ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

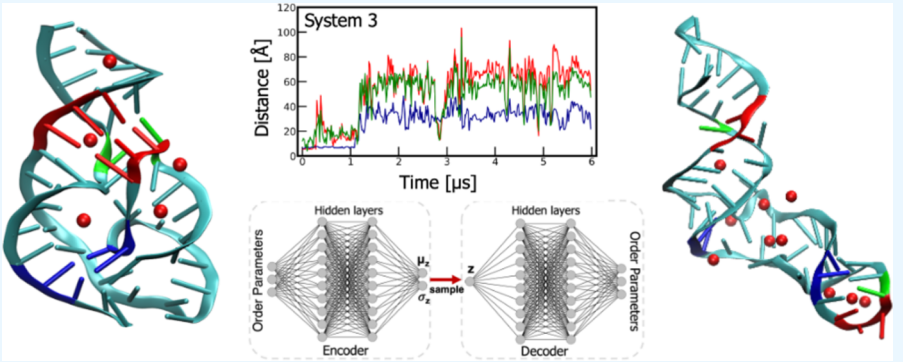
Influence of Mg^{2+} Distribution on the Stability of Folded States of the Twister Ribozyme Revealed Using Grand Canonical Monte Carlo and Generative Deep Learning Enhanced Sampling

Mert Y. Sengul and Alexander D. MacKerell, Jr.*

Cite This: *ACS Omega* 2023, 8, 19532−19546

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Metal ions, particularly magnesium ions (Mg^{2+}), play a role in stabilizing the tertiary structures of RNA molecules. Theoretical models and experimental techniques show that metal ions can change RNA dynamics and how it transitions through different stages of folding. However, the specific ways in which metal ions contribute to the formation and stabilization of RNA's tertiary structure are not fully understood at the atomic level. Here, we combined oscillating excess chemical potential Grand Canonical Monte Carlo (GCMC) and metadynamics to bias toward the sampling of unfolded states using reaction coordinates generated by machine learning allowing for examination of Mg^{2+}−RNA interactions that contribute to stabilizing folded states of the pseudoknot found in the Twister ribozyme. GCMC is used to sample diverse ion distributions around the RNA with deep learning applied to iteratively generate system-specific reaction coordinates to maximize conformational sampling during metadynamics simulations. Results from 6 μs simulations performed on 9 individual systems indicate that Mg^{2+} ions play a crucial role in stabilizing the three-dimensional (3D) structure of the RNA by stabilizing specific interactions of phosphate groups or phosphate groups and bases of neighboring nucleotides. While many phosphates are accessible to interactions with Mg^{2+}, it is observed that multiple, specific interactions are required to sample conformations close to the folded state; coordination of Mg^{2+} at individual specific sites facilitates sampling of folded conformations though unfolding ultimately occurs. It is only when multiple specific interactions occur, including the presence of specific inner-shell cation interactions linking two nucleotides, that conformations close to the folded state are stable. While many of the identified Mg^{2+} interactions are observed in the X-ray crystal structure of Twister, the present study suggests two new Mg^{2+} ion sites in the Twister ribozyme that contribute to stabilization. In addition, specific interactions with Mg^{2+} are observed that destabilize the local RNA structure, a process that may facilitate the folding of RNA into its correct structure.

INTRODUCTION

There is an increasing amount of research showing that a large portion of the genetic material in complex organisms, such as mammals, is in the form of RNA and that these RNAs have a variety of functions.[1] This makes them essential to fundamental biological studies and a promising target for new medical treatments.[2] One of the most important characteristics of RNAs is their ability to take on a wide range of structures from simple helices to complex, diverse folded shapes. To fully comprehend their functions, it is crucial to examine the folding and stabilization process that enables RNAs to adopt their three-dimensional (3D) structures.

Pseudoknotted RNAs are of particular interest due to their functional importance and as a model for studying the folding/unfolding pathways of complex structures.[3] This is due to these RNAs having both secondary and tertiary structures and containing complicated interactions such as noncanonical base

Received: February 12, 2023
Accepted: May 11, 2023
Published: May 24, 2023

pairs, triplexes, coaxial stackings, and sharp turns. Accordingly, the study of pseudoknots can aid in improving predictions of RNA tertiary structures, as the presence of noncanonical interactions presents significant challenges in developing sampling algorithms and free-energy estimation rules.[4]

It is well known that folding into compact tertiary structures requires overcoming the electrostatic repulsion of the negatively charged phosphate backbone of polynucleotides.[5] The ionic environment around the RNA helps to shield these charges, making folding possible.[6] Metal cations play a crucial role in the folding and catalytic activity of RNA, as highlighted in several reviews.[6−10] Because of this, researchers have been working to gain a deeper understanding of how metal ions interact with RNA and help stabilize its tertiary structure. However, it is difficult to fully understand the electrostatic forces that drive RNA folding, as the sampling of metal ions around the RNA and the folding processes are interconnected. There have been many notable experimental and theoretical studies on Mg^{2+} ions and RNA interactions.[11−25] Studies using all-atom classical molecular dynamics (MD)[12−15] are limited due to challenges in the sampling of Mg^{2+} ions as the exchange rates of Mg^{2+} (water−Mg^{2+} and phosphate−Mg^{2+} complexation) are in the micro- to millisecond time scale,[26] which are beyond timescales readily achievable by MD methods. As a result, these studies have not provided a complete understanding of Mg^{2+}−RNA interactions and their role in RNA folding. To address this issue, nucleic acid coarse-grained models have been used to examine the influence of Mg^{2+} ions in nucleating the folding.[13,27,28] While many useful insights have been gained from these studies, the impact of an extensive sampling of Mg^{2+} ions around the RNA and the interplay of Mg^{2+} ions and water interactions with RNA in RNA folding still requires additional investigation.

To gain a more detailed understanding of the effect of the distribution of Mg^{2+} ions on the stabilization of the folded state of RNA molecules, we combined metadynamics simulations and oscillating excess chemical potential, μ_{ex}, Grand Canonical Monte Carlo (GCMC) sampling of the ion distribution in a pseudoknot RNA system. The pseudoknot system used in this study is the Twister ribozyme, which was selected due to the wide range of available data, including experimental studies showing folding kinetics and self-cleavage activity in the presence of Mg^{2+} ions,[29] and has been shown to undergo 3D conformational transitions in relatively lower salt concentrations.[30] The secondary and tertiary structures of the Twister ribozyme are given in Figure 1, and a detailed description of the 3D structure can be found in the Methods section. Our workflow, as can be seen in Figure 2, starts with the GCMC sampling, which allows for the redistribution of Mg^{2+} ions, addressing the issue of exchange rates mentioned above, while the biased MD simulations via metadynamics allow the RNA molecules to explore a range of conformations including both unfolded and folded states. Generative deep learning and clustering techniques were used to predict the reaction coordinates to facilitate unfolding. Our results show that Mg^{2+} ions could stabilize tertiary structures via either bridging phosphate group−phosphate group as an outer-shell or inner-shell cation or phosphate group nucleobase of neighboring nucleotides as an outer-shell cation. If the Mg^{2+} ion bridges two nucleotides as an inner-shell cation, the strength of the tertiary contact is larger compared to outer-shell cation bridging, and according to our analysis, this increase in strength translates into a significant enhancement of the stability of the tertiary structure. Notably, the ability of Twister to sample conformations close to the experimental X-ray crystal structure is associated with multiple, specific ion

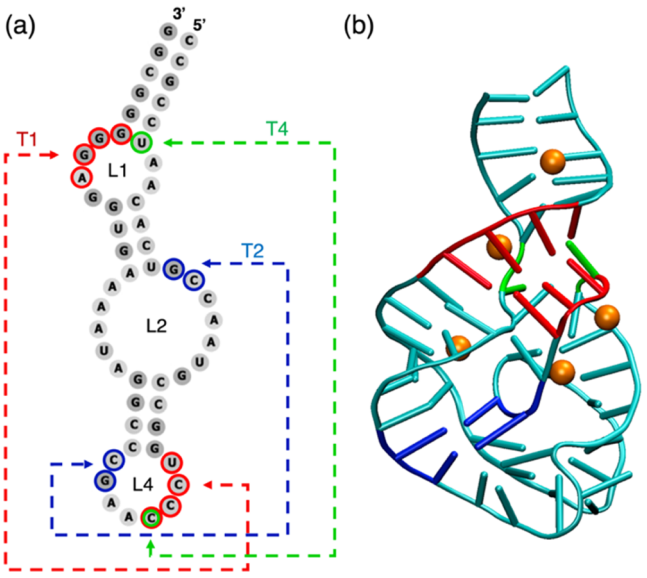


Figure 1. Secondary and tertiary structure of the Twister ribozyme. (a) Twister ribozyme secondary structure and the major tertiary contacts. (b) The tertiary structure of Twister ribozyme and the major tertiary contacts: T1, red; T2, blue; T4, green. T1 includes the C25−G49, C26−G48, C27−G47, and A28−A46 interactions, T2 involves G13−C31 and C14−G30 Watson−Crick interactions, and T4 represents adjacent nonbridging phosphate oxygens (NPBO) on U6 and C25.

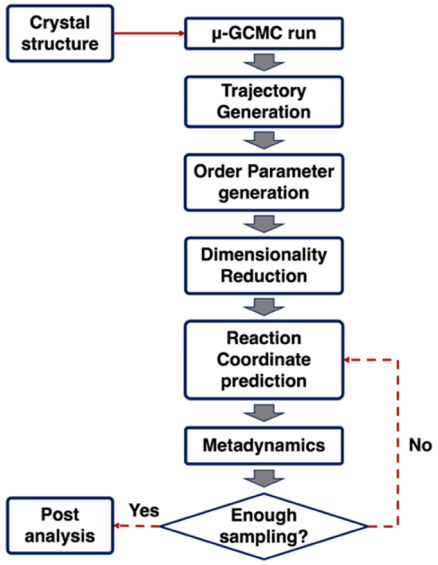


Figure 2. Workflow of the procedure used in this study. The workflow starts with a crystal structure of the RNA molecule and input parameters necessary for the oscillating μ_{ex} GCMC sampling, metadynamics, and machine learning (ML) stages. The trajectory generation through ML reaction coordinate prediction and is iterative until satisfactory sampling is achieved and needs to be terminated by the user. Initial trajectory generation is done using unbiased molecular dynamics, order parameter generation is done using an in-house developed code, and dimensionality reduction and reaction coordinate predictions are done using the AMINO and RAVE machine learning approaches. For details, see the Methods section.

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

sampling around the RNA. These results suggest Mg^{2+} ion locations for the Twister ribozyme that facilitate folding beyond those observed in the crystal structure as well as interactions that destabilize the local tertiary structure of RNA, events that may facilitate RNA folding into its correct structure.

