bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Title

Mg^{2+} Impacts the Twister Ribozyme through Push-Pull Stabilization of Non-Sequential Phosphate Pairs

Authors

A. A. Kognole and A. D. MacKerell, Jr.

Abstract

RNA molecules perform a variety of biological functions for which the correct three-dimensional structure is essential, including as ribozymes where they catalyze chemical reactions. Metal ions, especially Mg^{2+}, neutralize these negatively charged nucleic acids and specifically stabilize RNA tertiary structures as well as impact the folding landscape of RNAs as they assume their tertiary structures. Specific binding sites of Mg^{2+} in folded conformations of RNA have been studied extensively, however, the full range of interactions of the ion with compact intermediates and unfolded states of RNA is challenging to investigate and the atomic details of the mechanism by which the ion facilitates tertiary structure formation is not fully known. Here, umbrella sampling combined with oscillating chemical potential Grand Canonical Monte Carlo/Molecular Dynamics (GCMC/MD) simulations are used to capture the energetics and atomic-level details of Mg^{2+}-RNA interactions that occur along an unfolding pathway of the Twister ribozyme. The free energy profiles reveal stabilization of partially unfolded states by Mg^{2+}, as observed in unfolding experiments, with this stabilization being due to increased sampling of simultaneous interactions of Mg^{2+} with two or more non-sequential phosphate groups. Notably, the present results indicate a push-pull mechanism where the Mg^{2+}-RNA interactions actually lead to destabilization of specific non-sequential phosphate-phosphate interactions while other interactions are stabilized, a balance that facilitates the folding and stabilization of Twister including formation hydrogen bonds associated with the tertiary structure. The present study establishes a better understanding of how Mg^{2+} ion-interactions contribute to RNA structural properties and stability.

Keywords

RNA folding, Metal ions, Magnesium, Ion atmosphere, Grand Canonical, Monte Carlo, Umbrella sampling, CHARMM

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Statement of Significance

RNAs are biologically and therapeutically of great emerging interest such that it is critical to understand how RNA molecules fold into complex structures. While experiments yield information on the stabilization of RNA by ions they are limited in the atomic-level insights they can provide. A combination of enhanced sampling methods is applied to explore the compact intermediate states of RNA and their interactions with Mg^{2+} ions. Results reveal a picture of how Mg^{2+} overall stabilizes short phosphate-phosphate interactions thereby facilitating the stabilization of RNA, though doing so by both the stabilization and destabilization of specific interactions. The applied method will be applicable to exploring the impact of divalent ions on the conformational heterogeneity of a range of macromolecules.

Introduction

RNA molecules have highly diverse structures ranging from simple helices to highly heterogeneous folded conformations that are essential for their wide range of cellular functions (1-3). Specifically, ribozymes, a distinct class of enzymes, exhibit complex tertiary structures and catalyze self-cleavage or the cleavage of phosphodiester bonds of substrate RNA, with metal ions typically playing a central role in the catalytic activity (4, 5). To assume their tertiary structures RNAs must overcome large unfavorable electrostatic interactions associated with their polyanionic phosphodiester backbone (6). To facilitate this, positively charged ions screen the highly negative potential allowing the RNA secondary structures to collapse into compact tertiary conformations (7-10). Typically divalent ions, most often Mg^{2+}, facilitate the folding of RNA into tertiary interactions (11-13). However, the inability to visualize the ions during folding represents a key barrier to understanding the role of divalent ions in folding of RNA (14).

Studies have used classical MD or other theoretical approaches to investigate Mg^{2+}-RNA binding, but they were limited to native conformations due to their inability to overcome the issues associated with the Mg^{2+} exchange rates (15-21). The exchange rate of water complexed to Mg^{2+} is on the μs time scale (6.7×10^5 s^{-1}) and the exchange rate of Mg^{2+} with phosphate is on the ms time scale (0.5 to 2.5×10^3 s^{-1}) (22), which is beyond the time scale of typical atomistic MD simulations (19, 23) such that only limited insights into Mg^{2+}-RNA interactions are accessible (24). Alternatively, simulations using coarse-grained models of nucleic acids provided insights into how Mg^{2+} can serve to nucleate the folding of key tertiary interactions with the Mg^{2+}-RNA interactions being dominated by specific interactions even in unfolded states (18, 25). Reduced models were able to reproduce thermodynamics of Mg^{2+}-RNA interactions but were limited to

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

native states (26). Atomistic simulations of unfolding of a pseudoknot at high temperatures showed diverse intermediate states, although divalent ions were absent in those studies (27). Additionally, scientists have worked on improving the force-field parameters for positively charged metal ions to better simulate their interactions with biomolecules (28-30). Overall, a detailed picture of the interactions of Mg^{2+} in an explicit solvent environment with unfolded states of RNA has not yet been attained.

To investigate the impact of Mg^{2+} on the folding and stabilization of RNA at an atomic level of details we apply umbrella sampling MD simulations in conjunction with oscillating chemical potential (μ_{ex}) Grand Canonical Monte Carlo (GCMC) sampling (31). The application of GCMC allows for redistribution of Mg^{2+} ions thereby addressing the issue of exchange rates while the MD allows the RNA and waters to respond to the changes in ion positions along a folding pathway. The use of GCMC in combination with MD on RNA was first undertaken by Lemkul et. al. on four structurally distinct RNAs in their native conformations with restraints on the backbone and was shown to successfully identify experimental Mg^{2+} binding sites as well as predict new ion binding regions (32). In the present study we extend that approach by combining it with umbrella sampling to sample the distribution of Mg^{2+} around intermediate conformations along a folding pathway. Application of the method reveals an atomic picture of how Mg^{2+} lowers the free energy of partially folded states of RNA by increasing the sampling of specific non-sequential phosphate-phosphate interactions through simultaneous indirect interaction between two or more non-sequential phosphates. At the same time, other non-sequential phosphate-phosphate interactions are shown to be destabilized as required to allow for overall stabilization of the phosphate-phosphate interactions, representing a push-pull mechanism by which Mg^{2+} stabilizes RNA.

The Twister ribozyme was selected for the present study based on the availability of a range of structural and biochemical data (33-38). While Twister sequences are extremely widespread, the crystal structure (PDB 4OJI) of the ribozyme used in this study is based on Osa-1-4 sequence from *Oryza Sativa* (34). Figure 1 illustrates the secondary and tertiary interactions in Twister. The studied ribozyme has the conserved residues in loops L1, L2 and L4 corresponding to the self-cleavage site and major tertiary interactions (T1 and T2) associated with the double-pseudoknot structure of the ribozyme though it lacks the P3 and P5 stem-loops. We additionally refer to the contacts T0 and T3 as tertiary contacts, where T0 consists of a *trans* Watson-Crick-Hoogsteen base pair U24 – A29 situated in loop L4 and T3 consists of base pair C15 – G19 situated at the truncated P3 step-loop (34), since we find them important for the ribozyme to assume its tertiary structure (Figure 1b). A biophysical study on the same ribozyme, including single-molecule Fluorescence Resonance Energy Transfer (smFRET) experiments, reported folding kinetics and self-cleavage activity in the presence of Mg^{2+} at various concentrations (36). Results from that study showed Mg^{2+}

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

to facilitate the folding of the RNA although there was ambiguity around whether Mg^{2+} ions are involved in nucleolytic activity.

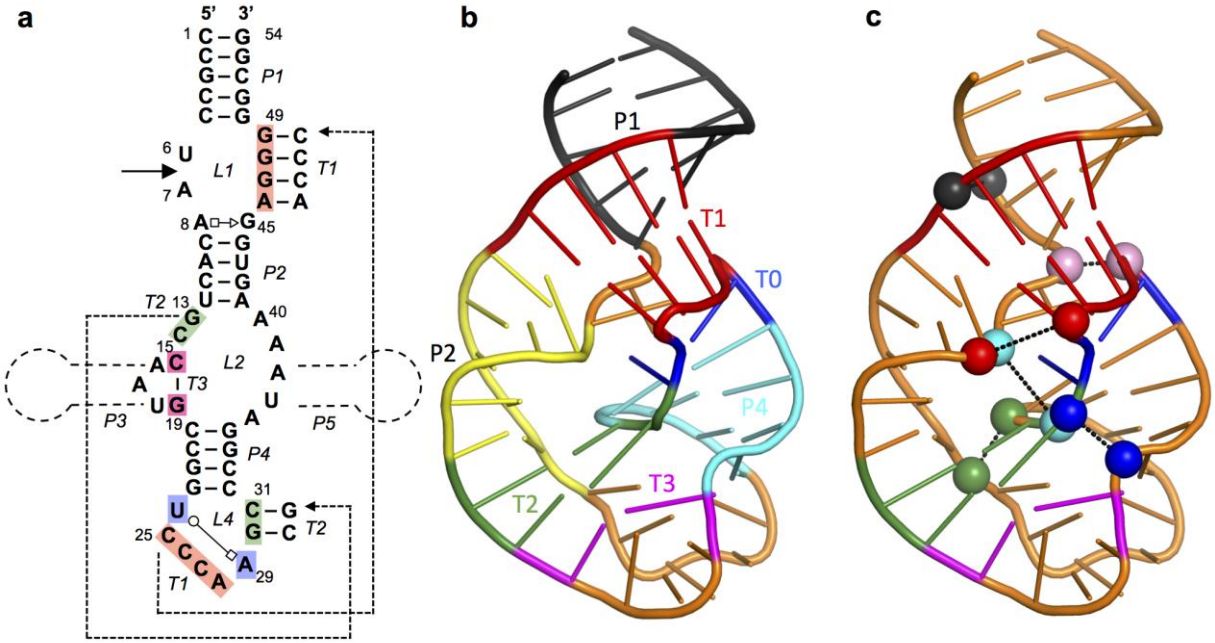

Figure 1. Structure of Twister ribozyme and spatially adjacent phosphate pairs. a) Twister Ribozyme secondary and tertiary contacts. The scissile bond is indicated with a solid arrow and the omitted P3 and P5 loops are indicated. b) Contacts are illustrated as cartoons colored as follows P1 – black, P2 – yellow, P4 – cyan, T0 – blue, T1 – red, T2 – green, T3 – magenta, rest of the RNA – orange. c) Twister ribozyme in its native state with spatially adjacent phosphate pairs that are present in the folded, tertiary structure (same-colored spheres connected with dashed line).

