Manuscript submitted to **Biophysical** *Journal*
**Article**

# Ion Condensation onto Ribozyme is Site-Specific and Fold-Dependent

Naoto Hori[1], Natalia A. Denesyuk[2], and D. Thirumalai[1,*]

[1]Department of Chemistry, University of Texas, Austin, Texas 78712, USA
[2]Biophysics Program, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA
[*]Correspondence: dave.thirumalai@gmail.com

ABSTRACT The highly charged RNA molecules, with each phosphate carrying a single negative charge, cannot fold into well-defined architectures with tertiary interactions, in the absence of ions. For ribozymes, divalent cations are known to be more efficient than monovalent ions in driving them to a compact state although $Mg^{2+}$ ions are needed for catalytic activities. Therefore, how ions interact with RNA is relevant in understanding RNA folding. It is often thought that most of the ions are territorially and non-specifically bound to the RNA, as predicted by the counterion condensation (CIC) theory. Here, we show using simulations of *Azoarcus* ribozyme, based on an accurate coarse-grained Three Site Interaction (TIS) model, with explicit divalent and monovalent cations, that ion condensation is highly specific and depends on the nucleotide position. The regions with high coordination between the phosphate groups and the divalent cations are discernible even at very low $Mg^{2+}$ concentrations when the ribozyme does not form tertiary interactions. Surprisingly, these regions also contain the secondary structural elements that nucleate subsequently in the self-assembly of RNA, implying that ion condensation is determined by the architecture of the folded state. These results are in sharp contrast to interactions of ions (monovalent and divalent) with rigid charged rods in which ion condensation is uniform and position independent. The differences are explained in terms of the dramatic non-monotonic shape fluctuations in the ribozyme as it folds with increasing $Mg^{2+}$ or $Ca^{2+}$ concentration.

## INTRODUCTION

RNA molecules are negatively charged polyelectrolytes (PEs) with each molecule carrying a total charge equal to $-Ne$ where $N$ is the number of nucleotides and $e$ is the elementary charge. The negative charges are localized on the phosphate groups. Because of the large electrostatic repulsion between the phosphate groups, RNA can only fold if the effective charge on the phosphate group is reduced or renormalized (1–5), which we know occurs by the counterion condensation (CIC) mechanism (6–9). In this process, described decades ago in remarkable studies by Oosawa and Manning (6, 7), the condensed ions are localized in an apparent volume occupied by the PE or any highly charged macroion. As a consequence, the effective charge on the negatively charged moieties in the PE (or the phosphate groups in DNA and RNA) is reduced from $-e$ to $-\beta e$ with $\beta < 1$ (6–8). The renormalized charge parameter, $\beta$, can be calculated for PEs with regular shapes (8). For example, the condition for ion condensation for a rod-like macroion is determined by the Manning parameter $\xi = \ell_B/b > 1/|z|$, where $\ell_B$ is the Bjerrum length, $b$ is the mean distance between charges on the polyelectrolyte, and $z$ is the valence of the counterion. The Bjerrum length is the distance at which the electrostatic energy of two univalent charges is equal to the thermal energy, $\ell_B = e^2\,(4\pi\varepsilon_0\varepsilon k_B T)^{-1}$, where $\varepsilon_0$ is the vacuum permittivity, $\varepsilon$ is the dielectric constant of the solvent, and $k_B T$ is the thermal energy. For highly charged rods, ion condensation occurs when $\xi \geq 1/|z|$, with the transition being sharp for long thin rods (high aspect ratio). A similar condition may be obtained for spherical macroions. In this case, ion condensation does occur with a fraction localized in the vicinity of the sphere, but unlike the situation of the cylindrical macroion, there is no critical value of $\xi$ at which counterion condensation is initiated (8).

There are two implications of these well-known aspects of the CIC theory, summarized above, that are relevant to the current study. (1) Monovalent ions ($|z| = 1$) are not as efficient in charge renormalization as higher-valent ions ($|z| > 1$). Based on this consideration, we calculated that about 90% of the phosphate charge in *Tetrahymena* ribozyme is neutralized in the presence of divalent and trivalent cations (10). (2) The condensation mechanism depends critically on the shape of the macroion. As noted above, the theoretical predictions are different for spheres and rods. Therefore, we expect that the CIC mechanism ought to be different for RNA, which undergoes substantial conformational fluctuations during the ion-driven folding process. The shape of *Azoarcus* ribozyme that is used as an example here, changes as the concentration of ions ($Mg^{2+}$) is increased from a low to a high value. This ribozyme, and many other RNA molecules, are not spherical even when folded (11). Moreover, even

at low $Mg^{2+}$ concentrations, *Azoarcus* ribozyme does not adopt globally cylindrical structures although individual helices may be rigid enough to be pictured as stiff small cylinders.

Because of the irregular shapes of RNA molecules at all ion concentrations, ion-RNA interactions, which are often thought to involve diffuse ions that are territorially (within a volume of RNA where the electrostatic potential is substantial) but non-specifically bound, require scrutiny. Motivated by the considerations described above, we answer the following question here: How does CIC occur in such molecules whose shapes are not only irregular but also change as they fold? We are able to quantitatively answer this question because of the development of an accurate coarse-grained force field, which has been used to quantitatively predict the thermodynamic properties of *Azoarcus* ribozyme and other RNA molecules (12) in the presence of $K^+$, $Mg^{2+}$ and $Ca^{2+}$.

There is an additional reason for undertaking this investigation. The physically appealing CIC mechanism qualitatively explains RNA folding both from the stability and kinetic perspective (10). The mechanism is based on the basic assumption that condensed ions are diffuse and freely move within the volume surrounding the RNA where the electrostatic potential is the largest. Indeed, the condition for ion condensation is obtained by assuming that the mobile but bound ions are in equilibrium with free ions (the chemical potential of these two species are identical) that explore the region outside the sphere of influence of the macroion (8). Such an assumption does qualitatively explain the success of even the $z$-dependent folding kinetics of *Tetrahymena* ribozyme (13, 14). However, our recent studies (12) suggest that the assumption of uniform condensation, often made in rationalizing electrostatic effects in RNA, may be qualitatively incorrect. In other words, nucleotide-specific interactions between ions (especially $Ca^{2+}$ and $Mg^{2+}$ ions) provide the key in quantitatively describing charge renormalization in ribozyme folding.

In order to fully understand the nature and consequence of ion condensation onto RNA, we undertook two types of simulations. We first probed the interaction of divalent cations with small finite-sized rigid rods of different linear charge density, expressed as $b = n/L$ where $n$ is the total number of charges and $L$ is the cylinder length. It follows that $b$ is the mean distance between charges. However, in our simulations the distance between the charges ($b$) is a constant. We also investigated the condensation of divalent cations onto *Azoarcus* ribozyme using an accurate model introduced recently (12). By comparing the results of these two systems, we elaborate on the importance of the architecture of the macroion (rod versus ribozyme) on the ion condensation mechanism. Our major finding is that the shape of the RNA matters for counterion condensation. In other words, ion-condensation to RNA depends exquisitely on the folded RNA architecture, and hence cannot be treated as a non-specific interaction that is used to understand CIC for macroion with definite simple shapes. A surprise in our study is that the specificity of ion binding plays a vital role in driving RNA folding even at very low divalent ion concentrations.

## MODEL AND METHODS

### Rigid rod simulations

A 5 nm-long rigid cylindrical rod is placed at the center of 15 nm-wide cubic box (Fig. 1a and b). The diameter of the rod is 0.42 nm, which is the size of the phosphate bead in the coarse-grained RNA model described below. Technically, the rod is constructed by overlapping spheres that are separated by 0.1 nm. The rod is completely rigid and fixed at the center of the simulation box, whereas ions are allowed to move freely around the rod. We use periodic boundary conditions. Three different charge densities on the rod, $\rho_{1D} = -1$, $-2$, and $-5\ e\,\mathrm{nm}^{-1}$ were simulated, which translates to the constant separation distance ($b$) between neighboring charges as 1.0, 0.5, and 0.2 nm, respectively. For comparison, the Bjerrum length ($\ell_B$) is 0.73 nm at the simulation temperature, $T = 37°C$. The interactions between the charges on the rod and the ions are the same as in the coarse-grained RNA-ion model (12). Ions ($K^+$, $Cl^-$, and $Mg^{2+}$) are modeled as spheres and interact with each other and with the charged rod via the excluded volume potential, which is identical to the one used in the RNA model,

$$U_{\mathrm{Ex}} = \varepsilon_{ij}\left[\left(\frac{1.6}{r_{ij}+1.6-D_{ij}}\right)^{12} - 2\left(\frac{1.6}{r_{ij}+1.6-D_{ij}}\right)^{6} + 1\right],\ r_{ij} \le D_{ij} \tag{1}$$

where $r$ is the distance of the two beads, $D_{ij} = R_i + R_j$, and $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$. The values of $R_i$ and $\varepsilon_i$ for the ions and the RNA sites are tabulated in Table S1 in the Supplementary Information (SI). The $R_i$ and $\varepsilon$ for the charged rod are the same as the phosphate site in the RNA model. If $r_{ij} > D_{ij}$ then $U_{\mathrm{EX}} = 0$. For any pair of charged beads, the standard Coulomb potential is used,

$$U_{\mathrm{Ele}} = \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}} \tag{2}$$

where $q_i$ is the charge of the $i^{\mathrm{th}}$ bead. The dielectric constant, $\varepsilon$, is temperature dependent and is 74 at the simulation temperature $T = 37°C$. A series of Langevin dynamics simulations were performed for $2 \times 10^7$ iteration time steps, in which ion positions are saved every 100 steps. Thus, in total $2 \times 10^6$ snapshots are used for analyses for each setup.

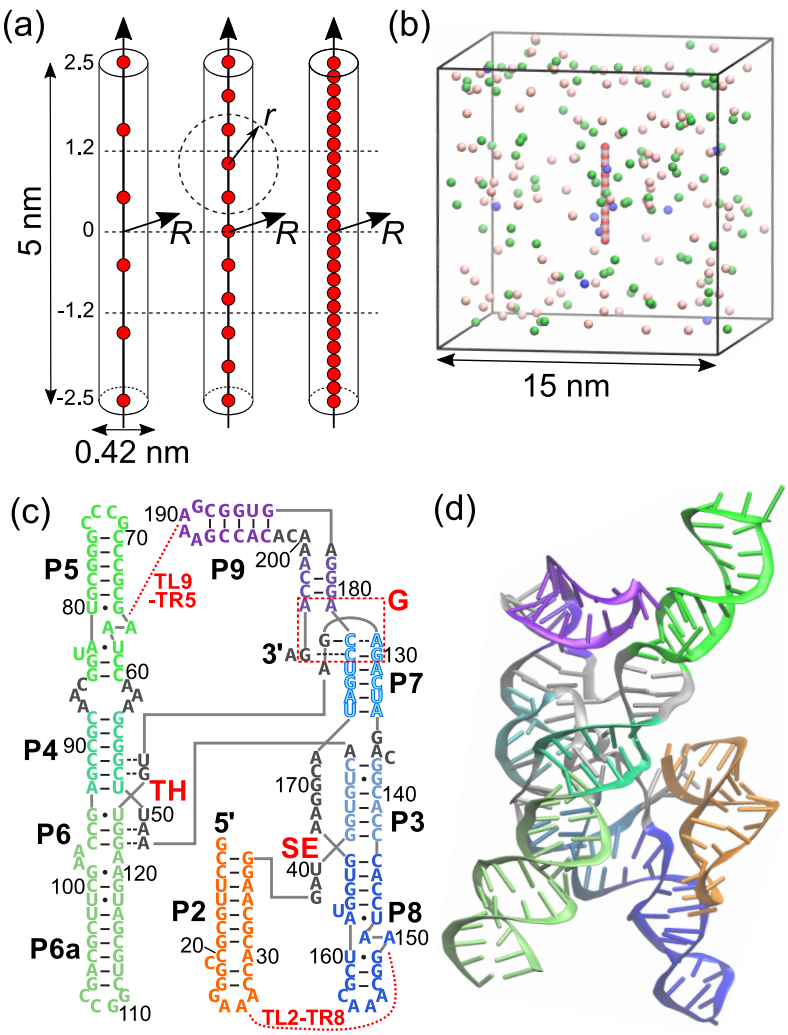


Figure 1: Schematic of charged rods and the *Azoarcus* ribozyme. **(a)** The dimensions of the rod, with aspect ratio 11.9, and positions of charges (in red) are illustrated. The diameter of the rod is 0.42 nm, which is the same as the diameter of the phosphate group in the coarse-grained RNA model. Negative charges are placed at regular interval with spacing that gives charge densities; −1.0 *e*/nm (left), −2.0 *e*/nm (middle), and −5.0 *e*/nm (right). *R* is the radial distance in the cylindrical coordinate, whereas *r* represents the distance from the point charge to a position in space in the spherical coordinate. **(b)** A snapshot of the rod simulations. Ions are colored in blue for $Mg^{2+}$, green for $K^{+}$, and pink for $Cl^{-}$. **(c, d)** The secondary (c) and tertiary (d) structures of *Azoarcus* group I intron ribozyme (PDB 1U6B (15)). The color codes in both structures are the same. The locations of the key tertiary interactions are indicated in red: TH, Triple Helix; SE, Stack-Exchange junction; G, G-binding pocket; TL-TR, tetraloop-tetraloop receptor.

