A Novel Implicit Solvent Model for Simulating the Molecular Dynamics of RNA

Yufeng Liu,[†] Esmael Haddadian,[‡] Tobin R. Sosnick,[§] Karl F. Freed,[¶*] and Haipeng Gong[†*]

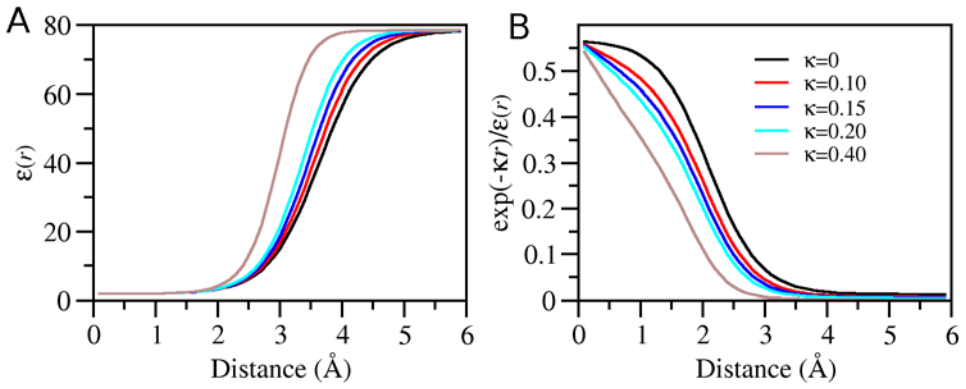

Figure S1. The influence of Debye-Hückle screening constant κ on the distance-dependent electrostatic screenings in our model. The dielectric function ε(r) and the overall screening term in electrostatic calculation (Eq. 13), exp(-κr)/ε(r), are plotted in the panel (**A**) and (**B**), respectively, for various κ values. κ = 0 (black), 0.1 (red), 0.15 (blue), 0.2 (cyan), and 0.4 (tan) correspond to the monovalent salt of 0, 0.1, 0.225, 0.4, and 1.6 M respectively.