Methods

Potential of Mean Force Calculations

Potential of mean force (PMF) calculations along a folding pathway were performed using umbrella sampling in combination with the oscillating μ_{ex} GCMC-MD method. System preparation involved setting up four Twister simulation systems at 0, 10, 20 and 100 mM $MgCl_2$ followed by subjecting each system to a classical MD simulation of 200 ns. The final coordinates from each of those simulations were used to generate conformations along the unfolding pathway. This was performed by rapidly unfolding the RNA based on the reaction coordinate going from 13 to 40 Å in 0.5 Å

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

increments. Each of the subsequent 55 windows was then subjected to a 10 ns MD simulation in the presence of the umbrella potential at the respective RC distances. The final snapshots from these simulations were used to initiate the oscillating μ_{ex} GCMC-MD PMF calculation. This PMF calculation was performed in 1 Å increments from 13 to 40 Å yielding a total of 28 windows. Specific details for the different aspect of these calculations follow.

MD simulations of the Twister ribozyme were performed in a 90 Å cubic waterbox to accommodate the unfolded conformations at 4 different concentrations of Mg^{2+} (0, 10, 20 and 100 mM of $MgCl_2$). The number of atoms in these systems is ~68000. The 10 mM $MgCl_2$ system corresponds to the number of Mg^{2+} ions in the Twister crystal structure (PDB 4OJI, 5 Mg^{2+} ions) and the initial positions of the ions were based on the crystal structure in this case. The two missing nucleotides in the crystal structure were added by using the internal coordinates in CHARMM (39). With 20 mM and 100 mM $MgCl_2$ the Mg^{2+} were randomly placed in the simulation system. In addition to $MgCl_2$, 100 mM KCl was used in all four systems. The systems were first minimized and equilibrated in CHARMM (39) with harmonic restraints on the backbone and base non-hydrogen atoms with force constants 1 and 0.1 kcal/mol/$Å^2$, respectively, for 100 ps in the NPT ensemble with a timestep of 1 fs. CHARMM36 additive force field was used to model RNA and the ions (40-42). Water molecules were modeled by the CHARMM modified-TIP3P force field (43, 44). Smooth Particle Mesh Ewald method was applied for calculation of electrostatic interactions with a real-space cutoff of 12 Å (45). The Lennard-Jones potential was force switched to zero between 10 and 12 Å. A Monte Carlo anisotropic barostat was used to maintain pressure at 1 bar. The lengths of all covalent bonds that involve a hydrogen atom were constrained using the SHAKE algorithm (46). Equilibration was followed by production runs in the NPT ensemble using the Langevin integrator (1 bar, 30 ˚C) with no restraints for 200 ns in OpenMM (47, 48) with a time step of 2 fs.

Umbrella sampling MD was used to determine the potential of mean force (PMF) of breaking the tertiary contacts in Twister (49). The distance between the center of mass of two groups of nucleotides served as the reaction coordinate (Supporting Figure S1). The reaction coordinate was selected to globally lead to perturbation of the major tertiary contacts, T1 and T2, in Twister (Figure 1). The selection also allows for the relative positions of the fluorophores in the FRET study on Twister (36) to be monitored as a function of the RC. To generate unfolded conformations of the RNA the RC distance was gradually increased in 0.5 Å increments from 15 Å (initial average distance in the native conformation from the 200 ns MD simulations) to 40 Å (partially unfolded conformation), yielding a total of 51 windows. In addition, windows from 15 to 13 Å were generated and sampled in a similar fashion to assure that the true energy minima in the PMFs were identified. The windows were generated using CHARMM by running 10 ps

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

MD simulations at each 0.5 Å step along the reaction coordinate with a force constant of 5000 kcal/mol/Å^2 on the RC distance. During this stage the Watson-Crick (WC) base pairs forming the secondary interactions (P1, P2 & P4) were maintained by restraining the distances between the non-hydrogen atoms involved in all WC base pair hydrogen bonds. A force constant of 4 kcal/mol/Å^2 was used to apply NOE restraints in CHARMM to maintain the distance between hydrogen bonding atoms of the bases within 2.7 Å to 3.0 Å. The final structure in each window was used as the starting structure for the next window. Each window was then simulated in the NPT ensemble for 10 ns in OpenMM (47) where the COM RC restraint was enforced using PLUMED using a force constant of 4.782 kcal/mol/Å^2 (2000 kJ/mol/nm^2) (50). No restraints were applied on the base pairs of secondary interactions and the RNA was allowed to propagate freely. The final frames from these 10 ns simulations were used to start the GCMC/MD simulation at alternate windows for the final PMF calculations. MD simulation conditions were the same in 200 ns production runs of the ribozyme in the native state.

GCMC/MD protocol

A PME oscillating excess chemical potential (μ_{ex}) GCMC algorithm was used to achieve enhanced sampling of the ions around the RNA (51) in conjunction with umbrella sampling to yield the final PMFs. At each window 2 separate PMFs were initially calculated. They each involved 5 cycles of oscillating μ_{ex} GCMC/MD with each cycle involving a resampling of the ion distribution in the simulation system using μ_{ex} GCMC, as described below, followed by 10 ns of NPT MD using the above protocol. From the 10 ns MD portions of the 10 oscillating μ_{ex} GCMC/MD cycles, the last 6ns of data from each of the 10 cycles were collected, yielding a total of 60 ns of sampling in each window from which the free energy profiles were calculated. Calculation of the PMF was performed using the Weighted Histogram Analysis Method (WHAM) and error analysis performed by block averaging over 5 12 ns blocks in each window and calculating the standard deviation. In total four systems (0, 10, 20 and 100 mM $MgCl_2$) were subjected to umbrella sampling PMFs each involving 10 GCMC/MD cycles with 10 ns MD per cycle for each of the 26 windows yielding a total of 2.6 microseconds of MD for the Twister ribozyme.

The oscillating μ_{ex} GCMC involved a scheme to oscillate the μ_{ex} of the ions to overcome the low acceptance ratios and avoid using pre-hydrated Mg^{2+} ions as previously performed (52). In this protocol restricted Monte Carlo moves were used to resample the ion distribution by draining and refilling the system with Mg^{2+} ions, while K^+ ions were added and drained simultaneously to neutralize the total charge on the system. For systems with no Mg^{2+} ions, the K^+ and Cl^- ions were drained and refilled together to achieve redistribution of the K^+ ions. One GCMC-MD cycle included seven

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

stages: 1) Delete Mg^{2+} and insert K^+ over 20000 MC steps; 2) Rotate and translate all ions and water for 80000 MC steps; 3) Repeat steps 1 and 2 until only one Mg^{2+} is left and the system is neutral; 4) Insert Mg^{2+} and delete K^+ over 20000 MC steps; 5) Rotate and translate all ions and water for 80000 MC steps; 6) Repeat steps 4 and 5 until the Mg^{2+} concentration is reached and the system is neutral; and 7) minimize the entire system, including RNA, for 5000 Steepest Descent steps. Once resampling of the ion distributions was complete the system was then equilibrated for 100 ps of NVT MD followed by the 10 ns of NPT MD in OpenMM (47). In the PME/GCMC approach the μ_{ex} values to delete and insert the ions were adopted from the recent study of ionic hydration free energy and are shown in Supporting Table S2 (52).

GFE maps

Occupancies of Mg^{2+} ions were calculated as the number of times a voxel (1 Å cubic unit of volume) around Twister ribozyme was occupied by Mg^{2+} during the simulation. At each window we analyzed 10000 frames and all frames were first aligned with respect to the backbone of the ribozyme. These occupancies were converted to "grid free energies" (GFE) according to $G = -k_B T \ln(P)$, where P is the probability of occupying a voxel (1 Å cubic unit of volume) on the RNA surface relative to the voxel occupancy of the same species in bulk solution, k_B is the Boltzmann constant, and T is the temperature (303.15 K). More detailed description of this procedure is provided by Lemkul et al (32) and elsewhere (53).

Results and Discussion

In the present study an enhanced sampling method for ion distributions is applied in combination with umbrella sampling to study atomic-level interactions of Mg^{2+} ions with fully and partially folded states of the Twister ribozyme. The protocol implemented includes GCMC simulations to redistribute the Mg^{2+}, K^+ and Cl^- ions in the simulation system. The GCMC is followed by minimization and MD simulations to allow the RNA, ions and water to relax, and to obtain conformational sampling of the RNA. The GCMC/MD approach overcomes the shortcomings of classical MD simulations due to the low exchange rates of Mg^{2+} ions with water and phosphate groups. Notably, the method achieves exchanges of ions around the phosphate groups at the inner-shell level, which is not currently feasible with classical MD. Validation of the method is detailed in Supporting Text 1. The combination of this approach in conjunction with umbrella sampling thus allows a direct relation of the atomic details of Mg^{2+}-RNA

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

interactions with the free energies associated with the stabilization of the RNA to be obtained as detailed in the remainder of this manuscript.

Potentials of Mean Force as a Function of $MgCl_2$ concentration

Twister ribozyme assumes a unique structure with a combination of secondary and tertiary interactions forming a double pseudoknot (34). The major tertiary contacts, T1 and T2, are long-range interactions that involve Watson-Crick (WC) base pairing. The nucleotides comprising the individual strands contributing to the two contacts are close in the primary sequence. Accordingly, the center of mass (COM) distance between two groups of nucleotides, regions 1 and 2, that approximately each represent half of the folded structure and each include the single strand regions that include T1 and T2 was chosen as the reaction coordinate (RC)(Supporting Figure S1). Regions 1 and 2 include the nucleotides U24 and G54, respectively, to which the individual fluorophores in the smFRET experiments were covalently linked (36). Along this RC a folding potential of mean force (PMF) of the Twister ribozyme was calculated using umbrella sampling with four different concentrations of Mg^{2+}. To obtain the impact of Mg^{2+} on the PMFs, the umbrella sampling MD simulations at each RC were periodically stopped and the ion distribution in the system resampled using the GCMC protocol followed by equilibration and additional production umbrella sampling from which the full PMF was obtained. It is noted that this RC does not allow for sampling of the full range of unfolded states, as discussed below, but represents a computationally accessible 1-D RC allowing the sampling for unfolded states of the RNA to investigate their interactions with Mg^{2+}.

