure 2.2d, pink beads)[2] Through quantitative thermodynamics analysis

[2] 21 coordinates of $Mg^{2+}$ ions are reported in the two asymmetric assemblies of 58mer (chain C and chain D in PDB 1HC8), which yield 13 positions for $Mg^{2+}$ ions after the chain C and D are aligned with each other. 8 of the 13 $Mg^{2+}$ positions are occupied in both copies, while 2

of $Mg^{2+}$ binding to 58-mer using nonlinear Poisson–Boltzmann model [43], only one $Mg^{2+}$ ion (Figure 2.2d, pink bead in black circle) is reported as a thermodynamically favorable site-bound ion with a binding free energy of -4.8 kcal/mol [43], which is also reported as the only chelated $Mg^{2+}$ ion in 58mer [10]. This chelated $Mg^{2+}$ ion is harmonically bound to the oxygen ligands in its chelation pocket in our MD simulations. SDFs of diffuse $Mg^{2+}$ (Figure 2.2a) and $K^+$ (Figure 2.2b) ions are calculated from the SMOG-ion model simulations of 58mer with [$MgCl_2$]=1 mM and [KCl]=150 mM. After applying Gaussian filter on the SDF of $Mg^{2+}$, we observed that 6 out of the rest 7 positions of $Mg^{2+}$ (excluding the chelation pocket) are captured in the high-density regions (Figure 2.2d, white isosurface corresponds to 500 mM density surface of $Mg^{2+}$ in the 1 mM simulation). The only position of $Mg^{2+}$ binding site is located next to the phosphate groups on the terminal G3P residue (Figure 2.2d, pink bead in dashed black circle), which is stable in the crystallographic structure of 58mer due to the interactions with other assemblies while highly flexible in the MD simulations. Besides the chelated $Mg^{2+}$ ions, the crystal structure of 58mer also resolved a position of chelated $K^+$, and two positions of $Os^{3+}$ ions (Figure 2.2d, dark blue beads). In the SMOG-ion simulations of 58mer, one $K^+$ ion is harmonically bound to the oxygen ligands in its chelation pocket (Figure B.2a) while the $Os^{3+}$ ions are not included. It is worth noting that the high-density regions of $Mg^{2+}$ ions calculated from the MD simulations also cover the positions of $Os^{3+}$ ions resolved in the crystal structure.

In addition to accurately predicting the regions of high ionic density, the SMOG-ion model also exhibits qualitative trends that are similar to previously-reported explicit-solvent simulations. Specifically, we find diffuse ions are primarily distributed in the major grooves of the RNA helices (Figure 2.2 a, b), as observed in explicit-solvent simulations [46, 152]. In addition, we observed that $Mg^{2+}$ ions are specifically favorable to the oxygen atoms (Figure 2.2c, red beads) on the phosphate group of

$Mg^{2+}$ positions are only occupied in chain C and 3 more only in chain D.

the nucleic residues, which is unique in comparison with $K^+$ ions, which is consistent with the observation of $Mg^{2+}$ ion SDF in the vicinity of SAM-I riboswitch from explicit-solvent simulations [47]. Together, the similarities between simulation methods suggest the qualitative aspects of the SMOG-ion model are comparable to current explicit-solvent models, while also exhibiting consistency with experimental measurements for the 58mer at a single ion concentration.

### 2.3.3. Model accurately predicts concentration-dependent ionic atmosphere

To quantitatively assess the reliability and transferability of the SMOG-ion model to describe the diffuse ionic environment, we compared the predicted concentration dependence of $\Gamma_{2+}$ with experimental values for two well-studied RNA systems: the 58-mer rRNA fragment [11] and an adenine riboswitch [44]. Comparing concentration-dependent values of $\Gamma_{2+}$ allows one to ask whether the modeled parameters accurately describe the competition of ionic species, which will result from a balance of energetic (i.e. stronger interactions with $Mg^{2+}$) and entropic (two monovalent vs. one divalent ion binding) factors. Further, since the modeled ion-RNA interaction strengths are based on comparison with a single explicit-solvent simulation and a single experimental value of $\Gamma_{2+}$, it is important to verify the model parameters are not overfit to a specific benchmark system/concentration. As described below, the model demonstrates a level of transferrability, since it can accurately predict $\Gamma_{2+}$ over a roughly one order-of-magnitude range of $[MgCl_2]$ values for both RNA molecules.

We first compared predicted and experimental [11] value of $\Gamma_{2+}$ for the 58-mer rRNA (Fig. 2.3a,b). While the model was calibrated with the 58mer at $[MgCl_2] = 1$ mM (and [KCl] = 150 mM), we find it accurately predicts $\Gamma_{2+}$ for $[MgCl_2]$ values ranging from 0.2 to 1.0 mM. Over these concentrations, the experimental $\Gamma_{2+}$ values change by 6.5, whereas the model predicts a change of 5.5. There is a systematic

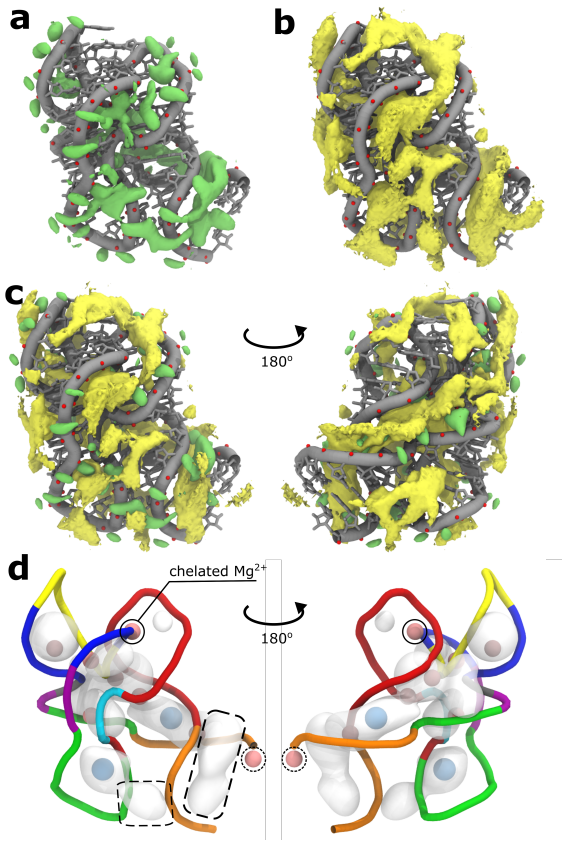


**Figure 2.2: Characteristics of monovalent and divalent ion-RNA association** The SMOG-ion model predicts spatial partitioning of $Mg^{2+}$ and $K^{+}$ ions. a) Isosurface of spatial distribution function (SDF, molar units) of diffuse $Mg^{2+}$ ions for the 58-mer ([$MgCl_2$]=1 mM and [KCl]=150 mM). The isosurface represents to a 1.3M concentration of $Mg^{2+}$ (1300-fold enrichment over the bulk value). b) SDF isosurface (1.3M) for diffuse $K^{+}$ ions. $K^{+}$ is more frequently found in the RNA major groove, consistent with its significant influence on the stability of secondary structure [1]. $Mg^{2+}$ is dominant along the RNA backbone, consistent with its contribution to tertiary structure formation [2, 92]. c) To describe the competition between monovalent and divalent ions, the difference between the SDFs was calculated : $\Delta\rho = \rho_{Mg^{2+}} - \rho_{K^{+}}$, where $\rho_i$ is the SDF of ion type i. The green isosurface shows preferential association of $Mg^{2+}$ ions ($\Delta\rho = 1.3$M), while yellow shows a preference for $K^{+}$ ions ($\Delta\rho = -1.3$M). d) SMOG-ion model predicts population of crystallographically-reported ionic densities. SDF calculated for diffuse $Mg^{2+}$ ions (chelated ion not included) after applying a Gaussian filter. Isosurface shown for 500 mM. Crystallographically assigned $Mg^{2+}$ ions (pink) and $Os^{3+}$ ions (blue) are within the predicted regions of high diffuse $Mg^{2+}$ densities, except for the chelated $Mg^{2+}$ position and one $Mg^{2+}$ ion at the terminal tail (dashed circle). The latter is likely stabilized by crystallographic contacts, in experiments, which rationalizes the the lack of density in the simulations. In addition to predicting the crystallographic ions, there are two additional regions of high density (dashed boxes) that may contribute to stability of tertiary structure.

underestimation ($\sim$ 1) of the $\Gamma_{2+}$ values at higher ion concentrations, though the uncertainty in the theoretical values is comparable to the difference. The implied differences in energetics between the model and experiments are also minimal, since the $\Gamma_{2+}$ values correspond to a difference in the free energy of ion association[3] that is only $\sim 1 k_B T$. In terms of molecular structure, our model explicitly stabilizes the crystallographic configuration. Thus, while minor concentration-dependent conformational shifts may occur in vitro, which would likely alter $\Gamma_{2+}$, such effects are likely to be attenuated in the current version of the model. Considering these limitations and uncertainties, one may consider the residual differences between the modeled and experimental $\Gamma_{2+}$ values to be minor.

We addressed the transferability of the model by applying it to the adenine riboswitch (Figure 2.3c). Since the riboswitch was not utilized for any aspect of parameter refinement, it serves as a true test of the predictive capabilities of the model. Further, since the experiments were performed at lower [KCl] values for the Ade riboswitch than for the 58-mer (50 mM [44] vs. 150 mM [11]), these comparisons also test the concentration-dependent influence of the monovalent ions. We find that the predicted $\Gamma_{2+}$ values agree very well with the experimental measurements, with differences that are less than 0.3 (Fig. 2.3d). This implies that the model predicts a free energy of association that is with 0.3 $k_B T$ of experimental measures. Combined with our analysis of the 58-mer, these comparisons demonstate the ability of the SMOG-ion model to accurately estimate diffuse ion energetics and distributions.

### 2.3.4. Diffuse ions control ribosome conformation

The demonstrated level of agreement with experiments for model systems, combined with the transferability of the SMOG-ion model (Figure 2.3), indicates that it a suitable tool for exploring monovalent and divalent ionic distributions around RNA

[3]Free energy of association calculated using: $\Delta G_{RNA-2+} = -RT \int_0^{c_{2+}} \Gamma_{2+} d[\ln c_{2+}]$ [11]

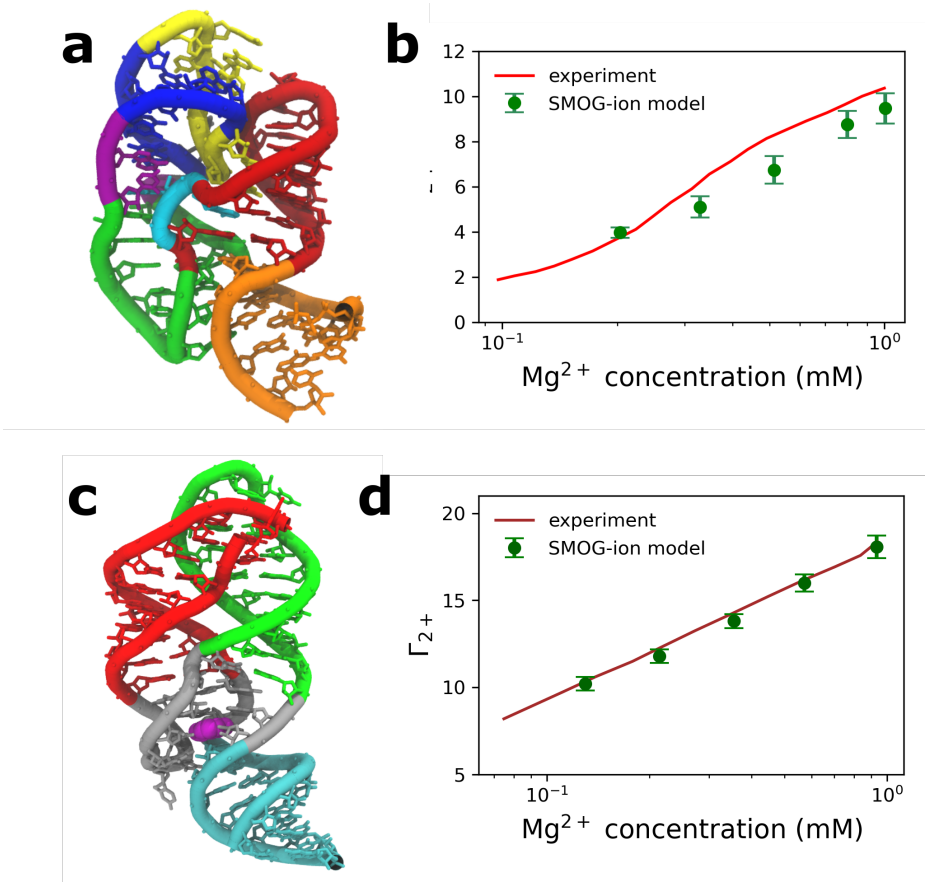


**Figure 2.3: Simplified model captures concentration-dependent ionic distributions** a) Tertiary structure of the 58-mer rRNA fragment [10] (colored as in Fig. 2.1c). b) Comparison of experimental [11] and simulated values of the preferential interaction coefficient ($\Gamma_{2+}$) for the 58-mer ([KCl] = 150mM). While the model was parameterized based on comparison with the experimental value of $\Gamma_{2+}$, obtained at [$MgCl_2$] = 1mM and [KCl] = 150mM, the predicted concentration dependence follows the experimental behavior. c) Tertiary structure of the adenine riboswitch [44], colored as in Fig. B.2b. d) Experimental [44] and predicted values of $\Gamma_{2+}$ for the adenine riboswitch ([KCl] = 50mM) exhibit excellent agreement. This represents a blind test of the model, since the riboswitch was not used during model parameterization. Further, the SMOG-ion model parameters were established using benchmark systems that were at higher value of [KCl] (100-150 vs. 50 mM). Accordingly, these comparisons illustrate the transferrability of the model to other RNA systems and ionic concentrations. Error bars in (b) and (d) represent the standard deviation of $\Gamma_{2+}$ calculated from 20 replicate simulations.

molecules. In addition, since the employed interaction potentials are additive pairwise terms, the simulations exploit available multi-level parallelization algorithms (Fig. B.11) to achieve long time trajectories of large assemblies (millisecond effective timescales). With these capabilities, we applied the model to ask how diffuse ions can affect the structure and dynamics of the ribosome (Fig. 2.4). While recent anomalous scattering data has been used to determine the positions of bound ions [59], the SMOG-ion model provides a complementary view of the diffuse ionic atmosphere.

For our initial application of the SMOG-ion model to the ribosome, we will focus on the dynamics of the L1 stalk (Fig. 2.4). During protein synthesis by the ribosome, tRNA molecules rapidly (10-100 milliseconds) traverse three distinct ribosomal binding sites (A, P and E). These rearrangements are facilitated by interactions with the L1 stalk, which can transiently bind and release the tRNA molecules. In particular, the stalk has been found to associate with tRNA molecules during the late stages of translocation and tRNA release [153–155]. Microsecond-scale explicit-solvent simulations highlight the flexibility of this region, where was found to spontaneously undergo nanometer-scale fluctuations [66]. This predicted flexibility is consistent with the body of structural [156] and smFRET [155, 157, 158] observations that indicate a balance between "open" and "closed" conformations is necessary for efficient tRNA translocation. While previous explicit-solvent simulations have detailed the L1-tRNA interface interactions [65, 68], the degree to which diffuse ions facilitate these interaction has not been reported.

To dissect the energetic and entropic factors that control movement of the L1 stalk, we performed simulations using multiple SMOG model variants, including SMOG-ion. In order to isolate the ionic-dependent properties of the stalk, this set of simulations was performed in the absence of the E-site tRNA molecule (FIG 2.4a). When the system was simulated with a conventional all-atom structure-based model (i.e. an electrostatics-free SMOG) [116, 143], there is a shift in L1 position from

its crystallographic structure position (Figure B.6a). Since the crystallographic position is explicitly defined to be the potential energy minimum, this shift reveals how configurational entropy favors more outward positions (Figure 2.4). The simulations were then repeated using a model for which Coulomb electrostatic interactions were included in the absence of ions (SMOG-coulomb model; Tab. 2.1). As expected, due to negative charge of the tRNA and rRNA residues, Coulomb repulsion leads to an exaggerated open-like configuration of the L1 stalk (Figure 2.4). Specifically, there is a ∼14 Å displacement of the stalk [4], relative to its crystallographic position. To partially account for ionic screening, we performed a third set of simulations with a Debye-Hückel potential. Similar to the pure Coulomb potential, the net repulsive character of RNA-RNA interactions favor more extended L1 positions, which leads to a ∼10.5 Å displacement of the stalk relative to its crystallographic position(Figure B.6b). However, when we explicitly included the diffuse ionic environment (SMOG-ion model), the stalk is found to favors more closed-liked conformations. In fact, rather than extending outward from the crystallographic configuration, it is displaces inward by ∼8 Å. Since the Coulomb and SMOG-ion models are identical, except for the diffuse ionic environment, these calculations directly implicate the diffuse ions as a critical factor that determines the position of the L1 stalk region. As discussed in the next section, this ion-induced inward motion suggests that diffuse ions allow the stalk to coordinate tRNA motion.

### 2.3.5. Diffuse ions mediate tRNA-ribosome interactions

Since our initial simulation of the ribosome suggest that diffuse ions can reposition the L1 stalk in order to favor the formation of tRNA interactions, we next applied the SMOG-ion model to determine whether diffuse ions mediate ribosome-tRNA interactions. To ask this question, we used multiple models to simulate the ribosome

[4] P atom in residue A2158 of 23S rRNA

**Table 2.1:** List of MD simulations presented in this work.

| Models | electrostatics | diffuse ions | system size (atoms+ions) | |
|---|---|---|---|---|
| | | | with out E-site tRNA | with E-site tRNA |
| SMOG | None | None | 149994 | 151620 |
| SMOG-coulomb | Coulomb | None | 149994 | 151620 |
| SMOG-DH | Debye-Hückel | None | 149994 | 151620 |
| SMOG-ion | Yes | Yes | 197794 | 199345 |

(see Table 2.1), with the E-site tRNA included. In order to investigate the effects of diffuse ions on the position of L1 stalk and avoid the stabilizing effects from the contacts between E-site tRNA and the L1 stalk on the close-like configuration of the stalk, we removed the contacts between E-site tRNA and the L1 stalk and L1 protein. When the ribosome was simulated with electrostatic-free SMOG model [116, 143], we observed an outward shift in L1 stalk position from its crystallographic position by a distance of ∼10 Å. This shift of L1 stalk is similar to that observed in the simulation without E-site tRNA, but with a longer distance away from the crystallographic position of L1 stalk, which also stretches the distance between L1 stalk and E-site tRNA[5] from 10.3 Å from the crystal structure to 14.6 Å. Then the simulation was repeated with SMOG-coulomb model in which Coulomb electrostatic interactions were included for the ribosome. Driven by the Coulomb repulsion, L1 stalk performed a displacement of ∼20 Å away from its crystallographic position, which further stretched the distance between L1 stalk and E-site tRNA to 19.4 Å. On the basis of SMOG-coulomb simulations, we performed simulations with Debye-Hückel potentials to account for the ionic screening at a concentration of 100 mM, from which we observed slightly less outward displacement of L1 stalk in comparison with simulations using pure Coulomb potential. When the ribosome was simulated with SMOG-ion model in 10 mM [$MgCl_2$] and 100 mM [KCl] explicit diffuse ionic

[5]Measured by the distance between P atom in residue G19 of E site tRNA and P atom in residue G2112 of 23S rRNA

environment, L1 stalk stayed close to its crystallographic position with only a 4.2 Å displacement, which also yielded the shorted average distance from the E-site tRNA (12.3 Å) among the four sets of simulations (2.2), which indicated the impact of diffuse ions on coordinating the interactions between L1 stalk and tRNA. In comparison with the simulations without the E-site tRNA, it could be observed that the electrostatic- or ion-induced displacement of L1 stalk is robust to the presence of E-site tRNA (Figure B.7).

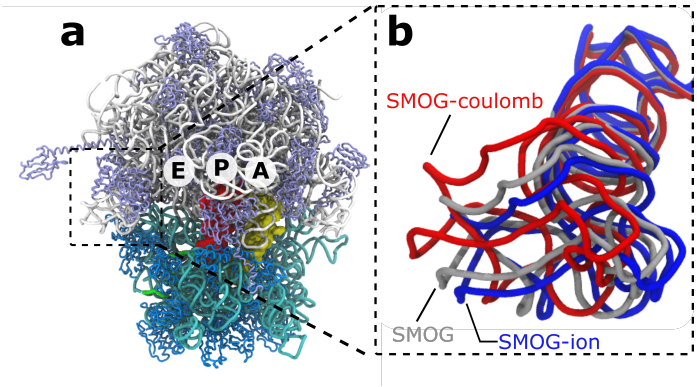


**Figure 2.4: Diffuse ions induce conformational rearrangements in the ribosome** a) Crystallographic structure of the bacterial ribosome (PDB: 6QNR [59], light grey). During elongation, tRNA molecules (red, yellow) sequentially bind the A, P and E sites. During tRNA translocation (P-to-E site displacements), L1 stalk (in black dashed box) can bind and facilitate tRNA movements. b) Average structures obtained with multiple variants of the all-atom structure-based model. When an electrostatics-free (gray) model is used, the L1 stalks maintains a position that is consistent with the crystallographic structure (Fig. B.6a, cyan). When electrostatic interactions are included, in the absence of diffuse ions, the L1 stalk adopts a more extended/outward configuration (red). This effect may be attributed to charge-induced backbone repulsion that disfavors the B-form RNA structure. In contrast, when diffuse ions are included with the SMOG-ion model, the stalk adopts a more inward conformation, relative to the electrostatics-free case. This reveals the strong effect that diffuse ions can have on the dynamics of extended structural elements in the ribosome.

To describe the effect of diffuse ions on the interactions between L1 stalk and E-site tRNA semi-quantitatively from the energetic perspective, we calculated the potential energy between L1 stalk and E-site tRNA ($U_{L1-tRNA}$), between L1 stalk and diffuse ions ($U_{L1_{ions}}$) as well as that between E-site tRNA and diffuse ions

**Table 2.2:** Summary of displacement of L1 stalk with the presence of E site tRNA in ribosome using varied simplified models.

| Models | without E-site tRNA | with E-site tRNA | |
|---|---|---|---|
| | displacement of L1 stalk(Å)[a] | displacement of L1 stalk(Å)[a] | $R_{\text{L1-tRNA}}$[b] |
| crystal structure | 0 | 0 | 10.3 |
| SMOG | 6.7 | 10.1 | 14.6 |
| SMOG-coulomb | 13.9 | 19.7 | 19.4 |
| SMOG-DH | 10.5 | 16.5 | 17.8 |
| SMOG-ion | -8.3[c] | 4.2 | 12.3 |

[a] Measured by the atomic distance of atom P in residue A2158 of 23S rRNA from its position in crystal structure.
[b] Measured by the distance between atom P in residue G19 of E site tRNA and atom P in residue G2112 in 23S rRNA. See Figure B.8 for the definition of $R_{\text{L1-tRNA}}$. See Figure B.9 for the statistical distribution of the $R_{\text{L1-tRNA}}$ value.
[c] Negative number means L1 stalk has inward displacement in comparison with its crystallographic position.

($U_{tRNA-ions}$) from the SMOG-ion model simulations of ribosome with the presence of E-site tRNA (Figure 2.5c). These potential energies were plotted as a function of the distance between L1 stalk and E-site tRNA[6]. When the L1 stalk approaches the E-site tRNA from 19 Å apart to 9 Å, $U_{L1-tRNA}$ monotonically increases by ∼12 $k_{\text{B}}T$, while $U_{L1_{i}ons}$ and $U_{tRNA-ions}$ decreases by ∼20 $k_{\text{B}}T$ and ∼8 $k_{\text{B}}T$, respectively. The potential energy comparison indicated that diffuse ions not only counterbalance the repulsive interaction between negatively charged L1 stalk and E-site tRNA, they also provide an attractive tendency between the L1 stalk and E-site tRNA.

To better understand which component of the diffuse ions modulate the interactions between the L1 stalk and the E-site tRNA, we used simulations with the SMOG-ion model to calculated the spatial distributions of $Mg^{2+}$ (Figure 2.5a) and $K^{+}$ ions (Figure B.10) in the vicinity of the L1 stalk . In sharp contrast to the finding of discrete chelated ions [59], we find that the $Mg^{2+}$ ions have scattered spatial

[6]Measured by the distance between P atom in residue G19 of E site tRNA and P atom in residue G2112 of 23S rRNA

distributions that span the major groove of the rRNA and tRNA. Figure 2.5a) shows $Mg^{2+}$ density surface of 3 M, which is 300-folds enrichment of the $Mg^{2+}$ above the bulk concentration of 10 mM. However, when the $K^+$ density surface of 3 M is highlighted (Figure B.10), these high density regions can be rarely seen in the vicinity of the L1 stalk or E-site tRNA. In addition to SDF, we also counted the number of $Mg^{2+}$ and $K^+$ that are within 1 nm range of both Helix 77 and 78 in the L1 stalk and the E-site tRNA elbow to anticodon region(Figure 2.5d), in which $Mg^{2+}$ shows significant population than the $K^+$ ions. This strong spatial heterogeneity also illustrates why mean-field approaches, such as the DH treatment tend to provide misleading evidence for the effect of ions on ribonucleoprotein assemblies.

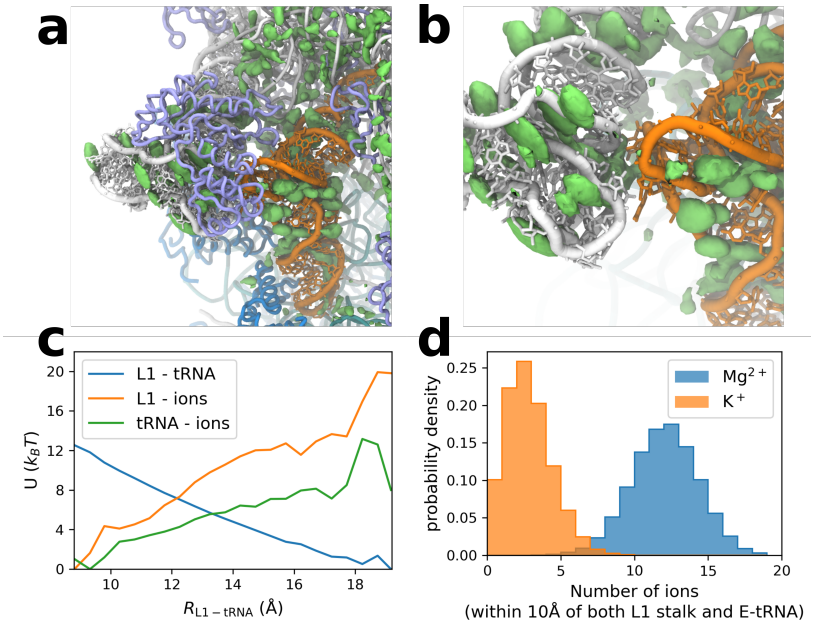


**Figure 2.5:** The SDF of $Mg^{2+}$ ions (panel a, green isosurface) and $K^+$ ions (panel b, yellow isosurface) in the vicinity of L1 stalk (cyan) and E-site tRNA (orange). The density surface of $Mg^{2+}$ and $K^+$ at 3 M are shown in panel a and b respectively. The 3 M density surface corresponds to 300-fold enrichment of $Mg^{2+}$ above its bulk concentration (∼10 mM) while 30-fold enrichment of $K^+$ above its bulk concentration (∼100 mM). Panel (c) shows the numbers of $Mg^{2+}$ ions and $K^+$ ions that are within 10 Å of both helix 77 and 78 in L1 stalk and the elbow to anticodon region in E-site tRNA are shown in terms of their probability density in production simulation trajectories.

## 2.4. Discussion

Savelyev *et al.* developed a coarse-grained model for aqueous solutions of monovalent ions (e.g. NaCl, KCl), in which both electrostatic and short-range hydration effects were taken into account [141]. This model was applied on simulations of double stranded DNA which generated realistic local motions of DNA and reproduced the large-scale chain dynamics [142]. In this study, we generalized the work to construct the simplified SMOG-ion model which include explicit electrostatic and implicit solvation/desolvation effects for both monovalent ($K^+$, $Cl^-$) and divalent ($Mg^{2+}$) diffuse ions, and where non-hydrogen atoms of RNA and protein as well as diffuse ions are explicitly represented. This simplified model not only accurately describes the metal ion condensation of RNA fragments, but also allows us to explore ion effects on the conformational dynamics of large-scale ribonucleoprotein assemblies.

### 2.4.1. Comparison with other models for excess ion atmosphere

It is worth comparing the predicted $\Gamma_{2+}$ values with that obtained from generalized Manning condensation models by Hayes *et al.* [45] on the same RNA systems. Hayes *et al.* [45] have proposed an electrostatic model for RNA by a statistical mechanical generalization of Manning theory of screening and condensed ions, treating $K^+$and $Cl^-$ ions implicitly and $Mg^{2+}$ ions explicitly. In the condensation model, the density of the condensed Manning ions is described by the sum of two normalized Gaussian distributions, the centers of which are located on the coordinate of each RNA charge. The free energy of mixing of the ions is regulated by the mixing Gaussians [45]. The accessibility of the ions near RNA is imposed by the hole Gaussians [45]. In implicit XCl condensation, the full atomic structure of the RNA is employed in which charges are placed on the phosphates. The charges interact by Debye-Hückel electrostatics.

These authors have synthesized the structure-based counterion condensation model of RNA with MD simulations [45].