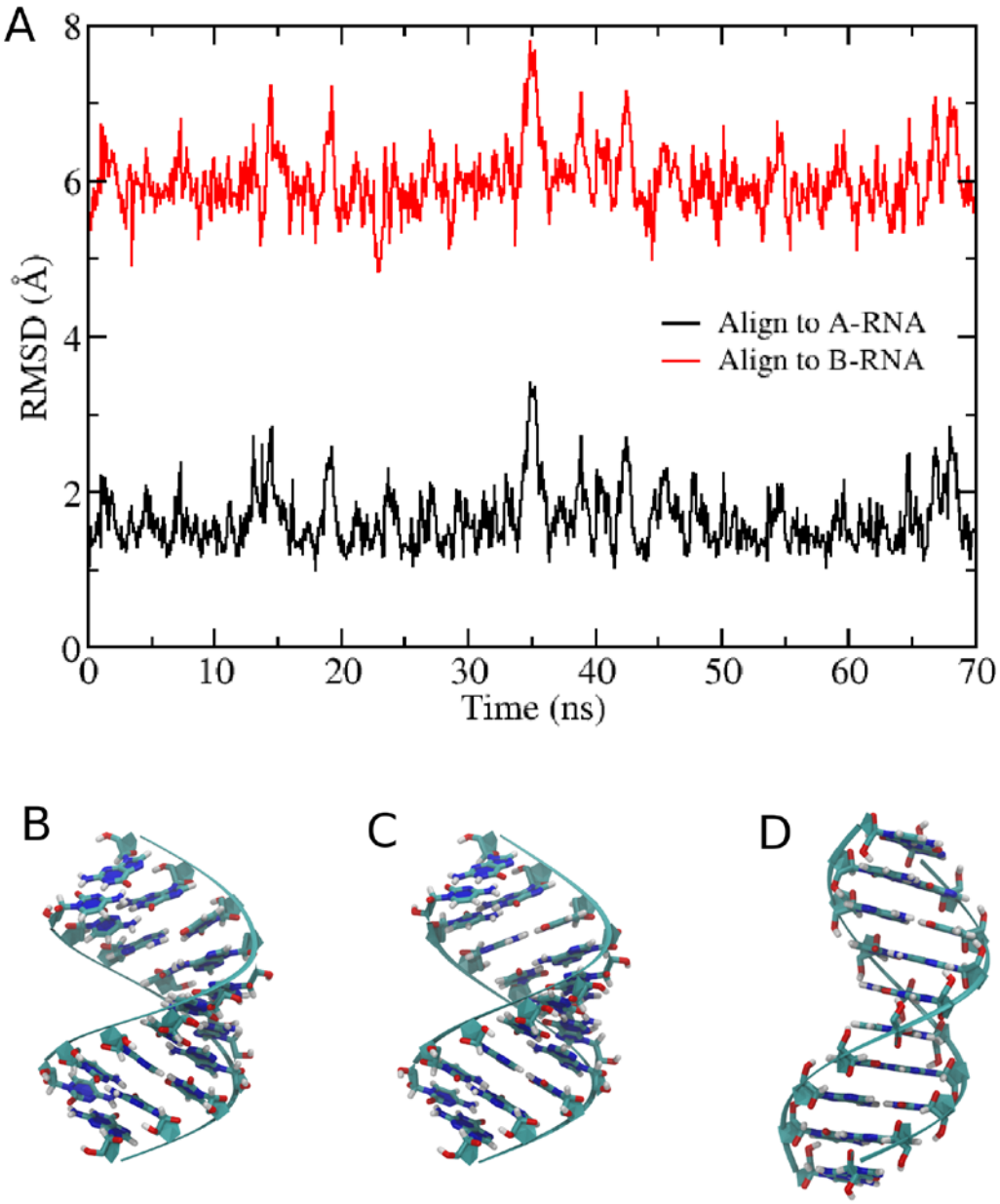

Figure S2. The RNA duplex retains the A-form conformation in the implicit solvent simulation. **(A)** The trajectory RMSDs of the simulated molecules when aligned to the canonical A-form (black) and B-form (red) conformations respectively. **(B-D)** The structural snapshot taken from the implicit solvent simulation (**B**) shows higher degree of structural similarity to the canonical A-form RNA duplex (**C**) than to the canonical B-form RNA duplex (**D**).

