4872–4882 *Nucleic Acids Research, 2018, Vol. 46, No. 10* *Published online 30 April 2018*
doi: 10.1093/nar/gky221

Influence of Na^+ and Mg^{2+} ions on RNA structures studied with molecular dynamics simulations

Nina M. Fischer[1], Marcelo D. Polêto[1,2], Jakob Steuer[1,3] and David van der Spoel[1,*]

[1]Uppsala Centre for Computational Chemistry, Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, SE-75124 Uppsala, Sweden, [2]Center of Biotechnology, Universidade Federal do Rio Grande do Sul, Bento Gonçalves 9500, BR-91500-970 Porto Alegre, Brazil and [3]Department of Chemistry, University of Konstanz, Universitätstraße 10, D-78457 Konstanz, Germany

Received August 14, 2017; Revised February 16, 2018; Editorial Decision March 09, 2018; Accepted April 23, 2018

ABSTRACT

The structure of ribonucleic acid (RNA) polymers is strongly dependent on the presence of, in particular Mg^{2+} cations to stabilize structural features. Only in high-resolution X-ray crystallography structures can ions be identified reliably. Here, we perform molecular dynamics simulations of 24 RNA structures with varying ion concentrations. Twelve of the structures were helical and the others complex folded. The aim of the study is to predict ion positions but also to evaluate the impact of different types of ions (Na^+ or Mg^{2+}) and the ionic strength on structural stability and variations of RNA. As a general conclusion Mg^{2+} is found to conserve the experimental structure better than Na^+ and, where experimental ion positions are available, they can be reproduced with reasonable accuracy. If a large surplus of ions is present the added electrostatic screening makes prediction of binding-sites less reproducible. Distinct differences in ion-binding between helical and complex folded structures are found. The strength of binding ($\Delta G^{\ddagger}$ for breaking RNA atom-ion interactions) is found to differ between roughly 10 and 26 kJ/mol for the different RNA atoms. Differences in stability between helical and complex folded structures and of the influence of metal ions on either are discussed.

INTRODUCTION

Positively charged ions play an essential role for the structural stability of RNA molecules. Especially, Mg^{2+} ions facilitate high structural complexity and folding arrangements that allow RNA molecules to perform various cellular functions (1). Apart from canonical functions assigned to RNA molecules such as being involved in protein synthesis, like messenger (m) or transfer (t) RNA, it is nowadays well established that RNAs act in many other biological processes. RNA molecules are for instance involved in gene regulation (e.g., small nuclear (sn), micro (mi) and small interfering (si) RNAs) and in enzymatic activity (e.g. ribozymes and ribonucleoprotein). In eukarya they also play a role in resistance to pathogenic and parasitic invaders (2,3). Some of these functions depend on the presence of metal ions. Mg^{2+} ions do not only stabilize specific RNA structures (4), but do also help to recognize binding partners and mediate catalytic processes (5–7). Hammerhead ribozymes are one well-known example that require metal ions to be present both for obtaining the correct three-dimensional fold and performing the ribozymes' function (8–12).

The need for positively charged ions in close proximity to RNA molecules is not surprising given their negatively charged backbone. Each RNA nucleotide contains one phosphate group that carries one negative charge. Positive ions shield negative charges on the RNA backbone by reducing repulsive forces, thereby allowing intramolecular interactions and compact RNA-biopolymers. RNA molecules are stabilized internally by hydrogen bonds between nucleotides in the same plane and by base stacking.

Positively charged ions can be divided into two main groups: (a) ions that bind to structurally well-defined sites in direct contact with or close to the RNA and (b) ions that form a cloud surrounding the RNA molecule (13). The classification of Mg^{2+} binding has been further refined (1,14,15): inner-sphere ions that form direct bonds with RNA atoms, outer-sphere ions that bind *via* a single hydration shell to the RNA, diffuse Mg^{2+} ions that bind *via* multiple hydration shells, and free ions where the RNA's charge has no direct effect on the ions. Monovalent ions can also be part of the first group and can bind sequence- specific to electronegative pockets formed by RNA structures (16–18); they should not just be considered to be part of a diffuse ionic cloud as described by Manning (19). Folding studies have shown that tRNA thermodynamic stability increases when adding monovalent (in particular Na^+ and K^+) and divalent (Mg^{2+}) ions (20–23). Mg^{2+} ions are however the

*To whom correspondence should be addressed. Tel: +46 18 471 4205; Fax: +46 018 530396; Email: david.vanderspoel@icm.uu.se

© The Author(s) 2018. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

most effective for stabilizing the native structure of RNA molecules (22,24).

For compensating negative RNA backbone charges, fewer divalent ions are required compared to the number of monovalent ions. In addition, divalent ions bind stronger and are able to bind several phosphate oxygen atoms at once. Finally, Mg^{2+} is small compared to other divalent ions and can bind to narrow well-defined pockets within RNA structures (13,25–30).

Despite the growing number of experimentally solved RNA structures, the number of structures with well-resolved ion binding sites is still limited (31,32). A non-redundant data set of RNA structures (May 2017) contains in total 1155 structures with a resolution better than 4 Å. Just 37% of the RNA structures, solved by X-ray crystallography include Mg^{2+} and 10% Na^+ ion positions. Less than one percent of the structures solved by NMR contain Mg^{2+} or Na^+ ions. An important reason for this is that Mg^{2+} ions, Na^+ ions, and water molecules have the same number of electrons. Thus, it is difficult to distinguish them from one another in electron density maps alone. They can only be assigned unambiguously in very high-resolution X-ray crystal structures (13,33–35). During the last years special NMR protocols have been developed to be able to study metal ion binding to RNA. However, a special sample preparation is needed to be able to detect Mg^{2+} and Na^+ ions (36).

Recent studies describe monovalent- and divalent ion binding and the influence of different ions on RNA structures. Molecular dynamics (MD) simulations of have been reported suggesting that many Mg^{2+} ions are strongly associated with RNA, but not directly bound (15). A review by Lipfert *et al.* (37) describes in detail the difference between direct binding of ions and longer range ('ion atmosphere') association to nucleic acids and how this influences structure and stability of RNA and DNA.

In this paper, we apply explicit solvent molecular dynamics simulations on a dataset of twenty-four RNA structures. Therefore, we can compare the implications of applying Na^+ or Mg^{2+} ions on helical and complex folded structures, solved by either X-ray crystallography or NMR techniques. By combining results from different structures and running simulations in triplicate we can statistically distinguish true effects of ions from stochastic fluctuations inherent to MD simulations. Obviously the findings are still dependent on force field quality and force fields for RNA have not been scrutinized (38) to the same extent as those for proteins (39).

MATERIALS AND METHODS

Dataset of RNA structures

A dataset of 24 RNA structures was selected consisting of twelve helical and twelve complex folded ones, (Table 1) from the Protein Data Bank (PDB, http://www.rcsb.org/pdb/) (40) (Supplementary Figure S1). In both groups six of the twelve structures were obtained by X-ray crystallography and six by NMR experiments. When present in experimental structures, we removed water molecules, ions, or other small molecules to ensure identical starting conditions for all structures in the dataset. In the case of 1D4R (41) we also deleted one separate single RNA strand (chain C) and in 1QC0 (42) the smaller RNA double helix (chain A and B).

We define helical structures as those that form one double helix composed of either one or two nucleotide strands. In these structures, there are only very few unpaired nucleotides present, i.e. single nucleotides at the strand's ends (1D4R (41), 4K31 (43) and 413D (44)). In the case of one nucleotide strand forming a double helix there are unpaired nucleotides in the loop region (1A4D (45), 2LPS (46), 2LV0 (47) and 2L2J (48)). In two structures, one single nucleotide is sticking out of the helical main structure (2LPS (46) and 2QEK (49)). In contrast to helical RNAs, complex folded RNA structures typically have a more globular shape, are often multi-helical RNAs, and may be categorized as, e.g., ribozymes or pseudoknots.

Ion selection and parameters

MD simulations that intend to mimic cellular conditions should use K^+ as monovalent ion, since it is the monovalent ion primarily found inside cells. However, we observed formation of salt crystals of K^+ and Cl^- ions in test simulations (data not shown). The same issue was previously described by other groups (65) when using K^+ instead of Na^+ ions. For this reason, we chose Na^+ as representative for monovalent ions in most of the simulations. Furthermore, we used for Mg^{2+} force field parameters refined by Allner *et al.* (66), that reproduce Mg^{2+} hydration free energies and exchange rates well.

