Riboswitch Folds to *Holo*-Form Like Structure Even in the Absence of Cognate Ligand at High Mg^{2+} Concentration

Sunil Kumar and Govardhan Reddy*

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bengaluru, Karnataka, India 560012

E-mail: greddy@iisc.ac.in

Phone: +91-80-22933533. Fax: +91-80-23601310

Abstract

Riboswitches are non-coding RNA that regulate gene expression by folding into specific three-dimensional structures (*holo*-form) upon binding by their cognate ligand in the presence of Mg^{2+}. Riboswitch functioning is also hypothesized to be under kinetic control requiring large cognate ligand concentrations. We ask the question under thermodynamic conditions, can the riboswitches populate *holo*-form like structures in the absence of their cognate ligands only in the presence of Mg^{2+}. We addressed this question using thiamine pyrophosphate (TPP) riboswitch as a model system and computer simulations using a coarse-grained model for RNA. The folding free energy surface (FES) shows that with the initial increase in Mg^{2+} concentration ($[Mg^{2+}]$), TPP AD undergoes a barrierless collapse in its dimensions. On further increase in $[Mg^{2+}]$, intermediates separated by barriers appear on the FES, and one of the intermediates has a TPP ligand-binding competent structure. We show that site-specific binding of the Mg^{2+} aids in the formation of tertiary contacts. For $[Mg^{2+}]$ greater than physiological concentration, AD folds into its *holo*-form like structure even in the absence of the TPP ligand. The folding kinetics shows that it populates an intermediate due to the misalignment of the two arms in the TPP AD, which acts as a kinetic trap leading to larger folding timescales. The predictions of the intermediate structures from the simulations are amenable for experimental verification.

Introduction

The mechanism of gene regulation by riboswitches in bacteria is a fascinating problem. Riboswitches are non-coding RNA present in the 5′ untranslated region of the mRNA and are composed of two domains: the aptamer domain (AD) and the expression platform (EP). An interplay of the structural transitions between the AD and EP play a critical role in gene regulation.[1–3] Structural transitions between the two domains are influenced by Mg^{2+} and the binding of a cognate ligand to the AD.

The Mg^{2+} influence RNA folding through diffuse and site-specific binding.[4] In diffuse binding, Mg^{2+} non-specifically binds to the negatively charged phosphate groups and renormalizes charge on the RNA chain[5] leading to the initial collapse. Whereas in site-specific binding, Mg^{2+} bind in the pockets created by the specific arrangement of nucleotides and aid in the stabilization of secondary and tertiary structures.[6,7] However, for the functioning of riboswitches, in addition to the Mg^{2+}, cognate ligand binding is equally essential. The cognate ligand binding to the AD leads to structural changes in both the AD and EP domains, which signal the attenuation or initiation of the transcription or translation process. Riboswitches are important drug targets since they are involved in the critical functional role of gene regulation in bacteria and also absent in the human genome.[8,9]

Experiments[10–22] and simulations[23–31] probing the folding of ADs in the presence of Mg^{2+} and their cognate ligands provide evidence that the ADs populate multiple intermediates in their folding pathways. Functional studies on TPP,[32] Adenine,[33] Mg^{2+},[34] c-di-GMP,[35] SAM-I[36] and FMN[37,38] riboswitches further demonstrate that higher concentrations of cognate ligands are required than their intrinsic affinity to effectively regulate gene expression, indicating that riboswitch functioning is under kinetic control rather than thermodynamic control. Despite significant progress in our understanding of the functioning of riboswitches, the role of cations and cognate ligands in the mechanism of structural transitions in the AD and EP is not completely clear. An interesting question that requires comprehensive understanding is if Mg^{2+} alone is sufficient to populate the *holo*-form like structures in the absence of the cognate ligand under thermodynamic conditions? To address this question, we studied the role of $[Mg^{2+}]$ on the stability of intermediates populated in the folding free energy surface (FES) of the AD of thiamine pyrophosphate (TPP) sensing riboswitch.

The TPP riboswitch is widely distributed along the three phylogenic branches: bacteria, archaea, and eukarya. It is found in 48 out of the 59 human bacterial pathogens making it the most abundant riboswitch in human pathogens.[39] The TPP AD is $\approx$ 80 nt long, and in the native state (*holo*-form), it is a junction-type riboswitch with a tuning-fork-like structure[40] composed of two arms (Figure 1A and S1 in supporting information (SI)). The P_{2-3} arm is composed of P_2-P_3 helices joined by the junction J_{32}, and the P_{4-5} arm is composed of P_4-P_5 helices joined by the junction J_{45}, respectively. The P_1 helix acts as a base holding

both the arms from P_2 and P_4 helices (Figure 1A). The P_{2-3} arm has the binding pocket for the 4-amino-5-hydroxymethyl-2-methyl-pyrimidine (HMP) moiety of the TPP ligand. The P_{4-5} arm binds to the pyrophosphate (PP) moiety of the ligand through the mediation of Mg^{2+}.

Single-molecule FRET[12,18] and SAXS[41–43] studies have shown that physiological [Mg^{2+}] is indispensable for the cognate ligand to bind to the TPP AD. Single-molecule FRET[18] study proposed that a Y-shaped intermediate mediated by Mg^{2+} gets populated during the initial stages of folding even in the absence of the cognate TPP ligand. Upon binding of the TPP ligand, the AD attains the prerequisite folded state that regulates the downstream gene expression. Studies on other riboswitches reveal that the AD domains can sample the *holo*-form like conformations even in the absence of cognate ligand but in [Mg^{2+}] greater than the physiological concentration.[44,45] Experiments studying TPP AD folding also suggest that multiple structural ensembles are populated depending on [Mg^{2+}] and TPP-ligand.[12] Further, the timescales associated with the transitions among these ensembles varied from μs to s, indicating significant variations in the barrier heights separating the basins.

Computer simulations using all-atom[7,30,46–51] and coarse-grained models[52–61] are playing an important role in elucidating various aspects of RNA folding. The stability of folded state compared to the unfolded state of RNA is very sensitive to the charge and the concentration of the cations.[62,63] Mg^{2+} is known to be remarkably efficient in facilitating the formation of native-like tertiary contacts and the folded state even in the millimolar concentration. In contrast, submolar to molar concentration of K^+ is required for the monovalent ion-driven folding of RNA.[63–68] Mg^{2+} preferentially accumulates around the phosphate groups (P sites) of specific nucleotides and initiates folding by stabilization of the secondary structures and subsequently leading to the hierarchical formation of tertiary contacts.[56,69,70] During various stages of RNA folding, the inherently hierarchical nature of structural reorganization can result in the population of intermediates with an increase in [Mg^{2+}].

We studied the effect of [Mg^{2+}] on the folding FES of TPP AD using molecular dynamics simulations and a coarse-grained RNA model.[56] We find that with the increase in [Mg^{2+}], AD initially undergoes barrierless compaction in size followed by the formation of intermediate states. We show that Mg^{2+} binding to the P sites of specific nucleotides leads to the formation of tertiary contacts that stabilize the *holo*-form like folded state even in the absence of the TPP ligand. Interestingly, folding kinetics show that the AD can populate a kinetic trap due to the misaligned orientation of the arms leading to a larger folding time.

Methods

We studied the effect of [Mg^{2+}] on the folding mechanism of TPP AD from *Escherichia coli thiM* mRNA using the three interaction site (TIS) RNA model[56,71] and Langevin dynamics simulations. The TIS model for the TPP AD is constructed using the crystal structure (PDB: 2GDI).[40]

TIS Model of RNA: We used the TIS model developed by Denesyuk and Thirumalai,[56,71] where each nucleotide is represented by three sites mimicking the phosphate (P), sugar (S), and base (B) groups. The centers of the sites are placed at the center of mass of the P, S, and B groups, respectively. In this model, the monovalent (K^+, Na^+, and Cl^-) and divalent ions (Mg^{2+}) are explicitly present in the simulation. The Hamiltonian for the TIS model has seven components and is given by

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

The components in the Hamiltonian correspond to the bond length (U_{BL}), bond angle (U_{BA}), excluded volume repulsion between different sites (U_{EV}), electrostatic interaction between charged sites (U_{EL}), single-strand base stacking interaction between consecutive bases (U_{ST}), hydrogen bonding interaction (U_{HB}) and tertiary stacking interaction between two non-consecutive bases (U_{TST}), respectively. The force field details and parameters can be found in the works of Denesyuk and Thirumalai.[56,71] Below we provide a brief description of the force field.

The harmonic bond length and bond angle potentials account for the RNA chain connectivity and stiffness. The excluded volume repulsion between a pair of interacting sites, either RNA or ions, is given by a modified Lennard-Jones potential. The electrostatic interaction between a pair of charged sites is given by the Coulomb potential scaled by the temperature-dependent dielectric constant of water. The charge is -1 for phosphate group, -1 for Cl^-, +2 for Mg^{2+}, +1 for K^+, and +1 for Na^+ ions, respectively. The consecutive bases in the RNA chain have base stacking interactions, which depend on the sequence. The list of base pairs involved in the native hydrogen bond network in the folded structure of TPP AD is obtained using the crystal structure (PDB: 2GDI) and WHAT IF web server (https://swift.cmbi.umcn.nl)[72](Table S1-S4). All the base pairs, which have a hydrogen bond between them, interact using the hydrogen bonding potential. Non-canonical base pair hydrogen bonds and tertiary base stacking interactions present in the crystal structure are also taken into account in the model (Table S5-S6). This model is successful in quantitatively accounting for the folding thermodynamics of various RNA systems such as *Azoarcus* ribozyme,[56] the central domain of 16S ribosomal RNA[70] and RNA pseudoknots[68]

demonstrating that it is reliable and transferable to study various other RNA systems.

