Theory and simulations for RNA folding in mixtures of monovalent and divalent cations

Hung T. Nguyen[a], Naoto Hori[a], and D. Thirumalai[a,1]

[a]Department of Chemistry, The University of Texas at Austin, Austin, TX 78712

Edited by Peter J. Rossky, Rice University, Houston, TX, and approved September 6, 2019 (received for review July 7, 2019)

RNA molecules cannot fold in the absence of counterions. Experiments are typically performed in the presence of monovalent and divalent cations. How to treat the impact of a solution containing a mixture of both ion types on RNA folding has remained a challenging problem for decades. By exploiting the large concentration difference between divalent and monovalent ions used in experiments, we develop a theory based on the reference interaction site model (RISM), which allows us to treat divalent cations explicitly while keeping the implicit screening effect due to monovalent ions. Our theory captures both the inner shell and outer shell coordination of divalent cations to phosphate groups, which we demonstrate is crucial for an accurate calculation of RNA folding thermodynamics. The RISM theory for ion–phosphate interactions when combined with simulations based on a transferable coarse-grained model allows us to predict accurately the folding of several RNA molecules in a mixture containing monovalent and divalent ions. The calculated folding free energies and ion-preferential coefficients for RNA molecules (pseudoknots, a fragment of the rRNA, and the aptamer domain of the adenine riboswitch) are in excellent agreement with experiments over a wide range of monovalent and divalent ion concentrations. Because the theory is general, it can be readily used to investigate ion and sequence effects on DNA properties.

RNA folding | free energy | ion preferential interaction coefficients | three-interaction site (TIS) model | reference interaction site model (RISM)

The lack of a rigorous and thermodynamically consistent treatment of interactions between counterions and RNA has impeded a quantitative description of the self-assembly of RNA molecules (1). Although many factors contribute to the stability of a folded RNA molecule, the interplay between monovalent and divalent cations and the highly correlated nature of ion–RNA interactions make it challenging to develop an accurate and tractable theory for RNA folding thermodynamics and kinetics. The effects of monovalent ions could be accurately accounted for by using the Debye–Hückel theory (2–4). However, theoretical and computationally tractable treatments of the effects of divalent cations, such as Mg^{2+}, which play an essential role in RNA structure, folding, and function (5–8), have not been fully developed. Accounting for the effects of divalent ions on RNA folding requires an approach that goes beyond the use of the Poisson–Boltzmann equation (9–11) to account for the ion size, ion–ion correlations, and the complex coordination with the phosphate groups. The simultaneous presence of monovalent and divalent ions introduces additional complexity that has to be dealt with to arrive at a reasonable predictive theory of RNA folding.

How ions modulate the RNA energy landscape has also been the subject of extensive experimental and theoretical studies (1, 9, 12–20). From the chemistry perspective, binding of divalent ions to the negatively charged phosphate groups could be conceptually classified into 2 categories: direct (or inner-sphere) contact where an atom (or more) of the RNA is part of the divalent ion coordination sphere and indirect (or outer-sphere) contact where the interaction is mediated by a water molecule (5). Recent surveys of the RNA structures in the Protein Data Bank (PDB) reported the frequencies of both inner and outer spheres Mg^{2+} binding to RNA atoms (21, 22), which suggests that a theoretical model must take these interactions into account to describe RNA folding. However, a complete knowledge of the distribution of ions around RNA in solution is still lacking although initial insights have been provided in a recent study (23).

To arrive at an accurate model, which reliably predicts the thermodynamic properties of large RNA molecules, we first began with a sequence-dependent three-interaction site (TIS) coarse-grained (CG) model for nucleic acids (24), which has been adopted to study a range of problems related to RNA folding (2, 10, 25–32). Even with this simplification, the inclusion of both monovalent ions (present in excess concentration relative to divalent cations) and divalent ions explicitly is computationally demanding, although folding simulations of a 195-nt *Azoarcus* ribozyme and pseudoknots have been carried out successfully (10, 33). Here, we report the folding thermodynamics of RNA molecules using simulations performed with a hybrid model in which monovalent ions are implicitly treated but divalent cations are explicitly included. To develop such a model, we resort to the reference interaction site model (RISM) theory (34–40) to obtain the potential of mean force (PMF) between divalent cations and phosphate groups. We show that this treatment is necessary to obtain accurate results for RNA thermodynamics, especially for Mg^{2+} ions, which are involved in both the inner and outer shell coordination with the negatively charged phosphate groups. Using our model and simulations, we calculate, with high accuracy, the thermodynamics of RNA folding in the presence of divalent and monovalent cations for several

Significance

RNA molecules require ions to fold. The problem of how ions of differing sizes and valences drive the folding of RNA molecules is unsolved. Here, we take a major step in its solution by creating a method, based on the theory of polyatomic liquids, to calculate the potential between divalent ions and the phosphate groups. The resulting model, accounting for inner and outer sphere coordination of Mg^{2+} and Ca^{2+} to phosphates, when used in coarse-grained molecular simulations, predicts folding free energies for a number of RNA molecules in the presence of both divalent and monovalent ions that are in excellent agreement with experiments. This work sets the stage for probing sequence and ion effects on DNA and synthetic polyelectrolytes.

Author contributions: H.T.N. and D.T. designed research; H.T.N. and N.H. performed research; H.T.N., N.H., and D.T. analyzed data; and H.T.N., N.H., and D.T. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: dave.thirumalai@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1911632116/-/DCSupplemental.

First published September 30, 2019.

RNA molecules in the folded, intermediate, and unfolded states. In the process, we establish that divalent ions interact strongly with RNA molecules in the intermediate structures (41), which provides a structural interpretation of site-specific interactions between RNA and ions. Our work also shows that accounting for both inner- and outer-sphere coordination of divalent cations with RNA is necessary to faithfully reproduce divalent cation–RNA interaction thermodynamics. The general framework, which integrates liquid-state theories with molecular simulations, is applicable to investigate folding of large RNA molecules over a broad range of salt concentrations, thus vastly expanding the scope of simulations to a variety of problems in RNA biology.

Theory

RNA Model. We adopt the TIS model, in which each nucleotide is modeled using three interaction sites located at the center of geometry of phosphate (P), sugar (S), and base (B) (2, 3, 10, 24, 31). The energy function, $U = U_{\mathrm{BA}} + U_{\mathrm{EV}} + U_{\mathrm{ST}} + U_{\mathrm{HB}} + U_{\mathrm{EL}}$, takes into account the bond length and angle constraints, excluded volume interactions, secondary stacking between consecutive bases and tertiary stacking (stacking between bases that are not consecutive in the structure), and both nonnative and native hydrogen bond interactions. Details of the force field are given in *SI Appendix*.

Previously (10), we treated all of the ions (including monovalent ions) explicitly, and therefore the phosphate charge was fixed at $Q = -e$. Here, we employ the Debye–Hückel (DH) equation to approximate the screening effect of monovalent ions. Thus, the electrostatic interactions between the P–P and the divalent cations are written as

$$U_e(r) = \frac{1}{2\epsilon}\sum_{i,j}\frac{Q_i Q_j \exp\left(-\frac{r}{\lambda}\right)}{r}, \qquad [1]$$

where $\lambda = (8\pi l_B \rho_1)^{-1/2}$ is the DH screening length that depends on the number density of monovalent ions, ρ_1, and $l_B = \frac{e^2}{4\pi\varepsilon k_B T}$ is the Bjerrum length. For divalent cations $Q_i = +2e$. The renormalized charge on the phosphate $Q_{\mathrm{P}}(T, C_1, C_2)$ depends on the concentrations of both the monovalent (C_1) and divalent ions (C_2), which is calculated using the counterion condensation (CIC) theory (42–44). The electrostatic interaction between the divalent ions and phosphate groups is treated precisely to account for water-mediated outer- and inner-shell interactions (below). Because the X^{2+}–P potential (X is Mg or Ca) includes the excluded volume interactions, we do not explicitly account for such interactions in the coarse-grained TIS force field (*SI Appendix*).

Phosphate Charge Renormalization. A consequence of CIC is that the effective charge on the phosphate is reduced from $-1e$, thus softening the overall electrostatic interactions enabling the compaction and folding of the RNA. Following our earlier studies (2, 3), we include ion condensation effects for the implicitly treated monovalent ion. Since we treat divalent ions explicitly and monovalent ions implicitly, a thermodynamically consistent treatment of ion effects is needed. As the divalent ion concentration increases, the condensed divalent ions outcompete monovalent ions for the phosphate groups. This occurs because for each condensed divalent ion, approximately 2 monovalent ions are released, which is favored because the overall entropy of the system is increased.