The resulting PMFs for Twister at the four $MgCl_2$ concentrations are shown in Figure 2a. The free energy surfaces show a sharp minimum at 15.5 Å, corresponding to the fully folded state. The free energy then increases rapidly to an inflection point at ~20 Å, following which the free energy gradually rises out to the fully extended state at RC = 40 Å. From the minima out to 19 Å the surfaces at the four $MgCl_2$ concentrations are similar. Beyond 19 Å the difference between the 0 and 100 mM $MgCl_2$ surfaces are significant with the inflection point occurring at ~23 kcal/mol for 100 mM $MgCl_2$ and ~30 kcal/mol for 0 $MgCl_2$, with that difference being ~7 kcal/mol throughout the 20 to 40 Å region of the PMF. The 10 and 20 mM $MgCl_2$ system PMFs are between the 0 and 100 mM $MgCl_2$ results beyond 20 Å out to 40 Å. To better understand these results with respect to experimental studies based on FRET and RNase T1 footprinting (36), analysis of the PMFs was undertaken specifically targeting the molecular phenomena directly related to those experiments.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

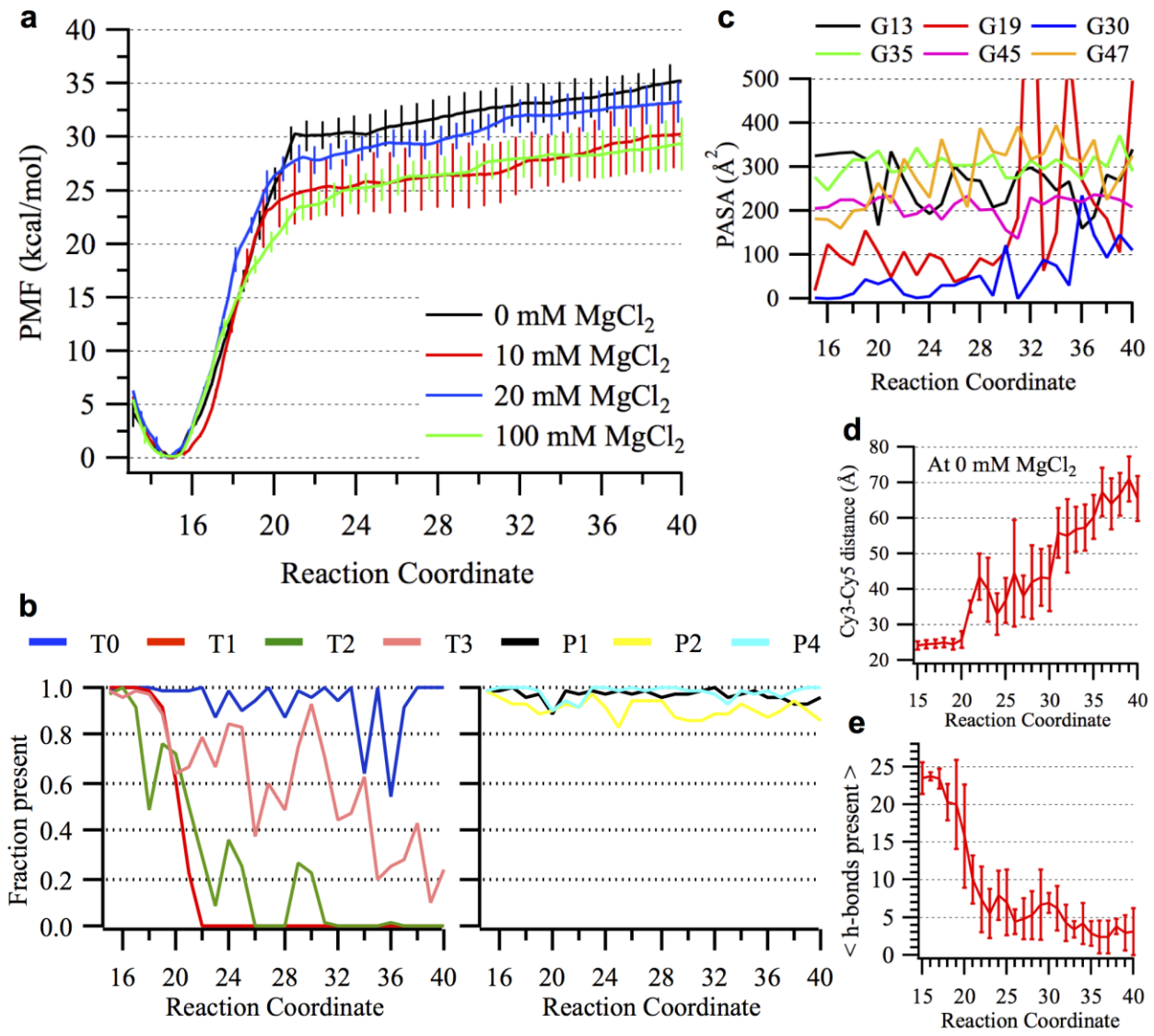

Figure 2. Analysis of Twister unfolding simulations. a) Potential of mean force (PMF) of Twister ribozyme at different concentrations of $MgCl_2$ based on the GCMC/MD simulations. The reaction coordinate (RC) corresponds to center of mass distance between two groups of residues that contribute to the T1 and T2 tertiary interactions (Supporting Figure S1). Native state corresponds to RC = 15.5 Å. PMFs were calculated on RC windows at 1 Å intervals from 13 to 40 Å with 60 ns of sampling per window with error bars based on standard deviations from five 12 ns samples extracted from the full 60 ns of sampling (see Methods). b) The fraction sampled of tertiary (T0, T1, T2 and T3) and secondary (P1, P2, P4) contacts in Twister present at a given RC. Data represents averages over the four $MgCl_2$ concentrations, with individual plots provided in Supporting Figures S5 and S6. A contact is present if the distance between the nucleotides is within 3 standard deviations of mean distance for the same contact

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

calculated from the MD simulations in the native state. Fraction is calculated based on the number of frames a contact is present out of all frames analyzed. c) Protein accessible surface area (PASA) for the guanine nucleotides protected from cleavage by RNase T1 calculated at each window along the reaction coordinate with 0 mM $MgCl_2$. PASA is calculated as the solvent accessible surface area using a probe radius of 10 Å as previously described (54). d) Interfluorophore distance (IFD) based on distance between C4' atoms of Ura 24 and Gua 54 along the reaction coordinate for Twister with 0 mM $MgCl_2$. e) Number of Watson-Crick hydrogen bonds present between nucleotides involved in tertiary interactions along the reaction coordinate averaged over the four $MgCl_2$ concentrations. Error bars indicate the standard deviation. A hydrogen bond was defined by donor-acceptor distance cutoff of 3.0 Å and donor-hydrogen-acceptor angle cutoff of 120˚.

FRET experiments were based on the fluorophores Cy3 and Cy5; Cy3 was inserted between nucleotides U24 and C25 while Cy5 was connected to the 3' termini via a ssDNA spacer (36). In the folded state the fluorophores are relatively close allowing for fluorescent energy transfer while upon unfolding the distance between the fluorophores increases leading to a loss of energy transfer. To model the spatial relationship of the fluorophores, distances between the C4' atoms of Ura 24 and Gua 54 were measured from which the average and standard deviations were obtained for each window in the PMFs. Shown in Figure 2d are the average values of the "Interfluorophore distance (IFD)" along the RC for 0 $MgCl_2$. Similar results were obtained with the other $MgCl_2$ concentrations (Supporting Figure S2). Over RC values out to 20 Å and beyond the IFD does not change, with IFDs beyond 24 Å not occurring until RC = 22 Å at which point the PMFs have passed their inflection points, attaining energies of 23 kcal/mol or more. Accordingly, significant changes in the actual distance between the fluorophores that would lead to a loss of fluorescence will not occur until this distance along the PMF. In addition, given the length of the linkers on the fluorophores (Supporting Figure S2), including the region of ssDNA linking Cy5 to the RNA, significant unfolding of the RNA may be required to observe the change of fluorescence. This suggests that the FRET experiments may be reporting events that are occurring beyond the global minimum energy wells in the PMFs. Accordingly, interpretation of the present simulation results in the context of the FRET data will focus on the regions of the PMF beyond the inflection point at RC = 20 Å. In this region the 100 mM $MgCl_2$ system is ~7 kcal/mol more favorable than the 0 mM system with the 10 and 20 mM results intermediate to those values. This indicates that the presence of 100 mM $MgCl_2$ lowers the free energy of Twister in these partially unfolded states. Supporting this interpretation is the presence of a small population of folded state at micromolar Mg^{2+} and a small population of unfolded state at and above 20 mM Mg^{2+} in the FRET experiments(36). The presence of small populations of the alternate states is

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

consistent with a free energy difference of ~ 7 kcal/mol versus a difference of over 20 kcal/mol between the global minima and the inflection at ~20 Å in the PMF.

A similar interpretation of the experimental results on unfolding based on RNase T1 footprinting experiments may be made (36). Shown in Figure 2c are protein accessible surface areas over the guanine bases in Twister known to undergo hydrolysis. The individual plots with error analysis for each nucleotide are shown in Supporting Figure S3. While the accessibilities of different nucleotides vary, the values for the individual nucleotides do not increase significantly to well beyond the inflection point at 20 Å in the PMF. This indicates that increased hydrolysis being observed in the footprinting experiments is also not monitoring events associated with transitions from the global free energy minimum to the partially unfolded states (R = 20-40 Å), but rather associated with events beyond that region that correspond with larger unfolding events.