Although it is known that the identity of the counter-ion matters to the composition of the ion cloud around nucleic acids (67) we have only used Cl^- ions in this work. Furthermore, we note that other cation force field parameters have been proposed for use in conjunction with nucleic acids (68), however rather than comparing many different ion parameter sets we here focus on comparing different RNA structures.

Molecular dynamics simulations

RNA topologies were built using the parm99 (69) force field with the GROMACS simulation package (version 4.6.7) (70). First, an *in vacuo* minimization was carried out. Then, each structure is placed in a rhombic dodecahedron box filled with TIP3P water molecules to reach a ratio of 550 water molecules/nucleotide. During this step four different systems are created (Supplementary Table S1), either with just counterions (CI) or at physiological salt (PS) concentration:

- Na^+_{CI}, just Na^+ counterions,
- Na^+_{PS}, Na^+ counterions plus 0.15 M/l NaCl,
- Mg^{2+}_{CI}, just Mg^{2+} counterions,
- Mg^{2+}_{PS}, Mg^{2+} counterions plus 0.15 M/l NaCl.

All ions were placed randomly in the simulation box. We monitored that Mg^{2+} ions maintain a specific initial distance (>2 Å to any RNA atom) to the RNA and that direct interactions did not occur during the minimization and first four ns of the equilibration phase.

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

Table 1. Dataset of 24 RNA structures, twelve X-ray (six helical and six complex folded) and twelve NMR (six helical and six complex folded) structures, taken from the Protein Data Bank (PDB, http://www.rcsb.org/pdb/) (40). The fraction of single nucleotides (Nuc.) is calculated for each structure. Helical structures that have only base pairs have a fraction of 0

Method/PDB id	Classification/system	Fraction unpaired	Nuc.	A	C	G	U	Mg^{2+}	K^+
X-ray	**Helical**								
1D4R (41)	Single recognition particle RNA	0.04	54	8	18	22	6	4	
1QC0 (42)	Plasmid copy control related RNA	0.00	38	6	13	13	6		
2QEK (49)	HIV-1 dimerization initiation site	0.04	46	14	12	12	8	1	3
4K31 (43)	rRNA A-site	0.05	44	8	12	14	10		
413D (44)	A-form RNA double helix	0.04	26	2	8	8	8		
420D (50)	RNA with A(anti)-G(syn) mispairs	0.00	32	10	6	8	8		
X-ray	**Complex folded**								
4B5R (51)	SAM-I riboswitch	0.30	94	28	22	32	12		
4FEJ (52)	Guanine riboswitch aptamer	0.31	67	16	17	17	17		
4FRG (53)	Cobalamin riboswitch aptamer	0.34	84	27	19	22	16	7	
4JF2 (54)	Class II preQ1 riboswitch	0.37	76	19	19	18	20	4	
4KQY (55)	S-box (SAM-I) riboswitch	0.34	119	40	27	33	19	2	
4P5J (56)	tRNA-like structure	0.36	83	17	28	22	16	2	
NMR	**Helical**								
1A4D (45)	Loop D/Loop E arm of 5S rRNA	0.07	41	6	10	17	8		
2D18 (57)	HIV-1 dimerization initiation site	0.00	34	8	10	12	4		
2KYD (58)	A-form RNA double helix	0.00	32	10	6	6	10		
2L2J (48)	R/G stem loop RNA	0.10	42	7	13	13	9		
2LPS (46)	ai5(gamma) group II intron	0.12	34	8	8	11	7		
2LV0 (47)	Stem–loop from 23S rRNA	0.17	24	8	4	6	6		
NMR	**Complex folded**								
1YMO (59)	Telomerase RNA pseudoknot	0.36	47	13	13	9	12		
2ADT (60)	Tetraloop–receptor complex	0.31	86	22	16	26	22		
2LKR (61)	U2/U6 snRNA	0.37	111	29	22	25	35		
2MHI (62)	CR4/5 domain of telomerase RNA	0.25	53	7	15	17	14		
2MTK (63)	Ribozyme's III-IV-V junction	0.36	47	7	13	16	11	6	
2M8K (64)	Telomerase RNA pseudoknot	0.30	48	12	9	6	21		
Total	**Average**		56	13	14	16	13		

All systems were minimized once more to eliminate any possible clashes and bad contacts. Subsequently, seven equilibration steps are carried out to provide a careful equilibration protocol. First, an NVT ensemble was conducted for 2 ns using position restraints with a force constant of 1000 kJ/(mol×nm^2)) to all heavy atoms. During this step the system was heated up to 300 K. Then, six NPT ensembles are conducted at 1 bar and 300 K for 26 ns in total. The number of restrained RNA atoms and the restraining force constant were gradually reduced while ions were given time to occupy preferred binding sites.

Finally, production runs were carried out for 50 ns at 300 K and 1 bar, with no restraints. An integration step of 2 fs was applied and all bonds were constrained using the LINCS (71,72) algorithm. A cutoff of 10 Å was used for Lennard–Jones and short-range Coulomb interactions and the particle mesh Ewald (PME) method (73) for long-range electrostatic interactions. Velocity rescaling (74) was used for temperature coupling with a time constant of 0.1 ps in order to ensure correct temperature fluctuations. For simulations at constant pressure we used the Parrinello-Rahman pressure coupling algorithm (75) with a time constant of 2 ps.

Each of the four systems was simulated three times in order to ensure statistical significance of our analyses. This resulted in 288 simulations and an overall simulation time of 14.4 μs for all production phases.

Force field evaluations

Since most of our simulations were done using a somewhat old force field, an updated set of force field parameters for nucleic acids was tested, namely parmbsc0 (82) in conjunction with parmOL (83,84) for a subset of four RNA struc-

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

tures. For this subset, we also evaluated the difference between K^+ and Na^+ counterions.

RMSD and ϵRMSD

The root mean square deviation (RMSD) was computed using GROMACS between the *in vacuo* minimized structure and the snapshots taken every 10 ps during the 50 ns production run (Supplementary Figures S4 and S5). The same was done to obtain the ϵRMSD values with a method developed by Bottaro *et al.* (76) (Supplementary Figures S6 and S7). Since each RNA structure was simulated three times in four different ionic conditions the mean and standard deviation of the RMSD values for each replica was determined. The mean RMSD values were subtracted from the individual values for all four systems for each structure. The 18 standardised RMSD/ϵRMSD values in each system Na^+_{PS} and Mg^{2+}_{PS} in all four structure groups were used to determine the p-values using a t-test between system Na^+_{PS} and system Mg^{2+}_{PS}. Finally, we calculated mean RMSD values over the three replicas and the error for each structure in each of the four ionic conditions.

Radial distribution functions

The radial distribution functions (RDFs) were determined using GROMACS and trajectories with structures taken every 10 ps of the 50 ns production run. For each of the four systems the RDFs are calculated between each of the RNA base (A-N1, A-N3, A-N6, A-N7, A-N9, G-O6, G-N1, G-N2, G-N3, G-N7, G-N9, C-O2, C-N1, C-N3, C-N4, U-O2, U-O4, U-N1 and U-N3), the two phosphate oxygen (O1P, O2P), or sugar oxygen (O2′, O3′, O4′ and O5′) atoms and positively charged ions (Na^+ or/and Mg^{2+}) present in the system. In all structures, O1P is the atom that points towards the solvent and O2P (particularly in helical structures) towards the minor groove. The average RDFs are calculated for each RNA atom over the 12 helical and 12 complex folded structures.

Free energy of activation

The same RNA atoms and positively charged ions as described in the RDF analysis were used to determine the free energy of activation $\Delta G^{\ddagger}$ for contact breaking. To specify the contact distance that is required as input parameter between an ion and a certain RNA atom we used the minimum between the first and second maxima from the corresponding RDF values for each structural replica. Similarly, the minimum between the second and third maxima determines the contact distance for second shell contacts. When no peak could be detected within a certain cutoff distance (3.5 or 6.0 Å) the contact distance for this RNA atom was calculated as the average over all minima of all other RNA atoms in this structure.

Ion binding sites in RNA structures

The occupancies of Na^+ and Mg^{2+} ions in close proximity of RNA structure were computed using the program MobyWat (80,81). Of the seven RNA structures with experimentally determined ion positions, structures were taken every 500 ps from the equilibration phase and every 250 ps from the production phase trajectory. These structures were superimposed to the experimental RNA structure while only considering atoms with <4 Å root mean square fluctuation (RMSF) values. RMSF values are obtained to the RNA structure closest to the average structure of the second half of the production run. The input parameters for MobyWat that differ to the default parameters are the following: the maximum and minimum distance limits were set to 6.0 and 1.0 Å and the clustering tolerance to 1.5 Å. The results are based on the MER clustering algorithm that yield the best results comparing experimental- and predicted ion binding sites. The top 50 predicted ion binding sites were used and RMSD values with respect to the experimental ones calculated for each of them. The predicted ion binding site with the smallest RMSD to an experimental ion binding site was considered as a potential binding site.