In the mixed ion system, we assume that 1 divalent ion replaces exactly 2 monovalent ions, and the total RNA charge neutralized in the process is equal to those in the monovalent salt alone. In other words, if θ_1 and θ_2 are the numbers of condensed monovalent and divalent ions per phosphate group, respectively, then

$$\theta_1 + 2\theta_2 = \theta = 1 - \frac{b}{l_B(T)}, \qquad [2]$$

with the mean spacing between phosphate charges, $b = 4.4$ Å, a value used in our previous studies (2, 31). Although a more complicated treatment based on the balance between the interaction energy of ion–phosphate and entropic effects is possible (45), we find that this simple approximation works well for a broad range of ion concentrations.

A relation between θ_1 and θ_2 can be derived by considering the entropic cost of localizing 1 divalent ion vs. 2 monovalent ions. By neglecting ion–ion correlation effects, we obtain

$$\ln\left(\frac{e\theta_2}{C_2 V_2}\right) = 2\ln\left(\frac{e\theta_1}{C_1 V_1}\right), \qquad [3]$$

where e is Euler's number, and C_i and V_i are, respectively, the bulk concentration and the effective condensation volume of ion i. We calculated V_i using (43, 45)

$$V_i = 4\pi e b^3 (1 + Z_i)(\xi - 1/Z_i), \qquad [4]$$

where Z_i is the bare charge of the ions ($Z_1 = +1$, $Z_2 = +2$), and the Manning parameter, $\xi = \frac{l_B}{b}$. From Eqs. **2** and **3**, one can determine both θ_1 and θ_2. Thus, the effective charge on the phosphate is $Q_{\mathrm{P}}(T, C_1, C_2) = 1 - \theta_1$, considering only the monovalent ion condensation. We account for the contributions from the divalent cations, θ_2, by treating them explicitly in the simulations. The electrostatic interactions involving the P groups are calculated using $Q_{\mathrm{P}}(T, C_1, C_2)$, as the effective charge on the phosphate.

Mg^{2+}–P Effective Potential. To compute the Mg^{2+}–P effective potential, $V_{\text{Mg-P}}(r)$, the PMF between Mg^{2+}–P derived from the RISM theory has to be modified because it is dependent on temperature and concentrations of both the monovalent and divalent ions. A number of studies have shown that it is difficult to capture the short-ranged electrostatic interactions, which has prompted others to propose several ways of separating the Coulomb potential into short- and long-ranged components (46–49). The short-ranged interactions are usually determined using molecular simulations. Our approach, which is closest in spirit to a more rigorous treatment by Weeks and coworkers (50, 51), is implemented as follows. The short-ranged part of the $V_{\text{Mg-P}}$ is taken to be identical to the PMF, while the long-ranged part is corrected based on the temperature and salt concentrations. We write the effective potential $V_{\text{Mg-P}}(r)$ as

$$V_{\text{Mg-P}}(r) = W(r) + [U_{\mathrm{DH}}(r) - W(r)]\exp\left(-\frac{a^2}{r^2}\right), \qquad [5]$$

where $W(r)$ is the PMF calculated using the RISM theory (details are in *SI Appendix*), and $U_{\mathrm{DH}}(r)$ is the DH potential between Mg^{2+}–P, accounting for the screening effect of monovalent ions. (We tried other functional forms to combine W and U_{DH} and found that Eq. **5** served our purposes in both maintaining the direct-contact interaction and smoothly merging V to U_{DH} at $r \gg a$.) The constant, $a = 5.0$ Å, was chosen to preserve the Mg^{2+}–P direct-interaction energy. Thus, at short distances a Mg^{2+} ion (or more generally, any spherical divalent cation) in proximity to the phosphate group would interact according to the PMF calculated theoretically using RISM. At

large values of r, the Mg^{2+} ion would experience a screened phosphate charge due to the presence of monovalent ions. An advantage of our approach is that we calculate the fully equilibrated PMF between the Mg^{2+} cation and phosphate using numerical solution of the RISM equations instead of relying on MD simulations (46–49). This is particularly important for divalent ions that have slow ion–water exchange rate and cannot be reliably implemented using conventional all-atom MD simulations (52–57).

The calculated PMF between Mg^{2+}–P is shown in Fig. 1 for a solution containing 1.0 mM magnesium monophosphate. It is worth pointing out that $V_{\text{Mg-P}}(r)$ (the red curve in Fig. 1) has 2 minima, one corresponding to the inner-sphere Mg^{2+} coordination to P, and the other is the outer-sphere coordination. Accounting for both the direct contact and water-mediated Mg^{2+}–P interactions using Eq. **5** is necessary to calculate Mg^{2+}-induced RNA folding accurately since the folding of most large RNA molecules requires both tightly bound and screening due to Mg^{2+} ions (5, 7, 10, 21). We used a similar procedure to obtain interactions for calcium $V_{\text{Ca-P}}(r)$.

Results

Determination of the Parameters in the TIS RNA Model. The 2 adjustable parameters in our RNA force field are U^0_{hb} that determines the strength of hydrogen bonds and ΔG_0 that dictates the balance between stacking and hydrogen bonding (details described in *SI Appendix*). Following our previous study (2), we determined the values of the 2 parameters by reproducing the experimental heat capacity curves of human telomerase RNA hairpin and a viral pseudoknot. The melting temperatures of the 2 RNA motifs are reproduced well by simulations using the model (largest deviation is $<5\,^{\circ}$C) (*SI Appendix*, Fig. S2). In the rest of this paper, we use this set of parameters, coupled with our treatment of divalent ion–phosphate interactions, to investigate the effects of a mixture of divalent and monovalent ions on folding of 3 RNA molecules. It is worth emphasizing that the same set of parameters is used for all RNA molecules over a wide range of ion concentrations.

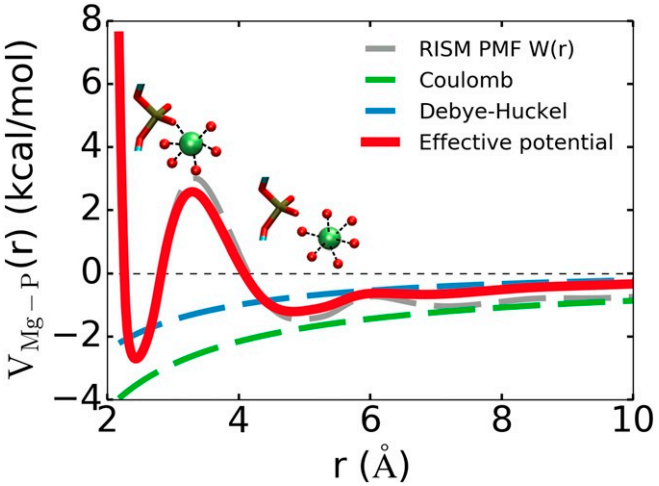

Fig. 1. Effective (Eq. **5**) Mg^{2+}–P potential (red) constructed by combining the short-ranged part of the PMF (gray dashed line) with the long-ranged Debye–Hückel potential (blue dashed line). Calculations of the PMF were performed at 1.0 mM magnesium monophosphate and 25 °C using the RISM theory (for details see *SI Appendix*). The first minimum represents the inner-shell interaction, where Mg^{2+} interacts directly with the phosphate groups. The second minimum at $r \approx 5$ Å represents the outer-shell interaction, where Mg^{2+} retains its first hydration shell.

Preferential Interaction Coefficient as a Function of Mg^{2+} Concentration. We first performed CG simulations to probe the binding of divalent cations to the RNA, expressed in terms of the experimentally measurable ion-preferential interaction coefficient, Γ_X (where X is Mg or Ca). In the simulations, a single RNA molecule was placed in a cubic simulation box in the presence of explicitly modeled divalent cations. After equilibration, we calculated Γ_X using the Kirkwood–Buff integral (58–63),

$$\Gamma_X = C_2 \int [g_X(\mathbf{r}) - 1]\, d\mathbf{r}, \qquad [6]$$

where $g_X(\mathbf{r})$ is the 3D distribution function of the divalent cations, reflecting the excess (or deficit) of X^{2+} relative to the bulk ion concentration, C_2, in the presence of the RNA.

In Fig. 2, we show Γ_{Mg} for beet western yellow virus pseudoknot (BWYV PK), a 58-nt fragment of the ribosomal RNA (rRNA), and the aptamer domain of adenine riboswitch at different concentrations of monovalent and magnesium ions. The BWYV PK folds in the presence of monovalent ions without Mg^{2+}. Our simulations show that if the divalent ion concentration is increased, more of them are attracted to the PK, which is in quantitative agreement with experimental data. We also capture a more subtle experimental finding that there is a decrease of Γ_{Mg} if the monovalent ion concentration is increased from 54 mM to 79 mM, thereby effectively enhancing the competition with Mg^{2+} binding.