For adenine riboswitch, Hayes *et al.* predict $\Gamma_{2+}$ to be 18.41 and 10.61 at 1.00 mM and 0.14 mM $Mg^{2+}$, respectively. For 58-mer rRNA, the ion preferential interaction coefficient $\Gamma_{2+}$ is 10.11 and 4.10 at 0.94 mM and 0.235 mM $Mg^{2+}$, respectively. For adenine riboswitch, our model predict $\Gamma_{2+}$ to be 18.06 and 10.21 at 0.93 mM and 0.13 mM $Mg^{2+}$, respectively, which is consistent with both Hayes *et al.* [45] and experiment results [44]. For 58-mer rRNA, our predicted values of $\Gamma_{2+}$ are 9.47 and 3.967 at 1.00 mM and 0.204 mM $Mg^{2+}$respectively, which is much closer result to experiments at the lower concentration.

Taubes *et al.* [28] obtained tight lower and upper bounds to the preferential interaction coefficient for polyions in aqueous solution containing univalent cations valid from low to intermediate concentrations. They exploited the observation that no function has a local minimum or a local maximum where its Laplacian is negative or positive, respectively. Their results are valid for rigid polyions of cylindrical symmetry but arbitrary length and linear charge density.

In the future, it will be fruitful to further investigate the ionic effects on the L1 stalk region using SMOG-ion model. The E-site tRNA reconfiguration can be examined under various ionic environments to reveal the stabilizing effects on its contacts with the L1 stalk in the tRNA release process. Besides, simulations can be conducted on different conformations of ribosome to demonstrate how the inward displacement of the L1 stalk would affect the formation of interactions between the P site tRNA along its movement towards the E site. In addition, with a crystal structure that has a well-resolved domain II of protein L1, SMOG-ion model provides the possibility of observing the interdomain motion in protein L1, which will deepen our understanding of its physiological significance.

## 2.5. Conclusion

The ionic environment plays an essential role in the large scale conformational dynamics and biological functions of ribonucleoprotein assemblies, such as ribosomes and introns. To facilitate the study of ionic effects on the 70S *Thermus thermophilus* ribosome, we developed a simplified SMOG-ion model based on energy landscape principles with all non-hydrogen atoms, explicit ions, Coulomb interactions between all charges, and implicit treatment of solvation/desolvation barriers of ions. The simplified energy landscape model describes the electrostatics in the outer-sphere in 58-mer rRNA and adenine riboswitch at the physiological ranges of both $K^+$ and $Mg^{2+}$ concentrations. Our analysis suggests that diffuse ions stabilize the contacts between the L1 stalk and E site tRNA. As biophysical studies continue to explore various facets of ribosomal dynamics, the current analysis combines experimental investigations with theoretical and computational approaches to provide insights into how electrostatic interactions and ionic effects perturb large-scale conformational dynamics of ribonucleoprotein assemblies.

# Chapter 3.

# Fluctuation Effects in the Adam–Gibbs Model of Cooperative Relaxation

*Reproduced with permission from Hutchison, C., Bhattarai, A., Wang, A. and Mohanty, U., 2019. "Fluctuation Effects in the Adam–Gibbs Model of Cooperative Relaxation." *The Journal of Physical Chemistry B*, 123(38), pp.8086-8090.

## 3.1. Introduction

There have been notable advances in recent years in our understanding of structural dynamics and thermodynamics of supercooled liquids. These liquids are in internal equilibrium in the temperature range between the glass-transition temperature and the melting temperature. Some unique characteristics of glass-forming liquids are the dramatic increase of relaxation times of the primary structural relaxation ($\alpha$-process) with lowering of temperature, the nonexponential time dependence of relaxation time in response to external perturbations, and the nonlinearity of the supercooled state [159–164].

The structural relaxation time $\tau_\alpha$ in glass-forming liquids is often approximated by the VFT relation, $\log(\tau_\alpha/s) = C + D/(T - T_0)$, where $C$ and $D$ are constants, and the temperature $T_0$ is a characteristic temperature greater than zero [159, 160, 165]. Various models for activation energies have been proposed that exhibit divergence of structural relaxation time at th VFT temperature $T_0$. These models include dynamical length scales [166] and the decrease of configurational entropy as the temperature is reduced toward the glass-transition temperature [167]. The entropy difference between the liquid and the crystal, i.e., the excess entropy $S_{\text{exc}}$ vanishes at the Kauzmann temperature ($T_{\text{K}}$) that is close to the VFT temperature $T_0$ [159, 168]. This implies an apparent connection between the vanishing of the excess entropy and the divergence of the relaxation times[168].

A quantitative model for relaxation time in glass-forming liquids is that due to Adam–Gibbs. In the Adam–Gibbs model [82], the smallest size of a cooperative region that allows a transition is inversely proportional to the configurational entropy, $z^* = s^* N_{\text{A}}/S_{\text{c}}$, where $N_{\text{A}}$ is the Avogadro number, $S_{\text{c}}$ is the configuration entropy of the supercooled liquid, and $s^*$ is the configurational entropy corresponding to the lowest size [82]. The Adam-Gibbs model predicts a linear relation between log (experimental relaxation time) and $(TS_{\text{c}})^{-1}$. There have been various generalizations

of the Adam-Gibbs model [83–86, 169–173].

Analysis of glass-forming liquids that span a range of fragilities signifies that the Adam–Gibbs relation is consistent with dielectric and configurational entropy data over the temperature range $T_g < T < T_B$, where $T_g$ is the glass-transition temperature and $T_B$ is a characteristic temperature where the temperature dependence of the relaxation time is observed to change from one VFT form to another VFT form [165, 171].

There are several weaknesses to the Adam–Gibbs mode. The model cannot account for the fact that relaxation of molecules exhibits dynamic heterogeneity and that relaxation near $T_g$ is well described by a stretched exponential decay. For $T > T_B$, the linear relation between $\log(\tau_\alpha)$ and $(TS_c)^{-1}$ breaks down, where $\tau_\alpha$ is the experimental primary structural relaxation time [165, 171]. The $\beta$ KWW parameter governs the breakdown [171]. THe observed correlation between the steepness index and the $\beta$ KWW parameter of the $\alpha$ relaxation time cannot be explained by the Adam–Gibbs model [171].

Adam–Gibbs assumed that the term that is maximal of the configurational partition function int he isothermal-isobaric ensemble can be associated with the logarithm of the number of configurations, i.e., proportional to the configurational entropy $S_c$ [82]. The excess entropy $S_{exc}$ of a glass former is defined as the difference btweent he entropy of the liquid and that of its crystal phase [82]. The excess entropy $S_{exc}$ incorporates vibrational contributions to the entropy of the glass-forming liquid. In the Adam–Gibbs model, the configurational entropy $S_c(T)$ is proportional to $S_{exc}(T)$ [82]. The configurational fraction is the fraction of the excess entropy that characterize the configurational contributions and is defined by $S_c(T) = fS_{exc}(T)$, where $f$ is assumed to be independent of temperature [86, 165, 174–176]. The characteristics of the configurational fraction $f$ in the vicinity of the glass-transition temperature provide insights into the underpinnings that link the excess entropy with kinetic properties described

in the Adam–Gibbs model.

Our work was motivated by several experimental studies. Dielectric spectroscopy has been used to measure the spectra of eight glass-forming liquids in the presence of direct current-bias electric fields [86]. Using the Adam–Gibbs model, after suitably modified, these authors find that field-induced relative change in the relaxation time is correlated with field-induced changes in the thermodynamic entropy [86]. Their work suggests that to generalize the Adam–Gibbs model to describe field effects, what is needed is the way to quantify the ration of the configurational and the excess entropies, i.e., the configurational fraction [86, 165].

Studies related to dynamics of supercooled liquids under high pressure provided another motivation for this work [174–176]. Analysis of the relaxation time for *ortho*-terphenyl (OTP) and *ortho*-phenylphenol mixtures as a function of the temperature at two pressure indicates that the relaxation time is a linear function of $(TS_{\mathrm{c}})^{-1}$ signifying the validity of the Adam–Gibbs model [175]. In another study, the primary relaxation time for *ortho*-tephenyl, triphenylchloromethane, and poly(methyl methacrylate) over a wide range of temperatures and pressures were measured that confirmed the validity fo the Adam–Gibbs model [176]. Analysis of dielectric relaxation times and thermal expansion data for these liquids indicates that both isothermal and isobaric configurational entropies are proportional to the excess entropy [176].

In this work, we describe the configurational fraction at $T_{\mathrm{g}}$ based on both the Adam–Gibbs model and a generalization of the Adam–Gibbs model, in which we have taken into account fluctuation effects of the number of molecules in a cooperative rearranging region. In this case, we quantitatively predict the configurational fraction at $T_{\mathrm{g}}$ of several glass-forming liquids that span a range of fragilities. For these liquids, we observe a connection between the configurational fraction and the $\beta$ KWW parameter at $T_{\mathrm{g}}$.

## 3.2. Methods

Structural constraints associated with order in the melt impose a characteristic size of for the smallest correlated volume $V$ for entropy fluctuations [177]. We assume that the volume V is large enough and migh contain a number of subregions of size $z^*$, each of which is able to relax to equilibrium independent of what is going on in the neighboring regions. In other words, we have not assumed, as did Adam–Gibbs [82], that the number of molecules inside the correlated volume $V$ is equal to the minimum number of molecules $z^*$ need for cooperative rearrangements into a different configuration.

The mean relaxation time of the rearranging region in the Adam–Gibbs model is [82]

$$\ln\tau = \ln A + \frac{z^*\Delta\mu}{k_{\mathrm{B}}T} \tag{3.1}$$

$\Delta\mu$ is the potential barrier per molecule to rearrangement region, $k_{\mathrm{B}}$ is the Boltzmann constant, and $A$ is proportional to the relaxation time at high temperature. We assume that the relaxation time $\tau_i$ of a local region $i$ if of the same form [177, 178] as Eq. 3.1

$$\ln\tau_i = \ln A + \frac{z_i^*\Delta\mu}{k_{\mathrm{B}}T} \tag{3.2}$$

Here, $z_i^*$ is the minimum number of molecules that permits cooperative rearrangements in local region $i$. The average value of $z_i$ is $z^*$.

The width of the distribution of relaxation time $\langle\Delta^2\ln\tau\rangle$ at a given temperature $T$ and pressure $P$ is defined to be $\langle(\ln\tau_i - \ln\tau)^2\rangle$. We assume that fluctuations in the equilibrium melt arise from fluctuations in the configurational entropy $S_c$. Since we as interested in the temperature dependence of $\langle\Delta^2\ln\tau\rangle$ at constant pressure, we write $z^* = z^*(S_c(T))$.

From eqs 3.1 and 3.2, the mean-square deviation of $\ln\tau_i$ from its average is found

to be

$$\langle \Delta^2 \ln\tau \rangle = \Big(\frac{\Delta\mu}{k_{\mathrm{B}}T}\Big)^2 \Big(\frac{\partial z^*}{\partial T}\Big)_P^2 \Big(\frac{\partial T}{\partial S_c}\Big)_P^2 \tag{3.3}$$

The last term in eq. 3.3 is simplified by the use of thermodynamic fluctuation theory [179]

$$\langle \Delta^2 S_c \rangle = \frac{k_{\mathrm{B}}\nu\Delta C_{\mathrm{P}}}{V} \tag{3.4}$$

where $\nu$ is the molar volume, $V$ is the volume of the correlated region, and $\Delta C_{\mathrm{P}}$ is the constant pressure configurational heat capacity.

We assume that the temperature dependence of the configurational entropy is approximated by the excess entropy, $S_c(T) \approx S_{exc}(T) f(T_g)$, where $f(T_g)$ is the configurational fraction [86, 165] at the glass-transition temperature $T_g$. Making use of eqs. 3.3, 3.4 and C.5 we find the configurational fraction to be expressible in terms of the excess heat capacity under constant pressure, $\Delta C_{\mathrm{P,\,exc}}$, wht width of the distribution of relaxation times, and the Tool–Narayanaswamy–Moynihan nonlinearity parameter $x$, which measures how far the system deviates from equilibrium [170]:

$$f(T_g) = \Big(\frac{\nu}{V}\Big)\Big(\frac{k_{\mathrm{B}}}{\Delta C_{\mathrm{P,\,exc}}}\Big)\Big(\frac{(1-x)^2}{\langle \Delta^2 \ln\tau \rangle}\Big)\Big(\frac{\mathrm{d}\ln}{\mathrm{d}\ln T}\tau\Big)^2. \tag{3.5}$$

The first term in eq. 3.5 is the number of structural units per mole. The last term in eq. 3.5 is related to the steepness index $m = \mathrm{d}\,\log_{10}\tau/\mathrm{d}(T_g/T)_{T=T_g}$ [164, 180, 181]. In deriving eq. 3.5, we have not assumed the Adam–Gibbs relationship between the critical size $z^*$ and the configurational entropy $S_c$.

If fluctuation effects are not included in the Adam–Gibbs model, the configurational fraction $f(T_g)$ at the glass-transition temperature is

$$f(T_g) = \frac{1}{m}\Big(\frac{C}{\ln T_g S_{exc}(T_g)}\Big)\Big(\frac{\Delta C_{\mathrm{P,\,exc}}(T_g)}{S_{exc}(T_g)}\Big) \tag{3.6}$$

Here, $C = \Delta\mu s^* N_{\mathrm{A}}/R$ and $m$ is the steepness index. In deriving eq. 3.6, we made

use of a relation between the steepness index and thermodynamic quantities [182].

## 3.3. Results

The glass-forming liquids *ortho*-terphenyl (OTP), salol, 2-methyltetrahydrofuran (MTHF), *n*-propanol (*n*POH), and 3-bromopentane (BPT) span a wide range of fragilities. For these glass-forming liquids, the temperature dependence of the excess entropy is described by $S_{\rm exc}(T) = S_\infty(1 - T_{\rm K}/T)$, and the corresponding excess heat capacity is inversely proportional to temperature, $\Delta C_{\rm P,exc} = S_\infty(T_{\rm K}/T)$, where $T_{\rm K}$ is the Kauzmann temperature and $S_\infty$ is a parameter that ensures that consistency of the two different modes of relaxation in glass-forming liquids, namely, the Adam–Gibbs model and the empirically fitted Vogel–Fulcher–Tamman model [183]. Substituting these relations in eq. 3.6. we obtain the configurational fraction at the glass-transition temperature

$$f(T_g) = \frac{B}{mT_g \ln(10)\left(1 - \frac{T_{\rm K}}{T_{\rm g}}\right)^2} \tag{3.7}$$

The quantities $T_{\rm K}$, $T_{\rm g}$, and $m$ and the parameter $B = C/S_\infty$ in eq. 3.7 for the five glass-forming liquids [183] are tabulated in Table 3.1. Our prediction of the configurational fraction at the glass-transition temperature for the five glass formers along with the experimental values for the $\beta$ KWW parameters [164, 184] is also shown in Table 3.1. For 1-propanol, and additional Debye peak is observed at a lower frequency than the $\alpha$ process via dielectric measurements. This peak was identified as the $\alpha$ process and incorrectly reported in the literature as $\beta$ KWW value of unity [164]. We have not included in value for 1-propanal in Table 3.1.

**Table 3.1:** Configurational Fraction $f(T_g)$ Computed at the Glass-Transition Temperature for Five Glass-Forming Liquids: *ortho*-terphenyl(OTP), salol, 2-methyltetrahydrofuran (MTHF), *n*-propanol (*n*POH), and 3-bromopentane (BPT).

| substance | $T_K$ (K) | $T_g$ (K) | $m$ | $B$ (K) | $\alpha$ | $f(T_g)$ |
|---|---|---|---|---|---|---|
| *ortho*-terphenyl | 204.2 | 246 | 81 | 684.0 | 0.57 [164], 0.55 [184] | 0.52 |
| salol | 175.2 | 220 | 63 | 823.5 | 0.60 [164], 0.58 [184] | 0.62 |
| MTHF | 69.3 | 91 | 65 | 406.6 | 0.62 [164], 0.64 [184] | 0.53 |
| *n*-propanol | 72.2 | 97 | 35 | 385.6 | 0.6 [184] | 0.77 |
| 3-bromopentane | 82.5 | 108 | 53 | 374.1 | 0.71 [164], 0.63 [184] | 0.51 |

[a] The Kauzmann temperature $T_K$, the glass-transition temperature $T_g$, the steepness index $m$ and the parameter $B$ were compiled from ref [183]. The $\beta$ KWW parameters $\alpha$ were obtained from refs [164, 184]

## 3.4. Discussion

The two basic assumptions [82, 83, 169] in the Adam–Gibbs model are (i) the inversion relationship between the minimum number of molecules $z^*$ that allows cooperative rearrangements and the configurational entropy of the glass-forming liquids, and (ii) the number of molecules inside a cooperative rearranging volume is equal to the minimum number that undergose independent rearrangements. The derivation leading to eq. 3.5 avoids both these assumptions.

We can obtain the configurational fraction by an alternative route. A lower bound to the number of correlated molecules or structural units has been obtained [185] by relating a four-point correlation function to two-point correlation functions $C(t)$ and by the use of fluctuation-dissipation theorem

$$\mathbb{N}_{\text{corr}}(T) = \frac{T^2}{\Delta C_p(T)/k_B}\left\{\max\frac{dC(t)}{dt}\right\}^2 \tag{3.8}$$

Assume that fluctuations in the equilibrium melt are due to configurational entropy $S_c$ and that the time dependence of the correlation function $C(t)$ is of the Kohlrausch–Williams–Watts form, $C(t) = C_o e^{-(t/\tau)^\alpha}$, $0 < \alpha \leq 1$. With these assumptions, we

find the configurational fraction to be

$$f(T_{\rm g}) \approx \frac{1}{\mathbb{N}_{\rm corr}(T)} \frac{(\alpha/e)^2}{\Delta C_{\rm P,\ exc}/k_{\rm B}} \left\{ \frac{{\rm dln}\tau}{{\rm dln}T} \right\} \tag{3.9}$$

The square of the $\beta$ KWW parameter $\alpha$ is inversely proportional to the width of the distribution of relaxation times $\langle \Delta^2 {\rm ln}\tau \rangle$. The essential difference between predictions of the configurational fraction based on eqs. 3.5 and 3.9 is the absence of the nonlinearity factor $(1 - x)$.

For the five glass-forming liquids, accurate estimates of both the number of structural units in a rearranging region and the widths of the distribution of relaxation time at or near $T_{\rm g}$ are not available for quantitative predictions of the configurational fraction based on eq. 3.6, i.e., fluctuation effects in the Adam–Gibbs model.

Moynihan and Schroder [177] have used assumptions (i) in the Adam–Gibbs model along with various approximations, such as that for the temperature dependence of the excess heat capacity and for the temperature dependence of the nonlinearity parameter, to obtain an expression that resembles eq. 3.6 but in which the configurational fraction has been set to unity.

A relationship between the configurational fraction and the *beta* KWW parameter $\alpha$ of glass-forming liquids (Table 3.1) can be unraveled through the coupling model. The basic idea in the coupling model is that the slow $\alpha$ relaxation time $\tau$ is related to the primitive relaxation time $\tau_{\rm p}$ [171]

$$\tau = (t_c^{-(\alpha+1)} \tau_{\rm p})^{1/\alpha} \tag{3.10}$$

Here, $t_c$ is the crossover time from the primitive to cooperative relaxation. For times less than the crossover time, correlation functions decay exponentially [171]. Continuity of the two corelation functions at the crossover time leads to eq. 3.10. On combining the coupling model with the Adam–Gibs model, one obtains the temper-

ature dependence of the relaxation time of the slow mode [171]

$$\tau = (t_c^{-(\alpha+1)} \mathrm{A}\ \mathrm{e}^{C/TS_{\mathrm{exc}}})^{1/\alpha} \tag{3.11}$$

Thus, within the framework of the coupling model, we observe that the configurational fraction at $T_{\mathrm{g}}$ can be identified with the $\beta$ KWW parameter $\alpha$. This prediction of the coupling model ought to be considered as approximate.

Predictions of the configurational fraction based on eq. 3.7 are qualitative estimates due to approximations inherent in the expression used for $S_{\mathrm{exc}}(T)$ and $\Delta C_{\mathrm{P,\ exc}}(T)$ [183]. By taking into account fluctuation effects in the Adam–Gibbs model of cooperative relaxation, we find that the relationship between configurational fraction at $T_{\mathrm{g}}$ and $\beta$ KWW parameter involves various factors such as the steepness index, the nonlinearity parameter, the excess heat capacity, and the number of structural units per mole in a cooperative rearranging region.

It would also be fruitful to generalize out approach to include fluctuations in the potential barrier per molecule to cooperative rearrangements and to explore isothermal and isobaric contributions to the configurational fractions for glass-forming liquids where experimental data is available [176, 186–189]. Unraveling the temperature dependence of the configurational fraction, i.e., generalization of eqs 3.5-3.7, is os importance to the underpinnings that connect excess entropy with kinetic properties of glass-forming liquids. Work is underway along these lines of thought.

## 3.5. Conclusions

The configurational fraction links the excess entropy with kinetic properties described in the Adam–Gibbs model. We have obtained an expression for the configurational fraction at the glass-transition temperature with and without fluctuation effects in the Adam–Gibs model In the latter case, we predict the configurational fraction for

five glass-forming liquids. A connection is observed between the $\beta$ KWW parameter and the configurational fraction at the glass-transition temperature.

# Chapter 4.

# The role of diffuse metal ions in the aa-tRNA accommodation process

## 4.1. Introduction

The ribosome is a large-scale RNA-protein assembly which performs an important role in the protein synthesis in living cells by translating the mRNA codon sequence into the order of amino acid residues. A critical step in the mRNA codon translation is the tRNA selection, which is carried out by the aminoacyl-tRNA (aa-tRNA) together with the mRNA and the conformational rearrangements of the two ribosomal subunits. The speed and accuracy of the translation process significantly alters the quality of the protein synthesis by the ribosome [54, 190].

The tRNA selection in the ribosome is considered to have two distinct steps, which contains the initial selection of aa-tRNA and followed by the proofreading step [191, 192]. At the beginning, aa-tRNAs are delivered to the aminoacyl (A) site of the small ribosomal subunit in a ternary complex with the elongation factor-Tu (EF-Tu) and GTP. During the initial selection, the codon on the mRNA is read by the aa-tRNA anticodon region. Proper base-paring between the tRNA anticodon and the mRNA codon region will form on the A-site of the small subunit during this step, which leads to the GTP hydrolysis and the conformational changes of EF-Tu as well as the aa-tRNA accommodation to the A-site in the large ribosomal subunit (the proofreading step).

Experimental and computational investigations have shown that many factors could influence the tRNA selection process during the ribosome elongation, including the internal effects such as the rate of the GTP hydrolysis [87, 193], the effect of bound $Mg^{2+}$ ions in the ribosome [194], and the thermo fluctuations of the structure [194], as well as external environmental effects such as the concentration of $Mg^{2+}$ ions [190, 195]. In these studies, while many insights was provided into the effects on the initial selection step, less is known about the regulating factors in the proofreading process. The proofreading process starts from a configuration that the aa-tRNA binds to the A site in the small ribosomal subunit and the EF-Tu simultaneously, which

is referred to as "A/T" state (Figure 4.1b), and it ends up with the 3'-CCA end of the aa-tRNA enters the aminoacyl site in the large subunit, which is referred to as "A/A" state (Figure 4.1a). When the aa-tRNA transits from A/T state to A/A state, a few intermediate states were observed for the aa-tRNA conformations [88, 196] (the elbow-accommodated (EA) conformation and the arm-accommodated conformation) and multiple pathways were reported for the accommodation process [88]. Molecular dynamics simulations[63, 91, 197, 198] have been performed to study the conformational dynamics of the aa-tRNA during its transition through these states, while the effects of electrostatics and diffuse ions on these conformational rearrangements remain elusive.

In the aa-tRNA accommodation process, the conformational rearrangement of the tRNA is not a monotonic energetically down-hill process towards the accommodated configuration. Instead, energetic barriers were reported for the transition of aa-tRNA from the A/T to EA configuration due to the steric effects from H89 [63]. In addition, both the aa-tRNA and the H89 are negatively charged, so as the P-site tRNA and 23S rRNA in the large subunit, so that strong electrostatic repulsion is expected along the transition pathway of the tRNA accommodation. Therefore, positively charged metal ions are expected to play significant roles to overcome the electrostatic repulsion and facilitate the conformational rearrangements of aa-tRNA towards the accommodated configuration.

To illustrate electrostatic and ionic effects on the aa-tRNA accommodation process, we performed a series of MD simulations with a simplified model (SMOG-ion). The model was constructed on the basis of an all-atom (non-H) structure-based (SMOG) model [116, 143], the key feature of which is that the energy minimum is defined as the experimentally identified native structure. In the model used here to investigate the ionic effect (SMOG-ion, details in Chapter 2), explicit treatment of diffuse ions are included with the explicit treatment of the electrostatics, while the

solvent molecules are implicitly represented and the solvation effects associated with diffuse ions are described with effective potentials. In the simulations formed in this work, the interatomic interactions within the ribosome assembly specifically stabilizes the A/A configuration, which is the end point of the aa-tRNA accommodation. As discussed below, two sets of MD simulations are conducted in this study, which address the target question from three perspectives: (1) the electrostatic repulsion between aa-tRNA and the ribosome avoids the accommodation process, (2) diffuse metal ions are critical and efficient in reducing the electrostatic repulsion and facilitates the tRNA accommodation, and (3) the accommodation process is sensitive to the $Mg^{2+}$ ion concentrations in the environment.

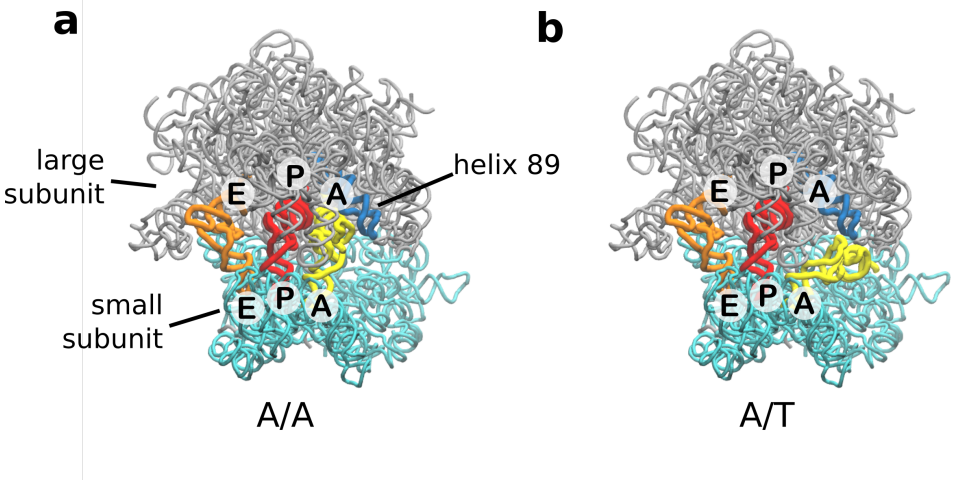


**Figure 4.1:** The start (A/T) and end (A/A) configurations of the aa-tRNA (yellow) accommodation. Helix 89 (dark blue) creates energy barriers along the accommodation pathway due to its steric effects with the aa-tRNA.

## 4.2. Methods

### 4.2.1. Structure Preparation.

The structure used in the simulations is an experimentally resolved *Thermus thermophilus* 70S ribosome (PDB: 6QNR) [59]. This structure was chosen due to its

detailed identification of the chelated ions inside the 70S ribosome, which was critical for the structural integrity of the ribosome and for accurate diffuse ion condensation on the ribosome structure. Three $^{\text{Phe}}$tRNA molecules are resolved in this structure, which are in A/A, P/P and E/E configuration respectively. This structure is considered as the end point configuration of the aa-tRNA accommodation. To obtain the starting point (A/T) of the aa-tRNA accommodation, a crystal structure of the *Thermus thermophilus* 70S ribosome (PDB: 4V5G [199]), which has three tRNA molecules in A/T, P/P and E/E configurations respectively, is used as a reference for structure alignment.

The structure alignment was performed using the VMD software [100]. The phosphorus atoms in the P-site tRNAs of both ribosome structures were selected, and a transformation matrix was calculated using the `measure fit` command in VMD according to the Cartesian coordinates of the phosphorus atoms in both P-site tRNAs. The atoms in the tRNA molecules in assembly 4V5G were transformed by applying the transformation matrix, so that the P-site tRNA in assembly 4V5G overlapped with the P-site tRNA in assembly 6QNR, while the A-site tRNA in assembly 4V5G aligned with the A-site tRNA in assembly 6QNR by the codon-anticodon region. After the transformation, the coordinates of the A-site tRNA in the 4V5G structure provided the reference position of the A/T configurations for the A-site tRNA in the assembly 6QNR. The A-site tRNA in the assembly 6QNR was gradually pulled toward the target A/T configuration in a MD simulation where the rest of the ribosome were held fixed. The aligning and pulling approach provided us a 70S ribosome structure with the aa-tRNA at A/T configuration. This structure was used as the initial structure in all MD simulations discussed in this study.

### 4.2.2. Structure-based (SMOG) model.

All simulations in this study employs a structure-based (SMOG) model variant. The central feature of the SMOG model is that the interatomic interactions are defined based on a preassigned structure. Here, the forcefield for each simulation was constructed such that the potential energy minimum of the ribosome is defined by the A/A conformation of the 70S ribosome (PDB: 6QNR). To avoid unphysical expansion of the ribosome structure, all contact distances were reduced by a factor of 0.96. Since the 3'-CCA contacts only form after the elbow accommodation, they are not included in the current models such that the simulations are focused on the elbow accommodation process. The stabilizing interactions that are unique to the A/A conformations were scaled by a factor of 0.375, similar to previous simulations[63, 88, 198], such that the aa-tRNA elbow can undergo reversible fluctuations between A/T and EA configurations.

Four variants of SMOG model with different treatment of electrostatics and diffuse ions were used in this study:

1. **SMOG**: All ribosome atoms and chelated ions in ribosome are zero-charged. No electrostatics effects or ionic effects are included.

2. **SMOG-coulomb**: Charges are included for ribosome atoms and chelated ions in the ribosome. Charge interactions are described by Coulomb interactions within a cutoff distance of 2 nm. Effects from diffuse ions are not included.

