Title

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

Running title: Push-Pull Stabilization of RNA by Mg^{2+}

Authors

Abhishek A. Kognole[1] and Alexander D. MacKerell, Jr.[1]*

[1] Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland, USA

* Corresponding author email: alex@outerbanks.umaryland.edu

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 (i.e. pushed apart) while other interactions are stabilized (i.e. pulled together), a balance that stabilizes unfolded states and facilitates the folding 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, Magnesium ions, Ion atmosphere, Grand Canonical Monte Carlo, Umbrella sampling, CHARMM

Statement of Significance

RNAs are biologically and therapeutically of great emerging interest such that it is critical to understand how RNA molecules fold into biologically active 3D 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 conformational and ion 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 (4-6). To assume their tertiary structures RNAs must overcome large unfavorable electrostatic interactions associated with their polyanionic phosphodiester backbone (7). To facilitate this, positively charged ions screen the highly negative potential allowing the RNA secondary structures to collapse into compact tertiary conformations (8-11). Typically divalent ions, most often Mg^{2+}, facilitate the formation of tertiary interactions required for the full folding of RNA (12-14). However, the inability to visualize the ions during folding represents a key barrier to understanding the role of divalent ions in folding of RNA (15).

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 (16-22). 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}) (23), which is beyond the time scale of typical atomistic MD simulations (20, 24) such that only limited

insights into Mg^{2+}-RNA interactions are accessible (25). 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 (19, 26). Reduced models were able to reproduce thermodynamics of Mg^{2+}-RNA interactions but were limited to native states (27). Atomistic simulations of unfolding of a pseudoknot at high temperatures showed diverse intermediate states, although divalent ions were absent in those studies (28). Additionally, scientists have worked on improving the force-field parameters for positively charged metal ions to better simulate their interactions with biomolecules (29-31). Overall, a detailed picture of the interactions of Mg^{2+} in an explicit solvent environment with partially 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 of the ion distribution (32). 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 GCMC approach was recently used to sample the distribution of Mg^{2+} ions in a highlighted study on the mu opioid receptor (33, 34). 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 (35). 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, effectively pulling those phosphates together, through simultaneous interaction between two or more non-sequential phosphates. At the same time, other non-sequential phosphate-phosphate interactions are shown to be destabilized and, therefore, pushed apart, in order to allow for overall stabilization of the tertiary 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 (6, 36-42). 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* (37). 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 stem-loop (37), 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 (39). Results from that study showed Mg^{2+} to facilitate the folding of the RNA although there was ambiguity around whether Mg^{2+} ions are involved in nucleolytic activity. Roth et. al. also showed that Twister is active at very low Mg^{2+} concentrations (43). It should be noted that both experimental studies included 100 mM KCl in the buffer, which likely also contributes to screening of the repulsive interactions potentially allowing Twister to assume the compact intermediate state along with stabilization of secondary interactions (44).

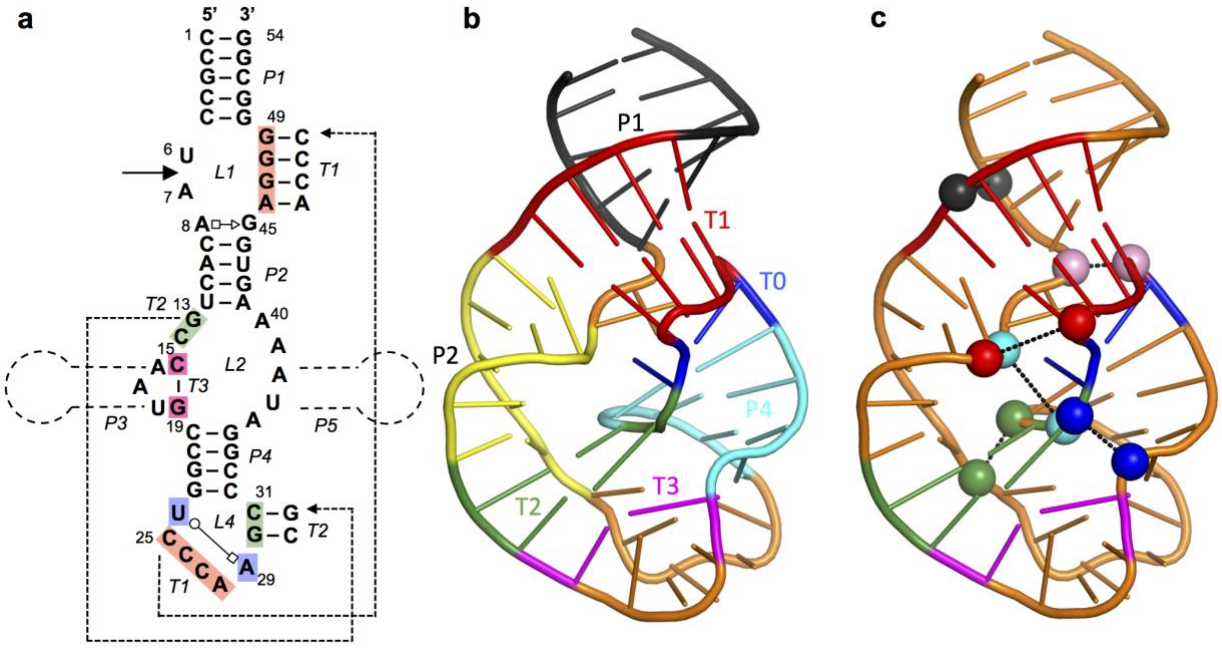


Figure 1. Structure of Twister ribozyme and spatially adjacent phosphate pairs. a) Twister Ribozyme secondary and tertiary contacts. The scissile bond is indicated with a solid arrow and the omitted P3 and P5 loops are indicated. b) Contacts are illustrated as cartoons colored as follows P1 – black, 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 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 (45). In the 20 mM and 100 mM $MgCl_2$ systems all the Mg^{2+} were randomly placed around Twister. In addition to $MgCl_2$, 100 mM KCl was used in all four systems. The systems were first minimized and equilibrated in CHARMM (45) 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 (46-48). Water molecules were modeled by the CHARMM modified-TIP3P force field (49, 50). Smooth Particle Mesh Ewald method was applied for calculation of electrostatic interactions with a real-space cutoff of 12 Å (51). 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 (52). 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 (53, 54) 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 (55). 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 (39) 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 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 (53) 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) (56). 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 simulations 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 (57) 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 oscillating 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) (58, 59). 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. Error analysis of the PMFs and other properties extracted from the PMFs was performed by block averaging over 5 12 ns blocks in each window and calculating the standard error of mean.

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 (60). In this protocol we apply Monte Carlo sampling moves, including insertions, deletions, rotations and translations of the ions and water. The GCMC sampling protocol was re-designed to resample the ion distribution by deleting and inserting Mg^{2+} ions, while K^{+} ions were inserted and deleted simultaneously to maintain the neutral total charge on the system. For systems with no Mg^{2+} ions, the K^{+} and Cl^{-} ions were deleted and inserted together to achieve redistribution of the K^{+} ions. Each single GCMC-MD cycle included seven stages: 1) Run 20000 MC steps for deletion of Mg^{2+} and insertion of K^{+}; 2) Run 80000 MC steps for rotation and translation of all ions and water; 3) Repeat steps 1 and 2 until only one Mg^{2+} is left and the system is neutral; 4) Run 20000 MC steps for insertion of Mg^{2+} and deletion of K^{+}; 5) Run 80000 MC steps for rotation and translation of all ions and water; 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, followed by equilibrated for 100 ps of NVT MD followed by the 10 ns of NPT MD in OpenMM (53). 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 S4 (60).

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 of the PMF we analyzed 10000 frames and all frames were first aligned with respect to the backbone of the ribozyme in the starting conformation for that window. 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 (35) and elsewhere (61).

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. Conformational sampling was performed using a computationally accessible 1-D RC in which tertiary interaction are lost allowing the sampling for partially unfolded states of the RNA to investigate their interactions with Mg^{2+}. 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 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 (37). 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 (39). 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.

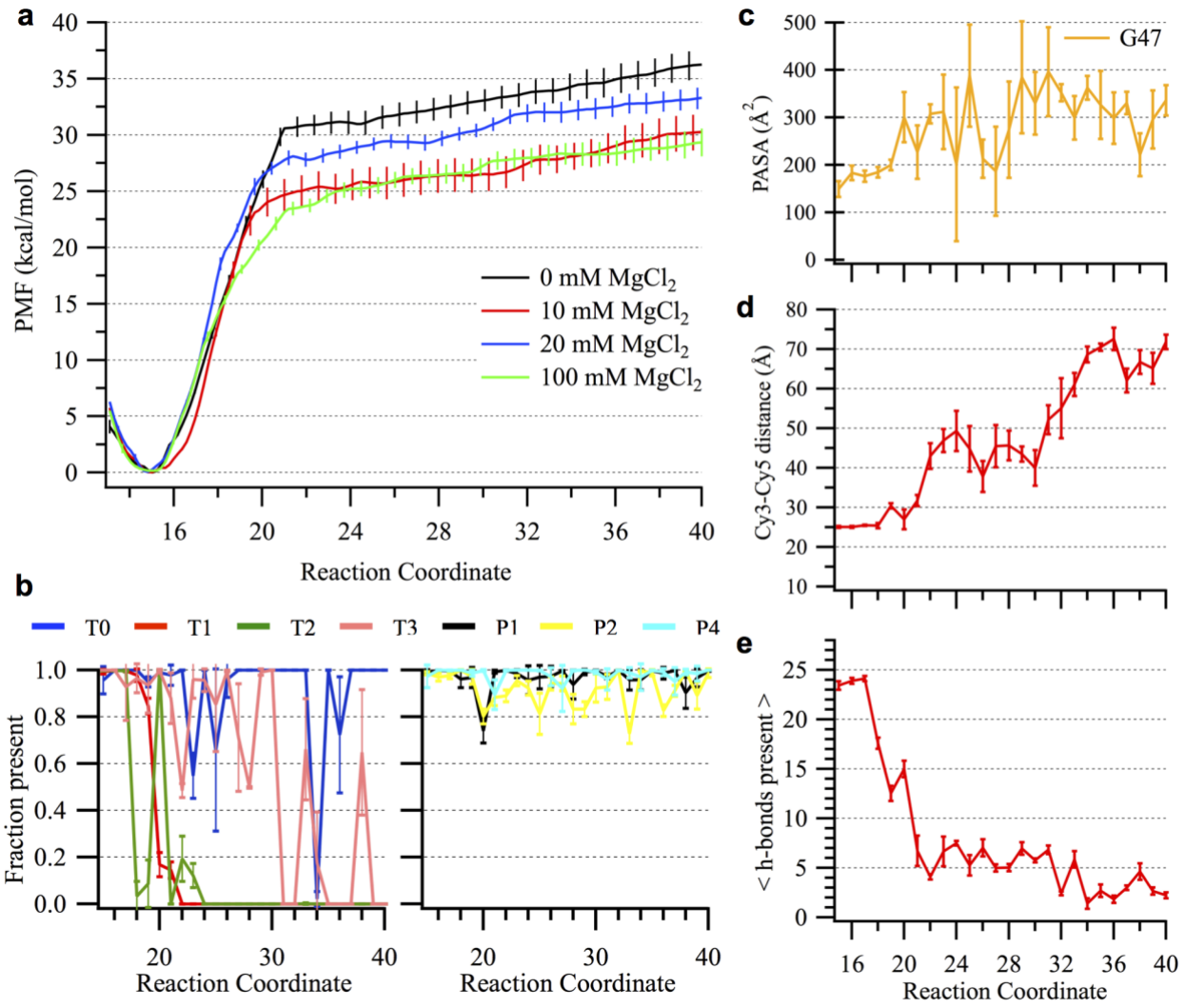


Figure 2. Analysis of Twister unfolding simulations. (a) Potential of mean force (PMF) of Twister ribozyme at different concentrations of $MgCl_2$ based on the GCMC/MD simulations. The reaction coordinate (RC) corresponds to center of mass distance between two groups of residues that contribute to the 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 error 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 for 100 mM $MgCl_2$ system. A contact is present if the distance between the nucleotides is within 3 standard deviations of mean distance for the same contact 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 nucleotide, G47, protected from cleavage by RNase T1 calculated at each

window along the reaction coordinate. PASA is calculated as the solvent accessible surface area using a probe radius of 10 Å as previously described (80). (d) Interfluorophore distance (IFD) based on distance between C4' atoms of Ura 24 and Gua 54 along the reaction coordinate for Twister. (e) Number of Watson-Crick hydrogen bonds present between nucleotides involved in tertiary interactions along the reaction coordinate. A hydrogen bond was defined by donor-acceptor distance cutoff of 3.0 Å and donor-hydrogen-acceptor angle cutoff of 120˚. For (b), (c), (d) and (e) the data are from the 100mM $MgCl_2$ system with results for the other three concentrations provided in the Supporting Material (Figure S2 to S6).