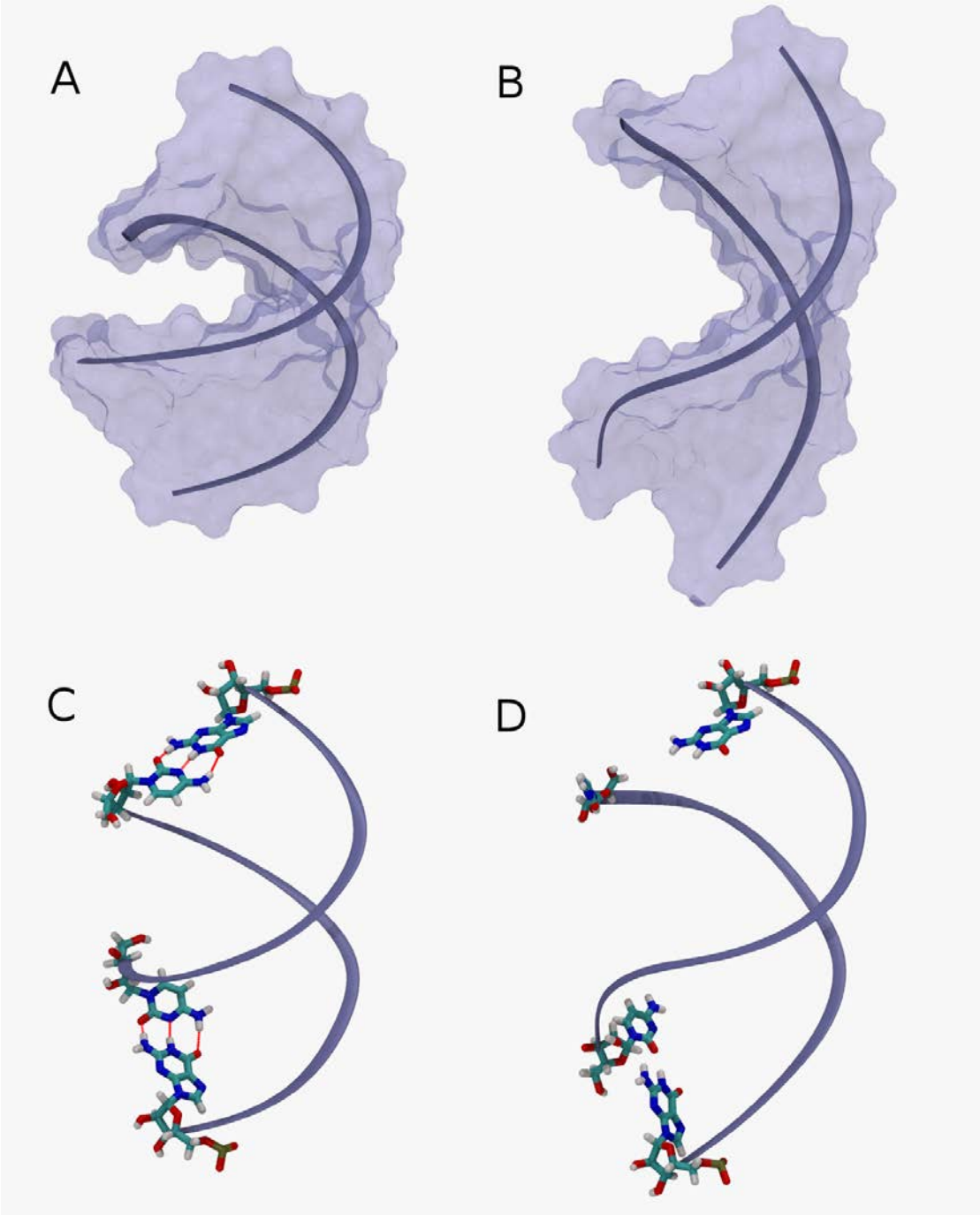

Figure S3. Duplex bending and the disruption of terminal G-C base pairs in the explicit solvent simulation of the A-form RNA duplex. Both bent (**A**) and straight (**B**) helical conformations are observed in the trajectory for frames 247 and 370, respectively, in the explicit solvent simulation. The terminal G-C pairs are preserved (**C**) in the early stage (frame 0) but become broken (**D**) at late stages (frame 460) in the explicit solvent simulation. All unrelated atoms are neglected to emphasize the terminal nucleotides.

