THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Mechanism of Fluoride Ion Encapsulation by Magnesium Ions in a Bacterial Riboswitch

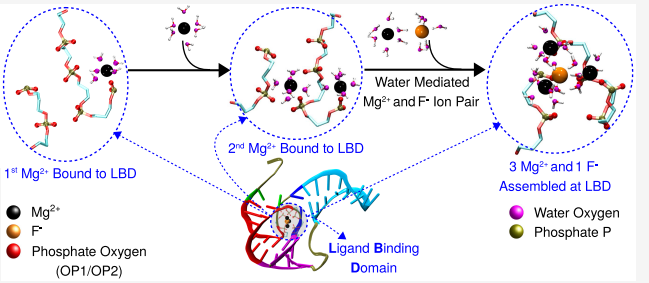
Sunil Kumar and Govardhan Reddy*

Cite This: *J. Phys. Chem. B* 2023, 127, 9267−9281

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Riboswitches sense various ions in bacteria and activate gene expression to synthesize proteins that help maintain ion homeostasis. The crystal structure of the aptamer domain (AD) of the fluoride riboswitch shows that the F^- ion is encapsulated by three Mg^{2+} ions bound to the ligand-binding domain (LBD) located at the core of the AD. The assembly mechanism of this intricate structure is unknown. To this end, we performed computer simulations using coarse-grained and all-atom RNA models to bridge multiple time scales involved in riboswitch folding and ion binding. We show that F^- encapsulation by the Mg^{2+} ions bound to the riboswitch involves multiple sequential steps. Broadly, two Mg^{2+} ions initially interact with the phosphate groups of the LBD using water-mediated outer-shell coordination and transition to a direct inner-shell interaction through dehydration to strengthen their interaction with the LBD. We propose that the efficient binding mode of the third Mg^{2+} and F^- is that they form a water-mediated ion pair and bind to the LBD simultaneously to minimize the electrostatic repulsion between three Mg^{2+} bound to the LBD. The tertiary stacking interactions among the LBD nucleobases alone are insufficient to stabilize the alignment of the phosphate groups to facilitate Mg^{2+} binding. We show that the stability of the whole assembly is an intricate balance of the interactions among the five phosphate groups, three Mg^{2+}, and the encapsulated F^- ion aided by the Mg^{2+} solvated water. These insights are helpful in the rational design of RNA-based ion sensors and fast-switching logic gates.



INTRODUCTION

Riboswitches are noncoding RNAs that modulate gene expression in response to ligand binding. Cations such as Mg^{2+}, Na^+, and K^+ condense on the riboswitch to counterbalance the negative backbone charge and stabilize the folded conformations relevant to its function. An anionic ligand has to breach the positively charged cloud of cations in the RNA vicinity to bind to the RNA. Further, the polyanionic nature of RNA excludes anions from its ionic atmosphere,[1−3] which makes the study of anionic ligand binding to RNA interesting and challenging. There are riboswitches such as TPP,[4] GlmS,[5] and FMN[6] where the cognate ligands are large anionic bio-organic molecules, and their binding is mediated through the Mg^{2+} electrostatic interaction and a large number of other stabilizing interactions involving dispersions and dipole and hydrogen bonds. However, a fluoride-sensing riboswitch is an exciting example of the RNA−anion interaction where the smallest biologically relevant monatomic anion, F^- ion, acts as a cognate ligand and binds to RNA through the cationic pocket composed of the biologically relevant smallest divalent cation, Mg^{2+} (Figure 1A,B), only through electrostatic interactions. The fluoride riboswitch aptamer domain (AD) has unparalleled selectivity toward only F^- among all of the halide anions.[7] Understanding the basic principles that dictate the high selectivity of riboswitches in binding their cognate ligands and the thermodynamic energy barriers involved in the structural transitions can aid in the rational design of nucleotide-based sensors and logic gates, which switch efficiently, enabling the reality of practical devices using these gates.[8,9]

We studied the mechanism of AD folding and F^- encapsulation by the cationic pocket composed of three Mg^{2+} ions, which are bound to the ligand-binding domain (LBD) located at the core of the AD taken from the *crcB* gene of *Thermotoga petrophila*. The AD is 52 nt long, and in the native state (holo form), it has the following tertiary structural elements (TSE)—helical motifs stem-1 (S_1) and stem-2 (S_2) joined by an internal loop that interacts with the 5′ nucleotides to form a pseudoknot (PK) (Figure 1A,B). To avoid confusion when we refer to the various phosphate groups, we used S_1 and S_2 to refer to the helices instead of P_1 and P_2, which is customary. The LBD comprises phosphate groups of five nucleotides, A6, U7, G8, U41, and G42 (LBD nucleotides),

Received: June 11, 2023
Revised: September 4, 2023
Published: October 18, 2023

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

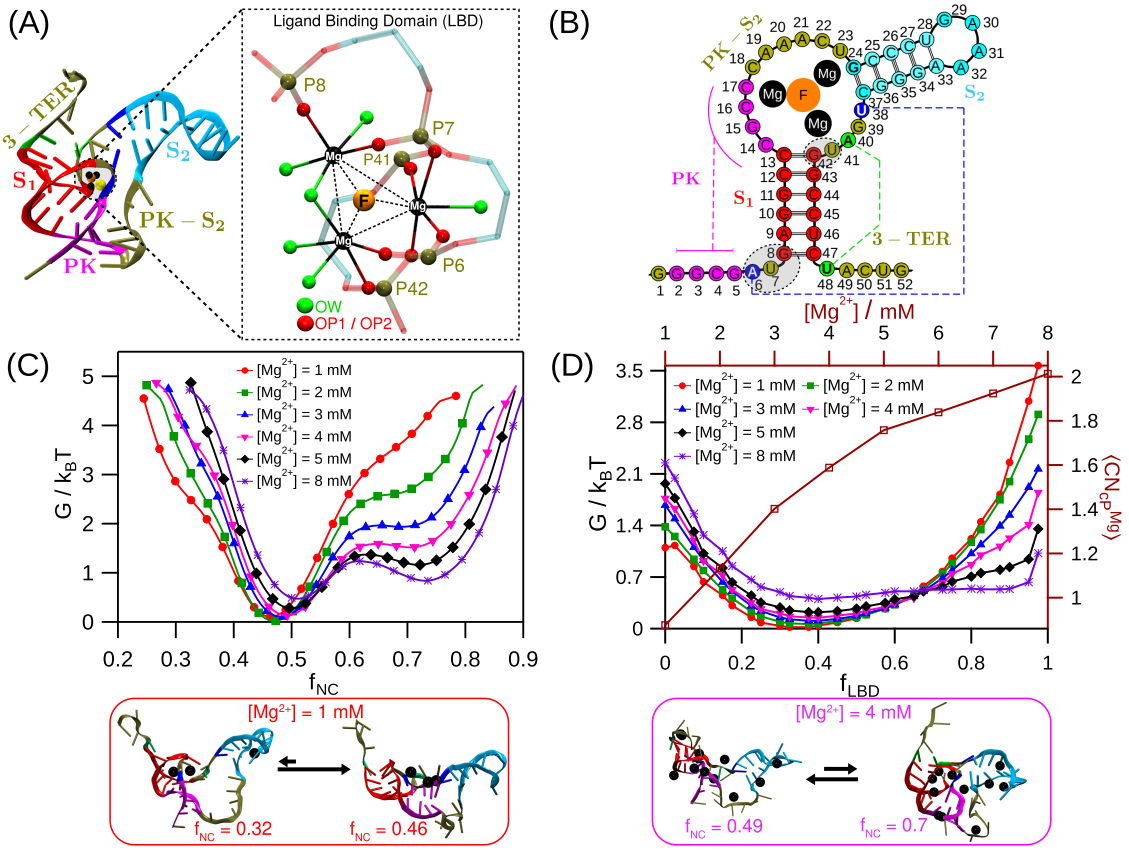


Figure 1. (A) Crystal structure of the fluoride aptamer domain (AD) (PDB: 4ENC)[7] shown in cartoon representation. The tertiary structural elements pseudoknot (PK), helical stem-1 (S_1), and stem-2 (S_2) are shown in magenta, red, and cyan, respectively. The linchpin hydrogen bond pairs A6/U38 and A40/U48 are shown in blue and green cartoon representations, respectively. The nucleotides located between PK and S_2 (referred to as PK-S_2) and at the 3′ terminal (referred to as 3-TER) are shown in tan. Mg^{2+} and F^- are shown as black and orange beads, respectively. K^+ near Mg^{2+} is shown as a yellow bead. Three Mg^{2+} interact with the O atoms of the phosphate group of nucleotides A6, U7, G8, U41, and G42, shown in the gray shaded area in the schematic. The ligand-binding domain (LBD) is highlighted inside the dotted rectangle. The backbone of the five nucleotides A6, U7, G8, U41, and G42 (referred to as LBD nucleotides) are shown with a translucent stick model. The LBD-forming RNA atoms (referred to as LBD RNA atoms) include two phosphate group oxygen atoms (OP) each from A6, U7, and G42 and one OP atom each from G8 and U41. All of the LBD RNA atoms are shown as red beads. All of the octahedral coordinating sites for each of the Mg^{2+} are occupied either by the phosphate O atom (shown as red beads) or the water O atom (shown as green beads). The P atoms of the LBD nucleotides are shown as tan beads. Three Mg^{2+} that are bound to LBD (referred to as LBD-3Mg) are located at the octahedral holes created by the LBD RNA atoms and water molecules. LBD-3Mg constitutes the trigonal base of the cationic pocket, and F^- occupies the fourth vertex of the trigonal pyramidal unit. (B) Two-dimensional structure of the AD is shown using circle and stick representation. The nucleotide regions are marked using a similar coloring scheme as mentioned before. The linchpin hydrogen bond pairs A6/U38 and A40/U48 are shown in blue and green circles, respectively. (C) Free energy surface (FES) of the AD is projected onto the fraction of native contacts, f_{NC}. The AD populates two major basins corresponding to the dominant unfolded and folded states. A sparsely populated unfolded state is observed in low $[Mg^{2+}]$. Representative conformation of the AD in each basin of the FES is shown in a cartoon representation. The equilibrium between conformations in different basins is shown for $[Mg^{2+}]$ = 1 and 4 mM. (D) FES of the AD (left axis) is projected onto the fraction of LBD native contacts, f_{LBD} (bottom axis), for different $[Mg^{2+}]$. The coordination number (right axis) of Mg^{2+} with respect to the LBD nucleotide phosphate beads, $\langle CN_{cP}^{LBD} \rangle$ as a function of $[Mg^{2+}]$ (top axis).

pointing toward the core of the AD. Three Mg^{2+} bind to the LBD through these five phosphate groups (Figure 1A) and create a cationic pocket where F^- binds with high specificity. AD has six tertiary stacking interactions (G5/G42, A6/G24, A6/G39, U7/G39, G8/A40, and A40/A49) that enhance the LBD's stability and hold the backbone of LBD nucleotides in a position conducive to Mg^{2+} and F^- binding (Figure S1).

It is challenging to simultaneously study riboswitch folding and ligand binding, as they span multiple time scales. As a result, an experimental technique can probe only a specific aspect of riboswitch folding and ligand binding. Experiments[10−23] and computer simulations[24−31] probed the effect of various cations and cognate ligands on the folding kinetics and thermodynamics of riboswitches. The ion-counting experiments show that cations preferentially accumulate, and anions are excluded from the ionic atmosphere of RNA.[1−3] The SHAPE-seq,[32] CEST-NMR,[33] and smFRET[34] studies on the AD of the fluoride-sensing riboswitch from *Bacillus cereus crcB* showed that in the absence of F^- binding at the LBD, the AD populates both the *holo*-like and unfolded states. The population of the *holo*-like folded state in the absence of F^- binding suggests that the AD follows a conformational selection mechanism. To dominantly populate the folded state, F^- binding at the LBD of the AD is essential. The CEST-NMR study[33] further suggested that in the *holo*-like conformations sparsely populated in the absence of F^- binding, the "linchpin" hydrogen bonds, which enable the terminator helix formation and signals the transcription termination, are

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

ruptured (Figure 1A,B). The smFRET[34] study demonstrated that the "linchpin" hydrogen bond between nucleotides A40/U48 forms and ruptures upon F^- ion binding and unbinding, respectively. The SHAPE-seq[32] and CEST-NMR[33] studies identified that the nucleotides located in between the pseudoknot (PK) and the helix S_2 and the 3′ tail undergo structural changes when the AD transitions between the unfolded and folded states (Figure 1A,B). In this study, using computer simulations and multiple RNA models, we probed the AD folding thermodynamics and probed the assembly mechanism of the LBD and F^- encapsulation by the three Mg^{2+} ions bound to the LBD, complementing the existing experiments.[32−34]

A single experimental technique cannot simultaneously track the folding of various secondary and tertiary structural elements in RNA and probe the contribution of multiple ions to the stability of these structures. Molecular dynamics (MD) simulations can simultaneously study the ion assembly and RNA folding. However, it is not trivial even for MD simulations to probe the folding and ion assembly in systems such as a fluoride riboswitch, as multiple time scales ranging from nanoseconds to seconds are involved. To overcome this problem in MD simulations, we studied the AD folding and F^- binding using both coarse-grained and all-atom models of RNA. Using coarse-grained models, we show that in the presence of Mg^{2+}, AD populates *holo* form-like structures with similar tertiary structural elements irrespective of its cognate ligand, F^- binding. The tertiary stacking interactions alone cannot stabilize the orientation of the five LBD phosphate groups without Mg^{2+} binding. The LBD region has a high affinity for Mg^{2+} binding and can accommodate two Mg^{2+} ions without F^- ion encapsulation. Using all-atom simulations, we show that three Mg^{2+} bind to the LBD sequentially by computing the Mg^{2+} binding lifetime at the LBD. Each bound Mg^{2+} should transition from a water-mediated outer-shell binding to an inner-shell binding to strengthen its interaction with the LBD. The third Mg^{2+} and F^- approach the LBD as a water-mediated ion pair for efficient binding, which allows charge neutrality in the whole assembly comprising the five phosphate groups, three Mg^{2+}, and encapsulated F^-. In agreement with the previous results, we also show that the linchpin hydrogen bond between A40 and U48, which plays a role in the formation of the terminator helix, preventing transcription, is unstable in the absence of F^- binding.