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. In the remainder of the manuscript analysis focuses on interpretation of the PMFs in terms of analysis of molecular details of the Mg^{2+}-RNA interactions with emphasis on differences between the 0 and 100 mM $MgCl_2$ systems supported by results from the 10 and 20 mM $MgCl_2$ systems. While monotonic differences going from to 0 to 100 mM were anticipated, given the limited number of ions in the 10 and 20 mM systems, 5 and 10, respectively, the overall convergence of those systems may require additional sampling. To better understand the results with respect to experimental studies based on FRET and RNase T1 footprinting (39), analysis of the PMFs was undertaken specifically targeting the molecular phenomena directly related to those experiments.

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 (39). 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 for each window in the PMFs at all four concentrations (Supporting Figure S4). Shown in Figure 2d are the values of the “Interfluorophore distance (IFD)” along the RC at 100 mM $MgCl_2$. 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, including the region of ssDNA linking Cy5 to the RNA (Supporting Figure S4), 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. A similar interpretation of the experimental results on unfolding based on RNase T1 footprinting experiments may be made (39, 62). Shown in Figure 2c are calculations at 100 mM $MgCl_2$ for protein accessible surface areas (PASA) of the guanine base, G47, in Twister known to undergo hydrolysis (39). The individual plots for the other guanine bases at all four $MgCl_2$ concentrations are shown in Supporting Figure S5. While the accessibilities of different nucleotides vary, the values for the individual nucleotides do not change significantly to well beyond the inflection point at 20 Å in the PMF. The G47 base is involved in formation of T1 interaction and significant increase in PASAs for this base are also observed after the inflection point at 20 Å. 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 (RC = 15-20 Å), but rather associated with events beyond that region that correspond with more global unfolding events (RC = 20-40 Å). 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(39). The presence of small populations of the alternate states is 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. Consistent with this model are β parameter results, $\beta_{fold} = 0.37$ and $\beta_{unfold} = 0.72$, indicating that the transition to the fully unfolded state observed in the FRET and footprinting experiments occurs closer to the unfolded state (39).

Further support for the experiments monitoring unfolding events beyond those obtained in the present calculations may be obtained from analysis of folding rates. The experimental FRET studies indicate that the folding rate of Twister is under 0.1 s^{-1} (39). Estimates of the folding rate based on number of nucleotides (N) for Twister, N = 54, is ~1200 s^{-1}(63). From the present study, based on an energetic barrier of ~20 kcal/mol determined from the PMF (Figure 2a) from the minima (Folded state, F) to the inflection

point (Intermediate state, I, representing an initial unfolded state) an estimate of the equilibrium constant, K_{eq}, for $F \rightleftharpoons I$ based on the vant Hoff equation yields 3.8×10^{-15}. Next, applying an Arrehnius model to calculate the unfolding rate associated with I (64, 65), $k_{F\rightarrow I} = k_o \exp\left(-\frac{E_a}{RT}\right)$, where E_a = 20 kcal/mol, and assigning a pre-exponential value, k_o, of 1.3×10^{11} s^{-1} associated with a harmonic well at the minima of the PMFs, an unfolding rate $k_{F\rightarrow I}$ of 4.9×10^{-4} s^{-1} is calculated. The folding rate $k_{I\rightarrow F}$ may then be obtained from $K_{eq} = k_{F\rightarrow I}/k_{I\rightarrow F}$ yielding 1.3×10^{11} s^{-1}. Thus, both the folding rate estimate based on chain length and on the PMF assuming the first intermediate at the inflection point being the unfolded state are orders of magnitude larger than the experimental estimate of <0.1 s^{-1}. These differences further suggest that the FRET and Footprinting experiments are monitoring unfolding events that are occurring in the region beyond the inflection point at ~20 Å in the PMF that may be associated with the complexity of a double pseudoknot like Twister that differs from that of typical RNAs. In this scenario the increased population of the folded state observed in the FRET experiments in the presence of Mg^{2+} is more consistent with the free energy of the 100 mM $MgCl_2$ system being ~7 kcal/mol more favorable than 0 mM $MgCl_2$. Furthermore, beyond the inflection region at ~20 Å in the PMF there is no indication of a transition to a further unfolded state and analysis of sampled conformations from the 100 mM $MgCl_2$ PMF in Figure 3 show that 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. Similar results are obtained at the other three concentrations (Figure S7). Thus, the present results indicate that the loss of all or some of the secondary structural interactions may lead to the unfolded states observed experimentally.

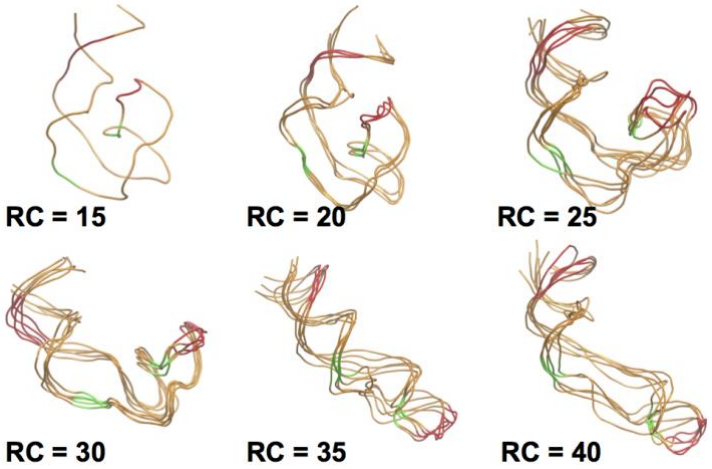


Figure 3. Central conformations of top 5 clusters from 100mM $MgCl_2$ PMF simulations at various values of the reaction coordinate. Backbone RMSD clustering analysis was performed with a cutoff of 2.5 Å from which the of top 5 clusters of conformations at each window were identified.

Analysis of fraction present of the tertiary and secondary contacts (Figure 1) at 100 mM MgCl2 system as a function of the RC is shown in Figure 2b, with results for the individual systems shown in Supporting Figure S2 and S3. 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 including those in the P1 region, are largely maintained throughout the PMFs (Supporting Figure S3), consistent with a report that secondary structures of RNA are independently stable of the tertiary interactions (66). 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.

Vušurovic et. al. present results showing that the crucial T1 interaction, whose formation must precede the P1 stem closing, shows reversible dynamics at physiological Mg^{2+} concentrations (42). However, they did not comment on formation or unfolding of the T2 tertiary interactions. Overall, both our sampling pathway and the experimental observations from Vušurovic et. al. are in agreement that formation of T1 is one of the last critical steps in folding which ultimately activates self-cleavage of Twister. The importance of P1 formation for folding is still a topic that needs further attention as the Vušurovic study points out that even after self cleavage and loss of the P1 duplex, the compact pseudoknot tertiary structure is maintained (42).

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. At 100 mM $MgCl_2$, 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). Similar changes in number of hydrogen bonds are observed at all other systems (Supporting Figure S6). Considering that each WC hydrogen bond corresponds to a stabilization free energy ranging from -0.7 to -2 kcal/mol (67, 68) 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 hydrogen bonds associated with the T1 and T2 tertiary interactions. Analysis of traces of the phosphodiester backbone of Twister at various stages along all four PMFs shows that at RC = 20 and 25 Å the overall 3D

conformation of the RNA is maintained (Figure 3 and Supporting Figures S7). 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 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. This unfolding is hypothesized to involve a transition from conformations being sampled beyond the inflection point at 20 Å in the PMFs and a more globally unfolded state. Accordingly, the free energy difference of ~7 kcal/mol between the partially unfolded states in the PMF at 0 and 100 mM Mg^{2+} represent the effective folded state being observed in the experimental studies, consistent with favoring of the sampling of folded state in the presence of Mg^{2+} (39, 42).

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

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 native and intermediate 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 previously performed (35). Figure 4 presents the distribution of Mg^{2+} around Twister from the RC = 15 Å PMF windows in the presence of 100, 20 and 10 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 4c 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 such that clear separation of

the 0 and 100 mM PMFs are evident with the 10 and 20 mM falling between those curves. Additional studies are required to address this issue.

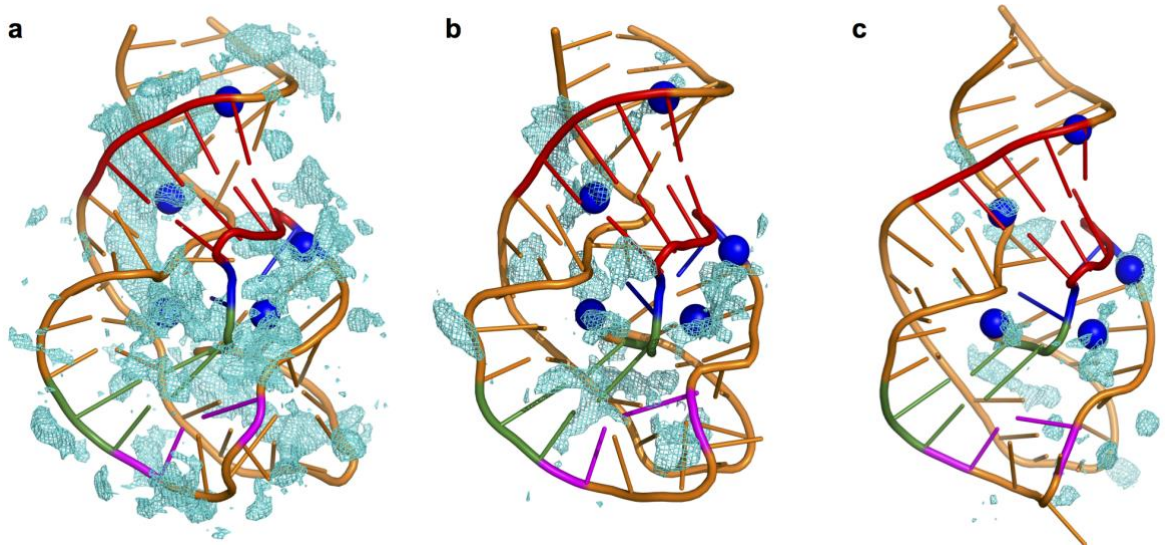


Figure 4. Occupancy maps for Mg^{2+} ion distribution around native conformation of the Twister ribozyme from the GCMC-MD simulations. a) 100 mM $MgCl_2$ b) 20 mM $MgCl_2$ and c) 10 mM $MgCl_2$. The cutoff for GFE is -2 kcal/mol. The crystallographic positions of Mg^{2+} binding sites are shown in blue 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 5). The cutoff of 9 Å is defined based on radial distribution functions (RDF) of the distance between NBPO and Mg^{2+} ions (Supporting Figure S8) 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 (9). 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 5). Analysis of Figure 5 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 regions of non-adjacent NBPO interactions that occur in the folded state that are identified in the top-left panel of Figure 5. 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.