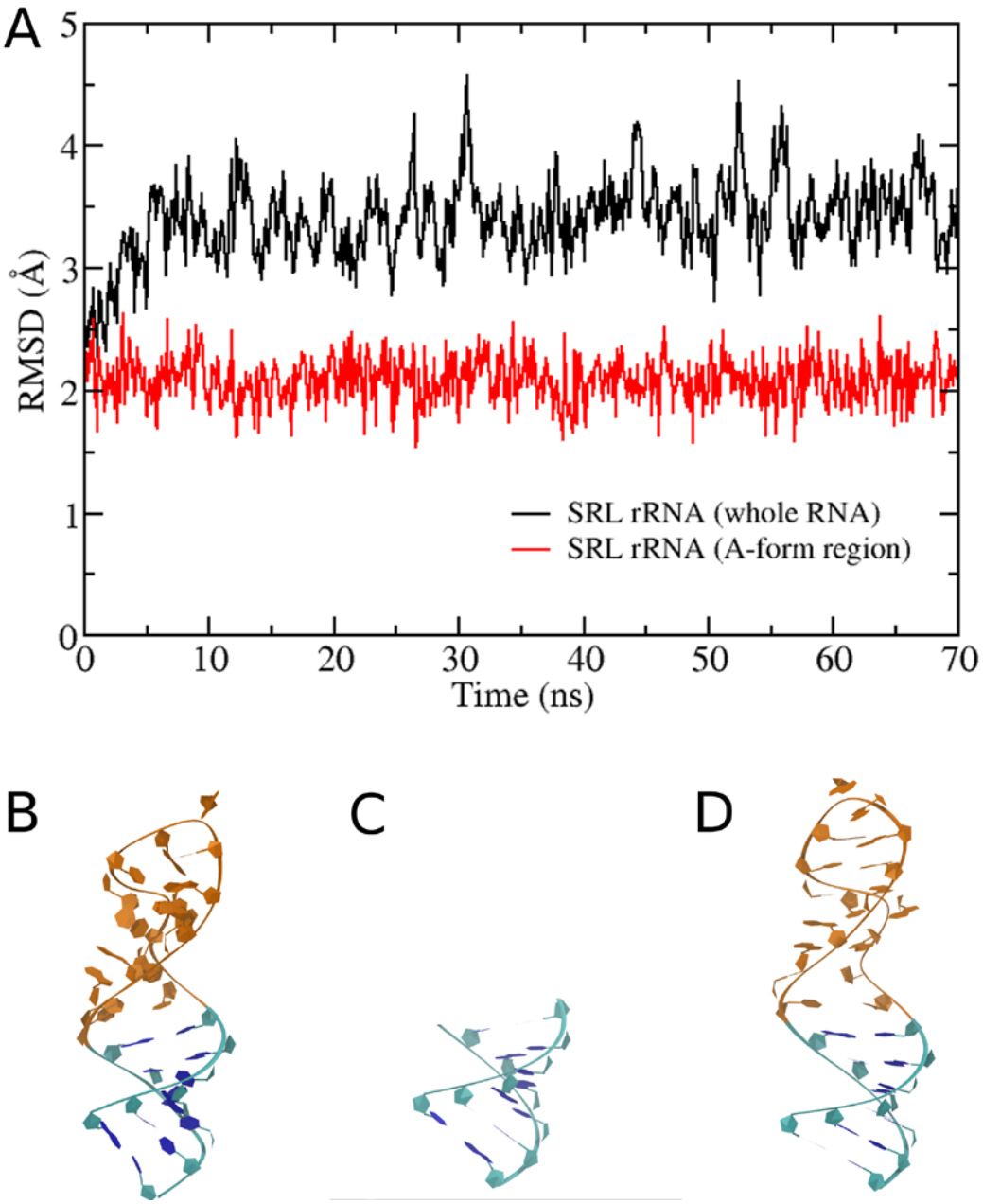

Figure S4. The Watson-Crick region of the SRL rRNA retains the A-form helical conformation in the implicit solvent simulation. (**A**) The trajectory RMSD of the Watson-Crick region (red) is greatly lower than that of the overall RNA molecule (black). **(B-D)** The structural snapshot taken from the implicit solvent simulation (**B**) contains a stable Watson-Crick region (cyan), which is structurally similar to both the canonical A-form RNA duplex (**C**) and the corresponding region in the crystal structure of the SRL rRNA (**D**).