METHODS

The crystal structure of the Twister ribozyme used in this study is based on Osa-1−4 sequence from Oryza sativa (PDB: 4OJI).[31] The secondary and tertiary interactions in Twister are given in Figure 1 where the shown structure has the missing nucleotides model built and subjected to energy minimization from our previous study.[32] The Twister ribozyme used in this study has 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, but it lacks the P3 and P5 stem-loops. The crystal structure is missing two nucleotides (nucleotides 17, 18), which were added by using CHARMM[33] and the internal coordinates in the CHARMM36 nucleic acid force field.[34−36] All systems were initially prepared using CHARMM-GUI[37] and were minimized and equilibrated in CHARMM[33] using harmonic restraints on the nonhydrogen atoms of the backbone and bases with force constants 1 and 0.1 kcal/mol/Å^2, respectively. The Twister ribozyme and ions were modeled using CHARMM36 additive force field,[34−36] including the CHARMM36 default Mg^{2+}, K^+, and Cl^- parameters as used in our previous studies.[32,38] Water molecules were modeled using the CHARMM-modified TIP3P model.[39,40] Cubic simulation box dimensions of 87.6 Å were used for all systems that included a total of 21920 water molecules. The number of ions was determined to set the magnesium chloride ($MgCl_2$) concentrations to 100 mM while yielding a neutral system, which resulted in 40 magnesium ions (Mg^{2+}), 42 potassium ions (K^+), and 69 chloride ions (Cl^-), corresponding to final concentrations of ∼100, 105 and 173 mM, respectively. We note that in the crystal structure of the Twister ribozyme (PDB 4oji) there are 5 Mg^{2+} ions/monomer. This corresponds to a total of 60 Mg^{2+} ions identified in the full unit cell (P 65 2 2 space group) that, based on the total volume of the unit cell, corresponds to a concentration of approximately 185 mM. The ions were distributed randomly in a simulation box and resampled using an oscillating μ_{ex} Grand Canonical Monte Carlo/Molecular Dynamics (GCMC/MD), which is explained in detail in the following section. All covalent bonds containing hydrogen atoms were constrained using SHAKE algorithm.[41] The Lennard-Jones potential was smoothed to zero between 10 and 12 Å using a force switch, and particle mesh Ewald (PME) method was used to calculate electrostatic interactions with a real space cutoff of 12 Å.[42] The equilibration simulations were performed for 100 ps in the NPT ensemble using a timestep of 1 fs. The production simulations were performed in the NPT ensemble using the Langevin integrator at 1 bar pressure and 300 K temperature in the OpenMM simulation package[43,44] using a timestep of 2 fs. The average pressure was kept constant using a Monte Carlo barostat. The simulation lengths were adjusted for different workflow stages, and details are given in corresponding sections in the main text.

In this study, the focus was on unfolding events of the overall tertiary structure such that secondary structural interactions were not investigated. Thus, Watson−Crick (WC) pairs or other interactions that stabilize the canonical secondary structure regions are not of direct interest; therefore, an external harmonic restraint was applied to WC pairs to maintain the canonical secondary structure. The restraint force is defined as $E = \frac{1}{2}k(x - x_0)^2$, where the force constant k is 40 kJ/mol/nm^2 and the equilibrium distance x_0 is 0.3 nm. Metadynamics simulations in which the WC pairs in the duplex regions were not restrained lead to significant overall opening of those regions that would not reassociate during the simulations (not shown). Accordingly, the inclusion of the WC restraints was performed to focus our efforts on the equilibrium of the unfolded and folded states of the tertiary structure.

Oscillating Excess Chemical Potential Grand Canonical Monte Carlo Sampling. The generation of unique ion distributions around the Twister ribozyme for each individual system was performed using a particle mesh Ewald (PME) oscillating excess chemical potential (μ_{ex}) grand canonical Monte Carlo (GCMC) algorithm.[45] The μ_{ex}-GCMC sampling incorporates an oscillation scheme that is applied to the ions in the system to overcome low acceptance ratios.[45] The sampling involves MC insertion, deletion, rotation, and translation moves applied to the ions and water molecules. The sampling used in this study resamples the ion distributions using MC deletion and insertion moves on Mg^{2+} while simultaneously deleting or inserting K^+ and Cl^- ions to maintain charge neutrality throughout the system.[32] In addition, moves of the water molecules were performed. The μ_{ex}-GCMC sampling was performed on nine systems containing one Twister ribozyme with identical ion concentrations. Each μ_{ex}-GCMC run involved ten cycles of resampling of the ion distribution and 10 ns NPT MD equilibration simulations after each sampling cycle. One cycle of μ_{ex}-GCMC sampling involves seven stages, which are as follows: (1) 20,000 MC steps of deletion moves on Mg^{2+} and simultaneous K^+ insertions; (2) 80,000 MC steps of rotation and translation moves on all ions and water; (3) repetition of stages (1) and (2) until all Mg^{2+} was removed; (4) 20,000 MC steps of insertion moves on Mg^{2+} and simultaneous K^+ deletion; (5) 80,000 MC steps of rotation and translation moves on all ions and water; (6) repetition of stages (4) and (5) until the desired Mg^{2+} concentration is obtained; (7) energy minimization of the whole system for 5000 steepest descent steps following by a 100 ps NVT. The details of μ_{ex}-GCMC sampling used in this study can be seen in ref 32. Each system was then subjected to a 500 ns MD simulation in the NPT ensemble which was used to initiate the reaction coordinate prediction.

Reaction Coordinate Prediction Method. Prediction of the reaction coordinate for the individual systems was performed using a two-stage protocol. The first stage is the determination of the specific order parameters (OPs) that represent the conformational transitions of the whole system. This was done using a clustering-based dimensionality reduction method called Automatic mutual information noise omission (AMINO),[46] which is specifically developed for use in macromolecular systems. AMINO uses a K-medoids clustering with a mutual information-based distance metric to reduce the OP space to a few clusters and finds the most representative OP of each cluster using rate-distortion theory, therefore reducing the high-dimensional OP space to a few dimensions while keeping most of the information, in this case, the conformational changes of the tertiary structure of Twister.

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

Table 1. Nucleotides Contributing to the Most Representative Order Parameters (OPs) Selected by AMINO for Each System[a]

system	OP-1	OP-2	OP-3	OP-4	OP-5
System 1	24,25,26,27−41,42,43,44	2,3,4,5−36,37,38,39	8,9,10,11−35,36,37,38	10,11,12,13−51,52,53,54	25,26,27,28−35,36,37,38
System 2	8,9,10,11−35,36,37,38	2,3,4,5−8,9,10,11	4,5,6,7−12,13,14,15	26,27,28,29−35,36,37,38	
System 3	2,3,4,5−8,9,10,11	42,43,44,45−49,50,51,52	8,9,10,11−18,19,20,21	23,24,25,26−38,39,40,41	
System 4	2,3,4,5−9,10,11,12	7,8,9,10−18,19,20,21	39,40,41,42−49,50,51,52		
System 5	16,17,18,19−45,46,47,48	2,3,4,5−10,11,12,13	24,25,26,27−37,38,39,40		
System 6	27,28,29,30−35,36,37,38	30,31,32,33−49,50,51,52	18,19,20,21−27,28,29,30	3,4,5,6−10,11,12,13	
System 7	17,18,19,20−45,46,47,48	3,4,5,6−30,31,32,33			
System 8	29,30,31,32−50,51,52,53	9,10,11,12−36,37,38,39	2,3,4,5−10,11,12,13		
System 9	3,4,5,6−12,13,14,15	7,8,9,10−33,34,35,36			

[a]The systems differ in terms of the initial ion distribution.

More details about the implementation of the method can be found in ref 46. The OPs input into AMINO are all the distances between the center of mass (COM) of consecutive multinucleotide groups, with each group being composed of four nucleotides. The Twister ribozyme used in this study is composed of 54 nucleotides, which yields 1275 OPs. Using AMINO, we could reduce the number of OPs to 2−5, depending on the system (Table 1).

The second stage of the protocol is the optimization of a reaction coordinate for each system using the most representative OPs determined in the first stage. The reaction coordinate predictions were done using a modified version of the generative deep learning-based reaction coordinate prediction method called reweighted autoencoded variational Bayes for enhanced sampling (RAVE).[47] The generative deep learning model used in this method is a variational autoencoder (VAE),[48] which learns a low-dimensional latent space distribution of a given high-dimensional training data. The OP trajectories selected in the AMINO clustering stage are used as input data for the VAE. The neural network architecture used in this study (Figure 3) is composed of one

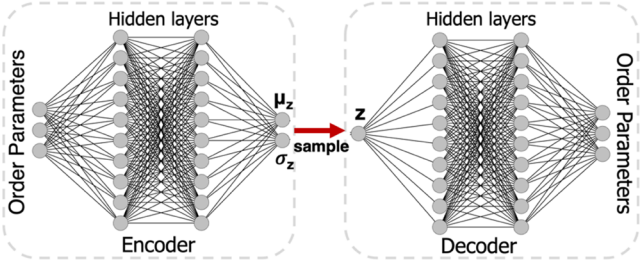


Figure 3. Variational autoencoder model is shown with a specific focus on the parameters of the deep neural networks used in this work. The encoder network converts a snapshot from order parameter trajectory into the mean and variance of a normal distribution. The decoder network takes a variable sample from the distribution and converts it into an order parameter sample similar to the input order parameter snapshot.

input layer, five encoder hidden layers, five decoder hidden layers, and one output layer. The number of nodes in the input and output layers is the same as the number of OPs selected by the dimensionality reduction stage (Table 1). The hidden layers in both the encoder and decoder parts are composed of 512 nodes. The exponential linear unit (ELU) activation function is employed for all layers,[49] and an adaptive learning rate scheme called Adaptive Moment Estimation (ADAM) is used.[50] We implemented an early stopping algorithm to avoid overfitting by stopping the training when the loss function criterion is satisfied. Using the OP trajectories, RAVE uses variational autoencoders to learn the latent distribution of OPs for a given system and parameterizes a reaction coordinate using the Kullback−Leibler divergence metric to reproduce the same OP distribution. Details on how VAE is implemented to optimize a reaction coordinate can be seen in ref 47. The reaction coordinate used in this study is a linear combination of OPs selected at the first stage, and the coefficient of each term is parameterized by RAVE. One advantage of RAVE is its ability to iterate through several molecular simulation rounds by producing the unbiased probability distribution via reweighting from previous biased rounds, so it has the ability to improve reaction coordinates at each iteration.