For the rRNA and the riboswitch, the situation is more complicated. Both of them partially unfold at low Mg^{2+} concentrations because tertiary interactions are disrupted. For these 2 RNA molecules, in addition to Mg^{2+} ion concentration, Γ_{Mg} also strongly depends on the state of the RNA. At low Mg^{2+} concentrations, the equilibrium shifts to extended states, which further decreases Γ_{Mg}. Interestingly, our simulations quantitatively reproduce Γ_{Mg} (Fig. 2) over a broad range of monovalent and divalent ion concentrations, for both the rRNA and riboswitch (see below for additional results for rRNA).

Divalent Ion-Dependent Folding Free Energy of BWYV Pseudoknot. In a typical titration experiment, one often measures Γ_X as a function of divalent ion concentrations in the presence of excess monovalent ions ($C_2 \ll C_1$) (41, 67, 70–73). If the RNA remains in a single state S (F, folded; I, intermediate; or U, unfolded) during the titration process, the free energy change due to the accumulation of divalent ions around RNA in state S is directly related to $\Gamma_{X,S}$. For concreteness, consider the following equilibrium reaction $RNA_S + nX^{2+} \rightleftarrows RNA_S.nX^{2+}$, showing that there is an uptake of n (need not be an integer) divalent cations by the RNA in the S state. Provided $C_2 \ll C_1$, the free energy change associated with the equilibrium reaction given above, $\Delta G_{X,S}$, is related to $\Gamma_{X,S}$ as

$$\Delta G_{X,S} = -k_B T \int_o^{C_2} \Gamma_{X,S}\, d \ln C. \qquad [7]$$

Note that Eq. **7** would not be valid if the RNA simultaneously populates different states during the titration process. For instance, if the RNA remains folded at high C_2 but unfolds at low C_2 (as in the rRNA and riboswitch cases), there is no obvious way to relate Γ_X and $\Delta G_{X,S}$, because Γ_X reflects the binding affinity to 2 (or more) states.

To calculate the free energy change upon folding, it is necessary to calculate $\Delta G_{X,S}$ for each state of the RNA separately, which can be done if the ensemble of RNA conformations is restricted to S. We choose BWYV PK for illustrative purposes because its (un)folding can be used to clearly distinguish

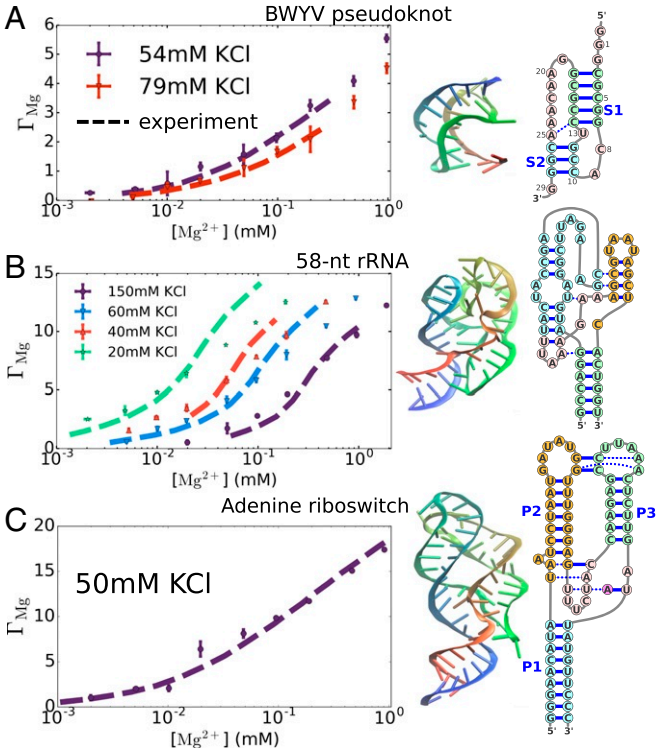

Fig. 2. (*A–C*) Comparison of the calculated and measured values of the preferential interaction coefficient Γ_{Mg} for (*A*) BWYV pseudoknot (64), (*B*) the 58-nt fragment of the large subunit ribosomal RNA (65), and (*C*) the aptamer domain of adenine riboswitch (66) at various monovalent salt concentrations. Experimental data are taken from refs. 67–69. The results of simulations are given by colored symbols with standard errors, and the dashed lines are data from experiments. In most cases, the error bars from the simulations are smaller than the symbol sizes. Only BWYV remains folded at all values of the concentrations of Mg^{2+}. The riboswitch and rRNA partially unfold at low Mg^{2+} concentrations. At high Mg^{2+} concentrations (up to ~1 mM) the RNA molecules are in the folded states, whose structures are shown in *A–C*, *Center*. Nucleotides are colored from 5′ to 3′ as red to blue. Secondary structures are shown in *A–C*, *Right*, displaying the sequences.

between the *F*, *I*, and *U* states. Furthermore, the availability of experimental data allows us to compare directly with our simulations (67, 74).

The folded structure of the BWYV PK (Fig. 2*A*) has 2 Watson–Crick stems (*S1* and *S2*) that are connected by 2 small loops. The *F* state is stabilized by tertiary interactions between the 2 loops and the 2 stems. The stem *S1* has 5 G-C base pairs, while *S2* has only 3 G-C base pairs. Therefore, we anticipate that *S2* should unfold first upon increasing the temperature or lowering the salt concentration, as predicted by the stability hypothesis (75). Both experiments (74) and our previous simulations (31) have shown that BWYV does unfold by 3 sequential equilibrium transitions as temperature is increased, thus populating 2 intermediates. In one of them, there is a loss of tertiary interactions but the stems are intact. However, the probability of formation of such a state is small. Thus, for practical purposes, the overall transition to the unfolded state occurs by populating one intermediate, $F \rightarrow I \rightarrow U$.

To calculate the free energies of the folding and unfolding transitions, we first generated an ensemble of unfolded structures. We performed simulations of the *I* state by disallowing interactions between base pairs in *S2*, while preserving the full interaction for *S1*, as shown in previous studies (31, 74). The ensemble of such structures coincides with what we observed in our thermal unfolding simulations (*SI Appendix*, Fig. S2). We surmise that the simulated ensemble is the one probed in the experiments (67), where all of the nine 3′-terminal nucleotides were mutated to uracil, thus preventing the formation of *S2*. In the *U* state, both the stems are unfolded, which can be mimicked by disrupting all of the specific interactions within the PK. This renders the PK essentially a polyelectrolyte dominated by Coulomb repulsions between the phosphate charges and secondary stackings between consecutive bases.

With the simulated ensemble of *U* and *I* structures, we calculated $\Gamma_{Mg,I}$ and $\Gamma_{Mg,U}$. The results are shown in Fig. 3 at two different KCl concentrations. The uptake of Mg^{2+} in the *I* and *U* states is less compared to the *F* state because they adopt more expanded conformations with spatially separated phosphate groups, thus weakening the electrostatic attraction. The calculations of $\Gamma_{Mg,S}$ are in quantitative agreement with the experimental data. The difference $\Delta\Gamma_{Mg} = \Gamma_{Mg,F} - \Gamma_{Mg,I}$ between the 2 states is the number of Mg^{2+} uptake in the $I \rightarrow F$ transition, which is shown in Fig. 3*B*, *Insets* at 2 monovalent concentrations. We find a slight increase in $\Delta\Gamma_{Mg}$ as the Mg^{2+} concentration increases, in agreement with the direct measurement of $\Delta\Gamma_{Mg}$ from the fluorescence dye method (67). This is in accord with other experiments, which have shown that $\Delta\Gamma$ for monovalent ions also exhibits a dependence on the salt concentration (44, 76). Our calculations, therefore, do not support the Wyman linkage analysis used to determine $\Delta\Gamma_{Mg}$, as this method yields a constant value for $\Delta\Gamma_{Mg}$ over the Mg^{2+} concentration range (67).