METHODS

We studied the effect of Mg^{2+} on the folding of the fluoride riboswitch aptamer domain (AD) from the *crcB* gene of *Thermotoga petrophila* using the coarse-grained three interaction site (TIS) RNA model[35,36] and Langevin dynamics simulations. We performed all-atom molecular dynamics simulations to understand the mechanism of Mg^{2+}-mediated F^- sensing by the AD.

Simulations. *TIS Model of RNA.* The TIS model for the AD is constructed using the crystal structure (PDB: 4ENC).[7] In the TIS model,[35,36] each nucleotide is represented using three sites representing the phosphate, sugar, and base groups. The positions of the phosphate, sugar, and base sites are the centers of mass of their respective chemical groups. The monovalent (K^+ and F^-) and divalent (Mg^{2+}) ions are explicitly present in the simulation. The Hamiltonian for the TIS model is given by

$$U_{TIS} = U_{BL} + U_{BA} + U_{EV} + U_{EL} + U_{ST} + U_{HB} + U_{TST} \quad (1)$$

The different components in the Hamiltonian correspond to the bond length (U_{BL}), the bond angle (U_{BA}), excluded volume repulsion between different sites (U_{EV}), the electrostatic interaction between charged sites (U_{EL}), the single-strand base stacking interaction between consecutive bases (U_{ST}), the hydrogen bonding interaction (U_{HB}), and the tertiary stacking interaction between two nonconsecutive bases (U_{TST}). The force field details and parameters for RNA sites are described in detail in the Supporting Information (SI) and can be found in the works of Denesyuk and Thirumalai.[35,36]

The list of base pairs involved in the native hydrogen bond network present in the folded structure of the AD (Tables S1b and S2) is obtained using the crystal structure (PDB: 4ENC) and WHAT IF web server (https://swift.cmbi.umcn.nl).[37] The base pairs with a hydrogen bond between them interact by using the hydrogen bonding potential (U_{HB}). The model takes into account the tertiary base stacking interactions present in the crystal structure (Table S3). The TIS RNA model allows the formation of non-native canonical hydrogen bonds but not the formation of non-native noncanonical hydrogen bonds.[36] In this study, we have not allowed the formation of any non-native hydrogen bonds. This model is successful in quantitatively accounting for the folding thermodynamics of multiple RNA systems,[30,36,38,39] demonstrating that it is reliable and transferable to study various other RNA systems.

Coarse-Grained Simulations Using the TIS RNA Model. Coarse-grained simulations using the TIS RNA model are performed to study the role of Mg^{2+} in the folding of AD as $[Mg^{2+}]$ is varied from 1 to 8 mM. The $[K^+]$ is fixed at 30 mM in all of the simulations. The simulations are performed in a cubic box of length 165 Å. The number of Mg^{2+} and K^+ ions in the simulation box is computed using their concentration and box volume. Only the cognate ligand for the AD, the F^- ions, are added to maintain charge neutrality in the simulation box. $[F^-]$ increases from 13 to 27 mM with the increase in $[Mg^{2+}]$ from 1 to 8 mM. All ions are explicitly modeled as beads with charge and excluded volume. The excluded volume force field parameters for K^+, F^-, and Mg^{2+} ions are taken from the modified AMBER ff14 force field.[40] SHAPE-seq experiments[32] were performed at $[Mg^{2+}]$ = 5 mM and $[K^+]$ = 50 mM. CEST-NMR and transcription assay experiments[33] used Mg^{2+} in a concentration range of 1−30 mM and $[K^+]$ = 50 mM. SmFRET experiments[34] were conducted at $[Mg^{2+}]$ = 1 to 10 mM and $[K^+]$ = 100 mM. The ion concentrations we used are similar to those used in the experiments.

We used Langevin dynamics simulations to study the folding dynamics of the AD at a temperature of T = 310 K. To efficiently compute the AD folding thermodynamic properties, we used 5% viscosity of water, $\eta = 5 \times 10^{-5}$ Pa·s so that the AD conformational sampling rate is enhanced. The equation of motion for the RNA sites and ions in Langevin dynamics is given by

$$m\ddot{\vec{r}}_i = -\zeta\dot{\vec{r}}_i + \vec{F}_{c,i} + \vec{\Gamma}_i \quad (2)$$

where m_i is the mass (in amu), ζ_i ($= 6\pi\eta R_i$) (in amu/fs) is the friction coefficient, $\vec{r}_i$ is the position, $\vec{F}_{c,i}\left(= -\frac{\partial U_{TIS}}{\partial \vec{r}_i}\right)$ is the deterministic force, R_i is the radius (in Å) of the ith site in the system, and $\vec{\Gamma}_i$ is the random force on the i^{th} site with a white-noise spectrum. The random force autocorrelation function is

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

Table 1. Details of Various Simulation Systems

system	N_{Mg}^{cry}	N_{F}^{cry}	N_{Mg}^{bulk}	N_{F}^{bulk}	N_K	N_{Cl}	$[Mg^{2+}]$ (mM)	$[F^-]$ (mM)	$[K^+]$ (mM)	$[Cl^-]$ (mM)	cumulative simulation time (μs)
AA[ϕ] (*apo*)	0	0	0	0	51	0	0	0	139	0	≈4
AA[0,0]	0	0	20	20	55	24	55	55	150	66	≈15
AA[1,0]	1	0	19	20	55	24	55	55	150	66	≈16
AA[2,0]	2	0	18	20	55	24	55	55	150	66	≈14
AA[3,0]	3	0	17	20	55	24	55	55	150	66	≈1.5
AA[3,1] (*holo*)	3	1	17	19	55	24	55	55	150	66	≈5

N_{Mg}^{cry} and N_{F}^{cry} are the number of bound Mg^{2+} and F^- to the LBD, respectively. N_{Mg}^{bulk} and N_{F}^{bulk} are the number of Mg^{2+} and F^- in the simulation box and away from the LBD, respectively. N_K and N_{Cl} are the number of K^+ and Cl^- in the simulation box, respectively. We ran four independent trajectories for the system with the AA[0,0] configuration and one trajectory for the AA[3,0] configuration. We ran two independent simulations for the rest of the system configurations.

given by $\langle \vec{\Gamma}(t) \cdot \vec{\Gamma}(t + nh) \rangle = 2\zeta k_B T \delta_{0,n}/h$, where $n = 0, 1, \cdots$, $\delta_{0,n}$ is the Kronecker delta function and k_B is the Boltzmann constant. The Langevin equation is integrated using the velocity Verlet algorithm with a time step h (= 2.5 fs).[36,41] System coordinates are saved after every 5000 steps (τ_f = 12.5 ps) to compute the properties. The initial 0.5 μs of simulation data are ignored in computing the properties. For each $[Mg^{2+}]$, at least ≈7 μs of simulation data is collected to compute the average properties. We used VMD to generate the three-dimensional structures of the AD.[42] The atomistic coordinates of the AD are generated using the coarse-grained coordinates and TIS2AA[43] program, which uses the fragment-assembly approach[44] and energy minimization embedded in AmberTools.[45]

All-Atom Simulations. We started the all-atom simulations of the AD using the crystal structure (PDB: 4ENC).[7] We performed all-atom simulations for five different AD systems in the presence and absence of crystal-bound Mg^{2+} and F^- (Table 1).

MD simulations were performed using GROMACS-5.1.8[46,47] with the modified AMBER ff14 force field (DESRES ff)[40] for the RNA and TIP4P-D water model.[48] Initially, the AD was placed at the center of a cubical box of length 85 Å and solvated with TIP4P-D water molecules. The required number of Mg^{2+}, K^+, F^-, and Cl^- ions were added to the simulation box (Table 1). The energy of the system is minimized using the steepest descent algorithm with a maximum force tolerance of 1000.0 kJ mol^{-1} nm^{-1}. The energy-minimized system was equilibrated in the following steps. (1) The positions of RNA and ions were kept fixed using harmonic position restraints with a spring constant of 1000.0 kJ $mol^{-1}nm^{-2}$, and water molecules were allowed to equilibrate in the NVT ensemble for 10 ns. This step ensured that hydration shells were formed for all of the ions. (2) Position restraints were applied only on RNA atoms, whereas ions and water molecules were allowed to equilibrate in the NVT ensemble for 10 ns so that the hydrated ions condense around the negatively charged RNA. (3) Finally, position harmonic restraints were eliminated on all of the atoms, and the system was allowed to equilibrate in the NPT ensemble for 10 ns.

The temperature T of the system is maintained at 310 K using a modified Berendsen thermostat[49] with a relaxation time τ_T = 0.1 ps. The pressure of the system is maintained at 1 atm using the Parrinello−Rahman barostat[50] with τ_P = 2.0 ps and isothermal compressibility of water. We have used 10 Å as the cutoff for both Lennard-Jones and Coulombic interactions. Long-range Coulombic interactions were computed using the particle mesh Ewald (PME)[51] sum method with a grid space of 1.6 Å. Long-range dispersion correction for energy and pressure was also incorporated. All bonds involving hydrogen atoms were constrained using the LINCS algorithm[52] throughout the production stage of the simulation. For all of the systems except AA[0,0] and AA[3,0], we performed two independent simulations, and the cumulative simulation time is reported in Table 1. For the AA[0,0] system, we performed four independent simulations.

We chose ion parameters from DESRES ff[40] for our all-atom simulations because (a) for a bulk Mg^{2+} to bind at the LBD and achieve crystal-like assembly, it has to transition to an inner-shell interaction from an outer-shell interaction with phosphate through dehydration from its first hydration shell, making it essential to model Mg−water interactions accurately. The DESRES ff ion parameters work reasonably well with the DESRES ff RNA parameters and capture RNA−ion interactions.[53,54] (b) Due to proximity among the 3 Mg^{2+} ions at the trigonal cationic cluster in such a small space of the LBD (Figure 1A), it is necessary to use an ion parameter that accurately models the nonbonded (LJ) interactions for Mg^{2+} ions. Since the DESRES ff has the LJ parameter of Mg^{2+} ($R_{min}/2$ = 1.185 Å) similar to the van der Waals radius of Mg (= 1.180 Å),[55,56] we used it. (c) The Mg−F interaction is vital for F^- capture, so for consistency, we used the parameters for Mg^{2+} and F^- obtained through similar schemes.

The 12−6−4 nonbonded parameters[57] that incorporate an ion-induced dipole to capture the polarizability can effectively model the covalent nature of divalent metal ions (Cd, Mn, Zn, Ni, Co) and nucleobase atom N7 interactions. Since Mg^{2+} has a lesser tendency to interact with the N7 atom (except in some specific RNAs),[58] we used the default ion parameters from the DESRES ff modeled only based on ion−water interactions to study the role of water in Mg^{2+} assembly at the LBD. The modified 12−6−4 LJ parameters (m12−6−4),[59] which are revised for the balanced interaction of divalent cations (Mg^{2+}, Mn^{2+}, Zn^{2+}, and Cd^{2+}) with specific RNA atoms (N7 atom of nucleobase guanosine and adenosine and the nonbridging phosphate oxygen of dimethyl phosphate) to model metal ion migration in a specific RNA system, can have an advantage over the existing set of ion parameters to model divalent ion and RNA interactions in water. A study by Panteva et al.[60] checked the capability of different water models and Mg^{2+} ion parameters to reproduce the thermodynamics and kinetics properties of the Mg^{2+} ion in water and showed that the 12−6−4 parameter in SPC/E-type water performs best. A similar study is yet to be performed in the presence of RNA to test which among the existing RNA force fields can emulate the behavior of RNA−Mg^{2+} interactions best in water.

