Biophysical Journal Volume 109 December 2015 2654–2665

Article

Predicting 3D Structure, Flexibility, and Stability of RNA Hairpins in Monovalent and Divalent Ion Solutions

Ya-Zhou Shi,[1] Lei Jin,[1] Feng-Hua Wang,[2] Xiao-Long Zhu,[3] and Zhi-Jie Tan[1,*]
[1]Department of Physics and Key Laboratory of Artificial Micro- and Nano-structures of the Ministry of Education, School of Physics and Technology, Wuhan University, Wuhan, China; [2]Engineering Training Center, Jianghan University, Wuhan, China; and [3]Department of Physics, School of Physics and Information Engineering, Jianghan University, Wuhan, China

ABSTRACT A full understanding of RNA-mediated biology would require the knowledge of three-dimensional (3D) structures, structural flexibility, and stability of RNAs. To predict RNA 3D structures and stability, we have previously proposed a three-bead coarse-grained predictive model with implicit salt/solvent potentials. In this study, we further develop the model by improving the implicit-salt electrostatic potential and including a sequence-dependent coaxial stacking potential to enable the model to simulate RNA 3D structure folding in divalent/monovalent ion solutions. The model presented here can predict 3D structures of RNA hairpins with bulges/internal loops (<77 nucleotides) from their sequences at the corresponding experimental ion conditions with an overall improved accuracy compared to the experimental data; the model also makes reliable predictions for the flexibility of RNA hairpins with bulge loops of different lengths at several divalent/monovalent ion conditions. In addition, the model successfully predicts the stability of RNA hairpins with various loops/stems in divalent/monovalent ion solutions.

INTRODUCTION

In early years, RNA was considered to be an intermediary in transcription and translation (1). However, in the last two decades, RNA has been shown to perform other crucial functions, such as catalyzing biological reactions and controlling gene expression (2,3). Understanding and utilizing these functions would require comprehensive knowledge of RNA structure and dynamics (4–7). Although RNA sequences are being discovered rapidly, only limited three-dimensional (3D) RNA structures have been determined through experimental methods such as x-ray crystallography, NMR spectroscopy, and cryo-electron microscopy (4,8). Simultaneously, for high efficiency and low cost, some computational models have been developed for predicting 3D structures or thermodynamics of RNA (8–18).

Some models based on fragment assembly, sequence alignment, and secondary structure, e.g., the MC-Fold/MC-Sym pipeline (19), are highly successful at predicting 3D structures even for large RNAs (19–39). However, these models are primarily designed to predict folded structures and would not give reliable predictions for the dynamic and thermodynamic properties of RNAs in three dimensions (16–18). Simultaneously, some other models have been developed that aim to predict RNA dynamics and thermodynamics. The Go-like coarse-grained (CG) three-interaction-site model can predict folding thermodynamics for hairpins and pseudoknots (40,41). Another CG model of oxRNA can capture the thermodynamic and mechanical properties of RNA structures with pairwise interaction potentials (42). However, neither the three-interaction-site model nor the oxRNA model could give reliable predictions for 3D structures of RNAs from their sequences (16,18). Although the three-bead CG model of iFoldRNA (43) and the six/seven-bead CG model of HiRE-RNA (16,44) can predict 3D structures of small RNAs including pseudoknots, the parameters of the two models may need further validation or adjustment for predicting thermodynamic and dynamic properties of RNAs (18,43,44). Furthermore, since RNAs are highly charged polyanionic polymers, RNA structures can be sensitive to ion conditions and temperature (15,45–53). However, none of the above models could predict 3D structures and thermodynamics of RNAs from their sequences over a wide range of ion concentrations and temperatures (16–18).

Very recently, to predict 3D structures and thermal stability of RNAs, we developed a CG model with three beads placed on the existing atoms P, C4′, and N9 for the purine (or N1 for the pyrimidine) (18,54). Combined with an implicit-salt/solvent force field and the Monte Carlo (MC) simulated annealing algorithm, the model not only can predict native-like 3D structures of small RNAs from their sequences at high salt concentration (e.g., 1 M NaCl), but also can give reliable predictions about the stability of RNA hairpins over a wide range of sequences and monovalent ion concentrations that are comparable to the extensive experimental data (54). However, compared with monovalent ions (e.g., Na^+), divalent ions such as Mg^{2+} can play a more special role in the stability and dynamics of RNA structures (55–60). For example, Mg^{2+} is ~1000 times more efficient in inducing

Submitted August 3, 2015, and accepted for publication November 6, 2015.
*Correspondence: zjtan@whu.edu.cn
Editor: Nathan Baker.

0006-3495/15/12/2654/12
http://dx.doi.org/10.1016/j.bpj.2015.11.006

the tertiary-structure folding of *Tetrahymena thermophila* ribozyme (45,49). Although a recent structure-based model with an explicit treatment of Mg^{2+} and an implicit treatment of K^+ can well capture the ion atmosphere around RNAs (61–63), there is still a need for a model that predicts 3D structures and stability of RNAs in divalent/monovalent ion solutions from their sequences. In this work, we further develop our previous model to enable it to predict the 3D structure and stability of RNAs in the presence of divalent ions.

In addition, the functions of RNAs may be not only related to the static 3D structures, but also influenced by the flexibility and stability of their structures (4-7,64–66). The flexibility of RNAs is rather important in the recognition of protein and in gene regulation (64–67). For example, the transactivator response element (TAR) for the transactivator (TAT) protein of the human immunodeficiency virus (HIV) can undergo large conformational changes through the small bulge during the binding of TAT proteins (66,67). Due to the polyanionic nature of RNAs, their flexibility would strongly depend on metal ions such as Mg^{2+} and loops (67–69). To examine the effects of metal ions (e.g., Mg^{2+}) and loops on RNA flexibility, we select HIV-1 TAR and HIV-2 TAR variants as two paradigms in this work. To predict the 3D structures and flexibility of RNA hairpins with bulge loops, such as HIV TAR variants, in divalent/monovalent ion solutions, we introduce, to our knowledge, a new implicit electrostatic potential and an indispensable coaxial stacking interaction between two helices at the junction. With the model presented here, we predict the flexibility of HIV-1 TAR and HIV-2 TAR variants in divalent/monovalent ion solutions to understand the effects of salt and bulge loops.

In this work, we essentially develop the model to simulate RNA folding in divalent/monovalent ion solutions by improving the implicit-salt electrostatic potential and involving a parameterized coaxial stacking potential. Afterward, we first show that the CG model presented here can predict 3D structures of RNAs with bulges/internal loops at given ionic conditions with higher accuracy. Second, the model is employed to investigate the effects of divalent/monovalent salts and bulge length on the flexibility of HIV TAR variant RNAs. Finally, the model is used to quantitatively examine the stability of various RNA hairpins in divalent ion solutions. Throughout the article, all predictions are compared with the extensive experimental data.

MATERIALS AND METHODS

CG structural model

Since CG models generally allow considerable extension of the accessible size and timescale in simulations of biological systems (70–74), we have proposed a three-bead CG model for RNAs where the three beads stand for phosphate, sugar, and base (54). The backbone phosphate (P) bead and sugar (C) bead are placed at the P and C4′ atom positions, respectively, whereas the base (N) beads are placed at the N9 position for purine or the N1 position for pyrimidine (see Fig. 1). The P, C, and N beads are treated as spheres with van der Waals radii of 1.9 Å, 1.7 Å, and 2.2 Å, respectively (54,75).

Force field

In our CG model, the implicit-solvent/salt force field includes eight energy potentials (54):

$$U = U_b + U_a + U_d + U_{exc} + U_{bp} + U_{bs} + U_{el} + U_{cs}. \quad (1)$$

The function forms for the eight energy potentials are described in detail in the Supporting Material, and we introduce them here only briefly, except in the case of the electrostatic interaction, U_{el}, and the coaxial stacking interaction, U_{cs}. The first three terms in Eq. 1 are the bonded potentials for covalent bonds (U_b), bond angles (U_a), and dihedral angles (U_d), respectively. The bonded potentials whose function forms have been described previously (54) were initially parameterized by the statistical analysis on the available 3D structures of RNA molecules in the Protein Data Bank (PDB; http://www.rcsb.org/pdb/home/home.do) (54,75–77). Since lots of native structures in the PDB are mostly A-form helix, the statistical parameters from these structures would not be reasonable to describe the nature of RNA free chains (77). Therefore, for bonded potentials, two sets of parameters are calculated for single strands/loops and stems, referred to as $Para_{nonhelical}$ and $Para_{helical}$, respectively (see the Supporting Material and our previous work (54) for details). The former are used to describe the folding of an RNA from a free chain, and the latter are only used for stems during structure refinement after the folding process. The remaining terms of Eq. 1, namely the nonbonded potentials, describe various pairwise nonbonded interactions. The excluded volume, U_{exc}, between CG beads is modeled by a purely repulsive Lennard-Jones potential (54,75). U_{bp} in Eq. 1 is employed to capture the basepairing interaction between Watson-Crick (G-C, A-U) and wobble (G-U) basepairs (43,54,78–80). U_{bs} in Eq. 1 is a temperature-dependent base-stacking potential that works between nearest-neighbor basepairs. The strength of U_{bs} was derived from the combined analysis of available sequence-dependent thermodynamic parameters (78–80) and the MC algorithm, and the details are shown in our previous work (54).

The electrostatic interaction, U_{el}, in Eq. 1, which is a newly refined term for the effect of divalent ions, is taken into account with the combination of the Debye-Hückel approximation and the concept of counterion condensation (CC) (78,81–83):

$$U_{el} = \sum_{i<j}^{N_P} \frac{(Qe)^2}{4\pi\varepsilon_0\varepsilon(T)r_{ij}} e^{-\frac{r_{ij}}{l_D}}. \quad (2)$$

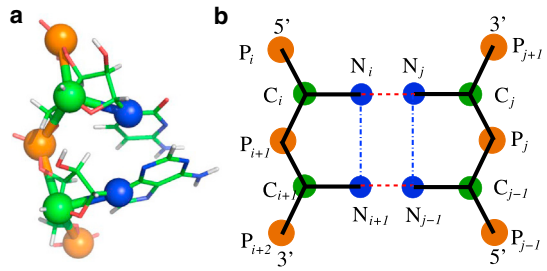

FIGURE 1 (*a*) Our coarse-grained representation for one fragment of an RNA superposed on an all-atom representation. Specifically, three beads are located at the atoms of phosphate (P), C4′ (C), and N1 for pyrimidine or N9 for purine (N), respectively. The structure is shown with the PyMol (http://www.pymol.org). (*b*) The schematic representation for basepairing (*dashed line*) and base-stacking (*dash-dotted line*). To see this figure in color, go online.

The summation is over all the phosphate beads, and r_{ij} is the distance between two phosphate beads i and j. ε_0 is the permittivity of vacuum. $\varepsilon(T)$ is an effective temperature-dependent dielectric constant (40,54,58). l_D is the Debye length of the ionic solution. Beyond our previous model (54), the effect of pure divalent ions and the competition between monovalent and divalent ions are also taken into account in the model presented here to study RNA folding in pure and mixed divalent ion solutions. Based on the CC theory (83), for a pure salt solution containing only one species of salt, such as NaCl or $MgCl_2$, the reduced charge fraction, Q, could be written as $Q = b/(\nu l_B)$ (40,54), where ν is the cation valence, b is the phosphate-phosphate spacing of an RNA, and l_B is the Bjerrum length (40,54). For a mixed Na^+/Mg^{2+} ion solution, we assume $Q = f_{Na^+}Q_{Na^+} + (1 - f_{Na^+})Q_{Mg^{2+}}$, where f_{Na^+} and $(1 - f_{Na^+})$ represent the contribution fractions from Na^+ and Mg^{2+}, respectively. f_{Na^+} can be approximately calculated by the empirical formula previously derived from the tightly-bound-ion (TBI) model, which could account for the divalent-ion-RNA interactions (15,60,84,85)

$$f_{Na^+} = \frac{[Na^+]}{[Na^+] + x[Mg^{2+}]}. \quad (3)$$

Here, $x = (8.1 - 64.8/N)(5.2 - \ln[Na^+])$ (15,60). $[Na^+]$ and $[Mg^{2+}]$ are the corresponding bulk concentrations in molar (M) and N is the chain length.