We then calculated the free energy changes for all of the states using Eq. **7**, and the results are shown in Fig. 3. Interestingly, Mg^{2+} ions are also localized near the *U* and *I* states, albeit to a lesser extent, demonstrating that it is important to characterize Mg^{2+}–RNA interactions not only in the *F* state but also in other

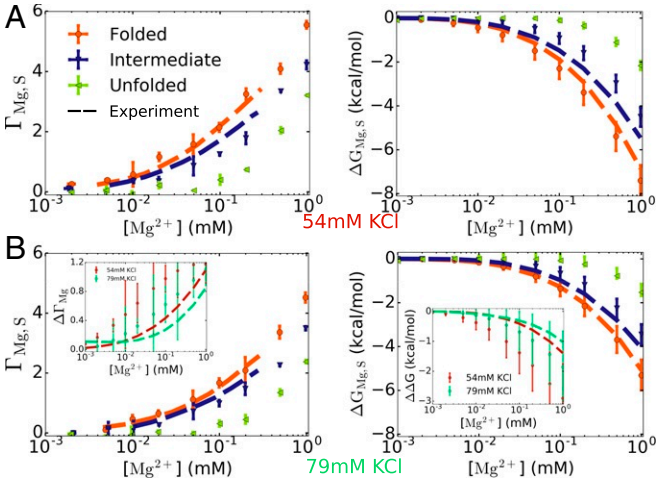

Fig. 3. (*A* and *B*) Ion preferential interaction coefficient $\Gamma_{Mg,S}$ (*Left*) and free energy of Mg^{2+}–pseudoknot interaction $\Delta G_{Mg,S}$ (*Right*) for BWYV in 54 mM KCl (*A*) and 79 mM KCl (*B*), at $T = 25$ °C. The calculations were performed for the folded, intermediate, and unfolded states (*S* = *F*, *I*, or *U*) (see main text and Fig. 4 for definition of the states). Experimental data for the *F* and *I* states are plotted as dashed lines (67). The results for the *U* state serve as a prediction of the simulations. The differences between the *F* and *I* states, $\Delta\Gamma_{Mg} = \Gamma_{Mg,F} - \Gamma_{Mg,I}$ and $\Delta\Delta G_{I-F} = \Delta G_{Mg,F} - \Delta G_{Mg,I}$, are plotted in *B*, *Insets*. $\Delta\Gamma_{Mg}$ (*Left Inset*) represents the number of Mg^{2+} ions released when BWYV transitions from the *F* to the *I* state. $\Delta\Delta G_{I-F}$ (*Right Inset*) is the change in the free energy of the *F* state relative to the *I* state upon addition of Mg^{2+} ions. The error bars in *Insets* are relatively large. However, it is clear that $\Delta\Gamma_{Mg}$ and $\Delta\Delta G_{I-F}$ are not constant in the range of $[Mg^{2+}]$ here.

relevant states along the folding pathway. For example, at 1 mM Mg^{2+} and 54 mM KCl, $\Delta G_{Mg,U} \sim -2$ kcal/mol, $\Delta G_{Mg,I} \sim -4$ kcal/mol, while $\Delta G_{Mg,F} \sim -7$ kcal/mol. One way to quantify the effect of Mg^{2+} addition on the folding process is to compute $\Delta\Delta G$. For instance, for the $I \rightarrow F$ transition, $\Delta\Delta G_{I-F} = \Delta G_{Mg,F} - \Delta G_{Mg,I}$. The stabilization of the F state relative to the I state or U state caused by Mg^{2+} addition ($\Delta\Delta G_{I-F}$ or $\Delta\Delta G_{U-F}$) is therefore ~ -3 and -5 kcal/mol, respectively. The relatively small value of the Mg^{2+} dependence on the folding free energy is likely due to the small size of this PK, whose folded state is stable even in the absence of Mg^{2+} (*SI Appendix*, Figs. S2*B* and S5).

Free Energy of BWYV Pseudoknot Using Thermodynamic Cycle. In Fig. 4, we illustrate how the free energy data shown in Fig. 3 in conjunction with the RNA folding free energy data obtained by varying the temperature could be used to construct a folding free energy diagram for BWYV at 0.2 mM Mg^{2+}. Similar diagrams at arbitrary concentration of Mg^{2+} can be generated. The vertical free energy differences (Fig. 4) are from WHAM analysis of multiple-temperature simulations with or without Mg^{2+} (see *SI Appendix* for additional details). The horizontal free energy differences are Mg^{2+}–RNA free energies as in Fig. 3. The Mg^{2+}-dependent free energy of stabilization is evaluated by 2 ways using the thermodynamic cycle,

$$\Delta\Delta G_{I-F} = \Delta G_{Mg,F} - \Delta G_{Mg,I} = \Delta G_{I-F,Mg} - \Delta G_{I-F} \quad [8]$$
$$\Delta\Delta G_{U-I} = \Delta G_{Mg,I} - \Delta G_{Mg,U} = \Delta G_{U-I,Mg} - \Delta G_{U-I},$$

where $\Delta G_{\alpha-\beta,Mg}$ and $\Delta G_{\alpha-\beta}$ are, respectively, the free energy difference between the α and β states (α, β is F, I, or U) in the presence and absence of Mg^{2+}. The calculated free energies are in remarkable agreement with experiments and are consistent with each other ($\Delta\Delta G$ values estimated by 2 different methods give similar values, with the errors ± 0.4 kcal/mol). Given that they are independently determined, it shows that our theory can be reliably used to study thermodynamics of Mg^{2+}-induced RNA folding.

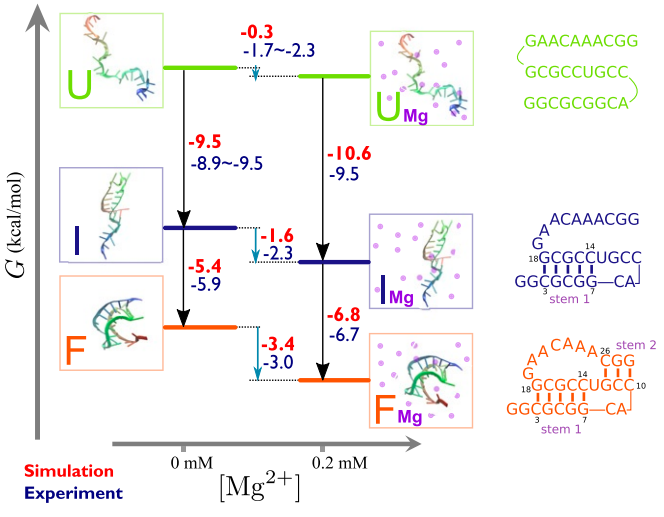

Fig. 4. Folding free energy (kcal/mol, 25 °C) diagram of BWYV pseudoknot at 54 mM KCl in the absence (*Left*) or presence (*Right*) of 0.2 mM Mg^{2+}. Theoretical values are in red and experimental values are in blue. Experimental data for $\Delta G_{U\text{-}I}$ (no Mg^{2+}) and $\Delta G_{Mg,U}$ are not available. For $\Delta G_{U\text{-}I}$, data for 40 mM and 74 mM KCl are reported with the hope that they should bracket the 54 mM KCl data. $\Delta G_{Mg,U}$ for experiment is then evaluated based on the other free energies in the cycle. Folding free energies (black arrows) are calculated from thermal denaturation simulations. Mg^{2+}–RNA free energies (blue arrows) are taken for each state at 0.2 mM Mg^{2+} from Fig. 3.

Mg^{2+}-Induced Folding of the 58-nt rRNA. We investigated the folding of the 58-nt rRNA, which, although folds at high (~1.6 M) monovalent ion concentrations, requires Mg^{2+} (68). We carried out simulations of the 58-nt rRNA as a function of different combinations of monovalent and divalent ion concentrations. Fig. 2*B* already shows that we can quantitatively account for the dependence of Γ_{Mg} as a function of both monovalent and divalent ion concentrations. To fully characterize the folding of the rRNA fragment, we calculated the radius of gyration, R_g, as a function of Mg^{2+} concentration at 4 concentrations of KCl. Fig. 5*A* shows that R_g values decrease continuously from ~25 Å at [Mg^{2+}] ~2 μM to ~16 Å at high Mg^{2+} concentrations, which is consistent with experimental measurements at 40 mM KCl (shown as red bars in Fig. 5) (68). Interestingly, our calculations show that the concentration of KCl has minimal effect on the dependence of R_g on Mg^{2+}, even at low Mg^{2+} concentrations at which the RNA is partially unfolded. This is because at these KCl concentrations, the secondary structure of the rRNA is fully formed (Fig. 5*B*). Therefore, adjusting the monovalent ion concentration does not considerably assist folding since it requires Mg^{2+} for tertiary structure formation. Fig. 5*B* shows the average fraction of native contacts for the rRNA, Q, as a function of Mg^{2+} concentrations. In accord with the R_g analysis, the Q values at low [Mg^{2+}] are small, fluctuating around ~0.35. At these Mg^{2+} concentrations, the secondary structure of the rRNA is completely stabilized without forming tertiary interactions. At high [Mg^{2+}], Q increases and reaches ~0.7. Just as found for R_g, we also observe a negligible dependence of Q on monovalent ion concentration. Thus, both the order parameters (R_g and Q) show that rRNA formation, which does not depend much on KCl concentrations, is only moderately cooperative.