Metadynamics Simulations. At the end of each reaction coordinate prediction cycle, a metadynamics[51] run is conducted (Figure 2) to bias the system to sample different conformations. Here, we used well-tempered metadynamics[52] with a bias factor (γ) of 15 and an initial hill height of 1.5 kJ. The Gaussian width (σ) was selected as the standard deviation of the reaction coordinate calculated using the initial unbiased MD trajectory. The biasing Gaussian energies are deposited each 1 ps, and each cycle of biased simulation is run for 200 ns. After each cycle, the RC is reevaluated using RAVE based on the last 200 ns of dynamics with the new RC used for the next 200 ns of metadynamics. Reevaluation of the RCs only involved reweighting of the contribution of the OPs; the same OPs listed in Table 1 were used throughout all the cycles. For all 9 systems, the production run was composed of 30 200 ns metadynamics simulations yielding a total of 6 μs of simulation time. Coordinates were saved every 20 ps for analysis.

Contact Map Analyses. Interactions between nucleotides in hairpins and cations are shown as heat maps, which are based on the distance between anionic phosphate oxygens (nonbridging phosphate oxygens (NBPO)) and cations that are less than 4.2 Å. This distance is used as the criterion for identifying outer-shell ions. The data used to create the heat maps is extracted by combining the last 150 ns of the first biased MD cycle and the full simulation time for the other cycles. The values in the heat maps indicate the percentage of the simulation during which a cation was present as an outer-shell ion.

RESULTS AND DISCUSSION

The present study is designed to investigate the role of the distribution of Mg^{2+} ions on the ability of the Twister RNA to sample conformations in the vicinity of the folded state. To achieve this, 9 unique distributions of the ions around the RNA were generated using the oscillating μ_{ex} GCMC methodology. Each system was then subjected to extensive

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

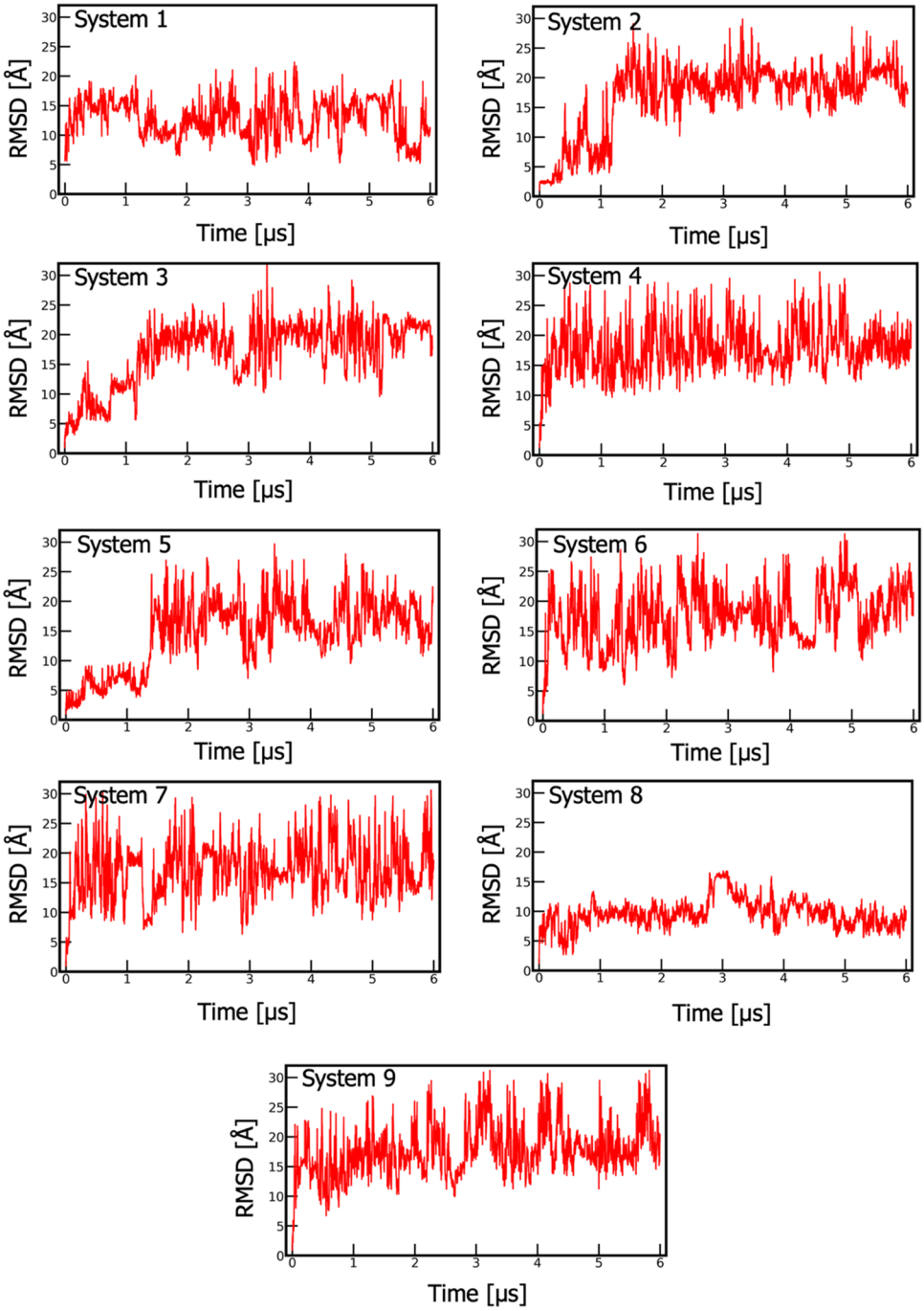


Figure 4. RMS difference for the RNA backbone nonhydrogen atoms as a function of time versus the X-ray crystal structure, 4OJI, following modeling and minimization of the missing two nucleotides. Each system is identical in terms of number of atoms, number of ions, and initial Twister ribozyme conformation. The systems differ in the Mg^{2+}, K^{+}, and Cl^{-} distributions generated by the GCMC method.

conformational sampling using metadynamics in which the RC was iteratively reevaluated to maximize conformational sampling. Given the timescales of the exchange rate of Mg^{2+} ions being from micro to milliseconds combined with the need for the RNA conformation to relax in response to the ion distribution, the ion distributions were not resampled during the metadynamics simulations, although the ions were allowed freely move and undergo significant displacements. An approximation made in the calculations was the placement of restraints on the known secondary structure of the RNA, consistent with experimental studies showing Mg^{2+} to dictate the formation of RNA tertiary structure. For example, with Twister formation of the T2 contact has been shown to be contingent on the formation of the P1 stem,[29] although in the SAM riboswitch the formation of helical structure has been shown to be the rate-limiting step in folding.[53] In addition, during preliminary metadynamics simulations without restraints on the secondary structure the RNA rapidly unfolded with respect to both tertiary and secondary structure and did not approach folded conformations in the context of either secondary or tertiary structure in any of the simulations indicating the need to maintain the secondary structure in the present study. This is likely due to the use of metadynamics leading to the loss of Watson−Crick and stacking interactions

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

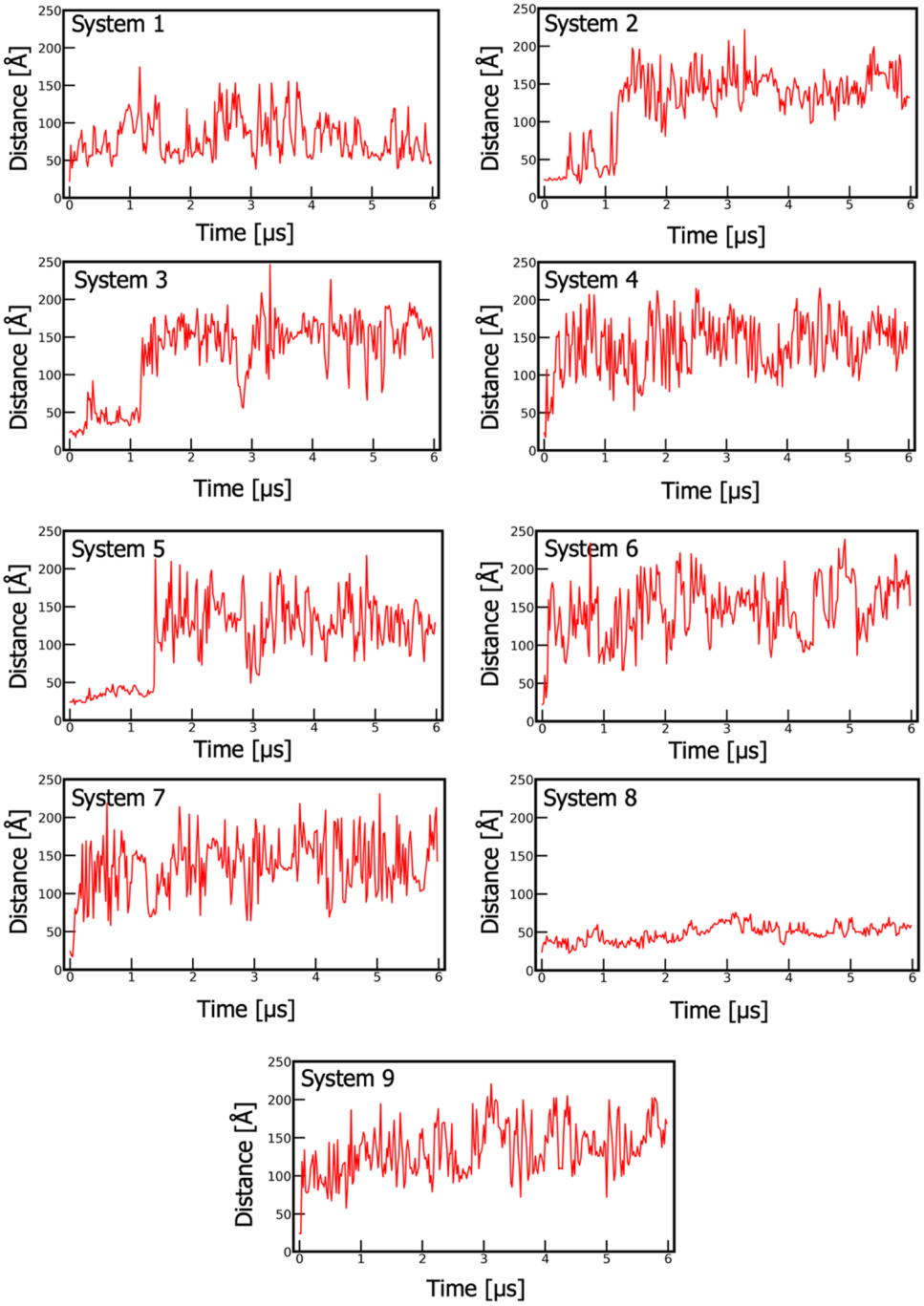


Figure 5. Change of a linear combination of T1, T2, and T4 contact COM distances with respect to time.

that cannot reform. It should be noted that standard simulations of Twister have shown the structure to be stable,[32,38] including during the present equilibration simulations. It should be further emphasized that the metadynamics method was designed to maximize conformational sampling in conjunction with AMINO and RAVE to identify unique RCs and continually reevaluated those RC throughout the simulations for each system. Thus, the present study focused on determining the distribution of Mg^{2+} ions required to sample conformations in the vicinity of the folded state. Finally, while free-energy profiles may be extracted from metadynamics simulations, such analysis in the present study is not possible due to the continual updating of the RCs for each of the 9 systems after every 200 ns of metadynamics.

