UC Irvine
UC Irvine Previously Published Works

Title

Structure-Based Model of RNA Pseudoknot Captures Magnesium-Dependent Folding Thermodynamics

Permalink

https://escholarship.org/uc/item/6mv141b5

Journal

The Journal of Physical Chemistry B, 123(7)

ISSN

1520-6106

Authors

Mandic, Ana
Hayes, Ryan L
Lammert, Heiko
et al.

Publication Date

2019-02-21

DOI

10.1021/acs.jpcb.8b10791

Supplemental Material

https://escholarship.org/uc/item/6mv141b5#supplemental

Copyright Information

Peer reviewed

Structure-Based Model of RNA Pseudoknot Captures Magnesium-Dependent Folding Thermodynamics

Published as part of The Journal of Physical Chemistry virtual special issue "Deciphering Molecular Complexity in Dynamics and Kinetics from the Single Molecule to the Single Cell Level".

Ana Mandic,[†] Ryan L. Hayes,[*,‡] Heiko Lammert,[†] Ryan R. Cheng,[*,†] and José N. Onuchic[*,†]

[†]Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States
[‡]Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States

Ⓢ *Supporting Information*

ABSTRACT: We develop a simple, coarse-grained approach for simulating the folding of the Beet Western Yellow Virus (BWYV) pseudoknot toward the goal of creating a transferable model that can be used to study other small RNA molecules. This approach combines a structure-based model (SBM) of RNA with an electrostatic scheme that has previously been shown to correctly reproduce ionic condensation in the native basin. Mg^{2+} ions are represented explicitly, directly incorporating ion−ion correlations into the system, and K^+ is represented implicitly, through the mean-field generalized Manning counterion condensation theory. Combining the electrostatic scheme with a SBM enables the electrostatic scheme to be tested beyond the native basin. We calibrate the SBM to reproduce experimental BWYV unfolding data by eliminating overstabilizing backbone interactions from the molecular contact map and by strengthening base pairing and stacking contacts relative to other native contacts, consistent with the experimental observation that relative helical stabilities are central determinants of the RNA unfolding sequence. We find that this approach quantitatively captures the Mg^{2+} dependence of the folding temperature and generates intermediate states that better approximate those revealed by experiment. Finally, we examine how our model captures Mg^{2+} condensation about the BWYV pseudoknot and a U-tail variant, for which the nine 3′ end nucleotides are replaced with uracils, and find our results to be consistent with experimental condensation measurements. This approach can be easily transferred to other RNA molecules by eliminating and strengthening the same classes of contacts in the SBM and including generalized Manning counterion condensation.

1. INTRODUCTION

Ionic conditions have a strong effect on RNA stability and dynamics.[1−3] This dependence arises because RNA has a negative charge for every nucleotide, and in order for RNA to fold into compact structures that perform biological function, the high density of negative charge must be offset by positive counterions. Because of the high charge density of RNA, Manning condensation[4] renders the Debye−Hückel approximation and other linear Poisson−Boltzmann approaches[5] inadequate. Even in low ion concentrations, counterions will condense onto the RNA to mitigate the high charge density.[4] These condensed counterions can lead to effective attractive interactions between phosphates.[6] Accounting for electrostatic effects is necessary for a quantitative understanding of RNA folding.

RNA tertiary structure frequently depends on Mg^{2+} in order to fold.[1] Mg^{2+} associates strongly with RNA because of its small size and divalent charge.[7−9] While Mg^{2+} ions can bind to specific chelation sites,[10,11] most associate nonspecifically.[2,12,13] The amount of excess Mg^{2+} that associates with a structural ensemble

(Γ^{2+}) as a function of Mg^{2+} concentration (c_{2+}) is directly related to the Mg^{2+}−RNA interaction free energy by

$$\beta \Delta G_{Mg^{2+}-RNA} = -\int_0^{c_{2+}} \Gamma_{2+} \, d(\ln c'_{2+}) \tag{1}$$

where $\beta = (k_B T)^{-1}$. The excess magnesium (Γ^{2+}) reflects the difference in the number of magnesium ions associated with the RNA with respect to the number of magnesium ions occupying a similar volume in the bulk. The relationship between the free energy and Γ^{2+} was previously derived in ref 14 using relations derived for describing how changing the ligand binding activity of a biomolecule affected the free energy difference between two biomolecular conformations.[15] Differences in this free energy between structural ensembles quantitatively describe the degree to which Mg^{2+} stabilizes particular ensembles or drives folding.[14]

Molecular dynamics simulations can provide valuable insights into the atomic details of biological processes. Simulations that
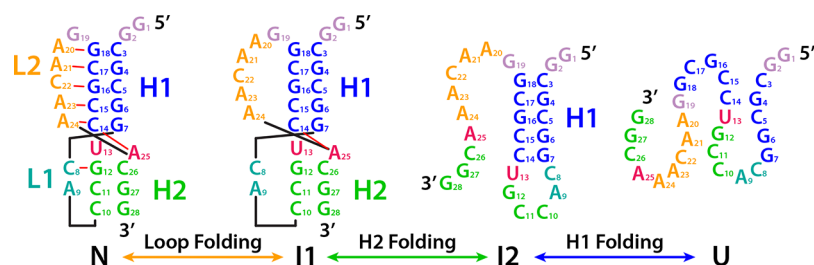
Figure 1. Cartoon of the unfolding order for the BWYV RNA pseudoknot. The folded BWYV pseudoknot is composed of four tertiary structure elements: helices H1 and H2 as well as loops L1 and L2. The relative order of unfolding was previously determined experimentally[39] (left to right) as starting with the loss of the L1 and L2 contacts, followed by H2, and finally the breaking of H1.