Consistent with the above model are β parameter results that are indicative of the region along the folding path where the transition state occurs leading to the fully unfolded state. With Twister β_{fold} = 0.37 and β_{unfold} = 0.72 indicating that the transition to the fully unfolded state observed in the FRET and footprinting experiments occurs closer to the unfolded state (36). In addition, the folding rate is under 0.1 s^{-1}, which is not consistent with the over 20 kcal/mol minimum observed in the PMFs in Figure 2a (55). These results suggest that the experimental results are monitoring folding events that are occurring in the region beyond the inflection point at ~ 20 Å in the PMF. In this scenario the increased population of the folded state observed in the FRET experiments in the presence of Mg^{2+} is consistent with the free energy of the 100 mM $MgCl_2$ system being ~7 kcal/mol more favorable that 0 mM $MgCl_2$. However, beyond the inflection region in the PMF there is no indication of a transition to a further unfolded state, suggesting that the present RC is not sampling the full folding surface of the Twister RNA. Analysis of example conformations from the 100 mM $MgCl_2$ PMF in Figure S4 of the supporting information show even at the longest RC = 40 Å, the RNA still contains a significant amount of Watson-Crick base pairing along with helical structure associated with the P1 loop. Loss of all or some of such interactions may lead to the unfolded states observed experimentally.

While the present PMF may not be representative of the full folding surface of Twister, the deep global minima may explain the ability of Twister to self cleave at very low Mg^{2+} concentrations. The presence of the deep minima, which are relatively insensitive to ion concentration, suggests that sampling of the folded state required for catalysis may occur in all conditions, which may be due to the background of 100 mM KCl (56) included in the present study to be consistent with the experimental study. Overall, the present results indicate a model where the global folded state observed experimentally is accessible even at low Mg^{2+} concentrations and that this conformation

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

is related to the catalytic activity. However, the experimental FRET and RNase T1 footprinting studies are monitoring unfolding events at RC distances beyond the inflection point at 20 Å in the PMFs, with the present RC used to define the folding landscape not accessing the fully unfolded conformations. Additional studies are needed to address this issue.

Analysis of the tertiary and secondary contacts (Figure 1) averaged over all the four systems as a function of the RC is shown in Figure 2b, with results for the individual systems shown in Supporting Figure S5. In the global minima there is some loss of T2 tertiary interactions, the second largest tertiary contact, in the vicinity of RC = 17 to 18 Å as the free energy rapidly increases. The larger T1 interaction starts to be perturbed at RC 19 Å but it is not until beyond the inflection point in the PMF that the T1 interaction is lost, with the remaining T0 and T3 interactions largely maintained through the inflection point out to larger RC values. The secondary contacts are largely maintained throughout the PMFs (Supporting Figure S6), consistent with a report that secondary structures of RNA are independently stable of the tertiary interactions (57). In the P2 secondary interaction a cis-Hoogsteen-sugar edge pair between A8-G45 is lost during initial stages when T1 is being broken. In later stages with T1 fully broken, P2 is slightly compromised in order to stabilize the newly freed bases A46:G49. The inclusion of 100 mM background KCl in the present study likely contributes to the stabilization of the secondary interactions throughout the PMF.

To understand the contributions to the energy increase ranging from 23 to 30 kcal/mol upon moving from the global minima to the inflection point in the PMF the Watson Crick (WC) hydrogen bonds associated with the tertiary interactions were monitored. The tertiary interactions, T1, T2 and T3 correspond to 11, 6 and 3 hydrogen bonds, respectively, associated with 7 WCs pairs. Upon going from the fully folded state to the inflection point in the PMF approximately 80% of the T1 and T2 interactions are lost by 22 Å corresponding to a loss of 17 hydrogen bonds (Figure 2e). Considering that each WC hydrogen bond corresponds to a stabilization free energy ranging from -0.7 to -2 kcal/mol (58, 59) their loss corresponds to -12 to -34 kcal/mol indicating that the deep free energy wells in the PMFs are associated with perturbation of the hydrogens bonds associated with the T1 and T2 tertiary interactions. Analysis of traces of the phosphodiester backbone of Twister at various stages along the 100 mM $MgCl_2$ PMF shows that at RC = 20 and 25 Å the overall 3D conformation of the RNA is maintained (Supporting Figures S4). The relatively small change in the overall conformation of the RNA upon going from the global minimum to the region of the inflection point in the PMF further supports the conclusion that the experimental FRET and RNase folding studies are not monitoring the initial loss of tertiary structure, but rather a larger scale unfolding event that is not being observed in the present calculations. This yields a scenario where the initial events in unfolding of Twister are dominated by local perturbations of

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

hydrogen bond interactions participating in WC interactions present in the tertiary contacts while the overall structure of the ribozyme is maintained. Once these hydrogen bonds are broken, associated with the change in free energy of 23 to 30 kcal/mol seen in the PMFs, then larger scale unfolding of the RNA can occur associated with the gradual increase in free energy beyond the inflection point at 20 Å in the PMFs.

Mg^{2+} distribution and non-sequential phosphate-phosphate interactions

While the present PMF doesn't sample fully unfolded states, the energetic differences between the 0 and 100 mM $MgCl_2$ concentrations beyond the PMF inflection point indicate that experimentally relevant partially unfolded states are being sampled. Accordingly, additional analysis focused on understanding the atomic details of the interaction of Mg^{2+} with Twister, including in both the folded state and partially unfolded conformations and how those interactions are leading to stabilization of the RNA. This analysis is based on the ability of the GCMC-MD protocol to sample a representative ensemble of Mg^{2+} ions around the RNA in both the folded and unfolded states.

Analysis of the distribution of Mg^{2+} ions around the RNA was initially undertaken for the folded state based on grid free energies (GFE) that are obtained directly from the ion probability distributions as described in the methods and previously performed (32). Figure 3 presents the distribution of Mg^{2+} around Twister from the RC = 15 Å PMF windows in the presence of 10, 20 and 100 mM $MgCl_2$ along with the crystallographic ion positions. It is evident that the simulations produce distributions that recapitulate the crystallographic locations of Mg^{2+} even at the lower $MgCl_2$ concentration. It shows that with decreasing concentration the Mg^{2+} ions still sample sites observed in the crystal structure along with additional regions. Figure 3c shows that even at 10 mM $MgCl_2$, Mg^{2+} ions are specifically distributed at crucial intersections, the majority of which correspond to the specific phosphate pairs discussed below. However, at the lower $MgCl_2$ concentrations it is likely the convergence of the ion distribution may require additional sampling, which may impact the free energy surfaces shown in Figure 2a. Additional tests are required to address this issue.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

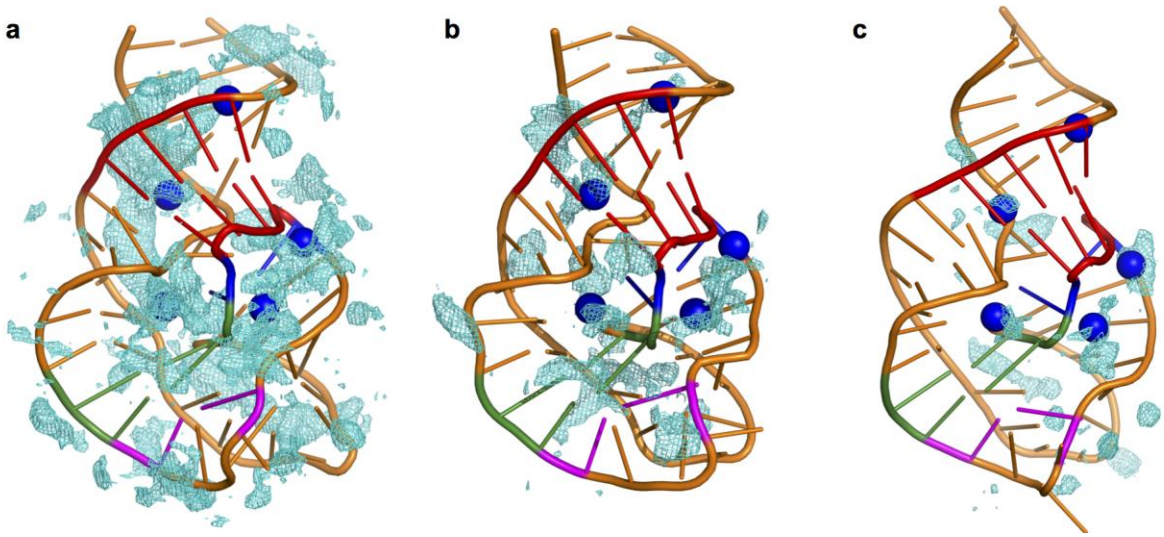

Figure 3. Occupancy maps for Mg^{2+} ion distribution around native conformation of the Twister ribozyme from the GCMC-MD simulations. a) 100 mM $MgCl_2$ b) 20 mM $MgCl_2$ and c) 10 mM $MgCl_2$. The cutoff for GFE is -2 kcal/mol. The crystallographic positions of Mg^{2+} binding sites are shown in blue spheres after aligning the structures.

Further analysis of interactions of the RNA with Mg^{2+} along the reaction coordinate focused on the presence of spatially adjacent non-bridging phosphate oxygens (NBPOs) of non-sequential nucleotides (*i.e*. phosphate moieties separated by one or more phosphates in the primary sequence) and their spatial relationship to Mg^{2+}. Phosphate-Mg^{2+} interactions, including both direct and indirect coordination, may stabilize both the folded and partially unfolded states by facilitating short phosphate-phosphate interactions. Non-sequential NBPO pairs with a high probability of being within 9 Å of each other were identified as pairs that may simultaneously interact with Mg^{2+} through outer shell interactions (Figure 4). The cutoff of 9 Å is defined based on radial distribution functions (RDF) of the distance between NBPO and Mg^{2+} ions (Supporting Figure S7) that show peaks at 1.95 Å, 4.2 Å and 6 Å corresponding to direct-interaction (Inner-Shell-Dehydrated: ISD), indirect-interaction (Outer-Shell-Dehydrated: OSD) and diffuse-interaction (Non-Dehydrated: ND), respectively (8). Here, we focus on the ISD and OSD interactions, which have been classified as strong interactions.