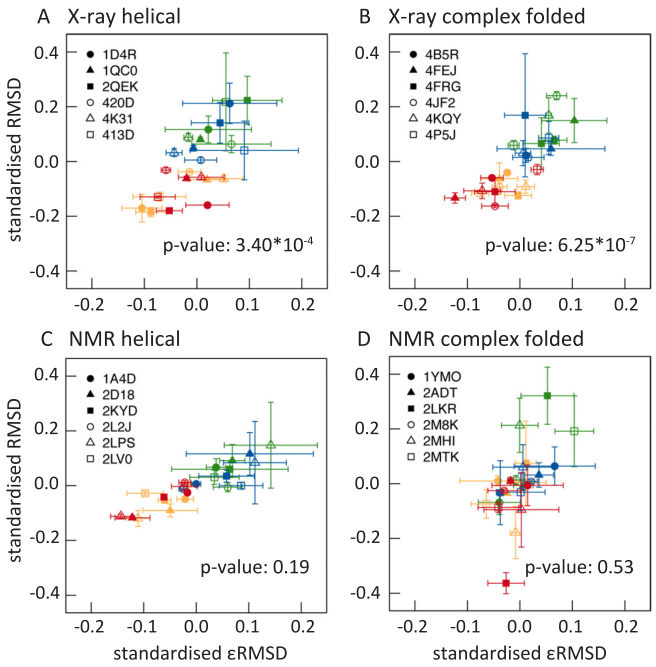

Figure 1. Standardized mean RMSD (nm) values are plotted against standardized mean ϵRMSD values for each RNA structure. RMSD and ϵRMSD values for helical X-ray and NMR as well as for complex folded X-ray and NMR structures are obtained during the production run for each replica and each system. First, an average value is determined over the production run RMSD values for all simulations. Second, a mean value over the three average replica values for each structure and system is calculated resulting in 4 x 24 data points with corresponding standard errors. In a last step, these data points are standardized by the mean over the four data points of each structure for each structure separately. This results in 4 x 24 data points with standard errors for RMSD and ϵRMSD values, which are plotted against each other. Each RNA structure is represented with a different symbol and the four systems are represented with: Na^+_{CI} (green), Na^+_{PS} (blue), Mg^{2+}_{CI} (orange), Mg^{2+}_{PS} (red). The p-values are calculated with a t-test between system Na^+_{PS} and system Mg^{2+}_{PS}.

RESULTS

Structural changes

Figure 1 illustrates how the surrounding environment, espe-

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

cially the presence or absence of Mg^{2+} ions, influences RNA structural changes during MD simulations. In agreement with other studies, root mean square deviation (RMSD) values are lower for RNA structures simulated with Mg^{2+} ions than without.

In addition to RMSD values, we calculated ϵRMSD values (76). This metric discriminates effectively between structurally and kinetically different RNA conformations. It directly describes variations in base-base interactions and therefore captures whether or not important structural characteristics, like base pairs, are preserved during the simulation. Bottaro *et al.* (76) showed that multiple different secondary RNA structures can be found within 4 Å RMSD of each other. Such two RNA structures with low RMSD values to a reference structure do not necessarily have the same secondary structures. Indeed, the base-base interactions could be completely lost in one structure and not in the other. This kind of structural differences is described by ϵRMSD values that takes structural information about base-pairing into account. An ϵRMSD of <0.8 indicates all base-base contacts are close to the native experimental structure and an ϵRMSD of >1 suggests non-native base-base contacts occur in the structure (76). There are more structures with an average (over three replicas) ϵRMSD >1 for systems simulated without Mg^{2+} than systems simulated with Mg^{2+} ions (Supplementary Table S2). The six complex folded NMR structures have almost always ϵRMSD values >1 regardless of the surrounding ionic environment, except for one structure (PDB id: 2ADT) and another structure (PDB id: 2MTK), when simulated with Mg^{2+} and NaCl. None of the structures obtained by X-ray crystallography have average ϵRMSD values >1 when Mg^{2+} ions were present during the simulation. In general, structures simulated with Mg^{2+} ions have lower RMSD and ϵRMSD values compared to structures simulated without Mg^{2+} ions.

We performed a statistical (kernel density) comparison test that compares the distributions of two-dimensional data points. It returns a p-value that is higher for better fits between the two distributions. Our null hypothesis is that the distributions are independent of whether Mg^{2+} ions are present in the simulations. When thus comparing simulations with a 0.15 M/l NaCl salt concentration with and without Mg^{2+} ions, the *P*-values for X-ray structures are both less than 0.001 (Figure 1 for individual *P*-values). This indicates that Mg^{2+} is significantly responsible for maintaining native base-base contacts during the simulations, at least for X-ray RNA structures. The *P*-values for NMR structures are higher and therefore statistically not significant.

Ion binding

In order to analyze where ions are located during the simulations, radial distribution functions (RDFs) were derived for positively charged ions present during the simulation and all RNA atoms. Direct contacts are identified between Na^+ and 16 RNA atoms (O1P, O2P, sugar oxygen atoms, A-N1, A-N3, A-N7, G-O6, G-N7, C-O2, C-N3, U-O2 and U-O4). Mg^{2+} ions form direct contacts in the simulations only twice with one of the RNA atoms. In one case Mg^{2+} binds directly to an O1P atom and in another structure to a cytosine oxygen (C-O2). All other Mg^{2+} interactions with RNA occur indirectly *via* water molecules. Clearly recognisable first and second shell contacts between positively charged ions and RNA atoms can be determined for seven RNA atoms in RDFs (Figure 2).

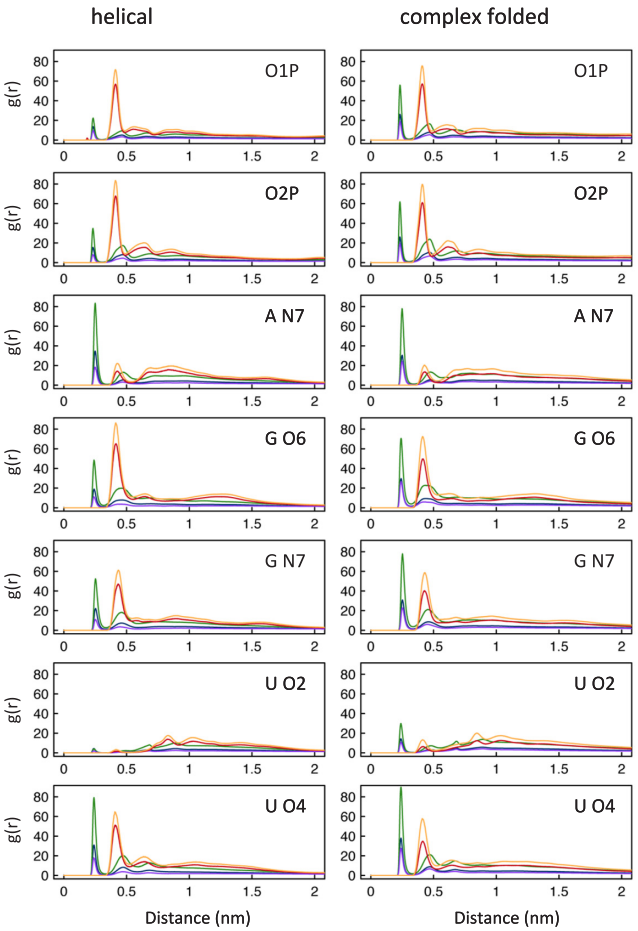

Figure 2. Average radial distribution functions (RDFs) for Na^+ and Mg^{2+} present in the simulations and seven RNA atoms (O1P, O2P, A N7, G O6, G N7, U O2, U O4). The average RDFs for 12 helical and 12 complex folded structures are colored according for each system: Na^+_{CI} (green), Na^+_{PS} (blue), Mg^{2+}_{CI} (orange) and in system Mg^{2+}_{PS} there are Mg^{2+} (red) and Na^+ (purple).

