Supporting information for: Evidence for the Role of Active Site Residues in the Hairpin Ribozyme from Molecular Simulations along the Reaction Path

Hugh Heldenbrand,[†] Pawel Janowski,[‡] George Giambaşu,[‡] Joseph E. Wedekind,[¶] and Darrin M. York[*,‡]

Department of Chemistry, University of Minnesota, 207 Pleasant Street SE, Minneapolis, MN 55455, USA, Center for Integrative Proteomics Research, BioMaPS Institute for Quantitative Biology and Department of Chemistry & Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA, and Department of Biochemistry and Biophysics, Center for RNA Biology, University of Rochester School of Medicine & Dentistry, 601 Elmwood Avenue, Box 712, Rochester, New York 14642, USA

E-mail: york@biomaps.rutgers.edu

Supporting Methods

We have completed three simulations of the RNA in a pre-catalytic conformation based on the crystal structure[1] (PDB ID 2OUE). Two 500 ns simulations, in the reactant state, have the A-1:O2′ nucleophile protonated, with A38 both in the neutral (R-A38^0) and protonated (R-A38$^+$) states. In addition, we completed a 150 ns activated precursor simulation with the A-1:O2′ de-protonated, the scissile phosphate protonated on the proR oxygen (G+1:O2P), and A38 protonated

[*]To whom correspondence should be addressed

[†]Department of Chemistry, University of Minnesota, 207 Pleasant Street SE, Minneapolis, MN 55455, USA

[‡]Center for Integrative Proteomics Research, BioMaPS Institute for Quantitative Biology and Department of Chemistry & Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA

[¶]Department of Biochemistry and Biophysics, Center for RNA Biology, University of Rochester School of Medicine & Dentistry, 601 Elmwood Avenue, Box 712, Rochester, New York 14642, USA

(AP-A38$^+$-O2P$^+$). We also performed two simulations, each 150 ns in length, of the hairpin ribozyme with transition state (TS) mimics in the active site and A38 in the protonated state. One (TS-P(V)-A38$^+$) has a pentacoordinate, dianionic scissile phosphate based on the vanadate TS mimic crystal structure[2] (PDB ID 2P7E). The other (TS-2'-5'-A38$^+$) is a TS mimic containing a 2' to 5' linkage of the scissile phosphate group, based on the crystal structure[3] (PDB ID 3CQS). Finally, we have performed an 85 ns simulation of the vanadate TS mimic-based structure crystal (TS-P(V)-A38$^+$-X) with 12 hairpin monomers modeled explicitly and arranged according to the experimental crystal space group symmetry.

Simulations were performed with NAMD version 2.9[4] (TS-P(V)-A38$^+$-X was performed with Amber12 CUDA[5]) in the NPT ensemble using the AMBER parm99 force field with the corrected α/γ torsional parameters[6] and sodium and chloride ions parametrized[7] for use with the TIP4P-ew[8] water model. We developed parameters for non-standard residues according to a protocol based on the development of the original AMBER parameter set.[9] Simulation temperature was maintained at 300 K using Langevin dynamics, with a damping coefficient of 1 ps^{-1}. Pressure was controlled using a Langevin piston, with a target pressure of 1 atm, period of 100 fs and decay time of 50 fs.

Each RNA simulation was initialized with a 15 Å truncated octahedral buffer of water molecules and enough Na^+ ions to neutralize the system. Na^+ and Cl^- ions were then added at random positions to bring the solvent concentration to 0.14 M. For the crystal simulations, enough solvent was added to ensure that unit cell volume would maintain the experimental value to within 0.3%.[10] Restraints based on crystallographic B-values were used on the RNA during two initial rounds of simulated annealing intended to accelerate equilibration of the ionic environment. Logarithmic removal of restraints and 10 ns of additional equilibration were performed prior to production simulation. The equilibration protocol used here has been applied to other ribozymes, and has been described in detail elsewhere.[11]

Charge parameters for protonated adenine were developed according to the following scheme: First, two sets of coordinates were obtained for 9-methyladenine, one protonated at N1 and

one neutral using the default structures available in Gaussview. Then each structure was geometry optimized at the HF/6-31G(d) level, and a set of electrostatic potentials was computed at various spherical gridpoints from each atom center using default density settings in the g09 electronic structure program. This set of electrostatic potentials was then used as the basis for computing atom-centered point charges using the restrained electrostatic potential (RESP) method implemented in AmberTools. During the RESP procedure, a restraint was placed on the 4 atoms that compose the methyl group attached to the N9 atom. These atoms were constrained to have a total charge that is the opposite of the sum of the charges for the adenine nucleobase of the adenosine nucleotide in the original (parm99) AMBER parameter set. This ensured that any derived charges could be added back onto the sugar/phosphate groups of the original AMBER parameter set with the correct total charge.