U_{cs} in Eq. 1 is newly introduced to model the coaxial stacking interaction at RNA junctions (4,86,87), and the coaxial stacking interaction between two discontinuous neighbor helices with interfaced basepairs i-j and k-l can be given by (86,87)

$$U_{cs} = \frac{1}{2}\sum_{i\text{-}j,k\text{-}l}^{N_{cst}} |G_{i\text{-}j,k\text{-}l}| \left\{ \left[1 - e^{-a(r_{ik} - r_{cs})}\right]^2 + \left[1 - e^{-a(r_{jl} - r_{cs})}\right]^2 - 2 \right\}, \quad (4)$$

where $G_{i\text{-}j,k\text{-}l}$ is the sequence-dependent base-stacking strength. $G_{i\text{-}j,k\text{-}l}$ is approximately taken as the stacking strength between the corresponding nearest-neighbor basepairs in an uninterrupted helix (79,86,87). r_{ik} (or r_{jl}) is the distance between two interfaced bases, $i(j)$ and $k(l)$, of two stems, and a represents the extent of distance constraint. r_{cs} is the optimum distance between two coaxially stacked stems. a and r_{cs} are directly obtained from the statistical analysis of the known structures in the PDB database (see Fig. S1). Here, we only include the cases where there is one base or less in at least one of the single-stranded chains between two neighbor helices (see Fig. S1), since the noncanonical basepairs would generally be formed when there are more than one base in each side between two stems (3,4,86).

The detailed descriptions of the potentials in Eq. 1 and all the parameters for the potentials have been described in the Supporting Material; see also our previous work (54) for the process of building the force field.

Simulation algorithm

We use the MC simulated annealing algorithm, which can effectively avoid the trap in local energy minima (34,54,88), to search near-native conformations for an RNA at a given solution condition. Based on a random chain generated for an RNA sequence, the MC simulated annealing algorithm is performed from an initial high temperature to the target temperature (e.g., 298 K) at a fixed ion condition. In the folding process, the $Para_{nonhelical}$ of bonded parameters and the efficient pivot moves for the RNA chain, as well as the standard Metropolis algorithm, are used to sample conformations of a free RNA chain (54,75,78). With gradual cooling of the system, the initial 3D near-native structures would be folded at room temperature for RNAs.

After the MC annealing process with the bonded parameters of $Para_{nonhelical}$ for the whole RNA chain, the secondary structure and native-like 3D structures are predicted from a given sequence. To better capture the geometry of helical parts, further structure refinement is performed for higher accuracy of 3D structures as follows: based on the final 3D structure predicted by the preceding annealing process, another MC simulation (generally 1×10^6 steps) is performed at the corresponding ion condition and room temperature, with the bonded parameters of $Para_{helical}$ and $Para_{nonhelical}$ for the basepairing regions (stems) and loops/single strands, respectively (54). As a result, an ensemble of refined 3D structures (~8000 structures) would be obtained over the last ~8×10^5 MC steps. The predicted 3D structures are evaluated by their RMSD values calculated over C beads from the corresponding C4′ atoms in the native structure in the PDB, though other parameters such as the template-modelling score and the base interaction network fidelity (89,90) can also be used to evaluate the predicted structures. Since the model described here generally predicts a series of native-like structures during the refinement process, we will use the mean RMSD (the averaged value over the whole structure ensemble in the refinement process) and minimal RMSD (corresponding to the structure closest to the native one in the refinement process) to evaluate the reliability of predictions of 3D structures (30,43,54). Although RNA hairpins are mostly A-form helix, since helical stems may deform from the standard A-form helix (48) and the relative orientation between stems at a junction may change (67) with changes in ionic conditions, we still use the RMSD of the whole RNA, including stem and loop, to evaluate the performance of 3D structure prediction (89).

RESULTS AND DISCUSSION

In this section, we employ the CG model to predict, from their sequences, the 3D structures of 32 RNA hairpins, most of which have bulges/internal loops ($\leq$77 nucleotides (nt)), at the respective experimental ion (Na^+/Mg^{2+}) conditions. Afterward, the CG model is used to predict the flexibility and stability of RNA hairpins at extensive divalent/monovalent ion conditions. Our predictions will be compared with the extensive experimental data.

Predicting RNA 3D structures at the respective experimental ion conditions

Going beyond our previous work, which focused on predicting RNA structures at 1 M NaCl (54), we predict the 3D structures of 32 RNA hairpins, most of which have bulges/internal loops. The 3D structures of these RNAs have been determined by NMR at certain ion conditions (see Table 1). For each RNA hairpin, we make two separate predictions using the model presented here at the experimental ion condition and the previous version of the model at 1 M NaCl (54). Table 1 summarizes the major information about the RNA molecules, the corresponding experimental monovalent/divalent ion conditions, and the predictions from our model and the MC-Fold/MC-Sym pipeline (19).

In monovalent solutions

First, we employ this model to predict 3D structures for RNA hairpins (with bulges/internal loops) at the corresponding monovalent ionic conditions listed in their PDB

TABLE 1 The 32 RNA Molecules for 3D Structure Prediction in This Work

RNAs	PDB[a]	Length (nt)	Type[b]	[1+/2+][c] (mM)	$RMSD_{pred.}$[d] (Å) (mean/minimum)	$RMSD^0_{pred.}$[e] (Å)	$RMSD_{MC\text{-}Sym}$[f] (Å)
1	2Y95[g]	14	H	100/0	2.0/1.0	2.2	2
2	2LP9[g]	16	B	65/0	2.6/1.1	2.7	2.8
3	1J4Y[g]	17	H	20/0	3.9/1.9	4	4.6
4	1YN2[g]	17	H	55/40	4.1/2.3	4.3	2.9
5	1Z30[g]	18	H	50/0	1.7/0.9	1.9	2.6
6	1U2A[h]	20	H	75/0	2.1/1.1	2.6	3
7	1QWA[g]	21	B	5/0	2.8/1.5	3	4.6
8	1D0U[g]	21	B	50/0	3.8/1.5	3.7	3.7
9	17RA[h]	21	B	20/0	3.7/1.3	4	4.9
10	1JUR[g]	22	B	100/0	2.9/1.5	2.9	3.4
11	1OSW[g]	22	I	25/0	3.7/1.5	4	4
12	2RO2[g]	23	H	12.4/0	2.8/1.4	2.9	2.6
13	1BGZ[h]	23	B&I	20/0	3.4/2.1	4.6	4.8
14	1S34[g]	23	B	35/0	3.5/1.7	4	2
15	1LC6[g]	24	I	50/0	3.7/2.0	4.1	5.7
16	2KEZ[g]	24	I	50/0	3.4/1.5	3.6	2.9
17	1M82[g]	25	B	20/0	2.1/1.2	2.4	2.2
18	2L5Z[g]	26	I	50/5	4.0/2.6	4.6	5.2
19	2AHT[g]	27	B	100/6	3.2/1.3	3.9	3.8
20	1F6X[g]	27	B	100/0	2.8/1.5	2.9	2.7
21	1XSH[g]	27	B	100/0	3.4/1.7	3.6	2.9
22	1NBR[g]	29	B	20/0	3.3/2.1	3.7	3.1
23	2JWV[g]	29	I	50/0	4.5/2.0	5	4.9
24	1YNE[g]	31	B	10/0	2.6/1.3	2.7	3.9
25	1JO7[g]	31	B&I	10/0	4.0/2.3	4.6	4.9
26	2LWK[g]	32	B&I	50/0	3.9/1.7	4.2	3.3
27	2JXV[g]	33	I	20/0	3.2/1.9	3.9	4.9
28	2KPV[g]	34	B&I	20/0	4.1/1.9	4.3	4.1
29	1ZC5[g]	41	B	10/0	3.5/1.8	3.4	5.9
30	2KUU[g]	48	B	20/0	4.5/2.3	5.1	5.5
31	2MQT[g]	68	B&I	10/0	6.4/3.8	6.8	7.1
32	1P5O[g]	77	B&I	100/5	11.0/8.7	13	10.8

[a]The 3D structures of these RNA hairpins have been determined by NMR at certain ion conditions.
[b]H, B, I, and B&I represent RNA hairpins without bulges/internal loops (H), with bulges (B), with internal loops (I), and with bulges and internal loops (B&I).
[c]The salt conditions of solutions in which RNA structure were experimentally determined.
[d]The mean/minimum RMSDs are calculated over C beads of the structures predicted by the model presented here from the corresponding C4′ atoms of the native structures.
[e]The mean RMSD values are calculated over C beads of structures predicted by our previous model at 1 M NaCl from the corresponding C4′ atoms of the native structures.
[f]The RMSD value is calculated over the C4′ atoms of the top single structure for each RNA predicted by the MC-Fold/MC-Sym pipeline (http://www.major.iric.ca/MC-Fold/) (19) from the native structure.
[g]These hairpins were experimentally determined after the year 2000.
[h]These hairpins were experimentally determined before the year 2000.

files. For example, the structure of stem loop IIa (PDB: 1U2A) is determined by NMR in a buffer consisting of 10 mM KH_2PO_4, 50 mM KCl, 15 mM NaCl, and 0.5 mM EDTA (91). We predict the 3D structures of stem loop IIa from its sequence in a solution of 75 mM monovalent salt. As shown in Table 1, the mean RMSD of the predicted 3D structures is ~2.1 Å, which is obviously smaller than that (~2.6 Å) of the structures predicted by the previous version of our model at 1 M NaCl. For the 28 tested RNA hairpins in the respective monovalent salt solutions, the overall mean RMSD between the structures predicted by this model at the respective experimental monovalent ionic conditions and the experimental structures is 3.37 Å, a smaller value than that (3.66 Å) predicted by the previous version of the model at 1 M NaCl (54); see Table 1.

In divalent solutions

In addition, one important feature of this model is that it combines CC theory and the results from the TBI model, and it can be employed to simulate RNA folding in mixed monovalent/divalent ion solutions. Four RNA hairpins with bulges/internal loops (PDB: 1YN2, 2L5Z, 2AHT, and 1P5O) were determined by NMR in solutions containing Mg^{2+}, and the corresponding monovalent/divalent ion conditions are listed in Table 1. The mean RMSD values for these RNAs predicted by the model presented here in the corresponding experimental mixed ion solutions are 4.1 Å,

4.0 Å, 3.2 Å, and 11.0 Å, respectively, which are clearly smaller than the values (4.3 Å, 4.6 Å, 3.9 Å, and 13.0 Å) predicted by our previous model, regardless of the salt effect and the coaxial stacking interactions (54).

Salt versus coaxial stacking

As shown in Table 1, for the 32 hairpins, this model can make predictions with a visibly lower overall mean RMSD compared to our previous model (3.64 Å vs. 4.02 Å; $p < 0.01$ from the two-tailed Student's *t*-test (12)), which suggests that the inclusion of monovalent/divalent salt and coaxial stacking in this model is effective for predicting RNA 3D structures in ion solutions. To clarify the contributions of the two improvements, we further perform two additional predictions for each hairpin using this model with coaxial stacking at 1 M NaCl and without coaxial stacking at the experimental ion conditions. We find that the two improvements both make positive contributions to the overall improved predictions and that the inclusion of the salt effect has a stronger contribution than that of coaxial stacking, as shown in Fig. S2. Due to the high charge density of the RNA backbone, the native-like structures of hairpins could be slightly stretched at low salt (48), and the difference between conformations at low and high salts should not be ignored (see Fig. S2). Although the coaxial stacking potential would only have a slight effect on hairpins with bulges/internal loops whose bases are stacking into neighbor helices (e.g., 1JUR and 1LC6), it would be indispensable for the formation of coaxially stacked states that are common in RNAs with bulges (e.g., 2KUU and 1P5O) and in the HIV-1 and HIV-2 TAR variants; see the subsection Flexibility of RNAs with bulges.