explicitly represent solvation by water and ions have contributed much to our understanding of RNA electrostatics.[13,16–19] While these simulations are fairly accurate, they are also quite slow, and biologically relevant processes such as folding are too slow to observe in such simulations. In contrast, coarse-grained simulations can access much longer time scales but can suffer from lower accuracy.[20–23] Structure-based models (SBMs) are an ideal coarse-grained approach to describe RNA folding. SBMs are based on energy landscape theory and the principle of minimal frustration, which together require that, on average, native contacts must be more energetically favorable than nonnative contacts for biopolymers to fold on biological time scales.[24–27] SBMs were originally conceived in the context of protein folding, but have been successfully applied to RNA as well.[28–32]

With a few notable exceptions,[33,34] coarse-grained models typically omit electrostatics or include the effects of ions through a Debye−Hückel potential,[22,23] which neglects important condensation effects. Recently, we introduced a coarse-grained model of RNA electrostatics[35] that describes the ion condensation around RNA with quantitative accuracy while remaining computationally inexpensive compared to explicit solvent molecular dynamics approaches. This model explicitly treats Mg^{2+} ions, capturing their important ion−ion correlations, while describing monovalent ions implicitly using a mean-field generalized Manning ion condensation model. This RNA electrostatics model was previously combined with a SBM of RNA to explore the native state dynamics of a pseudoknot. Here, we extend upon our earlier approach to explore the accuracy of our model beyond the folded basin in the context of RNA folding.

Energetic parameters of SBMs are well tuned for proteins[36,37] but still need calibration for RNA. In order to calibrate our model, we focus on the Beet Western Yellow Virus (BWYV) pseudoknot. BWYV pseudoknot folding has been carefully characterized experimentally,[38,39] showing the pseudoknot to disassemble piece by piece with increasing temperature through the two intermediates shown in Figure 1. Electrostatic and ionic effects have also been thoroughly measured for the BWYV pseudoknot. The Mg^{2+} and K^+ dependence of the transition temperatures between each intermediate has been measured directly.[38,39] Mg^{2+} condensation around the pseudoknot and its U-tail variant has been quantified through titration and fluorescence measurement techniques[14,39,40] for a range of ionic concentrations. In the U-tail variant, the nine 3′ nucleotides of the pseudoknot were mutated to uracil to mimic the partially unfolded ensemble.[39] The difference in Mg^{2+} condensation between the pseudoknot and U-tail can be used with eq 1 to obtain the Mg^{2+} dependence of the difference in stability between the folded and partially unfolded states. These

experiments provide a thorough characterization of folding and electrostatics for this RNA with which to constrain potential models.

In this paper, we begin by removing native contacts that overstabilize the backbone of the RNA. We then calibrate the energetics of the SBM to reflect the importance of relative helical stabilities in the folding of RNA. Finally, we combine our SBM with our electrostatic model to explore the ionic dependence of folding, both in the case of BWYV and in its 3′ end variant, U-tail BWYV, whose uracil-replaced tail segment inhibits the formation of tertiary structure and allows for the exploration of the extent to which the tail contributes to net stability in the molecular model. Our work is directly compared with published experimental findings throughout the Results and Discussion section.

2. METHODS

Simulations were performed using an all heavy atom SBM.[36] Such models have been used in RNA[28,29,31] but have typically used a 4 Å cutoff distance to define native contacts. Contact maps for this study were generated using the Shadow method, which includes additional contacts up to 6 Å but discards contacts occluded by an intervening atom.[41] We chose the Shadow contact map over the cutoff contact map because the shadow contact map automatically achieves the correct energetic balance between base pairing and base stacking contacts, while base stacking contacts must be scaled by 1/3 with a 4 Å cutoff contact map.[41]

Given a crystal structure of a native RNA fold, we define a structure-based potential as

$$V_{SBM} = \frac{\varepsilon_r}{2} \sum_i^{bonds} (r_i - r_{0i})^2 + \frac{\varepsilon_\theta}{2} \sum_i^{angles} (\theta_i - \theta_{0i})^2$$
$$+ \frac{\varepsilon_\chi}{2} \sum_i^{impropers\ or\ planar} (\chi_i - \chi_{0i})^2 + \varepsilon_{NC} \sum_{i,j}^{RNA-RNA\ noncontacts} \left(\frac{\sigma_{NC}}{r_{ij}}\right)^{12}$$
$$+ \varepsilon_\phi \sum_i^{proper\ dihedrals} F_D(\phi_i - \phi_{0i}) + \varepsilon_C \sum_{i,j}^{contacts} \left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right)$$

(2)

where

$$F_D(\phi) = (1 - \cos\phi) + \frac{1}{2}(1 - \cos 3\phi)$$

(3)