After point charges were computed both for the neutral and protonated 9-methyladenines, the difference in charge was computed for each atom in the system. This difference was then applied to the charges for the original AMBER parameter set to determine the new set of charges. Bond, angle and van der Waals parameters for the new hydrogen at position N1 were determined by analogy to other parameters in the set. The final set of charges is given in Table S1.

Atom types (which determine bond, angle, torsion and van der Waals parameters) were maintained at the standard (neutral) values for each atom in the protonated residue, with two exceptions. New van der Waals parameters were derived for the N1 and N6 (exocyclic amine) nitrogens to better reproduce interactions between the atoms (which now carry a higher charge) and water. The first step of the derivation of the new van der Waals parameters was to compute the binding energy of a single water molecule to the N1 and N6 atoms of both neutral and protonated 9-methyladenine at the M06-2X/6-311++G(3df,2p) level. The 9-methyladenine/water complexes were geometry optimized prior to computing the binding energy. The differential binding energy between the neutral and protonated residues was then computed, and used as the target for a chi-squared fit to determine new van der Waals parameters, using MM differential binding energies computed with AMBER. The MM water molecule was restrained to have rigid bond lengths and angles. If the van

der Waals potential between two atoms i and j is written as in equation Eq. (1) then only the "A" parameters were varied in our optimization.

$$V_{LJ} = \frac{A_i A_j}{r_{ij}^{12}} - \frac{B_i B_j}{r_{ij}^{6}} \tag{1}$$

Our optimized N1 van der Waals A parameter was 68% larger than the A parameter in a neutral parm99 adenine N1, and the N6 van der Waals parameter was 24% larger.

Supporting Tables

Table S1: Computed point charges for the protonated adenine residue used in this work.

H8	0.1965
N9	0.0961
N3	-0.5201
C8	0.2011
C2	0.4435
H61	0.4403
C6	0.5845
H62	0.4403
N1	-0.5776
C5	0.1136
N6	-0.8152
N7	-0.5569
H2	0.1307
C4	0.2681
H1	0.4310

Table S2: Comparison of TS-P(V)-A38$^+$-X, TS-P(V)-A38$^+$ and experimental structure results. Experimental parameters ("Exp.") are taken from PDB:2P7E.[2] Average solution ("Sol.") are parameters of the average structure from the solution simulation (TS-P(V)-A38$^+$). Average crystal ("Cryst") are parameters of the average structure from the crystal simulation (TS-P(V)-A38$^+$-X). Inspection of A38 active site geometry reveals four outliers in the set of twelve asymmetric units explicitly modeled in the crystal simulation (noted with asterisks). Average crystal without outliers ("Cryst.*") are parameters of the crystal simulation average structure calculated over all asymmetric units excluding outliers. Listed below are the parameters of the average structure from each individual asymmetric unit in the crystal simulation.RMSD values are provided for all heavy atoms in the hairpin monomer ("monomer") and for active site heavy atoms only ("active"). Key active site distances are given in Angstrom and degrees.