Comparisons with MC-Fold/MC-Sym pipeline

The MC-Fold/MC-Sym pipeline is a web service (http://www.major.iric.ca/MC-Pipeline/) for RNA secondary and tertiary structure predictions (19). To test the model presented here, we also make comparisons with the MC-Fold/MC-Sym pipeline. The RMSDs of the top 1 structures predicted by the pipeline online server (option: return the best 100 secondary structures and model_limit = 1000 or time_limit = 12 h) are calculated over C4′ atoms from the corresponding atoms in the experimental structures. As shown in Table 1 and Fig. 2, the overall predictions of 32 structures by this model (overall mean RMSD = 3.64 Å) appear visibly better than those from the MC-Fold/MC-Sym pipeline (overall mean RMSD = 4.12 Å), which is also suggested by the *p*-value of <0.01 from the two-tailed Student's *t*-test (12). Table 1 and Fig. 2 also show that this model can give reliable predictions for relatively large RNAs (>45 nt (54)). For the 48-nt and 68-nt RNAs (PDB: 2KUU and 2MQT), this model gives much better predictions than those made by the MC-Fold/MC-Sym pipeline, whereas for the 77-nt RNA (PDB: 1P5O), the two models give similar predictions (see Table 1 and Fig. 2). This may be attributed to the fact that this model still ignores the possible noncanonical basepairs, which are abundant in the 77-nt RNA.

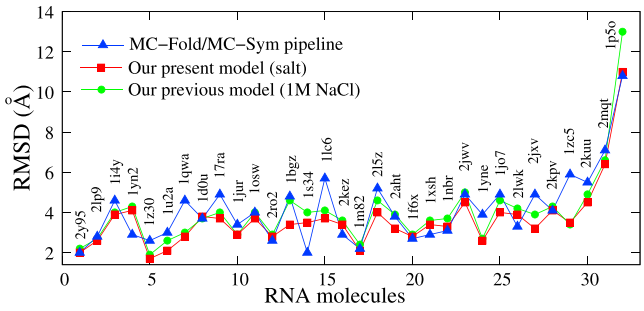

FIGURE 2 Comparison of the RMSD values of RNA 3D structures predicted by the previous version of our model and that presented here and by the MC-Fold/MC-Sym pipeline. For each of the 32 tested RNAs, we use the MC-Fold/MC-Sym pipeline online tool (http://www.major.iric.ca/MC-Fold/) (19) to test the accuracy of MC-Fold/MC-Sym and calculate the RMSD for the top single predicted structure over the C4′ atom in the backbone. The RMSDs of structures predicted by the model described here at the experimental ion conditions (see Table 1) and by its previous version (54) at 1 M NaCl are calculated over C beads from the corresponding C4′ atoms in the native structures. To see this figure in color, go online.

Flexibility of RNAs with bulges

RNAs are highly flexible biomolecules that can undergo dramatic conformational changes to fulfill their diverse functions, e.g., the structural reorganization of riboswitches or the hammerhead ribozyme (2,3). Generally, large conformational transitions of RNA structures are induced by the binding of ions, proteins, or ligands (45–47,65–67). For example, HIV TAR RNAs, which can bind TAT protein through conformational changes during viral replication, have been important paradigms for studying RNA dynamics (66,67). In this section, we employ the model presented here to study the flexibility of HIV-1/HIV-2 TAR variants at different ion conditions and make comparisons with the experimental data. The sequences of the HIV TAR variants and their secondary structures predicted by this model are shown in Fig. 3. The HIV-1 TAR variant with a 3-nt bulge is used to examine conformational changes at different ion conditions with this model, and HIV-2 TAR variants with different lengths of polyU (polyA) bulges are used to study conformational changes induced by various bulges. Our predictions are compared with results from the existing experiments for the HIV-1 TAR variant at extensive Na^+/Mg^{2+} concentrations and for the HIV-2 TAR variant with various bulge lengths at 5 mM Na^+ with/without 2 mM Mg^{2+}, respectively (67,68).

Bending versus monovalent/divalent salts

Recent studies have shown that RNA flexibility is strongly coupled to its ion condition (45–47,67–69,75,92). As experimentally measured by Casiano-Negroni et al., the bending

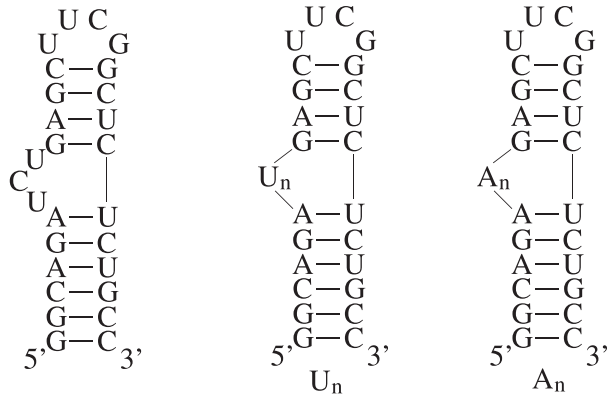

FIGURE 3 The sequences and the secondary structures predicted by model presented here for the HIV-1 TAR variant with a 3-nt bulge and the HIV-2 TAR variants with different lengths of polyU (polyA) bulges, which are used to study the flexibility of RNAs.

angle at the junction of the bulge is strongly dependent on salt (67). Here, we use our model to evaluate the bending at the bulge junction of the HIV-1 TAR variant (see Fig. 3) over a broad range of [Na^+] as well as [Mg^{2+}]. As shown in Fig. 4, *a* and *b*, the interhelical bend angle at the junction of the bulge decreases apparently with the increase of [Na^+] (and [Mg^{2+}]), and our predictions agree well with the corresponding experimental data (67). Such decrease of the bending angle with the increase of [Na^+]/[Mg^{2+}] is understandable. It comes from the competition between electrostatic repulsion, which plays a dominant role at low salt, and coaxial stacking interaction at the bulge, which plays a dominant role at high salt. A comparison between Fig. 4 *a* and Fig. 4 *b* shows that Na^+ and Mg^{2+} induce a similar structural transition from a bent state at low ion concentrations to a coaxially stacked state at high salt, although Mg^{2+} is much more efficient at inducing such structure transition due to the higher ionic charge (55–58).

To further clarify the effect of coaxial stacking, we use the model presented here without coaxial stacking to predict the 3D structures for the HIV-1 TAR variant at extensive Na^+/Mg^{2+} concentrations. A comparison of interhelical bend angles predicted by this model with and without coaxial stacking indicates that the inclusion of coaxial stacking can effectively capture coaxially stacked conformations that are closer to the experimental results, especially at high salt (67), as shown in Fig. S3.

Bending versus bulge length

Zacharias and Hagerman experimentally measured the bending angle for long RNA helices induced by bulges of various lengths (n = 1–6) and base compositions (A_n and U_n series) (68). For simplicity, we predict the 3D structures of the HIV-2 TAR variant (see Fig. 3) with different bulge lengths at 5 mM $NaPO_4$ in the absence and presence of 2 mM Mg^{2+} to study the bulge-induced bending of RNAs. As shown in Fig. 5, *a* and *b*, the bend angle at the bulge increases with an increase in the number of nucleotides in the bulge, which is in good accordance with the experimental data for the bulge of U_n (68). As shown in Fig. 5 *a*, the severe interhelical bending and its sharp increase for longer bulges come from the more extended structures with larger end-to-end distance for longer bulges and from the strong

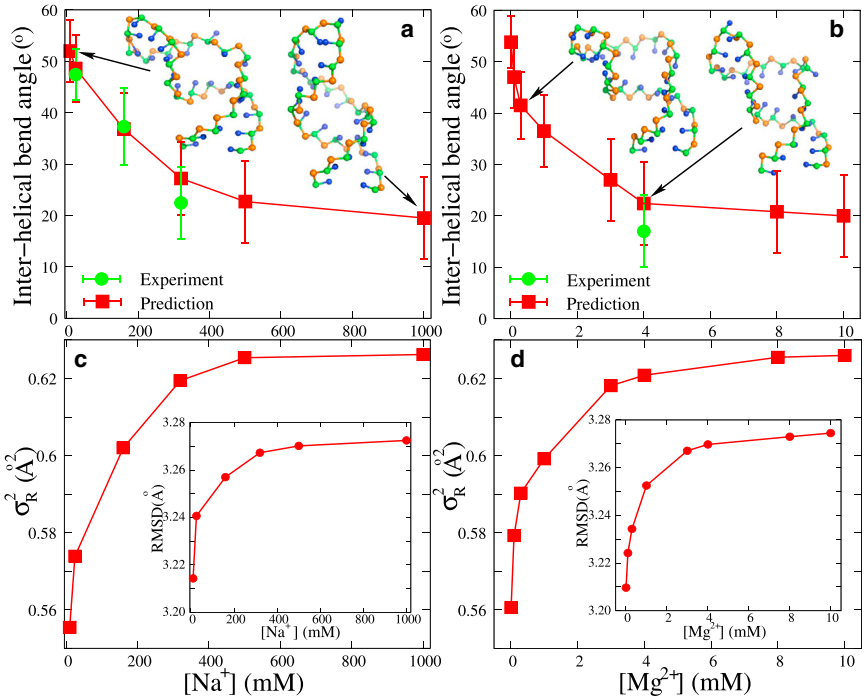

FIGURE 4 (*a* and *b*) The experimental (67) and predicted interhelical bend angle as functions of [Na^+] (*a*) and [Mg^{2+}] (*b*) for the HIV-1 TAR variant (see Fig. 3). The corresponding typical 3D structures predicted by the model presented here are shown with the PyMol (http://www.pymol.org). (*c* and *d*) The variances, σ_R^2, of the radius of gyration and the RMSD values from time-averaged reference structures as functions of [Na^+] (*c*) and [Mg^{2+}] (*d*) for the HIV-1 TAR variant (see Fig. 3). To see this figure in color, go online.

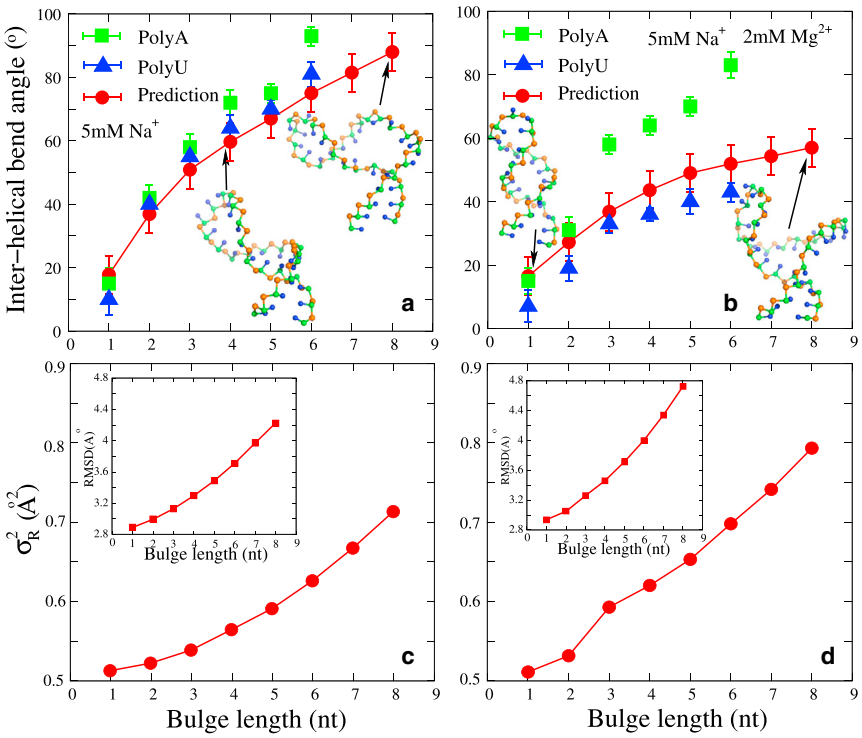

FIGURE 5 (*a* and *b*) The experimental (68) and predicted interhelical bend angles as functions of bulge length at 5 mM $NaPO_4$ without (*a*) and with (*b*) 2 mM Mg^{2+} for the HIV-2 TAR variant (see Fig. 3). The corresponding typical 3D structures predicted by the model presented here are shown with the PyMol (http://www.pymol.org). (*c* and *d*) The variances, σ_R^2, of radius of gyration and RMSD values from time-averaged reference structures as functions of bulge length at 5 mM $NaPO_4$ without (*c*) and with 2 mM Mg^{2+} (*d*) for the HIV-2 TAR variant (see Fig. 3). To see this figure in color, go online.

electrostatic repulsion between helices, which can offset the interhelical stacking interactions at low salt concentrations. Higher salt (2 mM Mg^{2+}) would reduce the electrostatic repulsion more strongly, and thus would promote interhelical stacking and reduce the bending angles (see Fig. 5 *b*). However, as shown in Fig. 5, *a* and *b*, the predictions on bending angle are apparently smaller than the experimental data for the HIV-2 TAR variant with an A_n bulge. This is because polyU behaves like a random coil, whereas polyA would exhibit strong intrachain self-stacking (69), which is not accounted for in the model presented here. Such self-stacking would enhance the rigidity of the single-stranded chain and consequently cause the large bending angle at the bulge.