We ought to emphasize that the results of rigid rod simulations for different purposes have been reported numerous times in the literature (see for example (16–18)) although new phenomena continue to be found in (19). The simulations presented here merely serve as a control in order to illustrate the dramatically different findings for the *Azoarcus* ribozyme.

## Simulations of group I intron RNA

Because the model and simulation methods of the ribozyme were described in detail previously (12), we give only a brief description here. We used a coarse-grained Three Interaction Site (TIS) RNA model that has three interaction sites, for each nucleotide, corresponding to phosphate, sugar and base moieties (20). All the ions are explicitly treated whereas water is modeled implicitly using the temperature-dependent dielectric constant. We considered concentration of $Mg^{2+}$ ions in the range

from 0 to 30 mM in the presence of either 50 mM or 12 mM of $K^+$ that is typically used in the Tris buffer (21). We added an appropriate number of anions ($Cl^-$) to ensure that the entire system is charge neutral. All the simulations were conducted using a periodic-boundary cubic box with each side being 35 nm. We performed a series of Langevin dynamics simulations at 37°C in order to efficiently sample conformations of the system containing the ribozyme and ions (22). For each salt concentration, several independent simulations were started from the crystal structure and trajectories were generated for sufficiently long time to ensure that equilibrium is reached before we obtained snapshots for analyses. For low $Mg^{2+}$/$Ca^{2+}$ concentrations, at which the ribozyme is unfolded state, we generated seven trajectories with 7.5 $\mu$s of equilibration followed by 15–22.5 $\mu$s of production runs. The length of these simulations are long enough to observe multiple unfolding/folding events in the low-friction Langevin dynamics. For higher $Mg^{2+}$/$Ca^{2+}$ concentrations, at which the ribozyme is predominantly in the folded conformations, we ran three trajectories with length 7.5 $\mu$s and discarded the first $10^5$ steps of each trajectory. Elsewhere (12) we provide more detailed descriptions of the model and the simulation protocol.

The crystal structure of *Azoarcus* group I intron (PDB 1U6B (15), Fig. 1c and d) was used as the reference structure of the native conformation. The numbering of nucleotides from G12 through G206 is consistent with the crystal structure (15) and the first nucleotide of the exon sequence is labeled 207.

## Data analyses

To quantify ion condensation onto the RNA and the rigid rod, we calculated the *local ion concentration*, a fingerprint of ion interaction with the macroion, around the $i^{\text{th}}$ phosphate (or point charges on the rod) defined as,

$$c_i^* = \frac{1}{N_A V_c} \int_0^{r_c} \rho_i(r)\, 4\pi r^2 \mathrm{d}r, \tag{3}$$

where $\rho_i(r)$ is the number of ions per unit volume (number density) of the ion at the distance $r$ from the $i^{\text{th}}$ phosphate, $V_c$ is the spherical volume of radius $r_c$, $N_A$ is the Avogadro's number to represent $c_i^*$ in molar units. We use the Bjerrum length as the cutoff distance, $r_c = \ell_B = 0.73$ nm because at that distance the favorable Coulomb attraction between a cation and an anion ($z = 1$) is $k_B T$.

We also calculated the fluctuations in the local ion concentration as the normalized root-mean-square deviation from the mean,

$$\Delta c_i^* = \frac{\left\langle \left(c_i^* - \langle c_i^* \rangle\right)^2 \right\rangle^{1/2}}{\langle c_i^* \rangle}, \tag{4}$$

where $\langle c_i^* \rangle$ is the ensemble average of the local ion concentration around the $i^{\text{th}}$ nucleotide. Thus, $c_i^* - \langle c_i^* \rangle$ is an instantaneous deviation from the ensemble average. The average ion concentration is,

$$\overline{\langle c^* \rangle} = \frac{1}{N} \sum_{i=1}^{N} \langle c_i^* \rangle, \tag{5}$$

where $N$ is the number of nucleotides. Note that the upper bar indicates an average over the sequence, whereas the bracket indicates the ensemble average. The extent of deviation of ion concentration fluctuation from $\langle c^* \rangle$ along the sequence was calculated using,

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \left( \langle c_i^* \rangle - \overline{\langle c^* \rangle} \right)^2. \tag{6}$$

In the results section, for simplicity, we omit the brackets for the ensemble average. Thus, all quantities are averaged over an ensemble of configurations in the system (ions and the charged RNA or the rod) unless otherwise stated.

For charged rods, we also computed the total charge per unit length due to contributions from each ionic species. The effective charge densities around the rod were calculated as the amount of accumulated charge per unit length for the ion species $m$,

$$Q_m(R) = z_m \int_0^R \rho_m(R')2\pi R' \,\mathrm{d}R', \tag{7}$$

where $\rho_m(R)$ is the number density of species $m$ at the radial distance $R$ from the axis of the cylinder, and $z_m$ is its valence. Note that $Q$ can be computed for each ion $Mg^{2+}$, $K^+$, and $Cl^-$ independently. Finally, the total effective charge per unit length is the sum of the rod charge and contributions from those ions,

$$Q_T(R) = \rho_{1D} + Q_{Mg^{2+}}(R) + Q_{K^+}(R) + Q_{Cl^-}(R). \tag{8}$$

Note that we use $R$ for the radial distance from the center of the rods in the cylindrical coordinate, whereas $r$ is for the distance from the point charges on the rod in the spherical coordinate (see Fig. 1a).

## RESULTS

### Charged cylindrical rods

We first describe condensations of ions onto charged rods with varying spacing between fixed charges, $b = -e/\rho_{1D}$, where $\rho_{1D}$ is the charge density of the rod. Based on the CIC theory, we expect that ions should condense uniformly along the rod. The condition for condensation in the CIC theory is that $\xi > {}^{1}/_{|z|}$, where $z$ is the valence of the ion species. In the absence of multivalent salt, $\xi > 1$ and thus we expect ions should condense on rods when $\rho_{1D} = -5\,e/\text{nm}$ ($\xi = 3.6$) and $\rho_{1D} = -2\,e/\text{nm}$ ($\xi = 1.5$) among the three values of charge densities we examine here. This expectation was borne out in the simulations even though the aspect ratio ($\approx 11.9$) of the rod was relatively small. In Fig. 2(a-c), the cumulative charges, $Q$ (Eq. 7), around the rod are plotted for each ion species and the net charge. The condensation of monovalent ions appears as a sharp increase in $Q_{K^+}$ around the vicinity of the rod ($R \sim 5$ nm), with a subsequent reduction of the net negative charge at longer distances (thick lines in gray). The amount of condensation is larger for the rod with $\rho_{1D} = -5\,e/\text{nm}$ than for one with $\rho_{1D} = -2\,e/\text{nm}$. In the case of the lowest charge density ($-1\,e/\text{nm}$ corresponding $\xi = 0.7$ ), condensation is not observed as predicted by the CIC theory (Fig. 2c).

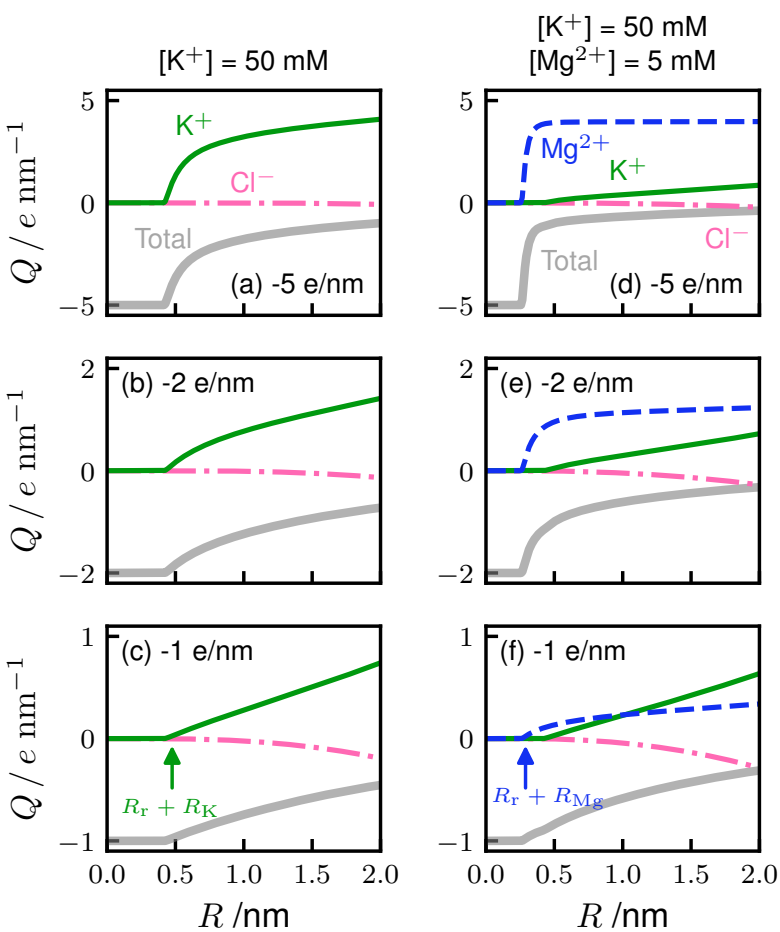


Figure 2: Comparison of the condensed ion species on charged rods with linear densities, $\rho_{1D} = -5\,e/\text{nm}$ (top panels), $-2\,e/\text{nm}$ (middle) and $-1\,e/\text{nm}$ (bottom) in the absence (a-c) and presence (d-f) of $Mg^{2+}$. The cumulative charges per unit length $Q$ (Eq. 7) is plotted as a function of $R$ for contribution from each ionic species (annotated in the top panels). The total cumulative charge, $Q_T(R)$ (thick gray line), is the sum of charges of the rod and surrounding ions (Eq. 8). The solution conditions are indicated on the top. The ranges of exclusion by the volume interactions (Eq. 1), $R_r + R_K = 0.48$ nm for $K^+$, and $R_r + R_{Mg} = 0.29$ nm for $Mg^{2+}$, are indicated by arrows in the bottom panels ($R_r$, $R_K$, $R_{Mg}$ are radii of the rod, $K^+$, and $Mg^{2+}$, respectively).