3. **SMOG-DH**: Charges of ribosome atoms and chelated ions are included. Debye-Hückel theory was applied on the electrostatic interactions to account for ionic effect implicitly. No explicit diffuse ions are included.

4. **SMOG-ion**: Charges of ribosome atoms and chelated ions are included. Explicit diffuse ions are included with corresponding charges assigned for each ion. Explicit electrostatics are included for the entire system.

Detailed descriptions of the models above are described in Chapter 2 and Appendix B.

### 4.2.3. Molecular Dynamics Simulations.

In this study, MD simulations were performed with the Gromacs-5.1.4 software package [107]. The temperature of the simulations were set to 60 Gromacs reduced units (0.5 reduced units in SMOG), which corresponds to 300 K in explicit solvent simulations. At the beginning of each simulation, the ribosome structure (A/T configuration) was placed in the center of a cubic box with length of 70 nm. In the cases without explicit representations of diffuse ions (SMOG, SMOG-coulomb, and SMOG-DH (see Chapter 2 for details)), four replicas of simulations were set for each model and all trajectories are used for statistical analysis. In the case with explicit treatment of diffuse ions (SMOG-ion) model, $Mg^{2+}$, $K^+$, and $Cl^-$ ions were added to the simulation box in addition to the ribosome structure. The number of ions and the target bulk concentration are listed in Table 4.1. At each concentration, the system was equilibrated for $1.5 \times 10^7$ timesteps with 0.002 reduced unit for each step. During the equilibration run, the ribosome structure was held fixed while the diffuse ions were allowed to move. Then a snapshot of the atoms and ion coordinates from equilibrated system was taken and set as the starting point of the production run. Five to nine replicas were set at each concentration in the production runs.

**Table 4.1:** Number of ions in the simulation system and the corresponding target bulk concentration.

| $Mg^{2+}$ conc.(mM) | $K^+$ conc.(mM) | num. of $Mg^{2+}$ | num. of $K^+$ | num. of $Cl^-$ |
|---|---|---|---|---|
| 0 | 100 | 0 | 23000 | 20226 |
| 1 | 100 | 500 | 20000 | 18226 |
| 5 | 100 | 1800 | 20000 | 20826 |
| 10 | 100 | 3500 | 20000 | 24226 |

## 4.3. Results and Discussion

To illustrate the effects from electrostatics interactions and diffuse ions on the aa-tRNA accommodation process, we made use of the MD simulations of the ribosome structure with different treatment of electrostatic and ionic effects using variants of SMOG model. Due to the complexity of the system, control experiments were carefully designed such that the effect of each factor (i.e., electrostatics, explicit representations of diffuse ions, ionic concentrations) could be clearly shown from the comparison. In addition to the comparison of simulation trajectories, statistical analysis was made for the conformational distribution, which reveals the impact of diffuse metal ions on the aa-tRNA conformations and its tendency in accommodation process.

### 4.3.1. Electrostatics and diffuse ions control aa-tRNA accommodation.

Comparison between simulations of aa-tRNA accommodation process with SMOG, SMOG-coulomb, SMOG-DH, and SMOG-ion model allowed us to incrementally understand the strength of the electrostatic repulsions between aa-tRNA and the ribosome on prohibiting the accommodation process, as well as the necessity of explicit treatment of diffuse ions in the investigations of ionic effects during this process. To quantify the transition of the aa-tRNA, a reaction coordinate, $R_{\text{elbow}}$, was used to measure the distance between the aa-tRNA and the P-site tRNA[197]. $R_{\text{elbow}}$ was defined to be the distance between the O3' atom in the residue U60 from the aa-tRNA and the O3' atom in the residue U8 from the P-site tRNA (Figure 4.2). When the aa-tRNA is in the A/T state, $R_{\text{elbow}}$ adopts a relatively high value $\sim$ 5–6 nm, while $R_{\text{elbow}}$ decreases to relatively small values ($\sim$ 2–3 nm) when the elbow of the aa-tRNA is accommodated.

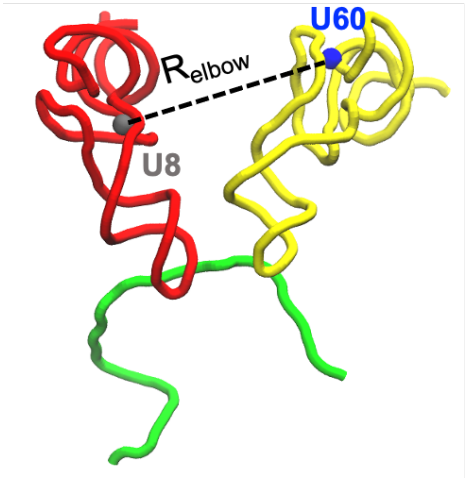


**Figure 4.2:** Defining the reaction coordinate $R_{\text{elbow}}$. It was calculated by the distance between O3' atom from the residue U60 on the aa-tRNA (yellow) and the O3' atom from the residue U8 on the P-site tRNA (red). The mRNA is shown in green.

In simulations with electrostatic-free and ion-free SMOG model, spontaneous and reversible transitions between the A/T and EA configurations of the aa-tRNA were observed from the simulation trajectories. One of the trajectories with SMOG model in this study is shown in Figure 4.3a in terms of the $R_{\text{elbow}}$ distance as a function of time. 7 barrier crossing events can be identified from the sample trajectory. The energy barrier is due to the steric effects from the helix 89 (H89) (Figure 4.1, dark blue helix) on the transition pathway between A/T and EA configurations of the aa-tRNA [63, 88, 91, 197]. The overall distribution of $R_{\text{elbow}}$ distance in the SMOG model simulation shows that the EA configuration is more preferred. With the electrostatics included for the ribosome structure (SMOG-coulomb model), only A/T ensemble of the aa-tRNA was observed and the elbow of aa-tRNA did not accommodate. This can be rationalized to the strong electrostatic repulsions between negatively charged aa-tRNA with the H89, P-site tRNA or A-site in the large-subunit of the ribosome. Since the anticodon region had base pairing with the codon on the mRNA and was stabilized by the intermolecular interactions defined in the forcefield, the aa-tRNA

did not dissociate from the ribosome. This intermolecular electrostatic repulsion is common in highly charged biopolymers, such as RNA and DNA. In addition to treating the electrostatics with direct Coulomb interactions, we also attempted to apply Debye-Hückel treatment on the electrostatics with SMOG-DH model to account for the screening effects from diffuse ions implicitly. However, the elbow-accommodation was not observed and the aa-tRNA remained in A/T ensemble (Figure 4.3b). By construction, while it is appropriate to use DH treatment for monovalent ions to describe interactions between opposing charges [138, 139], it cannot capture the ion-induced attractions between polyanionic molecules [140]. To address this shortcoming, we used SMOG-ion model (Figure 4.3c) to included the diffuse ions explicitly and repeated the simulation in the environment of $\sim$ 10 mM $Mg^{2+}$ ion and $\sim$ 100 mM $K^{+}$ ions. The simulation trajectories showed that the aa-tRNA, which started with A/T state, quickly accommodated into the EA configurations in all simulation replicas. No reversible transitions back to A/T state was observed in the presence of explicit diffuse ions.

Since the SMOG model and SMOG-coulomb model only differ in the presence of explicit electrostatic interactions, the comparison of the simulation trajectories from these models indicated that the electrostatic repulsions between the aa-tRNA and the ribosome create notable difficulties in achieving elbow accommodation. The simulation trajectories from the SMOG-ion model show that diffuse ions not only effectively reduced the electrostatic repulsion associated with the aa-tRNA, but also help to stabilize the elbow accommodated configuration.

### 4.3.2. $Mg^{2+}$ ions are critical for aa-tRNA accommodation.

In comparison with the monovalent $K^{+}$ ion, divalent $Mg^{2+}$ ions are smaller in size and hold higher charge density, which are also found much more efficient in reducing repulsions between negatively charged RNA phosphates. For example, millimo-

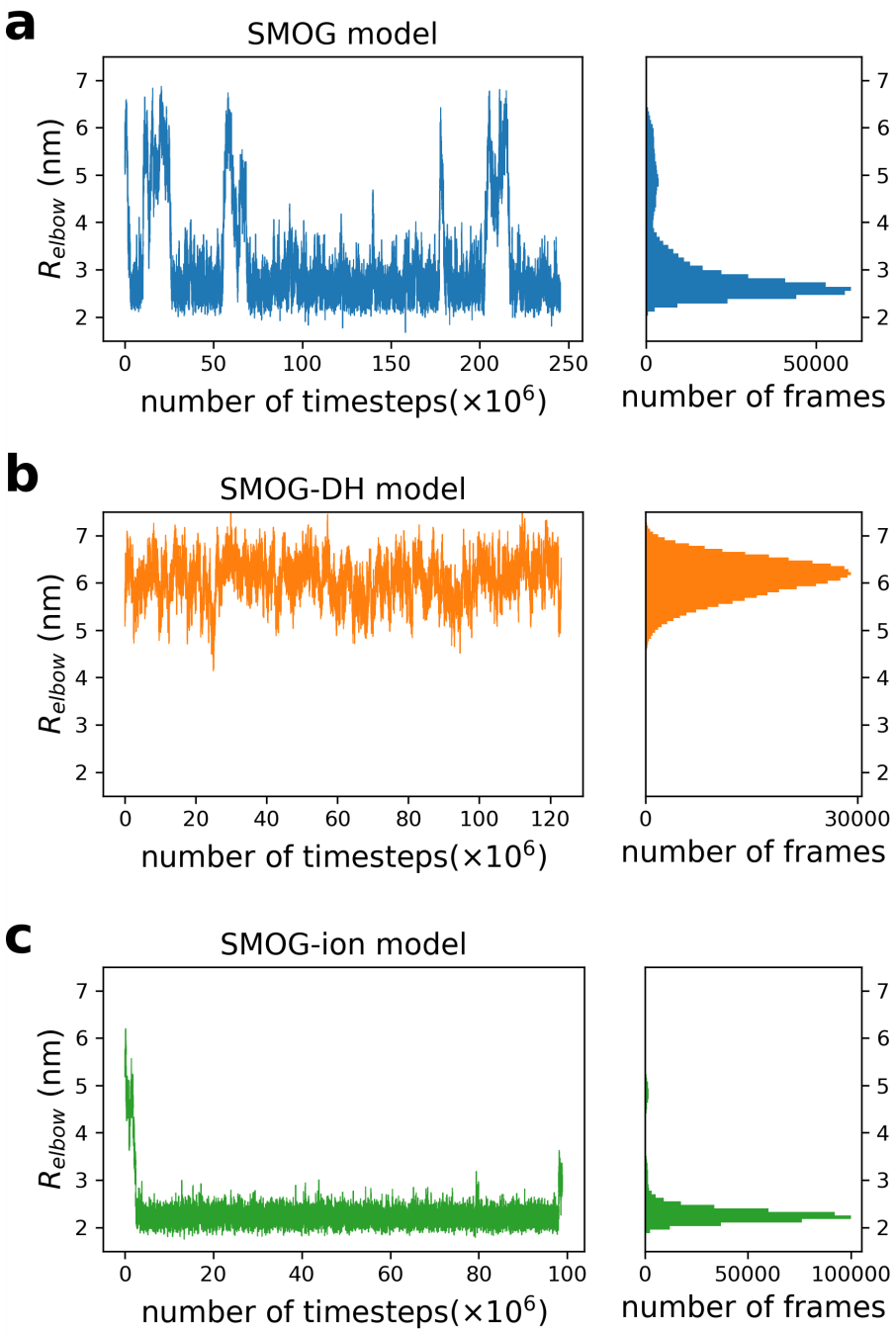


**Figure 4.3:** MD simulation trajectories demonstrated the electrostatic and ionic effects on the aa-tRNA accommodation process. Left panel shows the single trace of $R_{\text{elbow}}$ from one replica and the right panel shows the distribution of $R_{\text{elbow}}$ from all replicas of simulations with the same model. (a) Single trace of $R_{\text{elbow}}$ from one replica of simulations with SMOG model. aa-tRNA performed spontaneous and reversible transition between A/T and EA configurations. The distribution of $R_{\text{elbow}}$ shows the preference of EA state. (b) Single trace of $R_{\text{elbow}}$ from one replica of simulations with SMOG-DH model. aa-tRNA stayed in EA configurations. (c) aa-tRNA quickly transit from A/T state to EA state. Backward transition to A/T state was not observed.

lar concentrations of $Mg^{2+}$ ions are able to stabilize RNA structures that are only marginally stable at high monovalent metal ion concentrations [3, 5]. To illustrate the effect of $Mg^{2+}$ concentrations on the aa-tRNA accommodation process, a series of simulations were performed at fixed $K^+$ concentration ($\sim$ 100 mM) while different $Mg^{2+}$ concentrations.

MD simulations of the ribosome were performed with 4 concentrations of $Mg^{2+}$ ions in the presence of 100 mM $K^+$ ions. As discussed in the previous section, in the environment of 10 mM $Mg^{2+}$ ions and 100 mM $K^+$ ions, aa-tRNA performed elbow-accommodation shortly after the simulation started and stayed at EA configuration instead of returning to A/T state. When the $Mg^{2+}$ concentration was lowered to 5 mM (Figure 4.4a), aa-tRNA performed similarly to the 10 mM case, which strongly preferred EA configuration and no reversible transition was observed. When the $Mg^{2+}$ concentration was further lowered to 1 mM (Figure 4.4b), though the EA configuration was strongly preferred, backward transition to A/T started to appear in these simulations. In comparison with the 5 mM and 10 mM cases, the simulation at 1 mM $Mg^{2+}$ concentration indicated that the stabilizing effect from $Mg^{2+}$ ions on the EA configuration is related to the ion concentration and the tendency of conformational changes could be altered with the a few millimolar change of $Mg^{2+}$ concentrations. When the concentration of $Mg^{2+}$ ions was reduced to zero and only monovalent ions ($K^+$ and $Cl^-$) were present (Figure 4.4c), the aa-tRNA mainly stayed in the A/T configuration. In the simulations without $Mg^{2+}$ ions, several attempts were made to achieve elbow-accommodation, while the aa-tRNA were only maintained at the EA state for a few million second and quickly returned to A/T state. In comparison with the simulations at 1 mM $Mg^{2+}$ concentration, the $Mg^{2+}$-free case revealed the significance and efficiency of $Mg^{2+}$ ions in terms of stabilizing the accommodated configurations of aa-tRNA.

Summarizing the simulations with various concentrations of $Mg^{2+}$ ions, it clearly

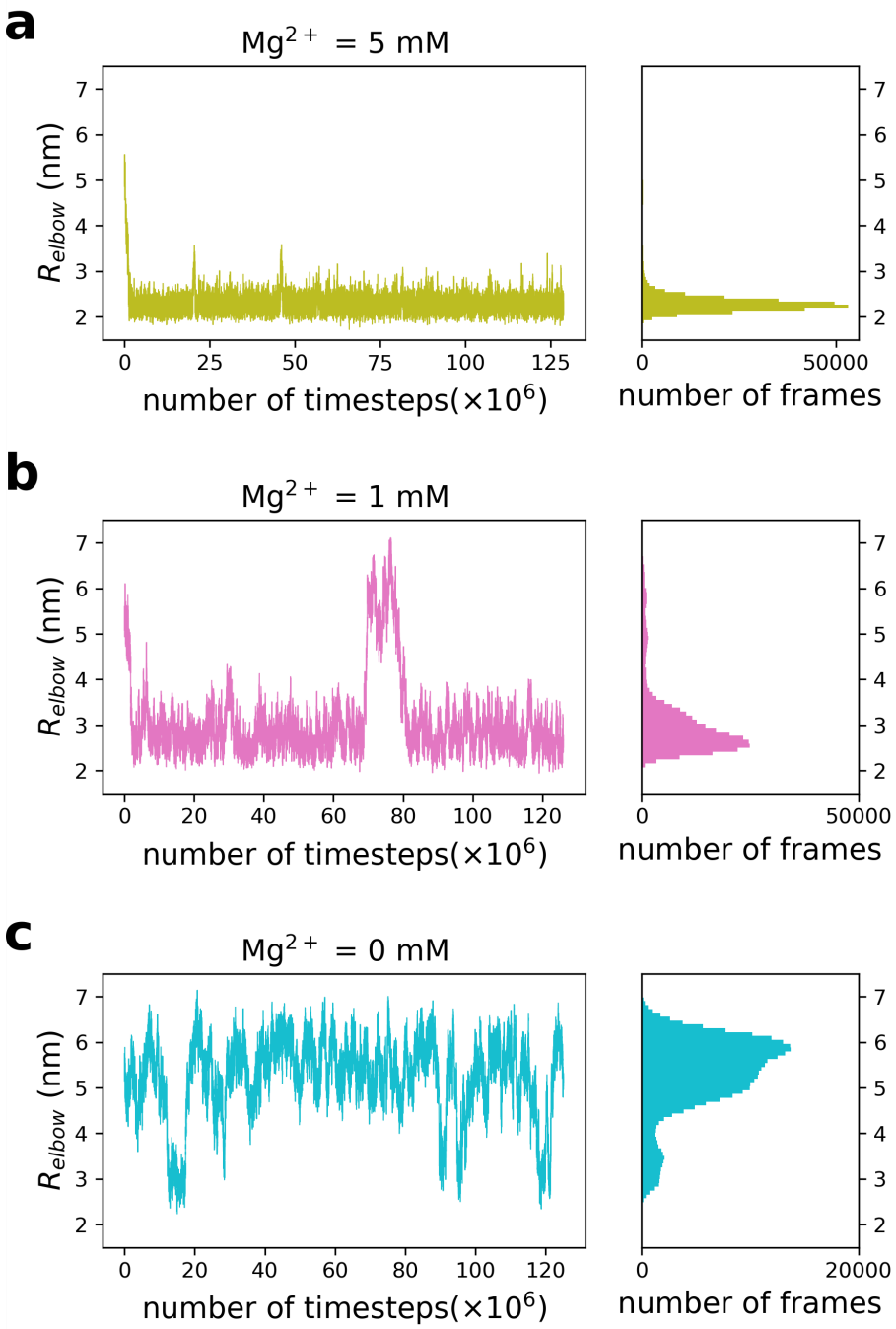


**Figure 4.4:** MD simulation trajectories demonstrated the dependence of aa-tRNA accommodation process on the $Mg^{2+}$ ion concentrations. Left panel shows the single trace of $R_{\text{elbow}}$ from one replica and the right panel shows the distribution of $R_{\text{elbow}}$ from all replicas of simulations with the same model. (a) With 5 mM of $Mg^{2+}$, aa-tRNA transits from A/T state to EA state shortly after the simulation starts. No backward transition was observed. (b) With 1 mM of $Mg^{2+}$, aa-tRNA is able to perform spontaneous and reversible transformation between A/T and EA state. The EA conformation is preferred. (c) With only monovalent ions, aa-tRNA was able to transit into EA state, while not stabilized at the accommodated state. A/T conformations are preferred under this condition.

shows that $Mg^{2+}$ ions are necessary in stabilizing the elbow-accommodated configurations of aa-tRNA. The simulations with 100 mM explicit $K^+$ ions and no $Mg^{2+}$ ions indicated that $K^+$ ions could facilitate the elbow-accommodation process. In comparison with SMOG-DH simulation described above, the $Mg^{2+}$-free simulations also revealed the significance of explicit treatment of diffuse ions in the investigations of ion-induced conformational changes.

## 4.4. Conclusion

As a significant step in the tRNA selection stage during the ribosome elongation cycle, the aa-tRNA accommodation is critical to the protein synthesis conducted by the ribosome. The role of diffuse ions in the aa-tRNA accommodation has remained elusive. While notable effort was made to reveal the mechanism and the regulator of the aa-tRNA accommodation process, investigating the ionic effects on this process remains challenging. As a step to this end, we employed an all-atom structure-based model with explicit treatment of electrostatics and diffuse ions and implicit treatment of solvent molecule in the MD simulations of a full ribosome structure. By comparing simulations with and without electrostatic interactions for the ribosome, we find that the electrostatic repulsion between aa-tRNA with the ribosome is able to prohibit the elbow-accommodation process of the tRNA. From a series of simulations with various concentrations of $Mg^{2+}$ ions at fixed $K^+$ concentration, we identify the critical roles of $Mg^{2+}$ ions in stabilizing the accommodated configurations. This study provided significant insights into the understanding of the role of metal ions in the conformational rearrangements of the ribosome. It will be fruitful to apply similar strategies to investigate the ionic effects in the conformational dynamics of the complex biomolecular assemblies.

# Appendix A.

# Appendix of Chapter 1: RNA as a Complex Polymer with Coupled Dynamics of Ions and Water in the Outer Solvation Sphere

## A.1. Structure of the RNA

## A.2. Preferential Interaction Coefficients

**Table A.1:** Preferential interaction coefficients from the explicit solvent MD simulations. Stated errors are computed from the variation between runs.

| i | $z_i$ | $N_i$ | $[i]^*$ / mM | $[i]$ / mM | $\Gamma_{2+}$ |
|---|---|---|---|---|---|
| K | 1 | 97 | 149.7 ± 0.6 | 146.4 ± 0.6 | +31.4 ± 0.3 |
| Cl | -1 | 59 | 144.2 ± 0.2 | 147.4 ± 0.2 | -7.0 ± 0.1 |
| Mg | 2 | 10 | 0.6 ± 0.2 | 0.5 ± 0.2 | +9.8 ± 0.1 |

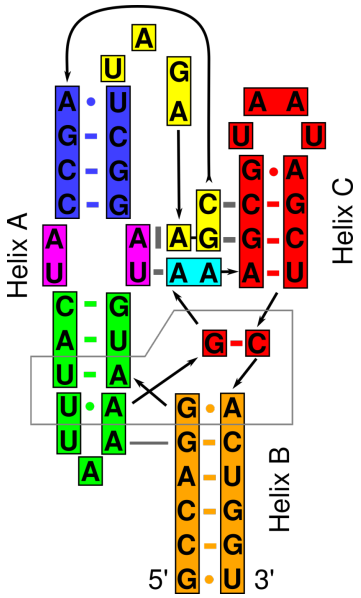


**Figure A.1:** Cartoon of the 58-mer RNA structure: Helices A,B,C are shown in green & blue, orange and red, respectively, and connecting elements are colored magenta, cyan and yellow, as in Fig. 1.1A. Helix base pairing contacts are shown in corresponding colors. The tertiary base pairing HB interactions and one base stacking interaction are shown as black lines.

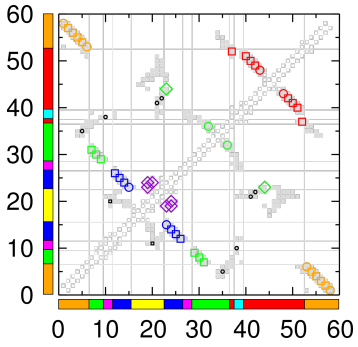


**Figure A.2:** Contact map of the 58-mer RNA structure. All base-base interactions shown in Fig. A.1 are marked in the contact matrix in corresponding colors. The tertiary BP interactions are represented by smaller black circles. Green and purple diamonds show base- base interactions that are mediated by the chelated Mg and K ions, respectively. Grey squares in the background show contacts determined based on structural proximity in the native crystal structure, using the SMOG2 tool.

## A.3. Coordination by Ions and Water

**Table A.2:** Lifetimes (and standard deviations of lifetimes) for coordination bonds by $Mg^{2+}$ and $K^+$ and for hydrogen bonds by water with different RNA atoms, in ps.

| atoms | $Mg^{2+}$ | $K^+$ | water |
|---|---|---|---|
| O1P | 167 (2300) | 46 (430) | 34 (61) |
| O2P | 311 (3000) | 50 (370) | 51 (176) |
| O5' | 43 (1000) | 34 (110) | 16 (17) |
| O3' | 79 (1800) | 50 (190) | 13 (11) |
| O2' | 121 (1400) | 90 (420) | 15 (9) |
| O4' | 28 (80) | 47 (390) | 12 (8) |
| triPO | 175 (864) | 56 (120) | 49 (98) |
| Base | 152 (2100) | 101 (420) | 23 (37) |
| all | 141 (1900) | 60 (360) | 26 (74) |

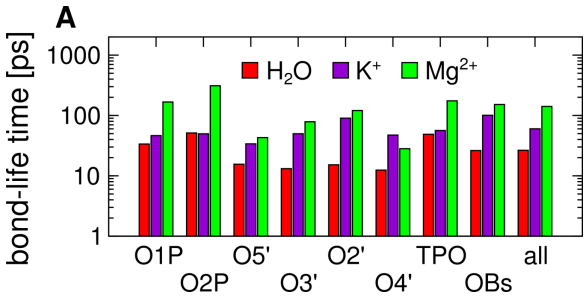


**Figure A.3:** Average bond life times between Mg, K or water and different species of RNA atoms. The same data are shown in different form in Fig. 1.5A and in Tab. A.2.

## A.4. Coordination Sites

## A.5. Principal Component Analysis

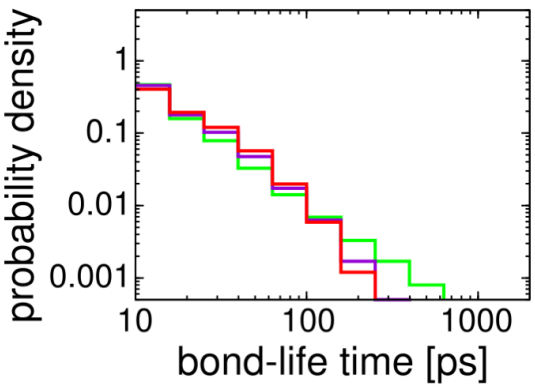


**Figure A.4:** Distribution of bond life-times, for coordination bonds between Mg, K or water and RNA atoms.

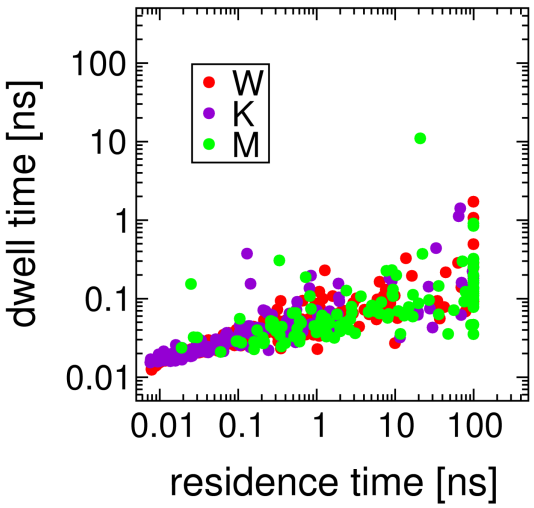


**Figure A.5:** Comparison of residence times and dwell times. Bond residence times, for which histograms are shown in Fig. 1.2D, are based on position correlation functions, and allow excursions from the site during residences. Dwell times are instead defined as times spent continuously in the site without any excursion.

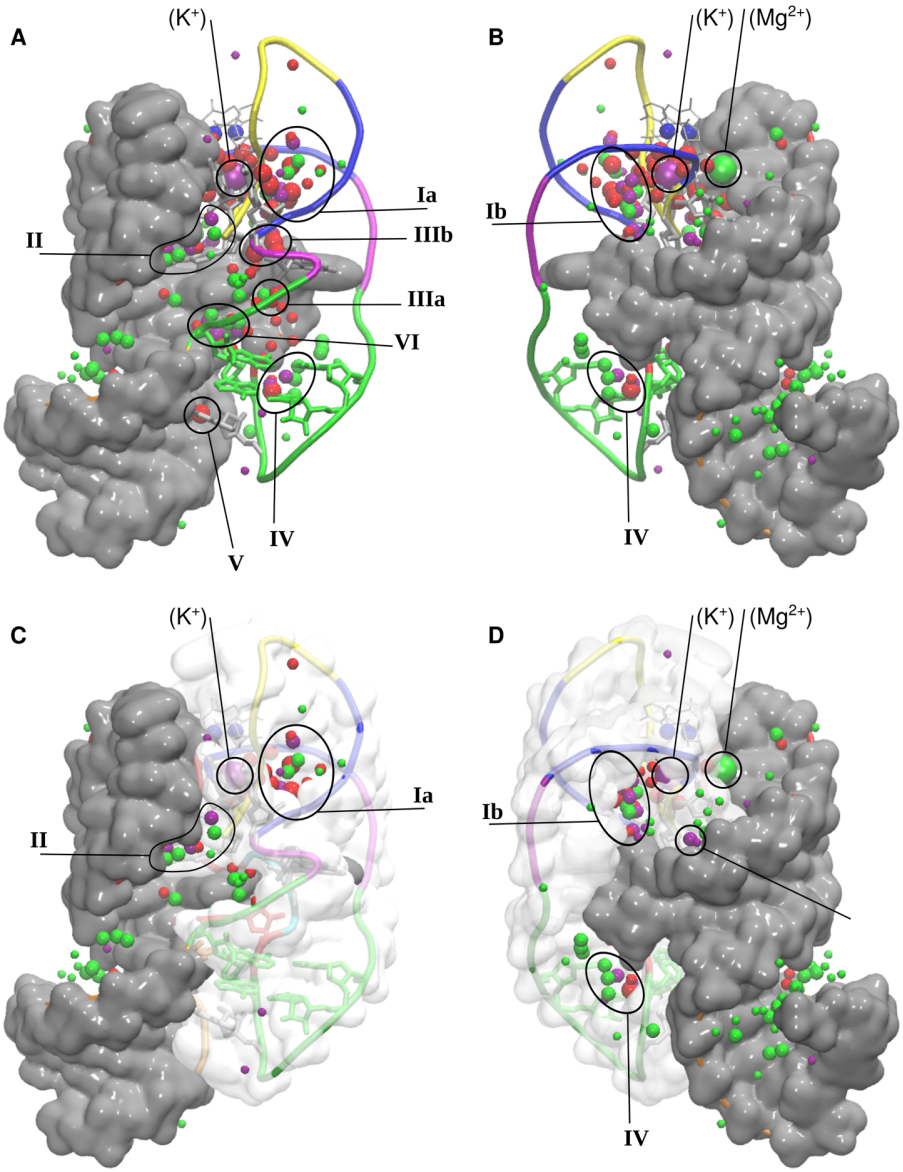


**Figure A.6:** Alternative views of slow sites around the RNA structure. Mg, K and water sites are shown as green, purple and red spheres, respectively. Sites with residence times >=10 ns are shown as larger spheres, sites with shorter residence times >=1 ns are shown as smaller spheres. The backbone of the RNA is colored as in Fig. 1.1 and A.1. The chelated Mg & K are shown as large green & purple spheres. A solid outer surface is shown in grey for helices B & C. Helix A is shown only as a tube in panels (A) & (B), which show corresponding views from opposite sides. Panels (C) & (D) repeat the same views with helix A also covered by a translucent white surface. The placement of sites at the interface between helices A & B/C and in the vicinity of the chelated ions is visible. Roman numerals label locations discussed in the text.