Recently, Mustoe et al. developed the CG model TOPRNA, which treats RNAs as collections of semirigid helices linked by freely rotatable single strands (74). Their model could nearly reproduce experimental bending angles of HIV-2 TAR with >2-nt bulges at low salt concentrations, whereas it did not give reliable predictions for HIV-2 TAR with a polyU bulge at high salt or for HIV-2 TAR with a 1-nt bulge, possibly because their model did not include the effects of salt and coaxial stacking and overestimated bulge rigidity for polyU (74).

Characterizing global structural fluctuation

The global size of an RNA can be characterized by its radius of gyration, R_g, the fluctuation of which can properly reflect RNA structural flexibility (64,93). Based on the conformational ensemble of an RNA, we have calculated the variance, σ_R^2, of R_g by $\sigma_R^2 = \overline{(R_g - \overline{R_g})^2}$; see Figs. S4 and S5, as well as the mean RMSD from the time-averaged reference structure. As shown in Fig. 4, *c* and *d*, σ_R^2 and the RMSD of the HIV-1 TAR variant increase with the increase in salt concentration, and such increase becomes saturated at high salt. This is because the higher salt can reduce the electrostatic repulsion in the RNA and would favor conformational fluctuation. At high salt concentrations, the coaxial stacking between the two stems can be formed, and consequently, σ_R^2 (and the RMSD) becomes saturated. Furthermore, as shown in Fig. 5, *c* and *d*, σ_R^2 (and the RMSD) of the HIV-2 TAR variant increases for longer bulge length in 5 mM $NaPO_4$ solution in the absence/presence of Mg^{2+}. This is because RNA single-stranded loops are distinctly more flexible than duplexes (69,75,93).

Characterizing local structural fluctuation

Furthermore, we calculated the root-mean-square fluctuation (RMSF) of backbone beads to analyze the local flexibility along an RNA chain (64). The RMSF for the *i*th C-bead is calculated as

$$\mathrm{RMSF}_i = \sqrt{\left\langle \left(\boldsymbol{r}_i(t) - \boldsymbol{r}_i^0\right)^2 \right\rangle_t}, \qquad (5)$$

where the term within the square root symbol is the average over time *t*, $\boldsymbol{r}_i(t)$ is the position of the *i*th C-bead at time *t*, and $\boldsymbol{r}_i^0$ is the time-averaged reference position of the *i*th C-bead.

Fig. 6 shows the RMSF of each C-bead of the HIV-1 TAR variant at 1 M NaCl. The nucleotides near the 5′ and 3′ ends exhibit stronger conformational fluctuations, since the terminal basepairs have fewer spatial constraints (93). As shown in Figs. 6 and 7, the bulge and hairpin loops are more flexible than the stems, which comes from the significantly higher flexibility of unpaired single-stranded chains, since the previous theoretical and experimental studies have suggested that the persistence length of the stem is ~60 times higher than that of the single-strand chain (64–69,75,93). When the length of bulges increases from 1 to 8, the fluctuations of nucleotides, especially at the bulge, are apparently enhanced (see Fig. 7). This is because longer bulges have higher flexibility and naturally cause the higher flexibility of the whole RNA. Additionally, Fig. 7 shows that the addition of Mg^{2+} could increase the local flexibility of RNAs, especially for the HIV-2 TAR variant with a longer bulge. For RNAs with small bulges, the salt effect on local flexibility would not be very obvious (see also Fig. S6). This is reasonable, since the addition of Mg^{2+} would bring the strong electrostatic screening and consequently increase the flexibility of RNAs, especially at bulges. However, for RNAs with small bulges ($\leq$~3 nt), coaxial stacking would be formed between two stems and the addition of Mg^{2+} would only have a slight effect on their flexibility.

RNA hairpin stability in divalent ion solutions

The RNA folded structure is stabilized by the interplay of diverse interactions such as basepairing, base-stacking, and electrostatic interactions. RNA stability at high salt concentrations (e.g., 1 M NaCl) can be predicted with a relatively simple nearest-neighbor model (79,80). However, the nearest-neighbor model cannot predict the 3D structure of RNAs at an arbitrary temperature and cannot predict RNA stability at ionic conditions departing from 1 M NaCl. However, due to its polyanionic nature, RNA stability is very sensitive to ionic condition (55-58,94–99), and Mg^{2+} ions are particularly efficient at stabilizing RNA tertiary structure (55–57). To address the effect of Mg^{2+} in RNA stability, we employ the model present here to study the stability of various RNA hairpins in divalent and mixed divalent/monovalent ion solutions. Generally, a hairpin is either in a folded state at low T or an unfolded state at high T, or is bistable at middle T around the melting temperature T_m (50–52,54). Based on the equilibrium value of the number of basepairs at each temperature, T, to obtain the T_m of a hairpin, the fraction of denatured basepairs, $f(T)$, can be calculated and fitted to a two-state model (54,79),

$$f(T) = 1 - \frac{1}{1 + e^{(T-T_m)/dT}}, \quad (6)$$

where dT is an adjustable parameter (54).

In pure divalent solutions

Recently, the loop-size dependence of the stability of an RNA hairpin (denoted as R0) was experimentally investigated in 2.5 mM Mg^{2+} (94). Fig. 8 *a* shows the sequences of hairpins R0, as well as the secondary structures predicted by the model presented here. Fig. 8 *a* also shows the predicted T_m for R0 in 2.5 mM Mg^{2+} as a function of the loop size (m = 4~34 nt), which is in good accordance with the experimental data (94). Due to the larger conformational entropy for longer loops, the hairpin stability would decrease when the hairpin loop becomes longer. In addition, we have studied the stability of hairpins with the same loop but with different stems in 0.7 mM Mg^{2+} solutions. Hairpins R1–R4 are four similar RNAs whose stems are slightly different in length or sequence (see Fig. 8 *b*). As shown in Fig. 8 *b*, the T_m values of four hairpins predicted by the model presented here at 0.7 mM $MgCl_2$ are in good agreement with the experimental data (95). The addition of a G-C or C-G basepair (from R1 to R2 and then to R3) can dramatically stabilize the RNA hairpins due to the strong basepairing/base-stacking interactions. The difference between the T_m values of R3 and R4 indicates that RNA stability is sensitive to sequence-dependent basepairing/base-stacking. The good agreement between our

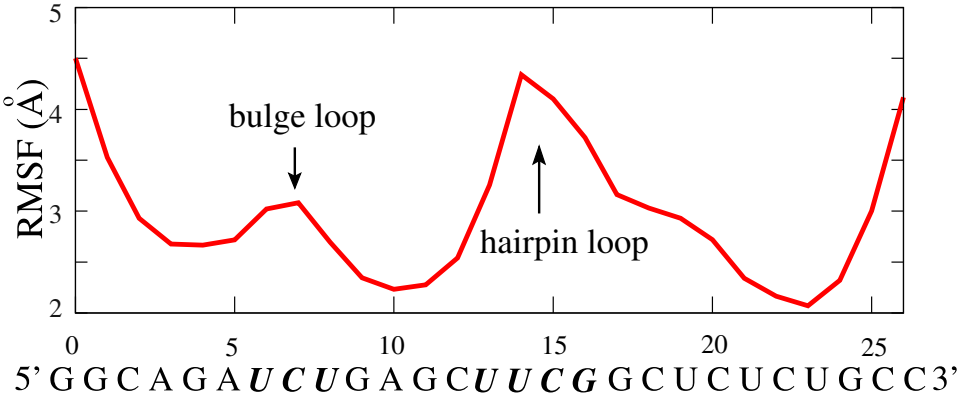

FIGURE 6 The RMSF for C-beads along the HIV-1 TAR variant (see Fig. 3) in 1 M NaCl solution. The sequences of bulge and hairpin loops are in italics. To see this figure in color, go online.

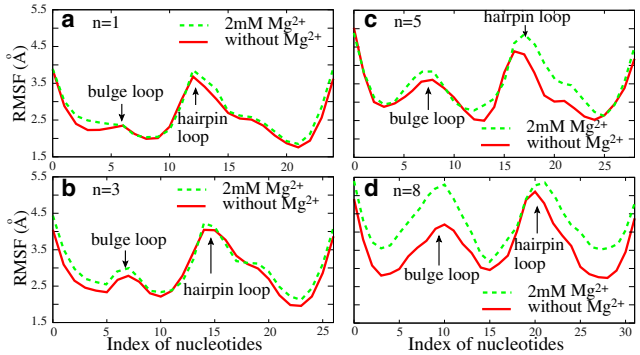

FIGURE 7 The RMSF for C-beads along the HIV-2 TAR variant (see Fig. 3) with different bulge length at 5 mM $NaPO_4$ with and without 2 mM Mg^{2+}. The lengths, *n*, of the bulge loop are 1 (*a*), 3 (*b*), 5 (*c*), and 8 (*d*), respectively. To see this figure in color, go online.

predictions and experiments (94,95) suggests that our model can well describe the folding stability of small RNAs in pure Mg^{2+} solutions.

In mixed divalent/monovalent solutions

Using the model presented here, we also examined the stability of RNA hairpins R5 and R6 in mixed K^+/Mg^{2+} solutions, and the secondary structures for R5 and R6 predicted by this model are shown in Fig. 8, *c* and *d*. As shown in Fig. 8, *c* and *d*, for T_m values of R5 and R6 at a fixed $[K^+]$ (100 mM), in addition to the general trend of increased stability for higher $[Mg^{2+}]$, the competition between K^+ and Mg^{2+} also leads to the following behavior of RNA stability. At low $[Mg^{2+}]$ ($\leq$0.1 mM), the stability of RNAs is dominated by (~100 mM) K^+ and the T_m values are close to that at the corresponding pure $[K^+]$. As $[Mg^{2+}]$ is increased, Mg^{2+} ions begin to play a role and the stability of RNAs begins to increase markedly due to the efficient role of Mg^{2+} in stabilizing RNAs. At very high $[Mg^{2+}]$, Mg^{2+} would become dominant and the T_m would become saturated (60). As shown in Fig. 8, *c* and *d*, the agreement with experimental data (98,99) indicates that the combination of CC theory and the results from the TBI model could give a good description of the competition between monovalent and divalent ions in stabilizing RNA hairpins, and the model presented here can give good predictions of the stability of RNA hairpins in mixed divalent/monovalent solutions.

CONCLUSIONS

In this work, we have developed our CG model to predict 3D structures and structural properties of RNAs with bulges/internal loops in the presence of divalent and monovalent ions. The major extensions of our CG model include improvement of the electrostatic potential to implicitly consider the effect of divalent ions and inclusion of coaxial stacking at the two-way junction. The improved CG model has been employed to examine the effects of divalent/monovalent salt on the 3D structure, flexibility, and stability of RNA hairpins with a two-way junction.