The condition for condensation, $\xi > 1/|z|$, also implies that the charged rod must be surrounded to a greater degree by $Mg^{2+}$ than $K^+$. This prediction was also borne out in the simulations. In Fig. 2 (d and e), $Mg^{2+}$ ions condense in the immediate vicinity of the rods taking over the role of $K^+$ in the charge renormalization. The replacement of $K^+$ by $Mg^{2+}$ in the vicinity of the rod is expected based on the counterion release mechanism (leads to an entropy of the whole system increase), which is another well-known consequence of the CIC theory. The condensation of cations, and the subsequent net reduction of negative charge on the rod, are more efficient in the presence of $Mg^{2+}$ than when $z = 1$ (compare panels (a) and (d) in Fig. 2). Since the condensation condition for $Mg^{2+}$ is $\xi > 0.5$, a distinct increase of $Q_{Mg^{2+}}$ is visible even at the lowest density case, $\rho_{1D} = -1$ $e$/nm ($\xi = 0.7$ ) albeit the extent of condensation is not as great as found in the two rod systems with higher density (Fig. 2f).

The results quoted above using the control simulations are entirely in accord with the CIC theory predictions. The extremely close correspondence between the CIC theory and simulations for charged rods motivated us to make several assessments of ion condensation phenomena in ribozyme folding, in which we use the same explicit-ion model. In order to make quantitative comparisons between the rod and RNA folding simulations, we need to devise a quantity to precisely measure the ion condensation. The radial distributions in cylindrical coordinates are not suitable for flexible RNA chains, for which the persistence length (1 ~ 2 nm for a single strand) and the Debye screening length (1.3 nm in the simulation conditions) are similar. With this in mind, we defined local ion concentrations ($c^*$) and the associated fluctuations (Eqs. 3-4) in the spherical coordinate for each point charge. The results of the rod simulations in terms of $c^*$ are shown in Fig. 3 (a and b) for the three rod charge densities in the presence of 5 mM $MgCl_2$ and 50 mM KCl. It is clear that the density of condensed ions around the charges in the rod is uniform; there is no position dependence along the rod. From panels (c) and (d) in Fig. 3, the fluctuation in $c^*$ is the lowest in the case of $Mg^{2+}$ condensed for the highest charge density rod. Both lowering the linear charge density and reducing the valence of cations ($Mg^{2+} \rightarrow K^+$) make the condensed state less favorable, thus increasing the magnitude of charge fluctuations. These quantitative descriptions for the charged rods set the stage for analyzing the completely contrasting behavior in ion-ribozyme interactions.

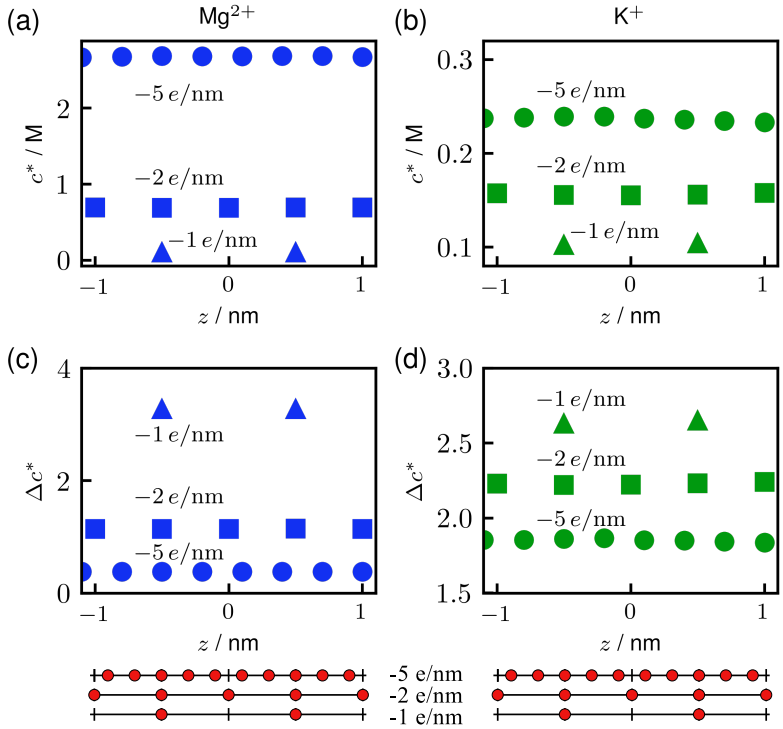


Figure 3: Condensation of $Mg^{2+}$ (a, c) and $K^+$ (b, d) for the three rods with differing charge density, $\rho_{1D}$, (Fig. 1a) in 5 mM $Mg^{2+}$ and 50 mM $K^+$ buffer. **(a, b)** Local ion concentrations of (a) $Mg^{2+}$ and (b) $K^+$. **(c, d)** Normalized fluctuations (Eq. 4) of the local ion concentrations of (c) $Mg^{2+}$ and (d) $K^+$ at the same condition as in (a) and (b). The positions of negative charges on the three rods are indicated at the bottom. In order to avoid end effects, only the middle region of the rod ($-1.1$ nm $\leq z \leq 1.1$ nm) was used in the analysis.

## [$Mg^{2+}$]-dependent folding of *Azoarcus* ribozyme

Before proceeding to examine the roles of ions in RNA folding, we summarize briefly the [$Mg^{2+}$]-dependent folding of the *Azoarcus* ribozyme, which has been investigated in a previous study using the elaborate version of the TIS model (12). The folding process of the ribozyme (Fig. 1) strongly depends on the concentration of $Mg^{2+}$. In Fig. 4a, the dependence of the radius of gyration ($R_g$) on [$Mg^{2+}$] is shown in gray lines for two different conditions of monovalent salt concentration, [$K^+$] = 12 and 50 mM. The global compaction of the RNA appears as a single transition in $R_g$ occurring at $C_m \approx 1.5$ mM, where we designate $C_m$ as the midpoint concentration of [$Mg^{2+}$] separating the folded and unfolded states. Comparison of 12 mM $K^+$ simulation data and SAXS experiments at the equivalent condition (21) (squares in Fig. 4) shows excellent agreement, thus establishing that our model captures monovalent- and divalent-ion dependence of folding of *Azoarcus* ribozyme. In the absence of $Mg^{2+}$, the key tertiary interactions (marked in red in Fig. 1c) are disrupted and only some of the helical domains, P2, P4, P5, and P8 are stably formed. As [$Mg^{2+}$] increases, other helices and tertiary interactions form in multiple steps. At [$Mg^{2+}$] $\sim$ 1 mM, tertiary interactions, TH, G site, and SE form, followed by folding of other peripheral interactions, TL2-TR8 and TL9-TR5, at [$Mg^{2+}$] $>$ 2 mM. The formations of these tertiary interactions lead to a compact intermediate state with $\langle R_g \rangle \sim 3.5$ nm. Some interesting cooperativity and anti-cooperativity, such as cooperative formation of P3 helix and SE, were found and discussed in the previous study (12). The complete native structure, including the active site, is spontaneously formed at [$Mg^{2+}$] $>$ 4 mM. In subsequent parts of this paper, we report detailed analyses of ion distributions around the RNA. Comparisons with the interaction of divalent ions with highly charged rods illustrate the importance of ribozyme architecture on ion–RNA interactions.

## Radial profiles of ion condensation

We calculated the density profiles of each ion species around phosphate, $\rho_i(r)$, where $r$ is the distance from the phosphate of the $i^{\text{th}}$ nucleotide. We also computed the profiles of phosphate–phosphate distance, which can be used to trace the compaction of the ribozyme. The ensemble-averaged profiles, for given $Mg^{2+}$ concentrations, are shown in Fig. 5. From the profiles of the phosphate (gray dotted lines in Fig. 5), we can track the folding of the ribozyme as follows. The peak at $r \approx 0.55$ nm is roughly at the distance between two consecutive phosphates, and the smaller second peak corresponds to the location of the second nearest-neighbor phosphates. The amplitudes of these two peaks do not show significant change as [$Mg^{2+}$] increases. However, the folding process can be discerned by the increase in $\rho_P(r)$ in the range of $0.7 < r < 0.9$ nm (red arrows in Fig. 5). At low [$Mg^{2+}$], there is practically no density in $\rho$ in this range ($\rho(r) \approx 0$). As [$Mg^{2+}$] increases, we find that the amplitude increases, which shows that $Mg^{2+}$ ions facilitate the approach of non-neighboring phosphates. The increase in $\rho(r)$ in this range coincides with the compaction of the ribozyme and subsequent folding as shown in Fig. 4(a).

In the absence of $Mg^{2+}$, $K^+$ cations are condensed around the ribozyme, regardless of the monovalent salt concentration, [KCl] = 12 mM or 50 mM (Figs. 5a and S1a). Under these conditions with [$Mg^{2+}$] = 0, all tertiary interactions are disrupted but helices, except P3 and P7, form secondary structures (Fig. 1c). Thus, $K^+$ cations must be condensed around the vicinity of phosphates to balance the electrostatic repulsion between negative charges on the phosphates. The distance to condensed $K^+$, $r \approx 0.4$ nm, is closer than the neighboring phosphate. In the presence of $Mg^{2+}$, $\rho(r)$ for $K^+$ decreases dramatically as the $Mg^{2+}$ concentration increases (Fig. 5 b-d). This shows, rather vividly, that $K^+$ ions condensed around RNA are replaced with $Mg^{2+}$ ions because divalent cations screen negative charges on the RNA more efficiently than monovalent cations. By replacing two $K^+$ ions by one $Mg^{2+}$, there is a net gain in translational entropy of the ions. This is the counterion release mechanism predicted by the CIC theory.

## Fingerprint of local ion concentration $c_i^*$

In order to examine the relationship between ion condensation and $Mg^{2+}$-driven RNA folding, we calculated the local ion concentration of the $i^{\text{th}}$ nucleotide $c_i^*$ (Eq. 3). The cutoff distance to define the local volume is chosen to be the Bjerrum length, $r_c = 0.73$ nm. The range of this distance covers the peaks in the radial distributions of both $K^+$ and $Mg^{2+}$ (Fig. 5). Accordingly, $c_i^*$ is expected to capture the amount of ion atmosphere around each nucleotide, and such a fingerprint is dependent on the architecture of the folded state of the RNA. The overall ion condensation of the entire molecule, $\overline{c^*}$, is calculated as the average of the local ion concentration along the RNA sequence (Eq. 5, see Models and Methods).

The changes in $\overline{c^*}$ as a function of [$Mg^{2+}$] is shown in Fig. 4(b). As noted previously, more $K^+$ cations condense at lower $Mg^{2+}$ than at higher $Mg^{2+}$ concentrations. Indeed, $\overline{c^*_{K^+}}$ monotonically decreases as [$Mg^{2+}$] increases (green lines). The folding midpoint of $Mg^{2+}$ concentration is around 1.5 mM. Above the midpoint, $Mg^{2+}$ ions are more dominantly found in the vicinity of the RNA than $K^+$. $\overline{c^*_{Mg^{2+}}}$ does not vary significantly as a function of $K^+$ concentrations for the entire range of [$Mg^{2+}$] (Fig. 4b, solid line for [$K^+$] = 50 mM and dotted line for 12 mM). The sharp increase in $\overline{c^*_{Mg^{2+}}}$ is well correlated with the global compaction of the ribozyme, that is evident as the decrease in $R_g$ in Fig. 4(a).