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

Data Analysis. *Fraction of Helix Formation.* The average fraction of helix formation for a given helix, H, is computed using the equation

$$\langle f_{HF} \rangle = \langle \frac{N^i_{HB,H}}{N^{cry}_{HB,H}} \rangle \tag{3}$$

where $N^i_{HB,H}$ and $N^{cry}_{HB,H}$ are the number of hydrogen bonds present in the helix H in ith conformation and crystal structure, respectively, and $\langle\ \rangle$ denotes the average over all of the conformations. A hydrogen bond is considered to be present if its energy is lower than the thermal energy (k_BT).[36] Similarly, the linchpin hydrogen bonds A6/U38 and A40/U48 are defined to be formed if the hydrogen bond interaction energy is less than k_BT.

Local [Mg^{2+}] Around Phosphate Sites. The local [Mg^{2+}] in the vicinity of the phosphate site of the ith nucleotide in molar units is computed[61] using the relation

$$c_i^* = \frac{1}{N_A V_c} \int_0^{r_c} \rho_i(r) 4\pi r^2\, dr \tag{4}$$

where $\rho_i(r)$ is the number density of Mg^{2+} at a distance r from the ith phosphate site, V_c is the spherical volume of the radius r_c, and N_A is Avogadro's number. The cutoff radius r_c is given by $r_c = R_{Mg} + R_P + \Delta r \approx 5$ Å where R_{Mg} (=1.185 Å) and R_P (= 2.1 Å) are the radii of Mg^{2+} and phosphate sites, respectively, and Δr is the margin distance. We used Δr = 1.7 Å to ensure that we take into account only the tightly bound local Mg^{2+} around the phosphate sites. We computed $\langle c_F^* \rangle$ using a similar equation as eq 4 with $r_c = R_P + R_{Mg} + R_F + \Delta r \approx 7.3$ Å, where R_F (=2.3 Å) is the radius of the F^- site.

Free Energy Surface (FES) Calculation. The total number of native contacts in the AD is computed using the TIS model of the folded crystal structure (PDB ID: 4ENC).[7] A pair of sites i and j in the crystal structure are defined to have a native contact between them if $|i - j| > 10$, and the distance between the sites r_{ij} is less than 15 Å (see SI Figure S2). The f_{NC} for the ith AD conformation is computed using the equation

$$f_{NC} = \frac{N^i_{nc}}{N^{cry}_{nc}} \tag{5}$$

where N^i_{nc} is the number of native contacts present in the ith conformation and N^{cry}_{nc} is the number of native contacts present in the crystal structure.

The FES corresponding to the AD folding (G) is projected onto the fraction of native contacts f_{NC}, and it is calculated using the equation

$$G = -k_B T \ln(P(f_{NC})) \tag{6}$$

where $P(f_{NC})$ is the probability distribution of f_{NC}.

To probe the role of C18−U23 nucleotides located between PK and S_2 (PK-S_2 nucleotides) in AD folding, we computed f_{PK-S_2}, a fraction of native contacts between the PK-S_2 nucleotides and their neighboring nucleotides (labeled as N_{PK-S_2}) that are located within 15 Å from PK-S_2 and at least three nucleotides away along the RNA contour in the AD crystal structure. Similarly, to probe the role of A49−G52 nucleotides located at the 3′-terminal (3-TER nucleotides), we computed $f_{3\text{-TER}}$, a fraction of native contacts between the 3-TER nucleotides and their neighboring nucleotides (labeled as $N_{3\text{-TER}}$) that are located within 15 Å from 3-TER and at least three nucleotides away along the RNA contour in the AD crystal structure.

Coordination Number Calculation. We computed the coordination number (CN^A_B) of particle i belonging to group A with particle j belonging to group B in an AD conformation using the COORDINATION module of the PLUMED library (v2.6.4).[62,63] CN^A_B is computed for every conformation using the following relation

$$CN^A_B = \sum_{i\in A} \sum_{j\in B} s_{ij} \tag{7}$$

where s_{ij} is a switching function and is given by

$$s_{ij} = \frac{1 - (r_{ij}/r^\circ_{ij})^6}{1 - (r_{ij}/r^\circ_{ij})^{12}} \tag{8}$$

where r_{ij} is the distance between the ith particle of group A and the jth particle of group B and r°_{ij} is the cutoff distance large enough to effectively capture the interaction between the particles i and j.

To account for the average number of Mg^{2+} bound to the LBD nucleotides in the CG simulations for a given [Mg^{2+}], we computed $\langle CN^{Mg}_{cP} \rangle$, which is the average coordination number of Mg^{2+} with the center of mass of phosphate sites of LBD nucleotides. We designated all of the Mg^{2+} present in the simulation box for a given [Mg^{2+}] as group A, and the center of mass of phosphate beads of the LBD nucleotides (A6, U7, G8, U41, and G42) as group B. The cutoff distance, r°_{ij} = 7.3 Å, which is the Bjerrum length for water at 300 K.

To track the binding of F^- to the LBD in the all-atom simulations, we computed the coordination number individually for each of the F^- ions with the LBD RNA atoms (CN^{LBD}_F). We designated the LBD RNA atoms as group A and a given F^- as group B. The cutoff distance, r°_{ij}, is 7.3 Å.

Root-Mean-Square Deviation (RMSD) and Root-Mean-Square Fluctuation (RMSF). We computed the RMSD of the conformations from the all-atom simulations with respect to the energy-minimized holo form AD crystal structure. We aligned all of the conformations with respect to the energy-minimized holo form structure of the AD before computing both RMSD and RMSF. To probe the changes in flexibility in the AD structure due to the absence of ligand binding, we computed RMSF in the N1 atom of all of the AD nucleotides, respectively (Figure S11B). Both RMSF and RMSD were computed using VMD.[42] We used the simulation data from the systems AA[3,1] and AA[ϕ] (Table 1) for the analysis of *holo* and *apo* states of AD, respectively.

Mg^{2+} Binding Lifetime. To probe the mechanism of binding of Mg^{2+} to the LBD of the AD, we computed the average conditional binding lifetime ($\tau(i, j)$) and the occupancy time (f^i_τ) of Mg^{2+} to the LBD in all-atom simulations. Mg^{2+} within the binding cutoff distance (R_b) from the phosphorus atoms (P atom) of all of the LBD nucleotides is labeled as bound to the LBD (Figure S9). The R_b is 11 Å for AA[0,0] and AA[1,0] and 10 Å for AA[2,0]. The R_b value is estimated from the mean of the average distances between P atoms of nucleotides A6/U41 and G8/G42.

Dynamical Cross-Correlation Map (DCCM). To probe dynamical cross-correlation[64] between the nucleobases, we first coarse-grained the all-atom AD conformations using the TIS-RNA bead description.[35,36] Each nucleotide is described by phosphate, sugar, and nucleobase beads, and the position of these beads is at the center of mass of their respective chemical

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

groups. The quantitative estimate for the correlated motion between a pair of beads mimicking nucleobases i and j is given by

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{(\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{1/2}} \quad (9)$$

where Δr_i is the displacement of the ith bead from its mean position. We used the MD-TASK package[65] to compute C_{ij}. For a given pair of beads i and j, C_{ij} takes values between −1.0 to 1.0. The value −1.0 represents complete anticorrelation (beads move in the opposite direction), the value 1.0 represents complete correlation (beads move in the same direction), and the value 0 represents no correlation (beads move in the perpendicular direction).

RESULTS

LBD and Tertiary Stacking Interactions are Unstable in the Absence of F^- Binding. Noncoding RNA molecules are known to fold to a unique native state to perform gene regulation. Using the coarse-grained model of RNA[35,36] and molecular dynamics simulations, we computed the free energy surface (FES) for the Mg^{2+}-driven folding of the AD (see eq 6 in Methods) to probe the role of Mg^{2+} in facilitating F^- binding (Figure 1C). In computing the FES, the cognate ligand F^- is used as a counterion to Mg^{2+} to maintain charge neutrality in the simulation box. The FES projected onto the fraction of native contacts (f_{NC}) (eq 6) shows that the aptamer is dominantly a two-state folder, as it populated two major basins, the unfolded state ($f_{NC} \approx 0.46-0.51$) and folded state ($f_{NC} \approx 0.70-0.73$). As we increased $[Mg^{2+}]$ from 1 to 8 mM, the minima of the major unfolded basin shifted from $f_{NC} \approx 0.46$ to ≈ 0.51, indicating that there is compaction in the aptamer size in the major unfolded ensemble (Figure 1C).

In low $[Mg^{2+}]$ (≤4 mM), the aptamer sampled a sparsely populated unfolded state ($f_{NC} \approx 0.32$) and a dominantly populated unfolded state ($f_{NC} \approx 0.46-0.49$) (Figure 1C). The shoulder in the FES corresponding to the sparsely populated unfolded state disappeared for $[Mg^{2+}] > 4$ mM. With the increase in $[Mg^{2+}]$ (>3 mM), the aptamer starts populating the *holo*-like folded state ($f_{NC} \approx 0.7-0.73$; see Table S4), separated by a barrier from the dominant unfolded state ($f_{NC} \approx 0.46$). The relative population of the folded state increased with the increase in $[Mg^{2+}]$ from 3 to 8 mM (Figure 1C). However, the dominant unfolded state ($f_{NC} = 0.46-0.51$) remained the global minimum, irrespective of $[Mg^{2+}]$, in the simulations using the coarse-grained model of RNA. CEST-NMR[33] and the smFRET[34] studies also demonstrated that in the presence of Mg^{2+}, the AD dynamically populates two states. One state lacks the essential tertiary contacts, and the other state is similar to the *holo* form. However, the two helices and the pseudoknot are formed in both states.

The tertiary structural elements (TSE) of the aptamer, the pseudoknot (PK), and helices (S_1 and S_2) are stable irrespective of $[Mg^{2+}]$ (Figures 1A and S3). This is in agreement with the experiment,[33] which showed that the aptamer in both the *apo* and *holo* forms have the same hydrogen bond network and similar tertiary structure (Figure S3). The stability of the TSE at low $[Mg^{2+}]$ implies that monovalent ions ($[K^+] = 30$ mM) are adequate to stabilize the native-like tertiary structure of the aptamer.

In the simulations using the coarse-grained RNA model, the dominant unfolded state of the AD is more stable compared to the folded state even in high $[Mg^{2+}]$ (Figure 1C) because we did not observe the encapsulation of F^- by three Mg^{2+} in the LBD as observed in the crystal structure[7] (Figure 1A). Due to the absence of interactions between the LBD nucleotides, bound F^-, and three Mg^{2+}, we observed that there are dynamic transitions between locally folded and unfolded states in the following regions of the aptamer: (a) the LBD domain of the aptamer (Figure 1D), (b) tertiary contacts formed by the nucleotides present between the pseudoknot PK and helix S_2, labeled as PK-S_2 nucleotides (Figures 1A,B, and S4B,C), (c) tertiary contacts formed by the nucleotides located at the 3′ terminal labeled as 3-TER nucleotides (Figures 1A,B and S4D), and (d) the six native tertiary stacking interactions present between the nucleotides: A6/G24, A6/G39, U7/G39, G8/A40, A40/A49, and G5/G42 (Figures S1 and S12A).

Nucleotides A6, U7, G8, U41, and G42 form the LBD and fold into an orientation with the phosphate groups pointing toward the cavity center, facilitating the binding of three Mg^{2+} and F^- (Figure 1A). The AD crystal structure shows that three Mg^{2+} (LBD-3Mg) form the base of the trigonal pyramidal LBD and facilitate the binding of F^- at the fourth vertex of the pyramid to form a 3Mg−1F cluster[7] (Figure 1A). The free energy for the LBD formation projected onto the fraction of native contacts between the LBD nucleotides, f_{LBD} (*see* eq S1 in the SI), shows that the aptamer populated conformations where the LBD nucleotides are in a *holo*-like orientation, which supports the cationic pocket formation for $[Mg^{2+}] \geq 4$ mM (Figure 1D; *see* Table S4). Although the population of conformations with the folded LBD increased with the increase in $[Mg^{2+}]$, the state with the partially folded LBD is the most stable. We hypothesize that a stable LBD domain requires the binding of three Mg^{2+} and an F^- (Figure 1A). However, in the simulations with the coarse-grained RNA model, we observed the binding of at most two Mg^{2+} and no F^- binding in the LBD (Figure 1D). The average coordination number of the LBD nucleotide phosphate beads with Mg^{2+}, $\langle CN_{cP}^{Mg} \rangle$ (*see* Methods), shows that the number of Mg^{2+} that bind to the LBD increased with $[Mg^{2+}]$ and approached the value 2 for $[Mg^{2+}] = 8$ mM (Figure 1D).