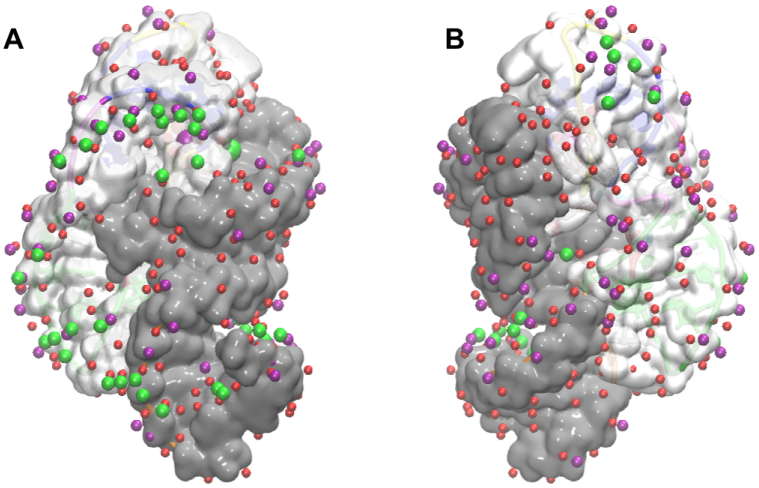


**Figure A.7:** Placement of sites with short residence times on the RNA structure. The outer surface of helix A is shown in white, the surface of helices B & C in grey. Sites for Mg, K and water are shown as green, purple and red spheres. Residence times are <1 ns for the shown Mg sites, <100 ps for the K and water sites. (A, B) show views from opposite sides of the molecule.

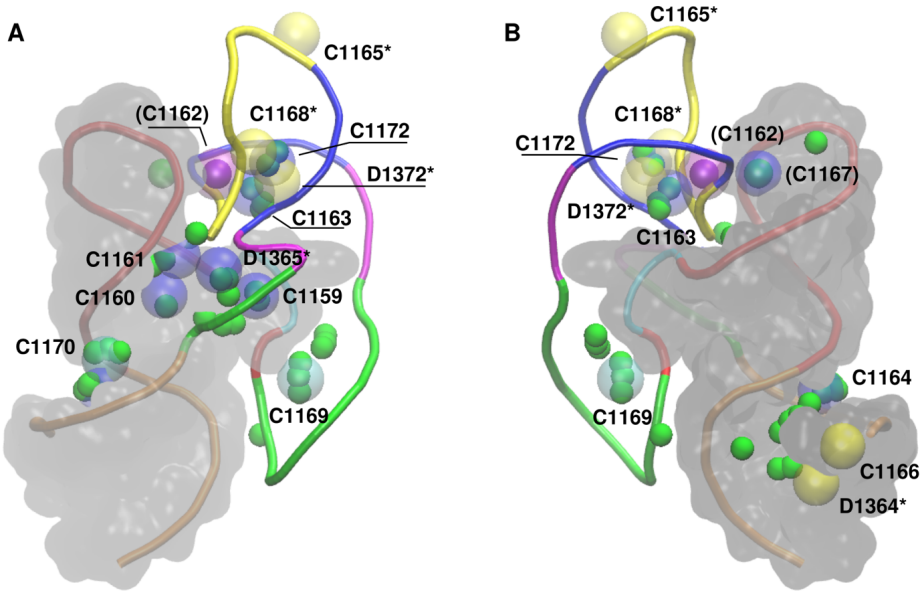


**Figure A.8:** Comparison of $Mg^{2+}$ sites from the simulations to ion positions from the x-ray structures. Simulation sites, shown as green spheres, have residence times for $Mg^{2+} \geq 10$ ns. Unmatched x-ray positions are shown yellow, x-ray positions with matching simulation sites are shown in blue. Lighter blue indicates two positions taken up by Osmium ions in the crystal. Labels give chain identifiers and residue numbers from the PDB. The two chelated $Mg^{2+}$ and $K^{+}$ ions are given in brackets. Asterisks mark ions that are only present in one of the two copies of the structure in the PDB.

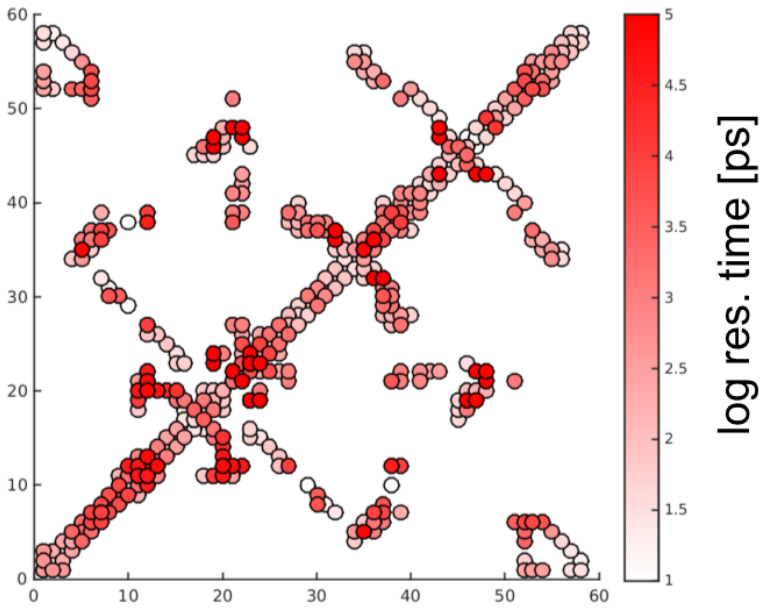


**Figure A.9:** Water-mediated contacts, showing the site with the longest residence time for each pair of RNA bases.

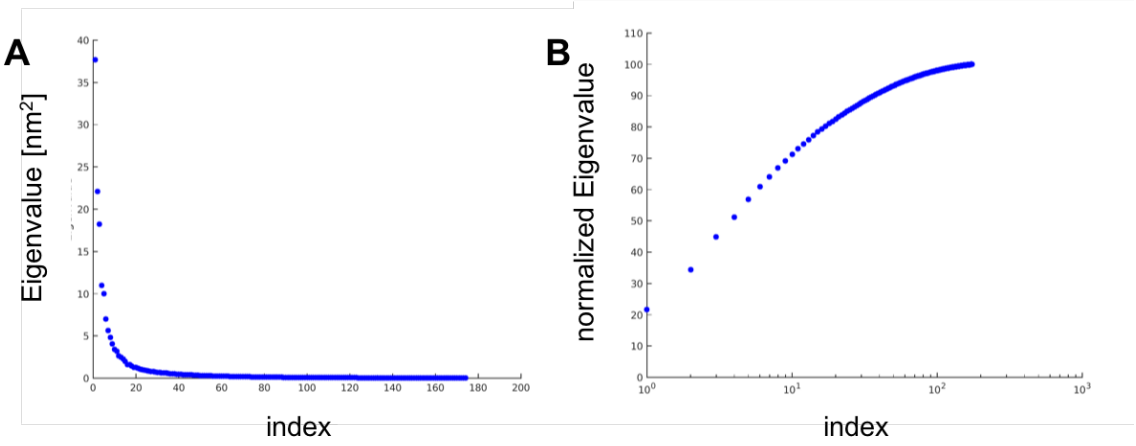


**Figure A.10:** Eigenvalues from the PCA of the RNA P atom positions. (A) individual eigenvalues, (B) normalized cumulative sum of all eigenvalues.

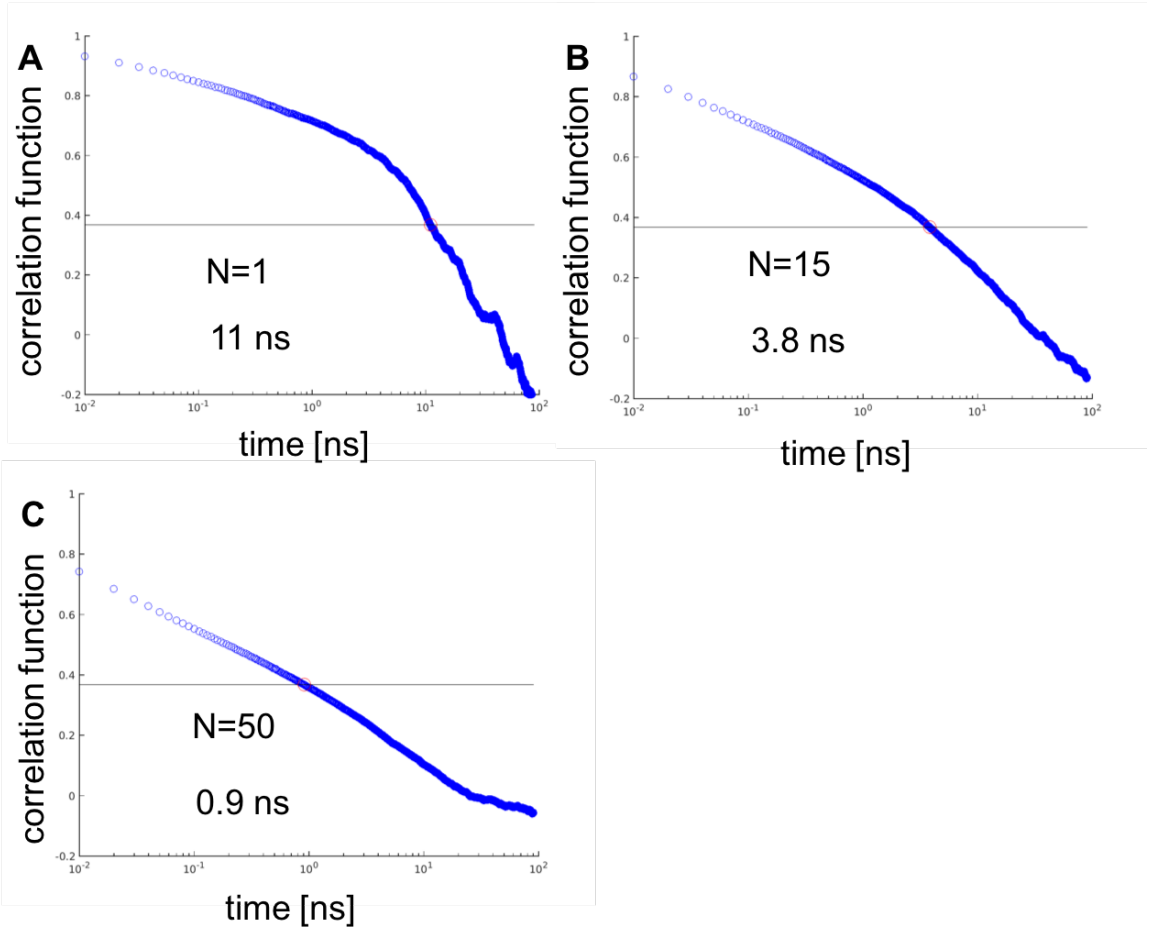


**Figure A.11:** Autocorrelation functions for the first N principal components of the RNA dynamics. (A) N=1, (B) N=15, (C) N=50. Autocorrelation times, determined where the function equals 1/e, are given in ns.

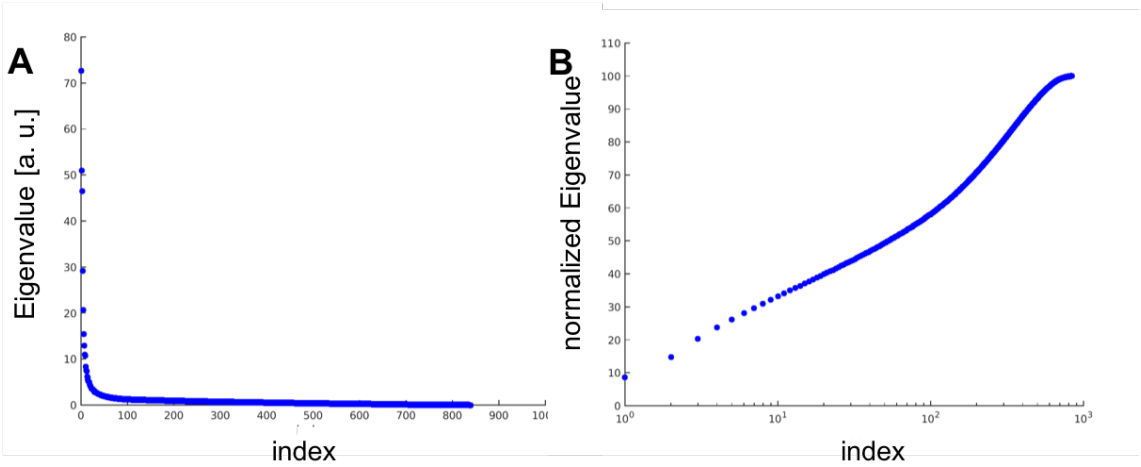


**Figure A.12:** Eigenvalues for the joint PCA of RNA P positions and Mg, K and water site occupations. (A) individual eigenvalues. (B) normalized cumulative sum of the eigenvalues.

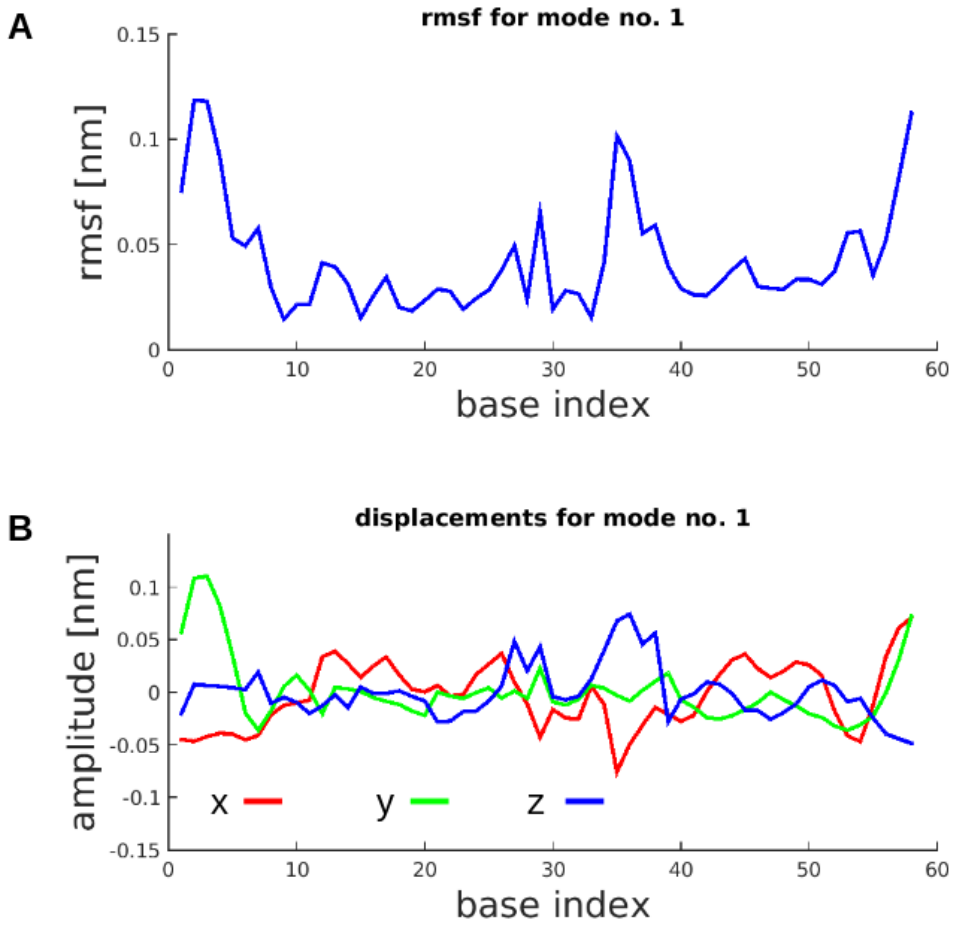


**Figure A.13:** Motions of RNA P atoms due to the first joint PC. (A) root mean square fluctuations of P atoms around the mean structure. (B) displacement components of the eigenvector in x, y, z directions, (scaled with the square root of the eigenvalue).

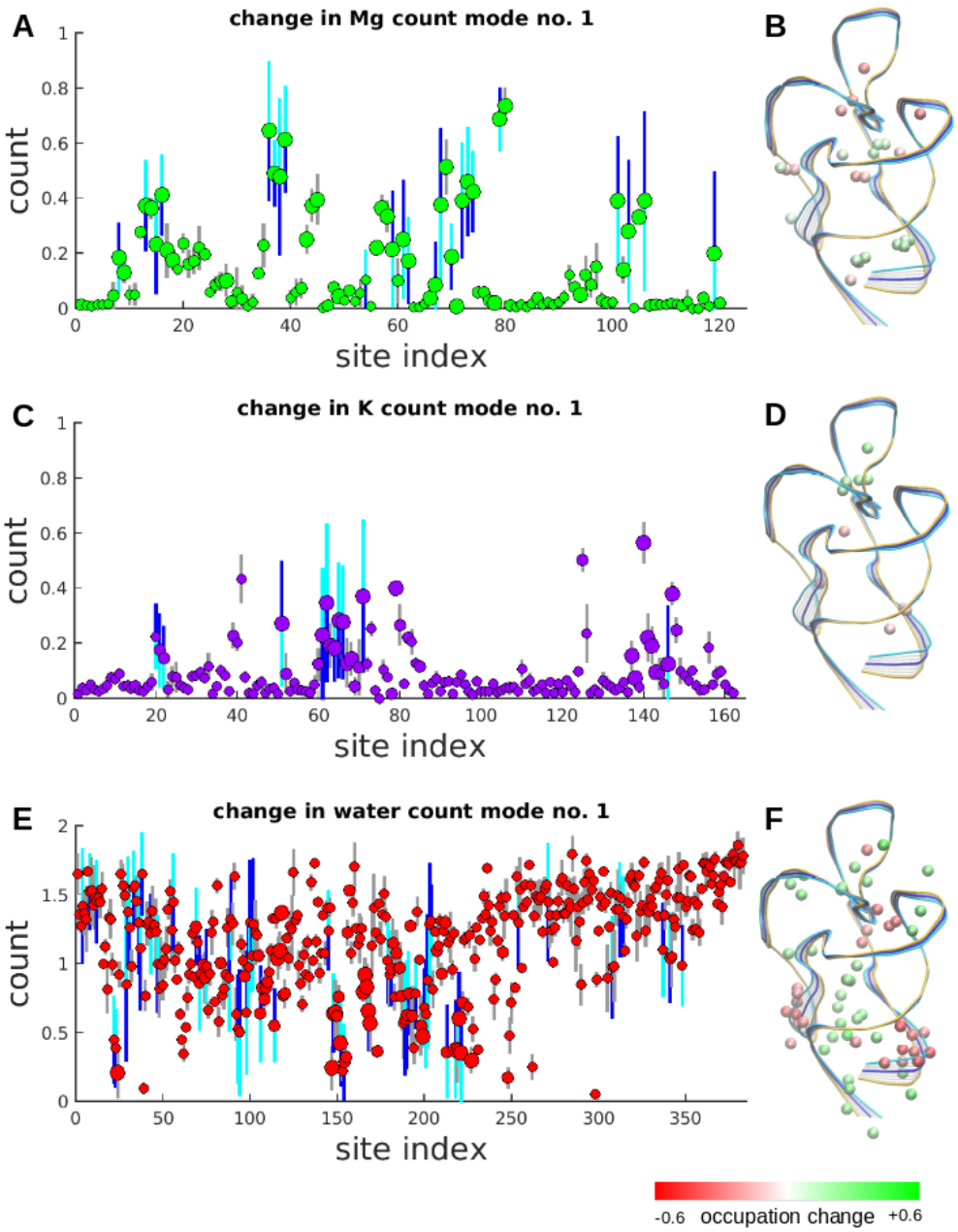


**Figure A.14:** changes in site occupation counts due to the first joint PC. (A) changes in occupation of Mg sites. Circles mark the mean occupations of sites, vertical bars indicate the changes in occupation count due to the first PC. Changes <0.1 ions are drawn in grey, larger changes are highlighted in blue and cyan, where cyan corresponds to the positive sign of the PC. Symbol size encodes residence times t>=10 ns, 10 ns > t >= 1ns and t < 1ns respectively. (B) sites with occupation changes >= 0.1 ions are shown in relation to the RNA structure. Occupation changes are color coded, with green positive and red negative. The blue trace gives the average RNA structure. Displacements of the RNA P atoms due to the first PC, scaled 2x, are given by cyan (positive) and yellow tubes (same as Fig. 1.3A) (C, D) the same information for potassium sites. (E, F) changes in occupation of water sites, highlighting changes >= 0.25.