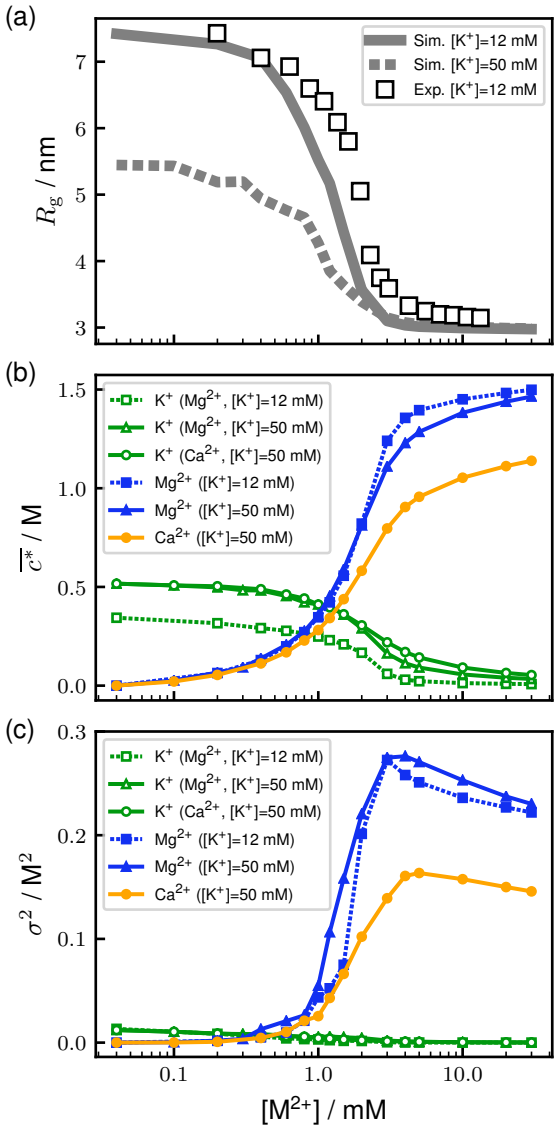


Figure 4: Ribozyme folding and ion condensation as a function of $Mg^{2+}$ concentration. **(a)** Average radius of gyration ($R_g$) as a function of the concentration of $Mg^{2+}$, that indicates the global compaction around the midpoint of $[Mg^{2+}] \sim 1.5$ mM. Squares are experimentally measured $R_g$ in a Tris buffer containing 12 mM monovalent salt (21). **(b)** $[Mg^{2+}]$-dependence of local ion concentrations averaged over the nucleotide sequence, $\overline{c^*}$, for $K^+$, $Mg^{2+}$, and $Ca^{2+}$. **(c)** Variances in the concentrations along the RNA sequence, $\sigma^2$.

## Ion condensation is site-specific

In Fig. 4c, we show variances, $\sigma^2$, of the local ion concentration over the RNA sequence (Eq. 6). This quantity indicates how much of each ion species are localized at particular positions along the sequence at each solution condition. As $Mg^{2+}$ concentration increases, $\sigma^2$ for $K^+$monotonically decreases, although it is much smaller compared to divalent cations for the entire range of $[Mg^{2+}]$. This means that $Mg^{2+}$ ions tend to be localized at specific positions more than $K^+$ ions. In particular, $\sigma^2$ for $Mg^{2+}$ sharply increases reaching a maximum around $[Mg^{2+}] \approx 3$ mM. This is because $Mg^{2+}$ ions are preferentially localized to specific places, which are, surprisingly, determined by the tertiary structure of the ribozyme. Beyond the midpoint $C_m$, the value of $\sigma^2$ decreases modestly because the specific sites of higher phosphate densities are occupied by existing localized $Mg^{2+}$. The newly added $Mg^{2+}$ ions start contributing to lessening the gap between specific and other non-specific sites.

Fig. 6 shows the distributions of the local ion concentration along the RNA sequence. This figure shows that all non-zero values of $Mg^{2+}$ condensation is highly specific. The extent of condensation depends on the nucleotide. Even at the lowest

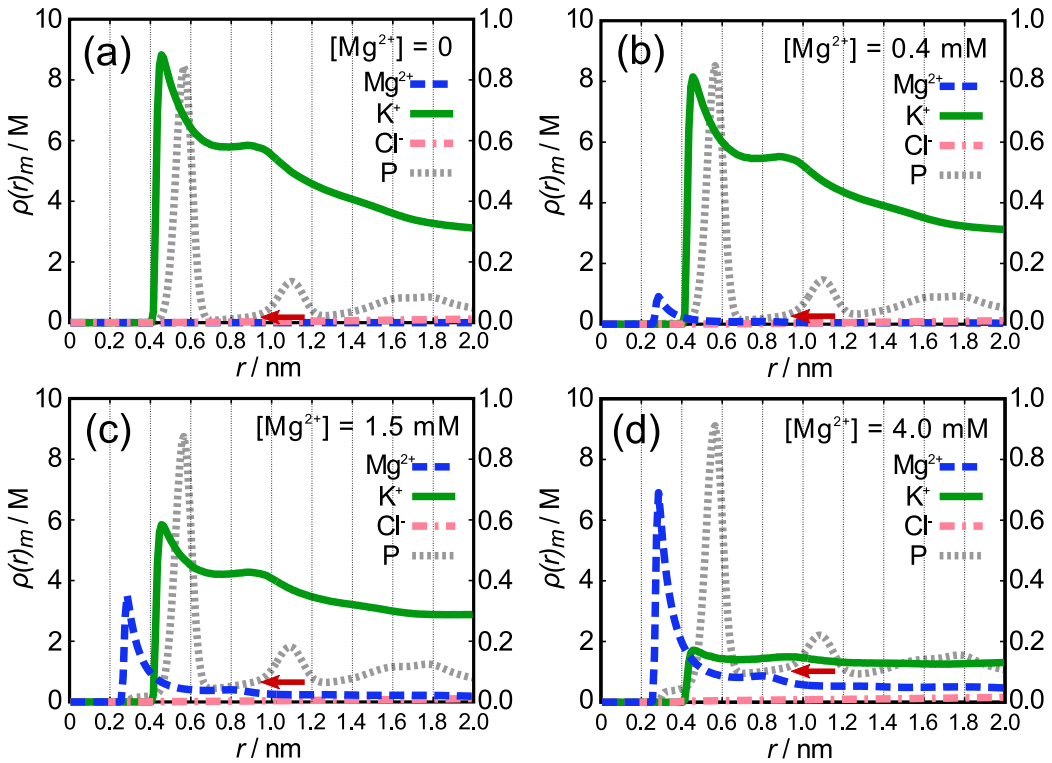


Figure 5: Radial profiles of ion distributions around the ribozyme, $\rho(r)$, for $Mg^{2+}$, $K^+$, $Cl^-$, and phosphate (P) at various $Mg^{2+}$ concentrations: (a) 0, (b) 0.4 mM, (c) 1.5 mM, and (d) 4.0 mM. The concentration of $K^+$ is 50 mM. An appropriate number of $Cl^-$ were added to the solution for the electroneutrality. Note that the scale for $\rho_{Mg^{2+}}$ and $\rho_P$ (left axis) is 10 times greater than $\rho_{K^+}$ and $\rho_{Cl^-}$ (right axis). It should be pointed out that the density profile for $K^+$ decreases substantially as $[Mg^{2+}]$ increases. The red arrows indicate the amplitude of non-neighboring phosphates. The results in 12 mM KCl are shown in Fig. S1 in the SI. The density profiles for $Mg^{2+}$ and P are roughly independent of the $K^+$ concentration.

$[Mg^{2+}]$ = 0.4 mM (much lower than $C_m$) there is a significant density of $Mg^{2+}$ around $i \approx 125$, which covers the triple helix (TH) of *Azoarcus* ribozyme. Thus, we conclude that $Mg^{2+}$ binding is highly specific. More importantly, charge neutralization resulting in $\beta < 1$, occurs at the lowest $[Mg^{2+}]$ not uniformly, as predicted by the CIC theory, but does so in a manner that initiates folding. This would mean that different regions of this ribozyme order at somewhat distinct $Mg^{2+}$ concentrations, as illustrated previously (see Fig. 2b in (12)). In other words, the midpoint concentration, $C_m$, of the transition measured by $R_g$ or other parameters is the (appropriately-weighted) mean of such individual $[Mg^{2+}]$ at which distinct structural elements achieve the native conformation.

In the absence of or at low $Mg^{2+}$ concentrations, the amount of condensed $K^+$ depends significantly on positions along the RNA sequence. At low $Mg^{2+}$ concentrations, the ribozyme is unfolded with some helices intact. To see if there is any relationship between $K^+$ condensation and such residual structure, we compare the distribution of $c^*$ with the fraction of secondary structure formation in Fig. 7(a). The profile of condensed $K^+$ (green in Fig. 7a) has several characteristic minima that correspond to loops in the stem-loop structures of P2, P5, P6, P8, and P9. This indicates that the distribution of $K^+$ reflects the shape of RNA that is mostly due to residual secondary structures at low $[Mg^{2+}]$.

At high $Mg^{2+}$ concentration ($[Mg^{2+}] > 1.5$ mM), the localization of condensed $Mg^{2+}$ ions depends on the tertiary structures of the RNA (Fig. 6c and d). The local concentration of $Mg^{2+}$ correlates with the local density of phosphates (compare blue and gray lines in Fig. 7b). Under such conditions, all the native secondary structures are formed with almost unit probability, but we do not find any secondary-structure dependence of condensed $Mg^{2+}$.

The relationship between ion condensation and local density of phosphates is vividly illustrated in Fig. 8(a, b). At low $[Mg^{2+}]$, the ribozyme is mostly unfolded, predominantly containing only secondary structures. Because there are no stable tertiary interactions, $c_P^*$ values are small (Fig. 8a, horizontal axis. $c_P^* \approx 3$ M). Accordingly, $c^*$ for $Mg^{2+}$ are relatively low and $K^+$ are predominantly condensed. However, there are six nucleotides that have higher $c^*$ for both P and $Mg^{2+}$ (see the rectangle of

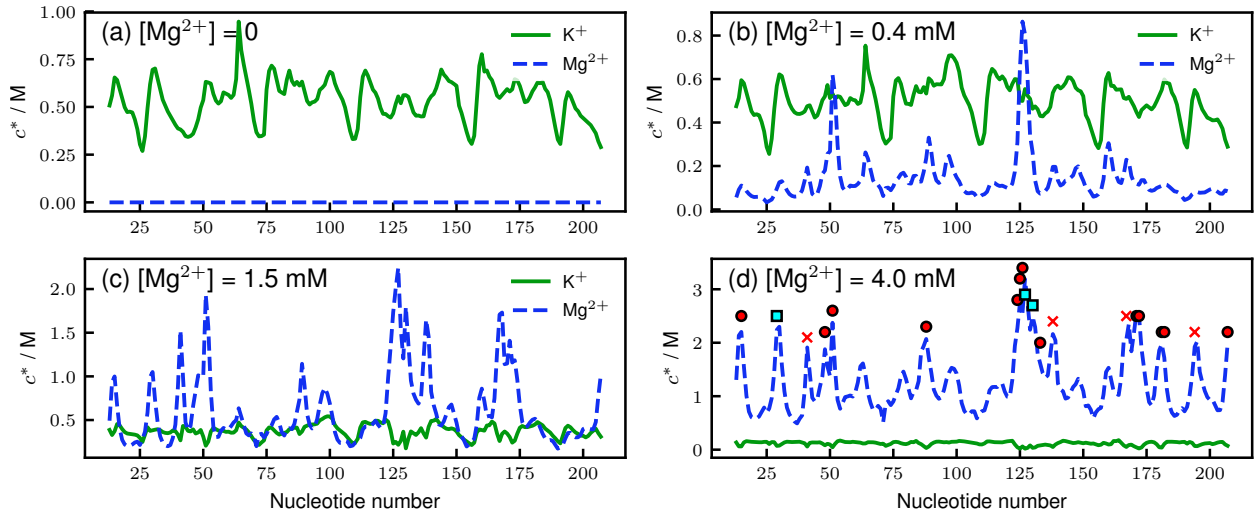


Figure 6: Fingerprint of ion coordination expressed as local ion concentrations, $c_i^*$ ($i$ is the nucleotide position), for $Mg^{2+}$ (blue) and $K^+$ (green). The local ion concentrations are calculated as the molarity of ions in the vicinity of each nucleotide ($r \leq \ell_B = 0.73$ nm). The peaks with varying heights at different nucleotides positions as ribozyme folds reflect the architecture of the folded state. Most importantly, $c^*$ is not uniform but varies dramatically along the nucleotides, which is very different from what is found for charged rods (see Fig. 3a). In the panel (d), positions of the $Mg^{2+}$-binding sites found in the crystal structure (15) are indicated by red circles (direct contact; nucleotides 15, 48, 51, 88, 124-126, 133, 171-172, 181-182, 207) and cyan squares (contact mediated by water molecules; nucleotides 29, 127, 130). The positions with high $c^*$ in the simulations but no $Mg^{2+}$ in the crystal structure are marked by red crosses (nucleotides 41, 138, 167, 194). The results at 50 mM $K^+$ are shown in Fig. S2 in the SI.