Simulations: We performed simulations to study TPP AD folding as $[Mg^{2+}]$ is varied from 1 mM to 6.5 mM. The $[K^+]$ is fixed at 30 mM in all the simulations. The simulations are performed in a cubic box of length 200 Å. The two Na^+ ions present in the crystal structure are added to the simulation box. The number of Mg^{2+} and K^+ ions in the simulation box are computed using their concentration and box volume. Cl^- ions are added to maintain charge neutrality in the simulation box. All ions are explicitly modeled as beads with charge and excluded volume. We used Langevin dynamics simulations to study the folding dynamics of TPP AD at temperature, $T = 310$ K. To compute the thermodynamic properties of TPP AD folding, we used 5% viscosity of water, $\eta = 5 \times 10^{-5}$ Pa·s to enhance conformational sampling. To study the TPP folding kinetics at $[Mg^{2+}] = 6.5$ mM, we used the viscosity of water, $\eta = 10^{-3}$ Pa·s. The equation of motion for the RNA sites and ions in Langevin dynamics is given by

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

where m_i is the mass (in amu), ζ_i $(= 6\pi\eta R_i)$ (in amu/fs) is the friction coefficient, 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 i^{th} site in the system. $\vec{\Gamma}_i$ is the random force on the i^{th} site with a white-noise spectrum. The random force auto-correlation function is given by $\left\langle \vec{\Gamma}(t) \cdot \vec{\Gamma}(t + nh) \right\rangle = 2\zeta k_B T \delta_{0,n}/h$, where $n = 0, 1, ...$, $\delta_{0,n}$ is 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$)$.[56,73] System coordinates are saved after every 5,000 steps ($\tau_f = 12.5$ ps) to compute the properties. The initial 1 μs of simulation data is ignored in computing the properties. For each $[Mg^{2+}]$, at least ≈ 13 μs of simulation data is collected to compute the average properties. The atomistic coordinates of the AD are generated using the coarse-grained coordinates and TIS2AA[74] program, which uses the fragment-assembly approach[75] and energy minimization embedded in AmberTools.[76] We used VMD to generate the three dimensional structures of the AD.[77]

Data Analysis: The radius of gyration R_g, of the TPP AD is computed using the equation, $R_g = \left(\frac{1}{2N^2}\sum_{i,j}\vec{r}_{ij}^2\right)^{1/2}$, where N $(= 240)$ is the total number of RNA sites in the TIS model of TPP AD, and $\vec{r}_{ij}$ is the vector connecting the sites i and j. The average fraction of helix formation for a given helix H, is computed using the equation

$$\langle f_{HF} \rangle = \left\langle \frac{N_{hb}^H(i)}{N_{hb}^H(cry)} \right\rangle, \tag{3}$$

where $N_{hb}^H(i)$ is the number of hydrogen bonds present in the helix H in i^{th} conformation

and $N_{hb}^{H}(cry)$ is the total number of hydrogen bonds present in the helix H in the crystal structure. $\langle\ \rangle$ denotes the average over all the conformations. A hydrogen bond is considered to be present if its energy is lower than the thermal energy ($k_{\mathrm{B}}T$).[56]

Local [Mg^{2+}] Around P sites: The local [Mg^{2+}] in the vicinity of the P site of the i^{th} nucleotide in molar units is computed[69] using the relation

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

where $\rho_i(r)$ is the number density of the Mg^{2+} at a distance r from the i^{th} P site, V_c is the spherical volume of radius r_c, and N_{A} is the Avogadro's number. The cutoff radius r_c is given by $r_c = R_{Mg} + R_P + \Delta r$, where R_{Mg} and R_P are the radii of Mg^{2+} and P sites, and Δr is the margin distance. We used $\Delta r = 1.7$ Å to ensure that we take into account only the tightly bound local Mg^{2+} around the P sites.

FES Calculation: The total number of native contacts in the TPP AD is computed using the TIS model of the folded crystal structure (PDB: 2GDI).[40] A pair of sites i and j in the folded 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 (Figure S1). The f_{NC} for i^{th} TPP AD conformation is computed using the equation

$$f_{NC} = \frac{N_{nc}(i)}{N_{nc(cry)}}, \tag{5}$$

where $N_{nc}(i)$ the number of native contacts present in i^{th} conformation and $N_{nc}(cry)$ is the number of native contacts present in the crystal structure.

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

$$G = -k_{\mathrm{B}}T \ln(P(f_{NC})), \tag{6}$$

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

Fraction of Tertiary Contacts (TCs): The TPP AD has two critical tertiary contact (TC) forming sites, the three-way junction (TC_{3WJ}) and the arm-tip (TC_{AT}). The TC_{3WJ} involves contacts between the nucleotide segments A12 - U20, C48 - U59, and A80 - G86 (Figure S2A). The TC_{AT} involves contacts between the nucleotide segments G21 - C24, and A69 - G72 (Figure S2A). Apart from the two TCs mentioned above, TPP AD has a third tertiary contact (TC_{RB}) between U54 (J_{24}) and U79 (J_{45}) located at the base of the P_{4-5} arm. The total number of tertiary interactions in a particular TC in the TPP AD native

state is equal to the number of native contacts between the sites belonging to the nucleotides involved in forming that specific TC. The fraction of tertiary contacts f_{TC} for a particular TC is computed using the equation

$$f_{TC} = \frac{N_{nc}^{TC}(i)}{N_{nc}^{TC}(cry)}, \tag{7}$$

where $N_{nc}^{TC}(i)$ is the number of native contacts present between the sites belonging to the nucleotides involved in the formation of that specific TC in the i^{th} conformation and $N_{nc}^{TC}(cry)$ is the total number of native contacts present in that specific TC in the crystal structure. We label a TC as formed (F) if $\langle f_{TC} \rangle$ is ≥ 0.5 and as ruptured (R) otherwise.

Role of Mg^{2+} in Tertiary Contact Formation: The contribution of Mg^{2+} binding to the P sites in the formation of TCs is inferred by computing the free energy difference ($\Delta\Delta G_{TC}$) in the TC formation with and without Mg^{2+} binding.[70] The free energy contribution to the TC formation due to Mg^{2+} binding to the i^{th} P site is given by

$$\Delta\Delta G_{TC}(i) = -k_{\mathrm{B}}T \ln \left(\frac{P_{TC,i}(\mathrm{F}, \mathrm{Mg^B})\ P_{TC,i}(\mathrm{R}, \mathrm{Mg^U})}{P_{TC,i}(\mathrm{R}, \mathrm{Mg^B})\ P_{TC,i}(\mathrm{F}, \mathrm{Mg^U})} \right), \tag{8}$$

where $P_{TC,i}(\mathrm{F}, \mathrm{Mg^B})$ is the joint probability that TC is formed (F), and Mg^{2+} is bound (Mg^B) to the i^{th} P site, $P_{TC,i}(\mathrm{R}, \mathrm{Mg^U})$ is the joint probability that TC is ruptured (R) and Mg^{2+} is not bound (Mg^U) to the i^{th} P site. Similarly $P_{TC,i}(\mathrm{R}, \mathrm{Mg^B})$ and $P_{TC,i}(\mathrm{F}, \mathrm{Mg^U})$ are defined. An Mg^{2+} is considered bound to the i^{th} P site if the distance between the centers of the P site and ion is less than 4.6 Å.

TPP AD Folding Kinetics: To decipher the kinetic intermediates populated in the TPP AD folding pathways, we computed an inter-arm angle (Ω_{IA}) between the helical arms to characterize their orientation. The arm with helices P_2 and P_3 is represented using a vector ($\vec{r}_1$) joining the center of mass of nucleotides C15 and G51, and the center of mass of nucleotides C22 and G37. Similarly, the other arm with helices P_4 and P_5 is represented using a vector ($\vec{r}_2$) joining the center of mass of nucleotides C57 and G82, and the center of mass of nucleotides U64 and A75. The inter-arm angle Ω_{IA} is defined to be the angle between the vectors $\vec{r}_1$ and $\vec{r}_2$ (Figure S2B). We computed a dihedral angle (Θ_J) to characterize the orientation of the junction J_{24}. Θ_J is computed using the position of the S sites of the nucleotides U14, G51, C57 and G86 belonging to P_1, P_2, P_4, and P_1 helices, respectively. The length of J_{24} motif ($R_{J_{24}}$) is the distance between the P sites of the nucleotides G51 and C57 (Figure S2B).

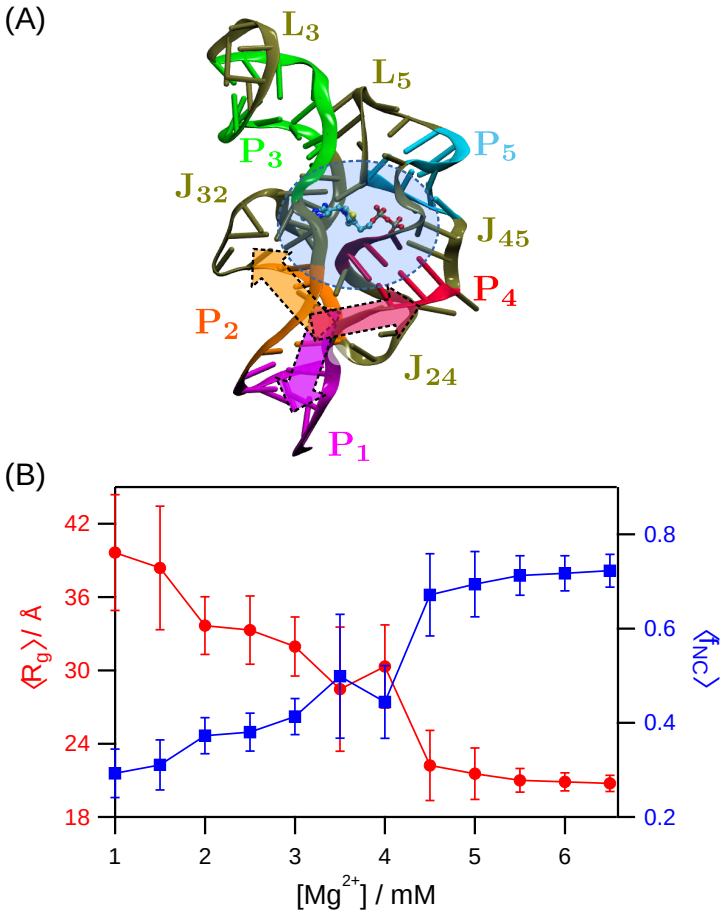

Figure 1: (A) Structure of the TPP AD (PDB: 2GDI[40]). Helices P_1 to P_5 are shown in magenta, orange, green, red, and cyan, respectively. Helical junctions (J_{32}, J_{45} and J_{24}) and loops (L_3 and L_5) are shown in tan. The core region of the AD containing the TPP ligand binding pocket is shown as a shaded area in blue. The bound TPP-ligand inside the core region is shown using ball-stick representation. The three-way junction (3WJ) is shown by the Y-shaped shaded area near the joint of P_1 and P_2 helices. (B) The average radius of gyration, $\langle R_g \rangle$ and the average fraction of native contacts, $\langle f_{NC} \rangle$ are plotted as a function of $[Mg^{2+}]$ using solid red circles and blue squares, respectively. Multiple transitions in both of the plots indicate that TPP AD is a multi-state folding system.