These observations are in agreement with the NMR experiments[33] on the AD, which has reported that the aptamer in the *apo* form without the bound F^- populates unfolded conformations where the pseudoknot PK and helices S_1 and S_2 are fully formed, but the local contacts around regions PK-S_2 and 3-TER are ruptured (see the SI for more details). These experiments further proposed that F^- binding is indispensable for the stability of the *holo*-like conformation. We did not observe stable F^- binding in the coarse-grained simulations (*see* below). As a result, local unfolding is observed in multiple regions of the aptamer, and the unfolded state is most apparent (Figure 1C).

Mg^{2+} Ions Preferentially Bind to the LBD. To understand the role of Mg^{2+} in aptamer folding, we computed the average local Mg^{2+} concentration ($\langle c_{Mg}^* \rangle$) in the aptamer vicinity at nucleotide resolution (see eq 4 in Methods)[61] using the data from coarse-grained simulations. Even at low $[Mg^{2+}]$, the ions are not randomly diffused along the RNA backbone.[30,36,38,61] Mg^{2+} preferentially binds to the phosphate beads of LBD nucleotides (A6 and G42) involved in the formation of an RNA scaffold that facilitates the F^- binding (Figure 2A). For $[Mg^{2+}] \leq 2$ mM, Mg^{2+} only binds around the nucleotides G5−G8 and U41−G43. This observation is in agreement with the experiment,[7] which demonstrated that

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

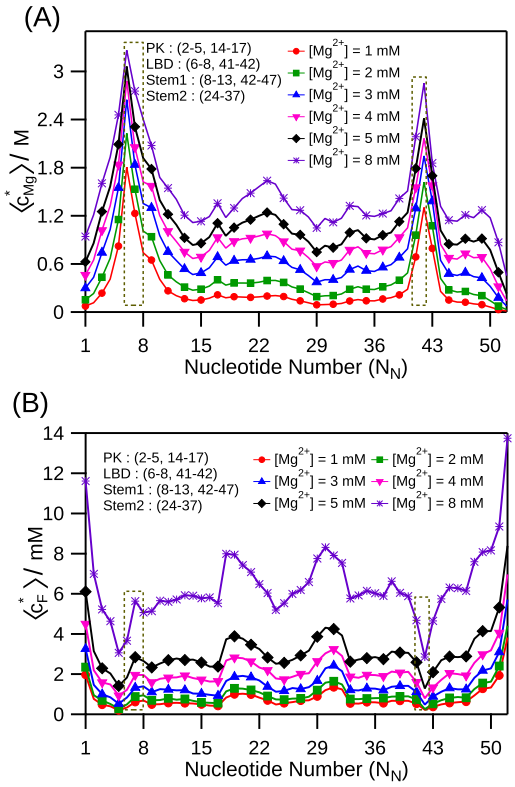


Figure 2. (A) Average local concentration of Mg^{2+} around individual phosphate sites ($\langle c^*_{Mg}\rangle$) (in molar units) is plotted as a function of the nucleotide number (N_N). The N_N range (starting from 1 as in the PDB[7]) for the pseudoknot (PK) and stems (S_1 and S_2) are in the annotation. For $[Mg^{2+}]$ = 1 mM, Mg^{2+} condenses around LBD nucleotides A6 and G42. As $[Mg^{2+}]$ increases, Mg^{2+} selectively accumulates around the same two nucleotides, indicating that Mg^{2+} preferentially binds to these two nucleotides to create the cationic LBD. (B) Average local concentration of F^- around individual phosphate sites ($\langle c^*_F\rangle$) (in millimolar units) is plotted as a function of the nucleotide number (N_N). The difference of three orders in magnitude in the local concentrations of Mg^{2+} and F^- ions is due to excess cationic accumulation of Mg^{2+} and preferential exclusion of F^- from the vicinity of RNA, respectively. The LBD nucleotides are highlighted with broken line rectangles.

Mg^{2+} binding in the vicinity of nucleotides A6−G8 and U41−G42 is essential for the cationic LBD formation. However, in the simulations using the coarse-grained RNA model, we observed at most two Mg^{2+} bind to the LBD nucleotides as indicated by the $\langle CN^{Mg}_{cP}\rangle$ values (Figure 1D).

In low $[Mg^{2+}]$ (≤2 mM), the lack of Mg^{2+} binding around the phosphate beads of PK-S_2 nucleotides (N_N = C18−U23) hinders the formation of tertiary contacts between PK-S_2 and G1−A6 nucleotides (part of the pseudoknot) due to the electrostatic repulsion (Figure S4C). Similarly, the lack of Mg^{2+} binding to the free 3-TER nucleotides (N_N = A49−G52) prevents the formation of native contacts between 3-TER and U38−U41 nucleotides, and also the formation of the linchpin hydrogen bond A40/U48 (Figures S3 and S4A). With the increase in $[Mg^{2+}]$ (≥3 mM), Mg^{2+} binds in the vicinity of the nucleotides A19−G24 (PK-S_2), G5−G8, and U41−G43, and stabilizes AD conformations with intact tertiary contacts between PK-S_2 and G1−A6 nucleotides, and local native contacts between 3-TER and U38−U41 nucleotides (Figure S4C,D). However, even for $[Mg^{2+}]$ = 8 mM, the AD conformations with the intact tertiary contacts are not the most stable state due to the absence of F^- binding to the LBD. The AD state with fully formed "linchpin" hydrogen bonding base pairs, A6/U38 and A40/U48, which stabilize the native-like state and play a role in ligand-binding signal transduction to the expression platform, is not the most stable state (Figure S3).[7,33] Although for $[Mg^{2+}] \geq 5$ mM, the AD significantly populates the folded state even in the absence of F^- binding, implying that $[Mg^{2+}]$ can drive the formation of conformations capable of initiating the transcription. The CEST-NMR and single-round transcription assay studies[33] also suggested that the *apo* form of the AD with a *holo*-like conformation can also activate transcription in the presence of $[Mg^{2+}] \geq 5$ mM.

F^- Transiently Binds to the LBD in the CG Simulations. The average local F^- concentration around the nucleotides ($\langle c^*_F\rangle$) shows that, unlike Mg^{2+}, F^- are diffused around the RNA and do not specifically bind to any particular set of nucleotides (Figure 2B). We find that the $\langle c^*_F\rangle$ value is 3 orders of magnitude less compared to the $\langle c^*_{Mg}\rangle$ value, implying the preferential exclusion of an anion around the RNA (Figure 2A,B). Ion-counting experiments[1,66] and previous molecular dynamics simulation studies[2] also support anion exclusion from the ionic atmosphere surrounding the nucleic acids. F^- binding to the AD is primarily governed by the correct orientation of LBD nucleotides where three Mg^{2+} bind to constitute the cationic LBD (LBD-3Mg) that captures the F^- from the solution (Figure 1A). At low $[Mg^{2+}]$ (= 1 mM), we found only 9 instances where F^- is transiently bound to the aptamer, and the LBD is partially folded in all of these conformations. Each F^- binding instance has a lifetime of 12.5 ps. In 8 out of 9 instances, F^- is bound to the LBD with the assistance of only one Mg^{2+} (Figure S8A). For $[Mg^{2+}]$ (= 4 mM), the number of instances where one F^- bound to the LBD increased to 21. Out of 21, in 7 instances, two Mg^{2+} capture an F^- to facilitate its binding to the partially folded LBD (Figure S8B). The number of F^- binding instances increased to 29 for $[Mg^{2+}]$ = 8 mM. We found 15 instances where two Mg^{2+} capture an F^-, and in 4 such instances, the aptamer is in native-like folded conformations with the fully folded LBD (Figure S8C). We found only one instance where three Mg^{2+} capture an F^- at LBD, and the aptamer is in the *holo* form (Figure S8D). However, the average lifetime of all of the F^--bound conformations observed in the simulations using the coarse-grained RNA model is ≤12.5 ps and it is not a stable state. The main conclusion from the simulations using the coarse-grained RNA model is that in high $[Mg^{2+}]$, AD can populate *holo*-like conformations even in the absence of F^- binding, which are competent to activate transcription (Figure S8D). The LBD has a high affinity for Mg^{2+}, and two Mg^{2+} are bound to the LBD nucleotides even in the AD unfolded state. F^- is transiently captured by the two Mg^{2+} bonds bound to the cationic LBD.

The transient nature of F^- binding to the cationic LBD, even at high $[Mg^{2+}]$ (≥5 mM), can be attributed to the lack of explicit solvent molecules in the coarse-grained simulations. We hypothesized that the lack of hydration shells around Mg^{2+} prevented the binding of more than two Mg^{2+} in the fully formed LBD due to the enhanced electrostatic repulsion among Mg^{2+} [7,31] (Figure 1A). Furthermore, due to the lack of a rigid hydration shell around Mg^{2+}, the cationic LBD is distorted and Mg^{2+} cannot effectively counterbalance the electrostatic repulsion among the negatively charged phosphate beads of LBD nucleotides. This resultant electrostatic

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

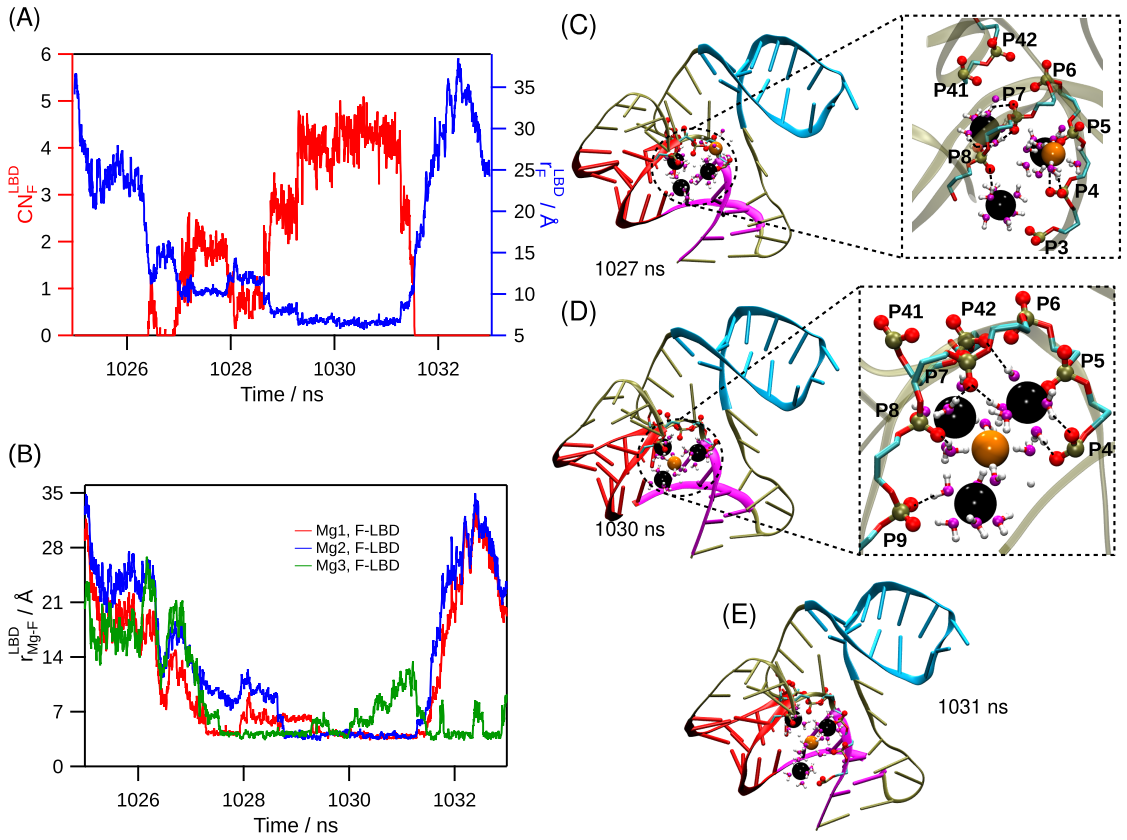


Figure 3. Kinetic pathway for the binding mode (I) obtained from the AA[1,0] simulations, where the three Mg^{2+} bind to the LBD nucleotides to form the cationic pocket. F^- later approaches the LBD and binds to the three Mg^{2+} ions. (A) Distance between the F^- ion, which binds to the LBD, and the center of mass of the LBD RNA atoms is plotted in blue (right axis). The coordination number of the same F^- with the LBD RNA atoms is plotted in red (left axis). (B) Distances between the bound F^- and the three Mg^{2+} are plotted. Distances for only those Mg^{2+} that are within the water-mediated binding distance, 5 Å from the bound F^-, are plotted. In panels C–E, different stages of F^- binding to the LBD are shown. The AD is shown in the cartoon representation. The pseudoknot PK and helices S_1 and S_2 are shown in magenta, red, and cyan, respectively. Mg^{2+}, F^-, phosphorus, and phosphate oxygen are shown as black, orange, tan, and red spheres, respectively. Water, oxygen, and hydrogen are shown as magenta- and white-colored balls, using a ball-and-stick representation, respectively. The backbone of LBD nucleotides and nucleotide G5 are shown using a stick representation. Cyan and red in the stick representation denote carbon and oxygen atoms, respectively. The nucleotide number (N_N) of the phosphates and nucleobase are provided in the annotation. The snapshot simulation time is in the annotation of each panel. (C) F^- approaches the LBD with three Mg^{2+} bound to the LBD. (D) F^- interacts with the three Mg^{2+} bound to the LBD through the water-mediated interaction to form the 3Mg–W-1F cluster. (E) F^- unbinding from the 3Mg–W-1F cluster.

repulsion among the negatively charged phosphate beads prevents the approach and binding of F^- to the LBD.