The structural parameters r_{0i}, θ_{0i}, χ_{0i}, σ_{ij}, and ϕ_{0i} are determined by their values in the crystal structure. The remaining energetic parameters have been previously calibrated and described.[35] The run input files were obtained using
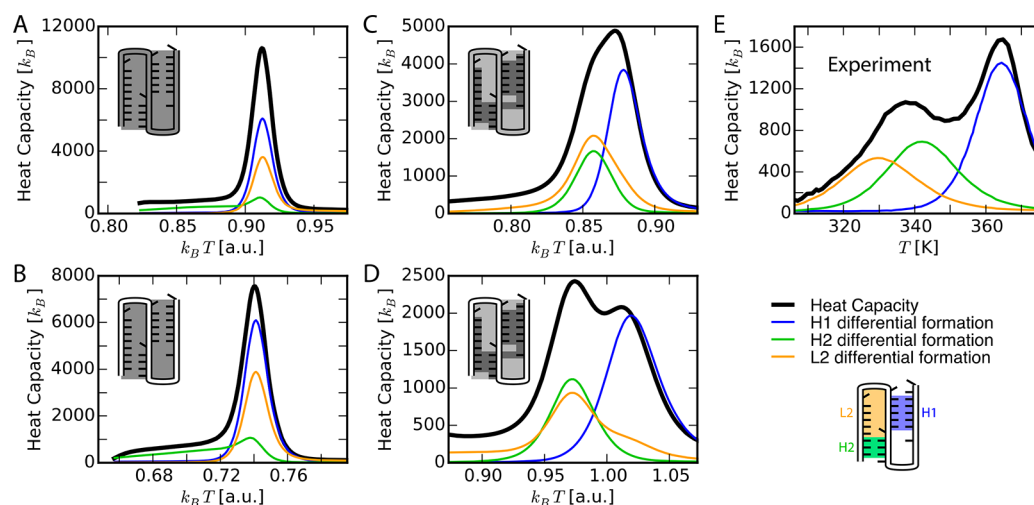
Figure 2. Effect of reducing backbone overstabilization and increasing relative stabilities of helicies on the heat capacity. (A−E) Result for all combinations of contact maps and BPS parametrizations. The simulated heat capacity is shown for (A) the full contact map SBM, (B) SBM with a reduced contact map (removal of contacts that lead to backbone overstabilization), (C) SBM with a full contact map and BPS contacts scaled with respect to all other native contacts by a ratio of 4:1, and (D) SBM with a reduced contact map and BPS contacts scaled with respect to all other native contacts by a ratio of 4:1. (E) As a reference, the experimental heat capacity is shown (ref 39). As can be observed in the first column (A,B), the exclusion of local contacts delivers a reduction in cooperativity as overenforcement of backbone rigidity is removed. When the system's nonbonding helical contact potentials are then rescaled (shown in the lower right-hand panel), a discretization of regional folding occurs. The H2 and L2 segments of the BWYV pseudoknot begin to unfold at a lower temperature than H1, which approximates with greater fidelity the assembly sequence documented to occur in previous studies[39] and is represented in Figure 1.

SMOG[42,43] and subsequently postprocessed to reweight the energies of various groups of contacts during model calibration.

SMOG produces SBMs that use reduced units that are described in the Supporting Information of ref 44. In this work, the SBM temperature unit T_0 is determined by matching the experimental unfolding temperature of the H1 helix in 1 mM $MgCl_2$ and 54 mM monovalent salt with the unfolding temperature of the calibrated SBM to give $99.4T_0 = 300$ K.

The U-tail construct of the BWYV pseudoknot[39] was constructed by replacing the nine nucleotides at the 3′ end of BWYV with uracil to prevent those residues from forming a tertiary structure. This was originally created[39] to mimic the I2 intermediate state of the BWYV pseudoknot (Figure 1). Structural parameters for the U-tail construct for eq 2 were obtained from the online nucleic acid builder server (http://structure.usc.edu/make-na/server.html).

Manning condensation plays a strong role in stabilizing nucleic acid structures.[4,6] The model we use to describe condensation and ionic strength dependence of RNA stability has been described in detail elsewhere.[35] In the model, Mg^{2+} ions are represented as an explicit particle with a charge of +2, and each phosphate is given a charge of −1. (In addition, the γ phosphate of the triphosphate tail is given a −2 charge, and the protonated N3 in C8 is given a + 1 charge.) Monovalent ions K^+ and $Cl^−$ are represented implicitly as condensed and screening ions. Condensed monovalent ions are represented by two Gaussian distributions centered on each phosphate; one Gaussian represents the condensation of monovalent ions about RNA, while the second Gaussian represents a "hole" by enforcing that no ions can get too close to the RNA. All electrostatic interactions in the model are described using Debye−Hückel interactions, including interactions between two screened phosphates as well as between a screened phosphate and a point charge (Mg^2). A more detailed description of this model can be found in ref 35. Because phosphates in RNA are mostly solvent exposed,[12] the dielectric constant of water

appropriate to the simulation temperature (eq 6 of ref 45) is used.

To explore the folding of our BYWV pseudoknot and U-tail models, we performed Langevin dynamics simulations. Simulations to calibrate the SBM before the addition of electrostatics were performed using the GROMACS molecular dynamics software package,[46] while simulations with the electrostatic model were performed using publicly available software that we have developed.[35] The free energy landscape and other thermodynamic characterizations of RNA folding were obtained from simulation trajectories using the Weighted Histogram Analysis Method (WHAM).[47]

For additional details on the model, refer to the Supporting Information text.