	RMSD		genAcid (A38)			genBase (G8)		
	monomer	active	H-O	N-O	N-H-O	H-O	N-O	N-H-O
Exp.	–	–	1.71	2.62	149.1	2.11	2.91	136.2
Sol.	2.965	1.543	2.02	2.95	151.6	2.38	3.15	136.2
Cryst.	0.890	0.581	2.69	3.55	146.0	1.88	2.81	162.5
Cryst.*	1.124	0.602	2.08	3.02	154.9	1.99	2.89	153.4
ASU01	1.264	0.936	*3.577	4.354	136.5	1.812	2.764	160.3
ASU02	1.274	0.755	2.046	2.974	153.4	1.930	2.863	155.4
ASU03	1.485	0.946	*4.333	5.037	130.5	1.887	2.879	177.3
ASU04	1.256	0.613	2.299	3.201	149.5	1.835	2.799	163.0
ASU05	1.663	0.708	2.094	3.043	157.8	1.981	2.896	152.4
ASU06	1.333	0.623	2.066	2.988	152.4	1.935	2.856	153.2
ASU07	1.246	0.695	2.034	2.965	154.1	1.950	2.872	153.4
ASU08	1.659	1.054	2.083	3.035	158.5	2.653	3.394	131.8
ASU09	1.189	1.327	*4.042	4.771	132.2	2.003	2.921	151.8
ASU10	1.526	0.715	2.042	2.974	154.2	1.866	2.807	157.5
ASU11	1.068	0.961	*4.011	4.792	137.2	1.831	2.806	165.8
ASU12	1.306	0.711	2.022	2.978	159.1	1.921	2.857	155.7

Supporting Figures

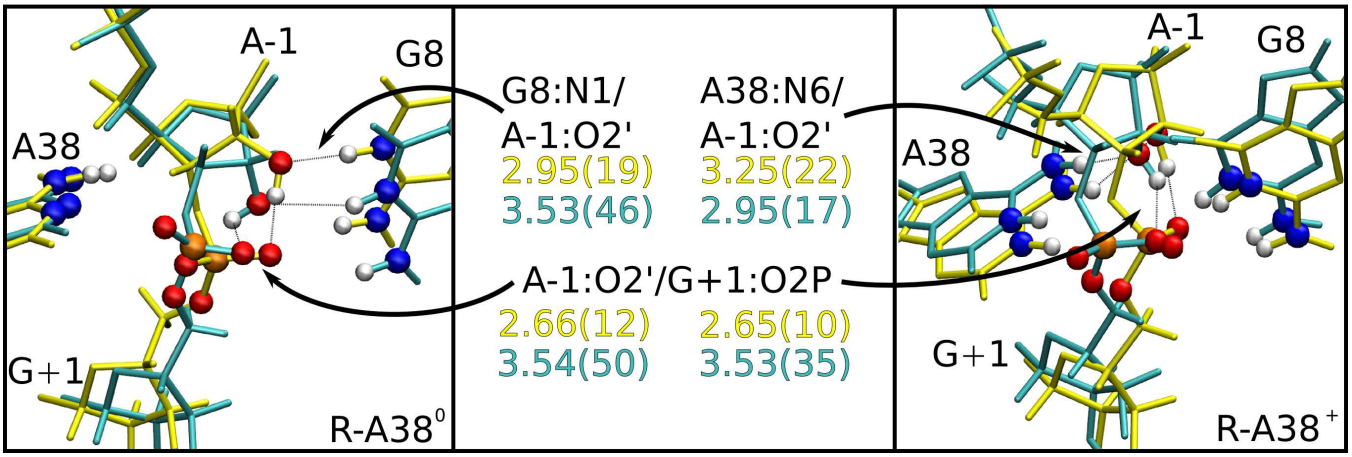

Figure S1: Overlays of the average active in-line attack (yellow) and inactive (cyan) active site structures from the R-A38^0 (left) and R-A38$^+$ (right) simulations. Distance values follow the same color scheme, with units in Å. In these all-RNA alignments, A38 and G8 are essentially superimposed, while the A-1 sugar pucker in R-A38^0 and the scissile phosphate in R-A38$^+$ adopt very different conformations in the active and inactive forms. Here we define an "active" in-line attack geometry to be one that has a O2'-P'O5' angle (θ_{inl}) more than 125 degrees and O2'-P distance (D_{inl}) less than 3.5 Å.