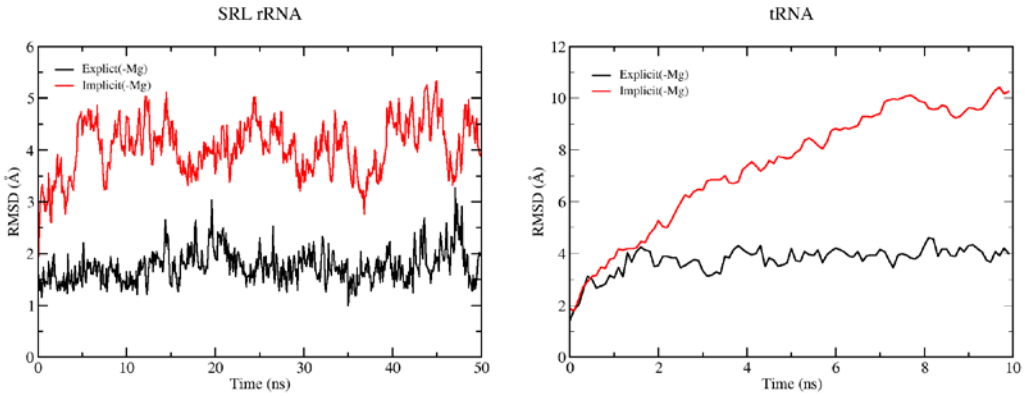

Figure S5. RMSD time series for the explicit (black) and implicit (red) simulations without explicit Mg^{2+} ions. Both SRL rRNA (left) and tRNA (right) simulations for implicit systems yield RMSDs remarkably larger than for explicit ones, which indicates the necessity to include Mg^{2+} in our implicit solvent model.

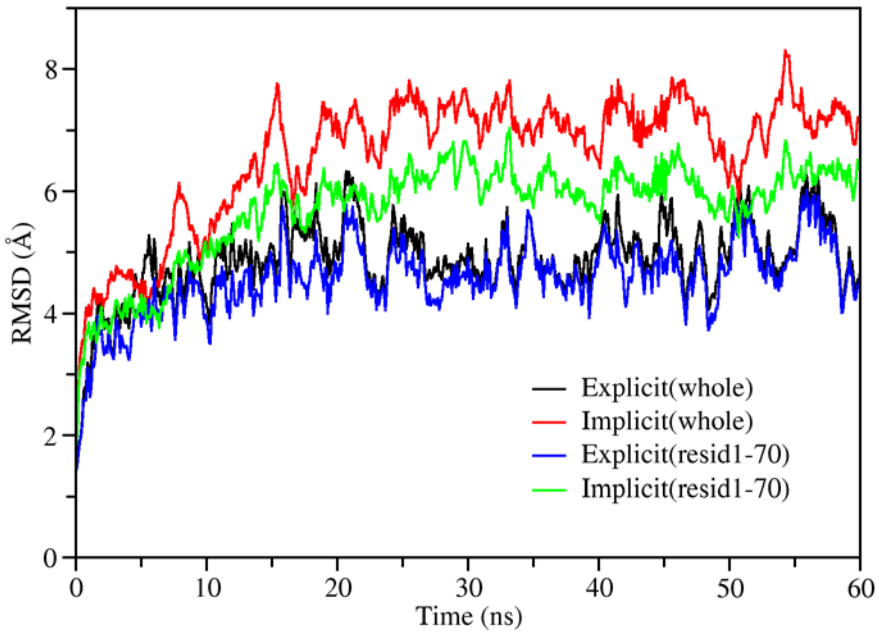

Figure S6. RMSD time series for tRNA with or without the five unpaired nucleotides located at the 3'-end. As labeled in the figure, the RMSD of entire molecule in implicit solvent (red) significantly decreases by ~1 Å when disregarding the residues at the 3'-end (green). On the contrary, explicit solvent simulations display little reduction of the RMSD if the unpaired terminal residues of the 3'-end are neglected (blue vs. black).