To identify interacting non-sequential NBPOs, adjacent NBPO probability analysis was performed for selected portions of the PMF by merging the data for four consecutive windows (Figure 4). Analysis of Figure 4 shows the presence of non-sequential NBPO pairs that are adjacent to each other that are far apart in the primary sequence as well as a number of NBPO pairs that are adjacent to the diagonal associated with phosphates separated by one or more nucleotides. Table 1 defines six

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

regions of non-adjacent NBPO interactions that occur in the folded state that are identified in the top-left panel of Figure 4. These are shown in Figure 1c using the same color scheme as used for the labels in Table 1. Of these a number are situated around the T1 and T2 tertiary interactions (Table 1). Notable of the pairs present in the folded state is the pair C9-A28 (Figure 1c, red spheres) in which A28 is situated at the end of T1, such that this pair coming together facilitates base-pair formation in T1. These NBPOs are exposed to the solvent allowing for Mg^{2+} to simultaneously coordinate with both of them. Additionally, the U6-C25 pair (Figure 1c, pink spheres) is situated on the other side of T1 further indicating the importance of bringing phosphates together to stabilize the pseudoknot structure of the ribozyme. Between G3:C5 and A46:G47 (Figure 1c, C4-A46 black spheres) there is a range of sites where Mg^{2+} ions may interact with more than 2 NBPOs. Formation of these pairs appear to stabilize the helical twist of P1 secondary motif while stabilizing nucleotides involved in T1 that may facilitate formation of that contact. The pairs A8-G30 and A8-C31 (Figure 1c, cyan spheres) provide an encapsulated space for Mg^{2+} that bridges nucleotides on T1 and T2. NBPO pairs between C32:C33 and A41:G42 (Figure 1c, C32-A41 green spheres) are adjacent to and appear to stabilize the T2 interaction. Finally, the C20-G30 pair (Figure 1c, blue spheres) is within 9 Å throughout the majority of the PMF and appears to stabilize nucleotides in the hairpin loop on the end of P4 that participate in both T1 and T2 as well as stabilizing the T0 (U24-A29) and T3 (C15-G19) interactions.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

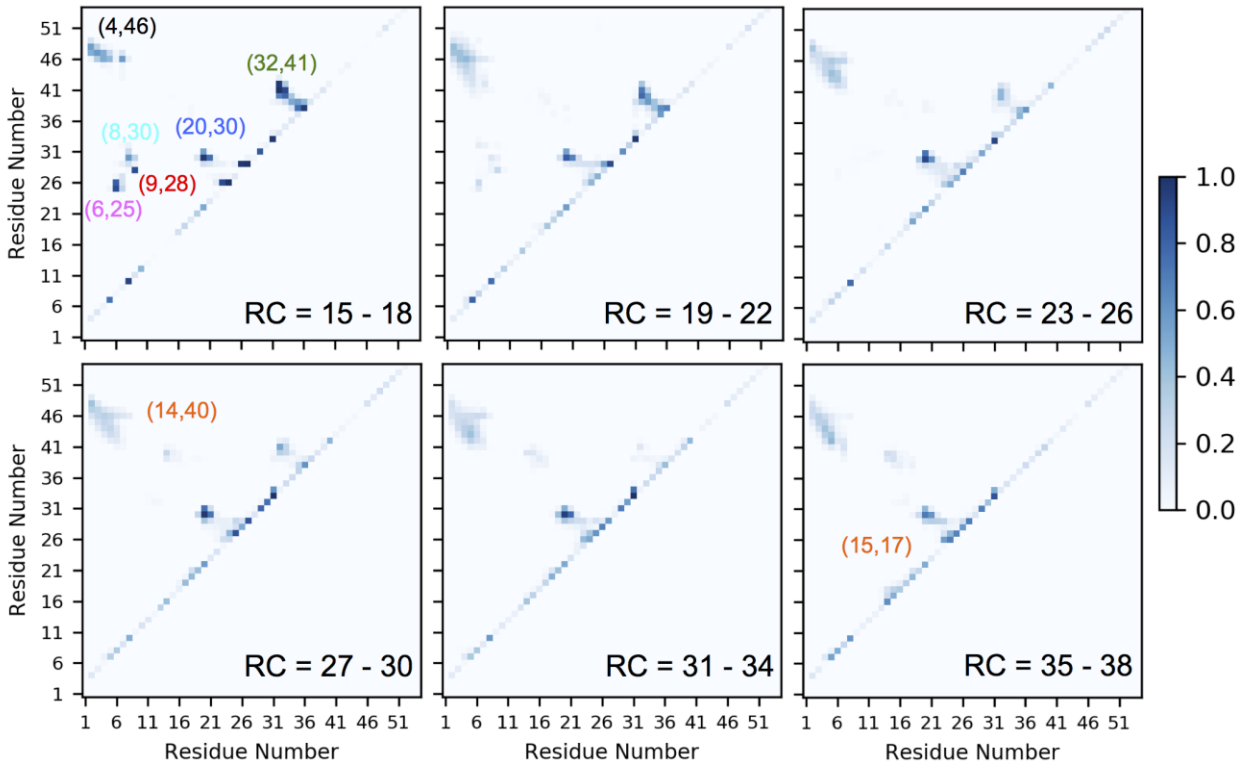

Figure 4. NBPO probability matrices from NBPO distance analysis for Twister from the GCMC-MD PMF at 100 mM $MgCl_2$. The probabilities were averaged over groups of 4 RC windows. RC = 15 – 18 corresponds to the fully folded structure. The interactions discussed in the text have been specified with circles.

Table 1. List of pairs that show high probabilities of their NBPOs being within 9.0 Å in the folded state.

	Pairs with NBPO correlation	Adjacent tertiary interaction
●●	C4 – A46	T1
●●	C9 – A28	T0, T1
●●	A8 – G30	T1
●●	U6 – C25	T0, T1
●●	C32 – A41	T2
●●	C20 – G30	T2, T3

Figure 5 shows the distribution of Mg^{2+} ions around the phosphate pairs listed in Table 1 using GFE maps at RC = 15 Å in the 100 mM $MgCl_2$ system. The presence of highly favorable regions of Mg^{2+} sampling directly adjacent to or between all the NBPO pairs is evident. For example, a large region is sampled around the pair 4-46 where phosphates are lined in parallel fashion with NBPOs facing each other. In

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

general, the presence of Mg^{2+} in the regions surrounding the interacting NBPO pairs will stabilize these interactions, which will lower the free energy of the partially unfolded states. In addition to the NBPO pairs discussed above, there are high probability pairs adjacent to the diagonal of the matrix in top left panel of Figure 4. These include 5-7, 8-10, 24-26, 27-29, 29-31, 31-33, and 36-38. These represent kinks in the backbone of the ribozyme that may play an important role in the formation of tertiary structure. It has been reported that such kinked structures are stabilized by divalent metal ions (60).

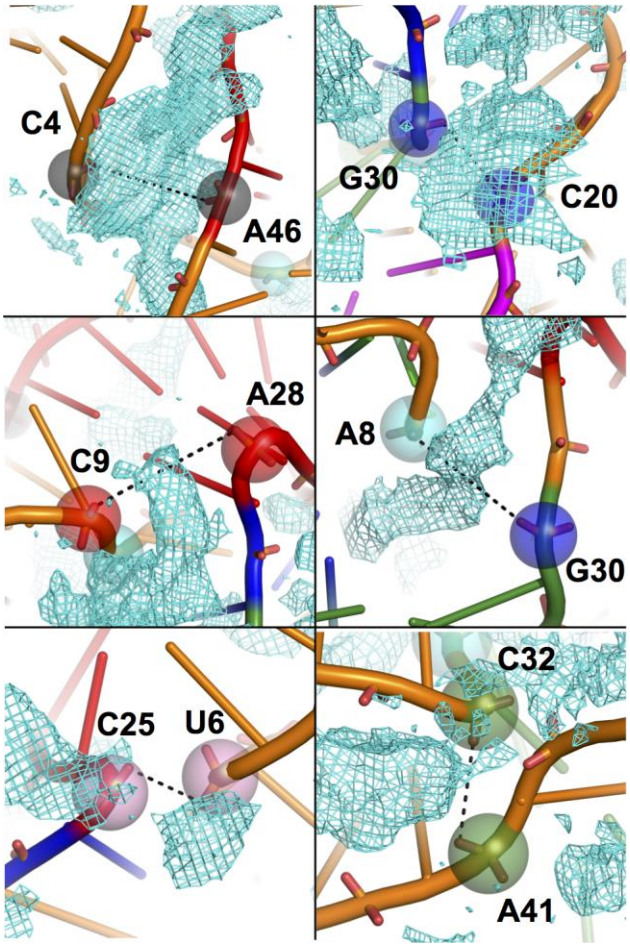

Figure 5. Distribution of Mg^{2+} ions around the NBPOs of residue pairs. From the RC = 15 Å native state of Twister ribozyme (cartoon) from the 100 mM $MgCl_2$ PMF. The phosphates are represented as same-colored transparent spheres. GFE maps are illustrated as cyan mesh at -2 kcal/mol cutoff.

At higher RC values a number of the pairs present in native state are lost while two new pairs appear. Pairs that are lost correspond to pairs whose individual partners are assigned to the two groups used to define the RC. For example, the 6-25, 8-30, and 9-28 NBPO pairs are lost. New interacting pairs at long RC values include the 15-17

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

pair that occurs at RC values greater than 30 Å and the 14-40 pair that first occurs in the 27-30 RC portion of the PMF.

The NBPO probability analysis was also performed for Twister with 0 mM $MgCl_2$ (Supporting Figure S8). In the absence of Mg^{2+} ions, near the folded state (RC = 15 to 18) the pattern of NBPO pair is similar to that at 100 mM $MgCl_2$, with the 9-28, 8-30, 6-25 and 32-41 pairs being present. However, some differences occur at larger RC values. The group of NBPOs near the 4-46 pair shows lower probabilities with the interaction pattern shifted and the interaction between 15-17 pair at RC = 35-38 is not observed. The similarity of the NBPO pairs in the folded state is consistent with experimental data indicating the ability of Twister to fold and self cleave at concentrations approaching 0 Mg^{2+} as well as with the present PMF (Figure 2a), while the difference at the larger RC values suggest a role of the ion in stabilizing the partially unfolded states.