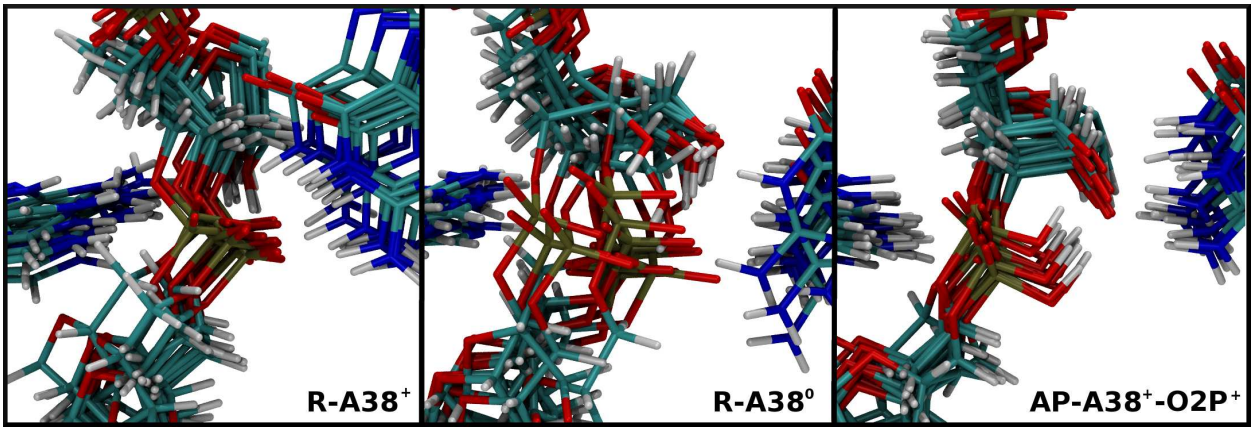

Figure S2: Overlays of 12 randomly selected active structures from each of the three precatalytic simulations. Structures are aligned based on all RNA atom coordinates. In the R-A38^{+} and AP-A38^{+}-O2P^{+} simulations, interactions between A38 or A-1:O2' and the scissile phosphate significantly reduce the available conformational space. In the R-A38^{0} simulation these interactions are absent, allowing much more flexibility to the scissile phosphate.

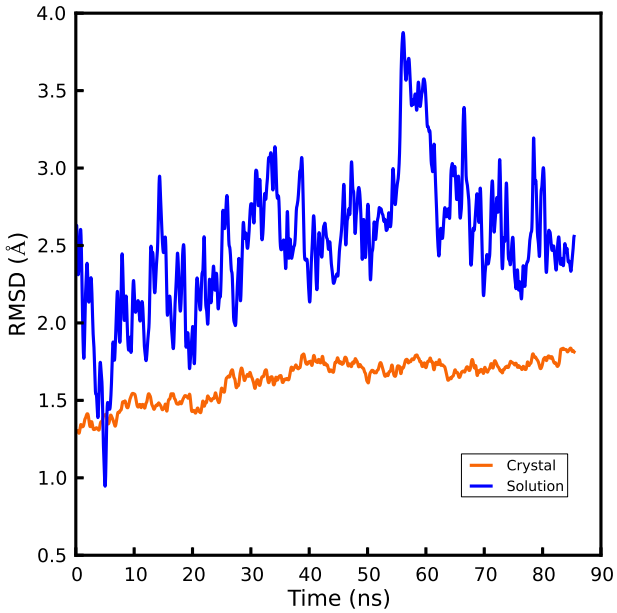

Figure S3: Comparison of solution (TS-P(V)-A38^{+}) and crystal (TS-P(V)-A38^{+}-X) simulation RMSD values. RMSD is measured after optimally aligning (quaternion method) all heavy atoms in the structure at each trajectory snapshot to the experimental structure.