To attain the goal of the present study an advanced sampling workflow was applied that combines deep learning, metadynamics, and GCMC simulations to investigate the influence of Mg^{2+} ions on the stability of the folded state of the Twister ribozyme (Figure 2). Ion sampling, which is limited due to the short timescales of classical MD simulations, is conducted using GCMC, and the conformational sampling is done via metadynamics using system-specific reaction coordinates predicted by clustering and generative deep learning-based algorithms. As the first stage of the workflow, the implemented protocol redistributes the Mg^{2+} as well as the K^+ and Cl^- ions in the simulation system using an oscillating μ_{ex} GCMC protocol. From this procedure, 9 individual systems were created that only differed in the ion distributions in each

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

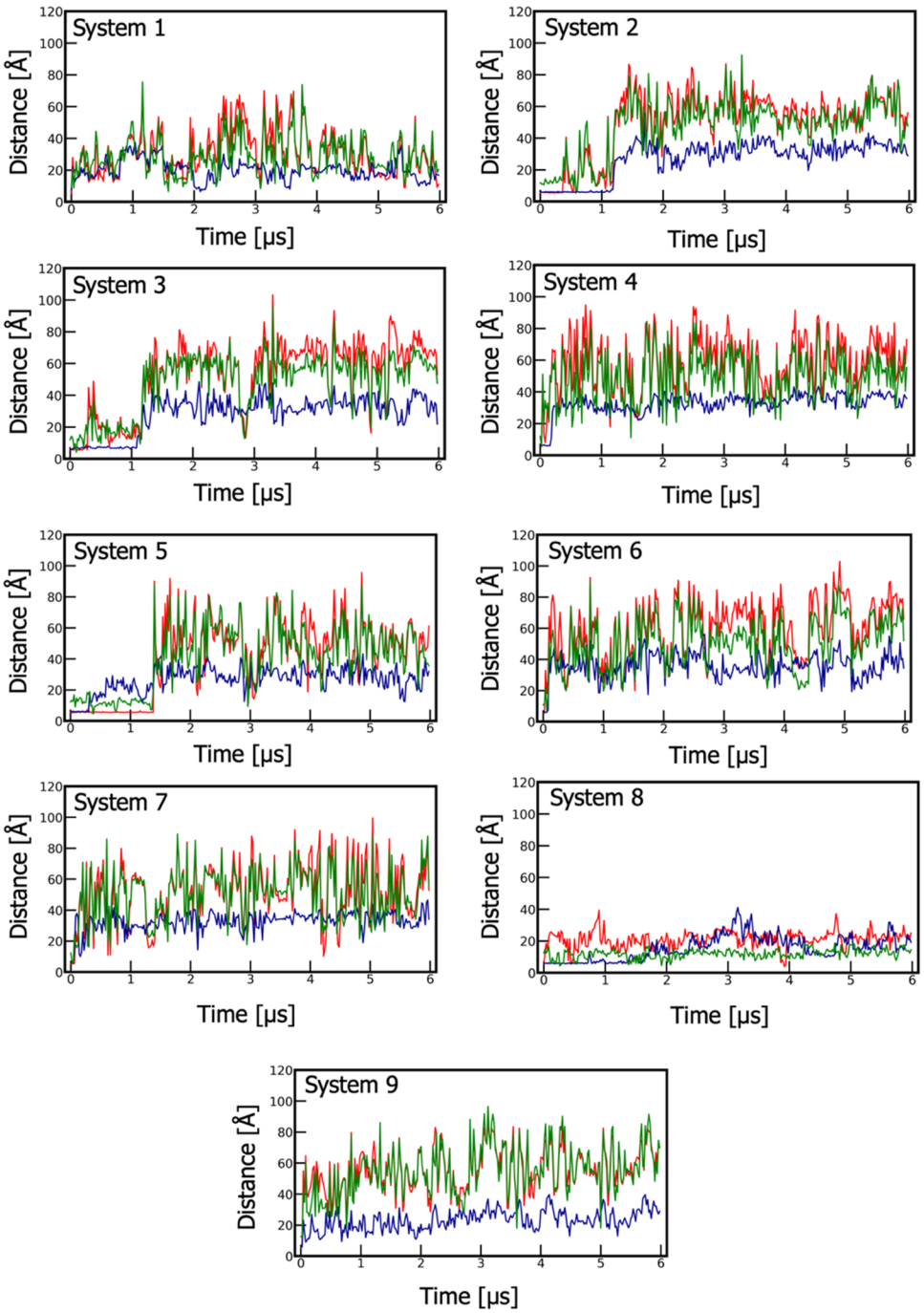


Figure 6. Change of major tertiary contact (T1, T2, T4) COM distances with respect to time for the Twister ribozyme. T1, T2, and T4 contact distances are represented by red, blue, and green lines, respectively.

system. All additional calculations were performed on each system separately. The sampling of the ions is followed by relaxation of the system and a 500 ns long unbiased MD simulation. From the resulting trajectories, the internucleotide distances were extracted, which are the basis of the OPs used to facilitate RNA conformational sampling and are then used as the training data for the dimensionality reduction stage. From this process, the order parameters that are the most important to capture the behavior of all OPs in the training trajectory are selected. The selected order parameter trajectories are then used as the training set for the variational autoencoder-based reaction coordinate prediction model, which parametrizes a linear combination of the selected order parameters to reproduce the latent space distribution of the order parameter trajectory. The reaction coordinate is then used to bias the system using metadynamics, with the RC updated every 200 ns to maximize conformational sampling. This was performed for 30 cycles yielding a total of 6 μs of sampling for each system. Overall, nine independent Twister ribozyme simulations were conducted where each system initially differed based on the distribution of Mg^{2+}, K^{+}, and Cl^{-} ions from the GCMC runs such that the generative deep learning approach identified different order parameters and subsequently different reaction coordinates for each system that yielded unique sampling of the folded and unfolded states of Twister via metadynamics. The selected order parameters for the individual systems are

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

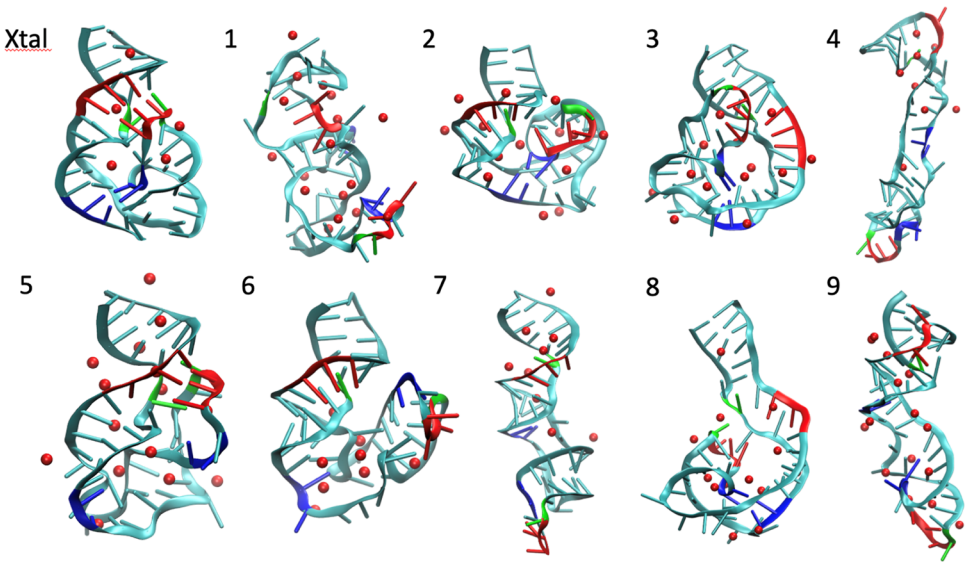


Figure 7. Conformations of Twister from the X-ray crystal structure (Xtal) and from 1 μs of the 9 metadynamics simulations indicated by the labels 1 through 9. Images in cartoon representation with the T1, T2, and T4 tertiary contacts are shown in blue, red, and green, respectively. Mg^{2+} ions within 5 Å of the Twister nonhydrogen atoms are shown as red vdW spheres.

shown in Table 1. The number of order parameters varies from 2 to 5. In addition, the sets of nucleotides also vary, though some nucleotides are in OPs common to all systems, such as 3, 4, and 5. However, there are differences in the second set of nucleotides that contribute to the second OP. For example, systems 1 and 7 show different nucleotides in the OP involving nucleotides 3, 4, and 5 versus the other systems. The OPs in Table 1 were then used in the context of the RAVE-determined RCs that were iteratively updated for the individual 6 μs metadynamics simulations with all analysis based on the metadynamic trajectories with the coordinates saved every 20 ps for analysis.

To determine the ability of the DNN-derived RCs in the context of metadynamics to sample a broad range of conformational space, the backbone root-mean-square difference (RMSD) of the backbone nonhydrogen atoms versus the crystal structure was calculated. As may be seen in Figure 4, increases in the RMSD occur in all of the systems though variations are present in the extent of the structural changes as well as their evolution over time. These results indicate the effectiveness of the reaction coordinates generated using AMINO and RAVE in facilitating the sampling of conformational space during the metadynamics as well as the different outcomes in that sampling as a function of different ion distributions.

The use of metadynamics is designed to maximize conformational sampling as the method drives the systems toward regions of OPs not significantly sampled. This leads to the changes in RMSDs in Figure 4. However, when using such an approach, by definition, the system will not significantly sample the folded state of the RNA. However, in certain systems, it is evident that the RNA maintains a conformation close to the folded state. This occurs in systems 2, 3, and 5 for the initial part of the trajectory while in system 8 the RNA stays close to the folded state throughout the simulation. To understand more details of the conformational sampling additional analysis focused on the three tertiary interactions, T1, T2, and T4, over the course of the simulations individually and as a linear combination (Figure 1). The extent of unfolding quantified using a linear combination of the T1, T2, and T4 tertiary contact distances differences are shown in Figure 5. Consistent with the RMS difference analysis of the 9 systems, five undergo rapid unfolding to combined values of > 50 Å, including systems 1, 4, 6, 7, and 9. Systems 2, 3, and 5 maintain the approximate folded state out to ~1.2 to 1.4 μs following which more significant unfolded states are sampled. System 8 samples conformations close to the folded state throughout the 6 μs metadynamics simulation. Notably, none of the reaction coordinates involve the exact T1, T2, and T4 contacts as one of the OPs (Table 1); however, for most of the systems, these contacts dissociated rapidly at the beginning of the simulations (Figures 5 and 6) and are highly correlated with the RMSD time series (Figure 4). Thus, the conformational sampling of folded vs. unfolded states may be largely described by the three tertiary contacts.