$c_P^* > 3.5$ M and $c_{Mg^{2+}}^* > 0.5$ M in Fig. 8a). We identified all those nucleotides to be ones involving the central triple helix (TH) formation; nucleotides 50, 51 and from 124 to 127. In accordance with the previous study (12), TH starts to form around $[Mg^{2+}]$ = 0.4 mM ($< C_m$). The data in Fig. 8 indicates that there are $Mg^{2+}$ ions specifically bound to the sites around TH. In contrast, at high $[Mg^{2+}]$, the entire ribozyme is folded and $c_P^*$ can be as high as ~7 M depending on the local architecture (Fig. 8b). The amount of condensed $Mg^{2+}$ is well correlated with the local concentration of phosphates. There is a weak anti-correlation between $K^+$ condensation and local phosphate concentration.

## Fluctuations of condensed ions

RNA folding is a stochastic process driven by thermal fluctuations and the solution conditions. Consequently, it would be interesting to see how these condensed ions fluctuate at different stages of folding. In Fig. S3, normalized fluctuations of the local ion concentrations, $\Delta c^*$ (Eq. 4), for each nucleotide are shown for the several solution conditions. Below or around the midpoint $Mg^{2+}$ concentration, where the ribozyme folding is not completed, fluctuations of condensed $K^+$ ions and $Mg^{2+}$ have similar position dependence. This correlation is more clearly seen in the logarithm plot of $\Delta c_{Mg}^*$ vs $\Delta c_K^*$ in Fig. 8(c). Comparing Figs. 6 and S3, we also find that nucleotide positions that have high $c^*$ values also have small $\Delta c^*$. This indicates that the specific positions, where more ions are condensed, bind those ions also more tightly. Once the ions reach the "equilibrium" positions, they tend to be localized.

The correlation between the fluctuations of $\Delta c_{Mg}^*$ and $\Delta c_K^*$ changes dramatically after the midpoint, $[Mg^{2+}] > 1.5$ mM (Fig. S3d and e). There is anti-correlation between $\log \Delta c_{Mg}^*$ and $\log \Delta c_K^*$ (Fig. 8d). As discussed in previous sections, there are specific positions where more $Mg^{2+}$ ions are condensed (Fig. 6) beyond the midpoint of $[Mg^{2+}]$. At these positions, there are smaller fluctuations in $c_{Mg}^*$ (smaller $\Delta c_{Mg}^*$) but higher fluctuations in $c_K^*$ (larger $\Delta c_K^*$), reflecting strong condensation of $Mg^{2+}$.

## Ion condensation in Calcium-driven folding

We have hitherto focused on the condensation of $K^+$ in the absence and presence of $Mg^{2+}$, and the condensation of $Mg^{2+}$ itself. It is known that $Ca^{2+}$ could also facilitate *Azoarcus* ribozyme folding although the folded state is not catalytically active and is

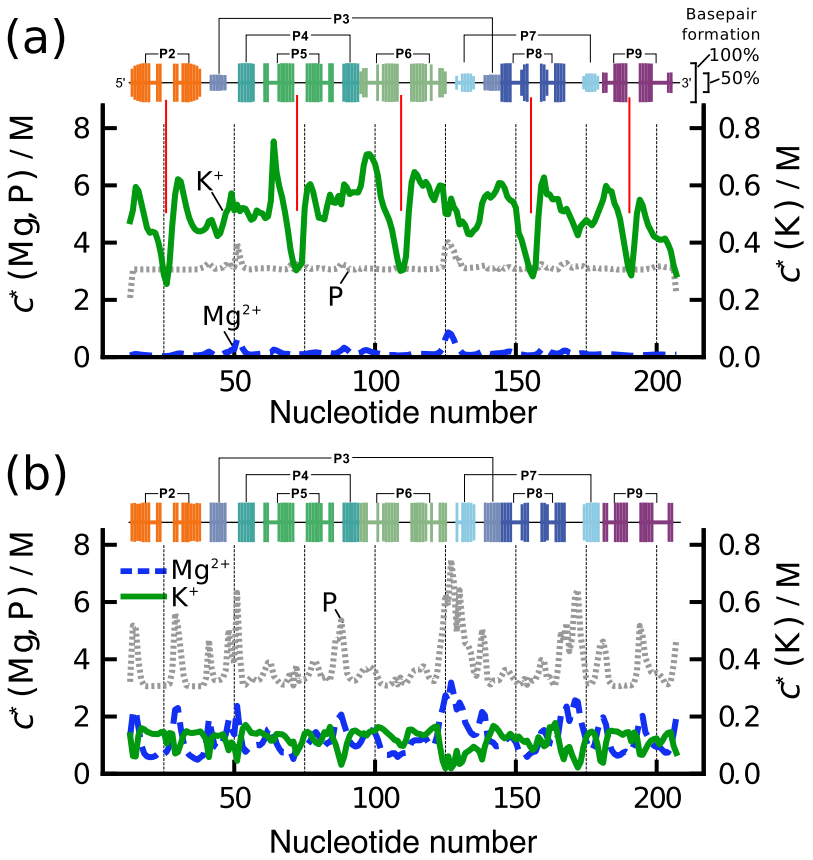


Figure 7: Relationship between $K^+$ condensation and residual structure. The fraction of secondary structure formation is shown on top as bars for each nucleotide. **(a)** At 0.4 mM $Mg^{2+}$, the profile of condensed $K^+$ (green) has several characteristic minima that correspond to loops in the stem-loop structures of P2, P5, P6, P8 and P9 (indicated by vertical red lines). See Fig. 1 for the structures. **(b)** At 4.0 mM $Mg^{2+}$, the profile of condensed $Mg^{2+}$ correlates with the profile of phosphate (gray dashed). Such a correlation does not exist at low $Mg^{2+}$ concentration. The secondary structures (various colored bars on top) are nearly fully formed.

modestly less compact than it is in the presence of $Mg^{2+}$ (12, 21). We repeated the same analyses in the case of $Ca^{2+}$-dependent folding instead of $Mg^{2+}$. Both average local concentration ($\overline{c^*}$) and its variance ($\sigma^2$) show similar trends as the case of $Mg^{2+}$ (Fig. 4b and c). At higher solution concentration ([Mg/$Ca^{2+}$] $\gtrsim$ 1 mM), the amount of condensation given by $\overline{c^*}$ is slightly less in the case of $Ca^{2+}$.

Fig. S4 shows the sequence dependence of the local concentration ($c^*$) and fluctuation ($\Delta c^*$) in the case of $Ca^{2+}$. Comparing them with Fig. 6 and S2, the condensation profiles of $Ca^{2+}$ and $K^+$ bear remarkable resemblances to the $Mg^{2+}$ case. At the highest concentration ([Mg/$Ca^{2+}$] = 30 mM), the local concentration of $Ca^{2+}$ has the same peak positions as $Mg^{2+}$, but has ~20 % lower values. Comparison of the fingerprint of ion concentration profiles between $Mg^{2+}$ and $Ca^{2+}$ clearly illustrates the efficacy of $Mg^{2+}$ of creating a compact and functionally competent ribozyme.

## DISCUSSION

The central result of this study is that the nature of ion condensation onto *Azoarcus* ribozyme differs drastically from theoretical predictions for regularly shaped polyanions. For example, the CIC theory shows that ions condense uniformly onto highly charged rods. The positively charged ions are localized at the positions of the charges on the rod. In sharp contrast, in *Azoarcus* ribozyme, and presumably in other RNA molecules as well, the condensation is highly non-uniform and nucleotide specific, and critically depends on the structure of the RNA. We note in passing that the prevailing view in much of the RNA literature is that most of the ions are uniformly condensed onto RNA although crystal structures have shown that there are site-specific localizations of $Mg^{2+}$. A natural explanation of our finding is that the electrostatic potential around the phosphate groups is non-uniform, with the folded structure with fully formed tertiary interactions containing pockets of highly negative potentials.

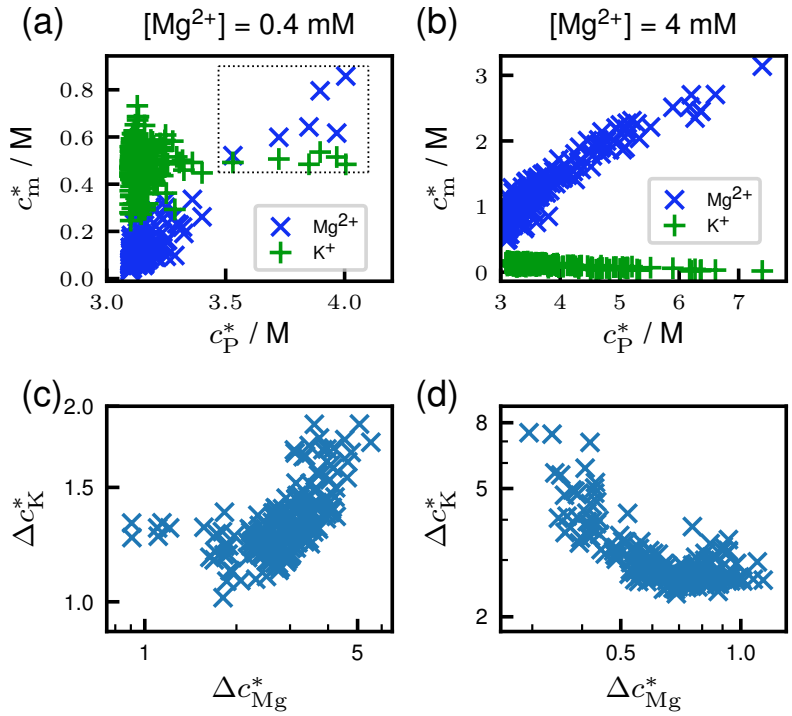


Figure 8: **(a, b)** Correlation between local ion concentrations of phosphate ($c^*_{\mathrm{P}}$) and the cations ($c^*_m$) where $m$ is either $K^+$ or $Mg^{2+}$. Each data point represents one nucleotide. There are 195 such data points in each panel. The solution conditions are (a) $[Mg^{2+}]$ = 0.4 mM, and (b) 4 mM, both in 50 mM KCl. In (a), the data points at $c^*_{\mathrm{P}} > 3.5$ M correspond nucleotides 50, 51, and 124-127 (inside the dotted rectangle), that are involving the tertiary helix. **(c, d)** Correlations of fluctuations, $\Delta c^*$, between $Mg^{2+}$ and $K^+$ at (c) $[Mg^{2+}]$ = 0.4 mM and (d) 4.0 mM in 50 mM KCl.

Consequently, ions are specifically drawn, with higher probability, to regions of large negative electrostatic potential with smaller ions approaching these regions easily. Indeed, this was first shown in our previous study using coarse-grained and atomic-detailed simulations (see Fig. S11 in (12)). However, what is surprising in this study is that the ions sense these regions even at extremely low divalent ion concentrations when there is no tertiary interaction present.