Results

Mg^{2+} Facilitates TPP AD Folding

In physiological conditions, noncoding RNA molecules fold to a unique native state to perform their regulatory activities. RNA molecules have a cooperative interaction network,[78,79] which guides folding to the native state and suppresses the formation of non-native structures from the early stages of folding.[80] The cooperativity in RNA folding further depends on the ions present in solution.[81–89] To study the effect of Mg^{2+} on TPP AD folding, we computed $\langle R_g \rangle$ and $\langle f_{NC} \rangle$ as a function of $[Mg^{2+}]$ (Figure 1B). As we increase the $[Mg^{2+}]$ from 1 mM to 6.5 mM, the TPP AD undergoes a transition from an unfolded state ($\langle R_g \rangle \approx$ 40 Å and $\langle f_{NC} \rangle \approx 0.3$) to a compact folded state ($\langle R_g \rangle \approx 21$ Å and $\langle f_{NC} \rangle \approx 0.7$). The plot shows two major folding transitions (Figure 1B). The first transition occurs in the range 1.5 mM $< [Mg^{2+}] <$ 2 mM, where $\langle R_g \rangle$ decreases from $\approx$ 38 Å to 34 Å ($\langle f_{NC} \rangle$ changes from 0.31 to 0.37), and the second transition occurs in the range 4 mM $< [Mg^{2+}] <$ 4.5 mM, where $\langle R_g \rangle$ decreases from $\approx$ 30 Å to 22 Å ($\langle f_{NC} \rangle$ changes from 0.44 to 0.67). The multiple transitions indicate that TPP AD populates intermediates when folding is initiated by increasing $[Mg^{2+}]$. The changes in $\langle R_g \rangle$ and $\langle f_{NC} \rangle$ with the increase in $[Mg^{2+}]$ shows that TPP AD can fold to its native-like state only beyond physiological $[Mg^{2+}]$ ($\approx$ 4 mM[90–92]) in the absence of the TPP ligand.

Folding Intermediates and *Holo*-Form Like Structure

The FES projected onto f_{NC} shows that AD populates four different states depending on the $[Mg^{2+}]$ (Figure 2). As $[Mg^{2+}]$ is increased, the equilibrium shifted towards populating compact intermediate states. Experiments have shown that Mg^{2+} contributes to both specific and non-specific compaction in the dimensions of the RNA unfolded state.[18,41,85,93,94] In the early stages of folding, we observed barrierless collapse of the AD as $[Mg^{2+}]$ is increased from 1 mM to 2 mM, where the FES retains its shape and only its minimum is shifted from f_{NC} = 0.29 to 0.38, indicating compaction in the AD size (Figure 2). The decrease in $\langle R_g \rangle$ and the disrupted tertiary contacts confirm compaction in the AD size without the formation of any tertiary structure (Figure 1B and 3A). For $[Mg^{2+}] \leq$ 2 mM, the AD unfolded state is the global minimum on the FES. Upon increasing $[Mg^{2+}]$ ($\geq$ 3 mM), intermediate states separated by energy barriers appear on the FES along with the shift in energy minima to higher f_{NC}. For $[Mg^{2+}] \geq$ 5 mM, the global minimum on the FES is the AD native-like folded state (*holo*-form) with $f_{NC} \geq 0.7$.

For $[Mg^{2+}]$ = 3 mM, the TPP AD exists in three states: unfolded ($f_{NC} = 0.32$), open-arm Y-shaped ($f_{NC} = 0.41$) and closed-arm Y-shaped ($f_{NC} = 0.55$) (Figure 2). As $[Mg^{2+}]$

increases to 5 mM, the FES shows that the TPP AD populates four states: unfolded (f_{NC} = 0.3), open-arm (f_{NC} = 0.44), closed-arm (f_{NC} = 0.58) and *holo*-form like folded state (f_{NC} = 0.71) (Figure 2). However, with the increase in [Mg^{2+}] from 3 mM to 5 mM, the most stable state shifts from the open-arm Y-shaped (f_{NC} = 0.41) to the *holo*-form state (f_{NC} = 0.71) (Figure 2). The increase in stability of the *holo*-form like state (f_{NC} = 0.71) for [Mg^{2+}] $\geq$ 4.5 mM is due to the formation of both TC_{3WJ} and TC_{AT} (Figure 3A). The transitions between the open and closed-arm Y-shaped intermediates indicate that at [Mg^{2+}] = 3 mM, AD populates conformations, which can either facilitate or hinder the formation of the ligand binding pocket.

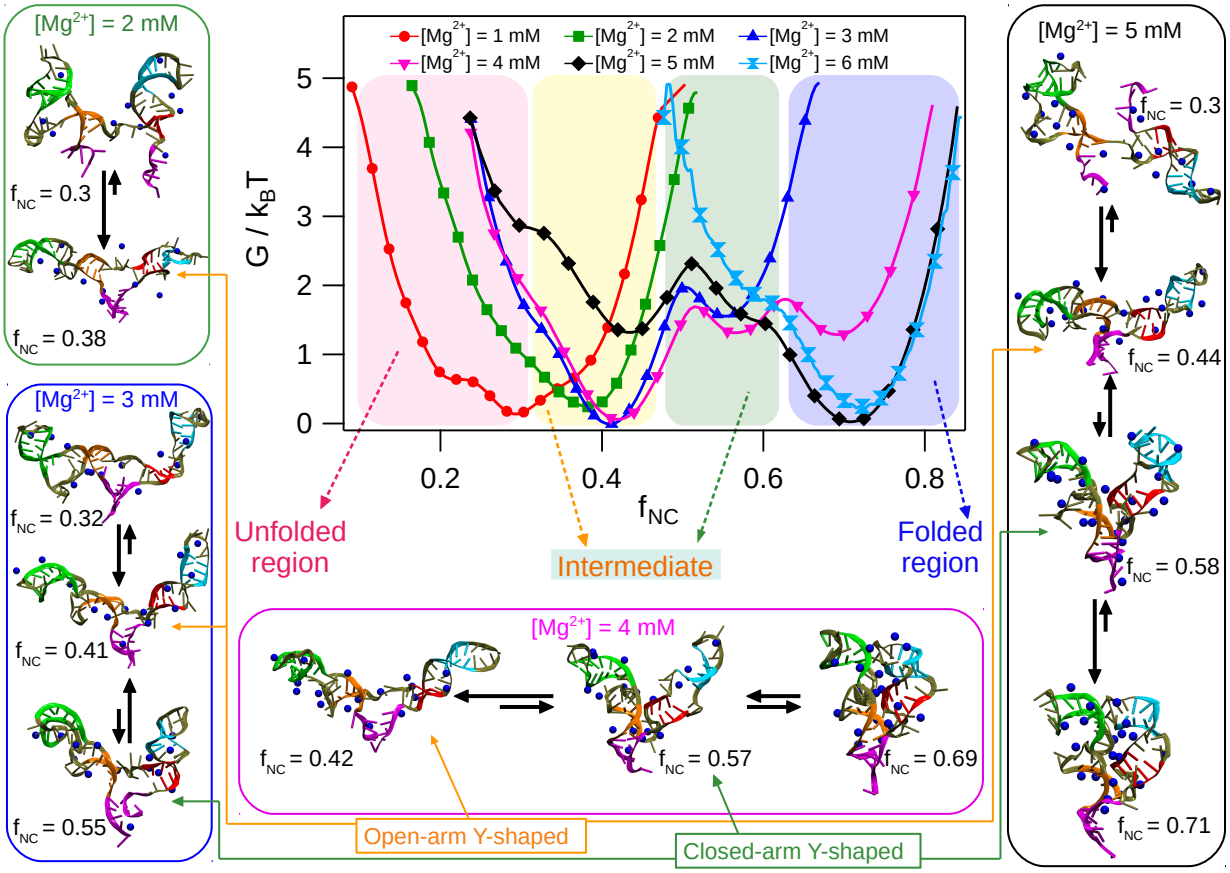

Figure 2: FES of the TPP AD projected onto the fraction of native contacts, f_{NC}. The FES for [Mg^{2+}] = 1 to 6 mM are plotted using solid red circle, green square, blue triangle, magenta inverted-triangle, black rhombus, and cyan hour-glass markers, respectively. The FES is divided into four regions. The unfolded and folded regions are shown in red and blue. The region showing intermediates with open and closed-arm Y-shaped structures are shown in yellow and green. Representative conformations of the AD in each basin of the FES are shown in cartoon representation. P_1 to P_5 helices are shown in magenta, orange, green, red, and cyan colors, respectively. The condensed Mg^{2+} on the TPP AD are shown as blue beads. The equilibrium between structures in different basins are shown for different [Mg^{2+}].

Stability of the Helical Arms

Out of all the five helices present in the AD, the P_1 helix located at the basal region, which holds the two helical arms together, is thermodynamically the least stable helix and remains unfolded for $[Mg^{2+}] < 2$ mM (Figure 1A, 2 and 3A). The lower stability of the P_1 helix was also observed in previous studies.[25,95] A single-molecule FRET experiment[18] has shown that the P_1 helix is highly dynamic and can exist in the unfolded form in the absence of Mg^{2+}. Unfolded P_1 helix leads to unzipping of the P_4 helix in low $[Mg^{2+}]$ (< 2 mM) (Figure 2). The two arms of the AD retain their shape as the helices P_2, P_3, and P_5 remain stable even in low $[Mg^{2+}]$ ($= 1$ mM) (Figure 3A). Monovalent cations ($[K^+] = 30$ mM), which neutralize the RNA backbone charge, are sufficient to stabilize the P_2, P_3 and P_5 secondary structures. However, the P_1 helix is unfolded in the presence of K^+ ions and low $[Mg^{2+}]$. When $[Mg^{2+}] \geq 2$ mM, the P_1 helix is stabilized, and it further stabilizes the P_{2-3} and P_{4-5} helical arms forming a Y-shaped structure (Figure 2). Fluorescence spectroscopy experiment[96] has shown that at physiological $[Mg^{2+}]$, P_1 helix orients the P_{2-3} and P_{4-5} arms forming a three-way-junction (3WJ).