First, we employed this model to predict 3D structures for 32 RNAs ($\leq$77 nt) at the respective monovalent/divalent

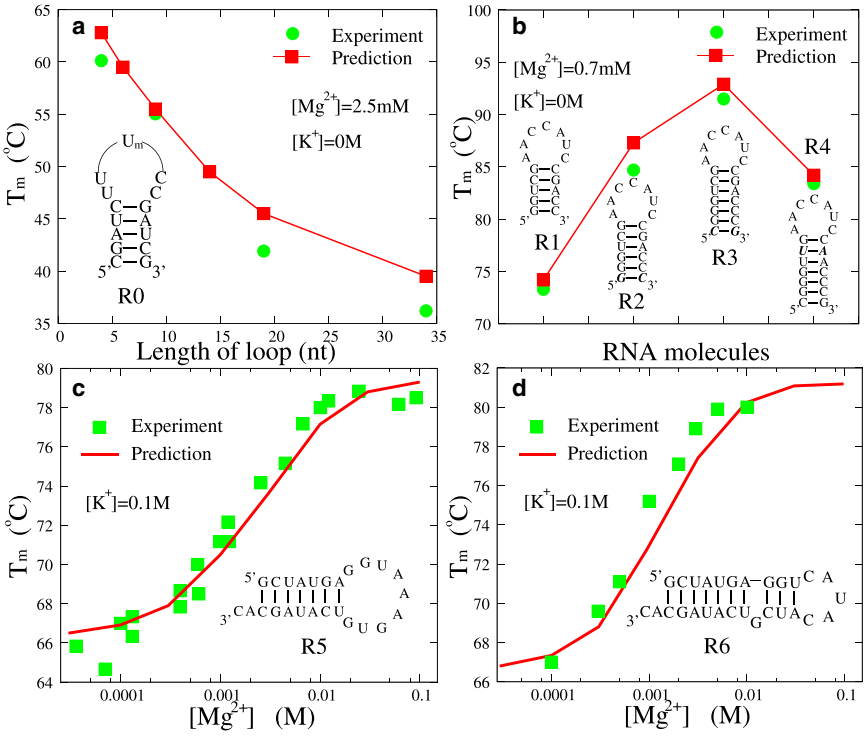

FIGURE 8 (*a*) The experimental (94) and predicted melting temperature as functions of length of hairpin loop for RNA hairpin R0 at 2.5 mM $[Mg^{2+}]$. (*b*) The experimental (95) and predicted melting temperatures of four RNA hairpins with different stems at 0.7 mM $[Mg^{2+}]$. (*c* and *d*) The experimental (98,99) and predicted melting temperatures, T_ms, as functions of $[Mg^{2+}]$ for the two RNA hairpins R5 (*c*) and R6 (*d*) in the presence of 0.1 M $[K^+]$. The sequences and the secondary structures predicted by the model presented here are also shown in (*a*)–(*d*). To see this figure in color, go online.

salt conditions in which the RNA structures were experimentally determined by NMR, and the overall mean RMSD of 3.64 Å between the predictions and the experimental structures is visibly smaller than those from predictions by the MC-Fold/MC-Sym pipeline and by previous version of our model at 1 M NaCl. Second, we studied the flexibility of RNA hairpins with varying bulge loops at extensive Na^+ and Mg^{2+} concentrations, and the predicted bending angles at the bulge for HIV-1 and HIV-2 TAR variants are in good agreement with the available experimental data for different salt conditions, as well as the different lengths of bulge loops. Third, we predicted the stability of RNA hairpins in divalent and mixed divalent/monovalent ion solutions, and the predictions agree well with the experimental data. Therefore, this improved model can provide the ensemble of probable 3D structures at extensive divalent/monovalent ion conditions and can make reliable predictions on structural properties such as flexibility and stability for small RNAs.

Despite the extensive agreement between our predictions and experiments, this model still involves some approximations and simplifications. First, in this model, the effect of divalent and monovalent salts is implicitly accounted for by the combination of CC theory and the TBI model. The good agreement with experimental data suggests that the combination of CC theory and TBI model can well capture the efficient role of divalent ions over monovalent ions, though more extensive experimental validation for larger RNAs is still required. Also, this model ignores the effect of specific ion binding, which might become important for large RNAs with complex structures (55–57). Of course, a more accurate treatment of salt is to explicitly consider the metal ions (75,97), which would bring a huge computational cost. Very recently, Hayes et al. proposed a generalized Manning CC model in an alternative way that reproduces the ion atmosphere around RNAs through the explicit representation of Mg^{2+} and implicit treatment of K^+ (61–63). Second, the CG model presented here only considers the canonical Watson-Crick (C-G, A-U) and wobble (G-U) basepairing, and ignores the possible noncanonical basepairs due to the lack of experimental thermodynamic parameters (79,80). Noncanonical basepairing can be included in this model with the corresponding thermodynamic parameters, which would further improve the accuracy of structure prediction for RNAs with loops (19,30). Third, although our model can predict 3D structures for RNAs beyond hairpins, e.g., small pseudoknots (54), it is still a challenge for the present version of the model to accurately and efficiently predict 3D structures of large RNAs with complex structures. Nevertheless, we are currently extending the model to predict 3D structures and stability for extensive RNA pseudoknots and complex structures, whereas 3D structure prediction for large RNAs from their sequences may require certain experimental constraints (34,39). Finally, the 3D structure predicted by this model is at the CG level, and consequently, it is still necessary to develop this model to reconstruct all-atomistic structures based on CG predictions. Nevertheless, this model can be a reliable predictive model for 3D structure ensemble of small RNAs in divalent/monovalent solutions and at arbitrary temperatures.

SUPPORTING MATERIAL

Supporting Materials and Methods, six figures, and two tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)01160-1.

AUTHOR CONTRIBUTIONS

Z.J.T. and Y.Z.S. designed the research; Y.Z.S., L.J., and F.H.W. performed the research; Z.J.T., Y.Z.S., and X.L.Z. analyzed data; and Y.Z.S. and Z.J.T. wrote the article.

ACKNOWLEDGMENTS

We are grateful to Professors Shi-Jie Chen (University of Missouri), Yang Zhang (University of Michigan), and Wenbing Zhang (Wuhan University) for valuable discussions.

This work was supported by grants from the National Key Scientific Program (973)-Nanoscience and Nanotechnology (2011CB933600), the National Science Foundation of China (11175132, 11374234, and 11575128), and the Program for New Century Excellent Talents (NCET 08-0408).

REFERENCES

1. Crick, F. 1970. Central dogma of molecular biology. *Nature.* 227:561–563.
2. Doherty, E. A., and J. A. Doudna. 2001. Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Biomol. Struct.* 30:457–475.
3. Edwards, T. E., D. J. Klein, and A. R. Ferré-D'Amaré. 2007. Riboswitches: small-molecule recognition by gene regulatory RNAs. *Curr. Opin. Struct. Biol.* 17:273–279.
4. Tinoco, I., Jr., and C. Bustamante. 1999. How RNA folds. *J. Mol. Biol.* 293:271–281.
5. Hall, K. B. 2012. Spectroscopic probes of RNA structure and dynamics. *Methods Mol. Biol.* 875:67–84.
6. Zhang, W., and S. J. Chen. 2002. RNA hairpin-folding kinetics. *Proc. Natl. Acad. Sci. USA.* 99:1931–1936.
7. Gong, S., Y. Wang, and W. Zhang. 2015. Kinetic regulation mechanism of pbuE riboswitch. *J. Chem. Phys.* 142:015103.
8. Sim, A. Y., P. Minary, and M. Levitt. 2012. Modeling nucleic acids. *Curr. Opin. Struct. Biol.* 22:273–278.
9. Rother, K., M. Rother, …, J. M. Bujnicki. 2011. RNA and protein 3D structure modeling: similarities and differences. *J. Mol. Model.* 17:2325–2336.
10. Laing, C., and T. Schlick. 2011. Computational approaches to RNA structure prediction, analysis, and design. *Curr. Opin. Struct. Biol.* 21:306–318.
11. Cruz, J. A., M. F. Blanchet, …, E. Westhof. 2012. RNA-Puzzles: a CASP-like evaluation of RNA three-dimensional structure prediction. *RNA.* 18:610–625.
12. Hajdin, C. E., F. Ding, …, K. M. Weeks. 2010. On the significance of an RNA tertiary structure prediction. *RNA.* 16:1340–1349.

13. Shapiro, B. A., Y. G. Yingling, …, E. Bindewald. 2007. Bridging the gap in RNA structure prediction. *Curr. Opin. Struct. Biol.* 17:157–165.
14. Zhang, Y. 2008. Progress and challenges in protein structure prediction. *Curr. Opin. Struct. Biol.* 18:342–348.
15. Tan, Z., W. Zhang, …, F. Wang. 2015. RNA folding: structure prediction, folding kinetics and ion electrostatics. *Adv. Exp. Med. Biol.* 827:143–183.
16. Cragnolini, T., P. Derreumaux, and S. Pasquali. 2015. Ab initio RNA folding. *J. Phys. Condens. Matter.* 27:233102.
17. Bailor, M. H., A. M. Mustoe, …, H. M. Al-Hashimi. 2011. Topological constraints: using RNA secondary structure to model 3D conformation, folding pathways, and dynamic adaptation. *Curr. Opin. Struct. Biol.* 21:296–305.
18. Shi, Y. Z., Y. Y. Wu, …, Z. J. Tan. 2014. RNA structure prediction: progress and perspective. *Chin. Phys. B.* 23:078701.
19. Parisien, M., and F. Major. 2008. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature.* 452:51–55.
20. Massire, C., and E. Westhof. 1998. MANIP: an interactive tool for modelling RNA. *J. Mol. Graph. Model.* 16:197–205, 255–257.
21. Jossinet, F., T. E. Ludwig, and E. Westhof. 2010. Assemble: an interactive graphical tool to analyze and build RNA architectures at the 2D and 3D levels. *Bioinformatics.* 26:2057–2059.
22. Martinez, H. M., J. V. Maizel, Jr., and B. A. Shapiro. 2008. RNA2D3D: a program for generating, viewing, and comparing 3-dimensional models of RNA. *J. Biomol. Struct. Dyn.* 25:669–683.
23. Rother, M., K. Rother, …, J. M. Bujnicki. 2011. ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res.* 39:4007–4022.
24. Flores, S. C., and R. B. Altman. 2010. Turning limited experimental information into 3D models of RNA. *RNA.* 16:1769–1778.
25. Popenda, M., M. Szachniuk, …, R. W. Adamiak. 2012. Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* 40:e112.
26. Zhao, Y., Z. Gong, and Y. Xiao. 2011. Improvements of the hierarchical approach for predicting RNA tertiary structure. *J. Biomol. Struct. Dyn.* 28:815–826.
27. Zhao, Y., Y. Huang, …, Y. Xiao. 2012. Automated and fast building of three-dimensional RNA structures. *Sci. Rep.* 2:734.
28. Huang, Y., S. Liu, …, Y. Xiao. 2013. A novel protocol for three-dimensional structure prediction of RNA-protein complexes. *Sci. Rep.* 3:1887.
29. Wang, J., Y. Zhao, …, Y. Xiao. 2015. 3dRNAscore: a distance and torsion angle dependent evaluation function of 3D RNA structures. *Nucleic Acids Res.* 43:e63.
30. Das, R., and D. Baker. 2007. Automated de novo prediction of native-like RNA tertiary structures. *Proc. Natl. Acad. Sci. USA.* 104:14664–14669.
31. Bida, J. P., and L. J. Maher, 3rd. 2012. Improved prediction of RNA tertiary structure with insights into native state dynamics. *RNA.* 18:385–393.
32. Zhang, J., Y. Bian, …, W. Wang. 2012. RNA fragment modeling with a nucleobase discrete-state model. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 85:021909.
33. Zhang, J., J. Dundas, …, J. Liang. 2009. Prediction of geometrically feasible three-dimensional structures of pseudoknotted RNA through free energy estimation. *RNA.* 15:2248–2263.
34. Seetin, M. G., and D. H. Mathews. 2011. Automated RNA tertiary structure prediction from secondary structure and low-resolution restraints. *J. Comput. Chem.* 32:2232–2244.
35. Tan, R. K. Z., A. S. Petrov, and S. C. Harvey. 2006. YUP: A molecular simulation program for coarse-grained and multiscaled models. *J. Chem. Theory Comput.* 2:529–540.
36. Jonikas, M. A., R. J. Radmer, …, R. B. Altman. 2009. Coarse-grained modeling of large RNA molecules with knowledge-based potentials and structural filters. *RNA.* 15:189–199.
37. Cao, S., and S. J. Chen. 2011. Physics-based de novo prediction of RNA 3D structures. *J. Phys. Chem. B.* 115:4216–4226.
38. Xu, X., P. Zhao, and S. J. Chen. 2014. Vfold: a web server for RNA structure and folding thermodynamics prediction. *PLoS One.* 9:e107504.
39. Xia, Z., D. R. Bell, …, P. Ren. 2013. RNA 3D structure prediction by using a coarse-grained model and experimental data. *J. Phys. Chem. B.* 117:3135–3144.
40. Denesyuk, N. A., and D. Thirumalai. 2013. Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B.* 117:4901–4911.
41. Hyeon, C., and D. Thirumalai. 2005. Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. USA.* 102:6789–6794.
42. Šulc, P., F. Romano, …, A. A. Louis. 2014. A nucleotide-level coarse-grained model of RNA. *J. Chem. Phys.* 140:235102.
43. Ding, F., S. Sharma, …, N. V. Dokholyan. 2008. Ab initio RNA folding by discrete molecular dynamics: from structure prediction to folding mechanisms. *RNA.* 14:1164–1173.
44. Pasquali, S., and P. Derreumaux. 2010. HiRE-RNA: a high resolution coarse-grained energy model for RNA. *J. Phys. Chem. B.* 114:11957–11966.
45. Woodson, S. A. 2005. Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol.* 9:104–109.
46. Chen, S. J. 2008. RNA folding: conformational statistics, folding kinetics, and ion electrostatics. *Annu. Rev. Biophys.* 37:197–214.
47. Lipfert, J., S. Doniach, …, D. Herschlag. 2014. Understanding nucleic acid-ion interactions. *Annu. Rev. Biochem.* 83:813–841.
48. Manning, G. S. 2015. The response of DNA length and twist to changes in ionic strength. *Biopolymers.* 103:223–226.
49. Takamoto, K., Q. He, …, M. Brenowitz. 2002. Monovalent cations mediate formation of native tertiary structure of the *Tetrahymena thermophila* ribozyme. *Nat. Struct. Biol.* 9:928–933.
50. Zhang, Y., J. Zhang, and W. Wang. 2011. Atomistic analysis of pseudoknotted RNA unfolding. *J. Am. Chem. Soc.* 133:6882–6885.
51. Bian, Y., J. Zhang, …, W. Wang. 2015. Free energy landscape and multiple folding pathways of an H-type RNA pseudoknot. *PLoS One.* 10:e0129089.
52. Xu, X., and S. J. Chen. 2012. Kinetic mechanism of conformational switch between bistable RNA hairpins. *J. Am. Chem. Soc.* 134:12499–12507.
53. Chen, J., and W. Zhang. 2012. Kinetic analysis of the effects of target structure on siRNA efficiency. *J. Chem. Phys.* 137:225102.
54. Shi, Y. Z., F. H. Wang, …, Z. J. Tan. 2014. A coarse-grained model with implicit salt for RNAs: predicting 3D structure, stability and salt effect. *J. Chem. Phys.* 141:105102.
55. Pabit, S. A., J. L. Sutton, …, L. Pollack. 2013. Role of ion valence in the submillisecond collapse and folding of a small RNA domain. *Biochemistry.* 52:1539–1546.
56. Leipply, D., and D. E. Draper. 2011. Effects of Mg^{2+} on the free energy landscape for folding a purine riboswitch RNA. *Biochemistry.* 50:2790–2799.
57. Meisburger, S. P., S. A. Pabit, and L. Pollack. 2015. Determining the locations of ions and water around DNA from x-ray scattering measurements. *Biophys. J.* 108:2886–2895.
58. Tan, Z. J., and S. J. Chen. 2006. Nucleic acid helix stability: effects of salt concentration, cation valence and size, and chain length. *Biophys. J.* 90:1175–1190.
59. Tan, Z. J., and S. J. Chen. 2010. Predicting ion binding properties for RNA tertiary structures. *Biophys. J.* 99:1565–1576.
60. Tan, Z. J., and S. J. Chen. 2007. RNA helix stability in mixed Na^{+}/Mg^{2+} solution. *Biophys. J.* 92:3615–3632.
61. Hayes, R. L., J. K. Noel, …, J. N. Onuchic. 2015. Generalized manning condensation model captures the RNA ion atmosphere. *Phys. Rev. Lett.* 114:258105.