Free Energy Changes upon Folding. To obtain the folding thermodynamics of rRNA we calculated $\Delta G_{Mg,S}$ for each state of the 58-nt rRNA by assuming that the folding transition occurs sequentially, $U \rightarrow I \rightarrow F$, where the I state is composed only of secondary structure of 3 stems connected by small loops (Fig. 6). It is possible that for such a complex RNA, more than one intermediate state could be populated during the (un)folding process. However, we chose only one intermediate state to separate the effect of divalent cations on secondary and tertiary interaction formation. The transition $U \rightarrow I$ involves only secondary structure formation, while $I \rightarrow F$ requires the formation of only tertiary interactions. As before, we performed simulations for each state by constraining the ensemble of RNA structures in such state. We emphasize that these simulations are completely different from the simulations above (from which we calculated Γ_{Mg}, R_g, and Q), which were performed using all of the sampled conformations. The constrained simulations were used only to calculate $\Gamma_{Mg,S}$, Mg^{2+}–RNA free energies, $\Delta G_{Mg,S}$, and $\Delta\Delta G_{\alpha-\beta}$ (where S, α, and β are F, I, and U).

In Fig. 6, we show the results for the stabilization free energies of the folding transitions upon addition of Mg^{2+} ions, $\Delta\Delta G_{\alpha-\beta}$, calculated using Eq. **8**. Data for $\Gamma_{Mg,S}$ and $\Delta G_{Mg,S}$, which were used to compute $\Delta\Delta G_{\alpha-\beta}$, can be found in *SI Appendix*, Fig. S7. $\Delta\Delta G_{\alpha-\beta}$ shows the change of the relative stability of the 2 states α and β upon addition of Mg^{2+} on the $\alpha \leftrightarrow \beta$ transition. We explicitly show only $\Delta\Delta G_{U-I}$ and $\Delta\Delta G_{U-F}$ curves, but

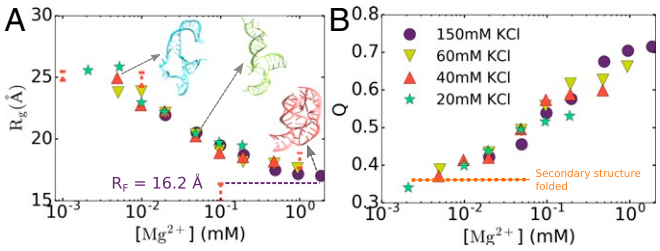

Fig. 5. Mg^{2+}-induced folding of the 58-nt fragment of ribosomal RNA. (*A*) Compaction of rRNA as the Mg^{2+} concentration increases. The colors correspond to 4 monovalent ion concentrations (shown in *B*). R_F is the value of R_g for the *F* state, calculated from the PDB structure (PDB ID 1HC8). Experimental measurements of R_g in 40 mM KCl are plotted as vertical red error bars (68). Some representative structures from the simulations are also shown. (*B*) Average fraction of native contacts vs. Mg^{2+} concentration. The horizontal dashed line at around $Q \approx 0.35$ indicates complete secondary structure formation with no tertiary interactions.

$\Delta\Delta G_{I-F}$ can be calculated as the difference between these 2 curves, as $\Delta\Delta G_{I-F} = \Delta\Delta G_{U-F} - \Delta\Delta G_{U-I}$. It is obvious that the higher the [Mg^{2+}] is, larger is the effect of Mg^{2+} ions on all 3 transitions since all $\Delta\Delta G$ values decrease (increase in magnitude) as [Mg^{2+}] rises. Therefore, higher [Mg^{2+}] induces a shift in the equilibrium toward more compact states ($U \rightarrow I \rightarrow F$). The magnitude of $\Delta\Delta G$ for rRNA is quite large compared to BWYV, indicating the dramatic dependence of the rRNA folding on Mg^{2+}. For rRNA in 60 mM KCl, adding 0.1 mM Mg^{2+} leads to $\Delta\Delta G_{I-F} \approx -8$ kcal/mol and $\Delta\Delta G_{U-F} \approx -13$ kcal/mol. In comparison, for BWYV at 54 mM KCl, those values are $\Delta\Delta G_{I-F} \approx -1.4$ kcal/mol and $\Delta\Delta G_{U-F} \approx -2.2$ kcal/mol, respectively. On the other hand, if one instead increases the concentration of monovalent ions, the values of $\Delta\Delta G$ become smaller (Fig. 6), which also happens in the BWYV PK (shown in Fig. 3*B*, *Insets*). At 150 mM KCl, adding 0.1 mM Mg^{2+} into the solution of rRNA leads only to $\Delta\Delta G_{I-F} \approx -4$ kcal/mol and $\Delta\Delta G_{U-F} \approx -6$ kcal/mol, respectively.

Comparison between Mg^{2+} and Ca^{2+} Ions. We also studied the effect of the cation size (Mg^{2+} vs. Ca^{2+}) on RNA folding. We computed the Ca^{2+}–P effective potential, $V_{\text{Ca-P}}(r)$, using the same procedure used to obtain $V_{\text{Mg-P}}(r)$. Fig. 7*A* compares the effective potentials of the 2 ions. A major difference between $V_{\text{Ca-P}}(r)$ and $V_{\text{Mg-P}}(r)$ is that the barrier separating the inner-shell and outer-shell binding in Ca^{2+} is much lower than in Mg^{2+}. This difference arises because the charge density of Mg^{2+} is much higher than that of Ca^{2+}, which results in orders of magnitude difference in the water exchange kinetics between Mg^{2+} and Ca^{2+} (77). It is also in accord with the observation that the interaction of Mg^{2+} with water in the first hydration shell is stronger than in Ca^{2+} (*SI Appendix*, Fig. S10).

Fig. 7*B* shows the radial distribution function, $g_{\text{X-P}}(r)$, between the divalent ions and phosphate groups in BWYV PK. It is obvious that our model accounts for both types of binding. The presence of inner- and outer-shell binding of Mg^{2+} to P is indicated by 2 visible peaks in $g_{\text{Mg-P}}(r)$. The first peak is located at $r \sim 2.4$ Å and the second one is at $r \sim 4.8$ Å. The slow decay of $g_{\text{Mg-P}}(r)$ toward 1 is due to the presence of other phosphates in the RNA. The peaks for Ca^{2+} are very similar and are shifted slightly upward toward longer distance (2.7 Å and 5.1 Å), indicating that Ca^{2+} has a comparable affinity for the phosphate groups.

We also investigated the folding of the rRNA in the presence of Ca^{2+} ions. Due to the similar affinity of the 2 ions toward phosphate groups, there is little difference in the folding behavior of rRNA between Ca^{2+} and Mg^{2+} in terms of ion accumulation, fraction of native contacts, and global size (Fig. 7 *C* and *D*). It is possible that the difference between the 2 ions is apparent only in the case of more complex RNAs, such as group I intron ribozyme (10), where the folded state is highly compact and there is not sufficient room in the core of the RNA to accommodate larger ions, and therefore replacing Mg^{2+} by Ca^{2+} in these cases would destabilize the folded state. Nonetheless, we find an interesting difference in the nature of binding of the 2 ions: Ca^{2+} dehydrates readily due to its lower charge density and binds the phosphate groups directly in the inner shell, while Mg^{2+} coordination, with a higher charge density, is roughly similar between the inner and outer shells. A more detailed study will be reported in a subsequent publication.

In addition, we also show in Fig. 7 data for Mg^{2+} assuming that the Mg^{2+}–P interaction is given by the DH potential (green curves). In the $g_{\text{Mg-P}}(r)$ plot (Fig. 7*B*), the DH potential completely misses the second peak and the first peak is also much lower compared to our model. This leads to lower affinity with the phosphate groups, resulting in fewer ions accumulated around the RNA. When applying the DH potential to rRNA folding, we find that the RNA does not fold, but rather adopts much more extended conformations, which is directly related to the small uptake of ions, at all Mg^{2+} concentrations (Fig. 7 *C* and *D*). Our model thus reveals the importance of treating Mg^{2+}–P interaction accurately to faithfully capture both structural and thermodynamic features of Mg^{2+}-assisted RNA folding.

Importance of Outer-Sphere Mg^{2+}-P Coordination. One of the key predictions of this work is that accurate predictions of RNA folding thermodynamics require a consistent description of both the inner- and outer-sphere coordination of divalent cations to phosphate groups. To assess the importance of the outer-sphere coordination, we created a potential that retains the inner-shell interaction between Mg^{2+}–P, while smoothly joining the outer-shell interaction with the DH potential (Fig. 8*A*). In so doing all of the outer-shell interactions between Mg^{2+}–P are eliminated. The barrier between inner sphere and outer sphere in the modified potential is lower than in the original potential. We believe that the smaller barrier would only alter the kinetics of water and ion exchange around Mg^{2+} and should not affect the folding thermodynamics quantities. The outer-sphere coordination interaction is only weakened ~0.2 kcal/mol in the modified potential.