The RDF peaks are higher for Na^+ ions in system Na^+_{CI} compared to Na^+ ions in system Na^+_{PS} and Mg^{2+}_{PS} and for Mg^{2+} ions in system Mg^{2+}_{CI} compared to Mg^{2+} ions in system Mg^{2+}_{PS}. In both systems (Na^+_{CI} and Mg^{2+}_{CI}) fewer positively charged ions are therefore present in the bulk water surrounding the RNA molecule. We use the same arbitrary definition of bulk water/ions (distance >20 Å to any RNA atom) as described by Hayes *et al.* (15). In system Mg^{2+}_{PS} when both Na^+ and Mg^{2+} ions are present more Mg^{2+} ions are found in the bulk solvent than in system Mg^{2+}_{CI}. Nevertheless, fewer Na^+ ions are found close to the RNA in system Mg^{2+}_{PS} compared to system Na^+_{PS} and also system Na^+_{CI}. When comparing helical versus complex folded RDFs, fewer Na^+ ions are in direct contact with helical phosphate oxygen atoms compared to complex folded ones. In both structure groups there is a preference

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

for O2P over O1P for Na^+ as well as Mg^{2+} ions. For the nitrogen atom in adenine (A-N7) we observe only differences of the RDFs of Na^+ ions in the second and forth system between helical and complex folded structures. For A-N7 there seems to be a preference for Na^+ first shell binding compared to Mg^{2+} second shell binding interactions. When comparing Na^+ direct binding between helical and complex folded structures for both favoured guanine atoms (G-O6 and G-N7), there are more occurrences in complex folded structures with a slight preference for G N7. This is in contrast to second shell Mg^{2+} interactions, where G-O6 is the preferred atom. The peak for helical G-O6 atoms and Mg^{2+} ions in system $Mg^{2+}{}_{PS}$ is the highest determined for all RDFs. For U-O2 atoms very few ion contacts were found, and most of them are found in complex folded structures. Since helical structures mostly form Watson-Crick base pairs U-O2 atoms lie in the minor groove and it is likely therefore they not easily accessible to ions (Figure2). The other uracil oxygen atom (U-O4) is slightly preferred by Na^+ ions in first shell interactions in complex folded structures and by Mg^{2+} ions in second shell binding for helical structures.

Ion binding energetics

An analysis method that was developed to study kinetics of hydrogen bond breaking and forming (77) and thermodynamics of hydrogen bond breaking in different environments (78) was used here for studying ion-binding energetics. This method yields the Gibbs energy of activation $\Delta G^{\ddagger}$ for contact breaking and was previously applied on RNA-ion contacts in a study of viral RNA (79). The highest energy for breaking first shell contacts was found between one phosphate oxygen atom (O1P) and a Mg^{2+} ion (Figure 3). This is the result of one of the two direct interactions between a Mg^{2+} ion and an RNA atom that occurred in the simulations, as also observed in the RDF analysis. The energy of first shell contacts between Na^+ ions and RNA phosphate oxygen atoms (O1P and O2P) is not as high compared to other RNA atoms (A-N7, G-O6, G-N7, C-O2, C-N3, U-O2 and U-O4). These results differ from the RDF results insofar that the RDF peaks for C-O2, C-N3, and U-O2 atoms are very low especially compared to the peaks of phosphate oxygen atoms. The main difference between helical and complex folded first shell contacts is for atoms that are only available for interactions in complex folded structures (A-N3, A-N9, G-N9, U-N1). A-N3 lies in the minor groove in helical structures and the other atoms are the base atoms that are closest to the sugar ring. Therefore, they are not easily accessible to ions in helical RNA structures.

Ion binding positions

To investigate whether Mg^{2+} ions find experimentally identified binding sites, when initially placed randomly in the solvent (with a distance >2 Å to any RNA atom), we determined the occupancy of Na^+ and Mg^{2+} ions during the simulation using the software MobyWat (80,81).

Figure 4 shows the top 10 predicted binding sites for Mg^{2+} and Na^+ ions for one of the three replicas of each system during the equilibration phase superimposed on the

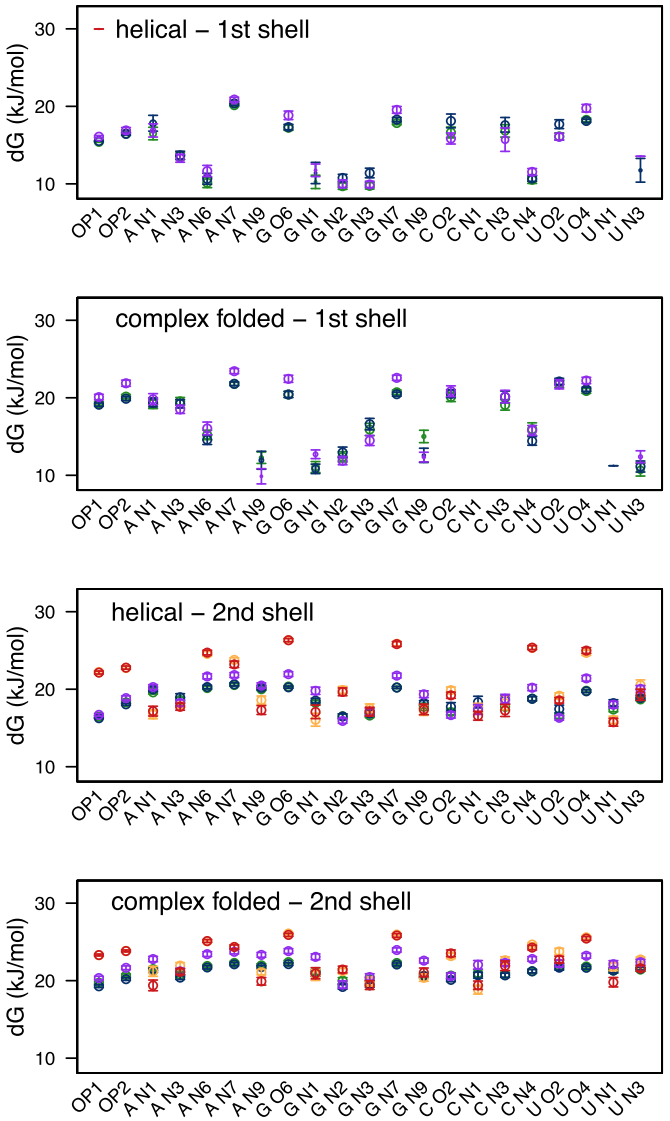

Figure 3. Gibbs energy of activation for contact breaking between RNA atoms and ions. The average energy values for 12 helical (**A**, **C**) and 12 complex folded (**B**, **D**) structures are determined for first shell interactions (direct bonds, A, B) and second shell interactions (C, D). Second shell energies are the sum of first and second shell interactions. The colors represent the corresponding ion in each of the four systems: $Na^+{}_{CI}$ (green), $Na^+{}_{PS}$ (blue), $Mg^{2+}{}_{CI}$ (orange) and in system $Mg^{2+}{}_{PS}$ Mg^{2+} (red) and Na^+ (purple).

X-ray structure of 2QEK (49). We chose 2QEK as example structure, because both monovalent (K^+) and divalent (Mg^{2+}) ions are present in this structure. When only Na^+ ions are present in the simulation both K^+ and Mg^{2+} binding sites are occupied (Figure 4A and B). This can also be observed for other structures (Table 2). In some cases it seems as if the binding site can be occupied by both Na^+ and Mg^{2+} ions. Mg^{2+} binding sites are more difficult to predict with MD simulations since the hydration layer around Mg^{2+} ions is almost never dismantled. When only Mg^{2+} are present in the simulation, ions are closer to RNA atoms compared to when both Mg^{2+} and Na^+ are present. The closest distance between experimentally pre-