Role of Mg^{2+} and Water in F^- Binding to the LBD. The coarse-grained simulations showed that the number of events where F^- binds to the LBD increased with [Mg^{2+}]. At high [Mg^{2+}], the dominant F^- binding mode to the AD is an F^- captured by two Mg^{2+} bound to the LBD. We performed all-atom molecular dynamics simulations to probe the role of water in Mg^{2+} mediated F^- binding to the LBD. We performed six independent all-atom simulations starting from the AD crystal structure (PDB: 4ENC)[7] but without any Mg^{2+} and F^- bound to the LBD. This system is referred to as AA[0,0] (Table 1). The bulk concentration of both Mg^{2+} and F^- is 55 mM, and the concentration of K^+ is 150 mM in the simulation box.

In two out of six AA[0,0] simulations, we found that during the three-step equilibration (*see* Methods), an Mg^{2+} water-solvated ion that was closer to the phosphate oxygen atom of the G5 nucleotide in the LBD vicinity formed a direct interaction (inner-shell interaction) with the phosphate oxygen atom of the G5 nucleotide. We treated the two systems where we observed this event as one Mg^{2+} bound near the LBD and referred to it as AA[1,0] (Table 1). We have four independent simulations using the system configuration AA[0,0]. The cumulative simulation time for both AA[0,0] and AA[1,0] configurations is tabulated in Table 1. The AA[0,0] and AA[1,0] system configurations enabled us to answer the following questions: (a) whether the cationic LBD pocket forms first due to the binding of three Mg^{2+}, followed by F^- binding in the pocket, (b) how does F^- approach and bind to the LBD? (c) what is the configuration of the bound Mg^{2+} and F^- cluster in the LBD? and (d) what is the role of water in F^- binding to the LBD?

To track F^- binding to the LBD, we computed the coordination number for each of the F^- with the LBD RNA atoms (CN_F^{LBD}) and the distance of each F^- from the center of mass of the LBD RNA atoms (r_F^{LBD}) (*see* Methods) (Figure 3A and 4A). We consider F^- to be bound to the LBD if $CN_F^{LBD} \geq 3$ and $r_F^{LBD} \leq 7.3$ Å. We identified Mg^{2+} that facilitates F^- binding by computing the distance between all Mg^{2+} with the bound F^- (r_{Mg-F}^{LBD}) (Figures 3B and 4B). Mg^{2+} that facilitated F^- binding are identified using the criteria $r_{Mg-F}^{LBD} \leq 5$ Å.

In AA[0,0] simulations, we found only one instance where an F^- is captured by three Mg^{2+} to form a water-mediated

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267–9281

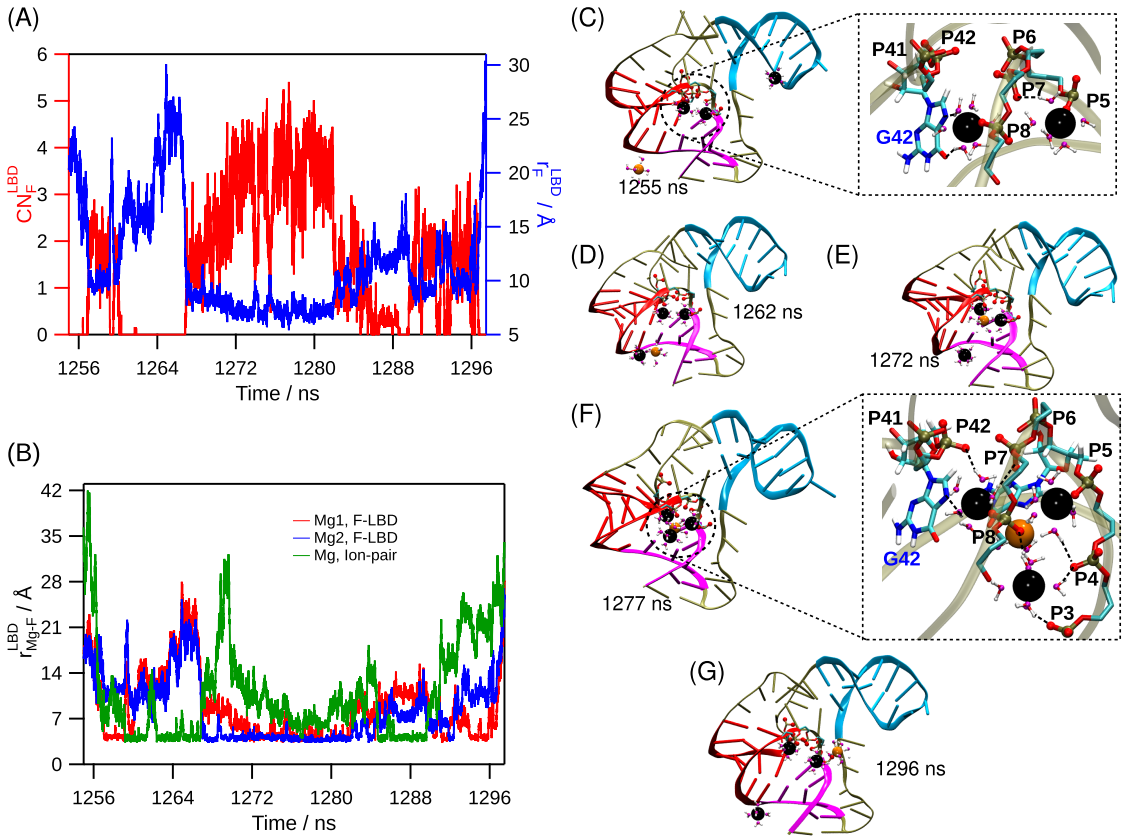


Figure 4. Kinetic pathway of F^- binding mode (II) obtained from the AA[1,0] simulations, where two Mg^{2+} bind to the LBD nucleotides and a third Mg^{2+} anchors the F^- to the LBD. (A) Distance between the F^- ion, which binds to the LBD, and the center of mass of the LBD RNA atoms is plotted in blue (right axis). The coordination number of the same F^- with the LBD RNA atoms is plotted in red (left axis). (B) Distances between the bound F^- and the three Mg^{2+} are plotted. Distances for only those Mg^{2+} that are within water-mediated binding distance, 5 Å from the bound F^-, are plotted. The distance between bound F^- and the Mg^{2+} of the ion pair is plotted with green color. In panels (C) to (G), the different stages of F^- binding to the LBD are shown. The AD is shown in a cartoon representation. The pseudoknot PK and helices S_1 and S_2 are shown in magenta, red, and cyan, respectively. Mg^{2+}, F^-, phosphorus, and phosphate oxygen are shown as black-, orange-, tan-, and red-colored spheres, respectively. Water, oxygen, and hydrogen are shown with magenta- and white-colored balls, using a ball-and-stick representation, respectively. The backbone of LBD nucleotides and the backbone of nucleotide G5 and the nucleobase of G42 are shown with a stick representation. Cyan, blue, red, and white colors in the stick representation are used to denote carbon, nitrogen, oxygen, and hydrogen atoms, respectively. The nucleotide number (N_N) of the phosphates and nucleobase are provided in the annotation. The snapshot simulation time is in the annotation of each panel. (C)−(E) Early stages of F^- binding to the cationic pocket and F^- approach to the LBD. (F) F^- bound conformation where 3 Mg^{2+} capture one F^- through a water-mediated interaction to form the 3Mg−W-1F cluster. (G) F^- unbinding from the 3Mg−W-1F cluster.

3Mg−W-1F cluster. In this F^- binding event, three Mg^{2+} bind to the LBD during the simulation, and F^- later approaches the LBD from bulk to form the 3Mg−W-1F cluster (Figure S10). Two out of three Mg^{2+} of the 3Mg−W-1F cluster are bound to the cavity of LBD through a water-mediated interaction with the phosphate O atoms of LBD nucleotides G8 and G42 (Figure S10C). The third Mg^{2+} in this cluster is bound at the periphery of the LBD through a water-mediated interaction with the phosphate O atoms of the G5 nucleotide and LBD nucleotides (U7 and G8), rendering itself exposed to the bulk (Figure S10C). All three Mg^{2+} interact with the bound F^- through a water-mediated interaction (Figure S10D). We discussed the implication of this solvent-exposed Mg^{2+} on the ionic assembly formation in the following section.

From the all-atom simulations with system configuration AA[1,0], we found four instances where F^- is bound to the cationic LBD composed of three Mg^{2+}. We observed the following two binding modes: (I) three Mg^{2+} bind to the LBD and form the cationic LBD pocket (Figure 3C) and later F^- from the bulk binds to this cationic LBD pocket (Figure 3D,E), and (II) two Mg^{2+} bound to the phosphate groups of the LBD nucleotides as observed in the coarse-grained simulations (Figure 4C). The third Mg^{2+} forms an ion pair with an F^- in bulk (Figure 4D), and the ion pair binds to the LBD (Figure 4E,F). In both the binding modes, one Mg^{2+} is directly bound (inner shell) to the phosphate group of G5 located in the LBD vicinity throughout the simulation. Water molecules are present in the rest of the five coordinating sites for this Mg^{2+}.

In the binding mode (I), the cationic pocket is preformed. Apart from the inner sphere bound Mg^{2+}, the other two Mg^{2+} form water-mediated interactions with the phosphate O atoms of nucleotides U7 and G8. The water-solvated F^- from the bulk approaches the cationic pocket and binds to it. There is only one hydration layer between Mg^{2+} and the bound F^- in the 3Mg−W-1F cluster. Computational studies have shown that the energy barrier for Mg^{2+} to lose a single water molecule from its first solvation layer is high (≈18.8 kcal mol^{-1}).[67] The incoming F^- loses its water solvation shell and binds to the Mg^{2+} of the cationic pocket, interacting with the water hydrogen atoms of the Mg^{2+} solvation shell to form a 3Mg−W-1F cluster (Figure 3D).