The simulations using the modified potential reveal a large impact on Mg^{2+}–RNA interactions and RNA folding. The

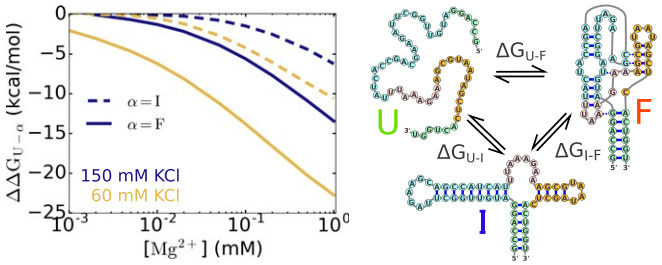

Fig. 6. Stabilization free energies, $\Delta\Delta G$ (defined in Eq. **8**), of the intermediate state (dashed lines) or the folded state (solid lines) relative to the unfolded state upon addition of Mg^{2+} ions for rRNA. $\Delta\Delta G_{I-F}$ is the difference between these 2 free energies $\Delta\Delta G_{I-F} = \Delta\Delta G_{U-F} - \Delta\Delta G_{U-I}$. In the *I* state, secondary structures are formed. Therefore, $\Delta\Delta G_{U-I}$ and $\Delta\Delta G_{I-F}$ are, respectively, the stabilization free energies of Mg^{2+} on secondary and tertiary structure formation.

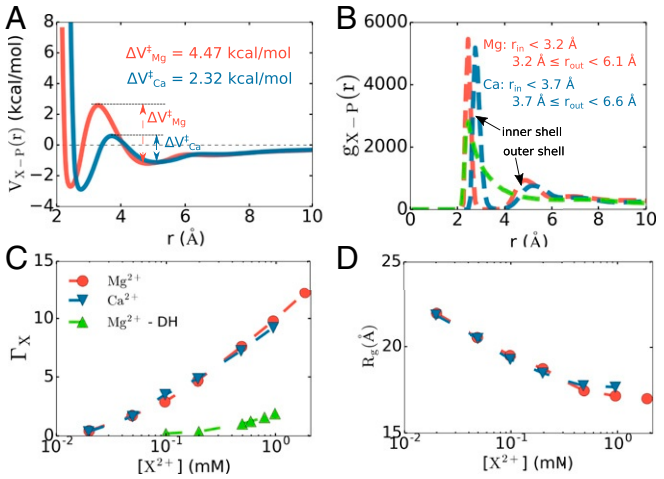

Fig. 7. Comparison between Mg^{2+} and Ca^{2+} binding to phosphate groups and RNA. (*A*) Effective potentials for Mg^{2+} and Ca^{2+} with the phosphate group show that the transition barrier between inner-shell–outer-shell coordination for Ca^{2+} is considerably lower than for Mg^{2+}. However, the depths of the inner- and outer-shell minima are comparable. (*B*) Radial distribution function between the ions and phosphate groups in BWYV. Also shown are the results for Mg^{2+} ions in which Mg^{2+}–phosphate interactions are modeled using the Debye–Hückel potential (green). (*C* and *D*) Preferential interaction coefficient Γ_X (*C*) and radius of gyration R_g (*D*) computed for 58-nt rRNA. Only data for 150 mM KCl are presented here; data for other KCl concentrations can be found in *SI Appendix*, Fig. S6. The rRNA reaches its native states at high divalent ion concentrations using our model, but does not easily fold at any Mg^{2+} concentration should the Debye–Hückel potential be used.

substantial changes are illustrated in Fig. 8*B*, comparing the values of Γ_{Mg} for BWYV in 54 mM KCl. A seemingly small softening of the outer-sphere coordination interaction leads to a significant decrease in Γ_{Mg}. This shows, rather vividly, that the Mg^{2+}–RNA interaction, and therefore the Mg^{2+}–RNA free energy, is extremely sensitive to the Mg^{2+}–P interaction. In rRNA, the situation is even more pronounced (Fig. 8*C*). Because the folding of the rRNA depends dramatically on Mg^{2+}, weakening the outer-sphere coordination interaction modestly prevents the folding of rRNA even at high $[Mg^{2+}]$ due to an insufficient number of condensed Mg^{2+}. Interestingly, the use of the modified potential results in insignificant compaction of rRNA even at the highest Mg^{2+} concentration (Fig. 8*C*, *Lower*). Surprisingly, there is negligible difference in the predictions for R_g between the predictions using the DH potential and the modified potential. The results in Fig. 8 *B* and *C* show that accurate predictions for RNA folding thermodynamics require accounting for both inner- and outer-sphere coordination of Mg^{2+} with phosphates.

Discussion

In this study, we have introduced a method to capture the impact of a solution containing a mixture of monovalent and divalent cations on RNA folding. In our model, based in part on liquid-state theory, divalent cations are treated explicitly while the screening effect of monovalent salt is treated implicitly. We reasoned that since the DH theory works well at long range (in fact, the DH theory gives asymptotically correct results at long distances), we could improve the divalent cation interactions by using theory to describe the short-ranged interactions. To obtain the short-ranged interactions, we used RISM theory to compute the PMF between the divalent cation–phosphate group and combined it with the long-ranged part of the DH potential to obtain the effective potential $V_{X\text{-}P}(r)$. Applications to 3 RNA molecules, with different sequences and structures, illustrate that our theory quantitatively reproduces the Mg^{2+} preferential interaction coefficients and Mg^{2+}–RNA free energies for different combinations of ion concentrations. The transition free energies, as the RNA traverses along the folding pathway, are also in remarkable agreement with experimental data. The simulations not only reproduce the thermodynamics of divalent cation binding to RNA but also recapitulate the correlation between divalent cation concentration and RNA folding. In addition, we presented the effects of divalent cations on secondary and tertiary structure formations for an RNA construct and its dependence on monovalent concentrations.

The difference in the divalent cation–RNA interactions between our model and the DH theory occurs only in the short-ranged part, $r_{X-P} \leq 10$ Å (little difference exists from $r \sim 6-10$ Å in Fig. 1). Nonetheless, such a difference proves to be very crucial in obtaining the correct divalent cation binding free energies, as it translates to at least ∼2.0 kcal/mol deviation of $\Delta G_{Mg,F}$ for BWYV (*SI Appendix*, Fig. S9). We also show that to obtain the correct divalent ion–RNA binding free energies, it is important to take into account both inner- and outer-shell interaction accurately. A deviation of only 0.2 kcal/mol in the divalent cation–P potential could lead to a substantial decrease in the number of bound ions. In complex RNAs whose folding depends on divalent ions, loss of bound ions could even cause the RNAs to be thermodynamically unstable even at elevated ion concentrations. It is also worth stating that since our model treats monovalent ions implicitly, it has a large advantage over fully explicit ion models in terms of simulation performance and could be used to study much larger RNA molecules including RNA–protein interactions. Indeed, our theory is sufficiently general that it can be applied to calculate ion (with arbitrary valence and size) effects on DNA as well as synthetic polyelectrolytes and polyampholytes.

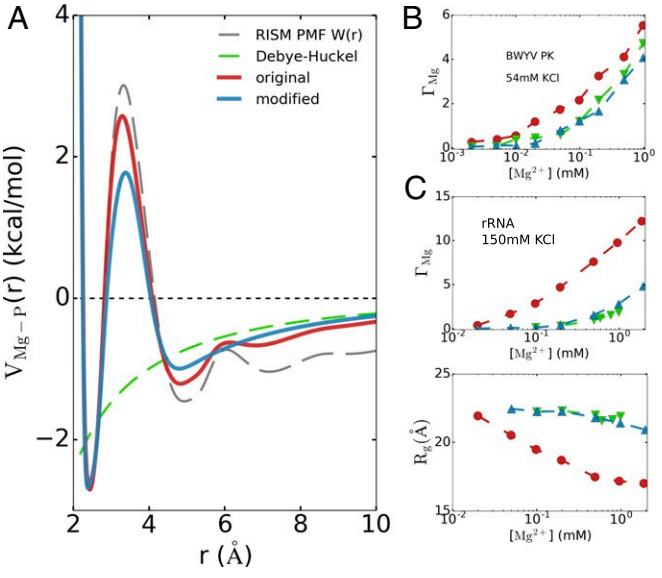

Fig. 8. Importance of treating both inner- and outer-sphere coordination. (*A*) Modified effective potential for Mg^{2+} with softened outer-sphere coordination interaction, while keeping the inner-sphere coordination interaction intact. (*B* and *C*) Comparison of (*B*) preferential interaction coefficient Γ_{Mg} for BWYV PK at 54 mM KCl and (*C*) Γ_{Mg} and R_g for 58-nt rRNA at 150 mM KCl using the modified potential. Despite a tiny difference in the outer-sphere coordination (∼0.2 kcal/mol), the values of Γ_{Mg} decrease significantly from those of the original model. In rRNA, the modified potential even cannot fold the rRNA at relevant Mg^{2+} concentrations, as can be seen in the high values of R_g.