3. RESULTS AND DISCUSSION

The objective of this study was to develop an RNA-specific structure-based modeling technique to guide the general modeling of small RNAs. SBMs have found wide application in the study of protein systems, but their use with RNA remains in its infancy. In our study, we emphasize the need for SBM design to provide due consideration to capturing the hierarchical unfolding mechanism documented to be characteristic of RNA,[48] and address how backbone overstabilization inhibits the formation of key intermediates observed in experiment. Then we turn our attention to incorporating the effects of counterions on folding by testing a successful, recently developed, electrostatic scheme.[35] Our simulated results are compared with experimental findings.[38,39]

3.1. Contact Map: The Key to Capturing Structural Transitions. We start by examining the uncalibrated SBM in the absence of electrostatics (Figure 2, panel A). In this system, we observe a highly cooperative folding transition characterized by a single, narrow peak in the calculated heat capacity. The second loop (L2) and major helices (H1 and H2) disassemble simultaneously, a result in direct contrast with the sequential

unfolding mechanism that has been established experimentally for the BWYV pseudoknot[39] (Figure 1).

Observing these inadequacies, two changes in the SBM were introduced that markedly improved its ability to represent the thermodynamic and structural characteristics of folding. Both of these modifications redistributed the energetic weighing scheme of the contact map in a manner that accounted for the key physical realities of backbone flexibility and hierarchical folding that in the past have often been overlooked in the interest of simplicity.

The first adjustment is the removal of local backbone interactions from the contact map. Local interactions between residues i and j within an $|i − j| \leq 3$ window are excluded from protein SBMs but were retained in RNA due to the importance of base stacking interactions. In this work, interactions between atoms on nucleotides i and j within an $|i − j| \leq 2$ window are excluded if the atoms are not involved in base pairing or stacking (BPS). This approach ensures that local backbone contacts are omitted without compromising the formation of a helical structure. Indeed, in early simulations, even weakening base stacking relative to base pairing resulted in severely nonphysical ensembles (see the SI); therefore, base stacking contacts must be preserved.

The motivation behind this design change is that the influence of contact potentials between nearby backbone atoms is redundant and leads to double-counting of local geometric terms. Already heavily constrained by strong, harmonic, geometric terms (eq 1), the inclusion of local Lennard-Jones interactions is unnecessary for ensuring the proper configuration of secondary structure as the molecule folds into its native state. Indeed, when we eliminate them completely in Figure 2B, there is little discernible change in the ability of the SBM to assemble into its native conformation.

A direct consequence of double-counting due to the inclusion of local contacts is a rigidification of the backbone that contributes to an inaccurate sequence of intermediates as well as to the residual structure in the unfolded ensemble. Local contacts have an overstabilizing effect in the loop regions that contributes to the nonphysical cooperativity between H1, H2, and L2. Though a local contacts exclusion scheme does not by itself succeed at decoupling H2 from L2, it does weaken the coupling between H1 and L2, as seen in the widening of the peak going from Figure 2A to B and the increased space between peaks going from Figure 4C to D. Furthermore, backbone rigidity leads to nonphysical kinks of the unfolded ensemble. Calculations of phosphate pseudodihedral entropy paint an unmistakable picture: as local contact interactions are eliminated, the configurational distribution of the pseudoknot becomes more uniform (Supporting Information Figure S1). Because local contacts are at best nonessential and at worst impede the accurate reproduction of intermediates, they are subsequently omitted.

The second modification identified BPS contacts and rescaled them to a ratio of 4:1 with respect to all other contacts. We did not calibrate the ratio between base pairing and base stacking contacts as this ratio had already been calibrated previously for a 4 Å cutoff contact map[28] as well as for a reduced shadow map,[41,42] which removed the overcounting of native contacts. The later use of the shadow map found that a ratio of 1:1 between base pairing and base stacking contacts yields the same balance between the total base pairing energy and the total base stacking energy.

In addition to these changes in strength aligning intuitively with the well-established reality that helical interactions are indeed the strongest in nucleic acids,[48] they are implemented to capture the hierarchical unfolding mechanism documented to be characteristic of small RNA molecules. Previous studies have established that the relative stability of helical structures is predictive of the order of assembly on the path toward the native state,[23] and rescaling the BPS group places greater energetic emphasis on the relative stability of the two helices. In this study, rescaling the BPS contacts in the simple SBM substantially improved the ordering of intermediates and automatically reveals that H1 is more stable than H2 (Figure 2C).

The 4:1 ratio was chosen as the product of conservative tuning of Lennard-Jones parameters in the system to disrupt the synchronicity of H1, H2, and L2 without excessively apportioning the available energy of the model to the BPS group contacts. Figure 3 shows the range of ratios that were tested and their effect on unfolding behavior. The 4:1 ratio in panel C was chosen because it is the smallest ratio that meets the goal of disrupting H1, H2, and L2 cooperativity and reproducing the dual peak in the heat capacity.

Perhaps the most striking result is seen in the combined effect of incorporating consideration of effective double-counting, nonhomogeneous interaction strengths, and hierarchical assembly in the design of an RNA SBM (Figure 2D). Results show large improvement in both the structural transitions and thermodynamic profile of the folding process. Newly emergent are two highly resolved peaks in the heat capacity, which are notably absent if either contact map adjustment is omitted.