62. Hayes, R. L., J. K. Noel, …, J. N. Onuchic. 2014. Reduced model captures Mg^{2+}-RNA interaction free energy of riboswitches. *Biophys. J.* 106:1508–1519.

63. Hayes, R. L., J. K. Noel, …, K. Y. Sanbonmatsu. 2012. Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. *J. Am. Chem. Soc.* 134:12043–12053.

64. Hagerman, P. J. 1997. Flexibility of RNA. *Annu. Rev. Biophys. Biomol. Struct.* 26:139–156.

65. Herschlag, D., B. E. Allred, and S. Gowrishankar. 2015. From static to dynamic: the need for structural ensembles and a predictive model of RNA folding and function. *Curr. Opin. Struct. Biol.* 30:125–133.

66. Mouzakis, K. D., E. A. Dethoff, …, S. E. Butcher. 2015. Dynamic motions of the HIV-1 frameshift site RNA. *Biophys. J.* 108:644–654.

67. Casiano-Negroni, A., X. Sun, and H. M. Al-Hashimi. 2007. Probing Na^{+}-induced changes in the HIV-1 TAR conformational dynamics using NMR residual dipolar couplings: new insights into the role of counterions and electrostatic interactions in adaptive recognition. *Biochemistry.* 46:6525–6535.

68. Zacharias, M., and P. J. Hagerman. 1995. Bulge-induced bends in RNA: quantification by transient electric birefringence. *J. Mol. Biol.* 247:486–500.

69. Chen, H., S. P. Meisburger, …, L. Pollack. 2012. Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci. USA.* 109:799–804.

70. de Pablo, J. J. 2011. Coarse-grained simulations of macromolecules: from DNA to nanocomposites. *Annu. Rev. Phys. Chem.* 62:555–574.

71. Zhou, H. X. 2014. Theoretical frameworks for multiscale modeling and simulation. *Curr. Opin. Struct. Biol.* 25:67–76.

72. Kikot, I. P., A. V. Savin, …, A. V. Onufriev. 2011. New coarse-grained DNA model. *Biofizika.* 56:396–402.

73. Daily, M. D., B. N. Olsen, …, N. A. Baker. 2014. Improved coarse-grained modeling of cholesterol-containing lipid bilayers. *J. Chem. Theory Comput.* 10:2137–2150.

74. Mustoe, A. M., H. M. Al-Hashimi, and C. L. Brooks, 3rd. 2014. Coarse grained models reveal essential contributions of topological constraints to the conformational free energy of RNA bulges. *J. Phys. Chem. B.* 118:2615–2627.

75. Wang, F. H., Y. Y. Wu, and Y. Y. Tan. 2013. Salt contribution to the flexibility of single-stranded nucleic acid offinite length. *Biopolymers.* 99:370–381.

76. Huang, S. Y., and X. Q. Zou. 2014. A knowledge-based scoring function for protein-RNA interactions derived from a statistical mechanics-based iterative method. *Nucleic Acids Res.* 42:e55.

77. Huang, S. Y., and X. Q. Zou. 2011. Statistical mechanics-based method to extract atomic distance-dependent potentials from protein structures. *Proteins.* 79:2648–2661.

78. Zhang, Y., H. Zhou, and Z. C. Ou-Yang. 2001. Stretching single-stranded DNA: interplay of electrostatic, base-pairing, and base-pair stacking interactions. *Biophys. J.* 81:1133–1143.

79. Xia, T., J. SantaLucia, Jr., …, D. H. Turner. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry.* 37:14719–14735.

80. Mathews, D. H., J. Sabina, …, D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* 288:911–940.

81. Thomas, D. G., J. Chun, …, N. A. Baker. 2013. Parameterization of a geometric flow implicit solvation model. *J. Comput. Chem.* 34:687–695.

82. Ren, P., J. Chun, …, N. A. Baker. 2012. Biomolecular electrostatics and solvation: a computational perspective. *Q. Rev. Biophys.* 45:427–491.

83. Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11:179–246.

84. Tan, Z. J., and S. J. Chen. 2005. Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. *J. Chem. Phys.* 122:44903.

85. Tan, Z. J., and S. J. Chen. 2011. Salt contribution to RNA tertiary structure folding stability. *Biophys. J.* 101:176–187.

86. Walter, A. E., D. H. Turner, …, M. Zuker. 1994. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl. Acad. Sci. USA.* 91:9218–9222.

87. Walter, A. E., and D. H. Turner. 1994. Sequence dependence of stability for coaxial stacking of RNA helixes with Watson-Crick base paired interfaces. *Biochemistry.* 33:12715–12719.

88. Kirkpatrick, S., C. D. Gelatt, Jr., and M. P. Vecchi. 1983. Optimization by simulated annealing. *Science.* 220:671–680.

89. Parisien, M., J. A. Cruz, …, F. Major. 2009. New metrics for comparing and assessing discrepancies between RNA 3D structures and models. *RNA.* 15:1875–1885.

90. Zhang, Y., and J. Skolnick. 2004. Scoring function for automated assessment of protein structure template quality. *Proteins.* 57:702–710.

91. Stallings, S. C., and P. B. Moore. 1997. The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA. *Structure.* 5:1173–1185.

92. Tan, Z. J., and S. J. Chen. 2008. Electrostatic free energy landscapes for DNA helix bending. *Biophys. J.* 94:3137–3149.

93. Wu, Y. Y., L. Bao, …, Z. J. Tan. 2015. Flexibility of short DNA helices with finite-length effect: from base pairs to tens of base pairs. *J. Chem. Phys.* 142:125103.

94. Kuznetsov, S. V., C. C. Ren, …, A. Ansari. 2008. Loop dependence of the stability and dynamics of nucleic acid hairpins. *Nucleic Acids Res.* 36:1098–1112.

95. Sehdev, P., G. Crews, and A. M. Soto. 2012. Effect of helix stability on the formation of loop-loop complexes. *Biochemistry.* 51:9612–9623.

96. Tan, Z. J., and S. J. Chen. 2008. Salt dependence of nucleic acid hairpin stability. *Biophys. J.* 95:738–752.

97. Wu, Y. Y., Z. L. Zhang, …, Z. J. Tan. 2015. Multivalent ion-mediated nucleic acid helix-helix interactions: RNA versus DNA. *Nucleic Acids Res.* 43:6156–6165.

98. Nixon, P. L., C. A. Theimer, and D. P. Giedroc. 1999. Thermodynamics of stabilization of RNA pseudoknots by cobalt^{III} hexaammine. *Biopolymers.* 50:443–458.

99. Nixon, P. L., and D. P. Giedroc. 1998. Equilibrium unfolding (folding) pathway of a model H-type pseudoknotted RNA: the role of magnesium ions in stability. *Biochemistry.* 37:16116–16129.