Analysis of the individual contact distances as a function of simulation time is shown in Figure 6. For the five systems that undergo rapid unfolding, 1, 4, 6, 7, and 9, all of the tertiary contacts are rapidly lost, with the exception of system 4 where T2 is maintained for approximately 0.2 μs and then lost for the remainder of the simulation. Once unfolded, the tertiary contacts are not regained for the remainder of the simulations indicating the RNAs to be fully unfolded; an exception exists with system 1, where some sampling of T2 close to the folded state occurs for short periods later in the simulation. With the systems that maintained the folded state during the early stage of the simulations, differential unfolding behavior was observed. In systems 2, 3, and 5, contacts T1 and T2 were initially maintained, with T1 lost before 0.5 μs in 2 and 3 with T2 maintained past 1 μs. In system 5, T2 is lost before 0.3 μs with T1 maintained out to ~1.3 μs. With all those systems, after the final tertiary contact is lost, full unfolding occurs and conformations near the folded states are not sampled throughout the remainder of the simulations. System 8 sampled conformations close to the folded state throughout the 6 μs metadynamics simulation. T2 is well maintained out to ~ 1.5 μs. Both T1 and T4 are lost at the beginning of the simulation, but the contacts stay in the vicinity of distances in

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

the folded state throughout the simulation, typically being within 20 Å during the majority of the simulation. These results further indicate the diversity of conformational space being sampled in the metadynamics simulations, with the RNA able to sample regions close to the folded state in some cases despite the presence of enhanced sampling along the RCs in the metadynamics simulations. Additional analysis of 2D distributions of T1 vs T2, T1 vs T4, and T2 vs T4 shown in Figure S1 in the Supporting Information further illustrates the range of conformations being sampled in the present study. As all of the simulations were initiated with the same conformation of Twister, subsequent analysis focused on the interactions of the Mg^{2+} ions with the RNA in the different systems, information that allows for an understanding of the contributions of the ions to sampling of conformations close to the folded state.

Snapshots at 1 μs from all 9 simulations along with the X-ray crystal structure are shown in Figure 7. Mg^{2+} ions within 5 Å of the RNA nonhydrogen atoms are included. As may be seen, there is a substantial difference in the conformations generated in the metadynamics simulations. The systems that were still partially folded at 1 μs, systems 2, 3, 5, and 8, maintain globular structures. In the systems that rapidly unfolded more extended conformations are observed, most notably with 4, 7, and 9. Systems 1 and 6 are less extended though they also rapidly unfolded indicating the variety of unfolded states that are accessible to Twister. In all systems, a substantial number of ions are in the vicinity of the RNA, as described in more detail below.

To understand the ability of Mg^{2+} to maintain the overall fold of Twister in the vicinity of the folded state, analysis of the Mg^{2+} distributions was carried out. To facilitate the analysis, the ion locations on the Twister crystal structure were identified along with the interacting nucleotides and specific atoms (Table 2). In the crystal structure, two ions, D and E,

Table 2. Mg^{2+} Ions Observed in the X-ray Crystal Structure, 4OJI, Nucleotides Close to Those Ions, and the Tertiary Contacts in the Vicinity of Those Ions

Xtal MG	minimum distance, Å	nucleotides, atoms, and minimum distances to Mg^{2+}, Å	tertiary contact
A	4.04	G22 N7 4.27, G23 O6 4.04, **U24 O4 4.34, A29 O2P 4.20**	T1′, T2′
B	4.28	A8 O2P 4.58, A10 N6 4.76, A41 N7 4.29, G42 O6 4.71	
C	4.72	**G49 O2P 4.72**, G50 N7 4.85	T1
D	2.04	**G23 O3′ 4.48, U24 O2P 2.13, C25 O3′ 3.28, C26 O1P 2.04**	T1, T4
E	2.06	**C5 O3′ 4.49, U6 O1P 2.06, A46 N6 4.27, G47 O6 4.39**	T1, T4

have direct contact with the phosphates of nucleotides involved in T1 and T4. T4 involves short phosphate−phosphate interactions of 5.2 Å in the crystal structure, which would be expected to be stabilized by the two Mg^{2+} ions in their vicinity. Ion B is located in a central cavity of Twister though it is not close to any individual residue, while C is located in the periphery of the RNA. Ion A is also located in a region between portions of the phosphodiester backbone whose bases are primarily hydrogen bonding with other bases in regions of the RNA distal to that surrounding the ion.

Nucleotides are selected based on having one or more nonhydrogen atoms within 5 Å of a Mg^{2+} ion. Bold indicates nucleotides involved in or adjacent to the tertiary contacts, and ′ indicates nucleotides that are adjacent to those in the tertiary contact definition. Residues involved in the tertiary contacts. T1: C25−G49, C26−G48, C27−G47, and A28−A46 interactions. T2: G13−C31 and C14−G30 Watson−Crick interactions. T4: Adjacent nonbridging phosphate oxygens (NPBO) on U6 and C25.

Analysis of the Mg^{2+}−RNA interactions first involved the probability of a Mg^{2+} being in the vicinity of the nonbridging phosphate oxygens over the entire simulations. Results in Figure 8 show that all of the nucleotides interact with Mg^{2+} ions at least in an outer-shell coordination during the simulations. This indicates that all of the Mg^{2+} ions are not locked into specific locations, which appears to be due to the extensive amount of conformational sampling of the RNA. However, high probabilities of 0.7 or more do occur. This is most evident in system 8 that stays closest to the folded state, but also occurs in systems 1, 2, 4, and 9. In the latter systems, the ions are interacting with nucleotides that are not close to Mg^{2+} in the crystal structure or are interacting with nucleotides close to the Mg^{2+} ion A that is not directly coordinated with the RNA. Only system 8 involves interactions of its nucleotides that participate in direct coordination with Mg_{2+} in the crystal structures (C25 with ion A and U6 with ion E, Table 2). These results indicate the importance of ion interactions with those sites for maintaining the conformation close to the folded state while the region occupied by ion A in the crystal structure does not appear to be important given the sampling of largely unfolded states in those systems.

Further analysis of the Mg^{2+}−RNA interactions and possible impact on conformational sampling involved contact maps between the NBPOs and the Mg^{2+} ions presented as heat maps (Figure 9). The heat maps show the total time spent by Mg^{2+} ions around the phosphate groups based on a distance between the cation and NBPOs of less than 4.2 Å. The red lines on the heat maps indicate the nucleotides in direct contact with Mg^{2+} ions in the crystal structure (U6, U24, C26, Table 2). For analysis purposes, we separate the nine systems into the three classes discussed above; those that rapidly unfolded, systems 1, 4, 6, 7, and 9; those that maintained substantial folded characteristics beyond 1 μs, systems 2, 3, and 5; and system 8 that sampled conformations close to the folded state throughout the entire simulation. Of the less stable systems, in system 4, MG18 is an outer-shell cation around C31 at the beginning of the simulation, which may help to initially stabilize T2 (Figure 6), but the cation migrated away from the RNA molecule as the simulation progressed. Similarly, in system 9, there is strong ion sampling in the vicinity of the T2 nucleotides, 13, 14, 30, and 31, which leads to the stabilization of that contact which samples partially folded states throughout the full simulation. However, there is a lack of ions in the vicinity of the remainder of the RNA consistent with sampling of mostly unfolded states in the system. The three other systems that quickly unfolded, 1, 6, and 7, show significantly different patterns of Mg^{2+} sampling. Systems 6 and 7 show significant ion sampling around only a limited number of nucleotides. The stable ions in system 6 are not in the vicinity of the tertiary contacts, consistent with the lack of stability of that system while with system 7 MG27 spends most of the time fluctuating between U24 to C27. However, while this is a region involved in the important T1 contact, during the initial phase of the simulation MG27 contacts the C26 NBPO as an inner-shell cation as well as having outer-shell

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

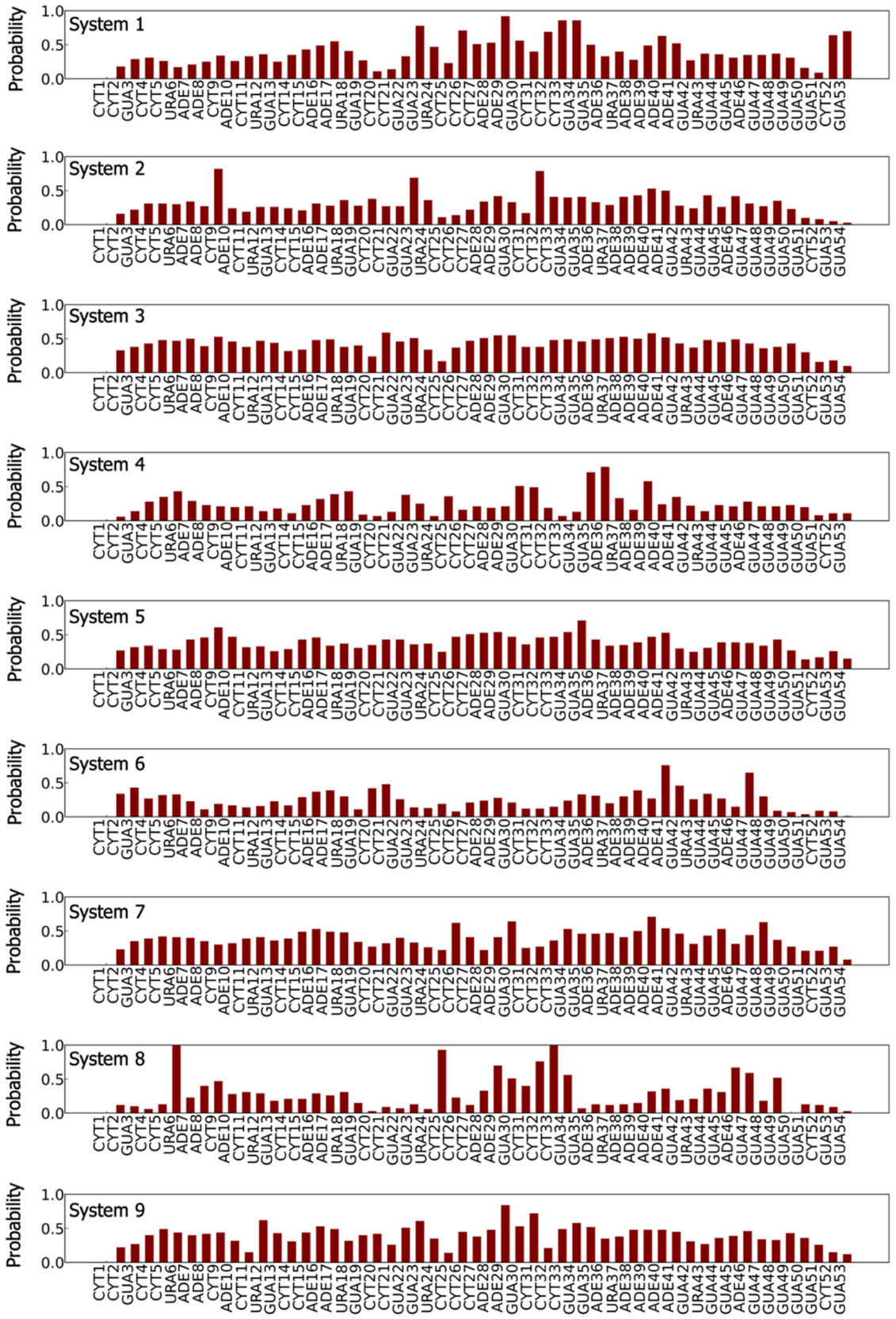


Figure 8. Probabilities of Mg^{2+} being within 4.2 Å of the nonbridging phosphate oxygens (NBPO)s over the entire metadynamics simulations.