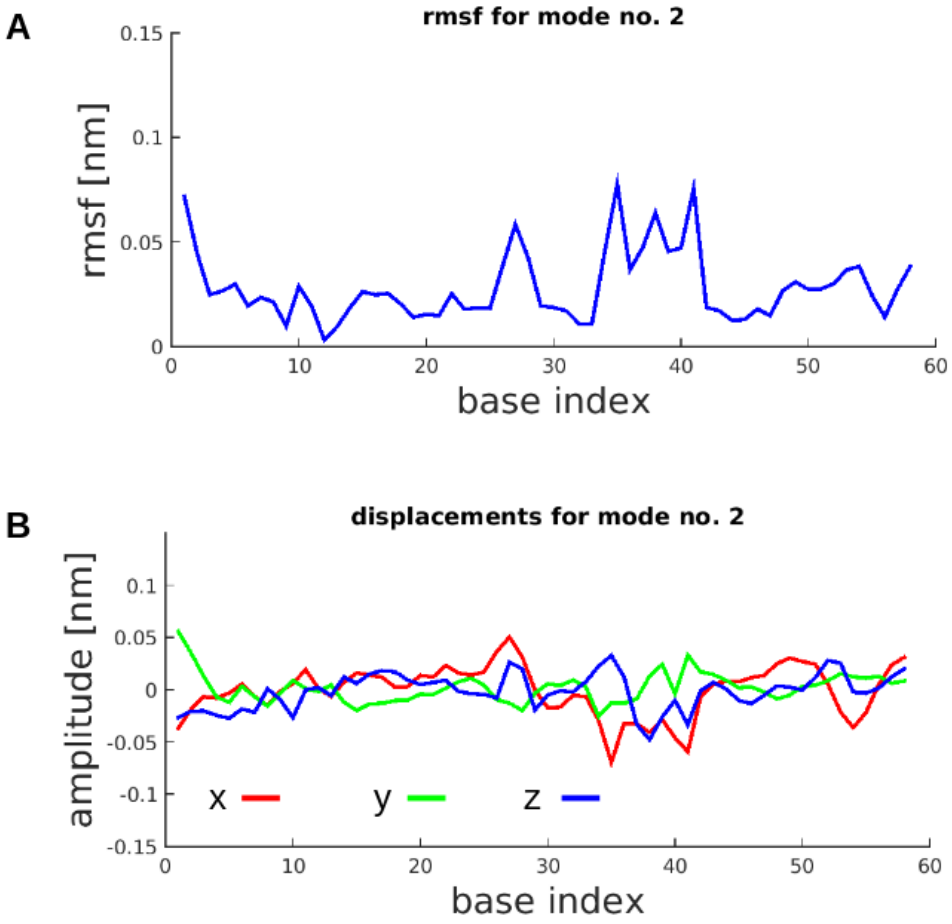


**Figure A.15:** Motions of RNA P due to joint mode 2. (same as Fig. A.13)

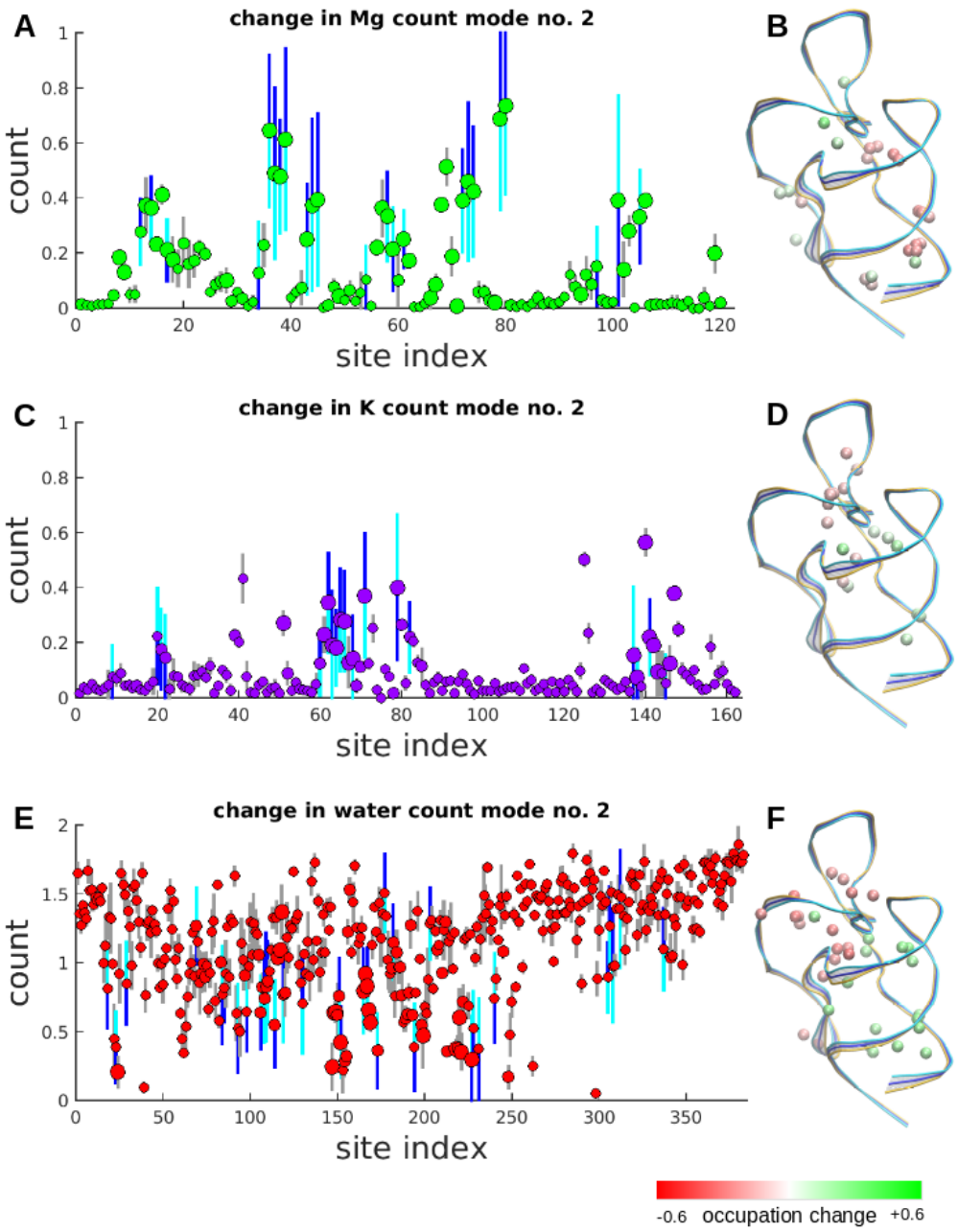


**Figure A.16:** Changes in site occupation counts due to joint PC 2. (same as Fig. A.14)

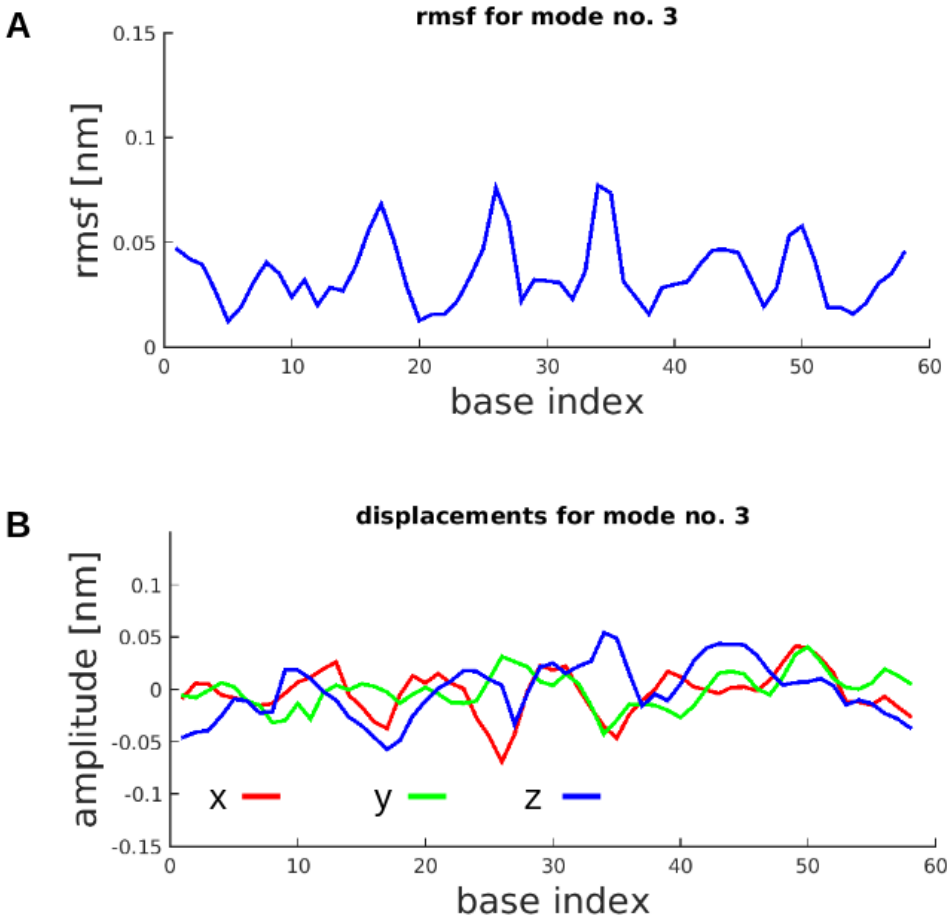


**Figure A.17:** Motions of RNA P due to joint mode 3. (same as Fig. A.13)

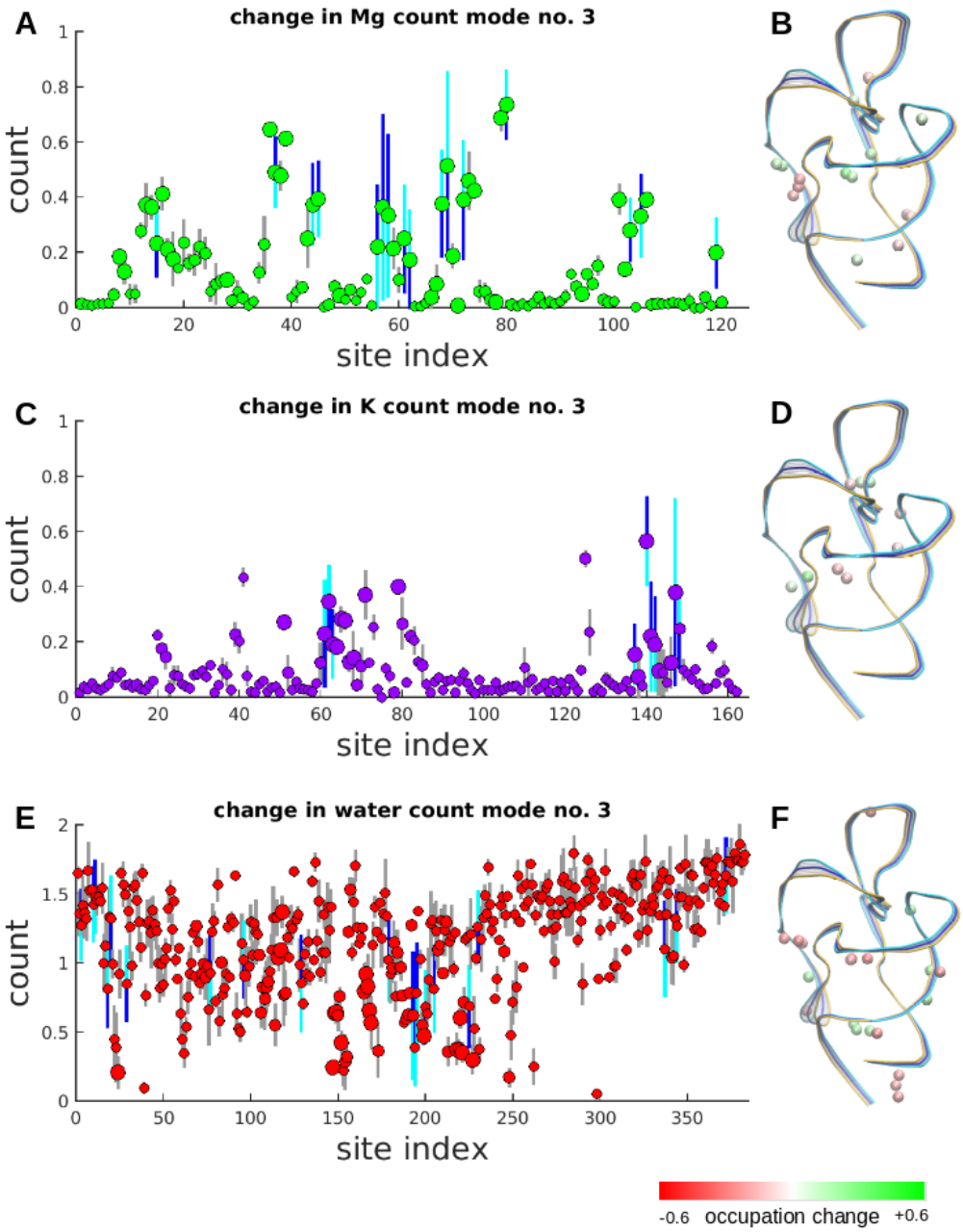


**Figure A.18:** Changes in site occupation counts due to joint PC 3. (same as Fig. A.14)

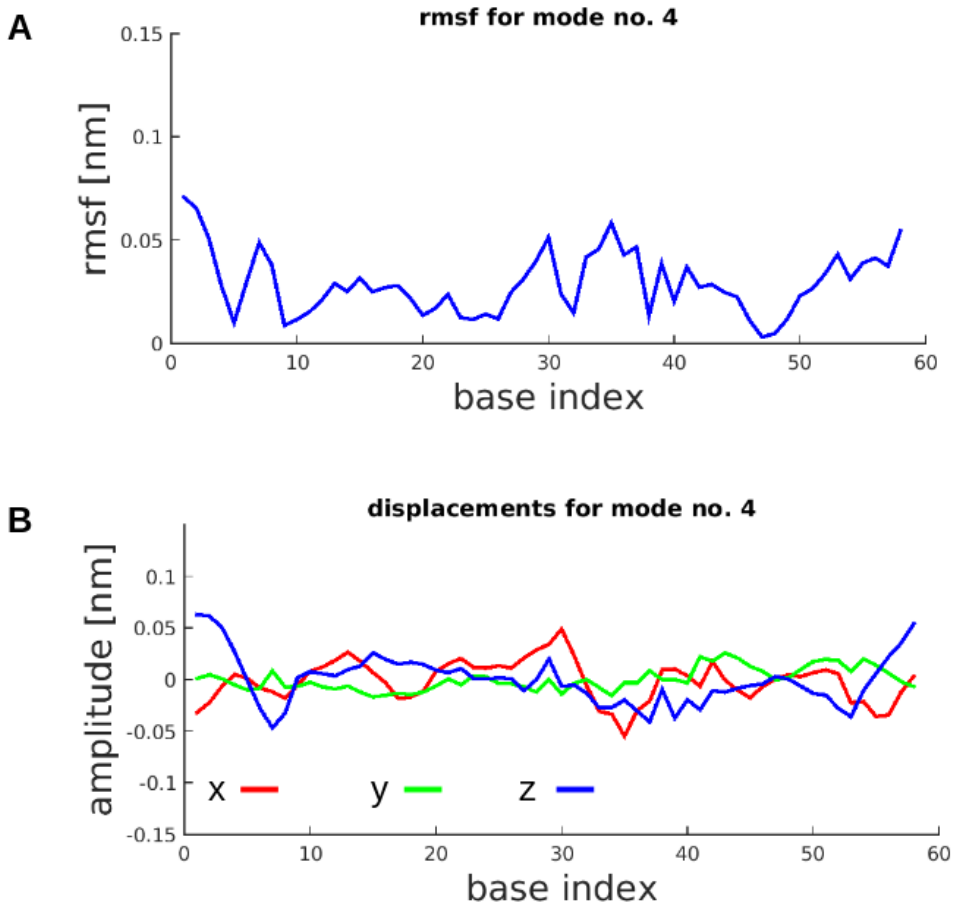


**Figure A.19:** Motions of RNA P due to joint mode 4. (same as Fig. A.13)

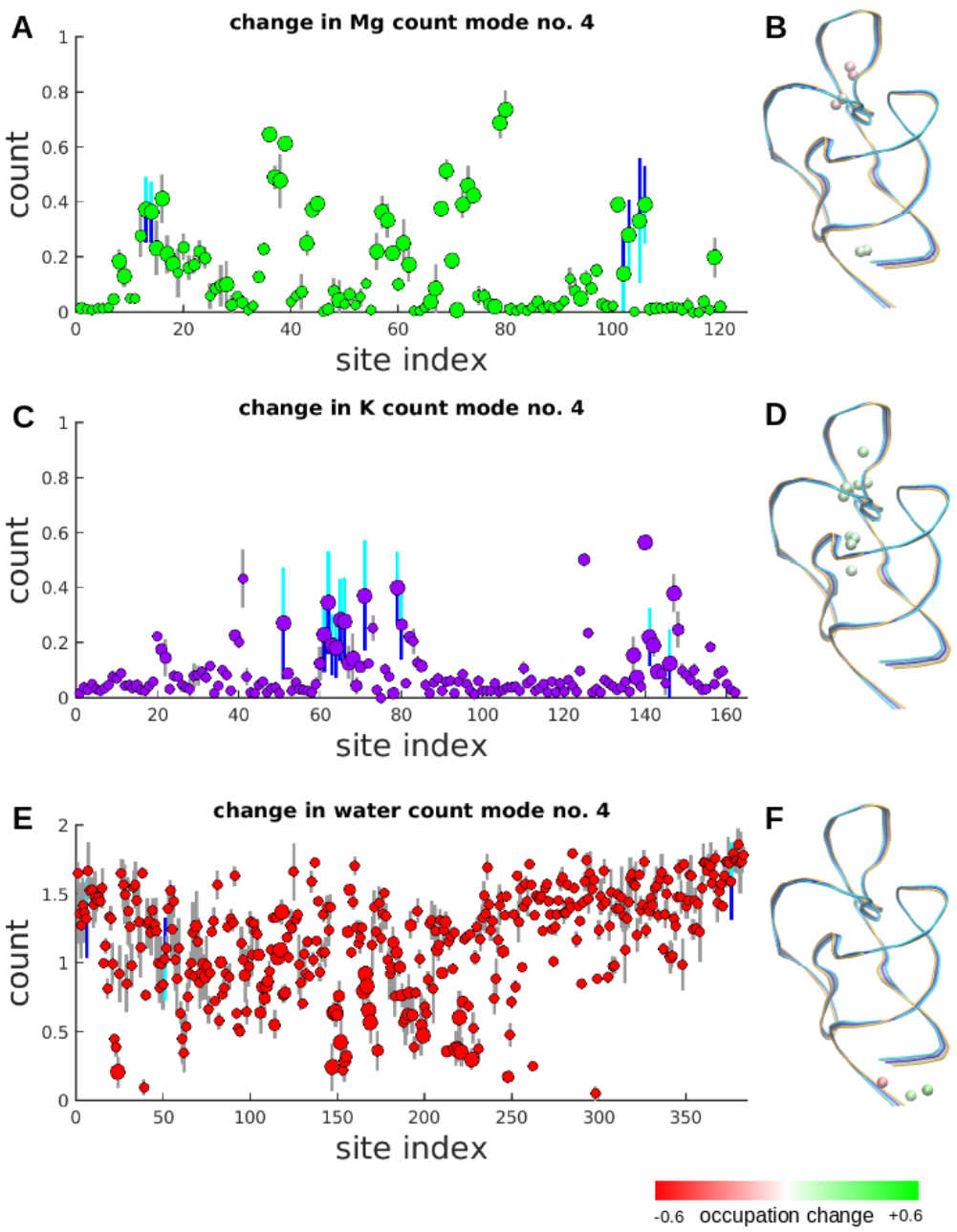


**Figure A.20:** Changes in site occupation counts due to joint PC 4. (same as Fig. A.14)

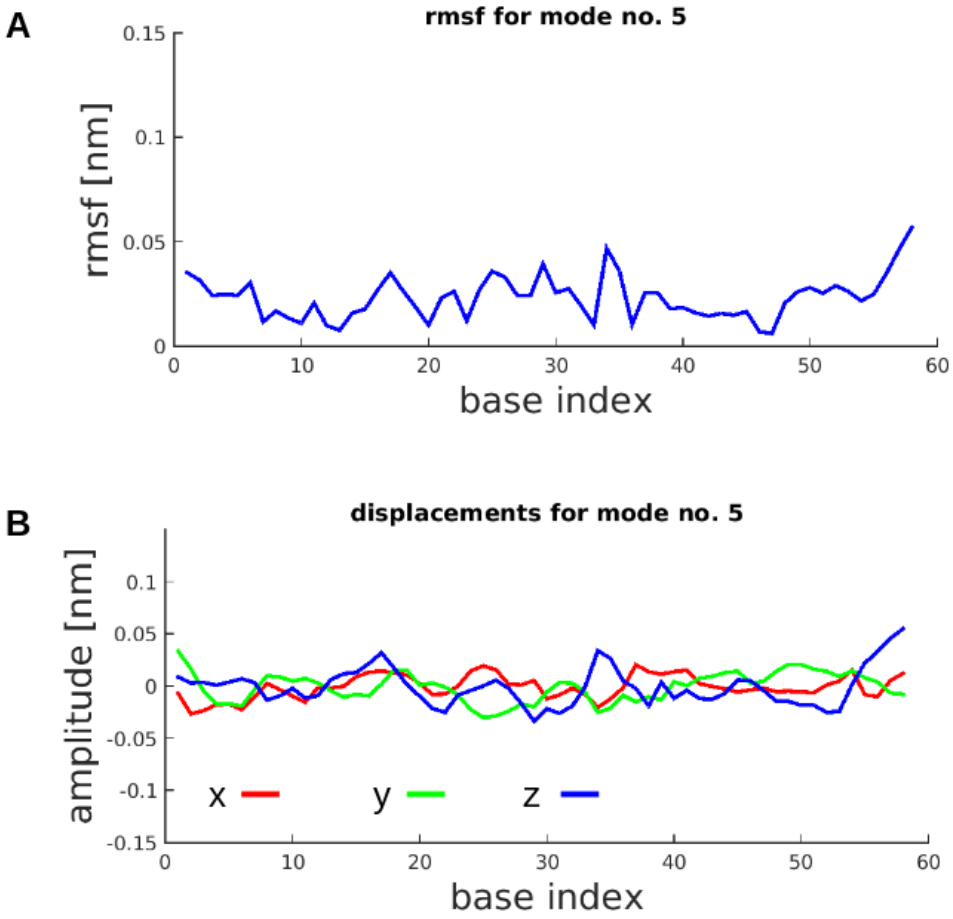


**Figure A.21:** Motions of RNA P due to joint mode 5. (same as Fig. A.13)

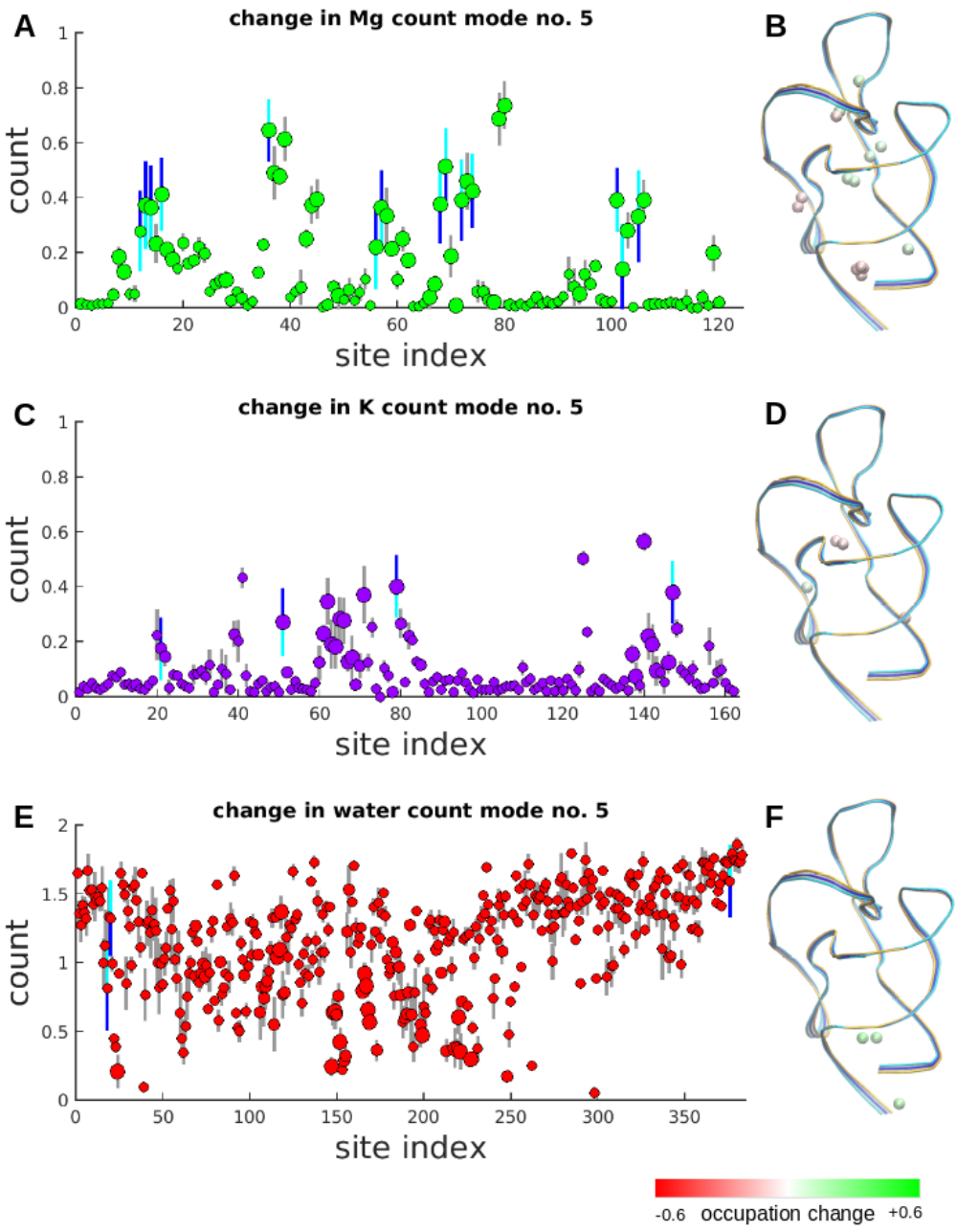


**Figure A.22:** Changes in site occupation counts due to joint PC 5. (same as Fig. A.14)

# Appendix B.

# Appendix for Chapter 2: Diffuse ions can coordinate dynamics in a ribonucleoprotein assembly

## B.1. Supporting Methods

### B.1.1. Parameter refinement of effective potentials for diffuse ion interactions

The parameters in the effective potential $V_{\mathrm{E}}$ (Eq.2.3 in main text) were initially set using explicit-solvent simulations as a reference. From explicit-solvent simulations, we calculated the radial distribution functions (RDF, Fig. B.4) for different types of ion-RNA, ion-ion and ion-protein interactions and then fit our potential to the potential of mean force (PMF) in order to provide initial estimates. Then, an iterative protocol was applied to refine the values of parameters in the effective potential. The details are described below.

### Initial parameters

Initial values for the parameters in the effective potential $V_{\mathrm{E}}$ $(A, B^{(k)}, C^{(k)}, R^{(k)})$ defined in Eq. 2.3 were determined by fitting the functional form of the effective potential to the potentials of mean force (PMF) (Fig. B.4b) for pairwise interactions (e.g. K-K, K-Cl, K-Mg, Mg-Cl, etc.), calculated from all-atom explicit-solvent MD simulations (Fig. B.4a). Each interaction was decomposed into a coulomb term , a effective excluded volume repulsive term and Gaussian terms that account for ionic solvation shells. First, we subtracted out the electrostatic potential (Debye-Hückel) from the PMF and then fit the remaining potential energy terms ($V_{\mathrm{ion-excl}}$ and $V_{\mathrm{sol}}$) to the residuals. The coefficient A in $V_{\mathrm{sol}}$ was initially set by fitting the function $\frac{A}{r^{12}}$ to the short distance region of the PMF associated with the corresponding interaction (Fig. B.4c, dashed blue line). In the functional form of $V_{\mathrm{sol}}$, the width ($C^{(k)}$), location ($R^{(k)}$) and amplitude ($B^{(k)}$) of each Gaussian were obtained by fitting to a corresponding wells and peaks of the PMF. The widths and positions of the Gaussian functions indicate the ranges and locations of solvation shells around each type of ion, and these values remain fixed during all additional parameter refinement steps. With $R^{(k)}$ and $C^{(k)}$ determined for the Gaussian-based interactions, $V_{\mathrm{E}}$ can be expressed as a linear combination of $\sum \frac{1}{r_{ij}^{12}}$ and $\sum e^{-C^{(k)}[r_{ij}-R^{(k)}]^2}$ with coefficients $A$ and $B^{(k)}$. These linear coefficients are then refined using the protocol described in the next section.

Through this procedure, we obtained the initial guess of parameters for ion-ion, ion-RNA and ion-protein interactions from explicit solvent simulations of 100 mM KCl and 10 mM $MgCl_2$ in water, helix 44 (h44, Fig. 2.1) from rRNA in 100 mM KCl and 10 mM $MgCl_2$ solution and ribosomal protein S6 (Fig. B.3) in 100 mM KCl and 10 mM $MgCl_2$ solution respectively.

### Iterative parameter refinement protocol

Starting from the initial guess of parameters $A$ and $B^{(k)}$ in Eq. 2.3, we performed an iterative parameter refinement strategy, where explicit-solvent simulations are used as a benchmark [141, 142]. While the complete protocol is described elsewhere [141], we provide an overview here, for completeness. This protocol allows for refinement of all parameters that are linear in the Hamiltonian. In this protocol, one considers a Hamiltonian of the form

$$H_{\mathrm{R}} = \sum_{\alpha=1}^{N} K_\alpha S_\alpha, \tag{B.1}$$

where a set of $N$ physical observables are denoted by $S_\alpha$ with associated weights $K_\alpha$. In our model, $K_\alpha$ corresponds to the parameters $A$ (in $V_{\mathrm{ion-excl}}$) and $B^{(k)}$ (in $V_{\mathrm{sol}}$), while $S_\alpha$ corresponds to $\sum_{i<j} \frac{1}{r_{ij}^{12}}$ and $\sum_{i<j} e^{-C^{(k)}[r_{ij}-R^{(k)}]^2}$. As described in the previous section, $C^{(k)}$ and $R^{(k)}$ remain fixed during refinement.

The form of the Hamiltonian is exploited to construct an iterative scheme to determine the coefficient $K_\alpha$ associated with each observable $S_\alpha$[141, 200, 201]. Let the expected value of the observable $S_\alpha$ be denoted by $\langle S_\alpha \rangle$. Thus, $\langle S_\alpha \rangle_{\mathrm{SBM}}$ and $\langle S_\alpha \rangle_{\mathrm{ex-sol}}$ denote the average values of the observables, as calculated from the SBM simulations and all-atom explicit-solvent simulation, respectively. To determine values of $S_\alpha$ that accurately describe observables in a reference system, we expand the expected value of each observable $\langle S_\alpha \rangle_{\mathrm{SBM}}$ in a Taylor series in $K_\alpha$ around some trial weights $\{K_\alpha^{(0)}\}$. The difference between the expected values of the observables, $\Delta\langle S_\alpha \rangle = \langle S_\alpha \rangle_{\mathrm{SBM}} - \langle S_\alpha \rangle_{\mathrm{ex-sol}}$, to a first-order perturbation, is[141, 200, 201]

$$\Delta\langle S_\alpha \rangle = \sum_\gamma \frac{\partial \langle S_\alpha \rangle}{\partial K_\gamma} \Delta K_\gamma \tag{B.2a}$$

$$= \frac{1}{k_B T} \sum_\gamma [\langle S_\alpha S_\gamma \rangle - \langle S_\alpha \rangle \langle S_\gamma \rangle] \Delta K_\gamma, \tag{B.2b}$$

where $\Delta K_\gamma^{(n)} = K_\gamma^{(n+1)} - K_\gamma^{(n)}$. Eq. (B.2b) follows from Eq. (B.2a) due to the

linearity of the reference Hamiltonian.

To obtain the corrections to the weights, $\Delta K_\alpha^{(n)}$, we iteratively performed SMOG-ion simulations and get the update of the weights in each iteration by solving the set of linear equations, i.e., Eq. (B.2b). The $n$th iteration, for example, defines the parameter set $K_\alpha^{(n+1)} = K_\alpha^{(n)} + \Delta K_\alpha^{(n)}$. The iterative simulation and updates of the weights are continued until $K_\alpha$ converges for $\alpha = 1, ..., N$. As $K_\alpha$ converges, the observable $S_\alpha$ obtained from the refined model converges to that from the reference system and the RDF from SMOG-ion model simulation converges to that from the explicit solvent simulation (Fig. B.12a). The values of coefficients $A$ and $B^{(k)}$ are obtained from the converged values of $K_\alpha$.

As employed in previous studies conducted by Savelyev and Papoian [141], the quality of the parameters can be measured by calculating the difference in free-energy ($|\delta F|$) between the explicit-solvent and structure-based model, defined as:

$$\delta F = \sum_\alpha |K_\alpha \Delta S_\alpha| \tag{B.3}$$

In the iterative simulations, $\delta F$ decreased significantly in the first few iterations and then levels off near zero (Fig. B.12). In the parameter refinement process, we stopped updating parameters when $\delta F$ reached a minimal value.

### B.1.2. Defining non-contact interactions in the SMOG-ion model

Any atom pairs that are not in contact in the predefined structure are defined as "non-contact interactions". In the SMOG-ion model, the non-contact interaction of a pair of atoms (i, j) in RNA or protein that are separated by $r_{ij}$ have a potential energy of the form $\frac{C_{18}}{r_{ij}^{18}} - \frac{C_{12}}{r_{ij}^{12}}$. This pair-wise potential accounts for the soft repulsive interactions between atoms due to the excluded volume. In order to mimic the

excluded volume description provided by all-atom explicit-solvent models, the coefficients $C_{18}$ and $C_{12}$ were calculated by fitting the 12-18 potential to the Lennard-Jones potential ($V_{\mathrm{Amber-LJ}}$) in the Amber99sb-ildn forcefield [144], which is given by $4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^{6}]$, where $\sigma$ and $\epsilon$ are provided in the force field parameter set. Two reference points, $(V_{\mathrm{Amber-LJ},1}, r_{ij,1})$ and $(V_{\mathrm{Amber-LJ},2}, r_{ij,2})$, are obtained for each pair of atoms from the Amber99sb-ildn forcefield[144], where $V_{\mathrm{Amber-LJ},1} = 1/2k_BT$ and $V_{\mathrm{Amber-LJ},2} = k_BT$. The value of coefficients $C_{18}$ and $C_{12}$ are calculated by fitting the functional form of the non-contact interaction to the reference points. An example of the curve fitting is shown in Figure B.1.

### B.1.3. Simulation details

#### *Explicit-solvent simulations*

Explicit-solvent MD simulations were used as a benchmark for initial parameterization of the SMOG-ion model. The explicit-solvent simulations were conducted using Gromacs v5.1.4 with the Amber99sb-ildn force field[144]. The simulated systems were solvated with SPC/E water molecules [112]. Modified $Mg^{2+}$ parameters described by Åqvist[36] were used in the simulation. The parameters for monovalent ions ($K^+$ and $Cl^-$) reported by Joung and Cheatham[37] were included. Periodic boundary conditions were used in the simulations. Particle-mesh Ewald (PME) was used for the evaluation of long-range electrostatics[114] with an Ewald radius of 10 Å. The Van der Waals cutoff was taken to be 10 Å. Neighbor lists were updated every 10 time steps. Equations of motion were integrated using a leap-frog integrator with a 2-fs time step.

In order to parameterize the ion-ion, ion-RNA and ion-protein interactions in the SMOG-ion model, the following explicit-solvent simulations were performed as benchmarks:

1. *Ions in aqueous solution.* 60 $K^+$, 6 $Mg^{2+}$, 72 $Cl^-$ ions and 32635 water molecules were added to a 10 nm cubic box to create a target concentration of 10 mM for [$MgCl_2$] and 100 mM for [KCl]. The water molecules and ions were energetically minimized with steepest descent followed by conjugate gradient methods. The system was initially equilibrated at 300 K and ambient pressure of 1 bar using the Berendsen thermostat and barostat[202] for 1 ns. After the equilibration step, 1$\mu$s of production runs was carried out with Nose-Hoover thermostat[203, 204] and Parrinello-Rahman barostat[205, 206]. This simulation provides benchmark reference for ion-ion interactions.

2. *Helix 44 in ionic solution.* A RNA fragment, helix 44 (h44), from 16S ribosomal RNA (PDB: 4V6F) was placed in the center of a $10\times12.2\times12.2$ $nm^3$ rectangular box with the principal axis of H44 aligned with the x-dimension of the box. 123 $K^+$, 14 $Mg^{2+}$, 108 $Cl^-$ and 47917 water molecules were added to the box to neutralize the negative charges from H44 and yield a bulk concentrations of approximately 10 mM for [$MgCl_2$] and 100 mM for [KCl]. To eliminate the influence from any conformational changes of h44 on ion-RNA interactions, every atom on H44 was position restrained to its initial coordinate with force constant 1000 kJ/mol/$nm^2$. The entire system was energetically minimized with steepest descent and then conjugate gradient method. Then we performed the equilibration of the system at 300 K and 1 bar with Berendsen thermostat and barostat[202], which was followed by 1 $\mu$s production run with Nose-Hoover thermostat[203, 204] and Parrinello-Rahman barostat[205, 206]. The first 20 ns trajectory was used for equilibrating ions with the h44 and was excluded from RDF analysis. This simulation provides the benchmark reference for ion-RNA interactions.

3. *Protein S6 in ionic solution.* Protein S6 was extracted from a structure of a

bacterial ribosome (PDB: 4V6F) and was placed in the center of a rectangular box of size $10\times16\times16\text{nm}^3$ with its first principal axis aligned to the x-dimension of the box. 15 $Mg^{2+}$, 154 $K^+$ and 185 $Cl^-$ ions were added to the box together with 88364 water molecules to create a bulk concentration of approximately 10 mM [$MgCl_2$] and 100 mM of [KCl]. Atoms in the protein molecule were position restrained to their initial coordinates with force constant 1000 kJ/mol/$\text{nm}^2$ during the simulations. The energy of the system was minimized with steepest descent and conjugate gradient methods. The entire system was equilibrated at 300 K and 1 bar with Berendsen thermostat and barostats[202], which was followed by the 1 $\mu$s production run with Nose-Hoover thermostat[203, 204] and Parrinello-Rahman barostat[205, 206]. The first 20 ns trajectory was excluded from the RDF analysis. This simulation provides the benchmark reference for ion-protein interactions.