To identify Mg^{2+} ions in the vicinity of the NBPO pairs along the PMF the probability of Mg^{2+} being present within 6.5 Å of the NBPO atoms of both phosphates in each pair along the RC was calculated. Data are presented for the pairs in Table 1 along with the 14-40 and 15-17 pairs seen in the partially unfolded states. Shown in Figure 6 are those values as a function of the RC along with the fraction of the NBPO pair that is within the 9 Å cutoff. It is evident that whenever the NBPOs are close to each other there is a significant probability of Mg^{2+} ion being in contact with both of them. The collection of NBPO pairs designated by 4-46 (Figure 4) was present at fraction probabilities of 0.2 or more along the RC with similar probabilities of Mg^{2+} in their vicinity. While this interaction is present for the full range of the PMF, it occurs at higher probabilities in the vicinity of the folded state with the Mg^{2+} probability following that trend. With pairs 9-28, 8-30 and 6-25, which involve nucleotides participating in T1, Mg^{2+} ions are interacting with both members while those interactions are present out to RC = 22 Å. The pair 32-41, which is situated around tertiary contact T2, also has Mg^{2+} extensively coordinated with this NBPO pair. As discussed above, pair 20-30 is maintained along the full PMF with Mg^{2+} ions consistently interacting with the pair. Notably, a similar pattern is observed with the NBPO pairs not present in the folded state. The presence of both the 14-40 and 15-17 pairs in the unfolded states along the PMF is strongly correlated with interactions with Mg^{2+}. Such transient interactions that are facilitated by Mg^{2+} may be important for guiding the RNA along the folding pathway, as previously discussed (18, 25), though further all-atom simulation studies are needed to directly address this possibility.

Interesting trends are seen with NBPO pairs that impact the T0 and T3 tertiary contacts. T0 is present throughout the majority of the PMF, though occurs at a lower probability in the most unfolded states (Figures 2b). This contact, which occurs in the L4

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

loop, is associated with the 20-30 NBPO pair maintained throughout the PMF. A similar trend is present with the T3 tertiary interaction (15-19 bulge) for which sampling at low levels occurs in the most unfolded states sampled in this study and gradually increasing at smaller RC values (Figure 2b). Analysis of the NBPO probability matrices shows the presence of a 15-17 interaction near the diagonal at RC=35-38, and Figure 6 shows that Mg^{2+} ions are present around this pair. It suggests that this interaction, which is associated with a kink in the RNA, facilitates formation of the T3 contact.

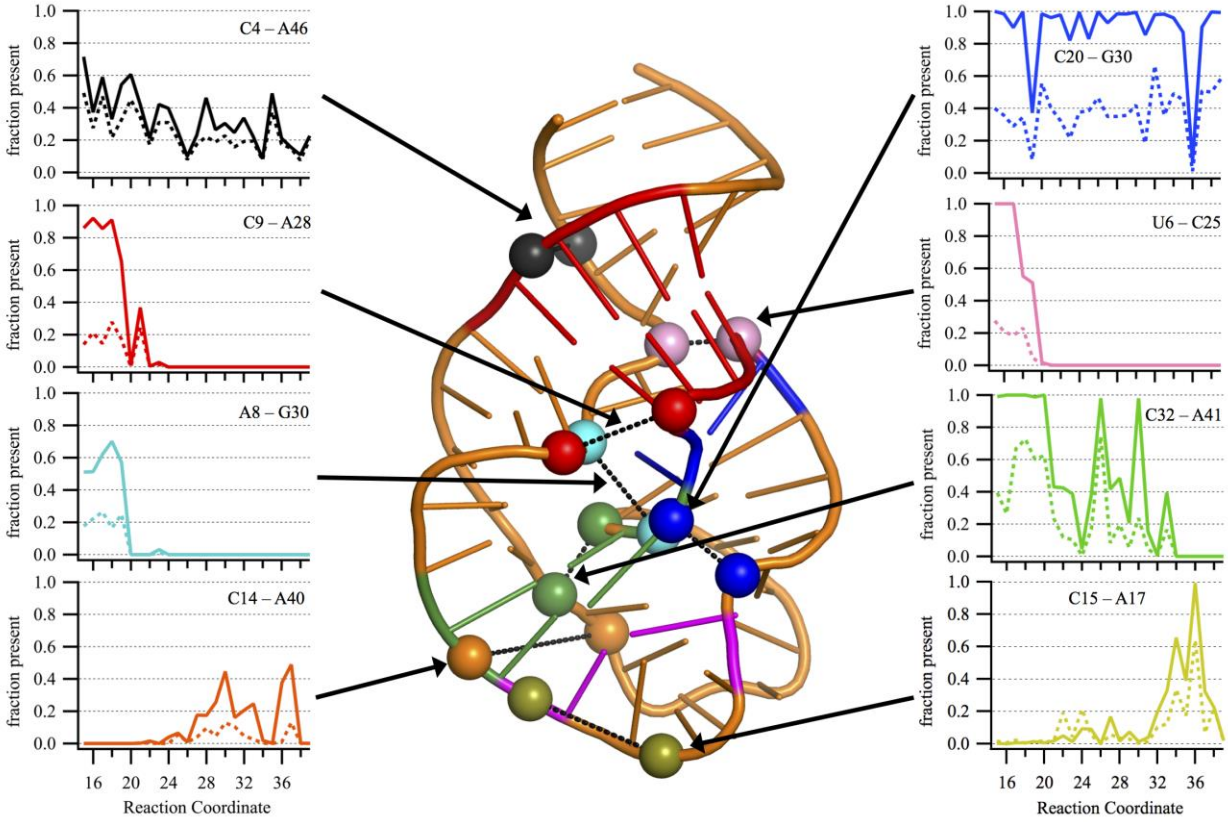

Figure 6. Adjacent NBPO pairs in contact with Mg^{2+} ions at 100 mM $MgCl_2$. The probability of different pairs of nucleotides with NBPOs within 9 Å (Solid line) and the probability of a Mg^{2+} ion within 6.5 Å of NBPOs of both phosphates coordinating the two NBPOs (Dashed line) along the reaction coordinate. The arrows point out the positions of NBPO pairs in the Twister structure shown in cartoon and the phosphate atoms shown as spheres with similar colors. Plots for 20mM and 10 mM $MgCl_2$ are provided in Supporting Figure S9.

To quantify the impact of the presence of Mg^{2+} on the NBPO pairs the overall probabilities of the presence of those pairs was calculated by integrating over all the individual pair probabilities excluding those in adjacent phosphates. The results from this analysis are presented in Table 2. Comparison of the results for 100 and 0 mM

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

show the probabilities to be systematically higher throughout the PMF at 100 mM $MgCl_2$, with 10 and 20 mM values typically falling between the 100 and 0 mM values consistent with the PMFs shown in Figure 2a. This difference indicates that while the majority of NBPO pairs are present in 0 $MgCl_2$, the presence of Mg^{2+} has an overall stabilizing effect on those interactions. This is consistent with the model of Mg^{2+} stabilizing the repulsion associated with phosphate-phosphate interactions occurring in the tertiary structure (7, 61). Thus, the presence of Mg^{2+} neutralizing phosphate-phosphate repulsion associated with tertiary interactions leads to an overall increase in the sampling of the non-sequential NBPO pairs in Twister. This increased sampling is suggested to contribute to the lowering of the free energy of the RNA as seen in the differences in the PMFs beyond the inflection point where the 100 mM $MgCl_2$ system is consistently 7 kcal/mol lower than the 0 mM $MgCl_2$ system.

Table 2. Summed probability of the non-sequential NBPO pairs being within the 9.0 Å cutoff at various stages of unfolding for systems at different Mg^{2+} concentrations. The analysis involves all phosphate-phosphate pairs that are not on adjacent sequential nucleotides in the RNA.

Reaction Coordinate	100 mM	20 mM	10 mM	0 mM
15 to 18	34.8 ± 0.6	31.1 ± 0.5	35.1 ± 0.4	29.9 ± 0.4
19 to 22	33.2 ± 1.9	29.7 ± 1.2	30.2 ± 0.6	28.1 ± 1.1
23 to 26	29.4 ± 1.5	28.7 ± 0.7	30.4 ± 1.7	28.5 ± 0.8
27 to 30	29.5 ± 1.2	29.9 ± 1.5	28.2 ± 2.3	25.7 ± 0.9
31 to 34	28.6 ± 1.6	26.4 ± 1.6	28.9 ± 1.0	25.4 ± 0.4
35 to 38	27.7 ± 1.3	24.1 ± 1.5	27.3 ± 0.9	25.2 ± 0.5

To investigate the impact of Mg^{2+} on the sampling of the individual NBPO pairs the summed probabilities were obtained over the different regions of non-sequential NBPO pairs shown in Figure 4 at 100 mM vs. 0 mM $MgCl_2$ and the differences determined. In addition, the contributions from all pairs adjacent to the diagonal were obtained. The differences between the two concentrations are shown in Table 3. As may be seen, in the vicinity of the folded state and the region leading to the inflection point in the PMF a number of pairs are more populated in 100 mM $MgCl_2$ including 8-30, 32-41, 4-46 and 20-30. The largest difference occurs with the 4-46 interactions with that increase occurring throughout the PMF. More sampling is also seen with the 15-17 pair at the RC values beyond 30 Å. For the interactions adjacent to the diagonal the difference are relatively small, being both higher and lower depending on the region of the PMF. Interestingly, for some non-sequential NBPO pairs there is a decrease in the

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

probability of sampling at 100 mM $MgCl_2$ indicating that while Mg^{2+} leads to an overall increase in sampling of non-sequential NBPO pairs, there is variability in the specific contributions at an atomic level.

Table 3. Difference between the summed probabilities at 100 and at 0 mM $MgCl_2$ for the individual non-sequential NBPO pairs and for the non-sequential pairs adjacent to the diagonal on Figure 4. The near diagonal values were calculated based on the total values shown in Table 2 minus the sum of the NBPO pairs listed in Table 3. Absolute values are shown in Supporting Table S1a and standard deviations in the differences are shown in Supporting Table S1b.