interactions with the phosphate of C25 (Figure 10). While these represent favorable phosphate−Mg^{2+} interactions, they contribute to the opening of the C25−G49 WC interaction and the loss of the T1 tertiary contact indicating specific ion interactions to contribute to local destabilization. As shown in Figure 10B,C, this conformational change corresponds to a flipping of C25 away from its WC partner G49 as well as loss of the C26−G48 hydrogen bond, which assumes a more

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

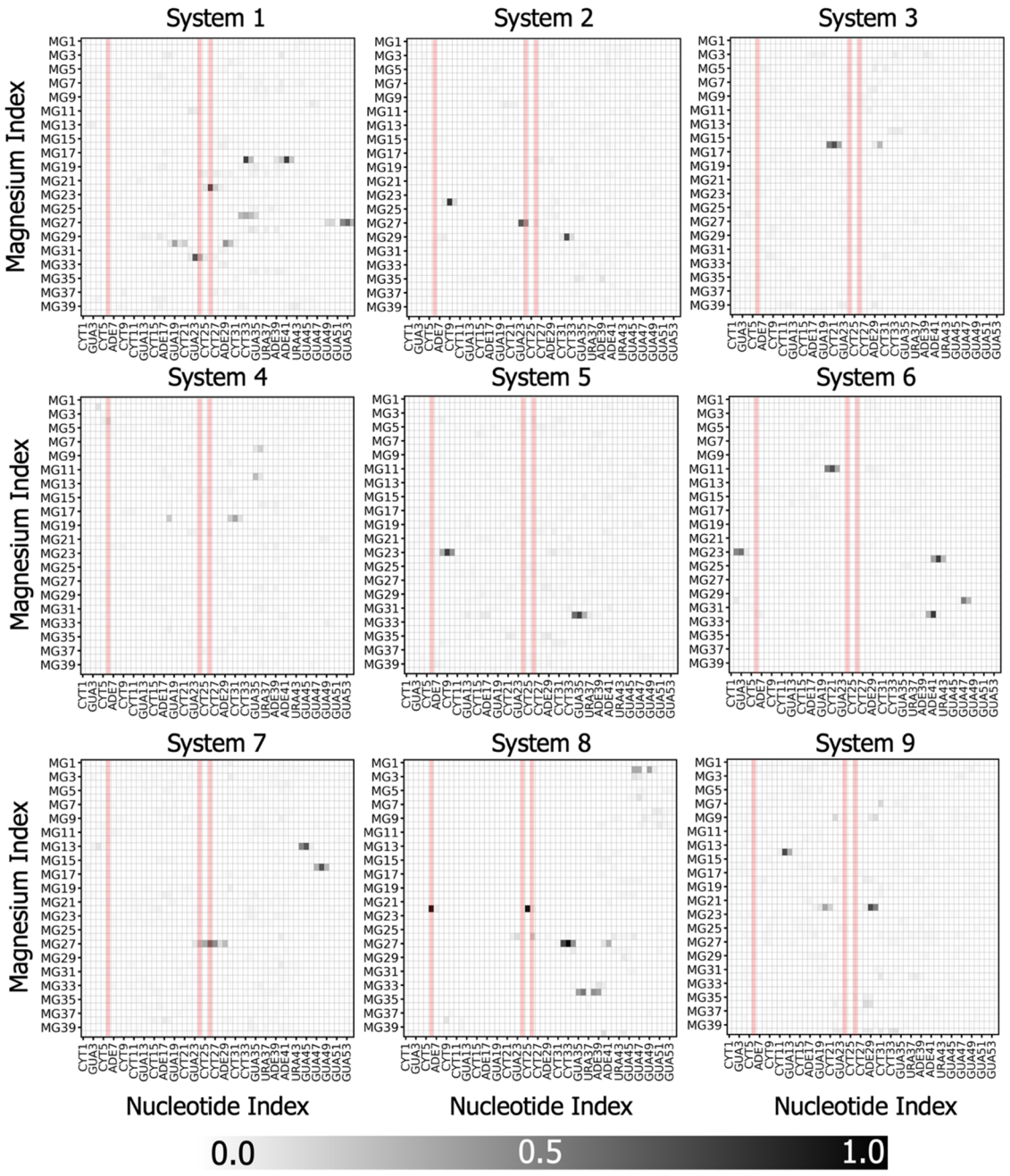


Figure 9. Contact maps between anionic phosphate oxygens of the nucleotides and the Mg^{2+} ions present in the system. The maps show the time spent by cations around the individual phosphate groups. Only outer-shell cations were sampled in the simulations based on a cutoff distance of 4.2 Å between the ions and the phosphate anionic oxygens. The heat map intensity represents the percentage of the whole trajectory that a cation spent as an outer-shell ion, as indicated by the bar. The red bars indicate the location of the Mg^{2+} ions in the crystal structure.

stacked interaction at 60 ns. The loss of these interactions ultimately led to the total loss of the T1 tertiary contact (Figure 6). In our previous study, the ability of Mg^{2+} to locally destabilize the RNA was observed in a number of cases indicating such events may contribute to the correct folding of RNA.[32] System 1 is interesting as there are a number of ions around the RNA for a majority of the simulation; however, ions are not located around U6 essential for T4 and only low sampling occurs in the vicinity of T2. While system 1 has its T1 contact fully dissociated the extent of unfolding is relatively small, which may be due to the inner-shell cation (MG22) located near C26. In this location, it can mediate energetically favorable interactions during the conformational changes; therefore, the range of T1 opening is limited. However, there is no sampling of Mg^{2+} around the T1 nucleotides 47 to 49 leading to the extent of unfolding that does occur. The lack of these as well as the ions around T4 is proposed to lead to the overall unfolding of the RNA in this system.

In the more stable systems, 2, 3, 5, and highly stable system 8, Mg^{2+} ions occupy regions in the vicinity of the three important tertiary contacts. For the T1 contact, system 5 is relatively stable compared to other systems, taking approximately 1.4 μs for unfolding to occur. This is interesting as this system does not have any Mg^{2+} ions around the crystal structure cation locations; however, MG23 is located between C9 and A28 with inner-shell coordination to C9 and outer-

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

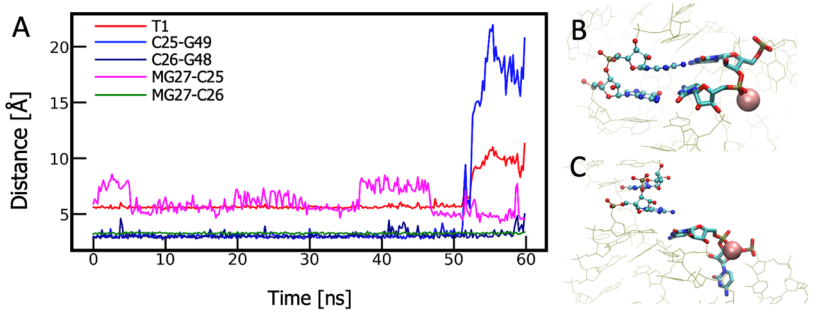


Figure 10. (A) Time series of selected interactions in system 7 during the first 60 ns of the 6 μs metadynamics simulation. Shown are the T1 tertiary COM to COM distance (red), the N1−N3 distances of the C25−G49 and C26−G48 base pair interactions, and the shortest on-bridging phosphate oxygen (NBPO) to Mg^{2+} 27 interactions. Snapshots at (B) 40 ns and (C) 60 ns showing the local RNA around G48 and G49 (CPK, atom-colored), C25 and C26 (stick, atom-colored), and Mg^{2+} 27 (sphere) with the remaining RNA as tan lines. Of the ions on Mg^{2+} 27 is shown for clarity.

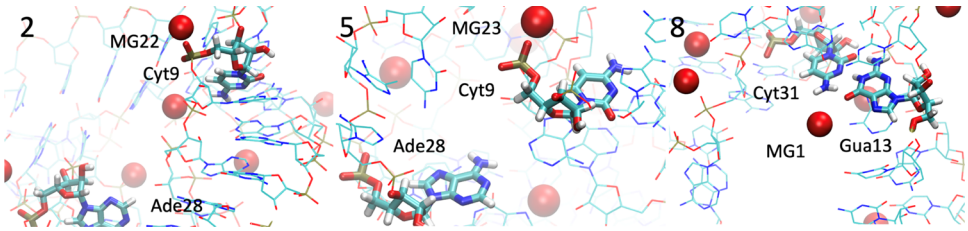


Figure 11. Local regions around bases Cyt9 and Ade28 in systems 2 and 5 and around Cyt31 and Gua13 in system 8 at 1 μs of the simulations. Mg^{2+} ions are shown as red spheres, and nucleic acid residues are represented as atom-colored lines with Cyt9, Ade28, Cyt31, and Gua13 shown in stick format.

shell coordination to A28 at the beginning of the simulation. Such an interaction is not observed in the crystal structure. Interestingly, the phosphate of C9 in system 2 also has direct interactions with Mg^{2+} and there is substantial interaction between Mg^{2+} and C9 in system 8 (Figures 8 and 9). An image of the Mg^{2+} adjacent to the C9 phosphate is shown in Figure 11 at 1 μs for systems 2 and 5. Direct coordination of the C9 phosphate is evident though the ion is no longer between C9 and A28 in system 5. The C9-A28 distance was reported to be one of the metrics correlated with the tertiary structure change in Twister ribozyme;[32] the present analysis further shows that it is a region that can stabilize the Twister ribozyme not identified in the crystal structure. For the T2 contacts, systems 2, 3, and 8 are more stable compared to others as a result of Mg^{2+} ion interactions with neighboring phosphate groups or nucleobases. As can be seen from the heat maps, system 2 has MG29 located near C32 and interacting with the base of A29, thereby stabilizing the T2 region. This interaction is stable throughout the simulation as shown in Figure 12. Similarly, system 8 has MG1 located near C31 and G13 thereby stabilizing T2 via interactions with N4 and O6, respectively, as shown in Figure 11. In contrast to systems 2 and 8, system 3 has MG16 bridging phosphate groups of C21 and G30, as seen from the heat map. The interaction with C21 is the inner shell and is stable throughout the simulation (Figure 12) and contributes to the stability of the T2 interaction (Figure 6). However, the contact with A30 is lost at ∼1 μs, which leads to dissociation of the T2 contact at that time (Figure 6) and appears to contribute to the subsequent further loss of the T1 and T4 contacts and full unfolding (Figures 4 and 5). The sampling of conformations close to the folded state throughout the entire simulation with system 8 appears to have important contribution from MG22 being in the vicinity of U6, U24, and C26 and, notably, that ion occupying that site for the whole simulation. However, MG22 is not observed in the crystal structure, but is an inner-shell cation bridging U6 and C25 and is at a highly stable position inside that region, hindering its ability to migrate to another location or away from the RNA (Figure 12).