3.2. Pseudoknot Electrostatics. RNA folding is highly coupled to its interaction with Mg^{2+} ions,[1,7−9] which act to stabilize the native and intermediate conformations. In this section, we apply the electrostatic model to our calibrated SBM with a reduced contact map and BPS contacts scaled 4:1 with respect to other native contacts. With the model calibrated to reflect a better balance between native contact strengths in different regions of the pseudoknot, we proceed to conduct an evaluation of how well ionic effects can be incorporated. The following results are obtained from molecular dynamics simulations that combine both the new contact map and the successful electrostatic scheme produced in past research[35] and summarized in the Introduction.

Our first observation is that the inclusion of electrostatics drives the simulated heat capacity of unfolding to more closely resemble the experimental curve. Experiment (Figure 2E) places the second (lower-temperature) transition peak at 7.3% below the BWYV pseudoknot's folding temperature (higher-temperature peak), an attribute well-approximated by the calibrated SBM with electrostatics, for which it is 7.1%. For comparison, the calibrated SBM without electrostatics had a distance of 3.7%.

We next examine the robustness of our calibrated SBM with electrostatics as the bulk ionic composition is varied, comparing our simulations directly with extensive experimental measurements.[39] Figure 4B shows the dependence of the folding temperature for tertiary elements H1, H2, and L2 on the bulk Mg^{2+} concentration for simulation and experiment. While the addition of charge interactions acts as a perturbation that reduces overall cooperativity, it nonetheless results in the retention of the tendency of regions H2 and L2 to unfold simultaneously, even as the composition of the solvent is varied. Experimenting with ionic parameters further, when K^+ is varied over concentrations spanning 50 mM−1 M (Figure 4C), H2 and L2 continue to share a melting point.
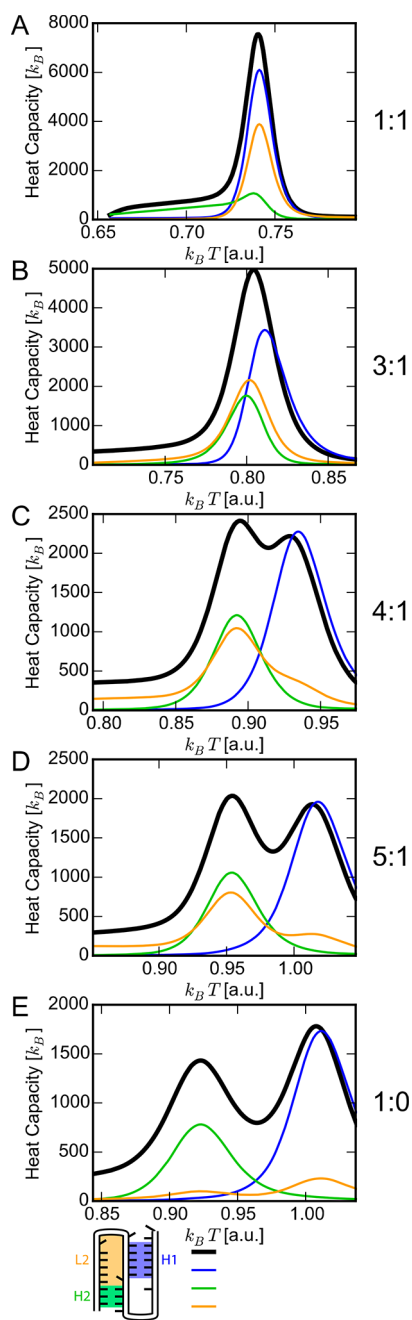
Figure 3. Effect of scaling BPS contacts relative to other native contacts on the heat capacity. The simulated heat capacity of our SBM is plotted for ratios of the BPS contact strengths with respect to other native contacts of (A) 1:1, (B) 3:1, (C) 4:1, (D) 5:1, and (E) 1:0. As the relative strength of BPS contacts is increased to reflect the importance of helical stabilities, the heat capacity begins to resemble the double-peaked heat capacity observed in experiment for BWYV[39] (shown in Figure 2E).

The overly high cooperativity of the H2 and L2 regions over a wide range of Mg^{2+} concentrations suggests that electrostatics in our model is not the origin of this high cooperativity. The synchronization of H2 and L2 plausibly arises from the idealized energy landscape of a SBM, which lacks energetic frustration. For example, U13-A25 could form a canonical base pair that would extend H2 from three residues to four residues and would disrupt the L2 structure (Figure 1). G19 also kinks very sharply back against H2 and is necessary for L2 formation, and G19 base

stacking interactions with G18 in H1 could compete with L2 formation. These plausible sources of energetic frustration are currently absent in our idealized SBM.