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

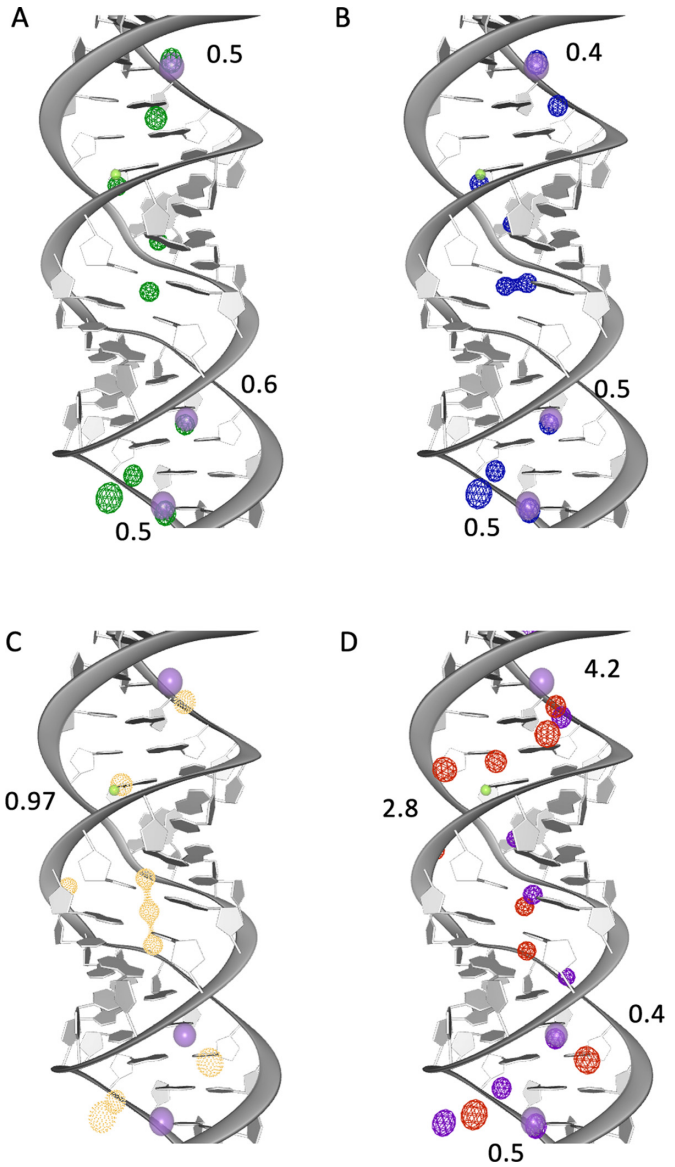

Figure 4. Experimental and predicted ion binding sites. The helical structure (PDB id: 2QEK) has one Mg^{2+} (solid green) and three K^+ (solid purple) binding sites. During the equilibration phase we observe Mg^{2+} and Na^+ in close proximity to the experimentally predicted binding sites. The RMSD between experimental and predicted binding sites are given in Å. The 10 top ranked ion binding sites predicted with MobyWat for one replica for each of the four systems is shown: (**A**) $Na^+{}_{CI}$ (green), (**B**) $Na^+{}_{PS}$ (blue), (**C**) $Mg^{2+}{}_{CI}$ (orange), (**D**) in $Mg^{2+}{}_{PS}$ there are Mg^{2+} (red) and Na^+ (purple).

dicted Mg^{2+} binding sites and those observed in our simulation is 1-2 Å. Overall, Mg^{2+} binding sites are predicted better than K^+ binding sites. This indicates a preference of Mg^{2+} ions to experimentally predicted Mg^{2+} ion binding sites. When both, Na^+ and Mg^{2+} ions are present in the simulations the distances to experimentally predicted binding sites are higher compared to other systems. This is surprising since it does not correlate with lower RMSD or ϵRMSD values for those structures. It indicates that although specific ion positions are not found during MD simulations the overall structure maintains a native-like fold, potentially due to there being a 'sufficient' amount of screening of electrostatic interactions. In general we observe Mg^{2+} ions present along the minor groove of the RNA and in some specific binding sites.

Table 2. Average RMSD values between experimental and predicted binding sites during the production phase. The position of ions are predicted with MobyWat (80,81). The resulting top 50 ion positions are considered for each replica

System	$Na^+{}_{CI}$	$Na^+{}_{PS}$	$Mg^{2+}{}_{CI}$	$Mg^{2+}{}_{PS}$	
Ion	Na^+	Na^+	Mg^{2+}	Na^+	Mg^{2+}
1D4R					
MG-90	2.8 ± 1.7	2.4 ± 1.0	1.0 ± 0.5	4.2 ± 1.2	1.1 ± 0.5
MG-91	6.0 ± 1.0	6.9 ± 2.1	5.0 ± 0.7	5.6 ± 1.5	4.4 ± 0.7
2MTK					
MG-48*	7.5 ± 0.9	7.7 ± 0.8	7.4 ± 3.2	8.1 ± 4.0	5.8 ± 1.9
MG-49	5.2 ± 1.2	4.3 ± 1.2	3.9 ± 0.9	5.6 ± 3.0	2.9 ± 1.6
MG-50	4.2 ± 1.0	4.6 ± 1.7	6.7 ± 3.0	4.7 ± 1.4	5.7 ± 2.3
MG-51	2.6 ± 0.7	1.7 ± 1.0	3.2 ± 0.8	3.9 ± 1.3	3.5 ± 0.4
MG-52	5.9 ± 0.4	4.1 ± 1.3	3.6 ± 0.5	3.9 ± 2.5	3.8 ± 2.4
MG-53	3.3 ± 1.2	3.1 ± 1.5	2.1 ± 0.4	3.2 ± 0.5	2.3 ± 1.1
2QEK					
K-47	3.3 ± 1.2	2.7 ± 0.3	2.1 ± 0.6	2.8 ± 2.4	1.8 ± 0.5
K-48	3.3 ± 0.7	3.4 ± 0.5	3.4 ± 0.1	3.7 ± 2.4	5.2 ± 2.4
MG-49	3.6 ± 0.6	2.4 ± 1.1	2.5 ± 0.2	2.9 ± 0.8	2.5 ± 1.3
K-50	1.6 ± 0.7	2.1 ± 0.6	2.0 ± 0.7	4.4 ± 3.7	2.5 ± 0.4
4FRG					
MG-179	2.2 ± 0.8	2.2 ± 0.3	2.4 ± 0.7	1.6 ± 0.2	4.4 ± 0.8
MG-180	1.5 ± 0.5	1.4 ± 0.2	2.4 ± 0.8	1.7 ± 0.2	5.3 ± 0.4
MG-181	2.5 ± 1.0	4.2 ± 2.8	2.8 ± 0.5	8.5 ± 0.8	1.4 ± 0.5
MG-182*	8.0 ± 0.6	9.1 ± 2.3	7.6 ± 0.5	11.5 ± 3.6	7.0 ± 0.6
MG-183	5.7 ± 1.2	4.7 ± 1.2	3.7 ± 1.5	3.5 ± 1.7	4.7 ± 2.9
MG-184	1.5 ± 0.6	1.6 ± 0.6	1.1 ± 0.3	3.5 ± 1.4	2.0 ± 1.5
MG-185	3.2 ± 1.3	3.4 ± 0.7	3.7 ± 1.3	2.1 ± 0.3	5.9 ± 1.4
4JF2					
MG-94	1.4 ± 0.7	1.5 ± 0.4	2.2 ± 1.1	2.5 ± 1.7	2.7 ± 1.0
MG-95*	7.0 ± 0.7	6.5 ± 1.8	3.2 ± 0.6	4.7 ± 1.8	4.6 ± 0.7
MG-96	1.8 ± 0.1	1.6 ± 0.8	2.5 ± 0.8	2.6 ± 0.3	2.9 ± 0.9
MG-97*	26.6 ± 7.7	31.1 ± 6.3	18.3 ± 2.7	24.9 ± 6.1	20.6 ± 0.8
4KQY					
MG-121	1.8 ± 0.5	1.6 ± 0.7	1.8 ± 0.4	0.8 ± 0.6	3.4 ± 0.9
MG-122*	3.4 ± 1.9	5.0 ± 0.4	1.8 ± 0.5	8.4 ± 2.3	4.2 ± 2.0
4P5J					
MG-85	2.3 ± 0.4	2.9 ± 0.6	1.1 ± 0.7	3.1 ± 1.4	1.8 ± 0.2
MG-86	2.2 ± 0.4	2.5 ± 0.2	2.5 ± 0.5	5.3 ± 3.2	3.5 ± 2.1

The predicted binding sites are in good agreement with experimentally identified ion locations in close proximity to the RNA (Table 2). There are some cases for which the experimental binding site was not detected, however, in particular for ions directly bound to RNA. This is expected, due to the high barrier for desolvation of Mg^{2+} ions (66). Most of these sites are located at the surface of the RNA and only one RNA atom can be identified as potential contact site in experimentally predicted structures. For this reason RMSD values to the experimental binding sites are marked with an asterisk in Table 2 for these ions.

DISCUSSION

The structural analyses indicate (at least in X-ray structures and most NMR structures) that Mg^{2+} ions have a stronger stabilizing effect for helical structures than for complex folded structures (Figure 1). Although ϵRMSD values of complex folded structures, when simulated with and without Mg^{2+} ions are comparable, they are in general high (above 1), indicating that these structures do not maintain their native fold (76). This might be due to the fact that the quality of complex folded NMR structures is not as good as that of X-ray structures (Supplementary Table S1), for instance because structures that are inherently more flexi-

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

ble and difficult to solve by X-ray crystallography instead are solved by NMR techniques. Especially, the RNA backbone seems not to be as well defined for NMR structures based on the validation results of X-ray and NMR structures (Supplementary Table S1).