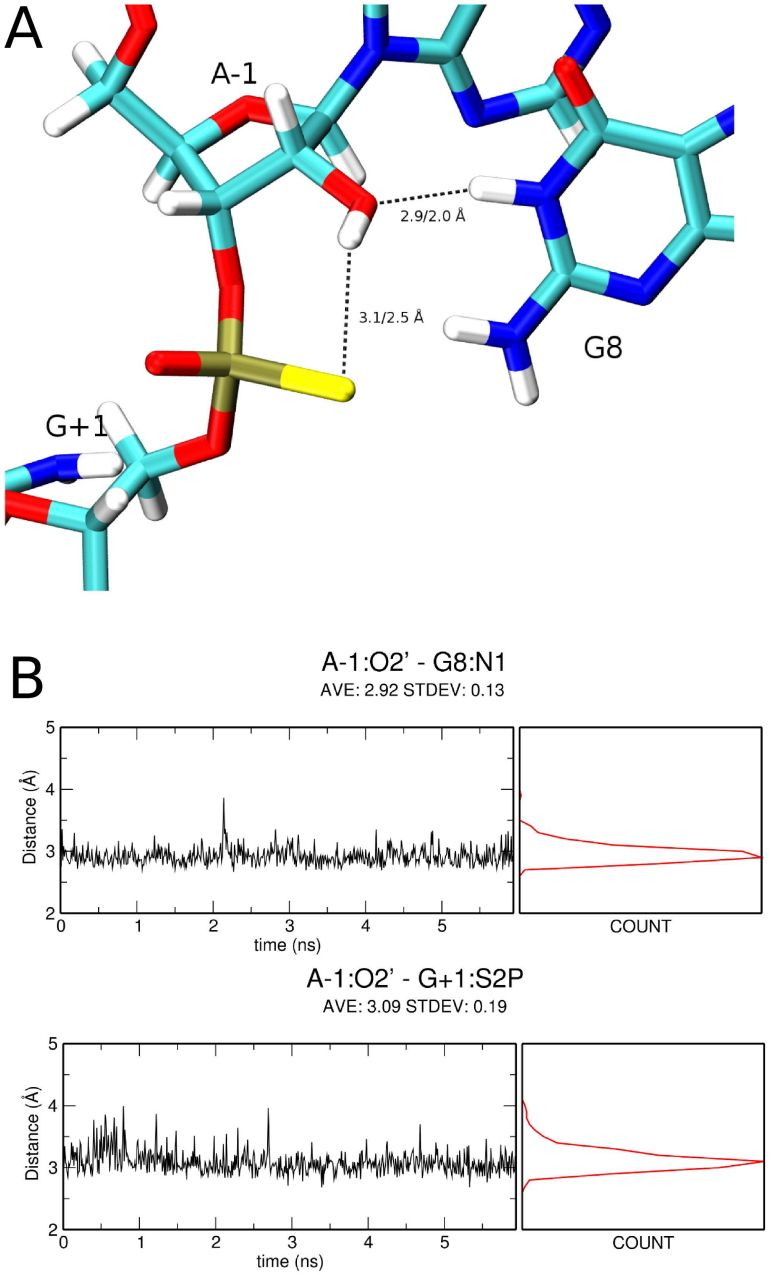

Figure S4: Impact of the proR-thio substitution at the scissile phosphate on the active conformation of the reactant state. A molecular dynamics simulation of the thio-substituted state has been run using as initial structure an equilibrated structure taken from the the wild type ribozyme simulation that has been modified to contain a thio-substitued scissile phosphate. To accommodate the structural change induced by the thio-substitution a 2 ns long equilibration has been run where all the heavy atoms of the ribozyme have been loosely restrained to their initial Cartesian coordinates with the exception of the active residues (A-1,G+1,G8,A38). During equilibration the restraints were removed in a stepwise manner by scaling the corresponding force constants. (A) Representative structure of the active site showing the average hydrogen bonding distances between G8:N1 and A-1:O2' and A-1:O2' and G+1:S2P. Both hydrogen bond donor–hydrogen bond acceptor and hydrogen–hydrogen bond acceptor distances are shown. (B) Time series and distributions of hydrogen bond distances corresponding to the G8:N1–A-1:O2' hydrogen bond (top) and A-1:O2'–G+1:S2P hydrogen bond (bottom).

Potential energy surfaces of the proton transfer between O2' and the nonbridging oxygen or sulfur