The extent to which Mg^{2+} ions condense onto the BWYV model and neutralize phosphate−phosphate repulsions is shown in Figure 5. At a fixed monovalent concentration of 54 mM K^+, the simulated excess Mg^{2+} associated with the RNA (Γ^{2+}) agrees well with experimentally measured values over a range of concentrations (Figure 5A). Calculations of Mg^{2+}−RNA interaction free energies (eq 1) show that the BWYV model overstabilizes by 3% compared with experiment, and when the 3′ end tail is replaced with uracils, we continue to see a reasonable (if slightly diminished) degree of accuracy, with an under-stabilization of 20%. Exploring further, when the net KCl concentration is then adjusted to 79 mM (Figure 5B), the models of the pseudoknot and its U-tail variant understabilize by 21 and 37%, respectively. We propose that the source of this understabilization in our simulations of the U-tail arises from all tertiary structural contacts for the nine nucleotides at the 3′ end that are simply removed in our model. It would be expected that this reduction in degree of compactness would lead to a significant, systematic decrease in net Mg^{2+} condensation and a poorer fit to experimental data.

4. CONCLUSIONS

In this work, we developed a coarse-grained, ion dependent SBM of the BWYV pseudoknot that reproduced experimental folding pathways and ionic effects. While this model was calibrated with experimental data, the modifications were motivated by physical intuition. We decreased the cooperativity of the model by removing overstabilizing backbone interactions and strengthened base pairing and stacking contacts relative to other native contacts to reflect the importance of helical stabilities in RNA folding. Consequently, these changes can be applied to SBM of other RNA molecules and should give improved folding behavior for those RNA molecules as well.

Two physically motivated changes were made to the initial shadow contact map that improved agreement with experimental thermal unfolding profiles. First, local backbone contacts were found to cause rigid kinks in the unfolded ensemble. While local contacts are excluded in proteins, these contacts were historically included in RNA due to the overwhelming importance of base stacking in RNA folding. We discriminate between critical base stacking contacts and nonphysical local backbone contacts by retaining all base−base contacts and excluding all backbone contacts between residues i and $i − 2$ to $i + 2$. Second, RNA folding is dominated by base pairing and stacking interactions, and we find that stabilizing these contacts relative to other contacts breaks cooperativity between H1 and H2 folding and improves agreement with experimental unfolding profiles. This modification also is likely to more generally reproduce the well-known hierarchical folding of RNA, where secondary structure folds first, followed by tertiary structure. Future attempts to model RNA should consider these two effects.

While we capture the broad unfolding behavior of the BWYV pseudoknot, we note that in our model L2 and H2 still fold cooperatively; thus, we are unable to capture the experimental I1 intermediate. SBMs neglect frustration due to non-native interactions; therefore, the presence of non-native interactions in this intermediate is a likely explanation.

Adding the generalized Manning counterion condensation model to this calibrated SBM enabled us to capture several ionic
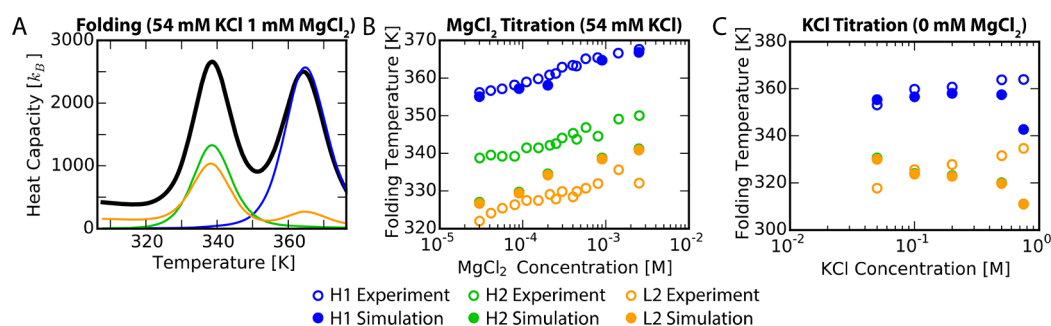
Figure 4. Dependence of the folding temperature on salt concentration. (A) The heat capacity of the simulated SBM is shown is shown at physiological salt concentration. (B) The dependence of the folding temperature for the H1, H2, and L2 structural elements on the Mg^{2+} concentration is shown in comparison to experiment (the experimental folding temperature is estimated from published data;[39] see the Supporting Information for more details). (C) The dependence of the folding temperature for the H1, H2, and L2 structural elements on the K^+ concentration is shown in comparison to experimental measurements.[38]
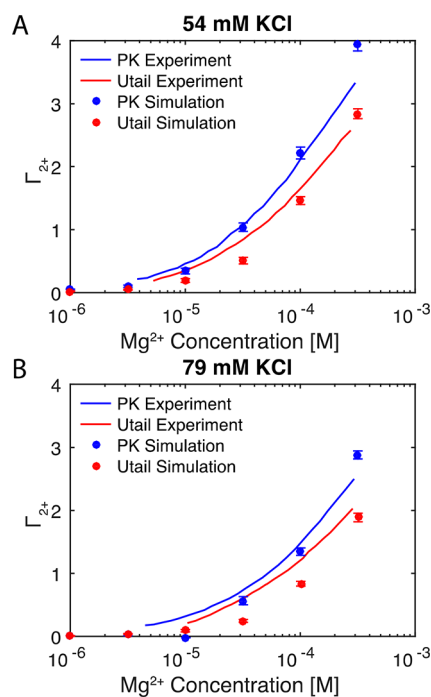

Figure 5. Excess Mg^{2+} associated with the ensemble of RNA structures plotted as a function of the bulk Mg^{2+} concentration for fixed monovalent ion concentrations of (A) 54 and (B) 79 mM KCl. The plots are shown for the simulated BWYV and U-tail variants in comparison with the experimental measurements[39] for those constructs.

effects in the BWYV. Previously, this electrostatic model had only been used within the folded basin. This work is the first to test its performance beyond the folded basin for RNA folding, and we find it to perform well. The Mg^{2+} dependence of the folding temperatures is captured almost perfectly, while K^+ dependence highlights the need for further improvement in the electrostatic model. Mg^{2+} condensation profiles also show reasonable agreement with experiment both within the folded basin and in the U-tail variant that models the I2 intermediate.