Site-Specific Mg^{2+} Binding Facilitates TC Formation

To understand the role of Mg^{2+} in TPP AD folding, we computed the average local Mg^{2+} concentration ($\langle c^* \rangle$) around the AD at nucleotide resolution (Eq. 4). We found that Mg^{2+} are not randomly diffused along the RNA backbone but are bound to specific nucleotides and assist the formation of tertiary contacts.[4,56,69,97,98] When $[Mg^{2+}] = 1$ - 2 mM, Mg^{2+} are dominantly distributed around the nucleotides, N_N (nucleotide number) = A43 to A45, which belong to the J_{32} junction (Figure 3B). Due to the lack of Mg^{2+} in the core region and the three-way-junction (3WJ) (Figure 1A, 3B), the electrostatic repulsion between the P_2, P_4 helices and J_{24} junction (basal inter-arm motif) hinders the close approach of the P_{2-3} and P_{4-5} arms to form TCs and the folded state (Figure 2).

As $[Mg^{2+}]$ is increased to 3 - 4 mM, Mg^{2+} accumulates preferentially around the core region (cavity formed by the P_3, J_{32}, P_2, J_{24}, P_4 and P_5 motifs) and the 3WJ region (N_N = G33 - A35; A43 - A45; C48 - C49; U54 - A56; C57 - C58; C74 - G76; U14 - G16, A84 - G86, respectively) (Figure 3B, S2B). The condensed Mg^{2+} decrease the effective negative charge on the P sites of the basal inter-arm motif weakening the electrostatic repulsion between the P_{2-3} and P_{4-5} arms. As a result, the helical arms approach each other, and the Y-shaped intermediate state is populated (Figure 2).

When $[Mg^{2+}]$ is large (≥ 5 mM), Mg^{2+} are condensed in the core and 3WJ regions with comparatively increased binding preference to the P_3 (N_N = C23 - C24), J_{32} (N_N = G19 - U20 and U39 - A41 and A43 - A45), P_2 (N_N = G17 - G18 and C48 - G51), J_{24} (N_N = C55

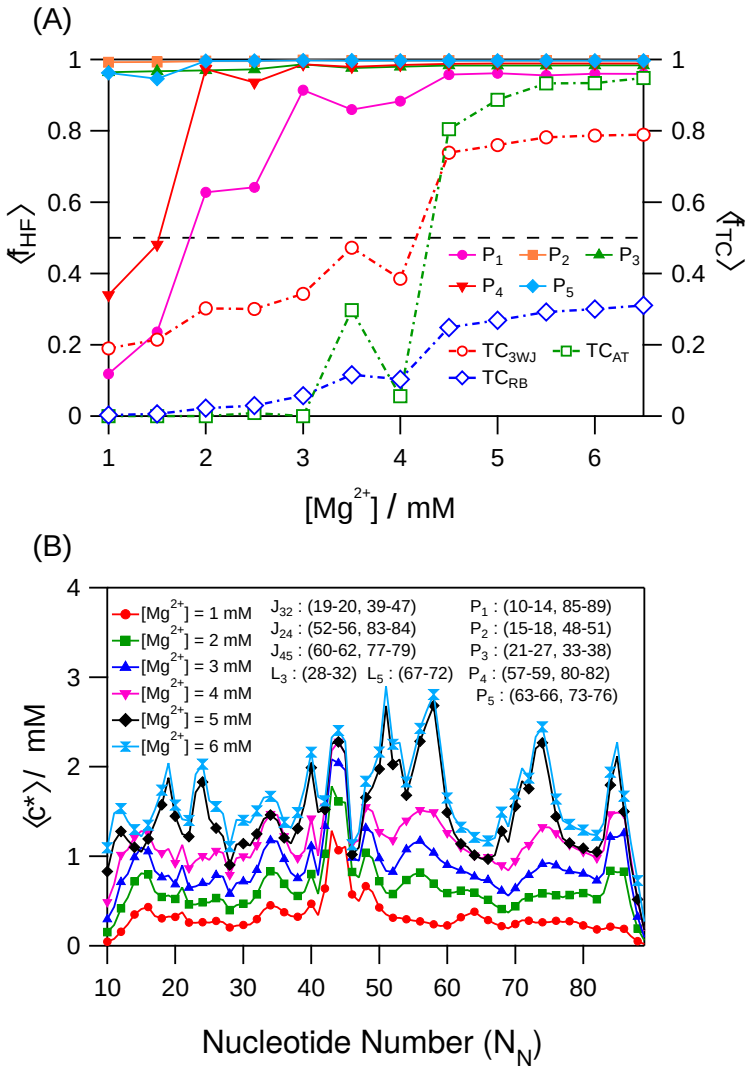

Figure 3: (A) Average fraction of helix formation ($\langle f_{HF} \rangle$) and average fraction of tertiary contact formation ($\langle f_{TC} \rangle$) plotted as a function of $[Mg^{2+}]$. $\langle f_{HF} \rangle$ for helices P_1 to P_5 is plotted with solid magenta circle, orange square, green triangle, red inverted triangle, and cyan rhombus markers, respectively. Helices P_2, P_3 and P_5 are stable for all $[Mg^{2+}]$. Helices P_1 and P_4 are stable for $[Mg^{2+}] \geq 2$ mM. $\langle f_{TC} \rangle$ for TC_{3WJ}, TC_{AT} and TC_{RB} are plotted with hollow red circles, hollow green squares, and hollow blue rhombus markers, respectively. (B) The average local concentration of Mg^{2+} around individual P sites ($\langle c^* \rangle$) is plotted as a function of the nucleotide number (N_N). The data for $[Mg^{2+}] = 1$ to 6 mM are shown in solid red circle, green square, blue triangle, magenta inverted triangle, black rhombus, and cyan hourglass markers, respectively. The N_N range (starting from 10 as in the PDB[40]) for the helices (P's), junctions (J's), and loops (L's) are in the annotation. For $[Mg^{2+}] = 1$ mM, Mg^{2+} condenses around the J_{32} junction on the P_{2-3} arms. As $[Mg^{2+}]$ increases, it accumulates around the core and 3WJ regions leading to the AD folded structure.

- A56), P_4 (N_N = C57 - U59), P_5 (N_N = C73 - A75), and L_5 (N_N = U71 - G72) motifs (Figure 3B). The increased accumulation of Mg^{2+} on the helical arms and the J_{24} junction further weakens the inter-arm electrostatic repulsion and allows the arms to form TCs, which stabilizes the *holo*-form like folded structure even in the absence of the TPP ligand.

The nucleotides with high preferential binding of Mg^{2+}, A43 and C74 to G76 are involved in the TPP ligand binding to the AD. A43 forms stacking interaction with the HMP group of the TPP ligand, and C74 - G76 bind to the pyrophosphate tail of TPP ligand bridged by Mg^{2+} (Figure S3).[40] Mg^{2+} prefers binding to A43 even in low [Mg^{2+}], whereas the ions show binding preference towards C74 - G76 when [Mg^{2+}] $\geq$ 3 mM. The fact that Mg^{2+} prefers to bind to the nucleotides involved in TPP ligand recognition and binding illustrates that Mg^{2+} aids in the formation and stabilization of the scaffolding for ligand binding pocket.[18,40] The SAXS experiments[42,43] on *E. coli thiM* mRNA show that in high [Mg^{2+}] ($\approx$ 10 mM), TPP AD samples compact conformations similar to the *holo*-form like structures, whose R_g is greater than the ligand-bound structures by $\approx$ 2 Å. The experiments provide evidence that the T-loop (U39 - A47) region where the TPP ligand binds (Figure S2B) could be partially unfolded in the absence of the ligand.

TC Formation is Cooperative

We computed the average fraction of TCs in TPP AD, $\langle f_{TC} \rangle$, to probe the cooperativity and hierarchy in TC formation with the variation in [Mg^{2+}]. The two TCs located at the 3WJ (TC_{3WJ}) and the arm-tip of AD (TC_{AT}) exhibit [Mg^{2+}] dependent formation (Figure 3A). Both the TCs form cooperatively and show a sharp transition in the range 4 mM $<$ [Mg^{2+}] $<$ 4.5 mM, which coincides with the second transition observed in both $\langle R_g \rangle$ and $\langle f_{NC} \rangle$ (Figure 1B) establishing that the compaction in AD dimensions is due to the formation of these TCs. The enhanced stability of the Y-shaped intermediates and the *holo*-form like state require the complete formation of these two TCs. The TC_{3WJ} forms first and stabilizes the two helical arms, which subsequently approach each other to form TC_{AT} (Figure 3A). The tertiary structures do not form in the early stages of folding but during the AD assembly to the native-like state in the late stages of folding shows that TPP AD follows the quasi-hierarchical folding model.[99] TC_{RB} remains disrupted even in high [Mg^{2+}]. We hypothesize that TPP ligand binding to the AD is essential to stabilize TC_{RB}. The crystal structure of the TPP AD shows that the pyrophosphate domain of the TPP ligand interacts with G78 (J_{45}) mediated by Mg^{2+} and water molecules (Figure S3C).