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

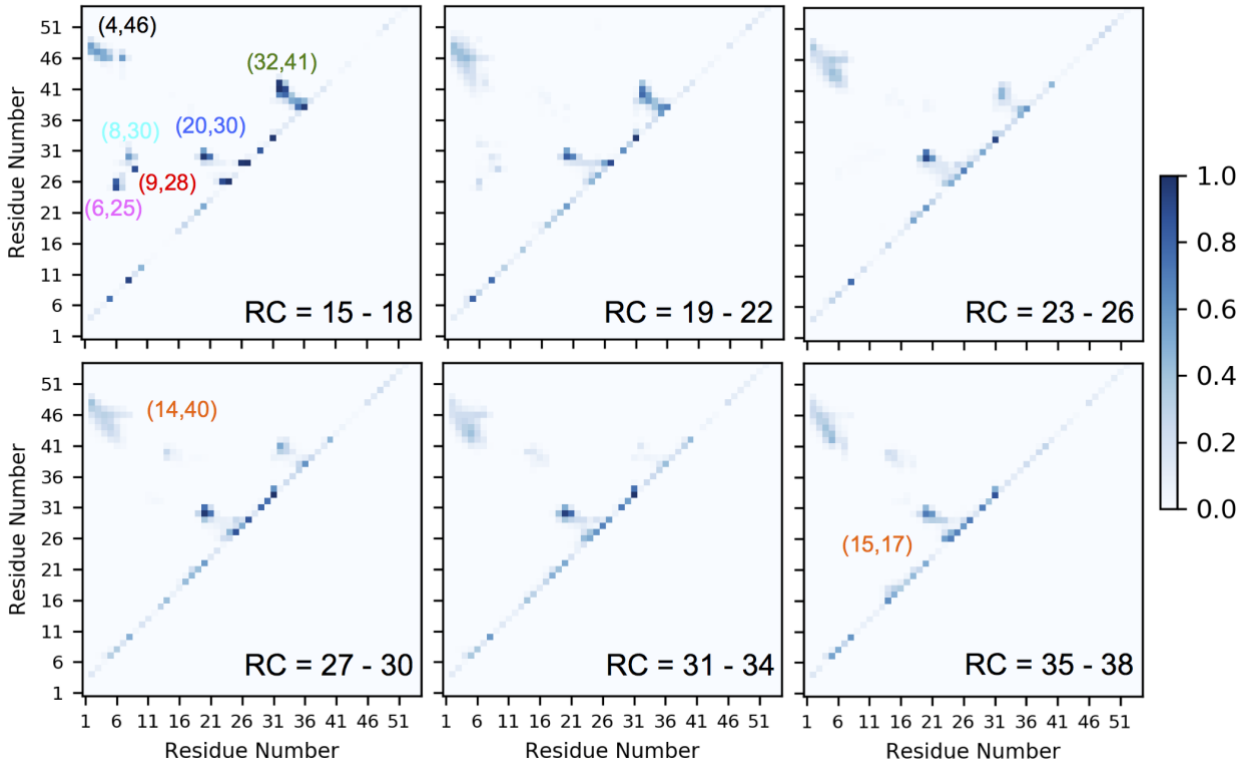


Figure 5. NBPO probability matrices from NBPO distance analysis for Twister from the GCMC-MD PMF at 100 mM $MgCl_2$. The probabilities were averaged over groups of 4 RC windows. RC = 15 – 18 corresponds to the fully folded structure. The interactions discussed in the text have been specified with circles.

Figure 6 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 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 5. 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 (69).

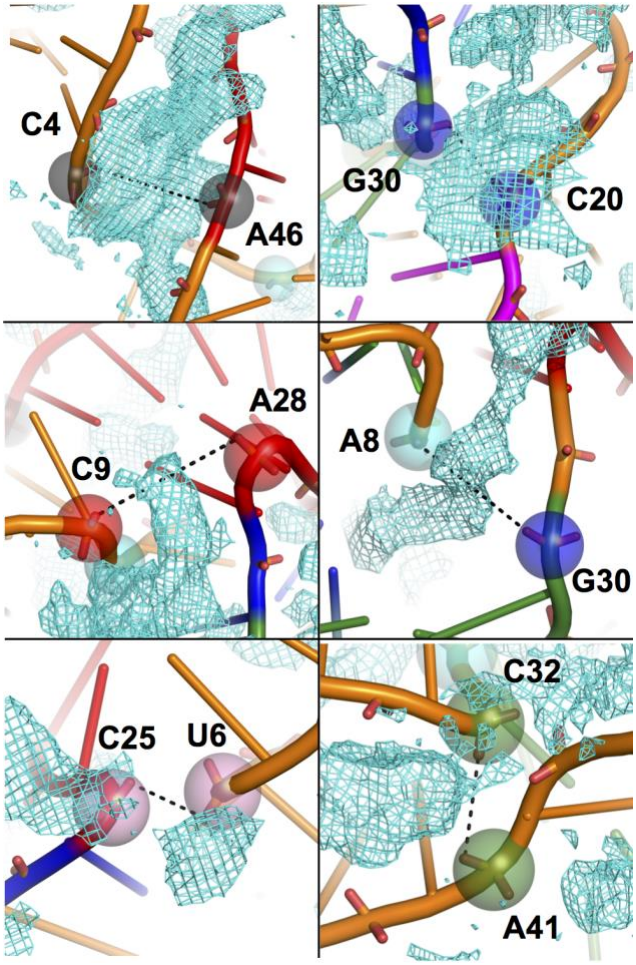


Figure 6. Distribution of Mg^{2+} ions around the NBPOs of residue pairs. From the RC = 15 Å native state of Twister ribozyme (cartoon) from the 100 mM $MgCl_2$ PMF. The phosphates are represented as same-colored transparent spheres. GFE maps are illustrated as cyan mesh at -2 kcal/mol cutoff.

The ability of bidentate chelation of Mg^{2+} ions with sequential phosphate groups has been studied quantum mechanically.(70) Intuitively, a bidentate chelation interaction at the A7-A8 pair would be expected to play a role in self-cleavage of the phosphodiester bond between U6-A7. However stereospecific phosphorothioate substitution studies of the A7 phosphate have shown that inner sphere coordination with divalent ion is not required for catalysis.(36, 38, 40, 71, 72). Interestingly, the GFE maps in the native state suggest that Mg^{2+} ions highly favor coordination between A8-G30 and A8-C31 and there is minimal Mg^{2+} sampling around the A7-A8 phosphate pair (Figure S9) showing that bidentate chelation of this sequential pair is not occurring in the present study, consistent with the experimental observations.

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 NBPO pairs at long RC values include the 15-17 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 S10). In the absence of Mg^{2+} ions, near the folded state (RC = 15 to 18) the pattern of NBPO pairs 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 quantify 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 7 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 5) 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 (19, 26).

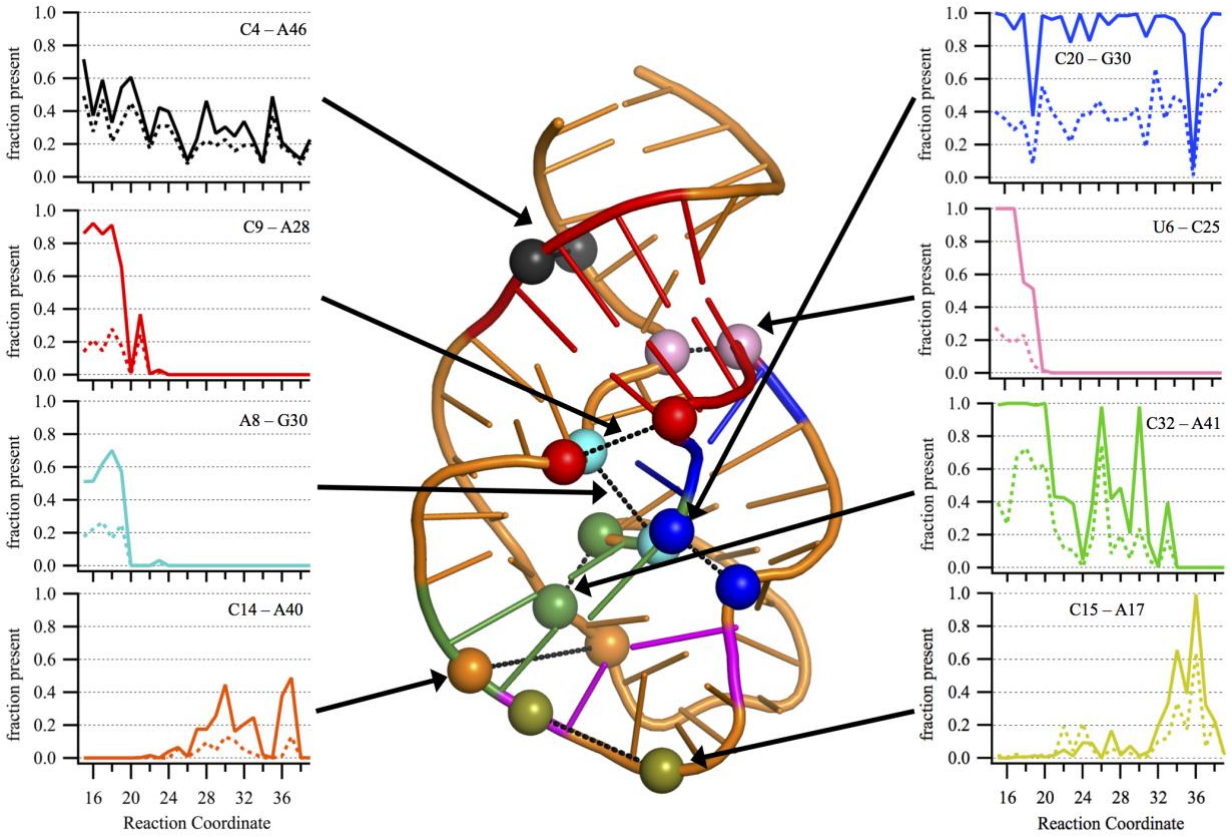


Figure 7. Adjacent NBPO pairs in contact with Mg^{2+} ions at 100 mM $MgCl_2$. The probability of different pairs of nucleotides with NBPOs within 9 Å (Solid line) and the probability of a Mg^{2+} ion within 6.5 Å of NBPOs of both phosphates coordinating the two NBPOs (Dashed line) along the reaction 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 S12.

Towards better understanding the role of NBPO pairs in the partially unfolded RNA both experimental and computational studies are suggested. Experimentally, phosphorothioate substitutions at these specific pairs could be performed followed by monitoring of the energetics of unfolded states using FRET or footprinting. Both the 14-40 and 15-17 pair interactions with Mg^{2+} include a significant number of direct interactions (Supporting Figure S11), such that it would be anticipated that the impact of Mg^{2+} on the sampling of the unfolded state would be decreased. Computationally, a repulsive potential between the specific phosphate groups and Mg^{2+} could be applied and the PMF recomputed. This would disallow stabilization of the specific NBPO pairs thereby altering the energetics of the unfolded states in the presence of Mg^{2+}.

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 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 7 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.

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 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 (8, 73). 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 5 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. 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 6) indicating that despite the presence of Mg^{2+} the increased sampling of other NBPO pairs occurs at the expense of these interactions. It is this balance of interactions that we refer to as the push-pull mechanism for the stabilization of RNA by Mg^{2+}, where push-pull refers to the presence of Mg2+ pulling some phosphate pairs together while simultaneously leading to other NBPO pairs being pushed apart with the overall outcome being the stabilization of the folded RNA structure despite some individual NBPO pairs being destabilized.

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 5. 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 S1 and standard error in the differences are shown in Supporting Table S2.

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

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 S3). 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 S13. 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 by “pulling” the RNA together while simultaneously leading to lowered sampling of short 9-28 and 6-25 NBPO pairs by “pushing” them apart in the inflection region of the PMF where the tertiary contacts are initially forming. This effect is suggested to allow for the T1 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 (19, 26). Notably, in the present study it is shown that Mg^{2+} leads to overall 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. Interestingly, Mg^{2+} is not seen in the P1 region of Twister in the crystallographic structures 4OJI and 4RGE (37, 41), while the FRET studies on both sequences show conflicting results on the importance of this stem in folding and activity as recently discussed and addressed by Gaines et. al. (39, 40, 42).

Other individual NBPO pairs are stabilized or destabilized to varying extents consistent with Mg^{2+} facilitating RNA folding and stability in a sequence specific fashion. Decreased 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 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 effectively being pushed apart to facilitate pulling the overall structure together yielding 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.

The deep global minima observed in the present PMFs, which are largely insensitive to ion concentration, may explain the ability of Twister to self cleave at very low Mg^{2+} concentrations and that sampling of the folded state required for catalysis may occur in all conditions. We recognize that the role of the background of 100 mM KCl (74), included in the present study to be consistent with the experimental study, requires further attention to understand the role of monovalent ions in RNA folding. We have provided analyses indicating that 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. Accordingly, the present RC used to define the folding landscape is not accessing the fully unfolded conformations. However, the 1-D RC selected for this study has allowed investigation of the impact of Mg^{2+} on compact intermediates and partially unfolded states using a novel GCMC ion sampling approach. Additional studies are needed to address the issue of choosing the best RC to investigate the full unfolding profile of Twister and recent advances in machine learning may help solve this problem (75). Overall, the study implemented a novel GCMC approach to investigate the highly complex problem of RNA folding and provides new insights on the impact of Mg^{2+} ions on both the folded and unfolded states of the Twister Ribozyme.