ASSOCIATED CONTENT

⊙ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b10791.

Detailed methods section and supporting figure showing the pseudodihedrals that reveal residual structure in an unfolded ensemble due to local backbone contacts (PDF)

AUTHOR INFORMATION

Corresponding Authors
*E-mail: rhayes@pointloma.edu (R.L.H.).
*E-mail: jonuchic@rice.edu (J.N.O.).
*E-mail: ryan.cheng@rice.edu (R.R.C.) .

ORCID
Ryan R. Cheng: 0000-0001-6378-295X
José N. Onuchic: 0000-0002-9448-0388

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Draper, D. E. A guide to ions and RNA structure. *RNA* **2004**, *10* (3), 335−343.

(2) Draper, D. E.; Grilley, D.; Soto, A. M. Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 221−243.

(3) Heilman-Miller, S. L.; Pan, J.; Thirumalai, D.; Woodson, S. A. Role of counterion condensation in folding of the Tetrahymena ribozyme II. Counterion-dependence of folding kinetics. *J. Mol. Biol.* **2001**, *309* (1), 57−68.

(4) Manning, G. S. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **1978**, *11* (2), 179−246.

(5) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (18), 10037−10041.

(6) Manning, G. S. Electrostatic free energy of the DNA double helix in counterion condensation theory. *Biophys. Chem.* **2002**, *101*, 461−473.

(7) Thirumalai, D.; Hyeon, C. *Non-Protein Coding RNAs*; Springer-Verlag: Berlin, Heidelberg; 2009; pp XI, 398.

(8) Heilman-Miller, S. L.; Thirumalai, D.; Woodson, S. A. Role of counterion condensation in folding of the Tetrahymena ribozyme. I.

Equilibrium stabilization by cations. *J. Mol. Biol.* **2001**, *306* (5), 1157−1166.

(9) Koculi, E.; Hyeon, C.; Thirumalai, D.; Woodson, S. A. Charge density of divalent metal cations determines RNA stability. *J. Am. Chem. Soc.* **2007**, *129* (9), 2676−2682.

(10) Klein, D. J.; Moore, P. B.; Steitz, T. A. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **2004**, *10* (9), 1366−1379.

(11) Leipply, D.; Draper, D. E. Evidence for a Thermodynamically Distinct Mg2+ Ion Associated with Formation of an RNA Tertiary Structure. *J. Am. Chem. Soc.* **2011**, *133* (34), 13397−13405.

(12) Draper, D. E. Folding of RNA Tertiary Structure: Linkages between Backbone Phosphates, Ions, and Water. *Biopolymers* **2013**, *99* (12), 1105−1113.

(13) Hayes, R. L.; Noel, J. K.; Mohanty, U.; Whitford, P. C.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y. Magnesium Fluctuations Modulate RNA Dynamics in the SAM-I Riboswitch. *J. Am. Chem. Soc.* **2012**, *134* (29), 12043−12053.

(14) Grilley, D.; Soto, A. M.; Draper, D. E. Mg2+-RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (38), 14003−14008.

(15) Wyman, J. Linked Functions and Reciprocal Effects in Hemoglobin - a 2nd Look. *Adv. Protein Chem.* **1964**, *19*, 223−286.

(16) Chen, A. A.; Draper, D. E.; Pappu, R. V. Molecular Simulation Studies of Monovalent Counterion-Mediated Interactions in a Model RNA Kissing Loop. *J. Mol. Biol.* **2009**, *390* (4), 805−819.

(17) Allner, O.; Nilsson, L.; Villa, A. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8* (4), 1493−1502.

(18) Kim, T.; Shapiro, B. A. The role of salt concentration and magnesium binding in HIV-1 subtype-A and subtype-B kissing loop monomer structures. *J. Biomol. Struct. Dyn.* **2013**, *31* (5), 495−510.

(19) Panteva, M. T.; Giambasu, G. M.; York, D. M. Force Field for Mg2+, Mn2+, Zn2+, and Cd2+ Ions That Have Balanced Interactions with Nucleic Acids. *J. Phys. Chem. B* **2015**, *119* (50), 15460−15470.

(20) Mustoe, A. M.; Al-Hashimi, H. M.; Brooks, C. L., III Coarse Grained Models Reveal Essential Contributions of Topological Constraints to the Conformational Free Energy of RNA Bulges. *J. Phys. Chem. B* **2014**, *118* (10), 2615−2627.

(21) Mustoe, A. M.; Brooks, C. L., III; Al-Hashimi, H. M. Topological constraints are major determinants of tRNA tertiary structure and dynamics and provide basis for tertiary folding cooperativity. *Nucleic Acids Res.* **2014**, *42* (18), 11792−11804.

(22) Hyeon, C.; Thirumalai, D. Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6789−6794.