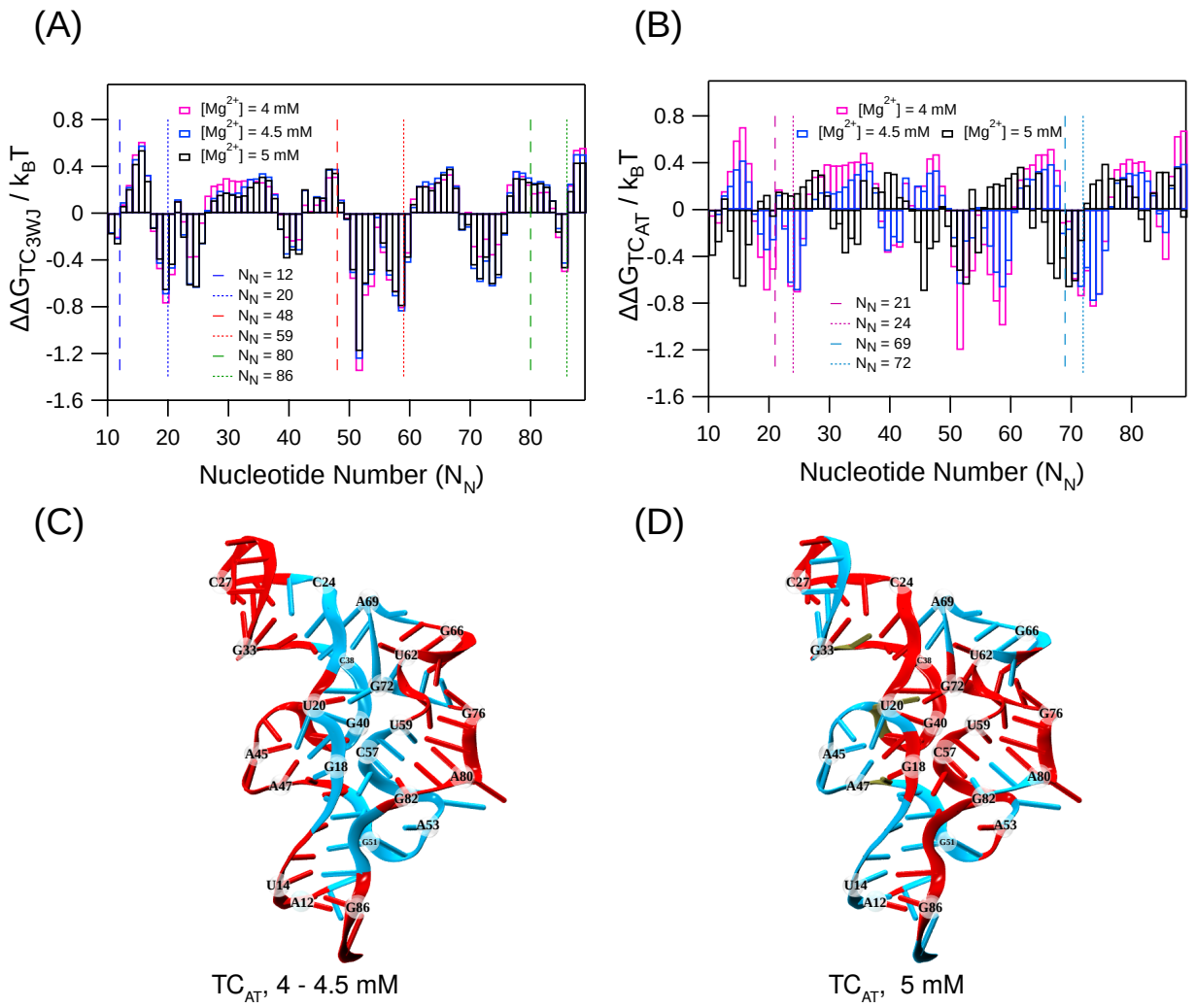

Figure 4: $\Delta\Delta G_{TC}$ for (A) TC_{3WJ} and (B) TC_{AT} are plotted as a function of N_N for $[\mathrm{Mg}^{2+}]$ = 4 mM, 4.5 mM and 5 mM in magenta, blue, and black bars, respectively. The dashed and dotted lines mark the range of nucleotides involved in the formation of a particular TC. Nucleotides, which facilitate ($\Delta\Delta G_{TC}(i) < 0$) and hinder ($\Delta\Delta G_{TC}(i) > 0$) the formation of TC_{AT} formation upon Mg^{2+} binding are shown in cyan and red colored cartoon representation for (C) $[\mathrm{Mg}^{2+}] = 4$ and 4.5 mM (D) $[\mathrm{Mg}^{2+}] = 5$ mM.

Mg^{2+} Binding to Specific P Sites Decreases the Free Energy for TC Formation

To decipher the contribution of site-specific Mg^{2+} binding to the formation of a TC, we computed the difference in the stability of that TC due to Mg^{2+} binding and unbinding to the P site of i^{th} nucleotide, $\Delta\Delta G_{TC}(i)$ (Eq. 8).[70] Negative value of $\Delta\Delta G_{TC}(i)$ for a TC indicates that Mg^{2+} binding to the i^{th} P site favors the formation of that TC. We computed $\Delta\Delta G_{TC}(i)$ for $[Mg^{2+}] = 4$ mM, 4.5 mM and 5 mM, which correspond to the second transition in AD folding where TC formation is observed (Figure 1B and 3A). For $[Mg^{2+}] = 4$ mM, conformations with both the TCs in the ruptured state are predominantly populated. With the increase in $[Mg^{2+}]$ ($\geq$ 4.5 mM), the AD dominantly populates conformations with the formed TCs (Figure S4).

Although, AD with the disrupted TC_{3WJ} is the most stable state at $[Mg^{2+}] = 4$ mM, Mg^{2+} are already bound to the P sites located in the 3WJ region (N_N = G19, C50 - A53, A56 - C58, A85) (Figure 3A,B). The bound Mg^{2+} significantly favor the TC_{3WJ} formation as $\Delta\Delta G_{TC}(i) <$ -0.5 k_BT for i belonging to the 3WJ nucleotides (Figure 4A). With the increase in $[Mg^{2+}]$ ($\geq$ 4.5 mM), the probability of TC_{3WJ} formation increased (Figure 3A). However, the $\Delta\Delta G_{TC}(i)$ values remain almost unchanged (Figure 4A), but the $\langle c^* \rangle$ around the N_N forming the TC_{3WJ} shows a significant increase with the increase in $[Mg^{2+}]$ ($\geq$ 4.5 mM) (Figure 3B, S5). We hypothesize that at $[Mg^{2+}] = 4$ mM, as lower number of Mg^{2+} accumulate in the 3WJ region, the backbone electrostatic repulsion may still prevail in the region and hinder the TC_{3WJ} formation. With increasing $[Mg^{2+}]$ ($\geq$ 4.5 mM), sufficient number of Mg^{2+} accumulate in the 3WJ region (Figure 3B) and diminish the backbone electrostatic repulsion to stabilize the TC_{3WJ} (Figure 3A, 4A).

During the second transition in AD folding ($[Mg^{2+}] = 4$ - 4.5 mM), TC_{AT} formation is favored by Mg^{2+} binding to the nucleotides located at the interior region mainly to the nucleotides located at both 3WJ (N_N = G19 - U20, C50 - A53, A56 - U59 and A85) and arm-tip (N_N = C22 - C23, A70 - G72) regions of the AD (Figure 4B,C). The lower value of $\Delta\Delta G_{TC_{AT}}$ ($<$ -0.9 k_BT) due to Mg^{2+} binding to the 3WJ nucleotides (N_N = G51 - C58) compared to the arm-tip nucleotides, indicates that Mg^{2+} binding in the 3WJ region can expedite the TC_{AT} formation, located at a distant site through allosteric interactions. At $[Mg^{2+}] = 4.5$ mM, AD dominantly populates the conformations with formed TC_{AT} (Figure 3A). Mg^{2+} binding to both 3WJ (N_N = G51, C58) and arm-tip (N_N = C23 - C24, A70 - G72) nucleotides almost equally contribute towards TC_{AT} formation. After the second transition ($[Mg^{2+}] = 5$ mM), TC_{AT} formation is governed by the Mg^{2+} binding to the nucleotides located at the periphery of AD (Figure 4B,D). Most of these peripheral nucleotides, N_N = C30 - U32 (L_3), A44 - A47 (J_{32}) and A67 - U71 (L_5), belong to motifs devoid of any three-dimensional structures. The favorable free energy for TC_{AT} formation due to Mg^{2+} binding

to the peripheral nucleotides can probably lead to the proper orientation of the junction and loop region so that AD can fold to its *holo*-form like conformation.

$\Delta\Delta G_{TC}(i)$ data reveals that the nucleotides located near the 3WJ region facilitate the AD folding to its *holo*-form like state by stabilizing both the TCs. The stabilization of TCs due to the site-specific Mg^{2+} binding to the nucleotides in the 3WJ region irrespective of $[Mg^{2+}]$ establishes that Mg^{2+} binds to specific nucleotides of an RNA molecule that drive and stabilize the TC formation leading to the RNA folded state.[70]

TPP AD Folds Through Slow and Fast Folding Pathways

We spawned 60 folding trajectories to study the AD folding kinetics in the absence of the TPP ligand starting from different unfolded conformations at $T = 310$ K and $[Mg^{2+}] = 6.5$ mM. The initial unfolded conformations are taken from the $[Mg^{2+}] = 1.0$ mM simulation, where the unfolded state is the most stable state (Figure 2). Upon initiating folding, we find that the AD folds through fast and slow folding pathways. We labeled trajectories where TPP AD folds in less than 0.5 μs as fast folding pathways and the other trajectories as slow folding pathways. 16 out of the 60 spawned folding trajectories follow the fast folding pathway and the rest fold through the slow folding pathway. In the slow folding pathway, an intermediate, which acts as a kinetic trap, is populated, leading to a longer folding time (Figure 5B,D).

The helices P_2, P_3 and P_5 are stable in the unfolded state at $[Mg^{2+}] = 6.5$ mM (Figure 3A). The Mg^{2+} mediated formation of the least stable switch helix P_1 is essential for the formation of 3WJ and open-arm Y-shaped intermediate for the folding of the AD in the absence of ligand. The initial barrier that AD has to overcome to fold is the relative orientation of the two helical arms P_{2-3} and P_{4-5} with respect to each other. In the early stages of folding, the arms fluctuate relative to each other as inferred from the inter-arm angle, Ω_{IA} (see Methods) (Figure 5A,B and 6A,A′). Stochastic fluctuations rotate the P_{4-5} arm around the J_{24} junction as an axis to align it parallel to the P_{2-3} arm (Figure 6B,B′). The parallel alignment of both the arms also facilitates the formation of P_1 helix at the base (Figure 6C_1,C′). The fluctuations in Ω_{IA} show that the parallel orientation of the helical arms is necessary but not sufficient for the AD to proceed to fold to its native state (Figure 5A,B).