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.

Supporting Citations

References (76-79) 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. Gaines, C. S., and D. M. York. 2016. Ribozyme Catalysis with a Twist: Active State of the Twister Ribozyme in Solution Predicted from Molecular Simulation. J. Am. Chem. Soc. 138(9):3058-3065.

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

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

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

10. 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.

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

12. 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.

13. 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.

14. 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.

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

16. 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.

17. 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.

18. 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.

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

20. 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.

21. 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.

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

23. 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.

24. 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.

25. 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.

26. 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.

27. 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.

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

29. 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.

30. 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.

31. 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.

32. 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.

33. Hu, X., D. Provasi, S. Ramsey, and M. Filizola. 2019. Mechanism of μ-Opioid Receptor-Magnesium Interaction and Positive Allosteric Modulation DOI:10.1016/j.bpj.2019.10.007. Biophys. J.:In press.

34. MacKerell, A. D. 2019. Ions Everywhere? Mg2+ in the μ-Opioid GPCR and Atomic Details of Their Impact on Function. Biophysical Journal.

35. 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.

36. 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.

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. Vusurovic, N., R. B. Altman, D. S. Terry, R. Micura, and S. C. Blanchard. 2017. Pseudoknot Formation Seeds the Twister Ribozyme Cleavage Reaction Coordinate. J. Am. Chem. Soc. 139(24):8186-8193.

43. 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. Nature Chemical Biology 10(1):56-60.

44. Draper, D. E. 2008. RNA folding: Thermodynamic and molecular descriptions of the roles of ions. Biophysical Journal 95(12):5489-5495.

45. 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.

46. 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.

47. 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.

48. 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.

49. 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.

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

51. 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.

52. 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.

53. 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.

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

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

56. 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.

57. 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.

58. Kumar, S., J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. 1995. Multidimensional free-energy calculations using the weighted histogram analysis method. J. Comput. Chem. 16(11):1339-1350.

59. Grossfield, A. WHAM: the weighted histogram analysis method , version 2.0.9. In http://membrane.urmc.rochester.edu/content/wham.

60. 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.

61. 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.

62. MacKerell, A. D., R. Rigler, L. Nilsson, U. Hahn, and W. Saenger. 1987. Protein dynamics: A time-resolved fluorescence, energetic and molecular dynamics study of ribonuclease T1. Biophysical Chemistry 26(2):247-261.

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

64. Laidler, K. J. 1987. Chemical kinetics. Harper & Row, New York.

65. Arrhenius, S. 1889. Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte. In Zeitschrift für Physikalische Chemie. 96.

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

67. 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.

68. 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.

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

70. Petrov, A. S., J. C. Bowman, S. C. Harvey, and L. D. Williams. 2011. Bidentate RNA-magnesium clamps: on the origin of the special role of magnesium in RNA folding. RNA 17(2):291-297.

71. Košutic, M., S. Neuner, A. Ren, S. Flür, C. Wunderlich, E. Mairhofer, N. Vušurovic, J. Seikowski, K. Breuker, C. Höbartner, D. J. Patel, C. Kreutz, and R. Micura. 2015. A Mini-Twister Variant and Impact of Residues/Cations on the Phosphodiester Cleavage of this Ribozyme Class. Angewandte Chemie - International Edition.

72. Breaker, R. R. 2017. Mechanistic Debris Generated by Twister Ribozymes. ACS chemical biology 12(4):886-891.

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

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

75. Wang, Y., J. M. L. Ribeiro, and P. Tiwary. 2019. Machine learning approaches for analyzing and enhancing molecular dynamics simulations. https://arxiv.org/abs/1909.11748. (preprint posted September 25, 2019).

76. 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.

77. 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.

78. 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.

79. Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. 1995. A smooth particle mesh Ewald method. The Journal of Chemical Physics 103(19):8577-8593.

80. 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.

SUPPORTING MATERIAL

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

Abhishek A. Kognole[1] and Alexander D. MacKerell, Jr.[1]*

[1] Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland, USA

* Corresponding author email: alex@outerbanks.umaryland.edu

SUPPORTING FIGURES

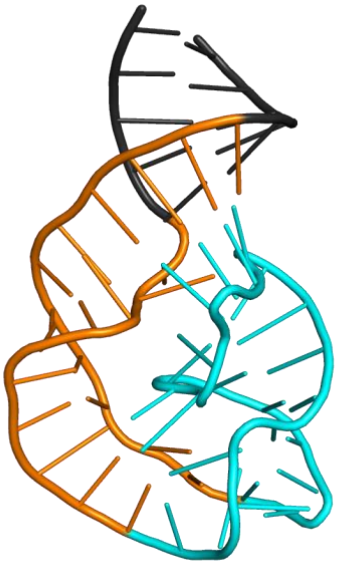


Figure S1. Reaction coordinate was calculated as the distance between the centers of mass of the non-hydrogen atoms in group1 (orange cartoon, nucleotides 6:15 + 40:49) and group 2 (cyan cartoon, residues 16:39).

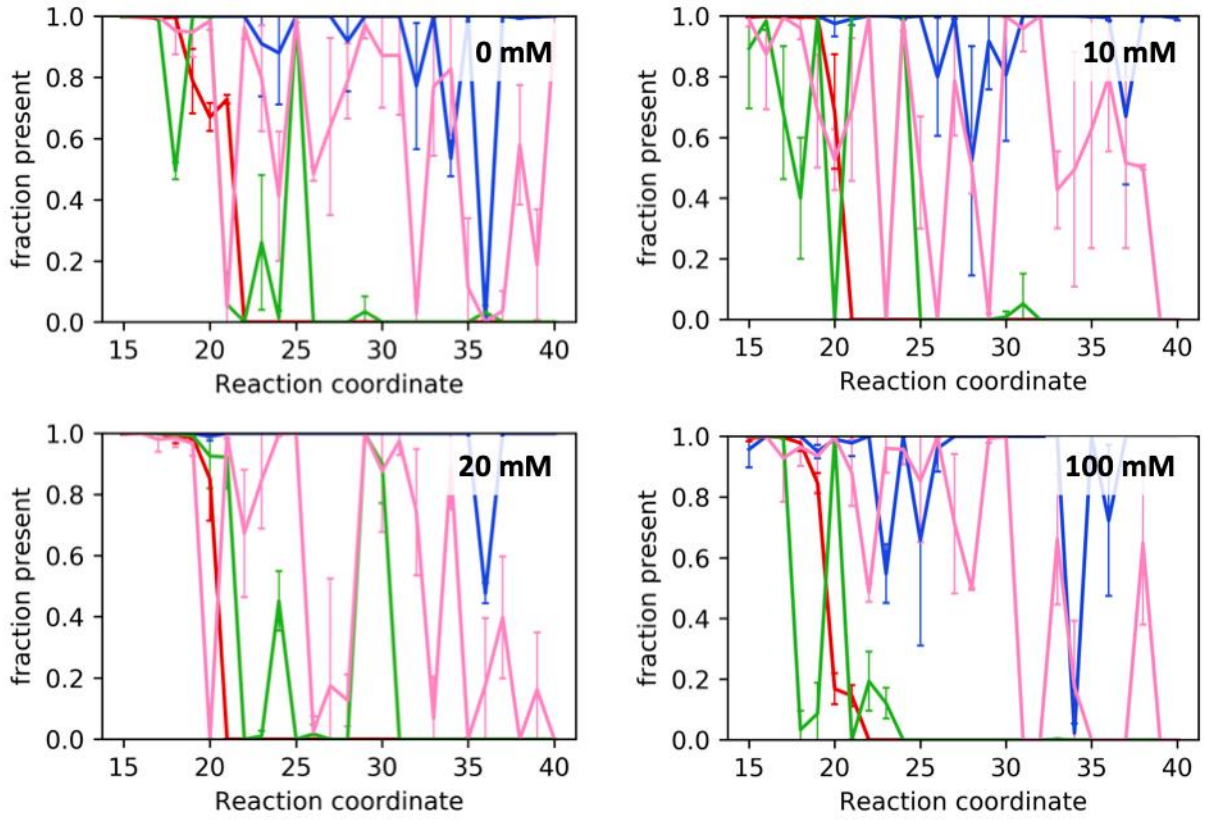


Figure S2. Fraction of tertiary contacts present along the reaction coordinate for the four different $MgCl_2$ concentrations. T1 – red, T2 – green, T3 – pink, T0 – blue.

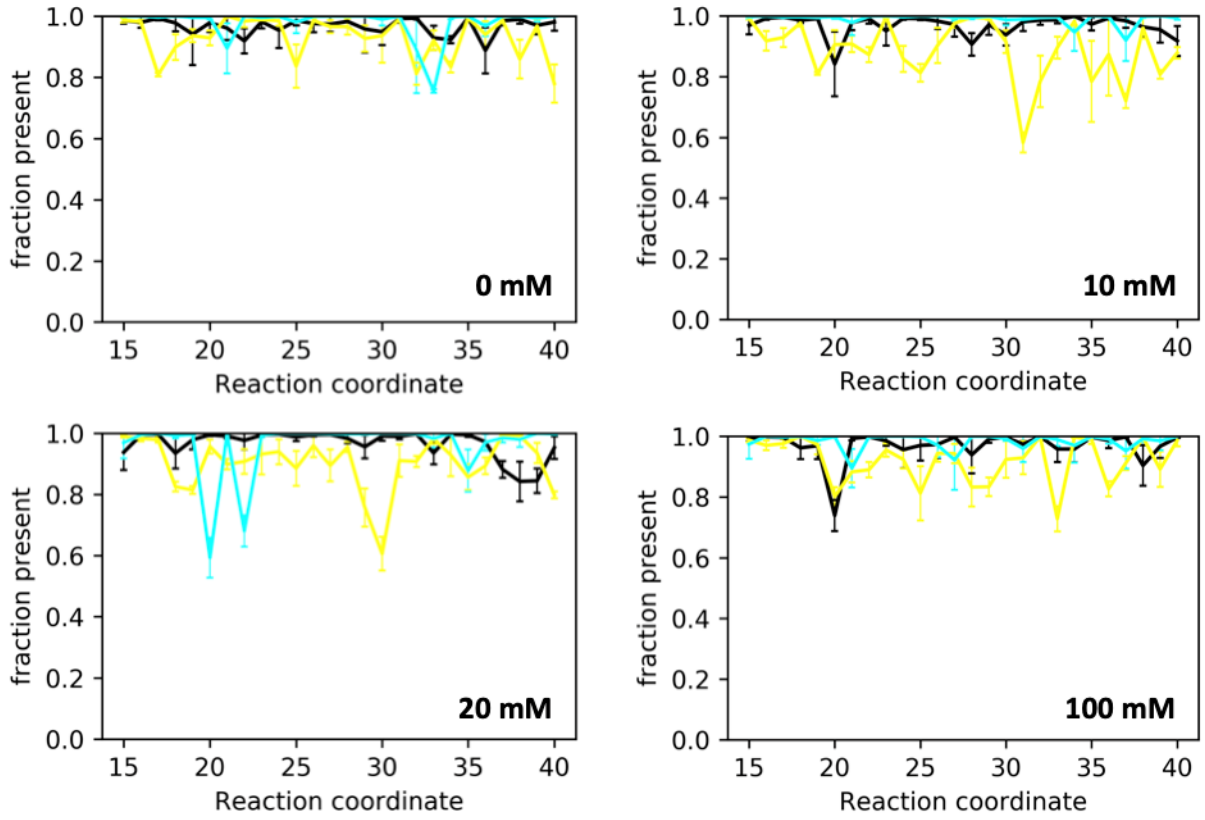


Figure S3. Fraction of secondary contacts present along the reaction coordinate for the four different $MgCl_2$ concentrations. P1 – black, P2 – yellow, P4 – cyan.