Supporting Material for

Predicting 3D Structure, Flexibility and Stability of RNA Hairpins in Monovalent and Divalent Ion Solutions

Ya-Zhou Shi[1], Lei Jin[1], Feng-Hua Wang[2], Xiao-Long Zhu[3], and Zhi-Jie Tan[1*]

[1]*Department of Physics and Key Laboratory of Artificial Micro- and Nano-structures of the Ministry of Education, School of Physics and Technology, Wuhan University, Wuhan, China*

[2]*Engineering Training Center, Jianghan University, Wuhan, China*

[3]*Department of Physics, School of Physics and Information Engineering, Jianghan University, Wuhan, China*

[*] To whom correspondence should be addressed. Email: zjtan@whu.edu.cn

The force-field of the coarse-grained model

1. Energy functions

The total energy *U* of the system in the present model includes eight distinct contributions (1):

$$U = U_b + U_a + U_d + U_{exc} + U_{bp} + U_{bs} + U_{el} + U_{cs}, \tag{S1}$$

where

$$U_b = \sum_{bonds} K_b(r - r_0)^2, \tag{S2}$$

$$U_a = \sum_{angles} K_\theta(\theta - \theta_0)^2, \tag{S3}$$

$$U_d = \sum_{dihedrals} \{K_\varphi[1 - cos(\varphi - \varphi_0)] + \frac{1}{2}K_\varphi[1 - cos3(\varphi - \varphi_0)]\}, \tag{S4}$$

$$U_{exc} = \sum_{i<j}^{N} \begin{cases} 4\varepsilon\left[\left(\frac{\sigma_0}{r_{ij}}\right)^{12} - \left(\frac{\sigma_0}{r_{ij}}\right)^{6}\right] & if\ r_{ij} \le \sigma_0 \\ 0 & if\ r_{ij} > \sigma_0 \end{cases}, \tag{S5}$$

$$U_{bp} = \sum_{i<j-3}^{N_{bp}} \frac{\varepsilon_{bp}}{1 + k_{NN}\left(r_{N_iN_j} - r_{NN}\right)^2 + k_{CN}\sum_{i(j)}\left(r_{C_iN_j} - r_{CN}\right)^2 + k_{PN}\sum_{i(j)}\left(r_{P_iN_j} - r_{PN}\right)^2}, \tag{S6}$$

$$U_{bs} = \frac{1}{2}\sum_{i,j}^{N_{st}} |G_{i,i+1,j-1,j}| \left\{\left[5\left(\frac{\sigma_{st}}{r_{i,i+1}}\right)^{12} - 6\left(\frac{\sigma_{st}}{r_{i,i+1}}\right)^{10}\right] + \left[5\left(\frac{\sigma_{st}}{r_{j,j-1}}\right)^{12} - 6\left(\frac{\sigma_{st}}{r_{j,j-1}}\right)^{10}\right]\right\}, \tag{S7}$$

$$U_{el} = \sum_{i<j}^{N_P} \frac{(Qe)^2}{4\pi\varepsilon_0\varepsilon(T)r_{ij}} \mathrm{e}^{-\frac{r_{ij}}{l_D}}. \tag{S8}$$

$$U_{cs} = \frac{1}{2}\sum_{i\text{-}j,k\text{-}l}^{N_{cst}} \left|G_{i\text{-}j,k\text{-}l}\right| \{\left[1 - e^{-a(r_{ik} - r_{cs})}\right]^2 + \left[1 - e^{-a(r_{jl} - r_{cs})}\right]^2 - 2\}. \tag{S9}$$

The first three terms in Eq. S1 are typical expressions for virtual bonds U_b, bond angles U_a and dihedrals U_d. The initial parameters including equilibrium distances/angles and corresponding energy strength in these terms are derived from the Boltzmann inversion of the corresponding atomistic distribution functions obtained by the statistical analysis on the experimental structures in the Protein Data Bank (PDB, http://www.rcsb.org/pdb/home/home.do) (1). For these bonded potentials, two sets

of parameters $Para_{helical}$ and $Para_{nonhelical}$ are provided for stems and single-strands/loops (1), respectively, due to the different structural features of stems (helical) and single-strands/loops (nonhelical) in RNAs. The $Para_{nonhelical}$ are used in the folding process to describe the folding of free RNA chains (1), while during the structure refinement based on the 3D structure predicted from the initial folding process, the $Para_{helical}$ and $Para_{nonhelical}$ are used for the stems and single-strands (1), respectively.

The remaining terms of Eq. S1 describe various pairwise, nonbonded interactions (1). U_{exc} represents the excluded volume interaction between the nonbonded CG beads to avoid their overlap. The interaction strength ε = 0.26 kcal/mol, and σ_0 is the sum of the radii of two interactional beads. U_{bp} is the orientation-dependent base-pairing interaction between bases in the canonical Watson-Crick base pairs (G-C and A-U) and the wobble base pairs (G-U), and the interaction strength $\varepsilon_{AU} = \varepsilon_{GU} = 0.5\varepsilon_{GC}$ (1). U_{bs} is the base-stacking interaction between two nearest neighbour base pairs, and the sequence-dependent strength of base-stacking energy $G_{i,i+1,j-1,j}$ can be estimated from the combination of the experimental thermodynamics parameters (2,3) and Monte Carlo algorithm; more details can be found in Ref. 1.

U_{el} is the electrostatic interaction between phosphate groups given by the Debye-Hückel approximation, and based on the counterion condensation theory (4) and the tightly bound ion model (5-7), the reduced charge fraction Q could be written as $Q = f_{\mathrm{Na^+}}\left(\frac{b}{l_B}\right) + (1 - f_{\mathrm{Na^+}})\left(\frac{b}{2l_B}\right)$ for mixed Na^+/Mg^{2+} solutions, where f_{Na^+} and 1-f_{Na^+} are the contribution fraction from Na^+ and Mg^{2+}, respectively. b is the phosphate-phosphate spacing of an RNA and is taken as 5.5Å (1). The empirical formula

$$f_{\mathrm{Na^+}} = \frac{[\mathrm{Na^+}]}{[\mathrm{Na^+}] + x[\mathrm{Mg^{2+}}]} \tag{S10}$$

derived by the tightly bound ion model is used for mixed divalent/monovalent ion solutions, and $x = (8.1 - 64.8/N)(5.2 - \ln[\mathrm{Na^+}])$, where $[Na^+]$ and $[Mg^{2+}]$ are the corresponding concentrations in molar (M) and N is the chain length (5-7).

Similarly to U_{bs}, U_{cs} is also the base-stacking interactions but between the bases which are the interfaces of two discrete stems, and the thermodynamics parameters from UV melting studies have shown that the coaxial stacking between two stems interrupted by one or more nucleotides is

approximately the same stable as the corresponding nearest-neighbour interaction in an uninterrupted helix of equivalent sequence (8,9). Since the noncanonical base pairs neglected by the present model are common when more than one base exists in each single-stranded chain between two discrete stems, we only consider the coaxial stacking for the cases that two stems are interrupted by less than two nucleotides in at least one side (see Fig. S1a) (8,9). To obtain the corresponding geometric constraint a and r_{cs}, we have calculated the distance distribution (Fig. S1b) between interfaced bases in coaxial stacked stems of known large structures including rRNAs, tRNAs and ribonuclease P RNAs.

2. Parameters

The initial parameters of the above described potentials are directly obtained from the statistical analysis on the known structures. The final parameters listed in Table SI and Table SII are derived through the comparisons between the predictions by the model and the experimental data (see more details in Ref. 1). All the parameters with the exception of ones newly introduced for U_{el} and U_{cs} are the same as the previous version of our model (1).

Table SI. The parameters of bonded potentials in Eqs. S2-S4.

Bond U_b				
	K_b (kcal/mol)		r_0 (Å)	
	$Para_{helical}$ [a]	$Para_{nonhelical}$ [b]	$Para_{helical}$	$Para_{nonhelical}$
P_iC_i	133.4	98.2	3.95	3.95
C_iP_{i+1}	75.0	42.5	3.93	3.93
C_iN_i	85.6	24.8	3.35	3.45
Angle U_a				
	K_θ (kcal/mol)		θ_0 (rad)	
	$Para_{helical}$	$Para_{nonhelical}$	$Para_{helical}$	$Para_{nonhelical}$
$P_iC_iP_{i+1}$	18.3	9.3	1.74	1.75
$C_{i-1}P_iC_i$	43.9	21.3	1.76	1.78
$P_iC_iN_i$	35.5	9.7	1.63	1.64
$N_iC_iP_{i+1}$	99.8	15.2	1.66	1.66
Dihedral U_d				
	K_φ (kcal/mol)		φ_0 (rad)	
	$Para_{helical}$	$Para_{nonhelical}$	$Para_{helical}$	$Para_{nonhelical}$
$P_iC_i\ P_{i+1}C_{i+1}$	2.8	1.1	2.56	2.51
$C_{i-1}P_iC_iP_{i+1}$	10.5	4.3	-2.94	-2.92
$C_{i-1}P_iC_iN_i$	3.8	0.8	-1.16	-1.18
$N_{i-1}C_{i-1}P_iC_i$	4.2	0.7	0.88	0.78

[a] The $Para_{helical}$ only are used in the processes of folded structure refinement for the base-pairing regions (stems) in the initially folded structure. [b] The $Para_{nonhelical}$ are used in RNA folding processes to possibly describe RNAs as free chains.

Table SII. The parameters of nonbonded potentials in Eqs. S5-S9.

Nonbonded				
U_{exc}	ε (kcal/mol)	0.26	σ_{st} (Å)	$R_i^a+R_j$
U_{bp}	$\varepsilon_{bp(\mathrm{GC})}$ (kcal/mol)	-3.5	$\varepsilon_{bp(\mathrm{AU/GU})}$ (kcal/mol)	-1.75
	k_{NN} (Å^{-2})	3.6	r_{NN} (Å)	8.9
	k_{CN} (Å^{-2})	1.9	r_{CN} (Å)	12.2
	k_{PN} (Å^{-2})	0.7	r_{PN} (Å)	13.9
U_{bs}	$G_{i,i+1,j-1,j}$ (kcal/mol)	Sequence-dependent[b]	σ_{st} (Å)	4.8
U_{el}	Q	Salt-dependent	b (Å)	5.5
U_{cs}	$G_{i\text{-}j,k\text{-}l}$ (kcal/mol)	Sequence-dependent[b]	a (Å^{-1})	0.4
			r_{cs} (Å)	5.0

[a] The radius of bead i, and the radii of P, C and N beads are 1.9 Å, 1.7 Å and 2.2 Å, respectively. [b] For different combinations of bases, $G_{sequence} = \Delta H - T(\Delta S - \Delta S_c)$, where ΔH and ΔS are the RNA thermodynamics parameters derived from the corresponding experiments for base-pair stacking (2,3) or coaxial stacking (8,9), and ΔS_c is equal to -9.7eu (see Ref. 1).

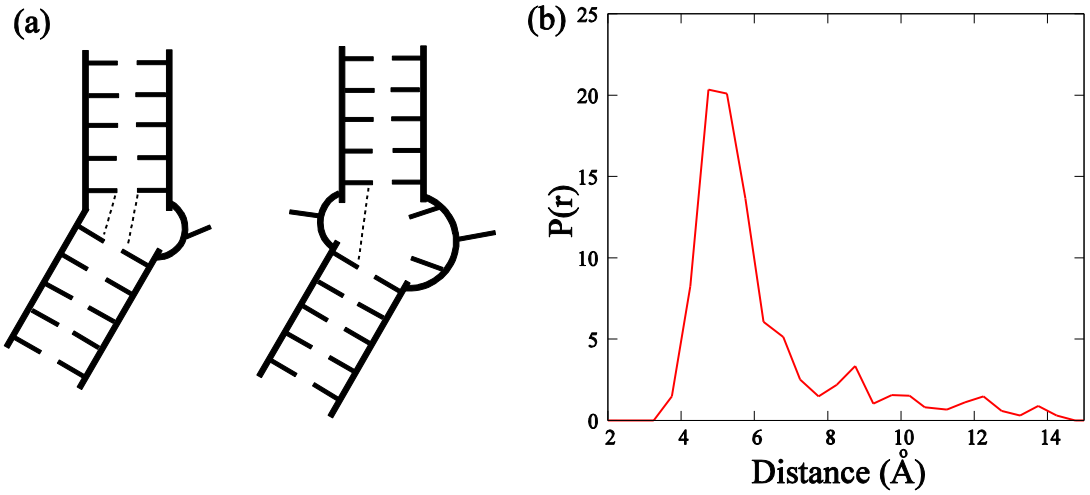

Figure S1. (a) A schematic representation of the coaxial stacking potential used in the model. (b) The normalized probability distribution $P(r)$ of the distances between two interfaced bases in coaxially stacked stems (dash line in (a)), which is obtained by the statistical analysis over the known structures of 18 large RNAs (PDB code: 1c2w, 1njn, 1p9x, 1y0q, 1ffk, 1fg0, 2a64, 2o44, 3g78, 3ize, 3j2h, 3j3c, 3j3d, 3j5z, 3l0u, 3rg5, 4c4q, 4p5j).