It has been reported (38) that helical RNA structures undergo irreversible structural changes in longer MD simulations (over 50 ns) when using parm99 and parmbsc0 (82). They change into a ladder–like structure, similar to what we observe in the majority of helical RNA structures with high RMSD values. The reason for this is that the glycoside torsion angle χ is shifted from the *anti* to the high-*anti* region. A specific force field parameter set, called parmOL (83,84), has been developed to eliminate this artifact. Since we did not use these parameters for most of the simulation, the correct backbone angle of some helical RNA structures was not maintained in this work. We did, however, use the combination of parmOL (83,84) and parmbcs0 (82) parameters specifically developed for RNA, for a subset of our structures. These structures undergo less structural changes and have lower RMSD and ϵRMSD values compared to structures that were simulated with the same ion conditions (Supplementary Figure S2). However, our observation that Mg^{2+} results in more stable simulations still holds. The comparison between Na^+ and K^+ as a counterion (Supplementary Figures S2 versus S3 and Supplementary Figures S8–S11) suggest potassium stabilizes the structures somewhat more than does sodium.

Both Na^+ and Mg^{2+} ions bind sequence specific and also to specific binding sites (Figure 4). In both helical and complex folded structures certain RNA atoms are preferred. In complex folded structures atoms are available for binding that are not sterically accessible for ions in helical RNA structures. For example, one of the oxygen atom in uracil (U-O2) is hidden in the minor groove of a helical RNA with classical Watson-Crick base pair interactions. In complex folded structures we find this atom to be more accessible to ions (Figure 2), consistent with findings reported by Kirmizialtin *et al.* (18). We think it is appropriate to distinguish between adenine and guanine N7 atoms unlike what was done in previous studies (15,18). Doing so reveals that more ions are close to the guanine N7 atom than can be explained based just on accessibility and indeed the distributions are quantitatively different for both atoms (Figure 2).

At low ion concentrations a larger fraction of the Na^+ and Mg^{2+} ions are in direct contacts with the RNA in our simulations than at higher concentrations (2). When, however, both Na^+ and Mg^{2+} ions are present, more Mg^{2+} ions are closer to the RNA (distance less than 10 Å) than Na^+. This is in agreement with the 'ion atmosphere' as described by Lipfert *et al.* (37). It seems therefore that the overall salt concentration should be factored in when considering the properties of the 'ion atmosphere'.

Zheng *et al.* (13) investigated Mg^{2+} ion binding sites experimentally, in particular the difference between first and second shell binding frequencies. Since we only observe two direct contacts for Mg^{2+} ions in our simulations we cannot compare our simulations with the first shell contact frequencies derived in that work (13). The main reason for this is that it is very difficult to replace the hydration shell around Mg^{2+} by direct contacts during explicit MD simulations (66). Although refined Mg^{2+} ion parameters (64) were used, the activation energy remains slightly higher and the ion–water exchange rate faster than experimental values (66,85).

When we compare our Gibbs activation energies for second shell dissociation/binding of Mg^{2+} ions (Figure 3) to the experimental frequencies reported in (13) we see a preference for the same RNA atoms. The calculations fit the results by Zheng *et al.* (13) remarkably well. The RNA atoms with the highest experimental frequencies are (starting from the highest): G-O6, G-N7, O2P, U-O4, A-N7, O1P and A-N6 (13). For helical structures the RNA atoms with highest $\Delta G‡$ are for system $Mg^{2+}{}_{PS}$ (starting from the highest): G-O6, G-N7, C-N4, U-O4, A-N6, A-N7, O2P, and O1P. For complex folded structures for system $Mg^{2+}{}_{PS}$ (starting from the highest): G-O6, G-N7, U-O4, A-N6, A-N7, C-N4, O2P, U-O2 and O1P. For helical structures the RNA atoms with highest free energies of activation $\Delta G‡$ are for system $Mg^{2+}{}_{CI}$ (starting from the highest): G-O6, G-N7, C-N4, U-O4, A-N6, A-N7, O2P and O1P. For complex folded structures for system $Mg^{2+}{}_{CI}$ (starting from the highest): G-O6, G-N7, U-O4, A-N6, A-N7, C-N4, O2P, C-O2 and O1P. Although the activation energy is higher for O2P than O1P it seems to be underestimated in all simulations compared to the energies calculated for other RNA atoms. The main difference between helical and complex folded structures is that in helical ones the activation energy is higher for C-N4. When we compare the activation energy of the Mg^{2+} ion directly in contact with O1P (30.0 kJ/mol) to experimentally predicted activation energies $\Delta G‡$ between Mg^{2+} ions and DNA (53.1–55.7 kJ/mol) (85) it is quantitatively underestimated.

After the equilibration phase we could reproduce all experimentally predicted ion binding sites with good accuracy (Table 2, Figure 4). Especially when Na^+ or Mg^{2+} ions are present (system $Na^+{}_{PS}$ and $Mg^{2+}{}_{PS}$) without any additional salt concentration the binding sites are reproduced well using the MobyWat (80,81) analysis (Supplementary Table S3). The reason for this is likely that in simulations at low ionic strength the ions are found in close proximity to the RNA. We find, however, that the occupancy of the experimental ion binding sites calculated from the simulation is not reproduced with the same accuracy as the positions. A similar study focused on ion-binding to helical DNA was able to reproduce experimental ion-counts quantitatively (86), possibly because of improved cation force field parameters (67).

An interesting study by Lemkul *et al.* (87) applied grand-canonical Monte Carlo-MD (GCMC-MD) in order to predict ion-binding for four different RNA molecules. Although this approach most likely is more suited to find first shell binding locations than the MD approach we used, the use of pure MD allows to deduce time-dependent properties such as $\Delta G‡$ for contact breaking (Figure3). A combination of the two techniques, prediction using GCMC-MD followed by regular MD would therefore yield a more complete picture of binding thermodynamics and kinetics. Nevertheless it seems the quality of binding site predictions is similar in both methods. Ion binding site prediction is inherently difficult for these systems with long exchange times.

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

It is likely as well that ion binding sites are missed by any structural analysis since ion-binding and conformational flexibility are interdependent. In fact, it is remarkable that Mg^{2+} ions are predicted so close to experimental binding sites in normal simulation, while they maintain their hydration shell.

In comparison to previous studies our dataset contains a large number (24) of structures yielding rigorous results. Binding site positions and kinetics can be studied, and the relative influence of different ions studied. Based on our results (e.g. Figure 1) there is no justification for using Na^+ ions rather than Mg^{2+} ions in RNA simulations, unless, as in this work, the purpose of the study is to investigate the difference in RNA properties due to the 'ion atomosphere' (37). Further improvement of force fields for RNA, water and ions remain needed to describe the complex energy landscape formed by these flexible biomolecules.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

FUNDING

Swedish research council [2013-5947 to D.S.]; eSSENCE e-Science collaboration [to N.M.F]; Coordination for the Improvement of Higher Education Personnel (CAPES) [to M.D.P.]; Swedish National Infrastructure for Computing (SNIC) at PDC Centre for High Performance Computing (PDC-HPC) [SNIC2016/34-44]. Funding for open access charge: Vetenskapsrådet, [2013-5947]; faculty funding.
Conflict of interest statement. None declared.