Figure S5 displays the relaxed potential energy surface (PES) of the proton transfer between O2' and the nonbridging oxygen as a function of $R_{O2'\text{-}H} - R_{H\text{-}O}$ distance and the corresponding $R_{O2'\text{-}H} - R_{H\text{-}S}$ PES upon thiosubstitution of the nonbridging oxygen. The PESs were constructed using two different models. The first model system, labeled "RNA mimic", consists of 290 atoms constructed by extracting all nonsolvent residues within 6 Å of the scissile phosphate from a molecular dynamics snapshot (see Figure S6). The PES of the $R_{O2'\text{-}H} - R_{H\text{-}O}$ reaction coordinate was performed with DFTB3 using a variable radii COSMO (VRCOSMO) implicit solvation model. The VRCOSMO method is equivalent to the smooth COSMO model introduced by York and Karplus,[12] but whose radii are allowed to vary as a function of atomic charge. The VRCOSMO parameters were optimized to reproduce the experimental relative solvation free energies of small molecules collected in Ref. 13. Only the A-1, G+1 and G8 residues (highlighted in Figure S6) were allowed to move in the geometry optimization so as to maintain the backbone scaffolding observed in the full RNA system. The scan was repeated a second time using the "small model" system shown in Figure S7(top), which consists of only 39 atoms. In this case, all atoms in the small model system were allowed to geometry optimize during the construction of the PES. The PES of the small model system was again performed upon thiosubstution of the nonbridging oxygen (see Figure S7(bottom)). We also performed DFT B3LYP/6-31++G** single point calculations of the small model system with PCM implicit solvation using UAKS radii. The RNA mimic and small model PESs are in good agreement; they both suggest that the proton transfer to the nonbridging oxygen is unfavorable by approximately 20 kcal/mol at the DFTB3 level. B3LYP single point calculations suggest that proton transfer to the nonbridging oxygen is unfavorable by approximately 24 kcal/mol.

Thiosubstitution at the nonbridging position makes proton transfer even more unfavorable. The DFTB3 energy is 31 kcal/mol higer in energy when the nonbridging sulfur is protonated than the O2'.

Subsequent DFT B3LYP/6-31++G** single point calculations using the PCM implicit solvent model similarly show a 30 kcal/mol difference in energy.

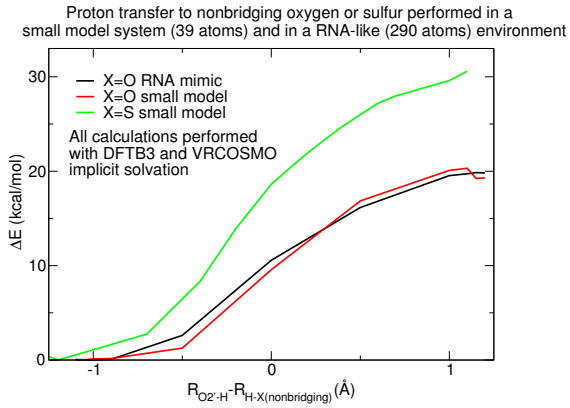

Figure S5: Potential energy surfaces of the proton transfer between O2' and the nonbridging oxygen (X=O) or the nonbridging sulfur (X=S). All calculations were performed with DFTB3 using the VRCOSMO implicit solvation method. The coordinates are provided in the files FigS5.RNAmimc.xyz.txt, FigS5.Omodel.xyz.txt, and FigS5.Smodel.xyz.txt, and the absolute energies are listed in Hartrees within the title section of the XYZ file format.

Potential energy surfaces of the proton transfer between O2' and the deprotonated Guanine N1

Figure S8 displays the relaxed potential energy surface (PES) of the proton transfer between O2' and the deprotonated Guanine N1 as a function of $R_{\text{O2'-H}} - R_{\text{H-N1}}$ distance.

The RNA model system described in the previous section is again used with DFTB3+VRCOSMO to construct a relaxed PES. The N1 proton has been removed and the highlighted atoms in Figure S9 are allowed to geometry optimize.

The DFTB3 PES displays two minima corresponding to covalent bonding of the proton to O2' and N1. Furthermore, the minima are nearly degenerate in energy and are separated by a transistion state barrier of approximately 4 kcal/mol.