Misaligned Orientation of P_2-$J_{24(L)}$-P_4 and P_1-P_2/P_4-$J_{24(S)}$-P_1 Motifs Leads to a Kinetic Trap

The kinetic trap in the slow folding pathway is due to the population of a conformation where although both the arms are aligned parallel to each other, their orientation is misaligned, and this conformation with a lifetime greater than 1 μs acts as a kinetic trap (Figure 5B,D and

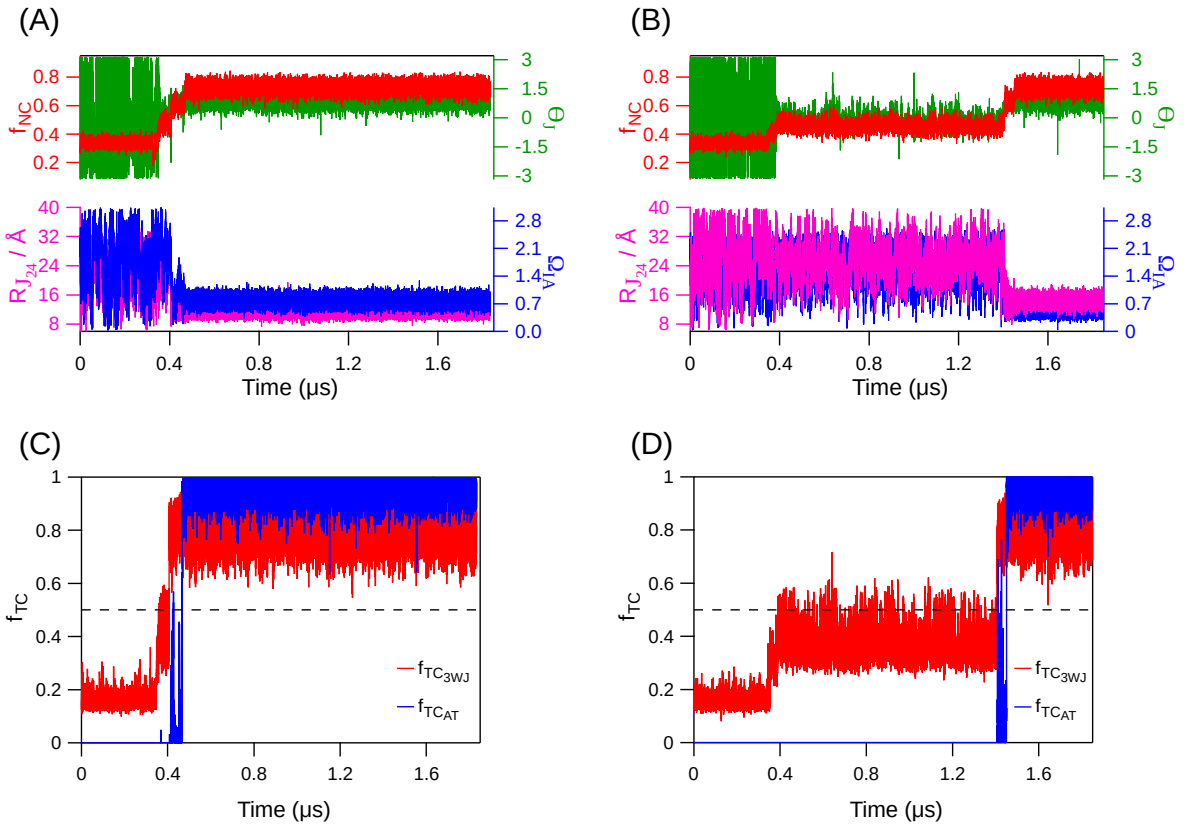

Figure 5: Fast and slow folding TPP AD pathways at $T = 310$ K and $[Mg^{2+}] = 6.5$ mM. Fraction of native contacts (f_{NC}), dihedral angle between the two strands of J_{24} motif (Θ_J), inter-helical arm angle (Ω_{IA}), and length of the J_{24} motif ($R_{J_{24}}$) are plotted for (A) fast and (B) slow folding pathways with red, green, blue, and magenta solid lines, respectively. For both the pathways, Ω_{IA} and $R_{J_{24}}$ exhibit large fluctuations till AD folds to the native state. Fluctuations in Θ_J decrease after the formation of the P_1 helix and fluctuate around 0.75 rad as the AD folds to the native state. Fraction of TC formation (f_{TC}) as a function of time is plotted for (C) fast and (D) slow folding pathways. $f_{TC_{3WJ}}$ and $f_{TC_{AT}}$ for both fast and slow folding trajectories are shown in red and blue solid lines, respectively. The TCs form during the late stages of folding and follow a hierarchical trend. In the slow folding trajectory, the kinetic trap is observed in the time interval 0.39 μs to 1.41 μs, where Θ_J fluctuates around -0.1 rad and f_{TC} fluctuates between 0.4 - 0.5 rad.

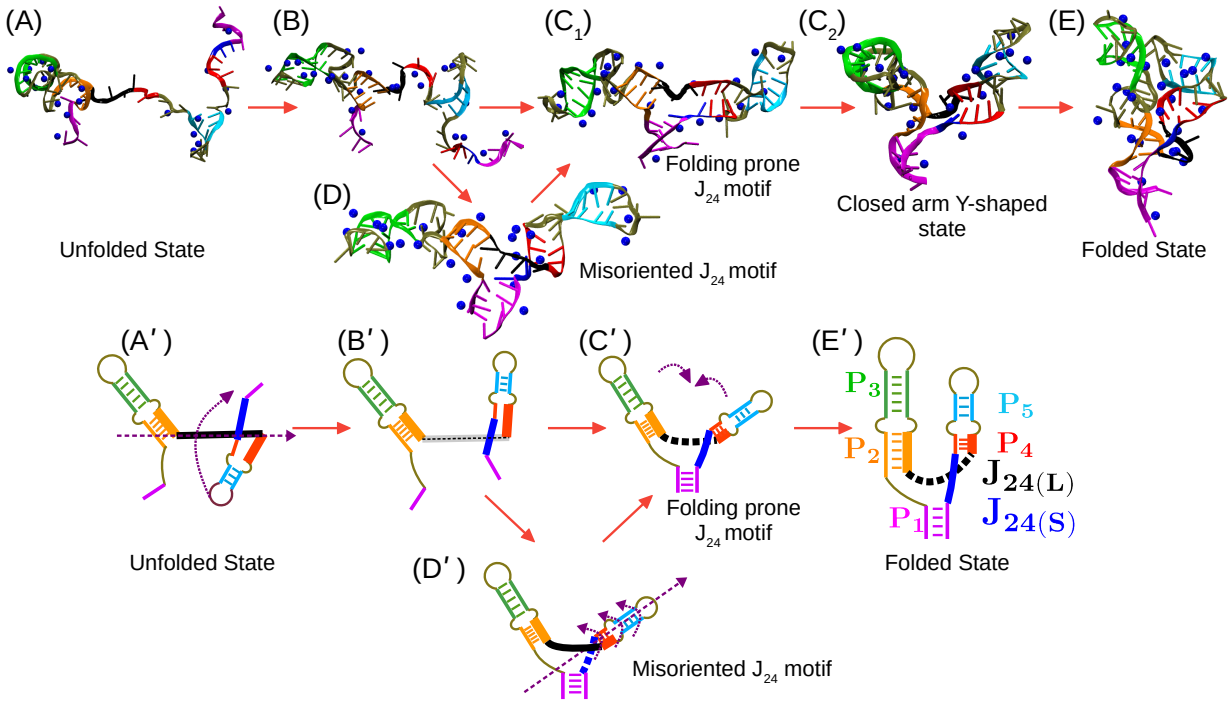

Figure 6: Intermediates populated during AD folding are shown in cartoon and schematic representations. Helices P_1 to P_5 are shown in magenta, orange, green, red, and cyan, respectively. The blue spheres are Mg^{2+} within 5 Å from the RNA. $J_{24(L)}$ and $J_{24(S)}$ are shown in black and blue lines, respectively. The rotational axis, the direction of rotation, and helical movement are shown as purple dashed and dotted arrows, respectively. The shaded black dot line in (B′) represents that the relative orientation between $J_{24(L)}$ and $J_{24(S)}$ motifs cannot be clearly demarcated as the P_4 helix is not formed at that stage of folding. The red arrows show the progression of folding events with time. (A) and (A′) Helical-arms are not in parallel orientation with respect to each other. The P_{4-5} arm rotates around the $J_{24(L)}$ motif to attain the folding prone conformation (B) and (B′). The AD at stage (B) and (B′) has neither adopted the folding prone conformations ($C_{1,2}$) and (C′) of $J_{24(L)}$ motif nor the misoriented (D) and (D′). In (C_1) and (C′), the P_2-$J_{24(L)}$-P_4 and P_1-P_2/P_4-$J_{24(S)}$-P_1 motifs (see SI for structural details) do not cross each other and the AD folds to form closed-arm Y-shaped conformation (C_2), which eventually leads to the *holo*-form like folded state (E) and (E′). (D) and (D′) The AD is stuck in the kinetic trap and to escape from the trap, $J_{24(L)}$ motif, P_4-P_5 helices rotate around the P_{4-5} arm as an axis to attain the folding prone structure.

6D,D′). In the kinetic trap, the nucleotides G51 - C57 (P_2-$J_{24(L)}$-P_4 motif) are misaligned with respect to the nucleotides 14U - 15C and G82 - A85 (P_1-P_2/P_4-$J_{24(S)}$-P_1 motif). We quantified this misalignment using the dihedral angle Θ_J (see Methods). For both the slow and fast folding pathways, Θ_J initially fluctuates due to the unfolded P_1 and P_4 helices. After the formation of the helices, the magnitude of the fluctuations in Θ_J decreased. If the AD is stuck in the kinetic trap, Θ_J fluctuates around -0.1 rad (Figure 5B). The misaligned arms prevent the formation of pre-organized 3WJ, which facilitates the formation of the TC_{AT} tertiary structure. To escape from the kinetic trap, $J_{24(L)}$, P_4 and P_5 rotate around the P_{4-5} arm (Figure 6C_1,D,C′,D′). After escaping from the kinetic trap, a U-shaped loop is formed by the longer part of the J_{24} junction ($J_{24(L)}$) joining the P_2 and P_4 helices, which allows the helical arms to come closer to acquire native-like folded state competent to form the ligand-binding pocket (Figure 6C_2,E,E′). The loop-formation by the folded $J_{24(L)}$ junction pulls the helical arms together leading to the formation of TC_{AT} and *holo*-form like folded state. The inter-arm angle, Ω_{IA} and $R_{J_{24}}$ show large fluctuations when the AD is unfolded, and fluctuations subside only after the AD is completely folded.