REFERENCES

1. Draper,D.E. (2008) RNA folding: thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.*, **95**, 5489–5495.
2. Aravin,A.A., Hannon,G.J. and Brennecke,J. (2007) The Piwi–piRNA pathway provides an adaptive defense in the transposon arms race. *Science*, **318**, 761–764.
3. Rollins,M.F., Schuman,J.T., Paulus,K., Bukhari,H.S.T. and Wiedenheft,B. (2015) Mechanism of foreign DNA recognition by a CRISPR RNA–guided surveillance complex from Pseudomonas aeruginosa. *Nucleic Acids Res.*, **43**, 2216–2222.
4. Misra,V.K. and Draper,D.E. (2001) A thermodynamic framework for Mg^{2+} binding to RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 12456–12461.
5. Jenner,L., Demeshkina,N., Yusupova,G. and Yusupov,M. (2010) Structural rearrangements of the ribosome at the tRNA proofreading step. *Nat. Struct. Mol. Biol.*, **17**, 1072–1078.
6. Brännvall,M. and Kirsebom,L.A. (2001) Metal ion cooperativity in ribozyme cleavage of RNA. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 12943–12947.
7. Bowman,J.C., Lenz,T.K., Hud,N.V. and Williams,L.D. (2012) Cations in charge: magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.*, **22**, 262–272.
8. Scott,W.G., Finch,J.T. and Klug,A. (1995) The crystal structure of an all–RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell*, **81**, 991–1002.
9. Sigurdsson,S.T. and Eckstein,F. (1995) Structure–function relationships of hammerhead ribozymes: from understanding to applications. *Trends Biotechnol.*, **13**, 286–289.
10. Schnabl,J. and Sigel,R.K.O. (2010) Controlling ribozyme activity by metal ions. *Curr. Opin. Chem. Biol.*, **14**, 269–275.
11. Wilson,T.J. and Lilley,D.M. (2015) RNA catalysis–is that it? *RNA*, **21**, 534–537.
12. Lilley,D.M.J. (2017) How RNA acts as a nuclease: some mechanistic comparisons in the nucleolytic ribozymes. *Biochem. Soc. Trans.*, **45**, 683–691.
13. Zheng,H., Shabalin,I.G., Handing,K.B., Bujnicki,J.M. and Minor,W. (2015) Magnesium–binding architectures in RNA crystal structures: validation, binding preferences, classification and motif detection. *Nucleic Acids Res.*, **43**, 3789–3801.
14. Draper,D.E. (2004) A guide to ions and RNA structure. *RNA*, **10**, 335–343.
15. Hayes,R.L., Noel,J.K., Mohanty,U., Whitford,P.C., Hennelly,S.P., Onuchic,J.N. and Sanbonmatsu,K.Y. (2012) Magnesium fluctuations modulate RNA dynamics in the SAM–I riboswitch. *J. Am. Chem. Soc.*, **134**, 12043–12053.
16. Auffinger,P. and Westhof,E. (2000) Water and ion binding around RNA and DNA (C,G) oligomers. *J. Mol. Biol.*, **300**, 1113–1131.
17. Auffinger,P. and Westhof,E. (2001) Water and ion binding around r(UpA)12 and d(TpA)12 oligomers–comparison with RNA and DNA (CpG)12 duplexes. *J. Mol. Biol.*, **305**, 1057–1072.
18. Kirmizialtin,S. and Elber,R. (2010) Computational exploration of mobile ion distributions around RNA duplex. *J. Phys. Chem. B*, **114**, 8207–8220.
19. Manning,G.S. (1978) The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.*, **11**, 179–246.
20. Urbanke,C., Römer,R. and Maass,G. (1975) Tertiary structure of tRNAPhe (yeast): kinetics and electrostatic repulsion. *Eur. J. Biochem.*, **55**, 439–444.
21. Leroy,J.L., Guéron,M., Thomas,G. and Favre,A. (1997) Role of divalent ions in folding of tRNA. *Eur. J. Biochem.*, **74**, 567–574.
22. Römer,R. and Hach,R. (1975) tRNA conformation and magnesium binding. A study of a yeast phenylalanine–specific tRNA by a fluorescent indicator and differential melting curves. *Eur. J. Biochem.*, **55**, 271–284.
23. Ha,B.–Y. and Thirumalai,D. (2003) Bending rigidity of stiff polyelectrolyte chains: a single chain and a bundle of multichains. *Macromolecules*, **36**, 9658–9666.
24. Stein,A. and Crothers,D.M. (1976) Equilibrium binding of magnesium(II) by Escherichia coli tRNAfMet. *Biochemistry*, **15**, 157–160.
25. Klein,D.J., Moore,P.B. and Steitz,T.A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA*, **10**, 1366–1379.
26. Lippert,B. (2000) Multiplicity of metal ion binding patterns to nucleobases. *Coord. Chem. Rev.*, **200**, 487–516.
27. Tinoco,I. and Kieft,J.S. (1997) The ion core in RNA folding. *Nat. Struct. Biol.*, **4**, 509–512.
28. Ennifar,E., Yusupov,M., Walter,P., Marquet,R., Ehresmann,B., Ehresmann,C. and Dumas,P. (1999) The crystal structure of the dimerization initiation site of genomic HIV–1 RNA reveals an extended duplex with two adenine bulges. *Structure*, **7**, 1439–1449.
29. Correll,C.C., Freeborn,B., Moore,P.B. and Steitz,T.A. (1997) Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell*, **91**, 705–712.
30. Petrov,A.S., Bowman,J.C., Harvey,S.C. and Williams,L.D. (2011) Bidentate RNA–magnesium clamps: on the origin of the special role of magnesium in RNA folding. *RNA*, **17**, 291–297.
31. Cooper,D.R., Porebski,P.J., Chruszcz,M. and Minor,W. (2011) X–ray crystallography: assessment and validation of protein–small molecule complexes for drug discovery. *Expert Opin. Drug Discov.*, **6**, 771–782.
32. Pozharski,E., Weichenberger,C.X. and Rupp,B. (2013) Techniques, tools and best practices for ligand electron–density analysis and results from their application to deposited crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **69**, 150–167.
33. Philips,A., Milanowska,K., Lach,G., Boniecki,M., Rother,K. and Bujnicki,J.M. (2012) MetalionRNA: computational predictor of metal–binding sites in RNA structures. *Bioinformatics*, **28**, 198–205.
34. Zheng,H., Chordia,M.D., Cooper,D.R., Chruszcz,M., Müller,P., Sheldrick,G.M. and Minor,W. (2014) Validation of metal–binding sites in macromolecular structures with the CheckMyMetal web server. *Nat. Protoc.*, **9**, 156–170.
35. Nayal,M. and Di Cera,E. (1996) Valence screening of water in protein crystals reveals potential Na^+ binding sites. *J. Mol. Biol.*, **256**, 228–234.