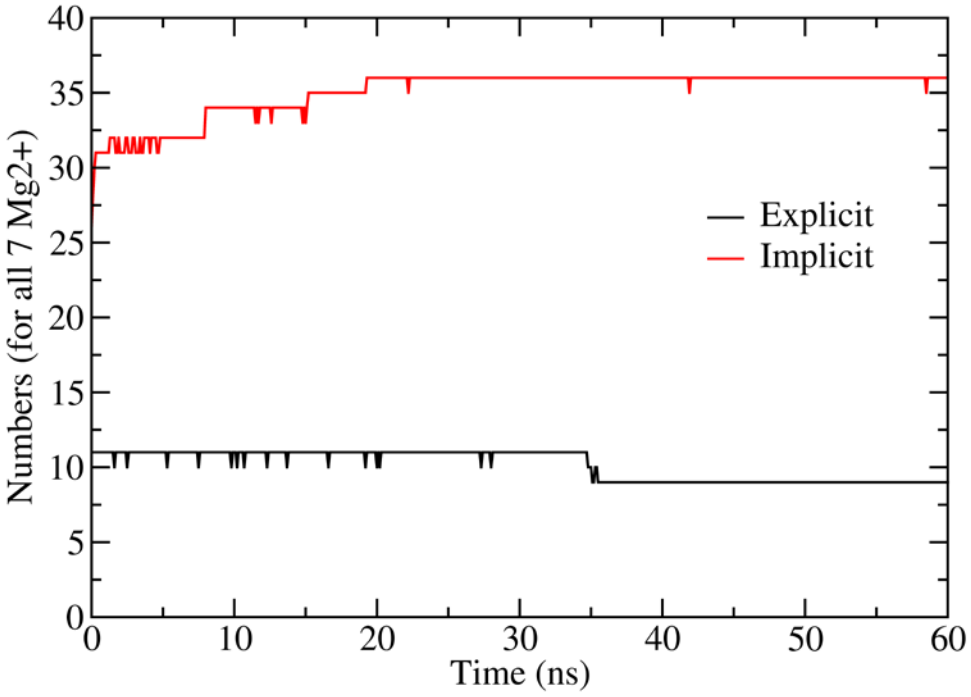

Figure S7. The overall coordination numbers around the seven Mg^{2+} in the simulation of tRNA using implicit (red) and explicit (black) solvent.