***SMOG-ion model simulations***

MD simulations with the SMOG-ion model were performed for different purposes: (1) refining the parameters A and B in the effective potential $V_{\text{E}}$ (Eq. 2.3) in comparison with explicit-solvent simulations, (2) investigating the preferential interaction coefficients ($\Gamma_{2+}$) of two RNA molecules with refined parameters to compare with experimental measurements, (3) studying the conformational dynamics of the ribosome structure under ionic effects. Simulation details are as followed.

For the purpose of parameter refinement, SMOG-ion simulations w ere performed with exactly the same number of ions in the same size of boxes as in the corresponding benchmark explicit-solvent simulations. The simulations were performed at a temperature of 60 Gromacs units (0.5 reduced units) which corresponds to 300 K in explicit-solvent simulations. In each iteration of parameter refinement, 10 replicas of 10 million timesteps were performed with the SMOG-ion model, and trajectories

from all replicas were used to determine the value of each observable.

After the parameter refinement process, SMOG-ion simulations were applied to the small 58-mer rRNA (PDB: 1HC8) and the adenine riboswitch (PDB: 1Y26) to investigate the ion preferential interaction coefficient ($\Gamma_{2+}$) for each system. In both cases, the small RNA molecule was placed in a cubic box of length 70 nm. Ions were placed randomly but not close to each other or the RNA molecule initially. For the 58mer rRNA, the number of $Mg^{2+}$ ions was varied from 45 to 216 to achieve the target bulk concentration of $Mg^{2+}$ ions (0.20-1.00 mM). 30983 $K^+$ ions were placed in the box to set the concentration of $K^+$ to 150mM. For the adenine riboswitch, 33 to 207 $Mg^{2+}$ ions were used with 10328 $K^+$ ions to create a $Mg^{2+}$ concentrations that range of 0.16 to 1.00 mM and a $K^+$ concentration of 50 mM. In both cases, $Cl^-$ ions were added to neutralize the charges in the system. The chelated ions were harmonically bound to the neighboring oxygen atoms in the chelation pocket. At each concentration, the system was equilibrated for $10^8$ timesteps with 0.002 reduced unit for each timestep. Then the equilibrium run was extended for another $2 \times 10^7$ and we took a snapshot of the atom coordinates after each $10^6$ timesteps. These 20 snapshots were used to initialize 20 replicas with random initial velocities. Each replica was then simulated for an additional $6 \times 10^7$ time steps, from which the first $10^7$ time steps were discarded during analysis.

In the simulation of the whole ribosome structure (PDB: 6QNR), the ribosome was placed in the center of a cubic box with length 70 nm. To avoid unphysical expansion of the ribosome molecule, all contact distances in the ribosome were reduced by a factor of 0.96. In the simulations with ions, 3500 $Mg^{2+}$ ions, 20000 $K^+$ ions and 24226 $Cl^-$ ions were added the box to neutralize the charges from the ribosome structure and target the bulk concentrations of 10 mM for [$MgCl_2$] and 100 mM for [KCl]. In the simulation of the ribosome structure without the presence of E site tRNA, 24301 $Cl^-$ ions were added with 3500 $Mg^{2+}$ and 20000 $K^+$ in the box. In

both simulations, we ran $1.25 \times 10^7$ timesteps of equilibration in which the ribosome was position restrained at its initial coordinates, which allowed the diffuse ions to equilibrate around the ribosome. Then a production run of $10^8$ timesteps without position restraints was performed. In simulations of the ribosome without explicit diffuse ions, the $10^8$-timestep production runs were performed after the equilibration run of $10^7$ timesteps.

## B.2. Supporting Results

### B.2.1. Refined effective potential parameters

As described in the Method section, the SMOG-ion model parameters were initially refined in comparison with explicit-solvent simulations. Then a few parameters of ion-RNA interactions were further adjusted by comparing with the experimental results and using the preferential interaction coefficient ($\Gamma_{2+}$) as a metric. The finalized parameter set s3 is shown in the tables below.

**Table B.1:** Refined parameters for ion-ion interactions are obtained from the simulation with 10mM $MgCl_2$ and 100mM KCl. $\epsilon = 2k_BT$ in the units of parameter A and B.

| Interaction | A [$\epsilon \cdot \mathrm{nm}^{12}$] | B [$\epsilon$] | | | | |
|---|---|---|---|---|---|---|
| | A | $B_1$ | $B_2$ | $B_3$ | $B_4$ | $B_5$ |
| $K^+$-$K^+$ | $2.510 \times 10^{-6}$ | -0.7371 | 0.1887 | -0.0029 | 0.111 | 0.0632 |
| $K^+$-$Cl^-$ | $4.484 \times 10^{-7}$ | -0.4908 | 0.7674 | -0.2481 | 0.0578 | -0.1116 |
| $K^+$-$Mg^{2+}$ | $1.639 \times 10^{-4}$ | -0.032 | 0.2484 | 0.0448 | 0.1482 | 0.1106 |
| $Cl^-$-$Cl^-$ | $4.689 \times 10^{-5}$ | -0.3862 | 0.348 | -0.0205 | 0.1262 | 0.0509 |
| $Mg^{2+}$-$Cl^-$ | $1.213 \times 10^{-5}$ | -0.4291 | 0.4266 | -0.2472 | -0.0106 | -0.1742 |
| $Mg^{2+}$-$Mg^{2+}$ | $9.215 \times 10^{-3}$ | - | - | - | - | - |

| Interaction | C [$\mathrm{nm}^{-2}$] | | | | | R [nm] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $R_1$ | $R_2$ | $R_3$ | $R_4$ | $R_5$ |
| $K^+$-$K^+$ | 284.3 | 1039 | 1458.5 | 427.8 | 235.5 | 0.424 | 0.5662 | 0.6673 | 0.7712 | 0.8986 |
| $K^+$-$Cl^-$ | 1125 | 304.3 | 571.8 | 1309.9 | 392.1 | 0.3247 | 0.3924 | 0.5386 | 0.6358 | 0.7441 |
| $K^+$-$Mg^{2+}$ | 1769.9 | 204.7 | 335.6 | 216.2 | 1815.1 | 0.5933 | 0.6973 | 0.8095 | 0.9056 | 1.0144 |
| $Cl^-$-$Cl^-$ | 770.8 | 299.9 | 1844 | 514.2 | 650.7 | 0.5245 | 0.638 | 0.768 | 0.8631 | 0.984 |
| $Mg^{2+}$-$Cl^-$ | 628.3 | 798.3 | 359.2 | 2041.6 | 404.1 | 0.4715 | 0.5465 | 0.6775 | 0.7775 | 0.8794 |
| $Mg^{2+}$-$Mg^{2+}$ | - | - | - | - | - | - | - | - | - | - |

**Table B.3:** Refined parameter for ion-RNA interaction. Atoms from RNA are grouped by their element and the value of their partial charges. Only the interactions between metal ions ($Mg^{2+}$ and $K^{+}$) with the negatively charged RNA atoms are parameterized. The negatively charged RNA atoms are grouped by their elements and the value of their partial charges. For atoms whose partial charges are more negative than -0.49e, the element symbols are superscripted with "<-0.5", otherwise, the element symbol will be superscripted with ">-0.5". In the table, "0" means corresponding parameters are defined while set to zero in the model, while "-" means corresponding parameters are not defined in the SMOG-ion model.

| Interaction | A [$\epsilon \cdot nm^{12}$] | B [$\epsilon$] | | | | |
|---|---|---|---|---|---|---|
| | A | $B_1$ | $B_2$ | $B_3$ | $B_4$ | $B_5$ |
| $K^{+}$-$O_{<-0.5}$ | $2.923 \times 10^{-8}$ | 0 | 0.2962 | 0.0339 | -0.0217 | -0.0973 |
| $K^{+}$-$N_{<-0.5}$ | $3.703 \times 10^{-8}$ | 0 | 0.085 | -0.2878 | - | - |
| $K^{+}$-$O_{>-0.5}$ | $3.688 \times 10^{-8}$ | 0 | 0.9727 | - | - | - |
| $K^{+}$-$C_{>-0.5}$ | $1.208 \times 10^{-6}$ | - | - | - | - | - |
| $K^{+}$-$N_{>-0.5}$ | $2.239 \times 10^{-7}$ | - | - | - | - | - |
| $Mg^{2+}$-$O_{<-0.5}$ | $2.522 \times 10^{-6}$ | -0.89 | 0.0157 | -0.1161 | 0.0093 | -0.0599 |
| $Mg^{2+}$-$N_{<-0.5}$ | $2.250 \times 10^{-6}$ | -1.02 | - | - | - | - |
| $Mg^{2+}$-$O_{>-0.5}$ | $1.465 \times 10^{-5}$ | - | - | - | - | - |
| $Mg^{2+}$-$C_{>-0.5}$ | $1.600 \times 10^{-5}$ | - | - | - | - | - |
| $Mg^{2+}$-$N_{>-0.5}$ | $1.162 \times 10^{-5}$ | - | - | - | - | - |

| Interaction | C [$nm^{-2}$] | | | | | R [nm] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $R_1$ | $R_2$ | $R_3$ | $R_4$ | $R_5$ |
| $K^{+}$-$O_{<-0.5}$ | 1095 | 501 | 438 | 1833 | 123 | 0.278 | 0.348 | 0.511 | 0.569 | 0.698 |
| $K^{+}$-$N_{<-0.5}$ | 1067 | 325 | 180 | - | - | 0.286 | 0.37 | 0.617 | - | - |
| $K^{+}$-$O_{>-0.5}$ | 954 | 255 | - | - | - | 0.285 | 0.38 | - | - | - |
| $K^{+}$-$C_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $K^{+}$-$N_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$O_{<-0.5}$ | 651 | 392 | 331 | 305 | 92.7 | 0.417 | 0.478 | 0.61 | 0.725 | 1.01 |
| $Mg^{2+}$-$N_{<-0.5}$ | 646 | - | - | - | - | 0.441 | - | - | - | - |
| $Mg^{2+}$-$O_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$C_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$N_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |

**Table B.5:** Refined parameter for ion-protein interaction. Atoms in the protein are also grouped by element and their assigned partial charges. Repulsive interactions between ion and protein atoms are parameterized with both excluded volume and Gaussian terms. For the attractive interactions, only excluded volume coefficients were refined. If an atom has a more negative charge than -0.49e, a subscription "<-0.5" will be added to the element symbol. If an atom has a charge between -0.49e and 0, a subscription ">-0.5" will be added to the element symbol. If an atom has positive partial charge in between - and 0.49e, its element symbol has a subscription "<0.5". If an atom has partial charge greater than 0.49e, its element symbol has a subscription ">0.5". There are a few special cases in this table: "$C_{<0}$" consists of both "$C_{<-0.5}$" and "$C_{>-0.5}$" atoms; "$N_{<0}$" consists of both "$N_{<-0.5}$" and "$N_{>-0.5}$" atoms.

| Interaction | A $[\epsilon \cdot \mathrm{nm}^{12}]$ | B $[\epsilon]$ | | | | |
|---|---|---|---|---|---|---|
| | A | $B_1$ | $B_2$ | $B_3$ | $B_4$ | $B_5$ |
| $K^+$-$O_{<-0.5}$ | $2.898 \times 10^{-8}$ | -1.1804 | 0.2875 | -0.0972 | - | - |
| $K^+$-$O_{>-0.5}$ | $7.092 \times 10^{-8}$ | -1.2356 | 0.371 | -0.0959 | - | - |
| $K^+$-$N_{<0}$ | $1.515 \times 10^{-6}$ | - | - | - | - | - |
| $K^+$-$N_{>0.5}$ | $3.340 \times 10^{-6}$ | - | - | - | - | - |
| $K^+$-$N_{<0.5}$ | $3.756 \times 10^{-7}$ | - | - | - | - | - |
| $K^+$-$C_{<0}$ | $1.635 \times 10^{-6}$ | - | - | - | - | - |
| $K^+$-$S_{>-0.5}$ | $4.297 \times 10^{-6}$ | - | - | - | - | - |
| $Mg^{2+}$-$O_{<-0.5}$ | $7.160 \times 10^{-7}$ | -1.3849 | 0.1479 | -0.0034 | - | - |
| $Mg^{2+}$-$O_{>-0.5}$ | $2.345 \times 10^{-6}$ | - | - | - | - | - |
| $Mg^{2+}$-$N_{<0}$ | $8.740 \times 10^{-5}$ | - | - | - | - | - |
| $Mg^{2+}$-$N_{>0.5}$ | $7.638 \times 10^{-4}$ | - | - | - | - | - |
| $Mg^{2+}$-$N_{<0.5}$ | $1.764 \times 10^{-5}$ | - | - | - | - | - |
| $Mg^{2+}$-$C_{<0}$ | $1.852 \times 10^{-5}$ | - | - | - | - | - |
| $Mg^{2+}$-$S_{>-0.5}$ | $9.425 \times 10^{-5}$ | - | - | - | - | - |
| $Cl^-$-$N_{>0.5}$ | $5.504 \times 10^{-7}$ | -0.944 | 0.1997 | -0.3535 | -0.0562 | -0.1387 |
| $Cl^-$-$N_{<0.5}$ | $6.004 \times 10^{-7}$ | -0.9644 | 0.0828 | -0.1301 | - | - |
| $Cl^-$-$C_{>0.5}$ | $4.707 \times 10^{-6}$ | - | - | - | - | - |
| $Cl^-$-$C_{<0.5}$ | $2.323 \times 10^{-6}$ | - | - | - | - | - |
| $Cl^-$-$C_{<0}$ | $2.420 \times 10^{-6}$ | - | - | - | - | - |
| $Cl^-$-$O_{<-0.5}$ | $8.554 \times 10^{-6}$ | - | - | - | - | - |
| $Cl^-$-$S_{>-0.5}$ | $1.193 \times 10^{-5}$ | - | - | - | - | - |

**Table B.6:** Refined parameter for ion-protein interaction (continued).

| Interaction | C [$nm^{-2}$] | | | | | R [nm] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $R_1$ | $R_2$ | $R_3$ | $R_4$ | $R_5$ |
| $K^+$-$O_{<-0.5}$ | 1127 | 376 | 481 | - | - | 0.27 | 0.351 | 0.498 | - | - |
| $K^+$-$O_{>-0.5}$ | 535 | 186 | 444 | - | - | 0.272 | 0.381 | 0.542 | - | - |
| $K^+$-$N_{<0}$ | - | - | - | - | - | - | - | - | - | - |
| $K^+$-$N_{>0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $K^+$-$N_{<0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $K^+$-$C_{<0}$ | - | - | - | - | - | - | - | - | - | - |
| $K^+$-$S_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$O_{>-0.5}$ | 524 | 1158 | 418 | - | - | 0.406 | 0.477 | 0.656 | - | - |
| $Mg^{2+}$-$O_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$N_{<0}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$N_{>0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$N_{<0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$C_{<0}$ | - | - | - | - | - | - | - | - | - | - |
| $Mg^{2+}$-$S_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Cl^-$-$N_{>0.5}$ | 1298 | 253 | 284 | 1719 | 307 | 0.337 | 0.395 | 0.546 | 0.652 | 0.758 |
| $Cl^-$-$N_{<0.5}$ | 1098 | 298 | 1130 | - | - | 0.338 | 0.395 | 0.558 | - | - |
| $Cl^-$-$C_{>0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Cl^-$-$C_{<0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Cl^-$-$C_{<0}$ | - | - | - | - | - | - | - | - | - | - |
| $Cl^-$-$O_{<-0.5}$ | - | - | - | - | - | - | - | - | - | - |
| $Cl^-$-$S_{>-0.5}$ | - | - | - | - | - | - | - | - | - | - |

## B.3. Supporting Figures

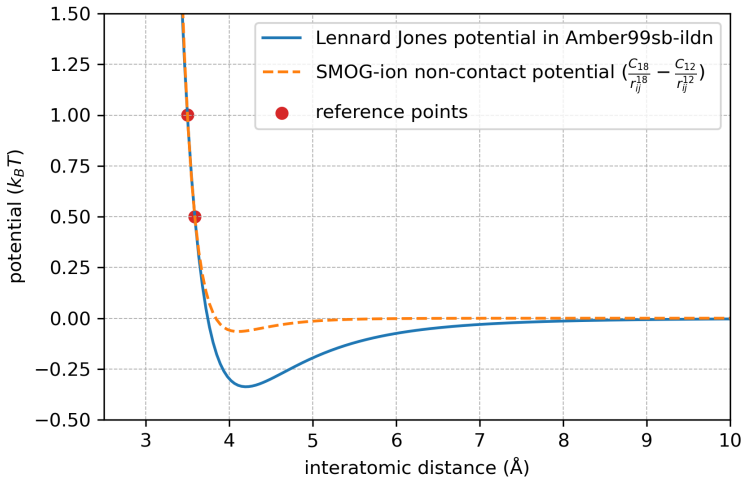


**Figure B.1: Characterize the excluded volume of atoms according to AMBER force field.** An example of determining the values of coefficients $C_{18}$ and $C_{12}$ for non-contact interactions (orange dashed line) in SMOG-ion model. The Lennard-Jones potential (blue solid curve) defined in Amber99sb-ildn forcefield[144] provides the reference points (red dots).

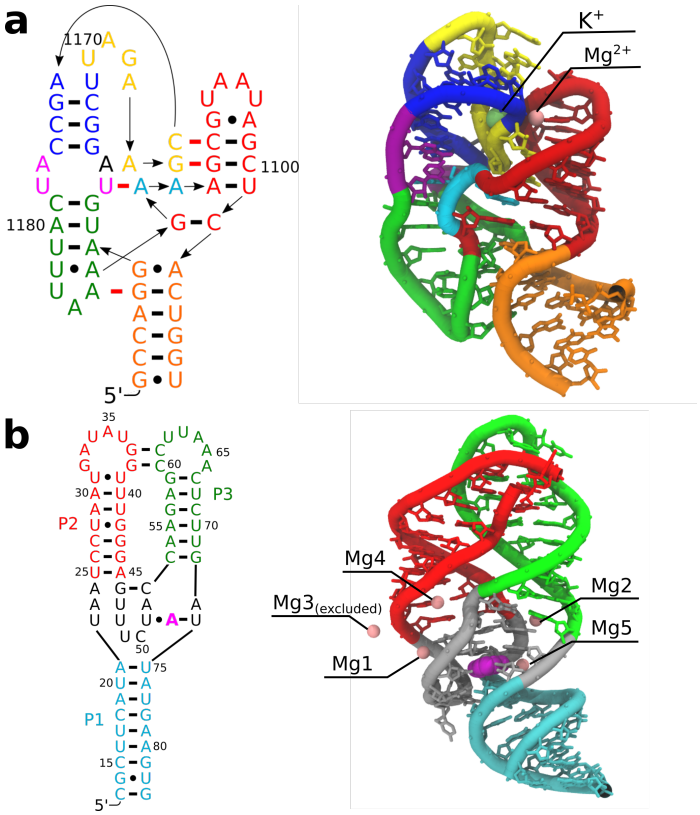


**Figure B.2: Structures of 58-mer rRNA fragment and adenine riboswitch.** (a) 58-mer rRNA fragment (PDB:1HC8). Color codes of residues on right panel are consistent with the left panel. In the left panel, base pairs are indicated by black horizontal lines, tertiary base-base hydrogen bonds in the folded RNA are shown in red bars. Arrows indicate the 5' to 3' direction of the backbone. Two chelated ions ($K^+$ and $Mg^{2+}$) are shown in beads. (b) Adenine riboswitch (PDB:1Y26). Color codes of residues in the tertiary structure on right panel are consistent with the left panel. 5 chelated $Mg^{2+}$ ions are shown in pink.

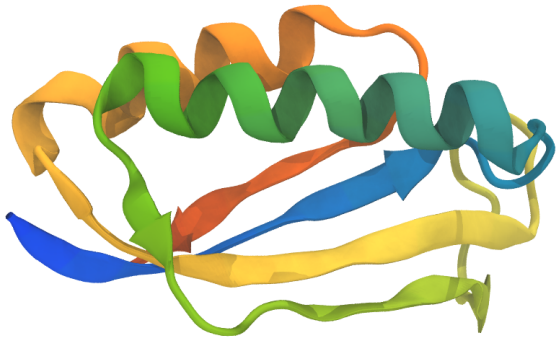

**Figure B.3: Structure of protein S6**. S6 is a small protein with 101 amino acid residues. It consists two $\beta-\alpha-\beta$ motifs with four-stranded anti-parallel $\beta$-sheet on one side and two $\alpha$-helices packed on the other side. Similar folding patterns were observed in other ribosomal proteins.[207]. Since our goal is to investigate the effects of diffuse ions on ribonucleoprotein assemblies, we decided to use this globular ribosomal protein as a reference when parameterizing the interactions between ions and protein atoms.

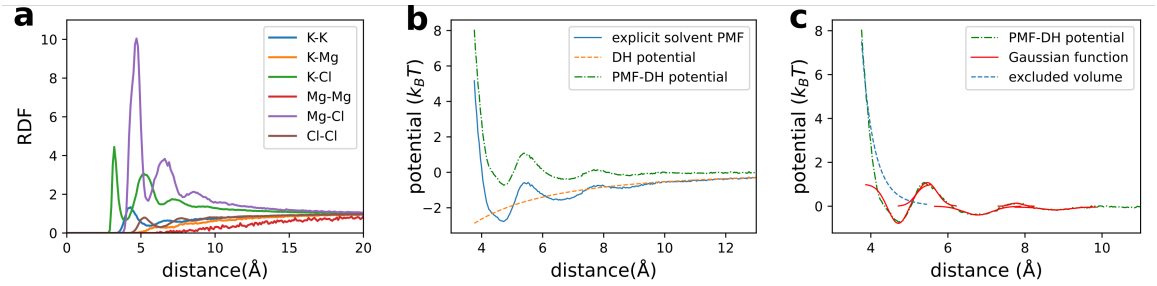


**Figure B.4: Explicit solvent simulations provide reference for ion associated interactions.**(a) The RDFs of ion-ion interactions from explicit solvent simulations of 10 mM $MgCl_2$ and 100 mM KCl in water. (b) The PMF of Mg-Cl interactions (blue), which is converted from the corresponding RDF values. The PMF based on the Debye-Hückel (DH) potential is shown as dashed line (orange). The PMF with DH contribution subtracted is shown as dashed dot line (green). (c) For initial fitting of the parameters in the Hamiltonian (see Eq.2.3) that describe Mg-Cl interactions, the DH contribution was subtracted from the PMF (dashed dot green line). The resulting PMF was approximated by the excluded volume repulsive term (dashed blue line) and the sum of five Gaussian functions (red line).

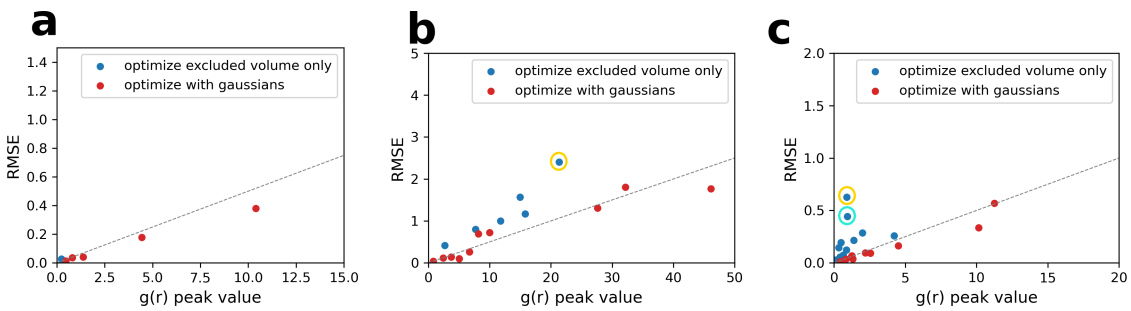


**Figure B.5: Consistency of parameterized interaction in SMOG-ion model with explicit-solvent model.** Root mean squared error (RMSE) of g(r) between SMOG-ion model (s1 parameter set, see main text for definition) simulation and the corresponding explicit solvent simulation versus the peak g(r) values of the latter are plotted for each type of interactions. Dashed line in each panel corresponds to the boundary of 5% relative error. Generally, interactions optimized with Gaussians in their effective potentials (red dots) benefit from the parameterization of the solvation effects, which results in lower relative error comparing to those optimized for excluded volume only (blue dots). a) The RMSE associated with ion-ion interactions from the simulations of ions in aqueous solution. b) The RMSE associated with ion-RNA and ion-ion interactions from the simulations of H44 in ionic solution. The blue dot in the yellow circle, which associates with the interaction between $Mg^{2+}$ ion and weakly negatively charged C atoms ($C_{>-0.5}$) in RNA residues, yields the highest relative error (11.2%) in this set. All other interactions have relative error within 10%. c)The RMSE associated with ion-protein and ion-ion interactions from the simulation of protein S6 in ionic solution. The interaction between $Mg^{2+}$ ion and weakly positively charged N atoms ($N_{<0.5}$, yellow circle) yield the highest relative error, because $N_{<0.5}$ only appears in ARG amino acid residues in protein S6 and has nearly zero (0.0329 $e$) charge, which result super weak interactions with $Mg^{2+}$ ions and obvious fluctuations in g(r). The second highest relative error associates with the interaction between $Mg^{2+}$ ion and the highly positively charged N atom ($N_{>0.5}$), because the latter only appears in N-terminal amino acid residues and LYS residue, which only appears 4 times in protein S6 and results insufficient samplings for precise parameterization.

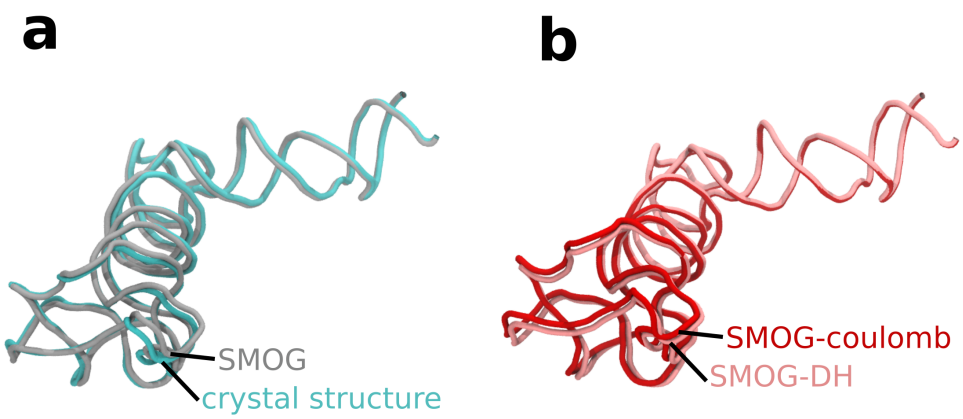


**Figure B.6: Similarity of L1 stalk position in variants of SMOG models.** Average structures of L1 stalk from simulations of ribosome (without E-site tRNA) with multiple variants of all-atom structure model and in the crystral structure of ribosome. a) When the electrostatic-free model (grey) is applied in the simulation, the average position of L1 stalk is consistent with the crystallographic structure (cyan). b) When the electrostatic interactions are included in the model, the average structure of L1 stalk with the Debye-Hückel treatment of implicit ions (pink) only shows sight difference from that without ions(red).

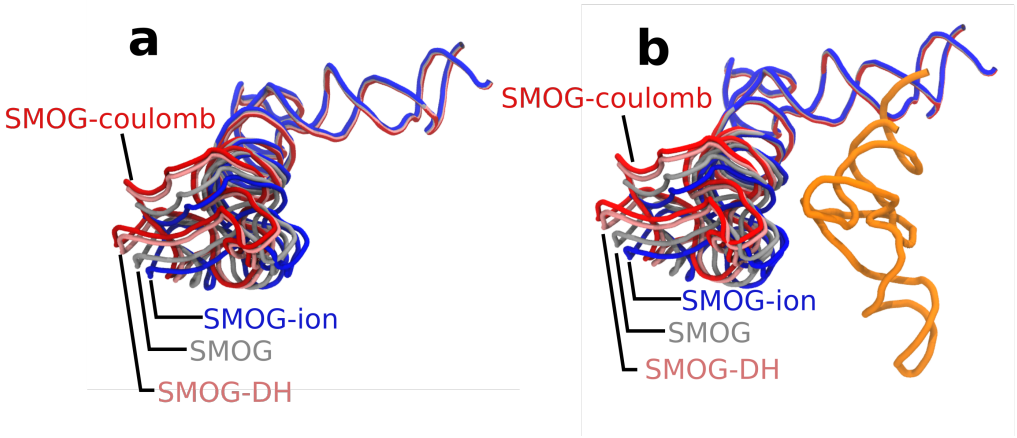


**Figure B.7: Ion-induced L1 stalk displacement is robust to the presence of E-site tRNA.** The electrostatic- and ion-dependent rearrangements of the L1 stalk are robust to the presence of E-site tRNA. The average structures of L1 stalk from simulations without the presence of E-site tRNA are shown in panel a. While panel b shows the average structures of L1 stalk with the presence of E-site tRNA in the simulation. See main text for details about the comparison of L1 stalk structures with variants of all-atom structure based models.