Mg^{2+} Binding to Specific P Sites Facilitates Folding to the Native State

To probe the effect of Mg^{2+} on TPP AD folding, we computed c^* around the P sites of nucleotides involved in the formation of TCs (P_{TC} sites) (Figure S6 and S7) and f_{TC} as a function of time (t_{kin}) (Figure 5C,D). In the slow folding pathway, during the early stages of folding ($t_{kin} \lessapprox 0.39$ μs), c^* varies between 0 - 1 mM for all the P_{TC} sites and none of the TCs are formed. After the population of the kinetic trap due to the misalignment of the two arms (0.39 μs $< t_{kin} < 1.41$ μs), TC_{3WJ} fluctuates between states with $f_{TC} = 0.4$ to 0.5, as the AD attempts to escape from the kinetic trap (Figure 5D). The f_{TC} plot shows hierarchy in the formation of the TCs and the formation of TC_{3WJ} is a prerequisite for the formation of TC_{AT} (Figure 5C,D). Both the TCs form after native-like reorganization of the J_{24} motifs at t_{kin} > 1.41 μs. Similar time-dependent trends are also observed in c^* for fast-folding pathways, where the formation of the TCs is facilitated by Mg^{2+} accumulation around the P_{TC} sites. Hierarchy in TC formation where TC_{3WJ} anchors the formation of TC_{AT} is also observed in the fast folding pathways as well. Both thermodynamics and kinetics folding studies confirm that AD folding follows a quasi-hierarchical model, where TCs are not formed during the early stages of folding, and they form only during the AD assembly in the late stages of folding.

Conclusion

In this study, we showed that TPP AD populates intermediates when its folding is induced by Mg^{2+} in agreement with the experiments.[12] We found that at physiological [Mg^{2+}] ($\approx$ 4 mM), AD populated a closed-arm Y-shaped and a *holo*-form like folded conformations. The *holo*-form like conformation is not the global minimum in the FES at around physiological [Mg^{2+}] but becomes the global minimum in higher [Mg^{2+}] ($\geq$ 5 mM) even in the absence of the TPP ligand. These results indicate that *holo*-form like riboswitch conformations are only marginally stable at physiological conditions in the absence of their cognate ligands. Therefore at physiological conditions, TPP-ligand binding to the AD is indispensable for gene regulation and further supports the hypothesis that gene regulation by riboswitches is under kinetic control. The population of *holo*-form like conformations in the absence of cognate ligand but in high [Mg^{2+}] is also observed in preQ$_1$, Fluoride and SAM-I riboswitch.[28,44,45,100,101] Experiments on preQ$_1$ riboswitch have shown that Mg^{2+} can further shift the aptamer-ligand sensing mechanism from induced-fit to conformation selection.[45] The unique ability of Mg^{2+} to stabilize the *holo*-form like AD conformations in the absence of the cognate ligand poses an interesting unanswered question whether only Mg^{2+} can lead to the population of *holo*-form like structures in the AD of all riboswitches without their cognate ligands and what is the optimum Mg^{2+} concentration required?

Acknowledgement

A part of this work is funded by the grant to G.R. by the National Supercomputing Mission (MeitY/R&D/HPC/2(1)/2014). S.K. acknowledges research fellowship from the Indian Institute of Science, Bangalore. We acknowledge 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.

Supporting Information Available

Structural details of P_2-$J_{24(L)}$-P_4 and P_1-P_2/P_4-$J_{24(S)}$-P_1 motifs is provided in supporting information. Figure S1: Native contact map of AD; Figure S2: 3D structure of the AD with highlighted TC_{3WJ} and TC_{AT} forming nucleotides and 2D schematic structure of the AD with definition of parameters used to characterize the kinetic results; Figure S3: Atomistic representation of RNA residues involved in the TPP ligand binding; Figure S4: Time profile of f_{TC} for both TC_{3WJ} and TC_{AT} showing the formation of TCs at $[Mg^{2+}] = 4$ to 5 mM; Figure S5: $\langle c^* \rangle$ for $[Mg^{2+}] = 4$ to 5 mM; Figure S6-S7: Time profile of c^* at TC forming nucleotide resolution for fast and slow folding trajectories; Table S1-S6: Various parameters used to model the TPP AD using TIS model is given in the tabular format in the supporting information.

References

(1) Montange, R. K.; Batey, R. T. Riboswitches: Emerging themes in RNA structure and function. *Ann. Rev. Biophys.* **2008**, *37*, 117–133.

(2) Serganov, A.; Nudler, E. A Decade of Riboswitches. *Cell* **2013**, *152*, 17–24.

(3) Breaker, R. R. Riboswitches and Translation Control. *Cold Spring Harbor Perspect. Biol.* **2018**, *10*, a032797.

(4) Draper, D. A guide to ions and RNA structure. *RNA* **2004**, *10*, 335–343.

(5) Trachman, R. J., III; Draper, D. E. Divalent ion competition reveals reorganization of an RNA ion atmosphere upon folding. *Nucleic Acids Res.* **2017**, *45*, 4733–4742.

(6) Bukhman, Y.; Draper, D. Affinities and selectivities of divalent cation binding sites within an RNA tertiary structure. *J. Mol. Biol.* **1997**, *273*, 1020–1031.

(7) 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.

(8) Blount, K. F.; Breaker, R. R. Riboswitches as antibacterial drug targets. *Nat. Biotechnol.* **2006**, *24*, 1558–1564.

(9) Panchal, V.; Brenk, R. Riboswitches as drug targets for antibiotics. *Antibiotics* **2021**, *10*, 45.

(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*, 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) Lin, J.-C.; Thirumalai, D. Relative stability of helices determines the folding landscape of adenine riboswitch aptamers. *J. Am. Chem. Soc.* **2008**, *130*, 14080–14081.

(24) Gong, Z.; Zhao, Y.; Chen, C.; Xiao, Y. Computational study of unfolding and regulation mechanism of preQ$_1$ riboswitches. *PLoS One* **2012**, *7*, e45239.

(25) Padhi, S.; Pradhan, M.; Bung, N.; Roy, A.; Bulusu, G. TPP riboswitch aptamer: Role of Mg^{2+} ions, ligand unbinding, and allostery. *J. Mol. Graph.* **2019**, *88*, 282–291.

(26) Suresh, G.; Srinivasan, H.; Nanda, S.; Priyakumar, U. D. Ligand-induced stabilization of a duplex-like architecture is crucial for the switching mechanism of the SAM-III riboswitch. *Biochemistry* **2016**, *55*, 3349–3360.

(27) Kesherwani, M.; Kutumbarao, N. H., V; Velmurugan, D. Conformational Dynamics of thiM Riboswitch To Understand the Gene Regulation Mechanism Using Markov State Modeling and the Residual Fluctuation Network Approach. *J. Chem Inf. Model.* **2018**, *58*, 1638–1651.

(28) 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.

(29) Huang, W.; Kim, J.; Jha, S.; Aboul-ela, F. The impact of a ligand binding on strand migration in the SAM-I riboswitch. *PLoS Comput. Biol.* **2013**, *9*.

(30) 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.

(31) 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.

(32) Guedich, S.; Puffer-Enders, B.; Baltzinger, M.; Hoffmann, G.; Da Veiga, C.; Jossinet, F.; Thore, S.; Bec, G.; Ennifar, E.; Burnouf, D. et al. Quantitative and predictive model of kinetic regulation by *E. coli* TPP riboswitches. *RNA Biol.* **2016**, *13*, 373–390.

(33) Wickiser, J.; Cheah, M.; Breaker, R.; Crothers, D. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **2005**, *44*, 13404–13414.

(34) Quarta, G.; Sin, K.; Schlick, T. Dynamic energy landscapes of riboswitches help interpret conformational rearrangements and function. *PLoS Comput. Biol.* **2012**, *8*.

(35) Smith, K. D.; Lipchock, S. V.; Ames, T. D.; Wang, J.; Breaker, R. R.; Strobel, S. A. Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1218–1223.

(36) Tomsic, J.; McDaniel, B. A.; Grundy, F. J.; Henkin, T. A. Natural variability in S-adenosylmethionine (SAM)-dependent riboswitches: S-box elements in *Bacillus subtilis* exhibit differential sensitivity to SAM in vivo and in vitro. *J. Bacteriol.* **2008**, *190*, 823–833.

(37) Wickiser, J.; Winkler, W.; Breaker, R.; Crothers, D. The speed of RNA transcription

and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell* **2005**, *18*, 49–60.

(38) Lin, J.-C.; Thirumalai, D. Gene regulation by riboswitches with and without negative feedback loop. *Biophys. J.* **2012**, *103*, 2320–2330.

(39) Pavlova, N.; Kaloudas, D.; Penchovsky, R. Riboswitch distribution, structure, and function in bacteria. *Gene* **2019**, *708*, 38–48.

(40) Serganov, A.; Polonskaia, A.; Phan, A. T.; Breaker, R. R.; Patel, D. J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **2006**, *441*, 1167–1171.

(41) Ali, M.; Lipfert, J.; Seifert, S.; Herschlag, D.; Doniach, S. The ligand-free state of the TPP riboswitch: A partially folded RNA structure. *J. Mol. Biol.* **2010**, *396*, 153–165.

(42) Baird, N. J.; Ferré-D'Amaré, A. R. Idiosyncratically tuned switching behavior of riboswitch aptamer domains revealed by comparative small-angle X-ray scattering analysis. *RNA* **2010**, *16*, 598–609.

(43) Baird, N. J.; Ferré-D'Amaré, A. R. Correction to "Idiosyncratically tuned switching behavior of riboswitch aptamer domains revealed by comparative small-angle X-ray scattering analysis". *RNA* **2010**, *16*, 1447.