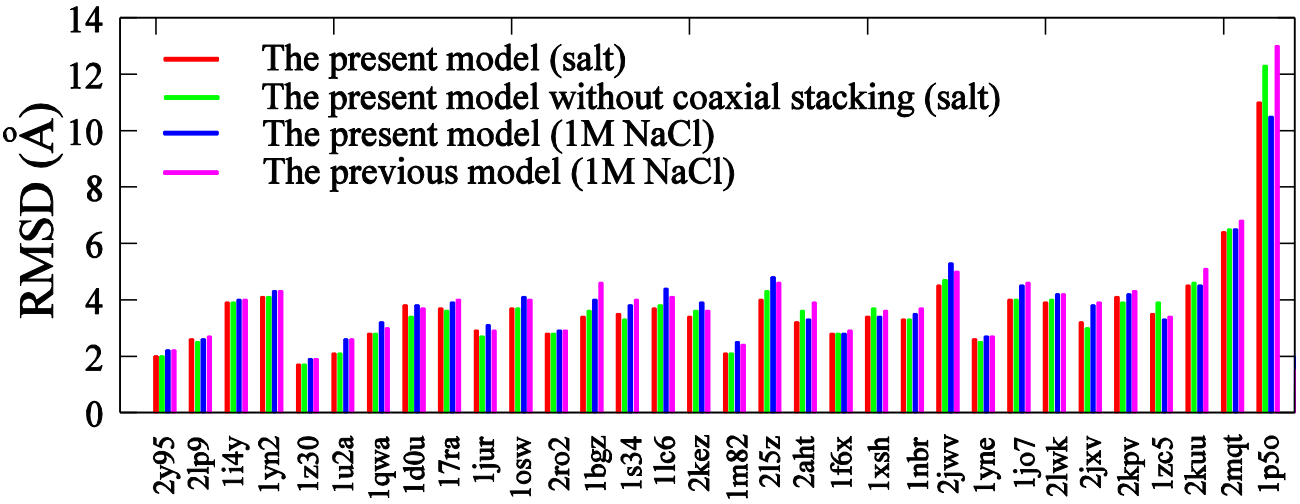

Figure S2. Comparison of the RMSDs of RNA 3D structures predicted by our present and previous models. The RMSDs of predicted structures are calculated over C beads from the corresponding C4' atoms in the native structures. Red: predicted by the present model with coaxial stacking at experimental ion conditions; green: predicted by the present model without coaxial stacking at experimental ion conditions; blue: predicted by the present model with coaxial stacking at 1M NaCl; pink: predicted by our previous model without coaxial stacking at 1M NaCl. For the four independent predictions, the overall mean RMSDs of 32 RNA hairpins are 3.64 Å, 3.71 Å, 3.89 Å and 4.02 Å, respectively.

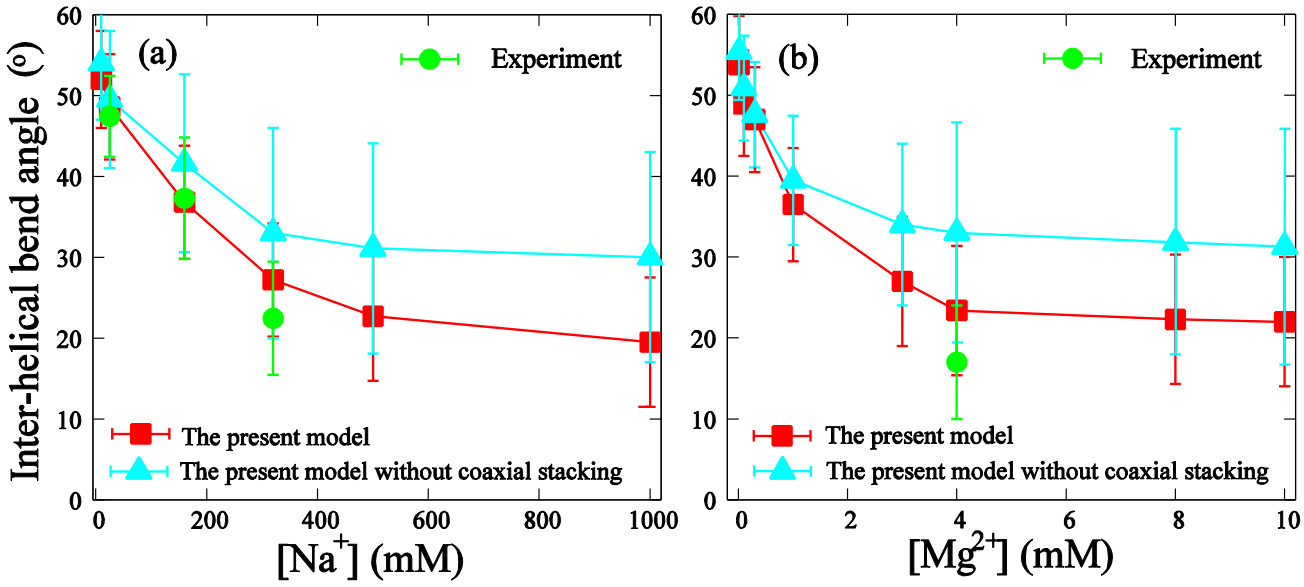

Figure S3. (a, b) The experimental (green; Ref. 10) and predicted inter-helical bend angle as functions of $[Na^+]$ (a) and $[Mg^{2+}]$ (b) for HIV-1 TAR variant (see Fig. 3). Red: predicted by the present model. Cyan: predicted by the present model without coaxial stacking.

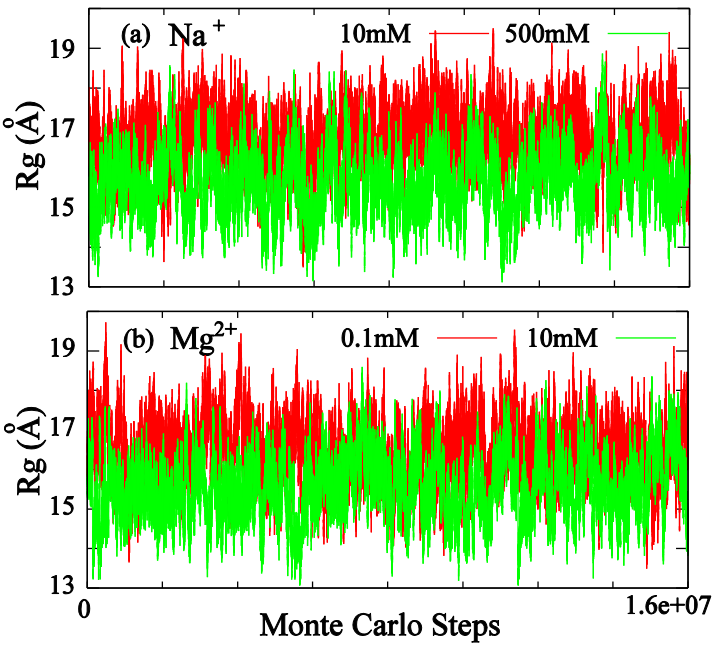

Figure S4. The time-evolution of the radius of gyration for HIV-1 TAR variant (shown in Fig. 3 in main text) at different Na^{+}/Mg^{2+} conditions.

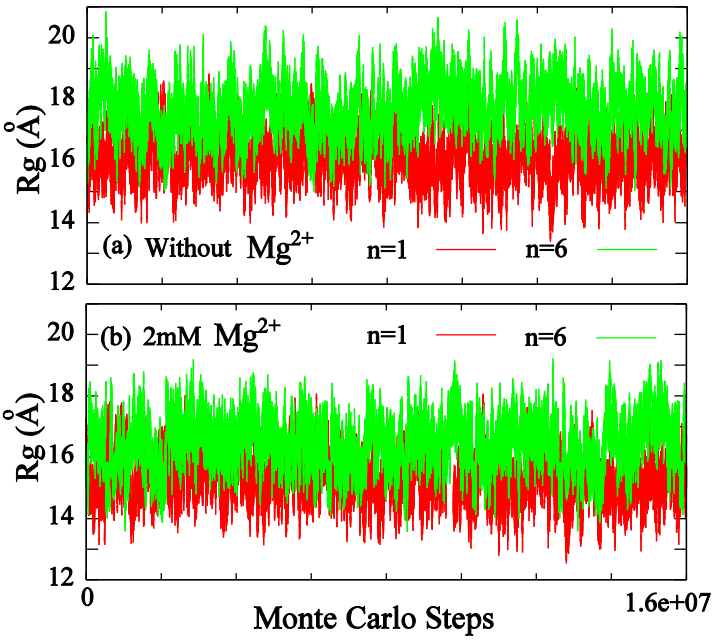

Figure S5. The time-evolution of the radius of gyration for HIV-2 TAR variant (shown in Fig. 3 in main text) with different lengths of bulge at 5 mM $NaPO_4$ without (a) or with (b) 2 mM Mg^{2+}.

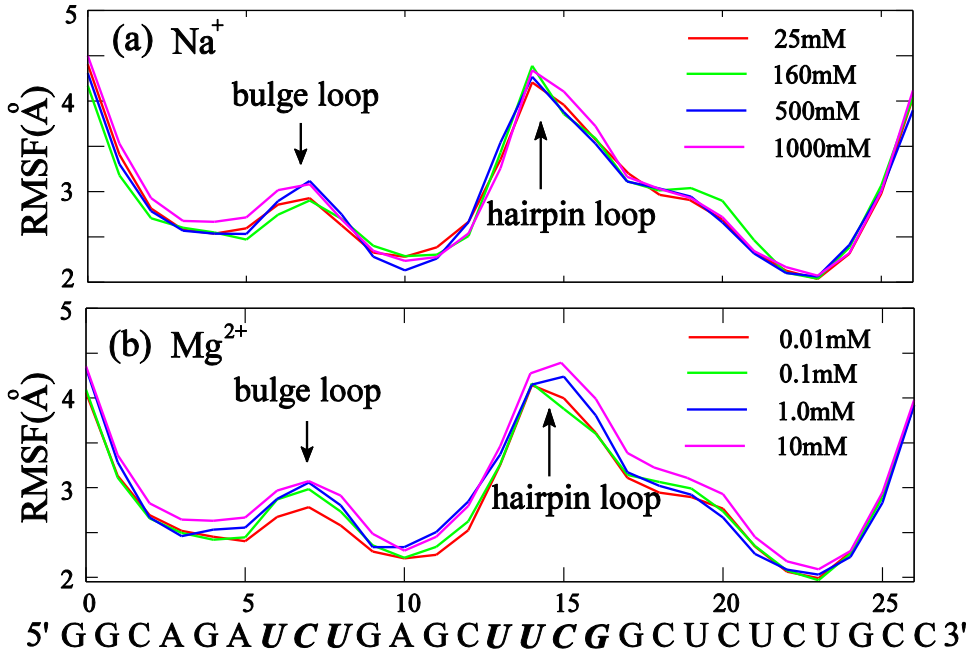

Figure S6. The RMSF for C-beads along HIV-1 TAR variant (shown in Fig. 3 in main text) in different Na^{+}/Mg^{2+} solutions. The sequences of loops are in italics.

SUPPORTING REFERENCES

1. Shi, Y. Z., F. H. Wang, Y. Y. Wu, and Z. J. Tan. 2014. A coarse-grained model with implicit salt for RNAs: predicting 3D structure, stability and salt effect. J. Chem. Phys. 141:105102.
2. Xia, T., J. SantaLucia, M. E. Burkand, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox, and D. H. Turner. 1998. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry. 37:14719-14735.
3. Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner. 1999. Expended sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. 288: 911-940.
4. Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. Q. Rev. Biophys. 11:179-246.
5. Tan, Z. J., W. Zhang, Y. Z. Shi, and F. H. Wang. 2015. RNA folding: structure prediction, folding kinetics and ion electrostatics. Adv. Exp. Med. Biol. 827:143-183.
6. Tan, Z. J., and S. J. Chen, 2007. RNA helix stability in mixed Na^+/Mg^{2+} solution. Biophys. J. 92:3615–3632.
7. Tan, Z. J., and S. J. Chen. 2005. Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. J. Chem. Phys. 122:44903.
8. Walter, A. E., D. H. Turner, J. Kim, M. H. Lyttle, P. Müller, D. H. Mathews, and M. Zuker. 1994. Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. Proc. Natl. Acad. Sci. 91:9218–9222.
9. Walter, A. E., and D. H. Turner. 1994. Sequence dependence of stability for coaxial stacking of RNA helixes with Watson-Crick base paired interface. Biochemistry. 33:12715–12719.
10. Casiano-Negroni, A., X. Sun, and H. M. Al-Hashimi. 2007. Probing Na^+-induced changes in the HIV-1 TAR conformational dynamics using NMR residual dipolar couplings: new insights into the role of counterions and electrostatic interactions in adaptive recognition. Biochemistry. 46:6525-6535.