CONCLUSIONS

We used oscillating μ_{ex} GCMC to generate different ion distributions in nine systems that are composed of the Twister RNA structure, and 100 mM $MgCl_2$ with additional neutralizing K^+ and Cl^- ions. While diverse distributions were achieved Mg^{2+} ion locations consistent with those observed in the crystal structure were identified showing that the GCMC method can successfully find these locations, consistent with our previous studies using the GCMC method to map known ion locations around RNA.[32,54] For each system, we then generated system-specific reaction coordinates using generative deep learning that were applied in metadynamics to bias RNA molecules to explore unfolded tertiary conformations. From the simulations, the majority of the systems fully unfolded with system 8 sampling conformations close to the folded state throughout the 6 μs simulation. Analysis of Mg^{2+} ion-nucleotide interactions in the biased trajectories showed the relationship between Mg^{2+} location to the stability of the RNA tertiary structures. Mg^{2+} ions directly contribute to the stability of tertiary contacts for systems that are relatively more stable compared to others. These contributions occur through interactions with specific nucleotides and can occur as outer-shell cations or inner-shell cations, including the bridging of phosphate groups, consistent with previous computational studies using both continuum[13,27,55,56] and explicit solvent[23,57−59] simulations. The lack of specific interactions allows for the rapid unfolding of systems 1, 4, 6, 7, and 9. However, while specific interactions

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

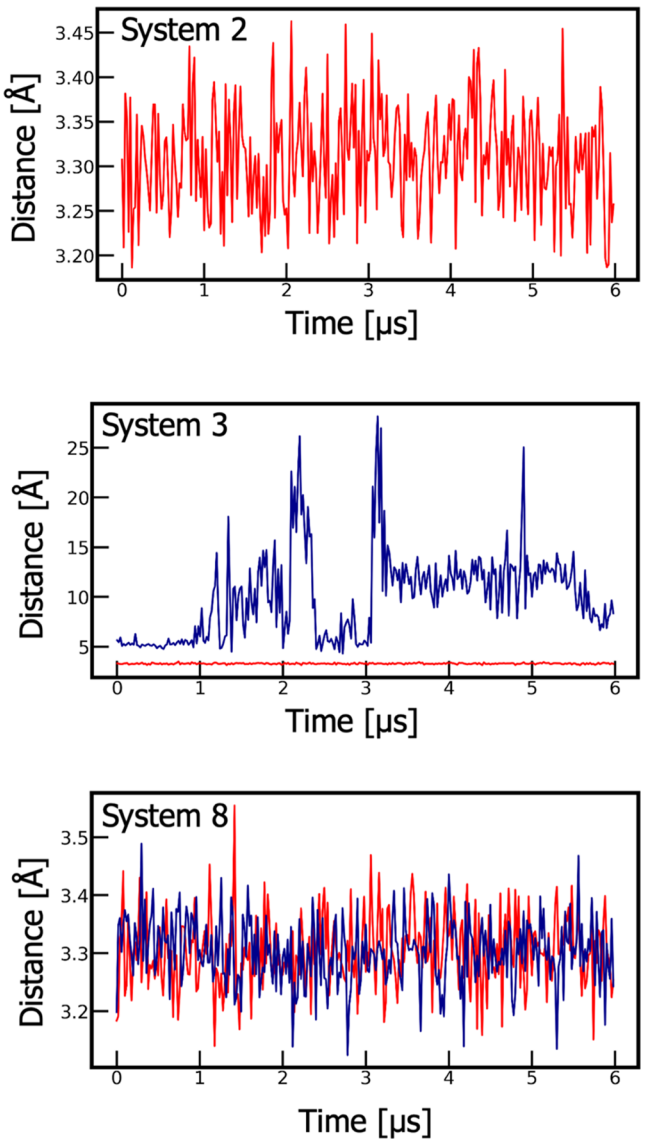


Figure 12. Selected P−Mg^{2+} distances for systems 2, 3, and 8. System 2 shows the M29-C32(P) distance, system 3 shows the inner-shell MG16-C21 (red) and outer-shell MG16-A30 (blue) distances, and system 8 includes the distance between MG22 and U6 (red) and U25 (blue).

lead to stabilization, it is essential that multiple, specific sites on the RNA be occupied to maintain the overall conformations close to the folded states as occurs with system 8. In addition, it is observed that specific Mg^{2+}-phosphate interactions can lead to destabilization of the local RNA structure as described in our previous study,[32] as well as in a recent study by Roy and co-workers.[60] Such destabilization may contribute to RNA overcoming local kinetic barriers during folding. For example, increased Mg^{2+} was shown to lead to an increase in accessing the folded state associated with self-cleavage over alternate catalytically inactive folded states as observed in photobleaching/FRET experiments on twister.[19] This is consistent with MD simulations studies indicating the need for rearrangement of the conformation of Twister to assume the catalytically active conformation.[61,62] More studies are required to address this possibility.

System 8 never undergoes full unfolding, including all major tertiary contacts. The T1 contact has a partial unfolding at the beginning of the simulation but then refolds. The T4 contact was stable during the whole simulation. The reason for this stability is MG22, which bridges T4 contact nucleotides, and, as it is an inner-shell cation, it is highly stable. In fact, it contributes to the reversal of the unfolding of T4 back to the folded state. The T2 contact in system 8 would likely dissociate as there are limited ions in its vicinity; however, due to the increased stability of the T1 and T4 contact sites and an additional contact near C31 and G13 close to T2 contributes to the Twister ribozyme never fully unfolding.

The majority of Mg^{2+} locations that contribute to stabilization are in the vicinity of those in the crystal structure. These include the direct coordination of Mg^{2+} ions with U6, U24, and C26. However, using the GCMC sampling, we identified additional Mg^{2+} locations that contribute to stability. There is a Mg^{2+} ion located between U6 and C25 that is proposed to increase the stability of the RNA in system 8 (Figure 12). In addition, interactions of Mg^{2+} with C9 were observed in three of the more stable systems and appear to contribute to the stability of the RNA (Figures 8, 9, and 11). Therefore, we suggest these previously unidentified Mg^{2+} ion locations are essential for the stabilization of the Twister ribozyme.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c00931.

2D distributions of the sampling of tertiary contacts (PDF)

Coordinates and scripts to perform the standard MD, AMINO/RAVE generative deep learning and metadynamics simulations (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Alexander D. MacKerell, Jr. − *Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, United States;* orcid.org/0000-0001-8287-6804; Email: alex@outerbanks.umaryland.edu

Author

Mert Y. Sengul − *Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, United States;* orcid.org/0000-0002-5309-0316

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c00931

Author Contributions

M.Y.S. and A.D.M. came up with the idea. The molecular dynamics simulations were run by M.Y.S. The analysis codes are written by M.Y.S. The analyses were conducted by M.Y.S. and A.D.M. The manuscript was written by M.Y.S. and A.D.M. All authors have given approval to the final version of the manuscript.

Notes

The authors declare the following competing financial interest(s): A.D.M. is co-founder and CSO of SilcsBio LLC.

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

■ ACKNOWLEDGMENTS

The authors acknowledge funding support from the National Institutes of Health under Award No. GM131710. They appreciate Dr. Abhishek Kognole for helpful discussions.