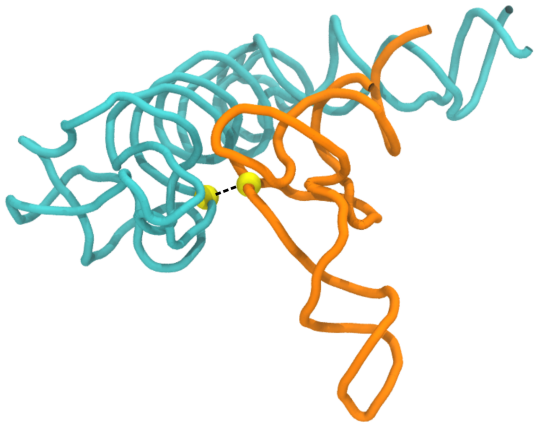

**Figure B.8: Defining $R_{\mathrm{L1-tRNA}}$.** The distance between P atoms (yellow) from the residue A2112 in L1 stalk (cyan) and residue G19 in E-site tRNA (orange) was used as the reaction coordinate to measure the distance between L1 stalk and E-site tRNA. The L1 stalk shown in cyan includes residue G2073 to A2238 in the large subunit of the ribosome structure.

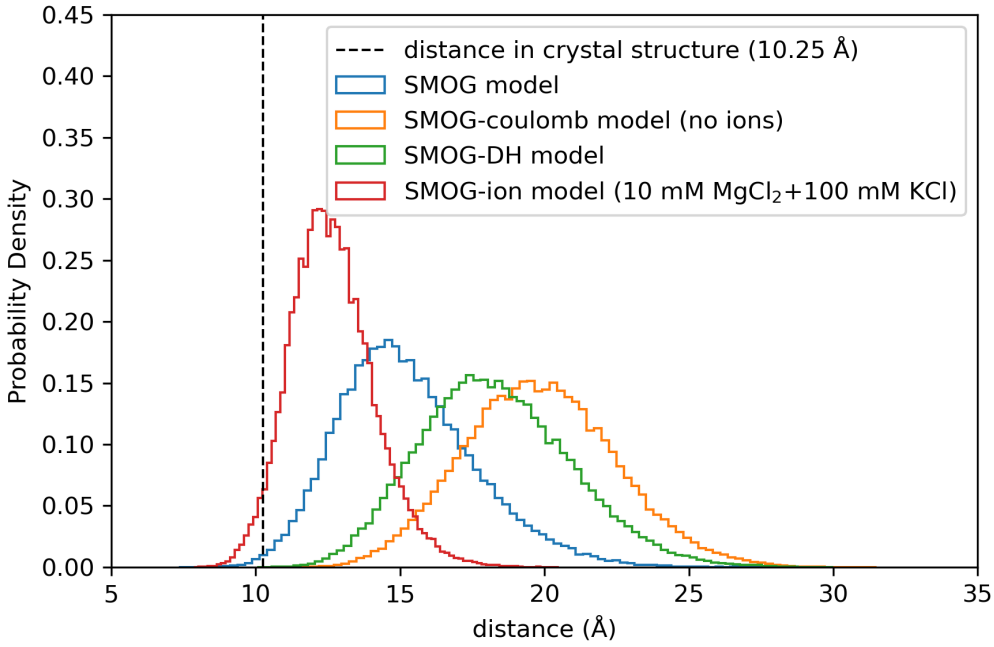


**Figure B.9: Distribution of $R_{\mathrm{L1-tRNA}}$ in simulations with variants of SMOG models.** Distance between L1 stalk and E site tRNA from simulations of different models with ribosome structure including the E-site tRNA. The distance was measured by the coordinates of atom P in residue G2112 in 23S rRNA and atom P in residue G19 in E site tRNA. Without considering the electrostatic effects and ionic effect, SMOG model revealed an average distance of 14.6Å(blue) between L1 stalk and E site tRNA. With the intra-electrostatic taken into considerations, SMOG-coulomb model revealed an average distance of 19.4Å(orange), which was due to the repulsions between L1 stalk rRNA and E site tRNA. With Debye-Hückel electrostatics included in the model, SMOG-DH simulation yielded an average distance of 17.8 Å(green). Under the effect of diffuse ions, SMOG-ion yielded an average distance of 12.3Å between L1 stalk and E site tRNA (red), which reflects the stabilizing effects from diffuse ions on the contacts between L1 stalk and E site tRNA.

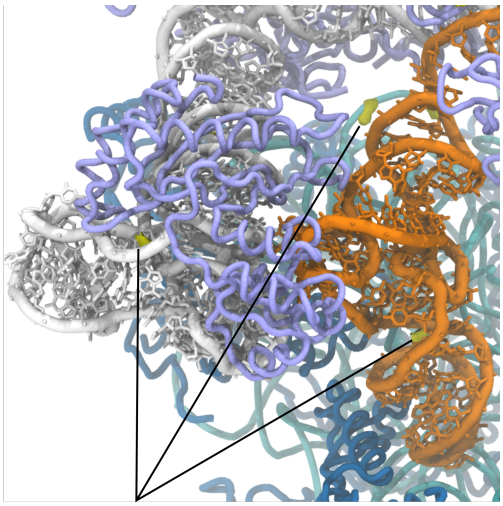


**Figure B.10: The spatial distribution function of $K^+$ ions in the L1 stalk and tRNA.** High density regions (conc. above 3 M) of $K^+$ ions near L1 stalk and tRNA are shown in yellow surface. In comparison with the SDF of $Mg^{2+}$ ions in the same region, the density of $K^+$ ions is significantly attenuated and only show three small regions where the density of $K^+$ ions is above 3 M.

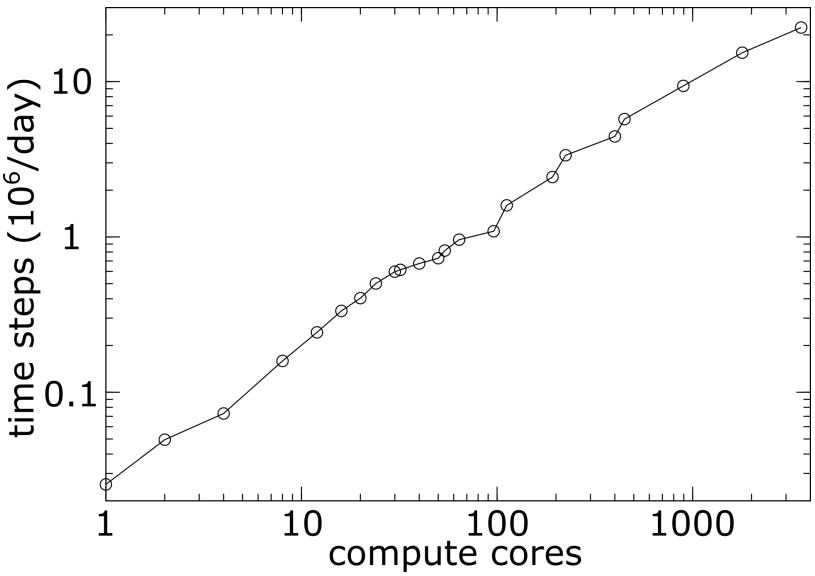


**Figure B.11: Simulation performance as a function of the number of compute cores utilized.** Scaling tests were performed for the ribosome with diffuse ions (198,654 atoms). Performance gains were obtained for up to 3,584 compute cores. Tests used Gromacs 5.1.4, with modified non-bonded kernels to accommodate the SMOG-ion model. Each compute node was equipped with two Intel Xeon Platinum 8276 processors (2.20GHz) with Infiniband interconnect between nodes. Larger core counts were not tested due to the availability of resources.

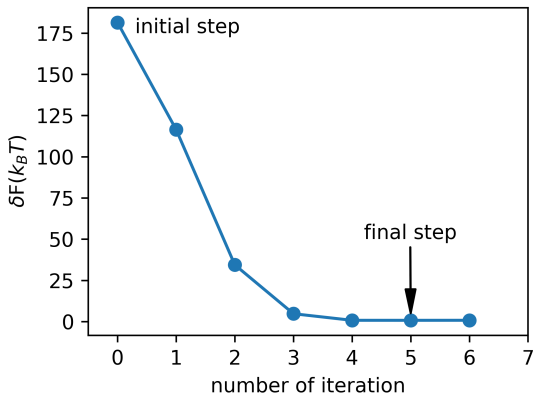


**Figure B.12: Iterative parameter refinement leads to energetic consistency with explicit solvent simulation.**The reduction of the total free energy difference ($\delta F$, Eq. B.3) between explicit solvent model and structure-based model along the parameter refinement of ion-RNA interactions. The parameter refinement is terminated when $\delta F$ stops decreasing and reached a minimal value.

# Appendix C.

# Appendix for Chapter 3: Fluctuation Effects in Adam–Gibbs Model of Cooperative Relaxation.

## C.1. Details of derivation leading to eq. 3.5

The slope of the critical size with respect to temperature at constant pressure will be evaluated by exploiting Tool–Narayanaswamy–Moynihan interpretation of eq. 3.1 [177, 208]. For the equilibrium liquid, the temperature dependence of the relaxation time is given by [177, 208]

$$\left(\frac{\mathrm{d}\ln\tau}{\mathrm{d}T}\right)_{eq} \equiv -\frac{\Delta h^*}{RT^2} \tag{C.1}$$

$$= \frac{\Delta\mu}{k_{\mathrm{B}}T}\left(\frac{\partial z^*}{\partial T}\right)_P - \frac{z^*\Delta\mu}{k_{\mathrm{B}}T^2}. \tag{C.2}$$

Here, $\Delta h^*$ is the activation enthalpy and $R$ is the gas constant. For glass like configuration, the temperature dependence of $\tau$ is [177, 208]

$$\Big(\frac{\mathrm{d}\ln\tau}{\mathrm{d}T}\Big)_{z^*} \equiv -\frac{x\Delta h^*}{RT^2} \tag{C.3}$$

$$= -\frac{z^*\Delta\mu}{k_{\mathrm{B}}T^2}. \tag{C.4}$$

$x$ is the Tool–Narayanaswamy–Moynihan non-linearity parameter [170, 177]. The quantity $x$ is experimentally measurable and denotes how far the system deviates from equilibrium. Making use of Equations C.1 and C.3 we get

$$\Big(\frac{\partial z^*}{\partial T}\Big)_P = -\Big(\frac{k_{\mathrm{B}}T}{\Delta\mu}\Big)\Big(\frac{(1-x)\Delta h^*}{RT^2}\Big). \tag{C.5}$$

On combining eq. C.1, and eq. 3.3 and eq. 3.4 we get eq. 3.5.

## C.2. Fluctuation and Adam–Gibbs Model

In this section we determine the effects of fluctuations on the configurational fraction by assuming that the smallest size $z^*$ of a cooperative region that allows transition is inversely proportional to the configuration entropy. In this case, the difference between the relaxation time $\tau_i$ of a local region and the mean relaxation time of the region is

$$\ln\tau_i - \ln\tau = \frac{\Delta\mu s^* N_{\mathrm{A}}}{k_{\mathrm{B}}T}\Big(\frac{1}{S_{\mathrm{exc,\,i}}f_{\mathrm{i}}} - \frac{1}{S_{\mathrm{exc}}f}\Big). \tag{C.6}$$

$\Delta\mu$ is the activation enthalpy, $S_{\mathrm{c}}$ is the configuration entropy of the supercooled liquid, $s^*$ is the configuration entropy that corresponds to the smallest size, and $N_{\mathrm{A}}$ is the Avogadro number. The terms in the brackets is rewritten as

$$\frac{1}{S_{\mathrm{exc,i}}f_{\mathrm{i}}} \approx \frac{1}{S_{\mathrm{exc}}f}\{1 - \frac{S_{\mathrm{exc,i}}f_{\mathrm{i}} - S_{\mathrm{exc}}f}{S_{\mathrm{exc}}f}\}. \tag{C.7}$$

To proceed further, we assume that the configurational fraction $f_{\mathrm{i}}$ is independent of local region $i$. This assumption can be bypassed. From equations C.6 and refeq:ch3-seq5, we obtain the desired explicit expression for the configurational fraction.

$$f = \Big( \frac{\Delta\mu s^* N_{\mathrm{A}}}{k_{\mathrm{B}} T S_{\mathrm{exc}}^2} \sqrt{\frac{\langle \Delta^2 S_{\mathrm{exc}} \rangle}{\langle \Delta^2 \ln \tau \rangle}} \Big) \tag{C.8}$$

# Bibliography


[1] A. M. Pyle, "Metal ions in the structure and function of RNA", J. Biol. Inorg. Chem. **7**, 679–690 (2002) 10.1007/s00775-002-0387-6.

[2] D. E. Draper, "A guide to ions and RNA structure", RNA **10**, 335–343 (2004).

[3] P. E. Cole, S. K. Yang, and D. M. Crothers, "Conformational Changes of Transfer Ribonucleic Acid. Equilibrium Phase Diagrams", Biochemistry **11**, 4358–4368 (1972) 10.1021/bi00773a024.

[4] C. URBANKE, R. RÖMER, and G. MAASS, "Tertiary Structure of tRNAPhe (Yeast): Kinetics and Electrostatic Repulsion", Eur. J. Biochem. **55**, 439–444 (1975) 10.1111/j.1432-1033.1975.tb02180.x.

[5] a. Stein and D. M. Crothers, "Conformational Changes of Transfer RNA. The Role of Magnesium(II)", Biochemistry **15**, 160–168 (1976).

[6] G. S. Manning, "The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides", Q. Rev. Biophys. **11**, 179–246 (1978) 10.1017/s0033583500002031.

[7] A. Spasic and U. Mohanty, "Counterion condensation in nucleic acid", Adv. Chem. Phys. **139**, 139–176 (2008) 10.1002/9780470259498.ch3.

[8] D. E. Draper, D. Grilley, and A. M. Soto, "Ions and RNA folding", Annu. Rev. Biophys. Biomol. Struct. **34**, 221–243 (2005) 10.1146/annurev.biophys.34.040204.144511.

[9] W. H. Braunlin, "NMR studies of cation-binding environments on nucleic acids", Adv. Biophys. Chem. **5**, 89–140 (1995).

[10] G. L. Conn, A. G. Gittis, E. E. Lattman, V. K. Misra, and D. E. Draper, "A compact RNA tertiary structure contains a buried backbone–$K^+$ complex", J. Mol. Biol. **318**, 963–973 (2002) 10.1016/S0022-2836(02)00147-X.

[11] D. Grilley, V. Misra, G. Caliskan, and D. E. Draper, "Importance of partially unfolded conformations for Mg2+-induced folding of RNA tertiary structure: Structural models and free energies of Mg2+ interactions", Biochem. **46**, 10266–10278 (2007) 10.1021/bi062284r.

[12] D. Thirumalai and C. Hyeon, "RNA and protein folding: common themes and variations.", Biochem. **44**, 4957–4970 (2005) 10.1021/bi047314+.

[13] R. Börner, D. Kowerko, H. G. Miserachs, M. F. Schaffer, and R. K. O. Sigel, "Metal ion induced heterogeneity in RNA folding studied by smFRET", Coord. Chem. Rev. **327**, 123–142 (2016) 10.1016/j.ccr.2016.06.002.

[14] A. A. Chen and A. E. Garcia, "High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations", Proc. Natl. Acad. Sci. USA **110**, 16820–16825 (2013) 10.1073/pnas.1309392110.

[15] J. C. Miner, A. A. Chen, and A. E. García, "Free-energy landscape of a hyperstable RNA tetraloop", Proc. Natl. Acad. Sci. USA **113**, 6665–6670 (2016) 10.1073/pnas.1603154113.

[16] B. G. Keller, A. Kobitski, A. Jäschke, G. U. Nienhaus, and F. Noé, "Complex RNA folding kinetics revealed by single-molecule FRET and hidden Markov models", J. Am. Chem. Soc. **136**, 4534–4543 (2014) 10.1021/ja4098719.

[17] S. E. McDowell, J. M. Jun, and N. G. Walter, "Long-range tertiary interactions in single hammerhead ribozymes bias motional sampling toward catalytically active conformations", RNA **16**, 2414–2426 (2010) 10.1261/rna.1829110.

[18] J. H. Roh, M. Tyagi, R. M. Briber, S. A. Woodson, and A. P. Sokolov, "The dynamics of unfolded versus folded tRNA: The role of electrostatic interactions", J. Am. Chem. Soc. **133**, 16406–16409 (2011) 10.1021/ja207667u.

[19] N. A. Denesyuk and D. Thirumalai, "Coarse-grained model for predicting RNA folding thermodynamics", J. Phys. Chem. B **117**, 4901–4911 (2013) 10.1021/jp401087x.

[20] S. Cao and S. J. Chen, "Predicting RNA folding thermodynamics with a reduced chain representation model", RNA **11**, 1884–1897 (2005) 10.1261/rna.2109105.

[21] J. Yoon, J. C. Lin, C. Hyeon, and D. Thirumalai, "Dynamical transition and heterogeneous hydration dynamics in rna", J. Phys. Chem. B **118**, 7910–7919 (2014) 10.1021/jp500643u.

[22] M. Egli, S. Portmann, and N. Usman, "RNA Hydration: A Detailed Look", Biochemistry **35**, 8489–8494 (1996) 10.1021/bi9607214.

[23] R. M. Fuoss, A. Katchalsky, and S. Lifson, "The potential of an infinite rod-like molecule and the distribution of the counter ions", Proc. Natl. Acad. Sci. USA **37**, 579–589 (1951) 10.1073/pnas.37.9.579.

[24] B. Honig and A. Nicholls, "Classical electrostatics in biology and chemistry", Science (80-. ). **268**, 1144–1149 (1995) 10.1126/science.7761829.

[25] J. Forsman, "A simple correlation-corrected poisson-boltzmann theory", J. Phys. Chem. B **108**, 9236–9245 (2004) 10.1021/jp049571u.

[26] G. S. Manning and U. Mohanty, "Counterion condensation on ionic oligomers", Phys. A Stat. Mech. its Appl. **247**, 196–204 (1997) 10.1016/S0378-4371(97)00413-5.

[27] U. Mohanty, A. Spasic, H. D. Kim, and S. Chu, "Ion atmosphere of three-way junction nucleic acid", J. Phys. Chem. B **109**, 21369–21374 (2005) 10.1021/jp050005o.

[28] C. H. Taubes, U. Mohanty, and S. Chu, "Ion atmosphere around nucleic acid", J. Phys. Chem. B **109**, 21267–21272 (2005) 10.1021/jp051590w.

[29] H. T. Nguyen, N. Hori, and D. Thirumalai, "Theory and simulations for RNA folding in mixtures of monovalent and divalent cations", Proc. Natl. Acad. Sci. USA **116**, 21022–21030 (2019) 10.1073/pnas.1911632116.

[30] R. Kjellander, "Modified Debye-Hückel approximation with effective charges: An application of dressed ion theory for electrolyte solutions", J. Phys. Chem. **99**, 10392–10407 (1995) 10.1021/j100025a048.

[31] H. Frusawa and R. Hayakawa, "Generalizing the Debye-Hückel equation in terms of density functional integral", Phys. Rev. E **61**, R6097–R6082 (2000) 10.1103/PhysRevE.61.R6079.

[32] D. Bashford and D. A. Case, "Generalized Born Models of Macromolecular Solvation Effects", Annu. Rev. Phys. Chem., 7–12 (2000).

[33] R. R. Netz and H. Orland, "Beyond Poisson-Boltzmann: Fluctuation effects and correlation functions", Eur. Phys. J. E **1**, 203–214 (2000) 10.1007/s101890050023.

[34] P. Attard, "Electrolytes and the electric double layer", Adv. Chem. Phys. **92**, 1–159 (1996) 10.1002/9780470141519.ch1.

[35] Y. Z. Shi, F. H. Wang, Y. Y. Wu, and Z. J. Tan, "A coarse-grained model with implicit salt for RNAs: Predicting 3D structure, stability and salt effect", J. Chem. Phys. **141**, 105102 1—13 (2014) 10.1063/1.4894752.

[36] J. Åqvist, "Ion-water interaction potentials derived from free energy perturbation simulations", J. Phys. Chem. **94**, 8021–8024 (1990) 10.1021/j100384a009.

[37] I. S. Joung and T. E. Cheatham, III, "Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations", J. Phys. Chem. B **112**, 9020–9041 (2008) 10.1021/jp8001614.

[38] A. A. Chen and R. V. Pappu, "Parameters of Monovalent Ions in the AMBER-99 Forcefield:Assessment of Inaccuracies and Proposed Improvements", J. Phys. Chem. B **111**, 11884–11887 (2007).

[39] S. Kirmizialtin, S. A. Pabit, S. P. Meisburger, L. Pollack, and R. Elber, "RNA and its ionic cloud: Solution scattering experiments and atomically detailed simulations", Biophys. J. **102**, 819–828 (2012) 10.1016/j.bpj.2012.01.013.

[40] J. C. Miner and A. E. García, "Concentration-dependent and configuration-dependent interactions of monovalent ions with an RNA tetraloop", J. Chem. Phys. **148**, 10.1063/1.5019939 (2018) 10.1063/1.5019939.

[41] A. Srivastava, R. Timsina, S. Heo, S. W. Dewage, S. Kirmizialtin, and X. Qiu, "Structure-guided DNA–DNA attraction mediated by divalent cations", Nucleic Acid Res. **48**, 7018–7026 (2020) 10.1093/nar/gkaa499.

[42] G. L. Conn, "Crystal Structure of a Conserved Ribosomal Protein-RNA Complex", Science (80-. ). **284**, 1171–1174 (1999) 10.1126/science.284.5417.1171.

[43] V. K. Misra and D. E. Draper, "A thermodynamic framework for $Mg^{2+}$ binding to RNA.", Proc. Natl. Acad. Sci. USA **98**, 12456–12461 (2001) 10.1073/pnas.221234598.

[44] D. Leipply and D. E. Draper, "Effects of $Mg^{2+}$ on the free energy landscape for folding a purine riboswitch RNA", Biochem. **50**, 2790–2799 (2011) 10.1021/bi101948k.

[45] R. L. Hayes, J. K. Noel, A. Mandic, P. C. Whitford, K. Y. Sanbonmatsu, U. Mohanty, and J. N. Onuchic, "Generalized Manning condensation model captures the RNA ion atmosphere", Phys. Rev. Lett. **114**, 258105 1—6 (2015) 10.1103/PhysRevLett.114.258105.

[46] H. Lammert, A. Wang, U. Mohanty, and J. N. Onuchic, "RNA as a complex polymer with coupled dynamics of ions and water in the outer solvation sphere", J. Phys. Chem. B **122**, 11218–11227 (2018) 10.1021/acs.jpcb.8b06874.

[47] R. L. Hayes, J. K. Noel, U. Mohanty, P. C. Whitford, S. P. Hennelly, J. N. Onuchic, and K. Y. Sanbonmatsu, "Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch", J. Am. Chem. Soc. **134**, 12043–12053 (2012) 10.1021/ja301454u.

[48] S. C. Abeysirigunawardena and S. A. Woodson, "Differential effects of ribosomal proteins and Mg2+ ions on a conformational switch during 30S ribosome 5'-domain assembly", RNA **21**, 1859–1865 (2015) 10.1261/rna.051292.115.

[49] H. D. Kim, G. U. Nienhaus, T. Ha, J. W. Orr, J. R. Williamson, and S. Chu, "Mg2+-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules", Proc. Natl. Acad. Sci. USA **99**, 4284–4289 (2002) 10.1073/pnas.032077799.

[50] D. J. Klein, P. B. Moore, and T. A. Steitz, "The contribution of metal ions to the structural stability of the large ribosomal subunit", RNA **10**, 1366–1379 (2004) 10.1261/rna.7390804.

[51] R. F. Gesteland, "Unfolding of Escherichia coli ribosomes by removal of magnesium", J. Mol. Biol. **18**, 356–371 (1966) 10.1016/S0022-2836(66)80253-X.

[52] H. D. Kim, J. D. Puglisi, and S. Chu, "Fluctuations of transfer RNAs between classical and hybrid states", Biophys. J. **93**, 3575–3582 (2007) 10.1529/biophysj.107.109884.

[53] W. Holtkamp, C. E. Cunha, F. Peske, A. L. Konevega, W. Wintermeyer, and M. V. Rodnina, "GTP hydrolysis by EF-G synchronizes tRNA movement on small and large ribosomal subunits", EMBO J. **33**, 1073–1085 (2014) 10.1002/embj.201387465.

[54] I. Wohlgemuth, C. Pohl, and M. V. Rodnina, "Optimization of speed and accuracy of decoding in translation", EMBO J. **29**, 3701–3709 (2010) 10.1038/emboj.2010.229.

[55] M. Johansson, J. Zhang, and M. Ehrenberg, "Genetic code translation displays a linear trade-off between efficiency and accuracy of tRNA selection", Proc. Natl. Acad. Sci. USA **109**, 131–136 (2012) 10.1073/pnas.1116480109.

[56] A. Tissières, J. D. Watson, D. Schlessinger, and B. R. Hollingworth, "Ribonucleoprotein particles from Escherichia coli", J. Mol. Biol. **1**, 221–233 (1959) 10.1016/S0022-2836(59)80029-2.

[57] F. C. Chao and H. K. Schachman, "The isolation and characterization of a macromolecular ribonucleoprotein from yeast", Arch. Biochem. Biophys. **61**, 220–230 (1956) 10.1016/0003-9861(56)90334-4.

[58] F. C. Chao, "Dissociation of macromolecular ribonucleoprotein of yeast", Arch. Biochem. Biophys. **70**, 426–431 (1957) 10.1016/0003-9861(57)90130-3.

[59] A. Rozov, I. Khusainov, K. El Omari, R. Duman, V. Mykhaylyk, M. Yusupov, E. Westhof, A. Wagner, and G. Yusupova, "Importance of potassium ions for ribosome structure and function revealed by long-wavelength X-ray diffraction", Nat. Commun. **10**, 2519 (2019) 10.1038/s41467-019-10409-4.

[60] L. V. Bock, M. H. Kolář, and H. Grubmüller, "Molecular simulations of the ribosome and associated translation factors", Curr. Opin. Struct. Biol. **49**, 27–35 (2018) 10.1016/j.sbi.2017.11.003.

[61] M. Levi, K. Nguyen, L. Dukaye, and P. C. Whitford, "Quantifying the relationship between single-molecule probes and subunit rotation in the ribosome", Biophys. J. **113**, 2777–2786 (2017) 10.1016/j.bpj.2017.10.021.

[62] A. Deckert, C. A. Waudby, T. Wlodarski, A. S. Wentink, X. Wang, J. P. Kirkpatrick, J. F. S. Paton, C. Camilloni, P. Kukic, C. M. Dobson, M. Vendruscolo, L. D. Cabrita, and J. Christodoulou, "Structural characterization of the interaction of $\alpha$-synuclein nascent chains with the ribosomal surface and trigger factor", Proc. Natl. Acad. Sci. U. S. A. **113**, 5012–5017 (2016) 10.1073/pnas.1519124113.

[63] J. K. Noel and P. C. Whitford, "How EF-Tu can contribute to efficient proofreading of aa-tRNA by the ribosome", Nat. Commun. **7**, 13314 (2016) 10.1038/ncomms13314.

[64] K. Nguyen and P. C. Whitford, "Steric interactions lead to collective tilting motion in the ribosome during mRNA-tRNA translocation", Nat. Commun. **7**, 1–9 (2016) 10.1038/ncomms10586.

[65] L. V. Bock, C. Blau, G. F. Schröder, I. I. Davydov, N. Fischer, H. Stark, M. V. Rodnina, A. C. Vaiana, and H. Grubmüller, "Energy barriers and driving forces in tRNA translocation through the ribosome", Nat. Struct. Mol. Biol. **20**, 1390–1396 (2013).

[66] P. C. Whitford, S. C. Blanchard, J. H. D. Cate, and K. Y. Sanbonmatsu, "Connecting the kinetics and energy landscape of tRNA translocation on the ribosome", PLoS Comput. Biol. **9**, edited by S.-J. Chen, e1003003 (2013) 10.1371/journal.pcbi.1003003.

[67] S. Arenz, L. V. Bock, M. Graf, C. A. Innis, R. Beckmann, H. Grubmüller, A. C. Vaiana, and D. N. Wilson, "A combined cryo-EM and molecular dynamics approach reveals the mechanism of ErmBL-mediated translation arrest", Nat. Commun. **7**, 10.1038/ncomms12026 (2016) 10.1038/ncomms12026.

[68] L. G. Trabuco, E. Schreiner, J. Eargle, P. Cornish, T. Ha, Z. Luthey-Schulten, and K. Schulten, "The role of L1 stalk-tRNA interaction in the ribosome elongation cycle", J. Mol. Biol. **402**, 741–760 (2010) 10.1016/j.jmb.2010.07.056.

[69] M. C. Small, P. Lopes, R. B. Andrade, and A. D. MacKerell, "Impact of Ribosomal Modification on the Binding of the Antibiotic Telithromycin Using a Combined Grand Canonical Monte Carlo/Molecular Dynamics Simulation Approach", PLoS Comput. Biol. **9**, 26–28 (2013) 10.1371/journal.pcbi.1003113.

[70] P. G. Debenedetti, T. M. Truskett, C. P. Lewis, and F. H. Stillinger, "Theory of supercooled liquids and glasses: Energy landscape and statistical geometry perspectives", Adv. Chem. Eng. **28**, 21–79 (2001) 10.1016/s0065-2377(01)28003-x.

[71] N. Goldenfeld, *Lectures on phase transitions and the renormalization group* (Westview Press, 1992).

[72] J. C. Dyre, "Colloquium: The glass transition and elastic models of glass-forming liquids", Rev. Mod. Phys. **78**, 953–972 (2006) 10.1103/RevModPhys.78.953.

[73] O. V. Mazurin, "Relaxation phenomena in glass", J. Non. Cryst. Solids **25**, 129–169 (1977) 10.1016/0022-3093(77)90092-8.

[74] G. W. Scherer, "Theories of relaxation", J. Non. Cryst. Solids **123**, 75–89 (1990) 10.1016/0022-3093(90)90775-H.

[75] G. Williams and D. C. Watts, "Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function", Trans. Faraday Soc. **66**, 80 (1970) 10.1039/tf9706600080.

[76] H. Vogel, "Das temperaturabhängigkeitsgesetz der viskosität von flüssigkeiten", Phys. Z. **22**, 645–646 (1921).