Difference: 100 - 0 mM $MgCl_2$									
RC	**pp9-28**	**pp6-25**	**pp8-30**	**pp32-41**	**pp4-46**	**pp20-30**	**pp14-40**	**pp15-17**	**Near Diagonal**
15 to 18	-0.08	-0.40	0.20	0.76	2.99	0.93	-0.13	0.11	0.56
19 to 22	-0.26	-0.71	0.18	1.33	2.74	0.95	0.00	-0.02	0.90
23 to 26	0.01	0.01	0.02	-0.65	3.65	-1.17	-0.12	-0.13	-0.75
27 to 30	0.00	0.00	0.00	0.25	3.54	-0.71	-0.40	0.04	1.11
31 to 34	0.00	0.00	0.00	-0.11	2.84	0.40	-0.73	0.24	0.60
35 to 38	0.00	0.00	0.00	0.00	3.05	-0.70	0.02	1.10	-0.96

While Mg^{2+} overall stabilizes Twister as well as other RNAs, the results in Table 3 indicate a more complex picture. Several pairs, including 8-30, 32-41, and 20-30 are stabilized in the vicinity of the inflection point in the PMF as well as in the fully folded states as expected. The 4-46 interaction is significantly stabilized throughout the PMF, indicating its impact on the stabilization of the RNA by Mg^{2+} in the full range of conformations studied. In contrast, decreases in the sampling of adjacent NBPO pairs occurs with 9-28 and 6-25 in the inflection and fully folded regions of the PMF while decreases at longer RC values occur with the 32-41, 20-30 and 14-40 pairs. With 14-40 the decreased sampling of close interactions is consistent that that pair not being present in the folded state, such that its destabilization would facilitate folding. With the 32-41 and 20-30 pairs the destabilization in the RC=23 to 26 region may help to avoid formation of tertiary interaction too early in the folding process, thereby facilitating the overall folding process. The largest decrease in the sampling of the 9-28 and 6-25 pairs occurs around the inflection point on the PMF at RC=19 to 22 Å. As these pairs are directly adjacent to the T1 tertiary contact, their destabilization may facilitate proper formation of the WC hydrogen bonds that dominate that tertiary contact. Notably, Mg^{2+} is sampling in the vicinity of these pairs (Figure 5) indicating that despite the presence of Mg^{2+} the increased sampling of other NBPO pairs occurs at the expense of these interactions consistent with a push-pull mechanism for the stabilization of RNA by Mg^{2+}.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

To further investigate the relationship of the 9-28 and 6-25 NBPO pairs to the T1 tertiary contact, correlation analysis was undertaken between the individual NBPO pairs and the number of T1 WC-associated hydrogen bonds in the region of the inflection point in the PMF (RC = 19 to 21, Table S1c). The Pearson correlation coefficients in the case of the 9-28 NBPO pair were -0.35 and -0.62 in 0 and 100 mM $MgCl_2$, respectively, while with the 6-26 NBPO pair the values were -0.37 to -0.65, respectively. Plots of the NBPO distances versus number of T1 WC hydrogen bonds are presented in Supporting Figure S10. These results show that the shorter NBPO distances correspond to a larger number of T1 hydrogen bonds, as expected given the spatial relationship of the two NBPO pairs to the T1 tertiary contact (Figure 1). Interestingly, the increase in the magnitude of these correlations in the presence of Mg^{2+} indicates how the ion influences the communication between these two classes of interactions. However, the decrease in the sampling of shorter NBPO distance in the presence of Mg^{2+} (Table 3) indicates that the formation of the T1 hydrogen bonds is actually decreased due to the presence of Mg^{2+}. This further indicates a complex push-pull type of mechanism where the presence of Mg^{2+} overall favors the tertiary structure but the lowered sampling of short 9-28 and 6-25 NBPO pairs in the inflection region of the PMF where the tertiary contacts are initially forming may allow for the hydrogen bonds to sample a wider range of conformational space. Such additional sampling would allow for formation of the correct pattern of hydrogen bonds associated with the WC interactions that occur in the fully folded state.

Conclusions

The central observation from the present calculations is the overall stabilization of the non-sequential NBPO interactions pairs by Mg^{2+}, an observation made in course-grained simulations studies of the *Azoarcus* ribozyme (18, 25). Notably, in the present study it is shown that Mg^{2+} leads to increased probability of the sampling of specific non-sequential NBPO pairs offering direct evidence of Mg^{2+} contributing to the energetic stabilization of the RNA observed in the PMFs. The most notable such effect occurred in the vicinity of pair 4-46. The region is involved in stabilization of the P1 helix along with the nucleotides that participate in the T1 tertiary contact. This suggests the Mg^{2+} is specifically interacting with the RNA to stabilize the conformation of this region thereby facilitating formation of the T1 contact upon full folding of the RNA. Previous studies have reported conflicting results on the importance of this stem in folding and activity (34, 37, 38). Interestingly, Mg^{2+} is not seen in this region of Twister in the crystallographic structures (34, 38).

Other NBPO pairs are stabilized or destabilized to varying extents consistent with Mg^{2+} facilitating RNA folding and stability in a sequence specific fashion. Decreased

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

sampling of specific non-sequential NBPO pairs in the presence of Mg^{2+} is a somewhat surprising result (Table 3). For example, in the vicinity of the inflection region of the PMF and in the folded states the sampling of the 9-28 and 6-25 pairs is decreased by Mg^{2+} as also occurs in the partially unfolded states with the 14-40 pair. This shows that the specific effects of Mg^{2+} include destabilization of selected non-sequential NBPO pairs as other phosphate-phosphate interactions are being stabilized in a push-pull mechanism. In the case of 14-40, as this interaction is not present in the folded states, its destabilization would facilitate access to those folded states. In the case of 9-28 and 6-25 the results indicate that partial destabilization of the pairs facilitates proper formation of the hydrogen bonds associated with the WC interactions of the T1 tertiary contact such that Mg^{2+} allows Twister to break and form the hydrogen bonds associated with the T1 tertiary contact. This appears to contribute to the smooth transition in the PMFs around inflection point in the presence of Mg^{2+} versus in its absence where a sharp transition occurs (Figure 2a). These results indicate a push-pull scenario of Mg^{2+} stabilization of RNA where, despite the general neutralization of phosphate-phosphate repulsion, the nature of the Mg^{2+} interactions leads to lowered sampling of specific non-sequential phosphate pairs that are being sacrificed to yield an overall stabilization of the RNA. While we anticipate that the identified push-pull mechanism contributes to the stabilization of other RNAs by Mg^{2+}, given the unique features of Twister where Mg^{2+} concentrations approaching zero still allow for self cleavage, it would appear that the role of Mg^{2+} in RNA structure and function varies for different systems.

Author Contributions

ADM Jr. conceived the overall project. AAK and ADM Jr. designed the methods and models for simulations. AAK built and performed the simulations. AAK and ADM Jr. analyzed the data and wrote the manuscript.

Conflict of Interest

ADM Jr. is co-founder and CSO of SilcsBio LLC.

Acknowledgements

We thank the National Institutes of Health [GM131710] for financial support for this work and the Computer-Aided Drug Design Center at the University of Maryland Baltimore for computing time. We thank Dr. Sarah Woodson and Dr. Darrin York for helpful discussions.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Supporting Citations

References (62-64) appear in the Supporting Material.

References

1. Fedor, M. J. 2002. The role of metal ions in RNA catalysis. Curr. Opin. Struc. Biol. 12(3):289-295.

2. Cech, T. R., and J. A. Steitz. 2014. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. Cell 157(1):77-94.

3. Morris, K. V., and J. S. Mattick. 2014. The rise of regulatory RNA. Nat. Rev. Genet. 15(6):423-437.

4. Pyle, A. M. 1993. Ribozymes: a distinct class of metalloenzymes. Science 261(5122):709-714.

5. Usman, N., L. Beigelman, and J. A. McSwiggen. 1996. Hammerhead ribozyme engineering. Curr. Opin. Struc. Biol. 6(4):527-533.

6. Tinoco, I., and C. Bustamante. 1999. How RNA folds. J. Mol. Biol. 293(2):271-281.

7. Draper, D. E., D. Grilley, and A. M. Soto. 2005. Ions and RNA Folding. Annu. Rev. Bioph. Biom. 34(1):221-243.

8. Draper, D. E. 2004. A guide to ions and RNA structure. RNA 10(3):335-343.

9. Heilman-Miller, S. L., D. Thirumalai, and S. A. Woodson. 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme. I. Equilibrium stabilization by cations. J. Mol. Biol. 306(5):1157-1166.

10. Woodson, S. A. 2010. Compact intermediates in RNA folding. Annu. Rev. Biophys. 39:61-77.

11. Fang, X., T. Pan, and T. R. Sosnick. 1999. A thermodynamic framework and cooperativity in the tertiary folding of a Mg2+-dependent ribozyme. Biochemistry-US 38(51):16840-16846.

12. Heilman-Miller, S. L., J. Pan, D. Thirumalai, and S. A. Woodson. 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme. II. Counterion-dependence of folding kinetics. J. Mol. Biol. 309(1):57-68.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

13. Misra, V. K., R. Shiman, and D. E. Draper. 2003. A thermodynamic framework for the magnesium-dependent folding of RNA. Biopolymers 69(1):118-136.

14. Pyle, A. 2002. Metal ions in the structure and function of RNA. J. Biol. Inorg. Chem. 7(7-8):679-690.

15. Tsui, V., and D. A. Case. 2001. Calculations of the Absolute Free Energies of Binding between RNA and Metal Ions Using Molecular Dynamics Simulations and Continuum Electrostatics.

16. Auffinger, P., L. Bielecki, and E. Westhof. 2003. The Mg2+ Binding Sites of the 5S rRNA Loop E Motif as Investigated by Molecular Dynamics Simulations. Chem. Biol. 10(6):551-561.

17. Kirmizialtin, S., Suzette A. Pabit, Steve P. Meisburger, L. Pollack, and R. Elber. 2012. RNA and Its Ionic Cloud: Solution Scattering Experiments and Atomically Detailed Simulations. Biophys. J. 102(4):819-828.