(23) Cho, S. S.; Pincus, D. L.; Thirumalai, D. Assembly mechanisms of RNA pseudoknots are determined by the stabilities of constituent secondary structures. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (41), 17349−17354.

(24) Bryngelson, J. D.; Wolynes, P. G. Spin-Glasses and the Statistical-Mechanics of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (21), 7524−7528.

(25) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein Folding Funnels - a Kinetic Approach to the Sequence Structure Relationship. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (18), 8721−8725.

(26) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of protein folding: The energy landscape perspective. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545−600.

(27) Onuchic, J. N.; Wolynes, P. G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 70−75.

(28) Whitford, P. C.; Schug, A.; Saunders, J.; Hennelly, S. P.; Onuchic, J. N.; Sanbonmatsu, K. Y. Nonlocal Helix Formation Is Key to Understanding S-Adenosylmethionine-1 Riboswitch Function. *Biophys. J.* **2009**, *96* (2), L7−L9.

(29) Whitford, P. C.; Geggier, P.; Altman, R. B.; Blanchard, S. C.; Onuchic, J. N.; Sanbonmatsu, K. Y. Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. *RNA* **2010**, *16* (6), 1196−1204.

(30) Suddala, K. C.; Rinaldi, A. J.; Feng, J.; Mustoe, A. M.; Eichhorn, C. D.; Liberman, J. A.; Wedekind, J. E.; Al-Hashimi, H. M.; Brooks, C. L., III; Walter, N. G. Single transcriptional and translational preQ1 riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure. *Nucleic Acids Res.* **2013**, *41* (22), 10462−75.

(31) Feng, J.; Walter, N. G.; Brooks, C. L., III Cooperative and Directional Folding of the preQ(1) Riboswitch Aptamer Domain. *J. Am. Chem. Soc.* **2011**, *133* (12), 4196−4199.

(32) Hori, N.; Takada, S. Coarse-Grained Structure-Based Model for RNA-Protein Complexes Developed by Fluctuation Matching. *J. Chem. Theory Comput.* **2012**, *8* (9), 3384−3394.

(33) Mak, C. H.; Henke, P. S. Ions and RNAs: Free Energies of Counterion-Mediated RNA Fold Stabilities. *J. Chem. Theory Comput.* **2013**, *9* (1), 621−639.

(34) Hinckley, D. M.; de Pablo, J. J. Coarse-Grained Ions for Nucleic Acid Modeling. *J. Chem. Theory Comput.* **2015**, *11* (11), 5436−5446.

(35) Hayes, R. L.; Noel, J. K.; Mandic, A.; Whitford, P. C.; Sanbonmatsu, K. Y.; Mohanty, U.; Onuchic, J. N. Generalized Manning Condensation Model Captures the RNA Ion Atmosphere. *Phys. Rev. Lett.* **2015**, DOI: 10.1103/PhysRevLett.114.258105.

(36) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K. Y.; Onuchic, J. N. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. *Proteins: Struct., Funct., Genet.* **2009**, *75* (2), 430−441.

(37) Karanicolas, J.; Brooks, C. L., III The origins of asymmetry in the folding transition states of protein L and protein G. *Protein Sci.* **2002**, *11* (10), 2351−2361.

(38) Nixon, P. L.; Giedroc, D. P. Energetics of a strongly pH dependent RNA tertiary structure in a frameshifting pseudoknot. *J. Mol. Biol.* **2000**, *296* (2), 659−671.

(39) Soto, A. M.; Misra, V.; Draper, D. E. Tertiary structure of an RNA pseudoknot is stabilized by "diffuse" Mg2+ ions. *Biochemistry* **2007**, *46* (11), 2973−2983.

(40) Grilley, D.; Soto, A. M.; Draper, D. E. Chapter 3 Direct Quantitation of Mg2+-RNA Interactions by Use of a Fluorescent Dye. *Methods in Enzymology*; Academic Press, 2009; Vol. 455, pp 71−94.

(41) Noel, J. K.; Whitford, P. C.; Onuchic, J. N. The Shadow Map: A General Contact Definition for Capturing the Dynamics of Biomolecular Folding and Function. *J. Phys. Chem. B* **2012**, *116* (29), 8692−8702.

(42) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nucleic Acids Res.* **2010**, *38*, W657−W661.

(43) Noel, J. K.; Levi, M.; Raghunathan, M.; Lammert, H.; Hayes, R. L.; Onuchic, J. N.; Whitford, P. C. SMOG 2: A Versatile Software Package for Generating Structure-Based Models. *PLoS Comput. Biol.* **2016**, *12* (3), e1004794.

(44) Hayes, R. L.; Noel, J. K.; Whitford, P. C.; Mohanty, U.; Sanbonmatsu, K. Y.; Onuchic, J. N. Reduced Model Captures Mg2+ -RNA Interaction Free Energy of Riboswitches. *Biophys. J.* **2014**, *106* (7), 1508−1519.

(45) Malmberg, C. G.; Maryott, A. A. Dielectric Constant of Water from 0-Degrees-C to 100-Degrees-C. *J. Res. Nat. Bur Stand* **1956**, *56* (1), 1−8.

(46) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4* (3), 435−447.

(47) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules 0.1. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011−1021.

(48) Brion, P.; Westhof, E. Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 113−137.