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

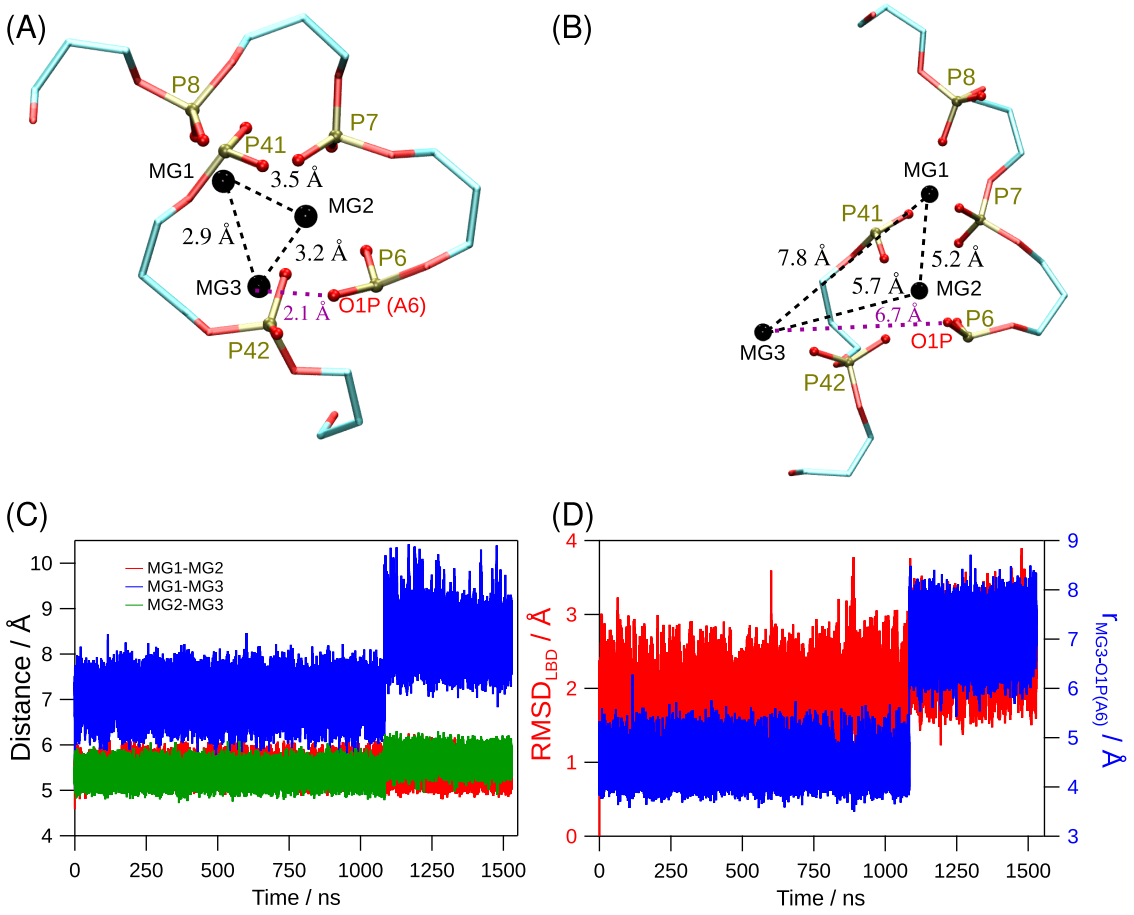


Figure 5. Data from AA[3,0] simulations. (A) and (B) Snapshot of the LBD in the crystal state and after ≈1139 μs of the simulation, respectively. Mg^{2+} ions are shown as black spheres. (C) Distance between a pair of LBD-bound Mg^{2+} ions is plotted. (D) RMSD of the LBD with respect to the crystal structure and distance between MG3 from phosphate of LBD nucleotide A6 are shown as red and blue lines, respectively.

In the binding mode (II), the second Mg^{2+} ion approaches the LBD and binds to the N7 atom of nucleobase G42 through a water-mediated interaction (Figure 4C). The third Mg^{2+} forms a water-mediated ion pair with hydrated F^- in the bulk, and this ion pair binds to the two prebound Mg^{2+} in the LBD, forming a water-mediated cluster of 3 Mg^{2+} and 1F^- (3Mg–W-1F) (Figure 4D,F). The cluster 3Mg–W-1F is held together by the water-mediated interactions between the Mg^{2+} and F^- where water atoms satisfy all six coordinating sites of Mg^{2+}, and the water H atoms interact with F^- (Figure 4F). The 3Mg–W-1F cluster interacts with the LBD through water-mediated interactions between Mg^{2+} and backbone O atoms of the LBD nucleotides (phosphate groups of U7, G8, and G42 and the nucleobase of G42) and the phosphate group of G5. The 3Mg–W-1F cluster has an average lifetime of ≈0.5 ns.

We observed that the third Mg^{2+}, which forms a water-mediated ion pair with F^- and forms part of the 3Mg–W-1F cluster, is not stable in the cluster (Figure 4E). This hydrated Mg^{2+} is weakly bound to the LBD compared to the other two Mg^{2+} because it cannot access the LBD cavity and is unable to form interactions with the LBD nucleotides (Figure 4F).

The incoming hydrated Mg^{2+} forms two water-mediated interactions with the phosphate O atoms of the 5′ terminal nucleotides (G3 and C4) and the LBD nucleotide (G8) (Figure 4F). These interactions are disrupted by thermal fluctuations, leading to the breaking away of Mg^{2+} from the 3Mg–W-1F cluster (Figure 4G). The inability of the third Mg^{2+} to form a strong interaction with the phosphate backbone contributes to the two distinctive binding F^- modes to the LBD, as discussed earlier.

In all of the F^- binding events obtained from both AA[0,0] and AA[1,0] simulations, the 3Mg–W-1F cluster is bound at the LBD periphery (Figure 4C) and not at the LBD core as observed in the crystal structure[7] (Figure 1A). In the all-atom simulations, the water-mediated interactions between Mg^{2+} and the phosphate O atoms of LBD nucleotides result in weak binding of the cationic pocket to the RNA. Even the interaction between Mg^{2+} and F^- is also water-mediated. In the crystal structure, Mg^{2+} and F^- interact directly and not through water-mediated interactions. In addition, all three Mg^{2+} also have a direct interaction with the phosphate O atoms (Figure 1A). To attain binding like the crystal-bound state, each Mg^{2+} in the 3Mg–W-1F cluster should shed at least three water molecules from its first solvation shell. Due to the high energy barrier for Mg^{2+} losing the first water molecule from its first solvation shell (≈18.8 kcal mol^{-1}),[67] we did not observe this event in the all-atom unbiased simulations even on a time scale of ≥15 μs. However, the magnitude of the energy barrier strongly depends on the water model and ion force field parameters.[68−70]

Three Mg^{2+} Ions Bound to the LBD in an Unstable Configuration in the Absence of a Bound F^- Ion. In the coarse-grained simulations, we observed the binding of at most two Mg^{2+} ions to the LBD nucleotides (Figure 1D). To check

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

if this result is an artifact due to the absence of explicit water molecules in the coarse-grained model and whether the configuration of three Mg^{2+} ions in the crystal-bound pose in the absence of F^- binding is stable or not, we performed all-atom simulations starting with the AA[3,0] configuration (*see* Table 1). In the AA[3,0] configuration, we removed the bound F^- ion from the crystal structure to obtain the AD with three inner-shell-bound Mg^{2+} to the LBD (Figure 5A).

The all-atom simulations of the AA[3,0] configuration show that in the absence of F^- binding, one of the Mg^{2+} ions is ejected out from the LBD due to the enhanced electrostatic repulsion between the three Mg^{2+} ions bound to the LBD (Figure 5B,C,D). The distance between the ejected Mg^{2+} and the other two Mg^{2+} bound to the LBD increased (Figure 5B,C). The orientation of the phosphate groups in the LBD is also distorted (Figure 5B,D). The MG3 ion moves away even from the O1P atom of LBD nucleotide A6 (right axis of Figure 5D), to which it is bound using the inner-shell interaction in the crystal structure (Figure 5A). This result from the all-atom simulations supports the result obtained from the coarse-grained simulations that the LBD can accommodate, at most, two Mg^{2+} ions in the absence of F^- binding. When two Mg^{2+} ions are bound to the LBD, the approach of the third Mg^{2+} to the LBD is hindered due to enhanced electrostatic repulsion from the two prebound Mg^{2+} ions. This result further supports the hypothesis that probably the third Mg^{2+} ion and the F^- ion simultaneously approach the LBD with the two prebound Mg^{2+} ions.

Plausible Mechanism for Mg^{2+} and F^- Assembly in the LBD. All-atom simulations have shown that the water-mediated cluster 3Mg−W-1F is transiently bound to the RNA due to a weak outer-shell interaction between the Mg^{2+} and phosphate groups of the LBD. For stable binding of the 3Mg−1F cluster at the LBD, Mg^{2+} has to directly interact with the O atoms of the phosphate groups of the LBD through inner-shell binding by shedding water molecules from its first solvation shell. A water molecule exchange from the Mg^{2+} first solvation shell has a high energy barrier,[67,70−72] which makes it highly unlikely that all three Mg^{2+} in the 3Mg−W-1F cluster will simultaneously shed water molecules and assemble concertedly to form the 3Mg−1F cluster at LBD. The system can avoid this improbable event if the Mg^{2+} binds hierarchically and transitions from the outer to the inner-shell interaction in the LBD to construct the cationic pocket.

The longer Mg^{2+} is bound to the LBD with its water solvation shell, the higher its probability of transitioning from an outer-shell to an inner-shell interaction. To show the plausibility of Mg^{2+} following the sequential assembly mechanism to form the 3Mg−1F cluster, we computed the average conditional lifetime and occupancy time of the three Mg^{2+} ions in the LBD. $\tau(i, j)$ is the average conditional lifetime that ith Mg^{2+} ion with its water-mediated shell is bound to the LBD given that j number of Mg^{2+} is already prebound to the LBD. f_τ^i is the fraction of simulation time at which at least i number of Mg^{2+} is bound to the LBD.

In the AA[0,0] system, the average lifetime of the first Mg^{2+} bound to the LBD is 27.31 ns, as it is a water-mediated interaction. For it to bind stably to the LBD, it should undergo a transition from an outer to an inner-shell interaction, which we did not observe on our simulation time scale due to the high energy barrier (Table 2a). The lifetime of the second and third Mg^{2+} bound to the LBD with their water solvation shells are even smaller, ≈1 and 0.04 ns, respectively. Even f_τ^i decreased with the increase in i, the number of Mg^{2+} bound to the LBD (Table 2b). Therefore, it is unlikely that the three bound Mg^{2+} will transition from an outer to an inner-shell interaction, with all of them bound to the LBD. A plausible mechanism is that the three Mg^{2+} transition from an outer to an inner-shell interaction occurs in a hierarchical manner, and the following analysis supports it.

Table 2. Mg^{2+} Binding Time with the LBD in All-Atom Simulations

(a) Average Conditional Binding Lifetime: $\tau(1,0)$ is the average binding lifetime of the first Mg^{2+} to the LBD. $\tau(2,1)$ represents the average binding lifetime of the second Mg^{2+} to the LBD, given one Mg^{2+} is already bound to the LBD. Similarly, $\tau(3,2)$ represents the average binding lifetime of the third Mg^{2+}, given two Mg^{2+} are already bound to the LBD. The large standard deviation values of $\tau(i, j)$ are due to the presence of both short-lived and long-lived binding events of Mg^{2+} to the LBD.

system	$\tau(1,0)$ (ns)	$\tau(2,1)$ (ns)	$\tau(3,2)$ (ns)
AA[0,0]	27.31 (±59.37)	1.04 (±1.11)	0.04 (±0.01)
AA[1,0]		4.16 (±15.53)	0.06 (±0.08)
AA[2,0]			0.05 (±0.04)

(b) Occupancy Time: f_τ^i represents the fraction of time at least i number of Mg^{2+} bound to LBD.

system	f_τ^1	f_τ^2	f_τ^3
AA[0,0]	0.95	0.085	0.0006
AA[1,0]	1	0.68	0.045
AA[2,0]	1	1	0.0283

The values of $\tau(2, 1)$ and f_τ^2 are larger for the AA[1,0] system compared to those of the AA[0,0] system (Tables 2a and 2b). This indicates that when the first Mg^{2+} ion is bound to the LBD using the inner-shell interaction, and then the frequency of the second Mg^{2+} ion binding to the LBD increases as well as its bound lifetime, enhancing its probability to transition from an outer to an inner-shell interaction. This supports the mechanism that initially, the first bound Mg^{2+} transitions from an outer to an inner-shell interaction in the LBD and facilitates the binding and transition of the second bound Mg^{2+} from an outer to an inner-shell interaction later. However, the $\tau(3, 2)$ and f_τ^3 values are similar for all of the systems, indicating that even with the two Mg^{2+} already bound to the LBD with an inner-shell interaction (AA[2,0] system), binding of the third Mg^{2+} with its water solvation shell is probably hindered due to the electrostatic repulsion from the other two bound Mg^{2+} (Figure 5) (Tables 2a and 2b). As a result, we hypothesize that the efficient way for the third Mg^{2+} to approach the LBD with two prebound Mg^{2+} is, as an ion pair with F^- to complete the 3Mg−W-1F cluster, as observed in binding mode II. The third Mg^{2+} will later shed water from its solvation shell to transition from the outer to the inner-shell interaction and complete the assembly of the 3Mg−1F cluster in the LBD.

Fluoride Binding to the LBD Stabilizes Linchpin Hydrogen Bonds Critical for Transcription Initiation. We performed simulations of the AD in AA[ϕ](*apo*) and AA[3,1](*holo*) configurations (Table 1) to identify the changes in the nucleotide fluctuations upon the 3Mg−1F cluster binding to the LBD. The root-mean-square fluctuation (RMSF) computed at the nucleobase resolution shows that in the *apo* state, the nucleobases of LBD nucleotides (A6, U7, G8, U41, and G42) are more flexible compared to the *holo* state (Figure S11B). The root-mean-square deviation (RMSD) of LBD nucleotides shows that in the *holo* state, the three bound Mg^{2+} to the phosphate backbone of LBD nucleotides

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

stabilize their corresponding nucleobases in native-like orientation (Figure S11C).