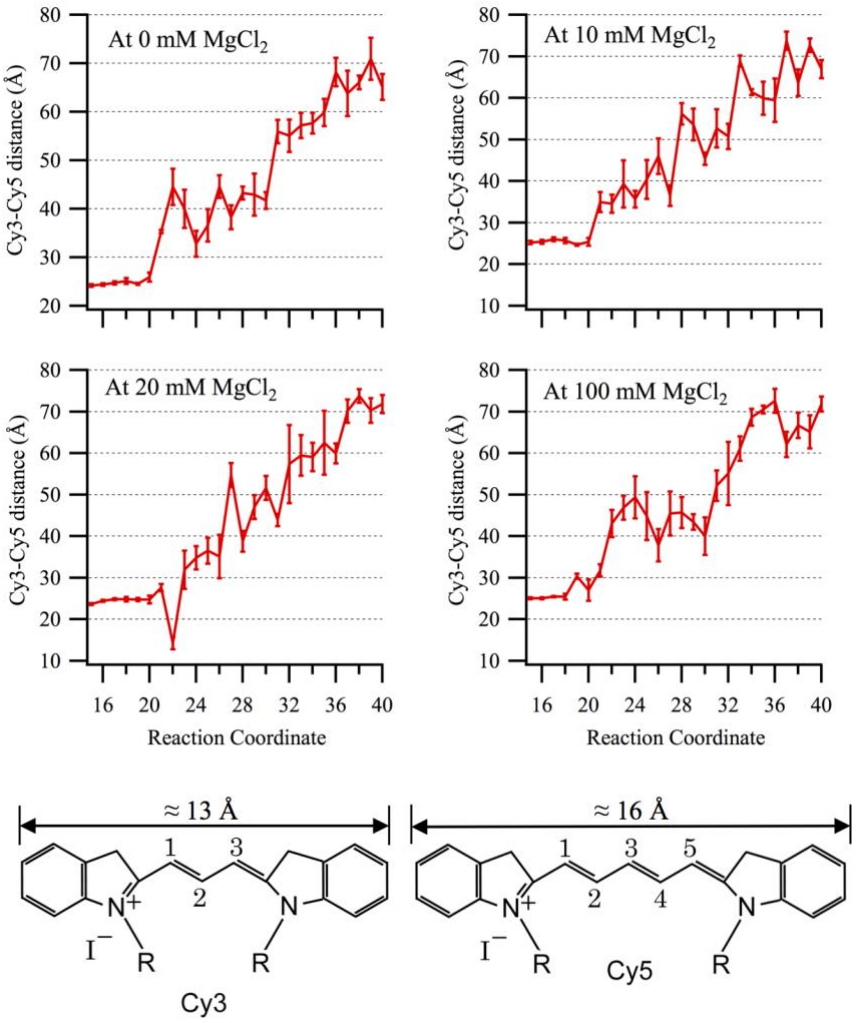


Figure S4. Interfluorophore distances based on the distance between C4' atoms of Ura 24 and Gua 54 along the reaction coordinate for Twister in 0, 10, 20 and 100mM $MgCl_2$. Bottom panels – Structures of core components of Cy3 and Cy5 fluorophores with approximate lengths, Cy3 ≈ 13 Å and Cy5 ≈ 16 Å. The backbone length of ssDNA with 5 nucleotides is approximately 18 Å.

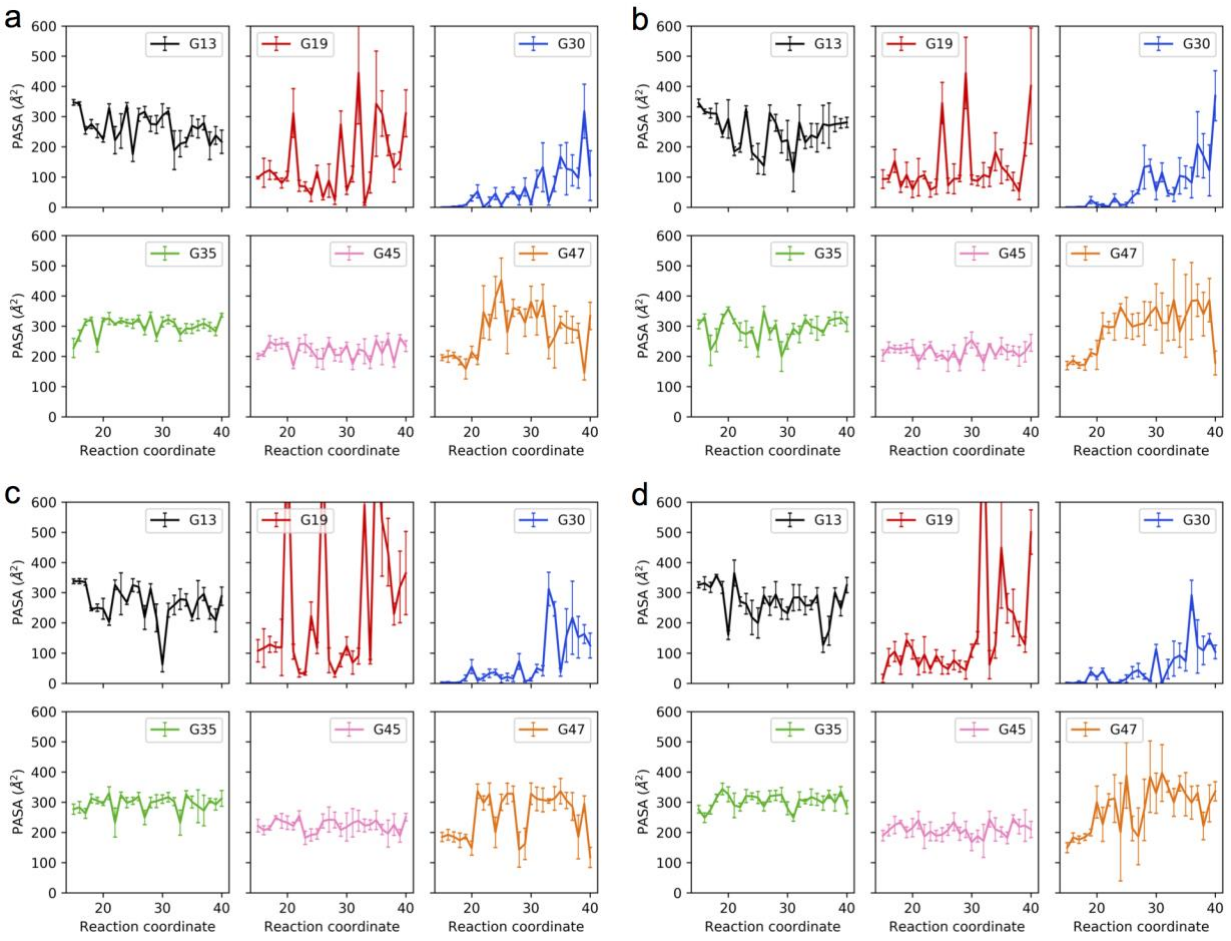


Figure S5: Protein accessible surface area (PASA) for the guanine nucleotides protected from cleavage by RNase T1 calculated at each window along the reaction coordinate for a) 0 mM b) 10 mM c) 20 mM and d) 100 mM $MgCl_2$ systems. PASA is calculated as the solvent accessible surface area using a probe radius of 10 Å.

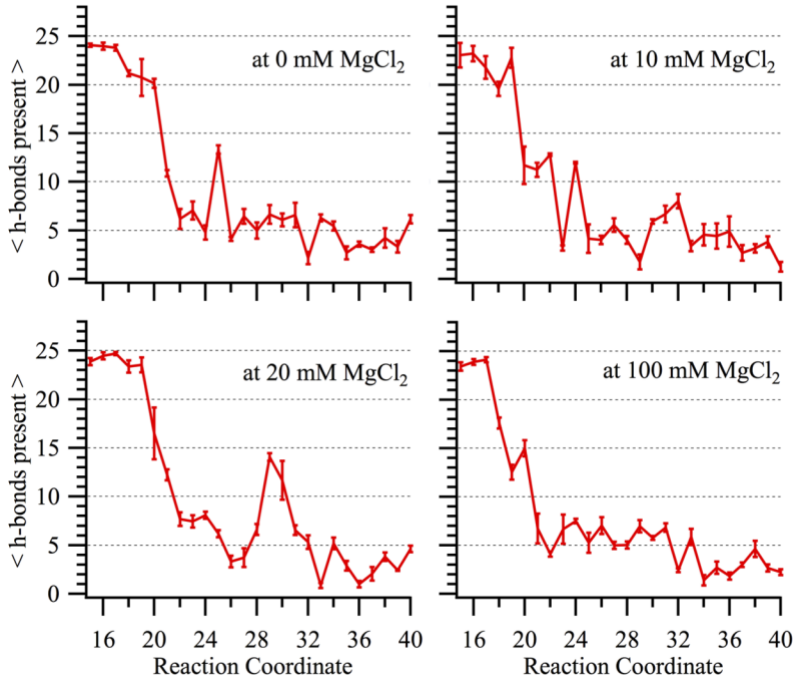
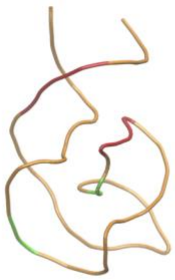
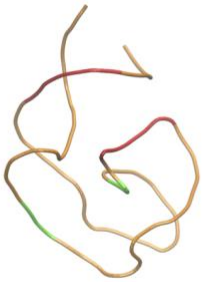


Figure S6: Number of Watson-Crick hydrogen bonds present between nucleotides involved in tertiary interactions along the reaction coordinate at four different $MgCl_2$ concentrations. A hydrogen bond was defined by donor-acceptor distance cutoff of 3.0 Å and donor-hydrogen-acceptor angle cutoff of 120˚.