■ REFERENCES

(1) Kampa, D.; Cheng, J.; Kapranov, P.; Yamanaka, M.; Brubaker, S.; Cawley, S.; Drenkow, J.; Piccolboni, A.; Bekiranov, S.; Helt, G.; Tammana, H.; Gingeras, T. R. Novel RNAs Identified from a In-Depth Analysis of the Transcriptome of Human Chromosomes 21 and 22. *Genome Res.* **2004**, *14*, 331−342.
(2) Disney, M. D. Targeting RNA with Small Molecules To Capture Opportunities at the Intersection of Chemistry, Biology, and Medicine. *J. Am. Chem. Soc.* **2019**, *141*, 6776−6790.
(3) Gupta, P.; Khadake, R. M.; Panja, S.; Shinde, K.; Rode, A. B. Alternative RNA Conformations: Companion or Combatant. *Genes* **2022**, *13*, 1930.
(4) Mlýnský, V.; Janeček, M.; Kührová, P.; Fröhlking, T.; Otyepka, M.; Bussi, G.; Banáš, P.; Šponer, J. Toward Convergence in Folding Simulations of RNA Tetraloops: Comparison of Enhanced Sampling Techniques and Effects of Force Field Modifications. *J. Chem. Theory Comput.* **2022**, *18*, 2642−2656.
(5) Woodson, S. A. Compact Intermediates in RNA Folding. *Annu. Rev. Biophys.* **2010**, *39*, 61−77.
(6) Draper, D. E. A Guide to Ions and RNA Structure. *RNA* **2004**, *10*, 335−343.
(7) Pyle, A. M. Metal Ions in the Structure and Function of RNA. *J. Biol. Inorg. Chem.* **2002**, *7*, 679−690.
(8) Draper, D. E. Reflections on 20 Years of RNA. *RNA* **2015**, *21*, 601−602.
(9) Thirumalai, D.; Lee, N.; Woodson, S. A.; Klimov, D. K. Early Events in RNA Folding. *Annu. Rev. Phys. Chem.* **2001**, *52*, 751−762.
(10) Draper, D. E. RNA Folding: Thermodynamic and Molecular Descriptions of the Roles of Ions. *Biophys. J.* **2008**, *95*, 5489−5495.
(11) Giambaşu, G. M.; Case, D. A.; York, D. M. Predicting Site-Binding Modes of Ions and Water to Nucleic Acids Using Molecular Solvation Theory. *J. Am. Chem. Soc.* **2019**, *141*, No. 2435.
(12) Ucisik, M. N.; Bevilacqua, P. C.; Hammes-Schiffer, S. Molecular Dynamics Study of Twister Ribozyme: Role of Mg2+ Ions and the Hydrogen-Bonding Network in the Active Site. *Biochemistry* **2016**, *55*, 3834−3846.
(13) Denesyuk, N. A.; Thirumalai, D. How Do Metal Ions Direct Ribozyme Folding? *Nat. Chem.* **2015**, *7*, 793−801.
(14) Auffinger, P.; Bielecki, L.; Westhof, E. The Mg^{2+} Binding Sites of the 5S RRNA Loop E Motif as Investigated by Molecular Dynamics Simulations. *Chem. Biol.* **2003**, *10*, 551−561.
(15) Kirmizialtin, S.; Pabit, S. A.; Meisburger, S. P.; Pollack, L.; Elber, R. RNA and Its Ionic Cloud: Solution Scattering Experiments and Atomically Detailed Simulations. *Biophys. J.* **2012**, *102*, 819−828.
(16) Cunha, R. A.; Bussi, G. Unraveling Mg++-RNA Binding with Atomistic Molecular Dynamics. *RNA* **2017**, *23*, 628−638.
(17) Gasser, C.; Gebetsberger, J.; Gebetsberger, M.; Micura, R. SHAPE Probing Pictures Mg2+-Dependent Folding of Small Self-Cleaving Ribozymes. *Nucleic Acids Re.s* **2018**, *46*, 6983−6995.
(18) Trachman, R. J.; Draper, D. E. Divalent Ion Competition Reveals Reorganization of an RNA Ion Atmosphere upon Folding. *Nucleic Acids Res.* **2017**, *45*, 4733−4742.
(19) Korman, A.; Sun, H.; Hua, B.; Yang, H.; Capilato, J. N.; Paul, R.; Panja, S.; Ha, T.; Greenberg, M. M.; Woodson, S. A. Light-Controlled Twister Ribozyme with Single-Molecule Detection Resolves RNA Function in Time and Space. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 12080−12086.
(20) Leipply, D.; Draper, D. E. Evidence for a Thermodynamically Distinct Mg 2+ Ion Associated with Formation of an RNA Tertiary Structure. *J. Am. Chem. Soc.* **2011**, *133*, 13397−13405.
(21) Vušurović, N.; Altman, R. B.; Terry, D. S.; Micura, R.; Blanchard, S. C. Pseudoknot Formation Seeds the Twister Ribozyme Cleavage Reaction Coordinate. *J. Am. Chem. Soc.* **2017**, *139*, 8186−8193.
(22) Yu, T.; Chen, S. J. Hexahydrated Mg2+ Binding and Outer-Shell Dehydration on RNA Surface. *Biophys. J.* **2018**, *114*, 1274−1284.
(23) Fingerhut, B. P.; Schauss, J.; Kundu, A.; Elsaesser, T. Contact Pairs of RNA with Magnesium Ions-Electrostatics beyond the Poisson-Boltzmann Equation. *Biophys. J.* **2021**, *120*, 5322−5332.
(24) Zhou, Y.; Chen, S. J. Graph Deep Learning Locates Magnesium Ions in RNA. *QRB Discov* **2022**, *3*, E20.
(25) Gasser, C.; Gebetsberger, J.; Gebetsberger, M.; Micura, R. SHAPE Probing Pictures Mg2+-Dependent Folding of Small Self-Cleaving Ribozymes. *Nucleic Acids Res.* **2018**, *46*, 6983−6995.
(26) Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 1493−1502.
(27) Hori, N.; Denesyuk, N. A.; Thirumalai, D. Ion Condensation onto Ribozyme Is Site Specific and Fold Dependent. *Biophys. J.* **2019**, *116*, 2400−2410.
(28) Hayes, R. L.; Noel, J. K.; Mandic, A.; Whitford, P. C.; Sanbonmatsu, K. Y.; Mohanty, U.; Onuchic, J. N. Generalized Manning Condensation Model Captures the RNA Ion Atmosphere. *Phys. Rev. Lett.* **2015**, *114*, 258105.
(29) Panja, S.; Hua, B.; Zegarra, D.; Ha, T.; Woodson, S. A. Metals Induce Transient Folding and Activation of the Twister Ribozyme. *Nat. Chem. Biol.* **2017**, *13*, 1109−1114.
(30) Roth, A.; Weinberg, Z.; Chen, A. G. Y.; Kim, P. B.; Ames, T. D.; Breaker, R. R. A Widespread Self-Cleaving Ribozyme Class Is Revealed by Bioinformatics. *Nat. Chem. Biol.* **2014**, *10*, 56−60.
(31) Liu, Y.; Wilson, T. J.; McPhee, S. A.; Lilley, D. M. J. Crystal Structure and Mechanistic Investigation of the Twister Ribozyme. *Nat. Chem. Biol.* **2014**, *10*, 739−744.
(32) Kognole, A. A.; MacKerell, A. D. Mg2+ Impacts the Twister Ribozyme through Push-Pull Stabilization of Nonsequential Phosphate Pairs. *Biophys. J.* **2020**, *118*, 1424−1437.
(33) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(34) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **2012**, *8*, 348−362.
(35) Beglov, D.; Roux, B. Finite Representation of an Infinite Bulk System: Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys.* **1994**, *100*, 9050−9063.
(36) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; MacKerell, A. D. 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.* **2011**, *32*, 1929−1943.
(37) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.
(38) Kognole, A. A.; Mackerell, A. D. Contributions and Competition of Mg 2+ and K + in Folding and Stabilization of the Twister Ribozyme. *RNA* **2020**, *26*, 1704−1715.
(39) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(40) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-Induced Forces between Two Hydrophilic Groups. *J. Phys. Chem* **1994**, *98*, 2198−2202.
(41) Lakkaraju, S. K.; Raman, E. P.; Yu, W.; Mackerell, A. D. Sampling of Organic Solutes in Aqueous and Heterogeneous Environments Using Oscillating Excess Chemical Potentials in

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546

Grand Canonical-like Monte Carlo-Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 2281−2290.
(42) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(43) Izaguirre, J. A.; Sweet, C. R.; Pande, V. S. Multiscale Dynamics of Macromolecules Using Normal Mode Langevin. *Pac. Symp. Biocomput.* **2010**, 240−251.
(44) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput Biol.* **2017**, *13*, No. e1005659.
(45) Sun, D.; Lakkaraju, S. K.; Jo, S.; Mackerell, A. D. Determination of Ionic Hydration Free Energies with Grand Canonical Monte Carlo/Molecular Dynamics Simulations in Explicit Water. *J. Chem. Theory Comput.* **2018**, *14*, 5290−5302.
(46) Ravindra, P.; Smith, Z.; Tiwary, P. Automatic Mutual Information Noise Omission (AMINO): Generating Order Parameters for Molecular Systems. *Mol. Syst. Des. Eng.* **2020**, *5*, 339−348.
(47) Ribeiro, J. M. L.; Bravo, P.; Wang, Y.; Tiwary, P. Reweighted Autoencoded Variational Bayes for Enhanced Sampling (RAVE). *J. Chem. Phys.* **2018**, *149*, No. 072301.
(48) Kingma, D. P.; Welling, M. Auto-Encoding Variational Bayes. 2013, arXiv:1312.6114. arXiv.org e-Print archive. https://arxiv.org/abs/1312.6114 (accessed Dec 10, 2022).
(49) Clevert, D.-A.; Unterthiner, T.; Hochreiter, S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). 2015, arXiv:1511.07289. arXiv.org e-Print archive. https://arxiv.org/abs/1511.07289 (accessed Feb 22, 2016).
(50) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. 2014, arXiv:1412.6980. arXiv.org e-Print archive. https://arxiv.org/abs/1412.6980 (accessed Jan 30, 2017).
(51) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(52) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603.
(53) Whitford, P. C.; Schug, A.; Saunders, J.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y. Nonlocal Helix Formation Is Key to Understanding S-Adenosylmethionine-1 Riboswitch Function. *Biophys. J.* **2009**, *96*, L7−L9.
(54) Lemkul, J. A.; Lakkaraju, S. K.; MacKerell, A. D. Characterization of Mg2+ Distributions around RNA in Solution. *ACS Omega* **2016**, *1*, 680−688.
(55) Roy, S.; Onuchic, J. N.; Sanbonmatsu, K. Y. Cooperation between Magnesium and Metabolite Controls Collapse of the SAM-I Riboswitch. *Biophys. J.* **2017**, *113*, 348−359.
(56) Hayes, R. L.; Noel, J. K.; Whitford, P. C.; Mohanty, U.; Sanbonmatsu, K. Y.; Onuchic, J. N. Reduced Model Captures Mg2+-RNA Interaction Free Energy of Riboswitches. *Biophys. J.* **2014**, *106*, 1508−1519.
(57) Cunha, R. A.; Bussi, G. Unraveling Mg 2+-RNA Binding with Atomistic Molecular Dynamics. *RNA* **2017**, *23*, 628−638.
(58) Sarkar, R.; Jaiswar, A.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y.; Roy, S. Chelated Magnesium Logic Gate Regulates Riboswitch Pseudoknot Formation. *J. Phys. Chem. B* **2021**, *125*, 6479−6490.
(59) Halder, A.; Kumar, S.; Valsson, O.; Reddy, G. Mg2+Sensing by an RNA Fragment: Role of Mg^{2+}-Coordinated Water Molecules. *J. Chem. Theory Comput.* **2020**, *16*, 6702−6715.
(60) Ramachandran, V.; Mainan, A.; Roy, S. Dynamic Effects of the Spine of Hydrated Magnesium on Viral RNA Pseudoknot Structure. *Phys. Chem. Chem. Phys.* **2022**, *24*, 24570−24581.
(61) Gaines, C. S.; York, D. M. Ribozyme Catalysis with a Twist: Active State of the Twister Ribozyme in Solution Predicted from Molecular Simulation. *J. Am. Chem. Soc.* **2016**, *138*, 3058−3065.
(62) Gaines, C. S.; Giese, T. J.; York, D. M. Cleaning up Mechanistic Debris Generated by Twister Ribozymes Using Computational RNA Enzymology. *ACS Catal.* **2019**, *9*, 5803−5815.

https://doi.org/10.1021/acsomega.3c00931
ACS Omega 2023, 8, 19532−19546