(44) 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.

(45) Suddala, K. C.; Wang, J.; Hou, Q.; Water, N. G. Mg^{2+} shifts ligand-mediated folding of a riboswitch from induced-fit to conformational selection. *J. Am. Chem. Soc.* **2015**, *137*, 14075–14083.

(46) Cunha, R. A.; Bussi, G. Unraveling Mg^{2+}-RNA binding with atomistic molecular dynamics. *RNA* **2017**, *23*, 628–638.

(47) Šponer, J.; Banáš, P.; Jurečka, P.; Zgarbová, M.; Kührová, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M. Molecular dynamics simulations of nucleic acids. From tetranucleotides to the ribosome. *J. Phys. Chem. Lett.* **2014**, *5*, 1771–1782.

(48) Miner, J. C.; Chen, A. A.; Garcia, A. E. Free-energy landscape of a hyperstable RNA tetraloop. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 6665–6670.

(49) Bao, L.; Wang, J.; Xiao, Y. Dynamics of metal ions around an RNA molecule. *Phys. Rev. E* **2019**, *99*.

(50) 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. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E1346–E1355.

(51) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jureacka, P. et al. RNA structural dynamics as captured by molecular simulations: A comprehensive overview. *Chem. Rev.* **2018**, *118*, 4177–4338.

(52) Hyeon, C.; Thirumalai, D. Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6789–6794.

(53) Sulc, P.; Romano, F.; Ouldridge, T. E.; Doye, J. P. K.; Louis, A. A. A nucleotide-level coarse-grained model of RNA. *J. Chem. Phys.* **2014**, *140*, 235102.

(54) Cragnolini, T.; Laurin, Y.; Derreumaux, P.; Pasquali, S. Coarse-grained HiRE-RNA model for ab initio RNA folding beyond simple molecules, including noncanonical and multiple base pairings. *J. Chem. Theory Comput.* **2015**, *11*, 3510–3522.

(55) Bell, D. R.; Cheng, S. Y.; Salazar, H.; Ren, P. Capturing RNA folding free energy with coarse-grained molecular dynamics simulations. *Sci. Rep.* **2017**, *7*, 45812.

(56) Denesyuk, N. A.; Thirumalai, D. How do metal ions direct ribozyme folding? *Nat. Chem.* **2015**, *7*, 793–801.

(57) Denesyuk, N. A.; Hori, N.; Thirumalai, D. Molecular simulations of ion effects on the thermodynamics of RNA folding. *J. Phys. Chem. B* **2018**, *122*, 11860–11867.

(58) 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.

(59) Poblete, S.; Bottaro, S.; Bussi, G. A nucleobase-centered coarse-grained representation for structure prediction of RNA motifs. *Nucleic Acids Res.* **2018**, *46*, 1674–1683.

(60) Zhang, D.; Li, J.; Chen, S.-J. IsRNA1: De novo prediction and blind screening of RNA 3D structures. *J. Chem. Theory Comput.* **2021**, *17*, 1842–1857.

(61) Hori, N.; Takada, S. Coarse-grained structure-based model for RNA-protein complexes developed by fluctuation matching. *J. Chem. Theory Comput.* **2012**, *8*.

(62) Li, P. T. X.; Vieregg, J.; Tinoco, I., Jr. How RNA unfolds and refolds. *Annu. Rev. Biochem.* **2008**, *77*, 77–100.

(63) Heilman-Miller, S.; Thirumalai, D.; Woodson, S. Role of counterion condensation in folding of the *Tetrahymena* ribozyme. I. Equilibrium stabilization by cations. *J. Mol. Biol.* **2001**, *306*, 1157–1166.

(64) Holmstrom, E. D.; Fiore, J. L.; Nesbitt, D. J. Thermodynamic origins of monovalent facilitated RNA folding. *Biochemistry* **2012**, *51*, 3732–3743.

(65) Cole, P.; Yang, S.; Crothers, D. Conformational changes of transfer ribonucleic acid. Equilibrium phase diagrams. *Biochemistry* **1972**, *11*, 4358–4368.

(66) Römer, R.; Hach, R. tRNA conformation and magnesium binding: A study of yeast phenylalanine-specific tRNA by a fluorescent indicator and differential melting curves. *Eur. J. Biochem.* **1975**, *55*, 271–284.

(67) Stein, A.; Crothers, D. Conformational changes of transfer RNA. The role of magnesium (II). *Biochemistry* **1976**, *15*, 160–168.

(68) 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.

(69) Hori, N.; Denesyuk, N. A.; Thirumalai, D. Ion condensation onto ribozyme is site specific and fold dependent. *Biophys. J.* **2019**, *116*, 2400–2410.

(70) 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.

(71) Denesyuk, N. A.; Thirumalai, D. Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B* **2013**, *117*, 4901–4911.

(72) Vriend, G. WHAT IF - a molecular modeling and drug design program. *J. Mol. Graph.* **1990**, *8*, 52–56.

(73) 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.

(74) Hori, N. TIS2AA. 2017; {https://doi.org/10.5281/zenodo.581485}.

(75) Humphris-Narayanan, E.; Pyle, A. M. Discrete RNA libraries from pseudo-torsional space. *J. Mol. Biol.* **2012**, *421*, 6–26.

(76) 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.

(77) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(78) Appignanesi, G.; Fernandez, A. Cooperativity along kinetic pathways in RNA folding. *J. Phys. A-Math. Gen.* **1996**, *29*, 6265–6280.

(79) Ralston, C.; He, Q.; Brenowitz, M.; Chance, M. Stability and cooperativity of individual tertiary contacts in RNA revealed through chemical denaturation. *Nat. Struct. Biol.* **2000**, *7*, 371–374.

(80) Behrouzi, R.; Roh, J. H.; Kilburn, D.; Briber, R. M.; Woodson, S. A. Cooperative tertiary interaction network guides RNA folding. *Cell* **2012**, *149*, 348–357.

(81) Celander, D.; Cech, T. Visualizing the higher order folding of a catalytic RNA molecule. *Science* **1991**, *251*, 401–407.

(82) Fang, X.; Pan, T.; Sosnick, T. A thermodynamic framework and cooperativity in the tertiary folding of a Mg^{2+}-dependent ribozyme. *Biochemistry* **1999**, *38*, 16840–16846.

(83) Dahm, S.; Uhlenbeck, O. Role of divalent metal ions in the hammerhead RNA cleavage reaction. *Biochemistry* **1991**, *30*, 9464–9469.

(84) Misra, V.; Draper, D. On the role of magnesium ions in RNA stability. *Biopolymers* **1998**, *48*, 113–135.

(85) Laing, L.; Gluick, T.; Draper, D. Stabilization of RNA structure by Mg ions: specific and non-specific effects. *J. Mol. Biol.* **1994**, *237*, 577–587.

(86) Pan, J.; Thirumalai, D.; Woodson, S. Magnesium-dependent folding of self-splicing RNA: Exploring the link between cooperativity, thermodynamics, and kinetics. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 6149–6154.

(87) Brannvall, M.; Kirsebom, L. Metal ion cooperativity in ribozyme cleavage of RNA. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12943–12947.

(88) Moghaddam, S.; Caliskan, G.; Chauhan, S.; Hyeon, C.; Briber, R. M.; Thirumalai, D.; Woodson, S. A. Metal ion dependence of cooperative collapse transitions in RNA. *J. Mol. Biol.* **2009**, *393*, 753–764.

(89) Woodson, S. A. Compact intermediates in RNA folding. *Annu. Rev. Biophys.* **2010**, *39*, 61–77.

(90) Goldberg, A. Magnesium binding by *Escherichia coli* ribosomes. *J. Mol. Biol.* **1966**, *15*, 663–673.

(91) Hurwitz, C.; Rosano, C. L. The intracellular concentration of bound and unbound magnesium ions in *Escherichia coli*. *J. Biol. Chem.* **1967**, *242*, 3719–3722.

(92) Alatossava, T.; Jütte, H.; Kuhn, A.; Kellenberger, E. Manipulation of intracellular magnesium content in polymyxin B nonapeptide-sensitized *Escherichia coli* by ionophore A23187. *J. Bacteriol.* **1985**, *162*, 413–419.

(93) Fang, X.; Littrell, K.; Yang, X.; Henderson, S.; Siefert, S.; Thiyagarajan, P.; Pan, T.; Sosnick, T. Mg^{2+}-dependent compaction and folding of yeast $tRNA^{Phe}$ and the catalytic domain of the *B. subtilis* RNase P RNA determined by small-angle X-ray scattering. *Biochemistry* **2000**, *39*, 11107–11113.

(94) Lambert, M.; Vocker, E.; Blumberg, S.; Redemann, S.; Gajraj, A.; Meiners, J.; Water, N. G. Mg^{2+}-induced compaction of single RNA molecules monitored by tethered particle microscopy. *Biophys. J.* **2006**, *90*, 3672–3685.

(95) Anthony, P. C.; Perez, C. F.; García-García, C.; Block, S. M. Folding energy landscape of the thiamine pyrophosphate riboswitch aptamer. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1485–1489.

(96) Lang, K.; Rieder, R.; Micura, R. Ligand-induced folding of the *thiM* TPP riboswitch

investigated by a structure-based fluorescence spectroscopic approach. *Nucleic Acids Res.* **2007**, *35*, 5370–5378.

(97) Rangan, P.; Woodson, S. Structural requirement for Mg^{2+} binding in the group I intron core. *J. Mol. Biol.* **2003**, *329*, 229–238.

(98) Misra, V.; Shiman, R.; Draper, D. A thermodynamic framework for the magnesium-dependent folding of RNA. *Biopolymers* **2003**, *69*, 118–136.

(99) Treiber, D.; Williamson, J. Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **1999**, *9*, 339–345.

(100) Haller, A.; Soulière, M. F.; Micura, R. The dynamic nature of RNA as key to understanding riboswitch mechanisms. *Acc. Chem. Res.* **2011**, *44*, 1339–1348.

(101) Rieder, U.; Kreutz, C.; Micura, R. Folding of a transcriptionally acting $PreQ_1$ riboswitch. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 10804–10809.

For Table of Contents Use Only