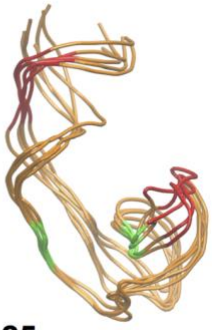
a



RC = 15

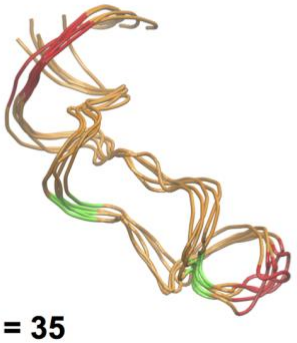


RC = 20

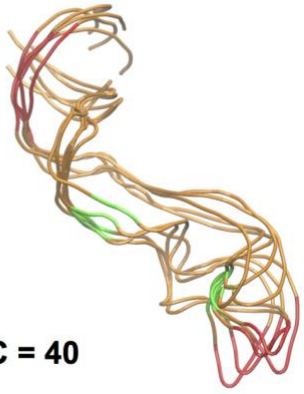


RC = 25

RC = 30

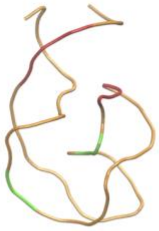


RC = 35

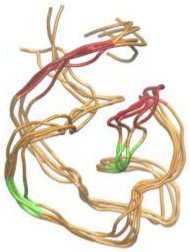


RC = 40

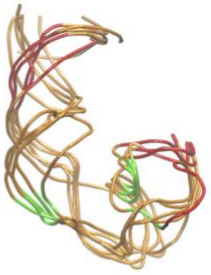
b



RC = 15

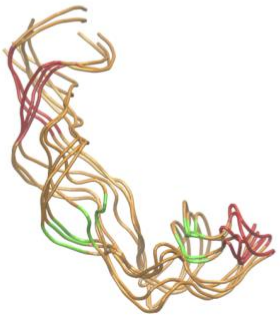


RC = 20



RC = 25

RC = 30



RC = 35

RC = 40

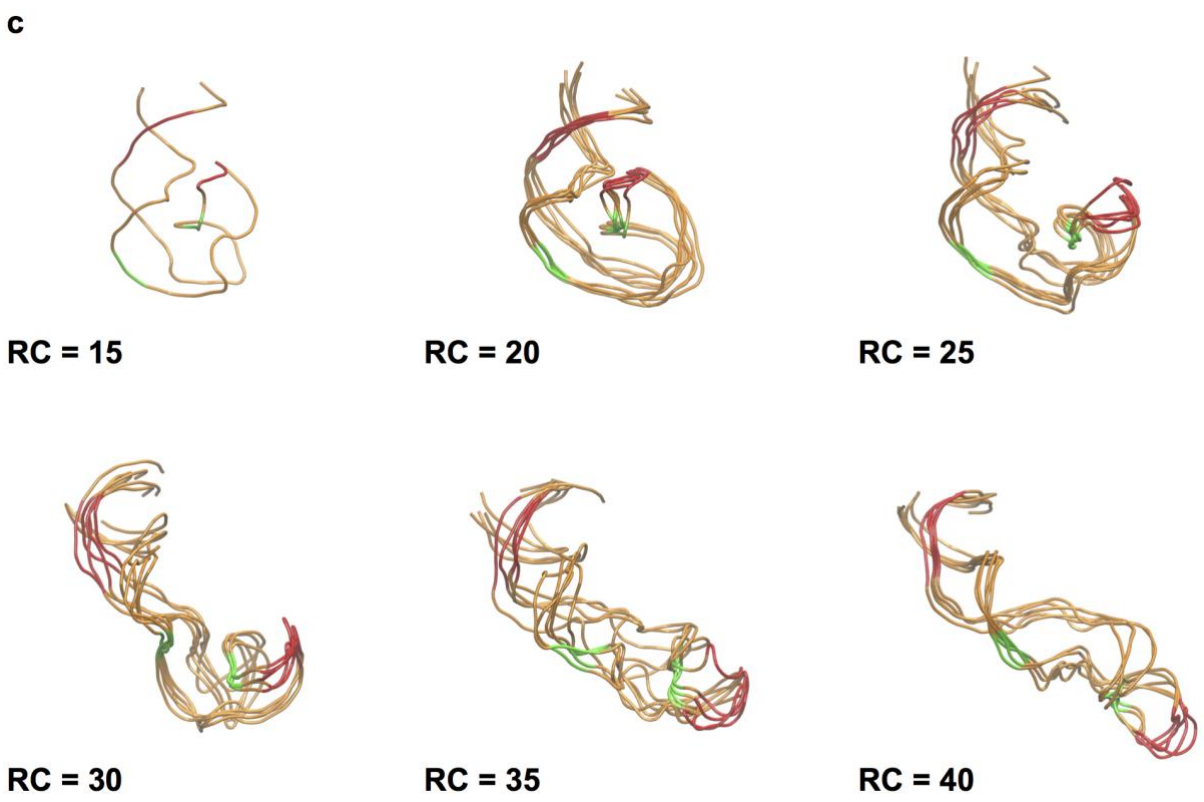


Figure S7. Central conformations of top 5 clusters from (a) 0mM (b) 10 mM and (c) 20 mM $MgCl_2$ simulations at various values of the reaction coordinate. RMSD clustering analysis was performed with a cutoff of 2.5 Å from which the of top 5 clusters of conformations at each window were identified.

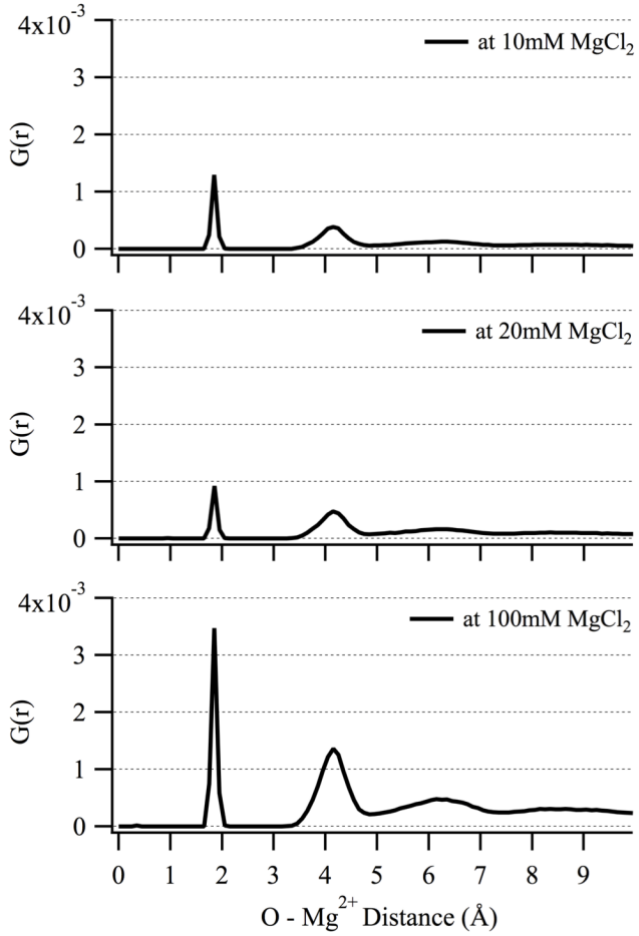


Figure S8. RDF of Mg^{2+} ions around non-bridging phosphate oxygens (NBPO) in the 10, 20 and 100 mM $MgCl_2$ systems. A 9 Å distance between NBPOs was used to define those NBPO pairs that could be simultaneously interacting with Mg^{2+} through outer shell interactions.

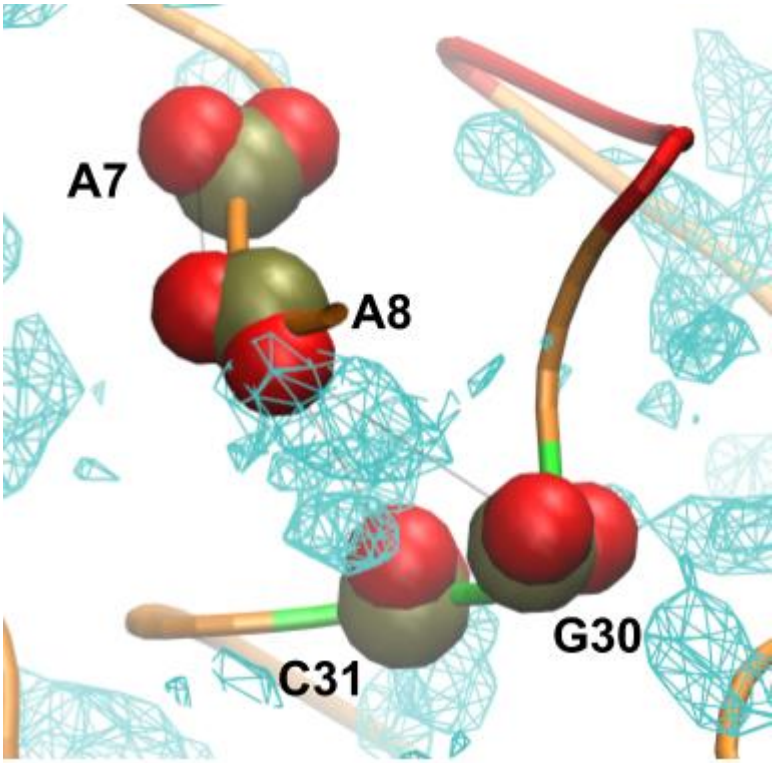


Figure S9. GFE map of Mg^{2+} (cyan wireframe) at 100 mM $MgCl_2$ at RC = 15.0 showing that Mg^{2+} ions highly favor coordination between A8-G30 and A8-C31 and there is minimal Mg^{2+} sampling around the A7-A8 phosphate pair. GFE level = -2.5 kcal/mol.

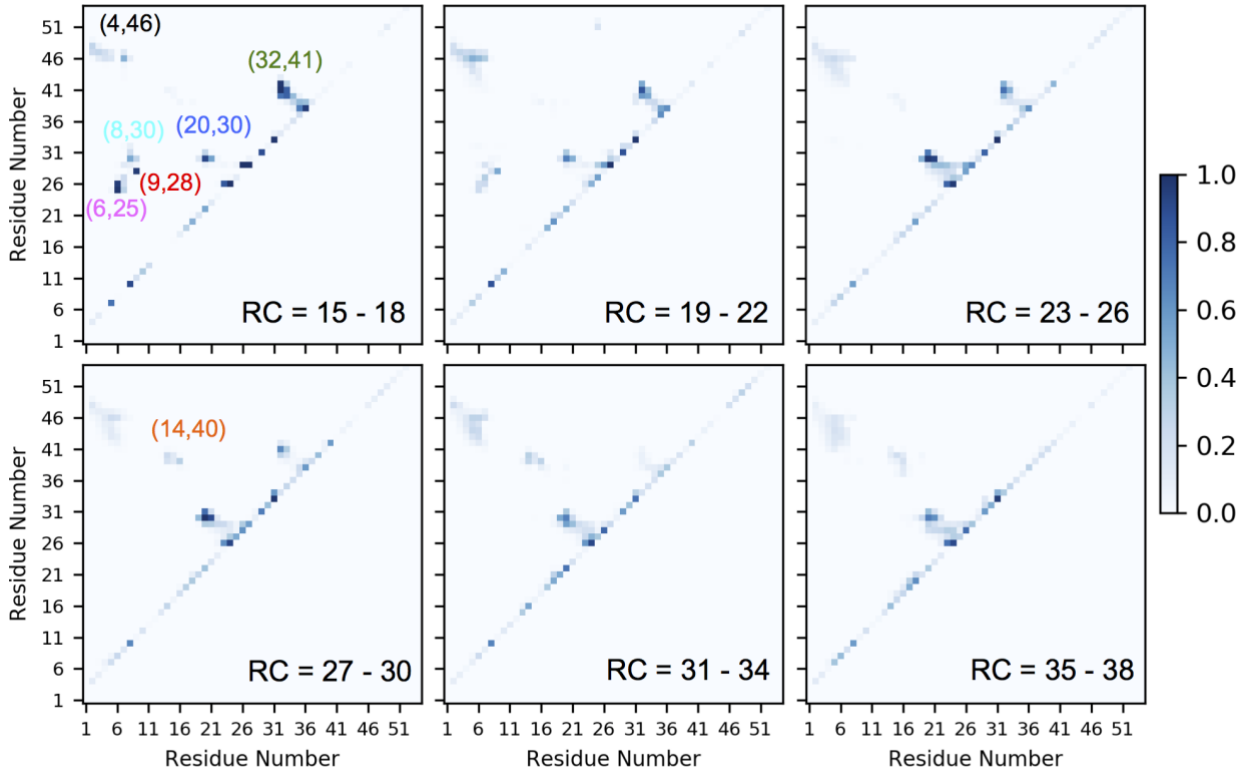


Figure S10. NBPO probability matrices from NBPO distance analysis for Twister from the GCMC-MD PMF at 0 mM $MgCl_2$. The probabilities were averaged over groups of 4 RC windows. RC = 15 – 18 corresponds to the fully folded structure.

PP-14-40

OSD	-	52.64 %
ISD	0.00 %	47.36 %
	ISD	OSD

PP-15-17

OSD	-	82.20 %
ISD	0.00 %	17.80 %
	ISD	OSD

Figure S11. Distances from Mg^{2+} were measured when NBPOs from the two corresponding phosphate pairs, PP-14-40 and PP-15-17, were within 4.5 Å of an Mg^{2+} for the windows at RC = 34 to 37. Distances were then classified into ISD (Inner shell) and OSD (outer shell) types of interactions to find the probabilities of combination of types. Significant percentage inner shell interactions were observed in case of both the pairs.