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

36. Gonzalez,R.L. Jr and Tinoco,I. Jr (2001) Identification and characterization of metal ion binding sites in RNA. *Methods Enzymol.*, **338**, 421–443.
37. Lipfert,J., Doniach,S., Das,R. and Herschlag,D. (2014) Understanding Nucleci Acid–Ion Interactions. *Annu. Rev. Biochem.*, **83**, 813–841.
38. Sponer,J., Otyepka,M., Banáš,P., Réblová,K. and Walter,N.G. (2012) Molecular Dynamics Simulations of RNA Molecules. In: Schlick,T (ed). *Innovations in Biomolecular Modeling and Simulations: Complete Set*. Royal Society of Chemistry, pp. 129–155.
39. Lange,O.F., van der Spoel,D. and de Groot,B.L. (2010) Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophys. J.*, **99**, 647–655.
40. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
41. Wild,K., Weichenrieder,O., Leonard,G.A. and Cusack,S. (1999) The 2 Å structure of helix 6 of the human signal recognition particle RNA. *Structure*, **7**, 1345–1352.
42. Klosterman,P.S., Shah,S.A. and Steitz,T.A. (1999) Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution. *Biochemistry*, **38**, 14784–14792.
43. Shalev,M., Kondo,J., Kopelyanskiy,D., Jaffe,C.L., Adir,N. and Baasov,T. (2013) Identification of the molecular attributes required for aminoglycoside activity against Leishmania. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 13333–13338.
44. Tanaka,Y., Fujii,S., Hiroaki,H., Sakata,T., Tanaka,T., Uesugi,S., Tomita,K. and Kyogoku,Y. (1999) A–form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC). *Nucleic Acids Res.*, **27**, 949–955.
45. Dallas,A. and Moore,P.B. (1997) The loop E–loop D region of Escherichia coli 5S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins. *Structure*, **5**, 1639–1653.
46. Henriksen,N.M., Davis,D.R. and Cheatham,T.E. (2012) Molecular dynamics re–refinement of two different small RNA loop structures using the original NMR data suggest a common structure. *J. Biomol. NMR*, **53**, 321–339.
47. Nikonowicz,E.P., Wang,J., Moran,S. and Donarski,J. (2012) Solution structure of Helix–35 stem–loop from E. coli 23S rRNA. *Biol. Magn. Res. Data Bank*, doi:10.13018/BMR18549.
48. Stefl,R., Oberstrass,F.C., Hood,J.L., Jourdan,M., Zimmermann,M., Skrisovska,L., Maris,C., Peng,L., Hofr,C., Emeson,R.B. and Allain,F.H.–T. (2010) The solution structure of the ADAR2 dsRBM–RNA complex reveals a sequence–specific readout of the minor groove. *Cell*, **143**, 225–237.
49. Freisz,S., Lang,K., Micura,R., Dumas,P. and Ennifar,E. (2008) Binding of aminoglycoside antibiotics to the duplex form of the HIV–1 genomic RNA dimerization initiation site. *Angew. Chem., Int. Ed. Engl.*, **47**, 4110–4113.
50. Pan,B., Mitra,S.N. and Sundaralingam,M. (1999) Crystal structure of an RNA 16–mer duplex R(GCAGAGUUAAAUCUGC)2 with nonadjacent G(syn).A^+(anti) mispairs. *Biochemistry*, **38**, 2826–2831.
51. Daldrop,P. and Lilley,D.M.J. (2013) The plasticity of a structural motif in RNA: structural polymorphism of a kink turn as a function of its environment. *RNA*, **19**, 357–364.
52. Stoddard,C.D., Widmann,J., Trausch,J.J., Marcano–Velazquez,J.G., Knight,R. and Batey,R.T. (2013) Nucleotides adjacent to the ligand–binding pocket are linked to activity tuning in the purine riboswitch. *J. Mol. Biol.*, **425**, 1596–1611.
53. Johnson,J.E., Reyes,F.E., Polaski,J.T. and Batey,R.T. (2012) B12 cofactors directly stabilize an mRNA regulatory switch. *Nature*, **492**, 133–137.
54. Liberman,J.A., Salim,M., Krucinska,J. and Wedekind,J.E. (2013) Structure of a class II preQ1 riboswitch reveals ligand recognition by a new fold. *Nat. Chem. Biol.*, **9**, 353–355.
55. Lu,C., Ding,F., Chowdhury,A., Pradhan,V., Tomsic,J., Holmes,W.M., Henkin,T.M. and Ke,A. (2010) SAM recognition and conformational switching mechanism in the Bacillus subtilis yitJ S box/SAM–I riboswitch. *J. Mol. Biol.*, **404**, 803–818.
56. Colussi,T.M., Costantino,D.A., Hammond,J.A., Ruehle,G.M., Nix,J.C. and Kieft,J.S. (2014) The structural basis of transfer RNA mimicry and conformational plasticity by a viral RNA. *Nature*, **511**, 366–369.
57. Baba,S., Takahashi,K.–i., Noguchi,S., Takaku,H., Koyanagi,Y., Yamamoto,N. and Kawai,G. (2005) Solution RNA structures of the HIV–1 dimerization initiation site in the kissing–loop and extended–duplex dimers. *J. Biochem.*, **138**, 583–592.
58. Tolbert,B.S., Miyazaki,Y., Barton,S., Kinde,B., Starck,P., Singh,R., Bax,A., Case,D.A. and Summers,M.F. (2010) Major groove width variations in RNA structures determined by NMR and impact of 13C residual chemical shift anisotropy and 1H–13C residual dipolar coupling on refinement. *J. Biomol. NMR*, **47**, 205–219.
59. Theimer,C.A., Blois,C.A. and Feigon,J. (2005) Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell*, **17**, 671–682.
60. Davis,J.H., Tonelli,M., Scott,L.G., Jaeger,L., Williamson,J.R. and Butcher,S.E. (2005) RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop–receptor complex. *J. Mol. Biol.*, **351**, 371–382.
61. Burke,J.E., Sashital,D.G., Zuo,X., Wang,Y.–X. and Butcher,S.E. (2012) Structure of the yeast U2/U6 snRNA complex. *RNA*, **18**, 673–683.
62. Kim,N.–K., Zhang,Q. and Feigon,J. (2014) Structure and sequence elements of the CR4/5 domain of medaka telomerase RNA important for telomerase function. *Nucleic Acids Res.*, **42**, 3395–3408.
63. Bonneau,E. and Legault,P. (2014) Nuclear magnetic resonance structure of the III–IV–V three–way junction from the Varkud satellite ribozyme and identification of magnesium–binding sites using paramagnetic relaxation enhancement. *Biochemistry*, **53**, 6264–6275.
64. Cash,D.D., Cohen–Zontag,O., Kim,N.–K., Shefer,K., Brown,Y., Ulyanov,N.B., Tzfati,Y. and Feigon,J. (2013) Pyrimidine motif triple helix in the Kluyveromyces lactis telomerase RNA pseudoknot is essential for function in vivo. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 10970–10975.
65. Auffinger,P., Cheatham,T.E. and Vaiana,A.C. (2007) Spontaneous formation of KCl aggregates in biomolecular simulations: a force field issue? *J. Chem. Theory Comput.*, **3**, 1851–1859.
66. Allnér,O., Nilsson,L. and Villa,A. (2012) Magnesium ion–water coordination and exchange in biomolecular simulations. *J. Chem. Theory Comput.*, **8**, 1493–1502.
67. Gebala,M., Giambasu,G.M., Lipfert,J., Bisaria,N., Bonilla,S., Li,G., York,D.M. and Herschlag,D. (2015) Cation-anion interactions within the nucleic acid ion atmosphere revealed by ion counting. *J. Am. Chem. Soc.*, **137**, 14705–14715.
68. Yoo,J. and Aksimentiev,A. (2012) Improved parametrization of Li^+, Na^+, K^+, and Mg^{2+} ions for all-atom molecular dynamics simulations of nucleic acid systems. *J. Phys. Chem. Lett.*, **3**, 45–50.
69. Wang,J., Cieplak,P. and Kollman,P.A. (2000) 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.
70. Pronk,S., Páll,S., Schulz,R., Larsson,P., Bjelkmar,P., Apostolov,R., Shirts,M.R., Smith,J.C., Kasson,P.M., van der Spoel,D. *et al.* (2013) GROMACS 4.5: a high–throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, **29**, 845–854.
71. Hess,B., Bekker,H., Berendsen,H.J.C. and Fraaije,J.G.E.M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.*, **18**, 1463–1472.
72. Hess,B. (2008) P–LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.*, **4**, 116–122.
73. Darden,T., Pearlman,D. and Pedersen,L.G. (1998) Ionic charging free energies: Spherical versus periodic boundary conditions. *J. Chem. Phys.*, **109**, 10921–10935.
74. Bussi,G., Donadio,D. and Parrinello,M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 014101.
75. Parrinello,M. and Rahman,A. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.
76. Bottaro,S., Di Palma,F. and Bussi,G. (2014) The role of nucleobase interactions in RNA structure and dynamics. *Nucleic Acids Res.*, **42**, 13306–13314.
77. Luzar,A. and Chandler,D. (1996) Effect of environment on hydrogen bond dynamics in liquid water. *Phys. Rev. Lett.*, **76**, 928–931.

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018

78. van der Spoel,D., van Maaren,P.J., Larsson,P. and Tîmneanu,N. (2006) Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B.*, **110**, 4393–4398.
79. Larsson,D.S.D. and van der Spoel,D. (2012) Screening for the location of RNA using the chloride Ion distribution in simulations of virus capsids. *J. Chem. Theory Comput.*, **8**, 2474–2483.
80. Jeszenői,N., Horváth,I., Bálint,M., van der Spoel,D. and Hetényi,C. (2015) Mobility-based prediction of hydration structures of protein surfaces. *Bioinformatics*, **31**, 1959–1965.
81. Jeszenői,N., Bálint,M., Horváth,I., van der Spoel,D. and Hetényi,C. (2016) Exploration of interfacial hydration networks of target-ligand complexes. *J. Chem. Inf. Model*, **56**, 148–158.
82. Perez,A., Marchan,I., Svozil,D., Sponer,J., Cheatham,T.E. III, Laughton,C.A. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of [alpha]/[gamma] conformers. *Biophys. J.*, **92**, 3817–3829.
83. Banas,P., Hollas,D., Zgarbova,M., Jurecka,P., Orozco,M., Cheatham,T.E. III, Sponer,J. and Otyepka,M. (2010) Performance of molecular mechanics force fields for RNA simulations. Stability of UUCG and GNRA hairpins. *J. Chem. Theory Comput.*, **6**, 3836–3849.
84. Zgarbova,M., Otyepka,M., Sponer,J., Mladek,A., Banas,P., Cheatham,T.E. III and Jurecka,P. (2011) 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.
85. Cowan,J.A., Huang,H.-W. and Hsu,L.-Y. (1993) Sequence selective coordination of Mg^{2+} (aq) to DNA. *J. Inorg. Biochem.*, **52**, 121–129.
86. Yoo,J. and Aksimentiev,A. (2012) Competitive binding of cations to duplex DNA revealed through molecular dynamics simulations. *J. Phys. Chem. B*, **116**, 12946–12954.
87. Lemkul,J.A., Lakkaraju,S.K. and MacKerell,A.D. (2016) Characterization of Mg^{2+} distributions around RNA in solution. *ACS Omega*, **1**, 680–688.

Downloaded from https://academic.oup.com/nar/article-abstract/46/10/4872/4990003 by Uppsala Universitetsbibliotek user on 20 September 2018