Although no other existing computational model can be used to calculate ion-dependent folding thermodynamic properties of RNA of arbitrary size and sequence accurately, as we have done here, our theory is not without limitations. For example, we considered only the divalent ion–phosphate interaction and neglected interactions with the bases, which might be important as the size of RNA molecules increases. It is suspected that the Mg^{2+} ion interacts with electronegative atoms in the base moiety (both inner shell and outer shell) (21, 22). However, it is unknown whether such interactions are relevant for RNA folding thermodynamics. In addition, our RNA force field includes nonnative interactions only in a limited manner. Nevertheless, the remarkable agreement between the theoretical predictions and experiments opens entirely different ways to quantitatively probe ion-induced folding of RNAs regardless of their sizes and sequences.

Conclusions

We have proposed a theory of divalent ion–phosphate interactions, based on concepts in liquid-state physics, for use in coarse-grained simulations of RNA folding in the presence of explicit divalent cations while the screening effect of monovalent salt is treated implicitly. Because our model accounts for both the inner- and outer-sphere coordination of divalent cations with the RNAs using RISM, the theory quantitatively reproduces divalent cation-dependent free energies for folding transitions and the correlation between the divalent cation binding and RNA folding. The success of our general method, which integrates liquid-state theories and a coarse-grained TIS model for RNA, is widely applicable to a variety of problems in RNA biology in which divalent cations play an important role. Finally, the theory could also be used to treat the effects of spherical and nonspherical ions on the conformations of RNA as well as DNA.

Materials and Methods

Full details of RISM theory, the calculation of divalent cation–phosphate potential, the RNA coarse-grained force field, and simulation details are provided in *SI Appendix*.

ACKNOWLEDGMENTS. We are grateful to Tom Record for insightful discussions. H.T.N. thanks Natalia Denesyuk for providing the source code for the previous model, and Mauro Mugnai and Debayan Chakraborty for several fruitful discussions. This work was supported by National Science Foundation Grant CHE 19-00093 and the Welch Foundation Grant F-0019 through the Collie–Welch chair. We are thankful to the Texas Advanced Computing Center for providing computational resources.

1. L. Z. Sun, D. Zhang, S. J. Chen, Theory and modeling of RNA structure and interactions with metal ions and small molecules. *Annu. Rev. Biophys.* **46**, 227–246 (2017).
2. N. A. Denesyuk, D. Thirumalai, Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B* **117**, 4901–4911 (2013).
3. N. A. Denesyuk, N. Hori, D. Thirumalai, Molecular simulations of ion effects on the thermodynamics of RNA folding. *J. Phys. Chem. B* **122**, 11860–11867 (2018).
4. Y. Zhu, S. J. Chen, Many-body effect in ion binding to RNA. *J. Chem. Phys.* **141**, 055101 (2014).
5. D. E. Draper, A guide to ions and RNA structure. *RNA* **10**, 335–343 (2004).
6. S. A. Woodson, Metal ions and RNA folding: A highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.* **9**, 104–109 (2005).
7. J. C. Bowman, T. K. Lenz, N. V. Hud, L. D. Williams, Cations in charge: Magnesium ions in RNA folding and catalysis. *Curr. Opin. Struct. Biol.* **22**, 262–272 (2012).
8. J. Lipfert, S. Doniach, R. Das, D. Herschlag, Understanding nucleic acid-ion interactions. *Annu. Rev. Biochem.* **83**, 813–841 (2014).
9. S. L. Heilman-Miller, D. Thirumalai, S. A. Woodson, Role of counterion condensation in folding of the tetrahymena ribozyme. I. Equilibrium stabilization by cations. *J. Mol. Biol.* **306**, 1157–1166 (2001).
10. N. A. Denesyuk, D. Thirumalai, How do metal ions direct ribozyme folding?*Nat. Chem.* **7**, 793–801 (2015).
11. L. Z. Sun, S. J. Chen, Predicting RNA-metal ion binding with ion dehydration effects. *Biophys. J.* **116**, 184–195 (2019).
12. X. W. Fang, T. Pan, T. R. Sosnick, Mg^{2+}-dependent folding of a large ribozyme without kinetic traps. *Nat. Struct. Mol. Biol.* **6**, 1091–1095 (1999).
13. J. Pan, D. Thirumalai, S. A. Woodson, Magnesium-dependent folding of self-splicing RNA: Exploring the link between cooperativity, thermodynamics, and kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6149–6154 (1999).
14. K. L. Buchmueller, A. E. Webb, D. A. Richardson, K. M. Weeks, A collapsed non-native RNA folding state. *Nat. Struct. Mol. Biol.* **7**, 362–366 (2000).
15. R. Russell, I. S. Millett, S. Doniach, D. Herschlag, Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nat. Struct. Mol. Biol.* **7**, 367–370 (2000).
16. D. Thirumalai, N. Lee, S. A. Woodson, D. Klimov, Early events in RNA folding. *Annu. Rev. Phys. Chem.* **52**, 751–762 (2001).
17. E. Koculi, C. Hyeon, D. Thirumalai, S. A. Woodson, Charge density of divalent metal cations determines RNA stability. *J. Am. Chem. Soc.* **129**, 2676–2682 (2007).
18. J. Noeske, H. Schwalbe, J. Wohnert, Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain. *Nucleic Acids Res.* **35**, 5262–5273 (2007).
19. R. K. O. Sigel, H. Sigel, A stability concept for metal ion coordination to single-stranded nucleic acids and affinities of individual sites. *Acc. Chem. Res.* **43**, 974–984 (2010).
20. G. C. Wong, L. Pollack, Electrostatics of strongly charged biological polymers: Ion-mediated interactions and self-organization in nucleic acids and proteins. *Annu. Rev. Phys. Chem.* **61**, 171–189 (2010).
21. H. Zheng, I. G. Shabalin, K. B. Handing, J. M. Bujnicki, W. Minor, Magnesium-binding architectures in RNA crystal structures: Validation, binding preferences, classification and motif detection. *Nucleic Acids Res.* **43**, 3789–3801 (2015).
22. F. Leonarski, L. D'Ascenzo, P. Auffinger, Mg^{2+} ions: Do they bind to nucleobase nitrogens?*Nucleic Acids Res.* **45**, 987–1004 (2017).
23. H. T. Nguyen, S. A. Pabit, L. Pollack, D. A. Case, Extracting water and ion distributions from solution X-ray scattering experiments. *J. Chem. Phys.* **144**, 214105 (2016).
24. C. Hyeon, D. Thirumalai, Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 6789–6794 (2005).
25. S. M. Gopal, S. Mukherjee, Y. M. Cheng, M. Feig, PRIMO/PRIMONA: A coarse-grained model for proteins and nucleic acids that preserves near-atomistic accuracy. *Proteins* **78**, 1266–1281 (2010).
26. S. Pasquali, P. Derreumaux, HiRE-rna: A high resolution coarse-grained energy model for RNA. *J. Phys. Chem. B* **114**, 11957–11966 (2010).
27. S. Cao, S. J. Chen, Physics-based de novo prediction of RNA 3d structures. *J. Phys. Chem. B* **115**, 4216–4226 (2011).
28. P. Sulc, F. Romano, T. E. Ouldridge, J. P. K. Doye, A. A. Louis, A nucleotide-level coarse-grained model of RNA. *J. Chem. Phys.* **140**, 235102 (2014).
29. A. M. Mustoe, H. M. Al-Hashimi, C. L. Brooks, Coarse grained models reveal essential contributions of topological constraints to the conformational free energy of RNA bulges. *J. Phys. Chem. B* **118**, 2615–2627 (2014).
30. M. J. Boniecki *et al.*, SimRNA: A coarse-grained method for RNA folding simulations and 3d structure prediction. *Nucleic Acids Res.* **44**, e63 (2016).
31. N. Hori, N. A. Denesyuk, D. Thirumalai, Salt effects on the thermodynamics of a frameshifting RNA pseudoknot under tension. *J. Mol. Biol.* **428**, 2847–2859 (2016).
32. D. R. Bell, S. Y. Cheng, H. Salazar, P. Ren, Capturing RNA folding free energy with coarse-grained molecular dynamics simulations. *Sci. Rep.* **7**, 45812 (2017).
33. N. Hori, N. A. Denesyuk, D. Thirumalai, Ion condensation onto ribozyme is site-specific and fold-dependent. *Biophys. J.* **116**, 2400–2410 (2019).
34. D. Chandler, H. C. Andersen, Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. *J. Chem. Phys.* **57**, 1930–1937 (1972).
35. F. Hirata, P. Rossky, An extended RISM equation for molecular polar fluids. *Chem. Phys. Lett.* **83**, 329–334 (1981).
36. F. Hirata, B. Pettitt, P. Rossky, Application of an extended RISM equation to dipolar and quadrupolar fluids. *J. Chem. Phys.* **77**, 509–520 (1982).
37. B. M. Pettitt, P. J. Rossky, Integral equation predictions of liquid state structure for waterlike intermolecular potentials. *J. Chem. Phys.* **77**, 1451–1457 (1982).
38. F. Hirata, "Theory of molecular liquids" in *Molecular Theory of Solvation, Understanding Chemical Reactivity*, F. Hirata, Ed. (Springer Netherlands, 2004), vol. 24, pp. 1–60.
39. T. Luchko, I. S. Joung, D. A. Case, "Chapter 4: Integral equation theory of biomolecules and electrolytes" in *Innovations in Biomolecular Modeling and Simulations* (The Royal Society of Chemistry, 2012), vol. 1, pp. 51–86.
40. E. L. Ratkova, D. S. Palmer, M. V. Fedorov, Solvation thermodynamics of organic molecules by the molecular integral equation theory: Approaching chemical accuracy. *Chem. Rev.* **115**, 6312–6356 (2015).
41. D. Grilley, A. M. Soto, D. E. Draper, Mg^{2+}–RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14003–14008 (2006).
42. F. Oosawa, *Polyelectrolytes* (M. Dekker, New York, NY, ed. 1, 1971).
43. G. S. Manning, The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **11**, 179–246 (1978).
44. M. T. Record, C. F. Anderson, T. M. Lohman, Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **11**, 103–178 (1978).
45. R. W. Wilson, D. C. Rau, V. A. Bloomfield, Comparison of polyelectrolyte theories of the binding of cations to DNA. *Biophys. J.* **30**, 317–325 (1980).
46. A. Savelyev, G. A. Papoian, Molecular renormalization group coarse-graining of electrolyte solutions: Application to aqueous NaCl and KCl. *J. Phys. Chem. B* **113**, 7785–7793 (2009).
47. J. W. Shen, C. Li, N. F. van der Vegt, C. Peter, Transferability of coarse grained potentials: Implicit solvent models for hydrated ions. *J. Chem. Theory Comput.* **7**, 1916–1927 (2011).