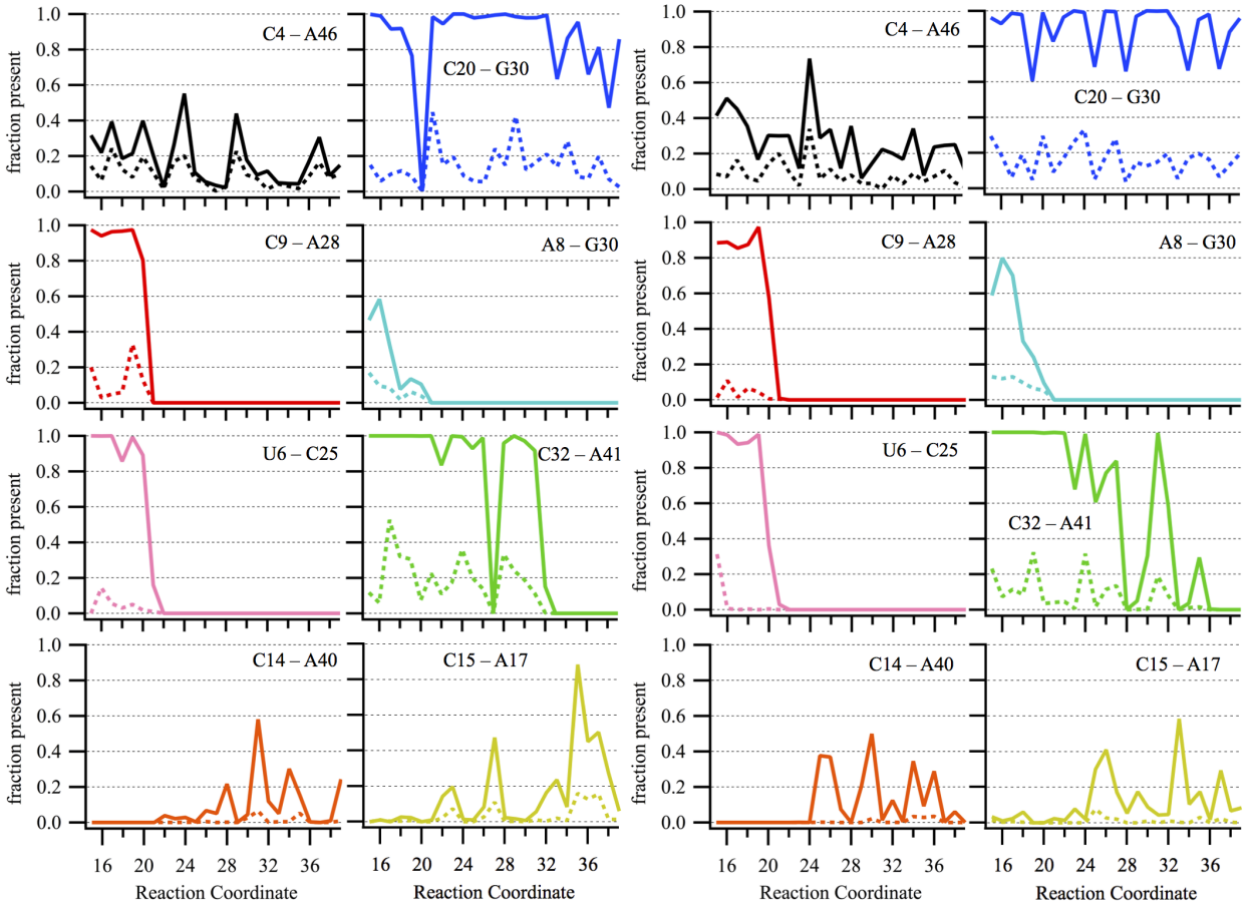


Figure S12. The probability of different pairs of nucleotides with NBPOs within 9 Å and the probability of a Mg^{2+} ion coordinating the two NBPOs along the reaction coordinate. Left two panels – 20mM $MgCl_2$ Right two panels – 10mM $MgCl_2$

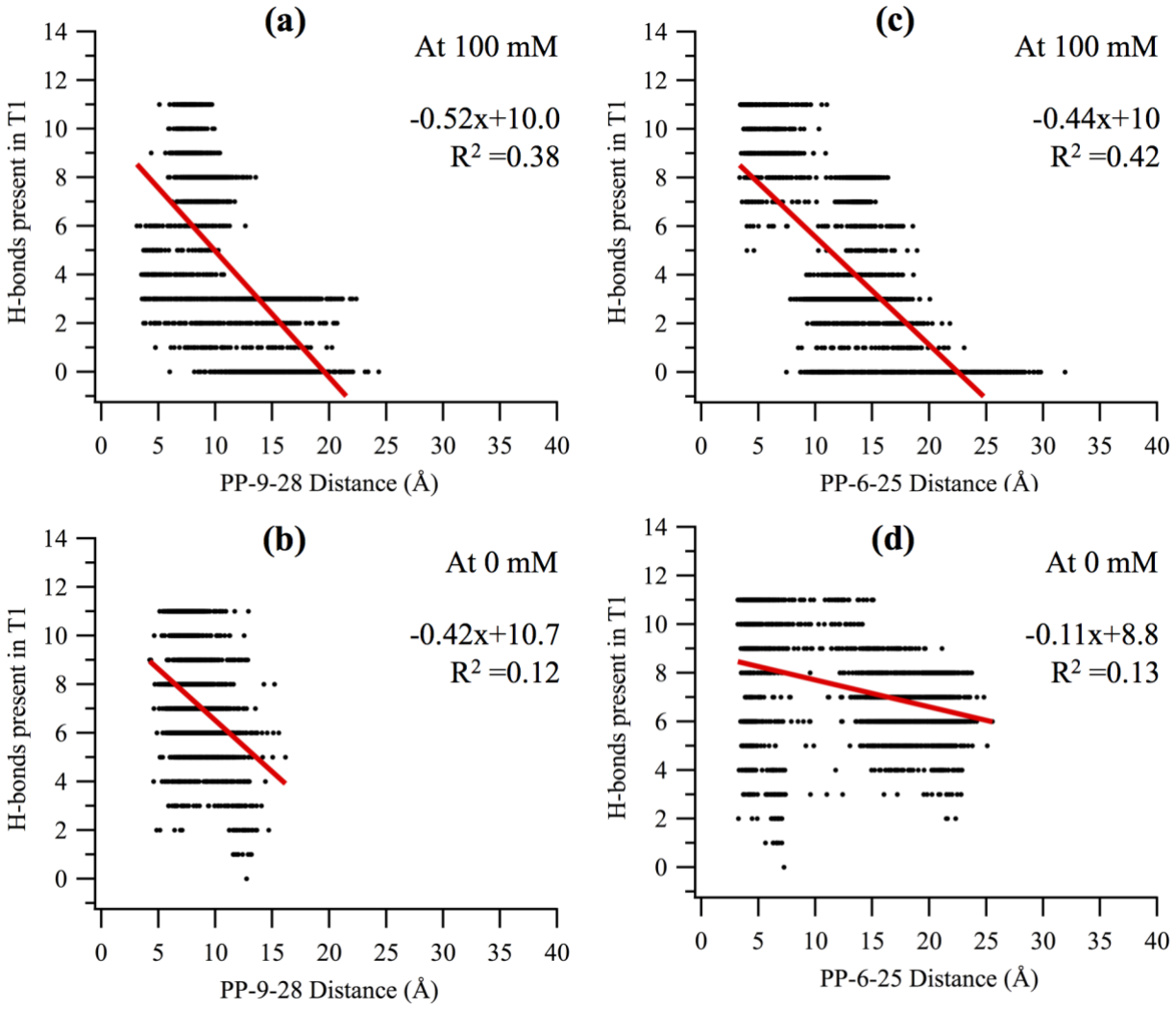


Figure S13. Correlation plots of the minimum NBPO distance versus the number of T1 tertiary contact WC hydrogen bonds for the pp-9-28 (a & b) and pp-6-25 (c & d) NBPO pairs. Data is from the RC = 19 to 21 windows of the PMF for the 100 mM and 0 mM $MgCl_2$ PMFs.

SUPPORTING TABLES

Table S1. Summed probability of selected regions of non-sequential NBPO pairs from the 0 and 100 mM $MgCl_2$ PMFs. Selected regions defining the NBPO pairs are listed in section 3 of the table. The values for the adjacent to the diagonal interactions (Figure 6, termed Near Diagonal) are based on the total summed probabilities (Table 2) minus the sum of the selected NBPO pair regions shown in this table.

100 mM									
RC	**pp9-28**	**pp6-25**	**pp8-30**	**pp32-41**	**pp4-46**	**pp20-30**	**pp14-40**	**pp15-17**	**Near Diagonal**
15 to 18	0.887	2.438	1.607	7.672	5.484	2.982	0.016	0.202	13.525
19 to 22	0.259	0.441	0.515	5.894	7.348	2.984	0.082	0.235	15.421
23 to 26	0.007	0.013	0.038	2.980	7.043	3.724	0.132	0.432	15.006
27 to 30	0	0	0.001	2.408	5.818	3.812	0.671	0.578	16.233
31 to 34	0	0	0	0.589	6.671	4.051	0.694	0.903	15.75
35 to 38	0	0	0	0.001	6.034	3.248	1.01	2.194	15.186
0 mM									
RC	**pp9-28**	**pp6-25**	**pp8-30**	**pp32-41**	**pp4-46**	**pp20-30**	**pp14-40**	**pp15-17**	**Diagonal**
15 to 18	0.962	2.833	1.409	6.914	2.490	2.055	0.144	0.092	12.963
19 to 22	0.520	1.147	0.335	4.569	4.613	2.033	0.082	0.254	14.525
23 to 26	0	0.001	0.014	3.634	3.397	4.897	0.248	0.559	15.752
27 to 30	0	0	0.001	2.155	2.276	4.524	1.073	0.541	15.123
31 to 34	0	0	0	0.695	3.832	3.655	1.419	0.661	15.155
35 to 38	0	0	0	0.004	2.988	3.944	0.988	1.095	16.144
Regions defining the NBPO pairs									

pp9-28 = P(9,28)	pp4-46 = P(2:8,41:51)
pp6-25 = P(5:8,23:27)	pp20-30 = P(18:22,28,32)
pp8-30 = P(6:9,29,32)	pp14-40 = P(12:16,38:42)
pp32-41 = P(31:35,38:43)	pp15-17 = P(14:17,16:18)

Table S2. Standard errors of difference between the summed probabilities at 100 and at 0 mM $MgCl_2$ for the individual non-sequential NBPO pairs for the results reported in Table 3. The standard errors were calculated over five 12-ns segments from the 60-ns sampling at the corresponding group of windows.

RC	pp9-28	pp6-25	pp8-30	pp32-41	pp4-46	pp20-30	pp14-40	pp15-17
15 to 18	0.02	0.39	0.18	0.59	0.21	0.30	0.15	0.06
19 to 22	0.04	0.18	0.08	0.72	0.51	0.49	0.05	0.14
23 to 26	0.01	0.02	0.04	0.40	0.90	0.40	0.16	0.18
27 to 30	0.00	0.00	0.00	1.13	0.68	0.36	0.27	0.13
31 to 34	0.00	0.00	0.00	0.27	1.29	0.30	0.20	0.24
35 to 38	0.00	0.00	0.00	0.00	1.12	0.37	0.73	0.14

Table S3. 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 6 over RC = 19 to 22 Å. The RC=22 window was excluded as the number of T1 hydrogen bonds was zero; exclusion of the RC=22 window led to the difference between the 100 and 0 mM 9-28 and 6-25 probabilities being -0.35 and -0.95, respectively, which are of larger magnitude than the values of -0.26 and -0.71 presented in Table 3 for RC = 19 to 22.

RC=19 to 21	**Total**	**pp9-28**	**pp6-25**	**pp8-30**	**pp32-41**	**pp4-46**	**pp20-30**	**pp14-40**	**pp15-17**	**Near Diagonal**
At 100mM	33.55	0.34	0.58	0.61	6.76	7.07	2.79	0.10	0.20	15.09
At 0mM	29.40	0.69	1.53	0.42	5.79	5.28	1.18	0.01	0.18	14.32
Diff (100-0)	4.16	-0.35	-0.95	0.19	0.97	1.79	1.62	0.09	0.02	0.77

Table S4. Hydration free energies of the ions and water and the minimum and maximum values of μ_{ex} used in the PME/GCMC protocol.