## Link to the shape fluctuations in the ribozyme

The non-uniform and site-specific ion-RNA interactions can be succinctly illustrated by considering the ion concentration-dependent changes in the shape parameters, asphericity $\Delta$ ($0 \leq \Delta \leq 1$) and prolateness $S$ ($-0.25 \leq S \leq 2$) (23, 24), which are readily computed from the inertia tensor that is related to the square of the radius of gyration. For reference, it is worth pointing out both $\Delta$ and $S$ are unity for rods, whereas $\Delta = S = 0$ for spheres. For long self-avoiding polymers, numerical results show that $\Delta = 0.55$ and $S = 0.92$ (a prolate ellipsoid) (25). In contrast to the values of the shape parameters for rods, the changes in $\langle\Delta\rangle$ and $\langle S\rangle$ as a function of $[Mg^{2+}]$ for the ribozyme, at the two values of $K^+$ concentration, exhibit unusual non-monotonic dependence as $[Mg^{2+}]$ is varied (inset in Fig. 9), which is not reflected in $R_g$. Fig. 9 also shows the distributions $P(\Delta)$ and $P(S)$ at three $Mg^{2+}$ concentrations. Surprisingly, the widths of $P(\Delta)$ and $P(S)$, which depend non-monotonically on $[Mg^{2+}]$, are relatively large. Therefore, shape fluctuations in the ribozyme are substantial. Two comments are worth making: (1) The values of $\Delta$ and $S$ depend greatly on the $K^+$ concentration, which is also reflected in the $R_g$ as a function of $Mg^{2+}$ concentration. (2) It is is interesting that the minima in $\langle\Delta\rangle$ and $\langle S\rangle$ are at $[Mg^{2+}] \approx 1.5$ mM, the midpoint of the folding transition, where the ribozyme almost resembles a sphere. At high and low $Mg^{2+}$ concentrations, the values of $\langle\Delta\rangle$ show that globally the ribozyme may be pictured as a prolate ellipsoid. It also implies that this ribozyme and others are not as densely packed as proteins (11). The irregular shape leads to the high specificity of ion–RNA interactions. It is likely that this finding would apply to any RNA molecules whose shape has grooves and whose surface are irregular. It is most interesting that there is a calculable fingerprint for ion-RNA interactions that is a reflection of the folded structure. A clear implication is that, even at very low ion concentration, binding is highly specific. This interpretation differs from the usual assertion that the stability of RNA is

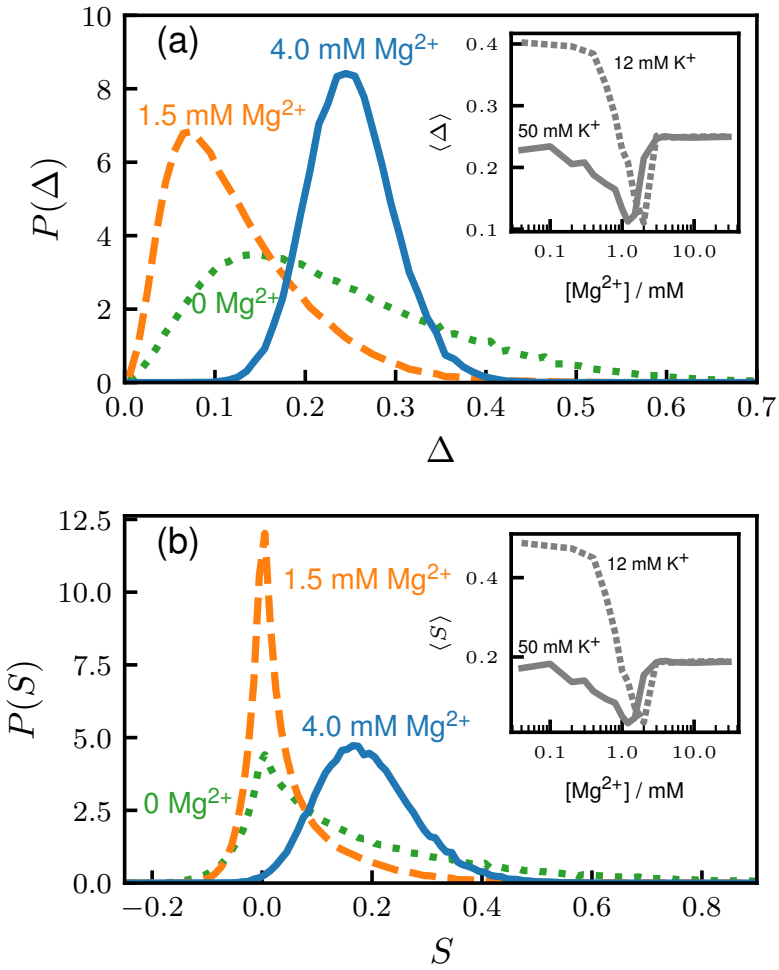


Figure 9: Probability distributions of shape parameters, (a) asphericity $\Delta$ and (b) prolateness $S$, for the *Azoarcus* ribozyme at three $Mg^{2+}$ concentrations, 0 mM (dotted green), 1.5 mM (dashed orange), and 4 mM (solid blue). If the shape is a perfect sphere then $\Delta = S = 0$, whereas a perfect rod corresponds to $\Delta = S = 1$. The insets show the average values of $\langle\Delta\rangle$ and $\langle S\rangle$ as a function of $[Mg^{2+}]$ at $[K^+]$ = 12 mM (dotted line) and 50 mM (solid line). The values of $\langle\Delta\rangle$ and $\langle S\rangle$ in the folded states are 0.24 and 0.17, respectively. Globally the shape of this ribozyme corresponds to a prolate ellipsoid.

determined largely by diffuse divalent cations (26).

## Correlation between ions at different RNA sites

We computed the correlations of $Mg^{2+}/K^+$ binding to two different nucleotides (Fig. S5). The correlations for a pair of nucleotides $(i, j)$ was calculated as the Pearson's coefficient of the number of $Mg^{2+}/K^+$ ions bound to site $i$ and site $j$ for the ensemble at a given $Mg^{2+}$ and $K^+$ concentrations. Overall, we did not find strong correlations involving $Mg^{2+}$ or $K^+$, except that some signatures appear as consequences of formations of secondary and tertiary structures. For example, there are strong correlations near the diagonal region because of helix domains P5, P6, and P8, as indicated by the red circles (Fig. S5a). In addition, there are high correlations at the interfaces of P3/P7 (see Figure 4 of Ref. (12)), and P2/P8. The observed high correlations arise because these nucleotides are close in space and share the same $Mg^{2+}$ ion when the secondary/tertiary interactions form. Interestingly, at the midpoint concentration $[Mg^{2+}] \sim 1.5$ mM, there are weak anti-correlations between P2 and P6a, and P2 and J8/7 (blue circles). Helix P2 is located at a peripheral part of the ribozyme, whereas P6a and J7/8 are close to the core region. These anti-correlations are a consequence of anti-cooperative folding of tertiary interactions, arising from topological frustration (12). Around the midpoint concentration of $[Mg^{2+}]$, there are still not enough amount of $Mg^{2+}$ ions to simultaneously satisfy all the folding elements. Thus, $Mg^{2+}$ binding to P2 reduces the probability of $Mg^{2+}$ binding to the core

region. The correlation maps between $K^+$-$K^+$ ions are less instructive.

### Flexible polyelectrolytes

Just as is the case for RNA, which does adopt a well-defined folded structure at high ion concentrations, the shapes of flexible PEs are also highly variable. Therefore, it follows that the CIC theory should not be quantitative in predicting the nature of ion condensation. Indeed, early simulations showed (27) that ion condensation is non-uniform in the presence of divalent and trivalent cations for PEs in poor solvents. Indeed, in certain cases, the divalent cations and the monomers are packed as in a crystalline arrangement, which would be inconsistent with the CIC theory. Using theory, Muthukumar (28) has shown that, even in monovalent salts ($z$=1), the condensation process in flexible PEs is non-uniform and could be understood by the variations in the local dielectric mismatches. A more recent simulation study (29) has also arrived at the conclusion that the condensation of monovalent ions depends crucially on the shape of the PE. In particular, they showed that non-uniform condensation is determined by the molecular topology. Thus, the conclusion that shapes determine ion condensation holds for both synthetic PEs and RNA. The latter has additional features (for example stacking interactions and hydrogen bonding), which make RNA different from a synthetic. In particular, the ion-induced shape changes in PEs are not as spectacular as they are in RNA.

## CONCLUSION

The Oosawa-Manning theory of ion condensation, which assumes that there is an equilibrium between bound and free ions, is the basis of our understanding of many thermodynamics properties of highly charged polyelectrolytes. As shown multiple times previously and in the simulations here, the predictions of the CIC theory are quantitatively correct if the charged polyelectrolytes adopt regular shapes, such as rods and spheres. This is not the case in RNA, which although is a polyanion, does not adopt a simple shape but undergoes large shape transitions as it folds. Our simulations of *Azoarcus* ribozyme in the presence of $Mg^{2+}$ or $Ca^{2+}$ in a buffer containing $K^+$ ions show that ion binding is highly specific. Surprisingly, the specificity of binding depends on the folded structure of the RNA. In other words, in this ribozyme with a complex fold, the charge neutralization of phosphates does not occur uniformly as predicted by the CIC theory for rod-like macroions. There are specific regions where condensation occurs even at very low concentrations of the divalent cations. Although we have arrived at this unexpected finding using simulations of one ribozyme, we expect that the main conclusion should be applicable for other RNA molecules as well.

We conclude with the following additional comments. (1) We find, perhaps not surprisingly, that monovalent ions unbind (termed as counterion release process) when RNA is titrated with divalent ions (Fig. 6). This validates one of the predictions of the CIC theory. The release of $K^+$, however, occurs from specific nucleotide locations on the RNA, which is not anticipated in the CIC theory. (2) It is unclear to us if our findings, especially the exquisite specificity of the predicted nucleotide-dependent binding of ions to RNA, could be tested in experiments because of the highly correlated and many-body nature of ion condensation. Multicolor FRET experiments may be used to label regions of high affinity for divalent cations. The conformational changes as a consequence of ion binding and unbinding at these locations could be used, most likely in conjunction with simulations, to determine the architecture-dependent interaction of ions with RNA. (3) Our results show that charge renormalization of phosphates occurs in a highly heterogeneous manner and not by the diffuse binding picture of ion-RNA interactions. The fingerprints of ion localizations show that the extent of charge neutralization depends on the nucleotide position – finding that has to be incorporated in calculating free energy changes of RNA as it folds. (4) The clear implication of our results is that, even when using coarse-grained models, ions (especially those with valence greater than unity) should be modeled explicitly to reflect their sizes and charge densities. It is possible that electrostatic interactions due to monovalent ions can be accounted for by using effective interactions (30). However, in order to obtain accurate thermodynamics of RNA folding, divalent ions have to be explicitly modeled (12, 31).

## AUTHOR CONTRIBUTIONS

NH, NAD, and DT designed the research; NH and NAD performed all the simulations; NH, NAD, and DT analyzed the data; NH and DT wrote the article.

## ACKNOWLEDGMENTS

We thank Dr. Hung T. Nguyen for useful discussions. This work was supported by a grant from the National Science Foundation (CHE 16-36424) and the Welch Foundation (F-0019) through the Collie–Welch Chair.