18. Denesyuk, N. A., and D. Thirumalai. 2015. How do metal ions direct ribozyme folding? Nat. Chem. 7(10):793-801.

19. Ucisik, M. N., P. C. Bevilacqua, and S. Hammes-Schiffer. 2016. Molecular Dynamics Study of Twister Ribozyme: Role of Mg ²⁺ Ions and the Hydrogen-Bonding Network in the Active Site. Biochemistry-US 55(27):3834-3846.

20. Giambasu, G. M., D. A. Case, and D. M. York. 2019. Predicting Site-Binding Modes of Ions and Water to Nucleic Acids Using Molecular Solvation Theory. J. Am. Chem. Soc. 141(6):2435-2445.

21. Cunha, R. A., and G. Bussi. 2017. Unraveling Mg 2+ –RNA binding with atomistic molecular dynamics. RNA 23(5):628-638.

22. Allner, O., L. Nilsson, and A. Villa. 2012. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. J. Chem. Theory Comput. 8(4):1493-1502.

23. Li, P., B. P. Roberts, D. K. Chakravorty, and K. M. Merz, Jr. 2013. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. J. Chem. Theory Comput. 9(6):2733-2748.

24. Hayes, R. L., J. K. Noel, U. Mohanty, P. C. Whitford, S. P. Hennelly, J. N. Onuchic, and K. Y. Sanbonmatsu. 2012. Magnesium Fluctuations Modulate RNA Dynamics in the SAM-I Riboswitch. J. Am. Chem. Soc. 134(29):12043-12053.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

25. Hori, N., N. A. Denesyuk, and D. Thirumalai. 2019. Ion Condensation onto Ribozyme Is Site Specific and Fold Dependent. Biophys. J. 116(12):2400-2410.

26. Hayes, Ryan L., Jeffrey K. Noel, Paul C. Whitford, U. Mohanty, Karissa Y. Sanbonmatsu, and José N. Onuchic. 2014. Reduced Model Captures Mg²⁺-RNA Interaction Free Energy of Riboswitches. Biophys. J. 106(7):1508-1519.

27. Zhang, Y., J. Zhang, and W. Wang. 2011. Atomistic analysis of pseudoknotted RNA unfolding. J. Am. Chem. Soc. 133(18):6882-6885.

28. Panteva, M. T., G. M. Giambasu, and D. M. York. 2015. Force Field for Mg(2+), Mn(2+), Zn(2+), and Cd(2+) Ions That Have Balanced Interactions with Nucleic Acids. J. Phys. Chem. B 119(50):15460-15470.

29. Yoo, J., and A. Aksimentiev. 2012. Improved Parametrization of Li+, Na+, K+, and Mg2+ Ions for All-Atom Molecular Dynamics Simulations of Nucleic Acid Systems. J. Phys. Chem. Let. 3(1):45-50.

30. Casalino, L., G. Palermo, N. Abdurakhmonova, U. Rothlisberger, and A. Magistrato. 2017. Development of Site-Specific Mg2+–RNA Force Field Parameters: A Dream or Reality? Guidelines from Combined Molecular Dynamics and Quantum Mechanics Simulations. J. Chem. Theory Comput. 13(1):340-352.

31. Lakkaraju, S. K., E. P. Raman, W. Yu, and A. D. Mackerell, Jr. 2014. Sampling of Organic Solutes in Aqueous and Heterogeneous Environments Using Oscillating Excess Chemical Potentials in Grand Canonical-like Monte Carlo-Molecular Dynamics Simulations. J. Chem. Theory Comput 10(6):2281-2290.

32. Lemkul, J. A., S. K. Lakkaraju, and A. D. Mackerell, Jr. 2016. Characterization of Mg 2+ Distributions around RNA in Solution. ACS Omega 1(4):680-688.

33. Roth, A., Z. Weinberg, A. G. Y. Chen, P. B. Kim, T. D. Ames, and R. R. Breaker. 2014. A widespread self-cleaving ribozyme class is revealed by bioinformatics. Nat. Chem. Biol. 10(1):56-60.

34. Liu, Y., T. J. Wilson, S. A. McPhee, and D. M. J. Lilley. 2014. Crystal structure and mechanistic investigation of the twister ribozyme. Nat. Chem. Biol. 10(9):739-744.

35. Wilson, T. J., Y. Liu, C. Domnick, S. Kath-Schorr, and D. M. J. J. Lilley. 2016. The Novel Chemical Mechanism of the Twister Ribozyme. J. Am. Chem. Soc. 138(19):6151-6162.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

36. Panja, S., B. Hua, D. Zegarra, T. Ha, and S. A. Woodson. 2017. Metals induce transient folding and activation of the twister ribozyme. Nat. Chem. Biol. 13(10):1109-1114.

37. Gaines, C. S., T. J. Giese, and D. M. York. 2019. Cleaning Up Mechanistic Debris Generated by Twister Ribozymes Using Computational RNA Enzymology. ACS Catal. 9(7):5803-5815.

38. Ren, A., M. Kosutic, K. R. Rajashankar, M. Frener, T. Santner, E. Westhof, R. Micura, and D. J. Patel. 2014. In-line alignment and Mg(2)(+) coordination at the cleavage site of the env22 twister ribozyme. Nat. Commun. 5:5534.

39. Brooks, B. R., C. L. Brooks, A. D. MacKerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. 2009. CHARMM: The biomolecular simulation program. J. Comput. Chem. 30(10):1545-1614.

40. Beglov, D., and B. Roux. 1994. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. J. Chem. Phys. 100(12):9050-9063.

41. Denning, E. J., U. D. Priyakumar, L. Nilsson, and A. D. Mackerell, Jr. 2011. Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. J Comput. Chem. 32(9):1929-1943.

42. Hart, K., N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, and A. D. Mackerell, Jr. 2012. Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. J Chem Theory Comput 8(1):348-362.

43. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79(2):926-935.

44. Durell, S. R., B. R. Brooks, and A. Bennaim. 1994. Solvent-induced forces between 2 hydrophilic groups. J. Phys. Chem. 98(8):2198-2202.

45. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. J Chem Phys 103(19):8577-8593.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

46. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of N-alkanes. Journal of Computational Physics 23(3):327-341.

47. Eastman, P., J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L. P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, and V. S. Pande. 2017. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS computational biology 13(7):e1005659.

48. Izaguirre, J. A., C. R. Sweet, and V. S. Pande. 2010. Multiscale dynamics of macromolecules using normal mode Langevin. Pac. Symp. Biocomput.:240-251.

49. Virnau, P., and M. Müller. 2004. Calculation of free energy through successive umbrella sampling. J. Chem. Phys. 120(23):10925-10930.

50. Tribello, G. A., M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi. 2014. PLUMED 2: New feathers for an old bird. Comput. Phys. Commun. 185(2):604-613.

51. Lakkaraju, S. K., E. P. Raman, W. Yu, and A. D. MacKerell, Jr. 2014. Sampling of Organic Solutes in Aqueous and Heterogeneous Environments Using Oscillating Excess Chemical Potentials in Grand Canonical-like Monte Carlo-Molecular Dynamics Simulations. J Chem Theory Comput 10(6):2281-2290.

52. Sun, D., K. Lakkaraju, S. Jo, and A. D. Mackerell, Jr. 2018. Determination of Ionic Hydration Free Energies with Grand Canonical Monte Carlo/Molecular Dynamics Simulations in Explicit Water. J. Chem. Theory Comput.

53. Raman, E. P., W. Yu, S. K. Lakkaraju, and A. D. MacKerell, Jr. 2013. Inclusion of multiple fragment types in the site identification by ligand competitive saturation (SILCS) approach. J Chem Inf Model 53(12):3384-3398.

54. Pancera, M., T. Zhou, A. Druz, I. S. Georgiev, C. Soto, J. Gorman, J. Huang, P. Acharya, G. Y. Chuang, G. Ofek, G. B. Stewart-Jones, J. Stuckey, R. T. Bailer, M. G. Joyce, M. K. Louder, N. Tumba, Y. Yang, B. Zhang, M. S. Cohen, B. F. Haynes, J. R. Mascola, L. Morris, J. B. Munro, S. C. Blanchard, W. Mothes, M. Connors, and P. D. Kwong. 2014. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature 514(7523):455-461.

55. Hyeon, C., and D. Thirumalai. 2012. Chain length determines the folding rates of RNA. Biophys. J. 102(3):L11-13.

56. Shiman, R., and D. E. Draper. 2000. Stabilization of RNA tertiary structure by monovalent cations11Edited by I. Tinoco. J. Mol. Biol. 302(1):79-91.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

57. Onoa, B., and I. Tinoco, Jr. 2004. RNA folding and unfolding. Curr. Opin. Struc. Biol. 14(3):374-379.

58. Stofer, E., C. Chipot, and R. Lavery. 1999. Free Energy Calculations of Watson−Crick Base Pairing in Aqueous Solution. J. Am. Chem. Soc. 121(41):9503-9508.

59. Turner, D. H., N. Sugimoto, R. Kierzek, and S. D. Dreiker. 1987. Free energy increments for hydrogen bonds in nucleic acid base pairs. J. Am. Chem. Soc. 109(12):3783-3785.

60. Lilley, D. M. 2012. The structure and folding of kink turns in RNA. WIRES RNA 3(6):797-805.

61. Draper, D. E. 2013. Folding of RNA tertiary structure: Linkages between backbone phosphates, ions, and water. Biopolymers 99(12):1105-1113.

62. Bonneau, E., and P. Legault. 2014. NMR Localization of Divalent Cations at the Active Site of the Neurospora VS Ribozyme Provides Insights into RNA–Metal-Ion Interactions. Biochemistry 53(3):579-590.

63. Egli, M., G. Minasov, L. Su, and A. Rich. 2002. Metal ions and flexibility in a viral RNA pseudoknot at atomic resolution. Proceedings of the National Academy of Sciences 99(7):4302.

64. Klein, D. J., T. E. Edwards, and A. R. Ferré-D'Amaré. 2009. Cocrystal structure of a class I preQ1riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. Nature Structural and Molecular Biology 16(3):343-344.

bioRxiv preprint doi: https://doi.org/10.1101/742940; this version posted August 21, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.