Fragment	Hydration FE (kcal/mol)	μ_{ex}^{min} (kcal/mol)	μ_{ex}^{max} (kcal/mol)

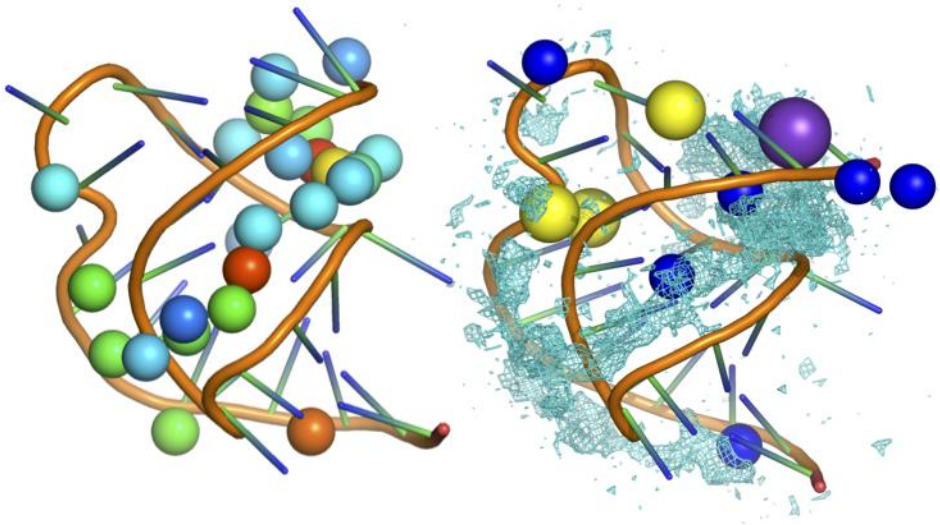
Mg^{2+}	-437.38	-837.38	-37.38
K^{+}	-70.51	-170.51	30.51
Cl^{-}	-81.26	-81.26	-81.26
Water	-5.60	-5.60	-5.60

Supporting Text 1. Validation of the PME GCMC-MD protocol

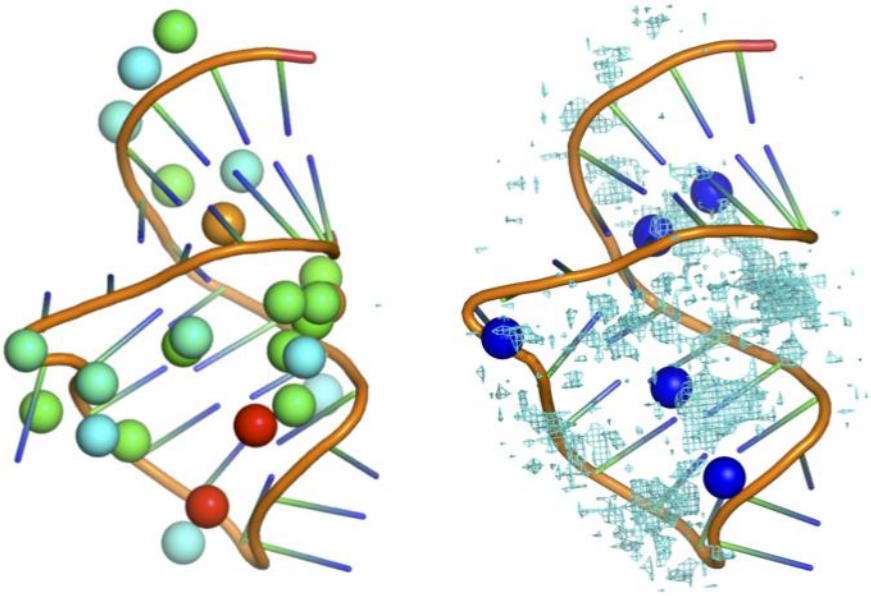
To validate the oscillating μ_{ex} GCMC Mg^{2+} ion sampling protocol we calculated the Mg^{2+} ion distribution in four RNA structures, VS ribozyme (*PDB 2MIS*) (1), BWYV pseudoknot (*PDB 1L2X*) (2), Twister ribozyme from *O. Sativa* (*PDB 4OJI*) (3), and preQ1 riboswitch (*PDB 3FU2*) (4). The former two were previously analyzed by our group for characterization of Mg^{2+} distributions around native conformations using a similar oscillating μ_{ex} GCMC-MD based approach (5). However, that study applied a non-bond truncation scheme for the electrostatic interactions. In the present work, a recently published extension of the oscillating μ_{ex} GCMC method (6) that includes particle mesh Ewald (PME) (7) for the treatment of long-range electrostatics was applied in conjunction with a new μ_{ex} oscillation scheme for the excess chemical potential to sample the distribution of Mg^{2+} around the RNA. To prepare the systems for validation of the new PME GCMC-MD protocol, all divalent ions and waters present in the crystal structures were removed. Other monovalent ions (Na^{+} and K^{+}) were retained in their experimental positions. For the VS ribozyme stem-loop VI solution NMR structure, the first model (out of 20 in the NMR ensemble) was used following the protocol from our previous study (5). For Twister ribozyme, two missing nucleotides in the crystal structure were added by using the internal coordinates in CHARMM (8). A stepwise restrained minimization of the coordinates was performed to relax the local conformation of the two nucleotides. For the preQ1 riboswitch the PRF ligand was removed. All four RNAs were solvated in a 65 Å cubic waterbox with 50 mM $MgCl_2$ and 100 mM KCl. The average size of the systems was ~31500 atoms. The systems were minimized and equilibrated with restraints on the backbone atoms using CHARMM (8). The same conditions were used for the equilibration MD simulations as described in the methods. The final conformations were used to start the GCMC-MD validation runs where the backbone atoms were harmonically restrained during the MD part to preserve

the native conformation. The validation runs involved 4 replicates of 50 cycles with 5 ns MD run per cycle. In total 1000 ns of MD simulation was performed for each RNA.

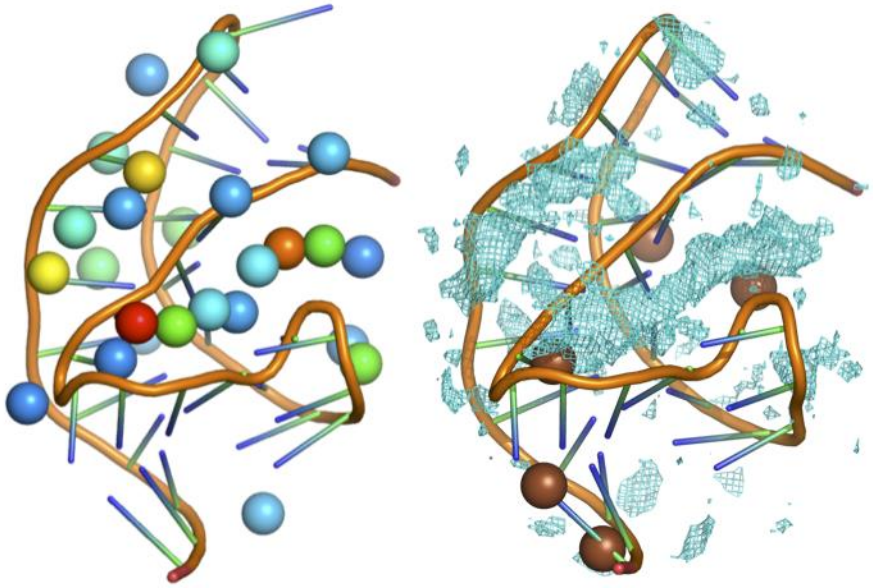
The validation simulations on the 4 RNAs show that the PME-based GCMC approach for sampling the ions around the RNA is more effective than the previous protocol (5). Analysis involved the distribution of Mg^{2+} ions over all 2000 frames from the 4 replicates × 50 cycles performed on each of the 4 systems. Following alignment of the frames with respect to the RNA backbone atoms of the experimental structures, clustering of the Mg^{2+} ion positions around the RNA was performed based on RMSD of ion positions. Figure T1 shows the clusters of Mg^{2+} ions colored with respect to their cluster size. Compared to the respective X-ray or NMR structures (1-4), almost all the experimental Mg^{2+} binding sites have been successfully sampled and represented by clusters with significant number of members. Table T1 shows the distance to the nearest cluster for crystallographic positions of Mg^{2+} for the Twister ribozyme. In addition, the Mg^{2+} positions from the GCMC-MD runs with Twister are well distributed with respect to inner-shell, outer-shell and weakly bound positions (Figure T2). This indicates that the PME GCMC-MD protocol is capable and efficient in finding binding sites associated with a distribution of Mg^{2+} coordination types.



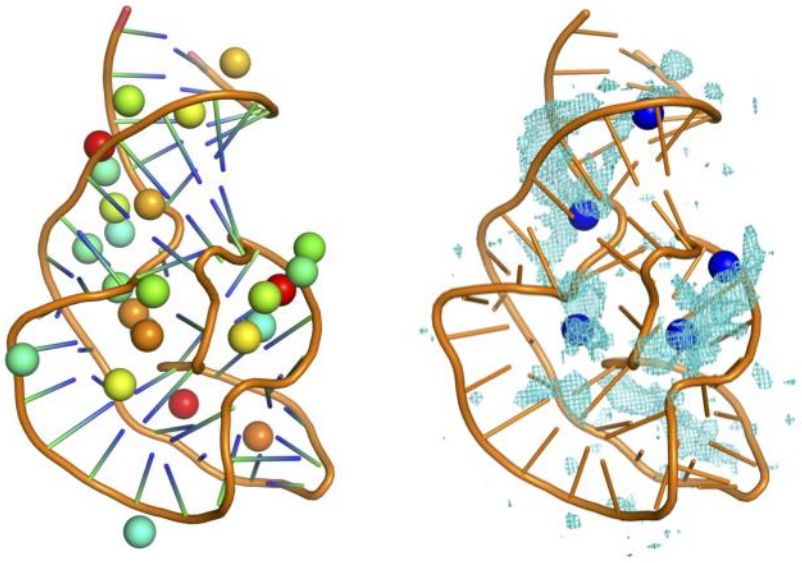
BWYV Pseudoknot



VS Ribozyme



PreQ1 Riboswitch



Twister Ribozyme

Figure T1. The left panels show centers of top 25 clusters calculated from the Mg^{2+} ion distribution from the GCMC-MD simulations around the 4 RNA systems. On the left panels, spheres representing the cluster centers are colored from red to blue. The right panels illustrate the binding sites of ions captured in X-ray or NMR experiments (1-4). Mg^{2+} occupancy maps with GFE cutoff of -2 kcal/mol are shown. On the right panels, Mg - Dark blue spheres, Na - Yellow spheres, K - Purple sphere, Ca - brown spheres.

Table T1. Minimum distance between experimental Mg^{2+} binding sites for Twister and cluster center for Mg^{2+} positions sampled in the 100 mM $MgCl_2$ GCMC/MD validation runs.

Crystal structure identifier for Mg ion	Distance to nearest cluster center (Å)
MG 101	3.16
MG 102	5.61
MG 103	2.02
MG 104	3.22
MG 105	2.10

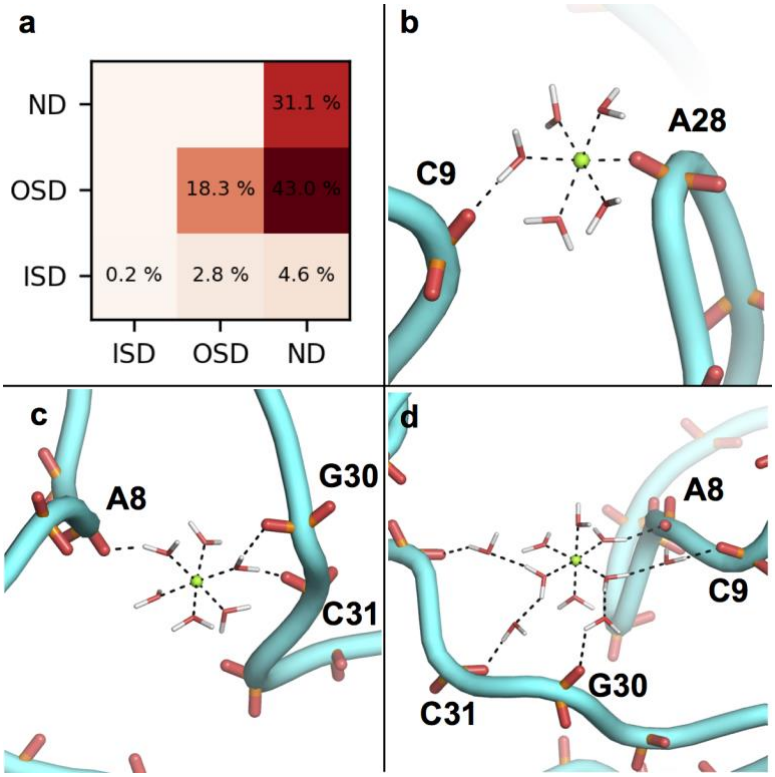


Figure T2. a) Distribution of combination of interactions between Mg^{2+} and NBPOs over all the simulations. Interaction types are direct-interaction (Inner-Shell-Dehydrated: ISD), indirect-interaction (Outer-Shell-Dehydrated: OSD) and diffused-interaction (Non-Dehydrated: ND). 70% Mg^{2+} ions were found to be participating in at least one strong interaction, either ISD or OSD, for one of the NBPOs in each pair. Very rarely both NBPOs are simultaneously coordinated at inner-shell level, while OSD+ND or ND+ND types of combination occur more frequently. Examples of the types of combinations include b) ISD with OSD interactions providing strong coordination with phosphate groups c) OSD with OSD interactions with phosphate groups and d) OSD with ND interactions with phosphate groups.

Supporting References:

1. 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.

2. 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.

3. 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.

4. 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.

5. 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.

6. 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.

7. 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.

8. 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.