## REFERENCES


1. Sun, L.-Z., D. Zhang, and S.-J. Chen, 2017. Theory and modeling of RNA structure and interactions with metal ions and small molecules. *Annu. Rev. Biophys.* 46:227–246.
2. Yu, T., and S.-J. Chen, 2018. Hexahydrated $Mg^{2+}$ binding and outer-shell dehydration on RNA surface. *Biophys. J.* 114:1274–1284.
3. Pollack, L., 2011. Time resolved SAXS and RNA folding. *Biopolymers* 95:543–549.
4. Kirmizialtin, S., S. A. Pabit, S. P. Meisburger, L. Pollack, and R. Elber, 2012. RNA and its ionic cloud: Solution scattering experiments and atomically detailed simulations. *Biophys. J.* 102:819–828.
5. Denesyuk, N. A., and D. Thirumalai, 2013. Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B* 117:4901–4911.
6. Oosawa, F., 1957. A simple theory of thermodynamic properties of polyelectrolyte solutions. *J. Polym. Sci.* 23:421–430.
7. Manning, G. S., 1969. Limiting laws and counterion condensation in polyelectrolyte solutions. I. Colligative properties. *J. Chem. Phys.* 51:924–933.
8. Oosawa, F., 1971. Polyelectrolytes. Marcel Dekker.
9. Alexander, S., P. Chaikin, P. Grant, G. Morales, P. Pincus, and D. Hone, 1984. Charge renormalization, osmotic pressure, and bulk modulus of colloidal crystals: Theory. *J. Chem. Phys.* 80:5776–5781.
10. Heilman-Miller, S. L., D. Thirumalai, and S. A. Woodson, 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme I. Equilibrium stabilization by cations. *J. Mol. Biol.* 306:1157–1166.
11. Hyeon, C., R. I. Dima, and D. Thirumalai, 2006. Size, shape, and flexibility of RNA structures. *J. Chem. Phys.* 125:194905–10.
12. Denesyuk, N. A., and D. Thirumalai, 2015. How do metal ions direct ribozyme folding? *Nature Chem.* 7:793–801.
13. Heilman-Miller, S. L., J. Pan, D. Thirumalai, and S. A. Woodson, 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme II. Counterion-dependence of folding kinetics. *J. Mol. Biol.* 309:57–68.
14. Thirumalai, D., N. Lee, S. A. Woodson, and D. Klimov, 2001. Early events in RNA folding. *Annu. Rev. Phys. Chem* 52:751–762.
15. Adams, P. L., M. R. Stahley, A. B. Kosek, J. Wang, and S. A. Strobel, 2004. Crystal structure of a self-splicing group I intron with both exons. *Nature* 430:45–50.
16. Grønbech-Jensen, N., R. J. Mashl, R. F. Bruinsma, and W. M. Gelbart, 1997. Counterion-induced attraction between rigid polyelectrolytes. *Phys. Rev. Lett.* 78:2477–2480.
17. Borukhov, I., D. Andelman, and H. Orland, 1997. Steric effects in electrolytes: A modified Poisson–Boltzmann equation. *Phys. Rev. Lett.* 79:435.
18. Deserno, M., C. Holm, and S. May, 2000. Fraction of condensed counterions around a charged rod: Comparison of Poisson–Boltzmann theory and computer simulations. *Macromolecules* 33:199–206.
19. Cha, M., S. Ro, and Y. W. Kim, 2018. Rodlike counterions near charged cylinders: Counterion condensation and intercylinder interaction. *Phys. Rev. Lett.* 121:058001.
20. Hyeon, C., and D. Thirumalai, 2005. Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. U. S. A.* 102:6789–6794.
21. Behrouzi, R., J. H. Roh, J. D. Kilburn, R. M. Briber, and S. A. Woodson, 2012. Cooperative tertiary interaction network guides RNA folding. *Cell* 149:348–357.
22. Honeycutt, J. D., and D. Thirumalai, 1992. The nature of folded states of globular proteins. *Biopolymers* 32:695–709.
23. Aronovitz, J. A., and D. R. Nelson, 1986. Universal features of polymer shapes. *Journal de physique* 47:1445–1456.

24. Dima, R. I., and D. Thirumalai, 2004. Asymmetry in the shapes of folded and denatured states of proteins. *J. Phys. Chem. B* 108:6564–6570.

25. Honeycutt, J. D., and D. Thirumalai, 1989. Static properties of polymer chains in porous media. *J. Chem. Phys.* 90:4542–19.

26. Soto, A. M., V. Misra, and D. E. Draper, 2007. Tertiary structure of an RNA pseudoknot is stabilized by "diffuse" $Mg^{2+}$ ions. *Biochemistry* 46:2973–2983.

27. Lee, N., and D. Thirumalai, 2001. Dynamics of collapse of flexible polyelectrolytes in poor solvents. *Macromolecules* 34:3446–3457.

28. Muthukumar, M., 2004. Theory of counter-ion condensation on flexible polyelectrolytes: Adsorption mechanism. *J. Chem. Phys.* 120:9343–9350.

29. Chremos, A., and J. F. Douglas, 2016. Counter-ion distribution around flexible polyelectrolytes having different molecular architecture. *Soft Matter* 12:2932–2941.

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

31. Hayes, R. L., J. K. Noel, A. Mandic, P. C. Whitford, K. Y. Sanbonmatsu, U. Mohanty, and J. N. Onuchic, 2015. Generalized Manning condensation model captures the RNA ion atmosphere. *Phys. Rev. Lett.* 114:258105–258106.

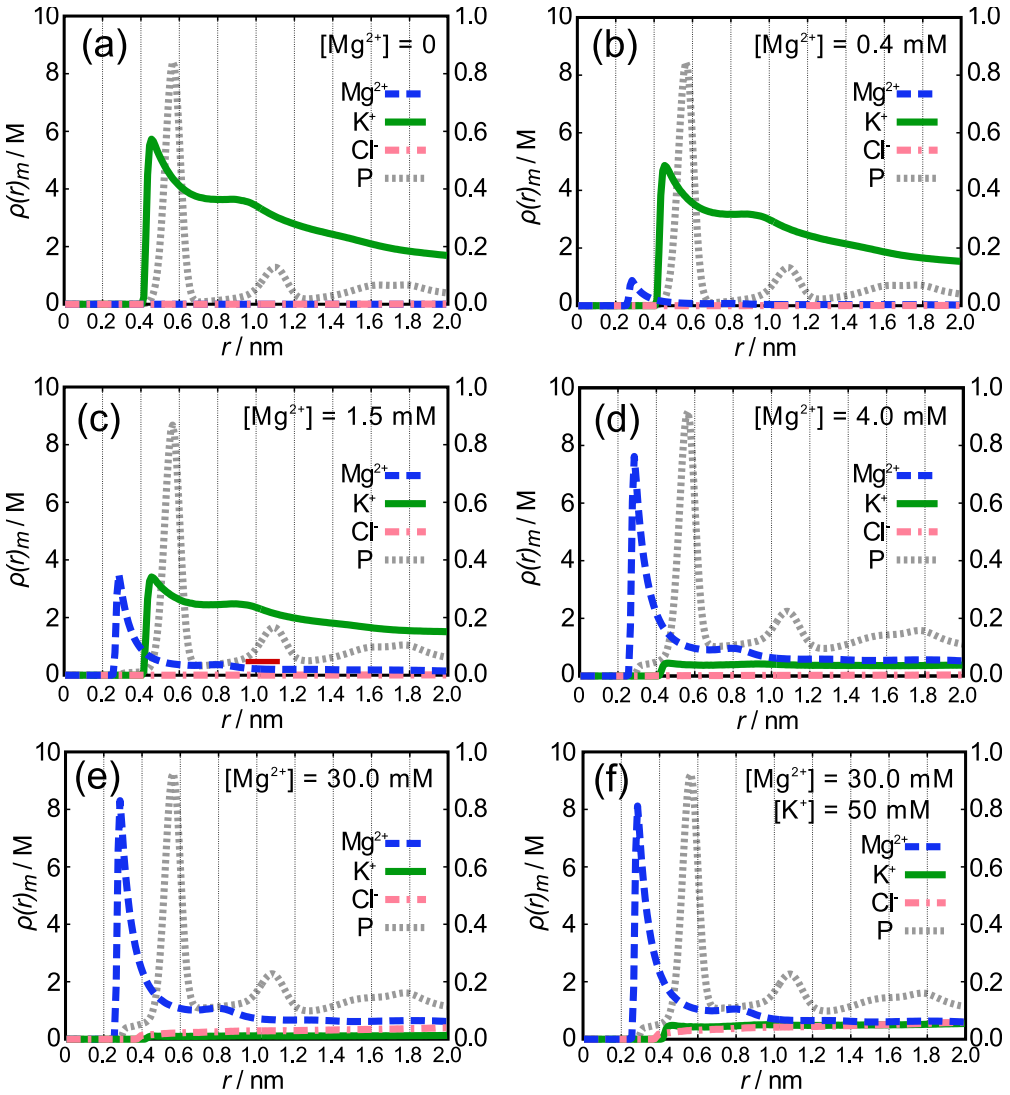


**Figure S1:** Radial profiles of local ion concentration, $\rho(r)$, for $Mg^{2+}$, $K^+$, $Cl^-$, and phosphate (P) at $[K^+] = 12$ mM and various $[Mg^{2+}]$: (a) 0, (b) 0.4 mM, (c) 1.5 mM, (d) 4.0 mM, and (e) 30 mM. (f) The profile at $[K^+] = 50$ mM and $[Mg^{2+}] = 30$ mM. Appropriate amount of $Cl^-$ were added to the solution for the electroneutrality. Note that the scale for $\rho_{Mg^{2+}}$ and $\rho_P$ (left axis) is 10 times greater than $\rho_{K^+}$ and $\rho_{Cl^-}$ (right axis).

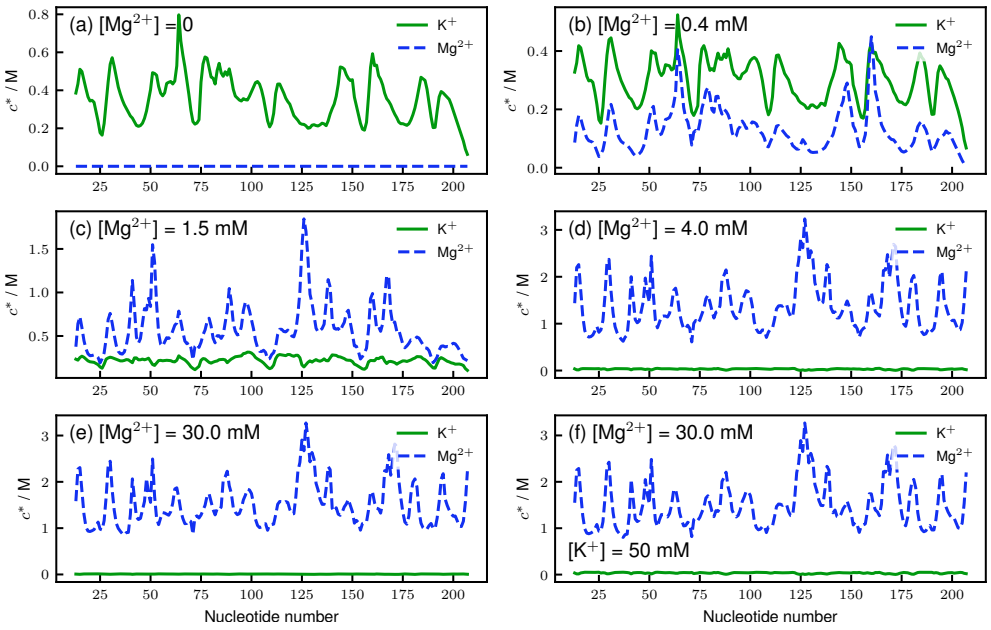


**Figure S2:** Local ion concentrations, $c_i^*$ ($i$ is the nucleotide position), for $Mg^{2+}$ (blue) and $K^+$ (green) at $[K^+] = 12$ mM and various $[Mg^{2+}]$: (a) 0, (b) 0.4 mM, (c) 1.5 mM, (d) 4.0 mM, and (e) 30 mM. (f) The profile at $[K^+] = 50$ mM and $[Mg^{2+}] = 30$ mM.