We computed the dynamic cross-correlation, C_{ij} of the AD, for AA[ϕ](*apo*) and AA[3,1](*holo*) states to quantify the impact of 3Mg−1F cluster binding on the relative motion between the nucleobases that are critical for transmitting the ligand-binding signal to the expression platform. We plotted the difference in the dynamical cross-correlation (ΔC_{ij}) (see eq 9 in Methods) between the nucleobases *i* and *j* by subtracting C_{ij} of the *holo* state (AA[3,1]) from the *apo* state (AA[ϕ]) (Figure 6). ΔC_{ij} shows that the following nucleotides become

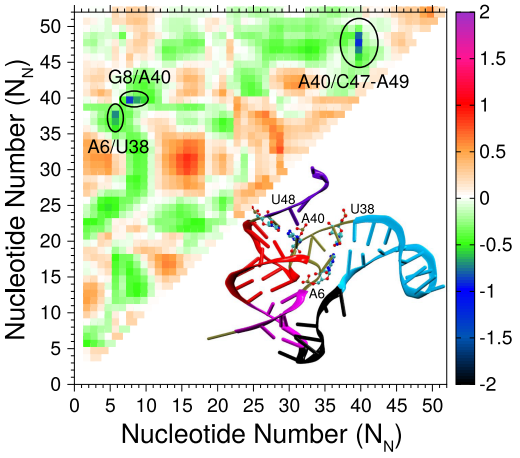


Figure 6. Heatmap of the difference in cross-correlations for nucleobases (ΔC_{ij}) obtained by subtracting the C_{ij} of the *holo* form (AA[3,1]) from the *apo* form (AA[ϕ]). ΔC_{ij} depicts the change in correlation between a pair of nucleobase beads upon binding of the 3Mg−1F cluster to the AD. The AD structure is shown in the inset by using the cartoon representation. The linchpin hydrogen bond pairs are shown using the ball−stick model. Magenta, red, cyan, black, and violet regions represent nucleotides belonging to pseudoknot, helix S_1, helix S_2, PK-S_2, and 3-TER, respectively.

anticorrelated due to the absence of 3Mg−1F cluster binding to the LBD in the *apo* form: (i) A40 and C47-A49, (ii) G8 and A40, (iii) and A6 and U38 (Figure 6).

We probed the structural changes in the RNA regions, which showed variations in the correlated motion (N_N = A6, G36-U38, A40, and U46−C50) upon 3Mg−1F cluster binding. The linchpin hydrogen bonds in the regions mentioned above break and only form in response to the 3Mg−1F cluster binding to the AD (Figure S12B). The linchpin hydrogen bonds with reverse Hoogsteen alignment (A40/U48) and reverse Watson−Crick alignment (A6/U38) are broken in the *apo* state (Figure S12B). This is in agreement with the experiments, which suggested that in the *apo* form, the linchpin nucleotide pair A40/U48 are not hydrogen-bonded to each other;[32] instead, U48 forms hydrogen bonds with A64 located in the expression platform at the terminator helix.[7,33] The terminator helix masks the polyU sequence from the RNA polymerase through strand invasion, and transcription is stopped.[7,33] We find that the binding of the 3Mg−1F cluster to the AD is imperative for forming both of the linchpin hydrogen bonds. Experiments also suggest that especially the formation of the A40/U48 bond facilitates the antiterminator helix formation, which exposes the polyU sequence to the RNA polymerase to activate the transcription (Figure S12B).[7,33]

CONCLUSIONS

The fluoride riboswitch performs the critical function of maintaining fluoride ion homeostasis in bacteria. The riboswitch folding and the coupled ion binding process is a multitime scale problem, and it is extremely challenging to model this process using a single RNA model or a computational technique. We performed simulations using multiple resolutions of RNA models to understand the assembly mechanism of the 3 Mg^{2+} ions in the LBD that encapsulate the F^- ion.

The crystal structure[7] shows that the five nucleotides that comprise the LBD orient pointing their phosphate groups toward the LBD center, leading to a concentration of negative charge, where the 3 Mg^{2+} ions are anchored to encapsulate the F^- ion. We show that the tertiary stacking interactions between the LBD nucleobases alone cannot stabilize the orientation of the phosphate groups and require condensation of the Mg^{2+} ions to stabilize them in that orientation. We show that in the absence of F^- ion binding, the LBD has a high affinity for Mg^{2+} binding and can accommodate the binding of two Mg^{2+} ions. Without F^- ion binding, the binding of three Mg^{2+} ions will lead to a concentration of positive charge in a small LBD region.

The computations support the sequential binding of the two Mg^{2+} ions to the LBD phosphate groups and the transition from an outer to an inner-shell interaction through dehydration. The third Mg^{2+} and F^- ion bind to the LBD in two ways to form the 3Mg−1F cluster. (I) The third Mg^{2+} binds to the LBD to form the trigonal cationic pocket and then captures the F^- and (II) the third Mg^{2+} and F^- form a water-mediated ion pair (Mg-water-F) in solution and simultaneously bind to the cationic pocket. The F^- ion most likely approaches the LBD by forming a water-mediated ion pair with the third Mg^{2+} ion, as this minimizes the electrostatic repulsion between the 3 Mg^{2+} ions in the cluster. The entire assembly comprising the 5 LBD phosphate groups, 3 Mg^{2+}, and the F^- ion points to an intricate interaction between the three charged components and the tertiary stacking interactions of the LBD nucleobases. The interactions between these multiple components can be tuned to devise synthetic RNA aptamers that can selectively sense other halides or anions.

The simulations further support the previous experimental findings, such as the fluoride riboswitch attains a *holo*-like structure irrespective of F^- binding, suggesting that it follows a conformational selection mechanism.[34] The *holo*-like structure is not globally stable in the absence of F^- binding and becomes the most stable structure only upon F^- binding. The simulations also show that the linchpin nucleotide pair A40/U48 is not hydrogen-bonded to each other in the *apo* form in agreement with the experiments.[7,32] Instead, the nucleotide U48 forms hydrogen bonds with A64 located in the terminator helix of the expression platform.[7,33]

NMR experiments can potentially be designed to verify the predictions from this work. The following predictions from the computations can be tested using experiments. (1) In the *holo*-like state of the riboswitch, which is populated in the absence of F^- binding, the simulations show that the LBD in this structure is unstable due to the ruptured tertiary stacking interactions (A6/G24, A6/G39, U7/G39, and G8/A40). (2) In low Mg^{2+} concentrations and the absence of F^- ions in the solution, the LBD can accommodate the binding of two Mg^{2+} ions. (3) The sequential binding of the initial two Mg^{2+} to the

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281

LBD.[73,74] (4) Test the prediction that the optimal way for the third Mg^{2+} and F^- to bind is to approach simultaneously as a water-mediated ion pair or arrive independently almost at the same time to the LBD.

The drawbacks of the current calculations are that the predictions involving the events related to the Mg^{2+} outer-to-inner-shell transitions are based on binding time calculations (Table 2) and are not observed in the all-atom simulations due to the high energy barrier for water exchange. Simulations using enhanced sampling techniques can be a viable option to probe these events. However, these simulations can be computationally intensive and not trivial as the number of possibilities for the three Mg^{2+} to interact with the five different phosphate groups and transition from an outer to an inner-shell interaction is large, and each transition has a high energy barrier, which makes it computationally challenging to sample all of the possibilities effectively. The coarse-grained and all-atom simulations show that the three Mg^{2+} ions, even with the inner-shell interaction with the LBD phosphate groups, are not stable without the bound F^- ion. The collective variables for enhanced sampling simulations to capture the simultaneous binding of F^- and the third Mg^{2+} to the LBD and transition from outer-to-inner-shell interactions may not be trivial. The high energy barrier for the outer-to-inner-shell transition could also be due to the imbalance between the ion−water and ion−RNA interactions in the all-atom force field, which uses 12−6 type of LJ parameters. It has been shown that the 12−6−4 type of LJ parameters can balance the ion−water and ion−RNA interactions for certain specific RNA sites.[59,60] However, accurate experiments are required to measure the transition rates of the Mg^{2+} ion from an outer to an inner-shell interaction with the phosphate groups in the nucleic acids.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.3c03941.

Additional methods details for the TIS-RNA model (eqs S1−S8) and coarse-grained parameters (Tables S1−S3); additional analyses details for the fraction of native contacts among LBD nucleotides (eq S8), the fraction of tertiary contacts (eq S9), and the fraction of tertiary stack formation (eq S10); additional results for local folding dynamics of the unstructured nucleotides, AD stabilization by tertiary stacks; comparison of ⟨RMSD⟩ between *holo*-like coarse-grained conformations and AA[3,1] all-atom conformations (Table S4); six native tertiary stacks in the AD (Figure S1); native contact map of the AD (Figure S2); average fraction of the tertiary structure elements and LBD (Figure S3); average fraction of native contacts in unstructured nucleotides and their FES (Figure S4); time profile for the fraction of native contacts and fraction of native contacts in unstructured regions at $[Mg^{2+}]$ = 1, 4, and 8 mM (Figures S5−S7); representative snapshots from coarse-grained simulations (Figure S8); schematic illustration of the criteria to decide whether Mg^{2+} is bound to the LBD in all-atom simulations (Figure S9); kinetic pathway of F^- binding in AA[0,0] simulations (Figure S10); RMSD and RMSF of various regions of the AD obtained from AA[ϕ] and AA[3,1] simulations (Figure S11); and average fraction of tertiary stack formation from coarse-grained and all-atom simulations and the fraction of linchpin hydrogen bond formation from all-atom simulations (Figure S12) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Govardhan Reddy − *Solid State and Structural Chemistry Unit, Indian Institute of Science, Bengaluru, Karnataka 560012, India;* orcid.org/0000-0002-9013-8040; Phone: +91-80-22933533; Email: greddy@iisc.ac.in; Fax: +91-80-23601310

Author

Sunil Kumar − *Solid State and Structural Chemistry Unit, Indian Institute of Science, Bengaluru, Karnataka 560012, India;* orcid.org/0000-0001-6830-5589

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.3c03941

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

G.R. acknowledges funding from the National Supercomputing Mission (MeitY/R&D/HPC/2(1)/2014). S.K. acknowledges a research fellowship from the Indian Institute of Science, Bangalore. We acknowledge the National Supercomputing Mission (NSM) for providing computing resources of "PARAM Brahma" at IISER Pune, which is implemented by C-DAC and supported by the Ministry of Electronics and Information Technology (MeitY) and Department of Science and Technology (DST), Government of India.

■ REFERENCES

(1) Bai, Y.; Greenfeld, M.; Travers, K. J.; Chu, V. B.; Lipfert, J.; Doniach, S.; Herschlag, D. Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.* **2007**, *129*, 14981−14988.

(2) Giambaşu, G.; Luchko, T.; Herschlag, D.; York, D. M.; Case, D. A. Ion counting from explicit-solvent simulations and 3D-RISM. *Biophys. J.* **2014**, *106*, 883−894.

(3) 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*, 2435−2445.

(4) Winkler, W.; Nahvi, A.; Breaker, R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **2002**, *419*, 952−956.

(5) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J.; Breaker, R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **2004**, *428*, 281−286.

(6) Winkler, W. C.; Cohen-Chalamish, S.; Breaker, R. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15908−15913.

(7) Ren, A.; Rajashankar, K. R.; Patel, D. J. Fluoride ion encapsulation by Mg^{2+} ions and phosphates in a fluoride riboswitch. *Nature* **2012**, *486*, 85−89.

(8) Win, M. N.; Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science* **2008**, *322*, 456−460.

(9) Miyamoto, T.; Razavi, S.; DeRose, R.; Inoue, T. Synthesizing Biomolecule-Based Boolean Logic Gates. *ACS Synth. Biol.* **2013**, *2*, 72−82.

(10) Lemay, J.-F.; Penedo, J. C.; Tremblay, R.; Lilley, D. M. J.; Lafontaine, D. A. Folding of the adenine riboswitch. *Chem. Biol.* **2006**, *13*, 857−868.