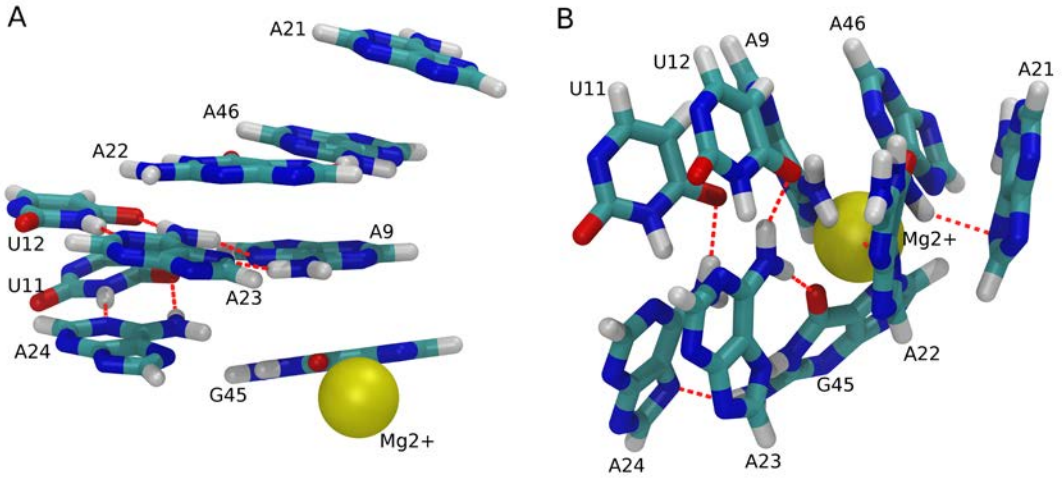

Figure S8. Structural disruption by Mg^{2+} in the implicit solvent simulation. Nine nucleotides (nucleotides A9, U11, U12, A21, A22, A23, A24, G45 and A46) as well as a nearby Mg^{2+} ion (residue index 79) are displayed here. (**A**) In the crystal structure, the Mg^{2+} stays besides the nucleotides and does not perturb the two base triplets (A9-A23-U12 and A9-U11-A24) which are stabilized by 7 hydrogen bonds (red dashed lines). (**B**) After 50 ns simulation, the ion is squeezed into the nucleotide core and thus breaks the previous triplets, leading to the reformation of two weakly interacting triplets (U12-A23-G45 and U11-A24-G45) with only 4 connecting hydrogen bonds. Thus, the overall orientation changes greatly (both A and B are taken from structurally aligned tRNA conformations). Additionally, the A9-G45 stacking is disrupted, leaving G45 unpacked, and A21 and G22 are also not well packed after simulation.

Table S1. Weight g_i of each atom type i for evaluating the solvation energy with the SASA model. Weights are obtained from fits to solvation energies.

Atom	g_i	Atom	g_i
C	0.001624	O2	-0.033022*
CA	-0.014606	O2-	-1.402429*
CT	-0.005197	OA	-0.075714
N	-0.093329	OH	-0.078084
N2	-0.065938	OS	-0.084605
NC	-0.085392	P	0.007208
O	-0.069924		

The atom-type denotations are as follows:

C: carbonyl carbon

CA: conjugated carbon in the base

CT: sp3 hybridized carbon

N: glycosyl nitrogen

N2: NH2

NC: conjugated nitrogen in the base

O: carbonyl oxygen

O2: double bonded oxygen in carboxylate or phosphate

O2-: negatively charged oxygen in the carboxylate or phosphate

OA: carbonyl oxygen in the base

OH: hydroxyl oxygen

OS: ether oxygen

P: phosphus.

*Since O2 and O2- are chemically identical, their g_i values are set to be the average of the fitted data, which gives g_i of -0.7177255.

Table S2. Comparison of our model with various GB models for simulations of tRNA.

System	RMSD in the last 1ns (Å)	
	With Mg^{2+}	Without Mg^{2+}
Our model	5.13±0.16	9.94±0.32
GB_{HCT}	14.76±0.75	15.92±0.39
GB_{OBC1}	20.87±0.92	19.61±0.76
GB_{OBC2}	14.51±0.74	17.17±0.58

RMSDs are calculated with respect to the crystal structure of tRNA for structural snapshots in the last 1 ns of the 10 ns trajectories with and without Mg^{2+} ions.

Table S3. The model compounds and their experimental solvation energies used in our model to derive the best estimates of g_i values for various atom types in RNAs.

Model compound	Solvation energy (kcal/mol)
trimethylphosphate	-8.70
triethylphosphate	-7.80
tripropylphosphate	-6.10
dihydrogenphosphate	-68.00
phosphine	0.60
pyridine	-4.70
4-methylpyridine	-4.94
4-ethylpyridine	-4.74
2-methylpyrazine	-5.57
2-ethylpyrazine	-5.51
aniline	-5.49
4-methylaniline	-5.55
3-aminoaniline	-9.92
benzene	-0.87
1-methylthymine	-10.40
9-methyladenine	-13.60
methanol	-5.11
ethanol	-5.01
cyclopentanol	-5.49
dimethylether	-1.92
diethylether	-1.76
tetrahydrofuran	-3.47
tetrahydropyran	-3.12
2-methoxyethanol	-6.77
acetamide	-9.71
N-methylacetamide	-10.00
benzamide	-10.90
urea	-13.80
1,4-dioxane	-5.05
acetone	-3.85
2-butanone	-3.64
ethanal	-3.50
propanal	-3.44
formic	-78.00
acetate	-80.00
aceticacid	-6.70
propionate	-78.00
propionicacid	-6.47
benzoicacid	-7.90
benzoate	-73.00
methane	2.00

ethane	1.83
cyclopentane	1.20