[77] G. S. Fulcher, "Analysis of recent measurements of the viscosity of glasses", J. Am. Ceram. Soc. **8**, 339–355 (1925).

[78] G. Tammann and W. Hesse, "Die Abhängigkeit der Viscosität von der Temperatur bie unterkühlten Flüssigkeiten", Zeitschrift für Anorg. und Allg. Chemie **156**, 245–257 (1926).

[79] G. W. Scherer, "Relaxation in glass and composites", (1986).

[80] I. M. Hodge, "Enthalpy relaxation and recovery in amorphous materials", Journal of Non-Crystalline Solids **169**, 211–266 (1994).

[81] I. Avramov, "Kinetics of structural relaxation of glass-forming melts", Thermochimica acta **280**, 363–382 (1996).

[82] G. Adam and J. H. Gibbs, "On the temperature dependence of cooperative relaxation properties in glass-forming liquids", J. Chem. Phys. **43**, 139–146 (1965) 10.1063/1.1696442.

[83] U. Mohanty, I. A. Oppenheim, and C. H. Taubes, "Low-temperature relaxation and entropic barriers in supercooled liquids", Sci. (80-. ). **266**, 425–427 (1994) 10.1126/science.266.5184.425.

[84] X. Xia and P. G. Wolynes, "Fragilities of liquids predicted from the random first order transition theory of glasses", Proc. Natl. Acad. Sci. U. S. A. **97**, 2990–2994 (2000) 10.1073/pnas.97.7.2990.

[85] J. C. Mauro, Y. Yue, A. J. Ellison, P. K. Gupta, and D. C. Allan, "Viscosity of glass-forming liquids", Proc. Natl. Acad. Sci. U. S. A. **106**, 19780–19784 (2009) 10.1073/pnas.0911705106.

[86] S. Samanta and R. Richert, "Electrorheological Source of Nonlinear Dielectric Effects in Molecular Glass-Forming Liquids", J. Phys. Chem. B **120**, 7737–7744 (2016) 10.1021/acs.jpcb.6b04903.

[87] U. Kothe and M. V. Rodnina, "Delayed release of inorganic phosphate from elongation factor Tu following GTP hydrolysis on the ribosome", Biochemistry **45**, 12767–12774 (2006) 10.1021/bi061192z.

[88] P. C. Whitford, P. Geggier, R. B. Altman, S. C. Blanchard, J. N. Onuchic, and K. Y. Sanbonmatsu, "Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways", RNA **16**, 1196–1204 (2010) 10.1261/rna.2035410.

[89] Y. S. Polikanov, A. L. Starosta, M. F. Juette, R. B. Altman, D. S. Terry, W. Lu, B. J. Burnett, G. Dinos, K. A. Reynolds, S. C. Blanchard, T. A. Steitz, and D. N. Wilson, "Distinct tRNA Accommodation Intermediates Observed on the Ribosome with the Antibiotics Hygromycin A and A201A", Mol. Cell **58**, 832–844 (2015) 10.1016/j.molcel.2015.04.014.

[90] P. C. Whitford, J. N. Onuchic, and K. Y. Sanbonmatsu, "Connecting energy landscapes with experimental rates for aminoacyl-tRNA accommodation in the ribosome", J. Am. Chem. Soc. **132**, 13170–13171 (2010) 10.1021/ja1061399.

[91] H. Yang, J. K. Noel, and P. C. Whitford, "Anisotropic Fluctuations in the Ribosome Determine tRNA Kinetics", J. Phys. Chem. B **121**, 10593–10601 (2017) 10.1021/acs.jpcb.7b06828.

[92] S. A. Woodson, "Metal ions and RNA folding: a highly charged topic with a dynamic future", Curr. Opin. Chem. Biol. **9**, 104–109 (2005).

[93] H. M. Al-Hashimi and N. G. Walter, "RNA dynamics: it is about time", Curr. Opin. Struct. Biol. **18**, 321–329 (2008) 10.1016/j.sbi.2008.04.004.

[94] Y. Wan, K. Qu, Z. Ouyang, M. Kertesz, J. Li, R. Tibshirani, D. L. Makino, R. C. Nutter, E. Segal, and H. Y. Chang, "Genome-wide measurement of RNA folding energies", Mol. Cell **48**, 169–181 (2012) 10.1016/j.molcel.2012.08.008.

[95] S. A. Pabit, J. L. Sutton, H. Chen, and L. Pollack, "Role of Ion Valence in the Submillisecond Collapse and Folding of a Small RNA Domain", Biochemistry **52**, 1539–1546 (2013) 10.1021/bi3016636.

[96] X. Shi, L. Huang, D. M. J. Lilley, P. B. Harbury, and D. Herschlag, "The solution structural ensembles of RNA kink-turn motifs and their protein complexes", Nat. Chem. Biol. **12**, 146–152 (2016) 10.1038/nchembio.1997.

[97] J. Lee, C. E. Vogt, M. McBrairty, and H. M. Al-Hashimi, "Influence of Dimethylsulfoxide on RNA Structure and Ligand Binding", Anal. Chem. **85**, 9692–9698 (2013) 10.1021/ac402038t.

[98] M. H. Bailor, X. Sun, and H. M. Al-Hashimi, "Topology links RNA secondary structure with global conformation, dynamics, and adaptation", Science (80-. ). **327**, 202–206 (2010) 10.1126/science.1181085.

[99] J. C. Bowman, T. K. Lenz, N. V. Hud, and L. D. Williams, "Cations in charge: magnesium ions in RNA folding and catalysis", Curr. Opin. Struct. Biol. **22**, 262–272 (2012) 10.1016/j.sbi.2012.04.006.

[100] W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics", J. Mol. Graph. **14**, 33–38 (1996) 10.1016/0263-7855(96)00018-5.

[101] A. Bleuzen, P. A. Pittet, L. Helm, and A. E. Merbach, "Water exchange on magnesium(II) in aqueous solution: A variable temperature and pressure 17O NMR study", Magn. Reson. Chem. **35**, 765–773 (1997) 10.1002/(SICI)1097-458X(199711)35:11<765::AID-OMR169>3.0.CO;2-F.

[102] O. Allnér, L. Nilsson, and A. Villa, "Magnesium ion-water coordination and exchange in biomolecular simulations", J. Chem. Theory Comput. **8**, 1493–1502 (2012) 10.1021/ct3000734.

[103] C. Bergonzo, K. B. Hall, and T. E. Cheatham, "Divalent Ion Dependent Conformational Changes in an RNA Stem-Loop Observed by Molecular Dynamics", J. Chem. Theory Comput. **12**, 3382–3389 (2016) 10.1021/acs.jctc.6b00173.

[104] C. A. Hanke and H. Gohlke, "Force Field Dependence of Riboswitch Dynamics", in *Methods enzymol.* (Elsevier, 2015), pp. 163–191, 10.1016/bs.mie.2014.10.056.

[105] A. C. Vaiana and K. Y. Sanbonmatsu, "Stochastic Gating and Drug-Ribosome Interactions", J Mol Biol **386**, 648–661 (2009) 10.1016/j.jmb.2008.12.035.

[106] K. Y. Sanbonmatsu, S. Joseph, and C.-S. Tung, "Simulating movement of tRNA into the ribosome during decoding", Proc. Natl. Acad. Sci. **102**, 15854–15859 (2005) 10.1073/pnas.0503456102.

[107] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers", SoftwareX **1-2**, 19–25 (2015) 10.1016/j.softx.2015.06.001.

[108] J. 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?", J. Comp. Chem. **21**, 1049–1074 (2000).

[109] A. Pérez, I. Marchán, 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**, 3817–3829 (2007) 10.1529/biophysj.106.097782.

[110] M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham, and P. Jurečka, "Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles", J. Chem. Theory Comput. **7**, 2886–2902 (2011) 10.1021/ct200162x.

[111] K. L. Meagher, L. T. Redman, and H. A. Carlson, "Development of polyphosphate parameters for use with the AMBER force field", J. Comput. Chem. **24**, 1016–1025 (2003) 10.1002/jcc.10262.

[112] 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) 10.1021/j100308a038.

[113] 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) 10.1021/ja00131a018.

[114] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems", J. Chem. Phys. **98**, 10089–10092 (1993) 10.1063/1.464397.

[115] Y. V. Bukhman and D. E. Draper, "Affinities and selectivities of divalent cation binding sites within an RNA tertiary structure 1 1Edited by I. Tinoco", J. Mol. Biol. **273**, 1020–1031 (1997) 10.1006/jmbi.1997.1383.

[116] J. K. Noel, M. Levi, M. Raghunathan, H. Lammert, R. L. Hayes, J. N. Onuchic, and P. C. Whitford, "SMOG 2: A versatile software package for generating structure-based models", PLoS Comput. Biol. **12**, edited by A. Prlic, e1004794 1—14 (2016) 10.1371/journal.pcbi.1004794.

[117] A. Luzar and D. Chandler, "Effect of Environment on Hydrogen Bond Dynamics in Liquid Water", Phys. Rev. Lett. **76**, 928–931 (1996) 10.1103/physrevlett.76.928.

[118] R. Kumar, J. R. Schmidt, and J. L. Skinner, "Hydrogen bonding definitions and dynamics in liquid water", J. Chem. Phys. **126**, 204107 (2007) 10.1063/1.2742385.

[119] A. E. García, "Large-amplitude nonlinear motions in proteins", Phys. Rev. Lett. **68**, 2696–2699 (1992) 10.1103/physrevlett.68.2696.

[120] A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, "Essential dynamics of proteins", Proteins Struct. Funct. Genet. **17**, 412–425 (1993) 10.1002/prot.340170408.

[121] S. Hayward and N. Go, "Collective Variable Description of Native Protein Dynamics", Annu. Rev. Phys. Chem. **46**, 223–250 (1995) 10.1146/annurev.pc.46.100195.001255.

[122] R. L. Hayes, J. K. Noel, P. C. Whitford, U. Mohanty, K. Y. Sanbonmatsu, and J. N. Onuchic, "Reduced model captures $Mg^{2+}$-RNA interaction free energy of riboswitches", Biophys. J. **106**, 1508–1519 (2014) 10.1016/j.bpj.2014.01.042.

[123] E. Koculi, C. Hyeon, D. Thirumalai, and S. A. Woodson, "Charge Density of Divalent Metal Cations Determines RNA Stability", J. Am. Chem. Soc. **129**, 2676–2682 (2007) 10.1021/ja068027r.

[124] J. A. Lemkul, S. K. Lakkaraju, and A. D. MacKerell, "Characterization of Mg2+ Distributions around RNA in Solution", ACS Omega **1**, 680–688 (2016) 10.1021/acsomega.6b00241.

[125] L.-Z. Sun, C. Kranawetter, X. Heng, and S.-J. Chen, "Predicting Ion Effects in an RNA Conformational Equilibrium", J. Phys. Chem. B **121**, 8026–8036 (2017) 10.1021/acs.jpcb.7b03873.

[126] A. Singh, L. Sethaphong, and Y. Yingling, "Interactions of Cations with RNA Loop-Loop Complexes", Biophys. J. **101**, 727–735 (2011) 10.1016/j.bpj.2011.06.033.

[127] R. Schroeder, A. Barta, and K. Semrad, "Strategies for RNA folding and assembly", Nat. Rev. Mol. Cell Biol. **5**, 908–919 (2004) 10.1038/nrm1497.

[128] R. M. Izatt, J. J. Christensen, and J. H. Rytting, "Sites and thermodynamic quantities associated with proton and metal ion interaction with ribonucleic acid, deoxyribonucleic acid, and their constituent bases, nucleosides, and nucleotides", Chem. Rev. **71**, 439–481 (1971) 10.1021/cr60273a002.

[129] M. H. Shamsi and H. B. Kraatz, "Interactions of Metal Ions with DNA and Some Applications", J. Inorg. Organomet. Polym. Mater. **23**, 4–23 (2013) 10.1007/s10904-012-9694-8.

[130] J. A. Cowan, "Coordination Chemistry of Mg2+ and 5S rRNA (Escherichia coli): Binding Parameters, Ligand Symmetry, and Implications for Activity", J. Am. Chem. Soc. **113**, 675–676 (1991) 10.1021/ja00002a046.

[131] S. S. Reid and J. A. Cowan, "Biostructural Chemistry of Magnesium Ion: Characterization of the Weak Binding Sites on tRNAPhe(yeast). Implications for Conformational Change and Activity", Biochemistry **29**, 6025–6032 (1990) 10.1021/bi00477a021.

[132] D. Kowerko, S. L. König, M. Skilandat, D. Kruschel, M. C. Hadzic, L. Cardo, and R. K. Sigel, "Cation-induced kinetic heterogeneity of the intron-exon recognition in single group II introns", Proc. Natl. Acad. Sci. U. S. A. **112**, 3403–3408 (2015) 10.1073/pnas.1322759112.

[133] M. Egli, S. Portmann, and N. Usman, "Rna hydration: A detailed look", Biochem. **35**, 8489–8494 (1996) 10.1021/bi9607214.

[134] C. Hyeon and D. Thirumalai, "Mechanical unfolding of RNA hairpins", Proc. Natl. Acad. Sci. USA **102**, 6789–6794 (2005) 10.1073/pnas.0408314102.

[135] Y. Liu, E. Haddadian, T. R. Sosnick, K. F. Freed, and H. Gong, "A novel implicit solvent model for simulating the molecular dynamics of RNA", Biophys. J. **105**, 1248–1257 (2013) 10.1016/j.bpj.2013.07.033.

[136] U. Mohanty, B. W. Ninham, and I. Oppenheim, "Dressed polyions, counterion condensation, and adsorption excess in polyelectrolyte solutions.", Proc. Natl. Acad. Sci. USA **93**, 4342–4344 (1996) 10.1073/pnas.93.9.4342.

[137] M. Schaefer and M. Karplus, "A comprehensive analytical treatment of continuum electrostatics", J. Phys. Chem. **100**, 1578–1599 (1996) 10.1021/jp9521621.

[138] A. Marcovitz and Y. Levy, "Frustration in protein-DNA binding influences conformational switching and target search kinetics", Proc. Natl. Acad. Sci. U. S. A. **108**, 17957–17962 (2011) 10.1073/pnas.1109594108.

[139] O. Givaty and Y. Levy, "Protein Sliding along DNA: Dynamics and Structural Characterization", J. Mol. Biol. **385**, 1087–1097 (2009) 10.1016/j.jmb.2008.11.016.

[140] G. C. L. Wong and L. Pollack, "Electrostatics of strongly charged biological polymers: Ion-mediated interactions and self-organization in nucleic acids and proteins", Annu. Rev. Phys. Chem. **61**, 171–189 (2010) 10.1146/annurev.physchem.58.032806.104436.

[141] A. Savelyev and G. A. Papoian, "Molecular renormalization group coarse-graining of electrolyte solutions: Application to aqueous NaCl and KCl", J. Phys. Chem. B **113**, 7785–7793 (2009) 10.1021/jp9005058.

[142] A. Savelyev and G. A. Papoian, "Chemically accurate coarse graining of double-stranded DNA", Proc. Natl. Acad. Sci. USA **107**, 20340–20345 (2010) 10.1073/pnas.1001163107.

[143] P. C. Whitford, J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu, and J. N. Onuchic, "An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields", Prot. Struct. Func. Bioinfo. **75**, 430–441 (2009) 10.1002/prot.22253.

[144] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the Amber ff99SB protein force field", Prot. Struct. Func. Bioinfo. **78**, 1950–1958 (2010) 10.1002/prot.22711.

[145] P. C. Whitford, W. Jiang, and P. Serwer, "Simulations of Phage T7 Capsid Expansion Reveal the Role of Molecular Sterics on Dynamics", Viruses **12**, 1273 (2020).

[146] J. K. Noel, P. C. Whitford, and J. N. Onuchic, "The shadow map: A general contact definition for capturing the dynamics of biomolecular folding and function", J. Phys. Chem. B **116**, 8692–8702 (2012).

[147] M. T. Record and C. F. Anderson, "Interpretation of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a two-domain model", Biophys. J. **68**, 786–794 (1995) 10.1016/S0006-3495(95)80254-7.

[148] J. Noeske, H. Schwalbe, and J. Wöhnert, "Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain", Nuc. Acids Res. **35**, 5262–5273 (2007) 10.1093/nar/gkm565.

[149] A. Serganov, Y.-R. Yuan, O. Pikovskaya, A. Polonskaia, L. Malinina, A. T. Phan, C. Hobartner, R. Micura, R. R. Breaker, and D. J. Patel, "Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs", Chem. Biol. **11**, 1729–1741 (2004) 10.1016/j.chembiol.2004.11.018.

[150] P. C. Whitford and J. N. Onuchic, "What protein folding teaches us about biological function and molecular machines", Curr. Opin. Struct. Biol. **30**, 57–62 (2015).

[151] M. Levi, J. K. Noel, and P. C. Whitford, "Studying ribosome dynamics with simplified models", Methods **162-163**, 128–140 (2019) 10.1016/j.ymeth.2019.03.023.

[152] K. Xi, F. H. Wang, G. Xiong, Z. L. Zhang, and Z. J. Tan, "Competitive binding of $Mg^{2+}$ and $Na^{+}$ ions to nucleic acids: from helices to tertiary structures", Biophys. J. **114**, 1776–1790 (2018) 10.1016/j.bpj.2018.03.001.

[153] J. Frank and R. L. Gonzalez Jr., "Structure and Dynamics of a Processive Brownian Motor: The Translating Ribosome", Annu. Rev. Biochem. **79**, 381–412 (2010).

[154] X. Agirrezabala, H. Y. Liao, E. Schreiner, J. Fu, R. F. Ortiz-Meoz, K. Schulten, R. Green, and J. Frank, "Structural characterization of mRNA-tRNA translocation intermediates", Proc. Natl. Acad. Sci. USA **109**, 6094–6099 (2012) 10.1073/pnas.1201288109.

[155] J. Fei, J. E. Bronson, J. M. Hofman, R. L. Srinivas, C. H. Wiggins, and R. L. Gonzalez, "Allosteric collaboration between elongation factor G and the ribosomal L1 stalk directs tRNA movements during translation", Proc. Natl. Acad. Sci. USA **106**, 15702–15707 (2009) 10.1073/pnas.0908077106.

[156] S. Mohan and H. F. Noller, "Recurring RNA structural motifs underlie the mechanics of L1 stalk movement", Nat. Commun. **8**, 1–11 (2017) 10.1038/ncomms14285.

[157] J. Fei, P. Kosuri, D. D. MacDougall, and R. L. Gonzalez, "Coupling of ribosomal L1 stalk and tRNA dynamics during translation elongation", Mol. Cell **30**, 348–359 (2008) 10.1016/j.molcel.2008.03.012.

[158] P. V. Cornish, D. N. Ermolenko, D. W. Staple, L. Hoang, R. P. Hickerson, H. F. Noller, and T. Ha, "Following movement of the L1 stalk between three functional states in single ribosomes", Proc. Natl. Acad. Sci. USA **106**, 2571–2576 (2009) 10.1073/pnas.0813180106.

[159] C. A. Angell, "The old problems of glass and the glass transition, and the many new twists.", Proc. Natl. Acad. Sci. **92**, 6675–6682 (1995) 10.1073/pnas.92.15.6675.

[160] K. F. Freed, "The Descent into Glass Formation in Polymer Fluids", Acc. Chem. Res. **44**, 194–203 (2011) 10.1021/ar100122w.

[161] H. Sillescu, "Heterogeneity at the glass transition: a review", J. Non. Cryst. Solids **243**, 81–108 (1999) 10.1016/s0022-3093(98)00831-x.

[162] A. N. Adhikari, N. A. Capurso, and D. Bingemann, "Heterogeneous dynamics and dynamic heterogeneities at the glass transition probed with single molecule spectroscopy", J. Chem. Phys. **127**, 114508 (2007) 10.1063/1.2768955.

[163] E. V. Russell and N. E. Israeloff, "Direct observation of molecular cooperativity near the glass transition", Nature **408**, 695–698 (2000) 10.1038/35047037.

[164] R. Böhmer, K. L. Ngai, C. A. Angell, and D. J. Plazek, "Nonexponential relaxations in strong and fragile glass formers", J. Chem. Phys. **99**, 4201–4209 (1993) 10.1063/1.466117.

[165] S. Samanta and R. Richert, "Dynamics of glass-forming liquids. XVIII. Does entropy control structural relaxation times?", J. Chem. Phys. **142**, 44504 (2015) 10.1063/1.4906191.

[166] R. H. Colby, "Dynamic scaling approach to glass formation", Phys. Rev. E **61**, 1783–1792 (2000) 10.1103/physreve.61.1783.

[167] W. Kauzmann, "The Nature of the Glassy State and the Behavior of Liquids at Low Temperatures.", Chem. Rev. **43**, 219–256 (1948) 10.1021/cr60135a002.

[168] C. A. Angell, "Entropy and fragility in supercooling liquids", J. Res. Natl. Inst. Stand. Technol. **102**, 171 (1997) 10.6028/jres.102.013.

[169] U. Mohanty, "Supercooled Liquids", in *Adv. chem. phys.* (John Wiley & Sons, Ltd, 1994), pp. 89–158, 10.1002/9780470141489.ch2.

[170] I. M. Hodge, "Adam-Gibbs formulation of enthalpy relaxation near the glass transition", J. Res. Natl. Inst. Stand. Technol. **102**, 195 (1997) 10.6028/jres.102.015.

[171] K. L. Ngai, "Modification of the Adam-Gibbs Model of Glass Transition for Consistency with Experimental Data", J. Phys. Chem. B **103**, 5895–5902 (1999) 10.1021/jp990594w.

[172] W.-S. Xu and K. F. Freed, "Generalized Entropy Theory of Glass Formation in Polymer Melts with Specific Interactions", Macromolecules **48**, 2333–2343 (2015) 10.1021/acs.macromol.5b00144.

[173] S. Capaccioli, G. Ruocco, and F. Zamponi, "Dynamically Correlated Regions and Configurational Entropy in Supercooled Liquids", J. Phys. Chem. B **112**, 10652–10658 (2008) 10.1021/jp802097u.

[174] R. Casalini and C. M. Roland, "Effect of entropy on the dynamics of supercooled liquids: new results from high pressure data", Philos. Mag. **87**, 459–467 (2007) 10.1080/14786430600936413.

[175] C. M. Roland, S. Capaccioli, M. Lucchesi, and R. Casalini, "Adam–Gibbs model for the supercooled dynamics in the ortho-terphenyl ortho-phenylphenol mixture", J. Chem. Phys. **120**, 10640–10646 (2004) 10.1063/1.1739394.

[176] D. Prevosto, M. Lucchesi, S. Capaccioli, R. Casalini, and P. A. Rolla, "Correlation between configurational entropy and structural relaxation time in glass-forming liquids", Phys. Rev. B **67**, 10.1103/physrevb.67.174202 (2003) 10.1103/physrevb.67.174202.

[177] C. T. Moynihan and J. Schroeder, "Non-exponential structural relaxation, anomalous light scattering and nanoscale inhomogeneities in glass-forming liquids", J. Non. Cryst. Solids **160**, 52–59 (1993) 10.1016/0022-3093(93)90283-4.

[178] U. Mohanty, "Inhomogeneities and relaxation in supercooled liquids", J. Chem. Phys. **100**, 5905–5909 (1994) 10.1063/1.467102.

[179] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, Oxford,New York, 1980).

[180] L.-M. Wang, C. A. Angell, and R. Richert, "Fragility and thermodynamics in nonpolymeric glass-forming liquids", J. Chem. Phys. **125**, 74505 (2006) 10.1063/1.2244551.

[181] U. Mohanty, "Fragility Metrics in Glass-Forming Liquids", in *Adv. chem. phys.* (John Wiley & Sons, Inc., Aug. 2004), pp. 71–103, 10.1002/047168077x.ch2.

[182] U. Mohanty, N. Craig, and J. T. Fourkas, "Relationship between kinetics and thermodynamics of supercooled liquids", J. Chem. Phys. **114**, 10577–10578 (2001) 10.1063/1.1381060.

[183] R. Richert and C. A. Angell, "Dynamics of glass-forming liquids. V. On the link between molecular dynamics and configurational entropy", J. Chem. Phys. **108**, 9016–9026 (1998) 10.1063/1.476348.

[184] K. L. Ngai, "Correlation between the secondary β-relaxation time atTgwith the Kohlrausch exponent of the primary α relaxation or the fragility of glass-forming materials", Phys. Rev. E **57**, 7346–7349 (1998) 10.1103/physreve.57.7346.

[185] L. Berthier, "Direct Experimental Evidence of a Growing Length Scale Accompanying the Glass Transition", Science (80-. ). **310**, 1797–1800 (2005) 10.1126/science.1120714.

[186] A. B. Bestul and S. S. Chang, "Excess Entropy at Glass Transformation", J. Chem. Phys. **40**, 3731–3733 (1964) 10.1063/1.1725086.

[187] S. S. Chang and A. B. Bestul, "Heat Capacity and Thermodynamic Properties ofo-Terphenyl Crystal, Glass, and Liquid", J. Chem. Phys. **56**, 503–516 (1972) 10.1063/1.1676895.

[188] C. A. Angell and S. Borick, "Specific heats Cp, Cv, Cconf and energy landscapes of glassforming liquids", J. Non. Cryst. Solids **307-310**, 393–406 (2002) 10.1016/s0022-3093(02)01500-4.

[189] R. Casalini, S. Capaccioli, M. Lucchesi, P. A. Rolla, and S. Corezzi, "Pressure dependence of structural relaxation time in terms of the Adam-Gibbs model", Phys. Rev. E **63**, 10.1103/physreve.63.031207 (2001) 10.1103/physreve.63.031207.

[190] J. C. Morse, D. Girodat, B. J. Burnett, M. Holm, R. B. Altman, K. Y. Sanbonmatsu, H. J. Wieden, and S. C. Blanchard, "Elongation factor-Tu can repetitively engage aminoacyl-tRNA within the ribosome during the proofreading stage of tRNA selection", Proc. Natl. Acad. Sci. U. S. A. **117**, 3610–3620 (2020) 10.1073/pnas.1904469117.

[191] R. C. Thompson and P. J. Stone, "Proofreading of the codon anticodon interaction on ribosomes", Proc. Natl. Acad. Sci. U. S. A. **74**, 198–202 (1977) 10.1073/pnas.74.1.198.

[192] T. Ruusala, M. Ehrenberg, and C. G. Kurland, "Is there proofreading during polypeptide synthesis?", EMBO J. **1**, 741–5 (1982).

[193] M. V. Rodnina, R. Fricke, L. Kuhn, and W. Wintermeyer, "Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome.", EMBO J. **14**, 2613–2619 (1995).

[194] Z. J. Guo, M. Gibson, S. Sitha, S. Chu, and U. Mohanty, "Role of large thermal fluctuations and magnesium ions in t-RNA selectivity of the ribosome", Proc. Natl. Acad. Sci. U. S. A. **108**, 3947–3951 (2011) 10.1073/pnas.1100671108.

[195] P. Geggier, R. Dave, M. B. Feldman, D. S. Terry, R. B. Altman, J. B. Munro, and S. C. Blanchard, "Conformational sampling of aminoacyl-tRNA during selection on the bacterial ribosome", J. Mol. Biol. **399**, 576–595 (2010) 10.1016/j.jmb.2010.04.038.

[196] A. B. Loveland, G. Demo, and A. A. Korostelev, "Cryo-EM of elongating ribosome with EF-Tu • GTP elucidates tRNA proofreading", Nature **584**, 640–645 (2020) 10.1038/s41586-020-2447-x.

[197] J. K. Noel, J. Chahine, V. B. P. Leite, and P. C. Whitford, "Capturing Transition Paths and Transition States for Conformational Rearrangements in the Ribosome", Biophys. J. **107**, 2881–2890 (2014) 10.1016/j.bpj.2014.10.022.

[198] H. Yang, P. Bandarkar, R. Horne, V. B. P. Leite, J. Chahine, and P. C. Whitford, "Diffusion of tRNA inside the ribosome is position-dependent", J. Chem. Phys. **151**, 10.1063/1.5113814 (2019) 10.1063/1.5113814.

[199] T. Martin Schmeing, R. M. Voorhees, A. C. Kelley, Y. G. Gao, F. V. Murphy IV, J. R. Weir, and V. Ramakrishnan, "The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA", Science (80-. ). **326**, 688–694 (2009) 10.1126/science.1179700.

[200] R. H. Swendsen, "Monte Carlo Renormalization Group", Phys. Rev. Lett. **42**, 859–861 (1979) 10.1103/PhysRevLett.42.859.

[201] A. P. Lyubartsev and A. Laaksonen, "Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach", Phys. Rev. E **52**, 3730–3737 (1995) 10.1103/PhysRevE.52.3730.

[202] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak, "Molecular dynamics with coupling to an external bath", J. Phys. Chem. **81**, 3684–3690 (1984) 10.1063/1.448118.

[203] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods", J. Chem. Phys. **81**, 511–519 (1984) 10.1063/1.447334.

[204] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions", Phys. Rev. A **31**, 1695–1697 (1985) 10.1103/PhysRevA.31.1695.

[205] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method", J. Appl. Phys. **52**, 7182–7190 (1981) 10.1063/1.328693.

[206] S. Nosé and M. L. Klein, "Constant pressure molecular dynamics for molecular systems", Mol. Phys. **50**, 1055–1076 (1983) 10.1080/00268978300102851.

[207] M. Lindahl, L. A. Svensson, A. Liljas, S. E. Sedelnikova, I. A. Eliseikina, N. P. Fomenkova, N. Nevskaya, S. V. Nikonov, M. B. Garber, T. A. Muranova, A. I. Rykonova, and R. Amons, "Crystal structure of the ribosomal protein S6 from Thermus thermophilus", EMBO J. **13**, 1249–1254 (1994) 10.1002/j.1460-2075.1994.tb06376.x.

[208] U. Mohanty, "Entropic Approach to Relaxation Behavior in Glass-Forming Liquids", in *Acs symp. ser.* Vol. 676 (Sept. 1997), pp. 95–109, 10.1021/bk-1997-0676.ch007.