48. Z. Cao, J. F. Dama, L. Lu, G. A. Voth, Solvent free ionic solution models from multiscale coarse-graining. *J. Chem. Theory Comput.* **9**, 172–178 (2013).
49. D. M. Hinckley, J. J. de Pablo, Coarse-grained ions for nucleic acid modeling. *J. Chem. Theory Comput.* **11**, 5436–5446 (2015).
50. J. M. Rodgers, J. D. Weeks, Local molecular field theory for the treatment of electrostatics. *J. Phys. Condens. Matter* **20**, 494206 (2008).
51. R. C. Remsing, S. Liu, J. D. Weeks, Long-ranged contributions to solvation free energies from theory and short-ranged models. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2819–2826 (2016).
52. O. Allner, L. Nilsson, A. Villa, Magnesium ion-water coordination and exchange in biomolecular simulations. *J. Chem. Theory Comput.* **8**, 1493–1502 (2012).
53. A. Saxena, D. Sept, Multisite ion models that improve coordination and free energy calculations in molecular dynamics simulations. *J. Chem. Theory Comput.* **9**, 3538–3542 (2013).
54. F. Duarte *et al.*, Force field independent metal parameters using a nonbonded dummy model. *J. Phys. Chem. B* **118**, 4351–4362 (2014).
55. M. T. Panteva, G. M. Giambasu, D. M. York, Comparison of structural, thermodynamic, kinetic and mass transport properties of Mg^{2+} ion models commonly used in biomolecular simulations. *J. Comput. Chem.* **36**, 970–982 (2015).
56. R. L. Hayes *et al.*, Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. *J. Am. Chem. Soc.* **134**, 12043–12053 (2012).
57. J. A. Lemkul, S. K. Lakkaraju, A. D. MacKerell, Characterization of Mg^{2+} distributions around RNA in solution. *ACS Omega* **1**, 680–688 (2016).
58. S. Kirmizialtin, S. Pabit, S. Meisburger, L. Pollack, R. Elber, RNA and its ionic cloud: Solution scattering experiments and atomically detailed simulations. *Biophys. J.* **102**, 819–828 (2012).
59. R. L. Hayes *et al.*, Generalized Manning condensation model captures the RNA ion atmosphere. *Phys. Rev. Lett.* **114**, 258105 (2015).
60. G. Giambasu, T. Luchko, D. Herschlag, D. York, D. Case, Ion counting from explicit-solvent simulations and 3d-RISM. *Biophys. J.* **106**, 883–894 (2014).
61. H. T. Nguyen, S. A. Pabit, S. P. Meisburger, L. Pollack, D. A. Case, Accurate small and wide angle X-ray scattering profiles from atomic models of proteins and nucleic acids. *J. Chem. Phys.* **141**, 22D508 (2014).
62. J. M. Schurr, D. P. Rangel, S. R. Aragon, A contribution to the theory of preferential interaction coefficients. *Biophys. J.* **89**, 2258–2276 (2005).
63. P. E. Smith, Equilibrium dialysis data and the relationships between preferential interaction parameters for biological systems in terms of Kirkwood-buff integrals. *J. Phys. Chem. B* **110**, 2862–2868 (2006).
64. M. Egli, G. Minasov, L. Su, A. Rich, Metal ions and flexibility in a viral RNA pseudoknot at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4302–4307 (2002).
65. G. L. Conn, A. G. Gittis, E. E. Lattman, V. K. Misra, D. E. Draper, A compact RNA tertiary structure contains a buried backbone–K^+ complex. *J. Mol. Biol.* **318**, 963–973 (2002).
66. A. Serganov *et al.*, Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11**, 1729–1741 (2004).
67. A. M. Soto, V. Misra, D. E. Draper, Tertiary structure of an RNA pseudoknot is stabilized by "diffuse" Mg^{2+} ions. *Biochemistry* **46**, 2973–2983 (2007).
68. D. Grilley, V. Misra, G. Caliskan, D. E. Draper, Importance of partially unfolded conformations for Mg^{2+}-induced folding of RNA tertiary structure: Structural models and free energies of Mg^{2+} interactions. *Biochemistry* **46**, 10266–10278 (2007).
69. D. Leipply, D. E. Draper, Effects of Mg^{2+} on the free energy landscape for folding a purine riboswitch RNA. *Biochemistry* **50**, 2790–2799 (2011).
70. Y. Bai *et al.*, Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.* **129**, 14981–14988 (2007).
71. C. V. Bizarro, A. Alemany, F. Ritort, Non-specific binding of Na^+ and Mg^{2+} to RNA determined by force spectroscopy methods. *Nucleic Acids Res.* **40**, 6922–6935 (2012).
72. A. Dittmore, J. Landy, A. A. Molzon, O. A. Saleh, Single-molecule methods for ligand counting: Linking ion uptake to DNA hairpin folding. *J. Am. Chem. Soc.* **136**, 5974–5980 (2014).
73. S. A. Pabit *et al.*, Counting ions around DNA with anomalous small-angle X-ray scattering. *J. Am. Chem. Soc.* **132**, 16334–16336 (2010).
74. P. L. Nixon, D. P. Giedroc, Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J. Mol. Biol.* **296**, 659–671 (2000).
75. S. S. Cho, D. L. Pincus, D. Thirumalai, Assembly mechanisms of RNA pseudoknots are determined by the stabilities of constituent secondary structures. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17349–17354 (2009).
76. D. R. Jacobson, O. A. Saleh, Quantifying the ion atmosphere of unfolded, single-stranded nucleic acids using equilibrium dialysis and single-molecule methods. *Nucleic Acids Res.* **44**, 3763–3771 (2016).
77. Y. Lee, D. Thirumalai, C. Hyeon, Ultrasensitivity of water exchange kinetics to the size of metal ion. *J. Am. Chem. Soc.* **139**, 12334–12337 (2017).