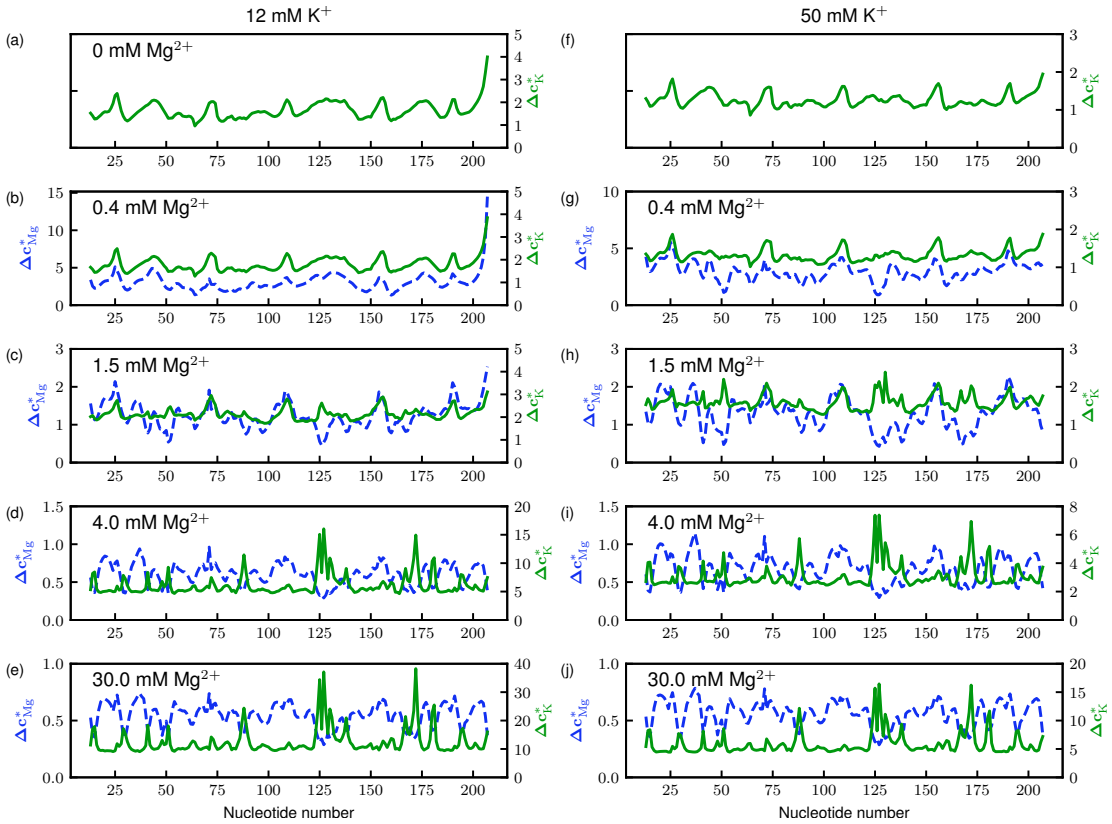


**Figure S3:** Normalized fluctuations of local ion concentrations, $\Delta c^*$ (Eq. 4), for $Mg^{2+}$ (dotted blue) and $K^+$ (solid green), in 12 mM KCl (left column; a–e) and 50 mM KCl (right column; f–j).

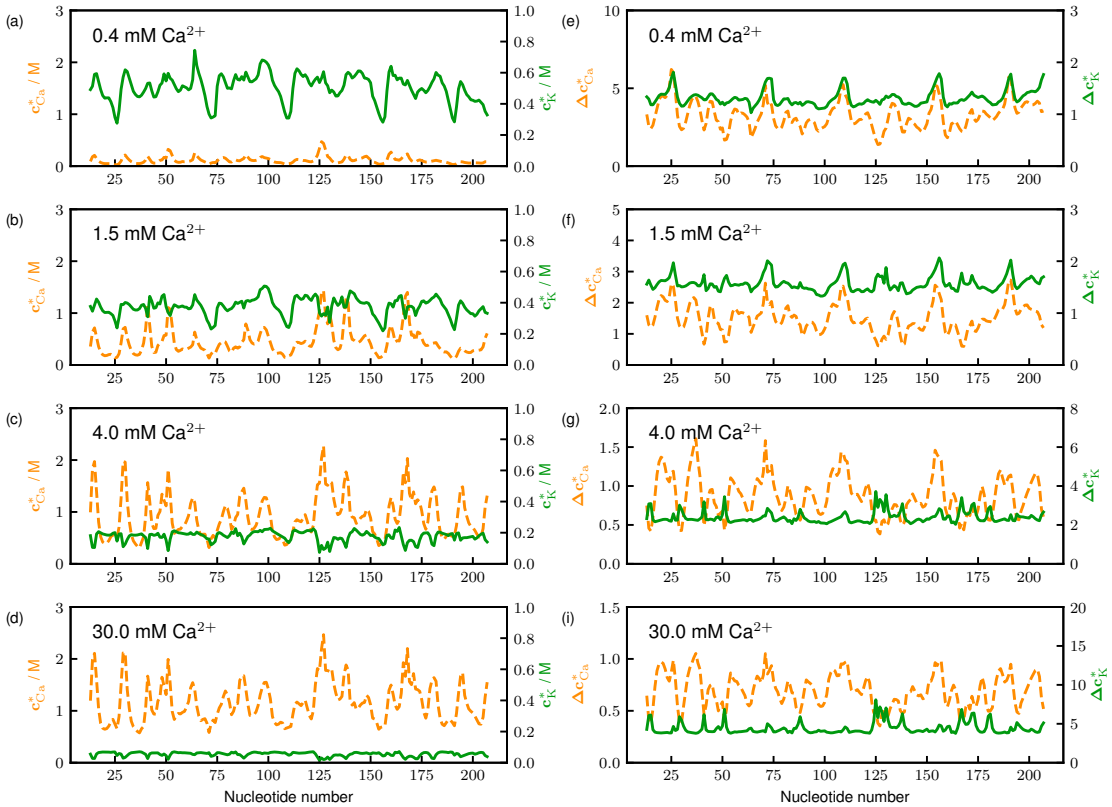


**Figure S4:** Sequence dependence of the local ion concentration (left column, $c^*$) and its fluctuations (right column, $\Delta c^*$) in the presence of $Ca^{2+}$ ions instead of $Mg^{2+}$. The concentration of $Ca^{2+}$ is labeled in each panel. [KCl] is fixed at 50 mM.

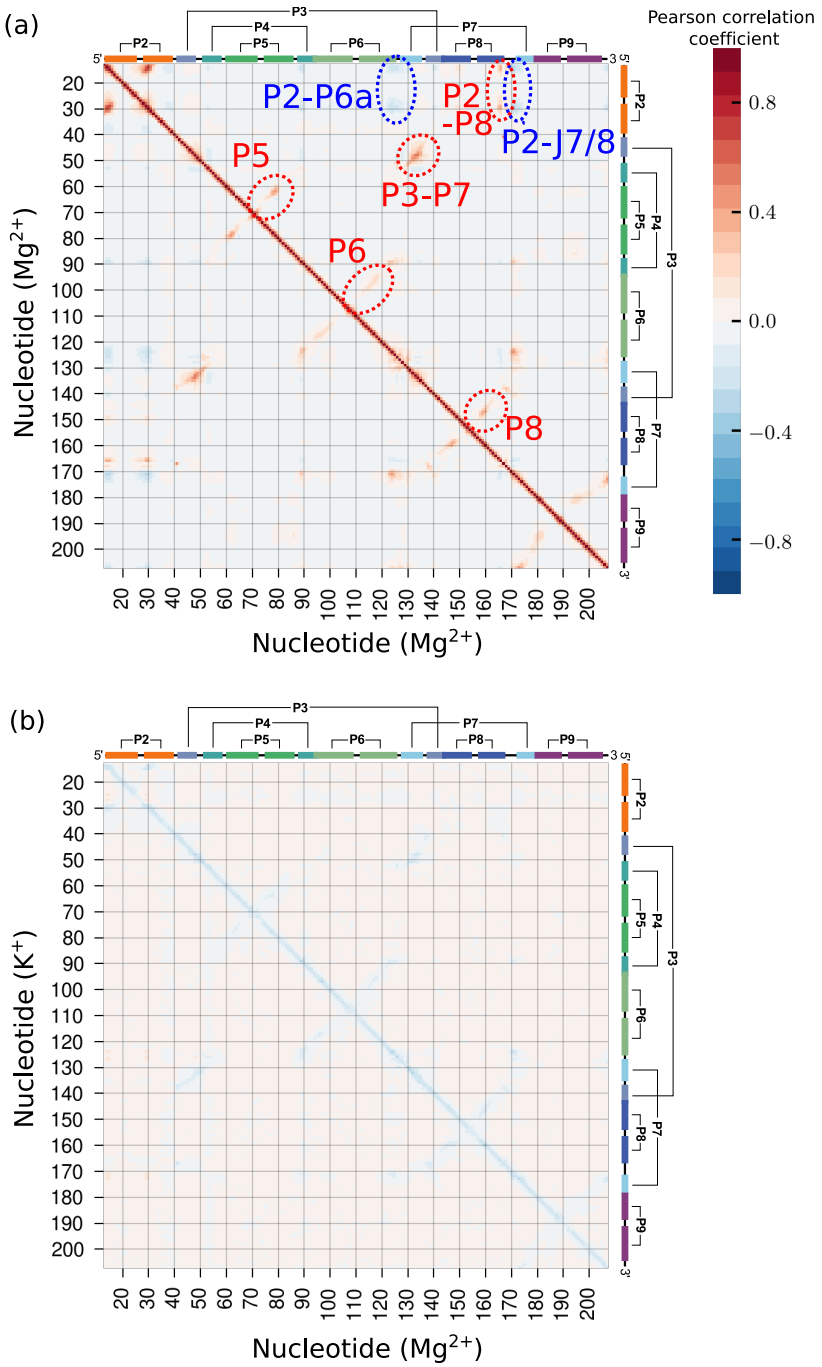


**Figure S5: (a)** Correlations between $Mg^{2+}$ condensations to two different nucleotides. The horizontal and vertical axes are the nucleotide positions $i$ and $j$. Color in each cell $(i, j)$ represents the Pearson correlation coefficient (scale on the right) of the number of bound ions at site $i$ and $j$. The $Mg^{2+}$ concentration is 1.5 mM and $[K^+] = 50$ mM. **(b)** Correlations between $Mg^{2+}$-$K^+$ condensations is most prominent along the diagonal. This implies that $K^+$ ion is released upon $Mg^{2+}$ binding.

**Table S1: Parameters for excluded volume and electrostatic interactions.**

| Bead type | $R_i$/nm * | $\varepsilon_i$/kcal mol$^{-1}$ | $z_i/e$ |
|---|---|---|---|
| Phosphate | 0.21 | 0.2 | -1 |
| Rod charge | 0.21 | 0.2 | -1 |
| Sugar | 0.29 | 0.2 | 0 |
| Base A | 0.28 | 0.2 | 0 |
| Base G | 0.30 | 0.2 | 0 |
| Base C | 0.27 | 0.2 | 0 |
| Base U | 0.27 | 0.2 | 0 |
| $Mg^{2+}$ | 0.08 | 0.9 | 2 |
| $Ca^{2+}$ | 0.17 | 0.5 | 2 |
| $Cl^{-}$ | 0.19 | 0.3 | -1 |
| $K^{+}$ | 0.27 | 0.0003 | 1 |

* If both interacting sites are RNA sites, we take $R_i + R_j = 0.32$ nm.