(11) Sung, H.-L.; Nesbitt, D. J. Sequential folding of the nickel/cobalt riboswitch is facilitated by a conformational intermediate: Insights from single-molecule kinetics and thermodynamics. *J. Phys. Chem. B* **2020**, *124*, 7348−7360.
(12) Ma, J.; Saikia, N.; Godar, S.; Hamilton, G. L.; Ding, F.; Alper, J.; Sanabria, H. Ensemble Switching Unveils a Kinetic Rheostat Mechanism of the Eukaryotic Thiamine Pyrophosphate Riboswitch. *RNA* **2021**, *27*, 771−790.
(13) Hua, B.; Jones, C. P.; Mitra, J.; Murray, P. J.; Rosenthal, R.; Ferré-D'Amaré, A. R.; Ha, T. Real-time monitoring of single ZTP riboswitches reveals a complex and kinetically controlled decision landscape. *Nat. Commun.* **2020**, *11*, No. 4531.
(14) St-Pierre, P.; Shaw, E.; Jacques, S.; Dalgarno, P. A.; Perez-Gonzalez, C.; Picard-Jean, F.; Penedo, J. C.; Lafontaine, D. A. A structural intermediate pre-organizes the *add* adenine riboswitch for ligand recognition. *Nucleic Acids Res.* **2021**, *49*, 5891−5904.
(15) Greenleaf, W. J.; Frieda, K. L.; Foster, D. A. N.; Woodside, M. T.; Block, S. M. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **2008**, *319*, 630−633.
(16) Neupane, K.; Yu, H.; Foster, D. A. N.; Wang, F.; Woodside, M. T. Single-molecule force spectroscopy of the *add* adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res.* **2011**, *39*, 7677−7687.
(17) Brenner, M. D.; Scanlan, M. S.; Nahas, M. K.; Ha, T.; Silverman, S. K. Multivector fluorescence analysis of the *xpt* guanine riboswitch aptamer domain and the conformational role of guanine. *Biochemistry* **2010**, *49*, 1596−1605.
(18) Haller, A.; Altman, R. B.; Soulière, M. F.; Blanchard, S. C.; Micura, R. Folding and ligand recognition of the TPP riboswitch aptamer at single-molecule resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 4188−4193.
(19) Heppell, B.; Blouin, S.; Dussault, A.-M.; Mulhbacher, J.; Ennifar, E.; Penedo, J. C.; Lafontaine, D. A. Molecular insights into the ligand-controlled organization of the SAM-I riboswitch. *Nat. Chem. Biol.* **2011**, *7*, 384−392.
(20) Haller, A.; Rieder, U.; Aigner, M.; Blanchard, S. C.; Micura, R. Conformational capture of the SAM-II riboswitch. *Nat. Chem. Biol.* **2011**, *7*, 393−400.
(21) Soulière, M. F.; Altman, R. B.; Schwarz, V.; Haller, A.; Blanchard, S. C.; Micura, R. Tuning a riboswitch response through structural extension of a pseudoknot. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E3256−E3264.
(22) Wood, S.; Ferré-D'Amaré, A. R.; Rueda, D. Allosteric tertiary interactions preorganize the c-di-GMP riboswitch and accelerate ligand binding. *ACS Chem. Biol.* **2012**, *7*, 920−927.
(23) Suddala, K. C.; Price, I. R.; Dandpat, S. S.; Janecek, M.; Kuhrova, P.; Sponer, J.; Banas, P.; Ke, A.; Walter, N. G. Local-to-global signal transduction at the core of a Mn^{2+} sensing riboswitch. *Nat. Commun.* **2019**, *10*, 4304 DOI: 10.1038/s41467-019-12230-5.
(24) Lin, J.-C.; Thirumalai, D. Relative Stability of Helices Determines the Folding Landscape of Adenine Riboswitch Aptamers. *J. Am. Chem. Soc.* **2008**, *130*, 14080.
(25) Gong, Z.; Zhao, Y.; Chen, C.; Xiao, Y. Computational study of unfolding and regulation mechanism of preQ$_1$ riboswitches. *PLoS One* **2012**, *7*, e45239.
(26) Roy, S.; Hennelly, S. P.; Lammert, H.; Onuchic, J. N.; Sanbonmatsu, K. Y. Magnesium controls aptamer-expression platform switching in the SAM-I riboswitch. *Nucleic Acids Res.* **2019**, *47*, 3158−3170.
(27) Hayes, R. L.; Noel, J. K.; Mohanty, U.; Whitford, P. C.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y. Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. *J. Am. Chem. Soc.* **2012**, *134*, 12043−12053.
(28) Di Palma, F.; Colizzi, F.; Bussi, G. Ligand-induced stabilization of the aptamer terminal helix in the add adenine riboswitch. *RNA* **2013**, *19*, 1517−1524.
(29) Halder, A.; Kumar, S.; Valsson, O.; Reddy, G. Mg^{2+} sensing by an RNA fragment: Role of Mg^{2+}-coordinated water molecules. *J. Chem. Theory Comput.* **2020**, *16*, 6702−6715.
(30) Kumar, S.; Reddy, G. TPP riboswitch populates holo-form-like structure even in the absence of cognate ligand at high Mg^{2+} concentration. *J. Phys. Chem. B* **2022**, *126*, 2369−2381.
(31) Chawla, M.; Credendino, R.; Poater, A.; Oliva, R.; Cavallo, L. Structural stability, acidity, and halide selectivity of the fluoride riboswitch recognition site. *J. Am. Chem. Soc.* **2015**, *137*, 299−306.
(32) Watters, K. E.; Strobel, E. J.; Yu, A. M.; Lis, J. T.; Lucks, J. B. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat. Struct. Mol. Biol.* **2016**, *23*, 1124−1131.
(33) Zhao, B.; Guffy, S. L.; Williams, B.; Zhang, Q. An excited state underlies gene regulation of a transcriptional riboswitch. *Nat. Chem. Biol.* **2017**, *13*, 968−974.
(34) Yadav, R.; Widom, J. R.; Chauvier, A.; Walter, N. G. An anionic ligand snap-locks a long-range interaction in a magnesium-folded riboswitch. *Nat. Commun.* **2022**, *13*, 207 DOI: 10.1038/s41467-021-27827-y.
(35) Denesyuk, N. A.; Thirumalai, D. Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B* **2013**, *117*, 4901−4911.
(36) Denesyuk, N. A.; Thirumalai, D. How do metal ions direct ribozyme folding? *Nat. Chem.* **2015**, *7*, 793−801.
(37) Vriend, G. WHAT IF- a molecular modeling and drug design program. *J. Mol. Graph.* **1990**, *8*, 52−56.
(38) Hori, N.; Denesyuk, N. A.; Thirumalai, D. Shape changes and cooperativity in the folding of the central domain of the 16S ribosomal RNA. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, e2020837118.
(39) Roca, J.; Hori, N.; Baral, S.; Velmurugu, Y.; Narayanan, R.; Narayanan, P.; Thirumalai, D.; Ansari, A. Monovalent ions modulate the flux through multiple folding pathways of an RNA pseudoknot. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E7313−E7322.
(40) Tan, D.; Piana, S.; Dirks, R. M.; Shaw, D. E. RNA force field with accuracy comparable to state-of-the-art protein force fields. *Proc. Nat. Acad. Sci. U.S.A.* **2018**, *115*, E1346−E1355.
(41) Veitshans, T.; Klimov, D.; Thirumalai, D. Protein folding kinetics: Timescales, pathways and energy landscapes in terms of sequence-dependent properties. *Fold. Des.* **1997**, *2*, 1−22.
(42) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.
(43) Hori, N. TIS2AA; Zenodo, 2017, https://doi.org/10.5281/zenodo.581485 (accessed 25 March, 2021).
(44) Humphris-Narayanan, E.; Pyle, A. M. Discrete RNA libraries from pseudo-torsional space. *J. Mol. Biol.* **2012**, *421*, 6−26.
(45) Case, D.; Aktulga, H.; Belfon, K.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham, T., III; Cisneros, G.; Cruzeiro, V. et al. *AMBER 2018*; University of California: San Francisco, CA, 2018.
(46) Berendsen, H.; van der Spoel, D.; van Drunen, R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(47) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.
(48) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113−5123.
(49) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101 DOI: 10.1063/1.2408420.
(50) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(51) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: an N·log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(52) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(53) Kührová, P.; Mlynsky, V.; Zgarbova, M.; Krepl, M.; Bussi, G.; Best, R. B.; Otyepka, M.; Sponer, J.; Banas, P. Improving the Performance of the Amber RNA Force Field by Tuning the Hydrogen-Bonding Interactions. *J. Chem. Theory Comput.* **2019**, *15*, 3288−3305.
(54) Kuhrova, P.; Mlynsky, V.; Zgarbova, M.; Krepl, M.; Bussi, G.; Best, R. B.; Otyepka, M.; Sponer, J.; Banas, P. Improving the Performance of the Amber RNA Force Field by Tuning the Hydrogen-Bonding Interactions (vol 15, pg 3288, 2019). *J. Chem. Theory Comput.* **2020**, *16*, 818−819.
(55) Stokes, R. H. The van der Waals radii of gaseous ions of the noble gas structure in relation to hydration energies. *J. Am. Chem. Soc.* **1964**, *86*, 979−982.
(56) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M., Jr. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for + 2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9*, 2733−2748.
(57) Li, P.; Merz, K. M., Jr. Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J. Chem. Theory Comput.* **2014**, *10*, 289−297.
(58) Leonarski, F.; D'Ascenzo, L.; Auffinger, P. Mg2+ ions: do they bind to nucleobase nitrogens? *Nucleic Acids Res.* **2017**, *45*, 987−1004.
(59) Panteva, M. T.; Giambasu, G. M.; York, D. M. Force field for Mg^{2+}, Mn^{2+}, Zn^{2+}, and Cd^{2+} ions that have balanced interactions with nucleic acids. *J. Phys. Chem. B* **2015**, *119*, 15460−15470.
(60) Panteva, M. T.; Giambasu, G. M.; York, D. M. Comparison of Structural, Thermodynamic, Kinetic and Mass Transport Properties of Mg^{2+} Ion Models Commonly used in Biomolecular Simulations. *J. Comput. Chem.* **2015**, *36*, 970−982.
(61) Hori, N.; Denesyuk, N. A.; Thirumalai, D. Ion condensation onto ribozyme is site specific and fold dependent. *Biophys. J.* **2019**, *116*, 2400−2410.
(62) 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.
(63) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banas, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.
(64) McCammon, J. A. Protein Dynamics. *Rep. Prog. Phys.* **1984**, *47*, 1−46.
(65) Brown, D. K.; Penkler, D. L.; Amamuddy, O. S.; Ross, C.; Atilgan, A. R.; Atilgan, C.; Bishop, O. T. MD-TASK: a software suite for analyzing molecular dynamics trajectories. *Bioinformatics* **2017**, *33*, 2768−2771.
(66) Gebala, M.; Johnson, S. L.; Narlikar, G. J.; Herschlag, D. Ion counting demonstrates a high electrostatic field generated by the nucleosome. *eLife* **2019**, *8*, e44993 DOI: 10.7554/eLife.44993.
(67) Schwierz, N. Kinetic pathways of water exchange in the first hydration shell of magnesium. *J. Chem. Phys.* **2020**, *152*, 224106.
(68) Kurnikov, I. V.; Kurnikova, M. Modeling electronic polarizability changes in the course of a magnesium ion water ligand exchange process. *J. Phys. Chem. B* **2015**, *119*, 10275−10286.
(69) Falkner, S.; Schwierz, N. Kinetic pathways of water exchange in the first hydration shell of magnesium: Influence of water model and ionic force field. *J. Chem. Phys.* **2021**, *155*, 084503 DOI: 10.1063/5.0060896.
(70) Grotz, K. K.; Schwierz, N. Optimized magnesium force field parameters for biomolecular simulations with accurate solvation, ion-binding, and water-exchange properties in SPC/E, TIP3P-fb, TIP4P/2005, TIP4P-Ew, and TIP4P-D. *J. Chem. Theory Comput.* **2022**, *18*, 526−537, DOI: 10.1021/acs.jctc.1c00791.
(71) Bock, C. W.; Kaufman, A.; Glusker, J. Coordination of water to magnesium cations. *Inorg. Chem.* **1994**, *33*, 419−427.
(72) Grotz, K. K.; Schwierz, N. Magnesium force fields for OPC water with accurate solvation, ion-binding, and water-exchange properties: Successful transfer from SPC/E. *J. Chem. Phys.* **2022**, *156*, 114501.
(73) Bleuzen, A.; Pittet, P.-A.; Helm, L.; Merbach, A. E. Water exchange on magnesium (II) in aqueous solution: a variable temperature and pressure 17O NMR study. *Magn. Reson. Chem.* **1997**, *35*, 765−773.
(74) Neely, J.; Connick, R. Rate of water exchange from hydrated magnesium ion. *J. Am. Chem. Soc.* **1970**, *92*, 3476−3478.

https://doi.org/10.1021/acs.jpcb.3c03941
J. Phys. Chem. B 2023, 127, 9267−9281