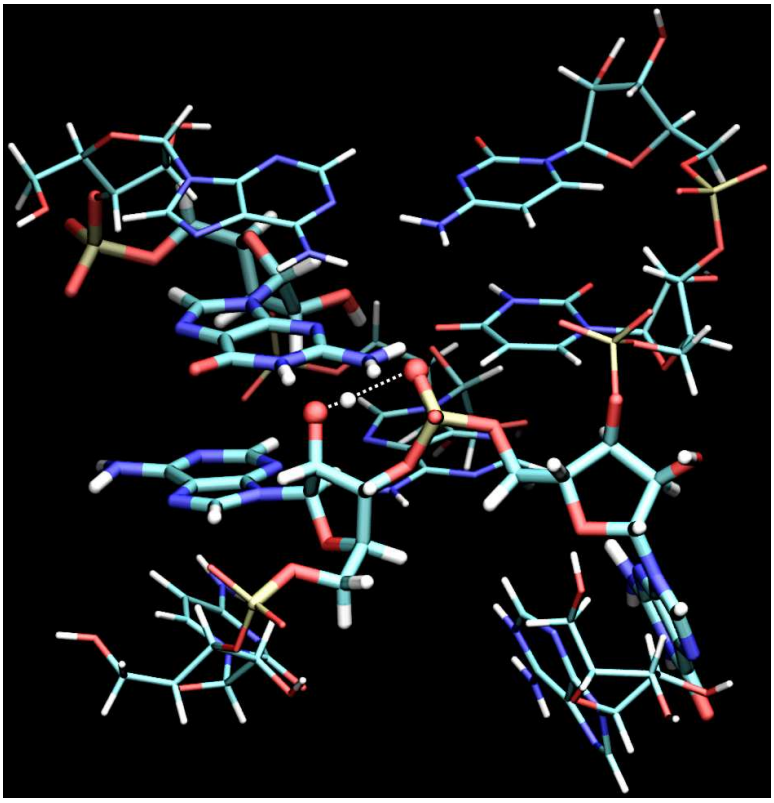

Figure S6: The RNA mimic system used to construct the potential energy surface corresponding to the proton transfer between O2’ and the nonbridging oxygen. The highlighted “thick” atoms are allowed to geometry optimize.

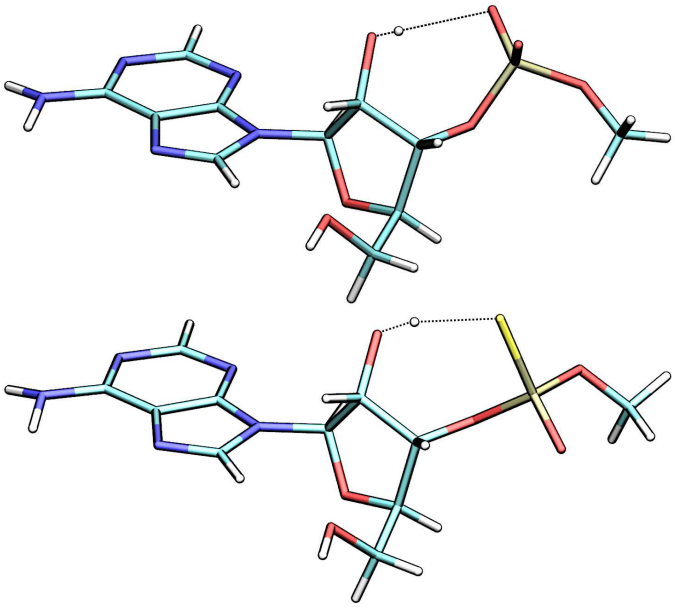

Figure S7: The “small model” system used to construct potential energy surfaces corresponding to the proton transfer between O2’ and the nonbridging oxygen (top) or sulfur (bottom).

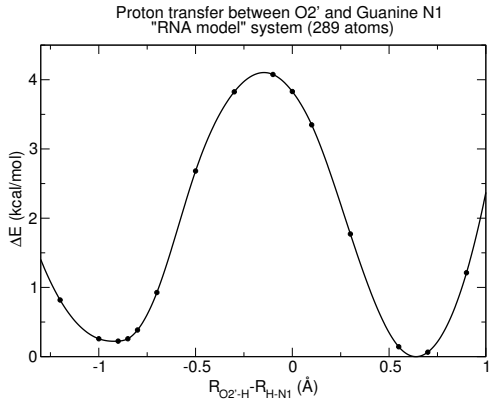

Figure S8: Potential energy surfaces of the proton transfer between O2' and the deprotonated Guanine N1. All calculations were performed with DFTB3 using the VRCOSMO implicit solvation method. The coordinates are provided in the file FigS8.RNAmimic.xyz.txt, and the absolute energies are listed in Hartrees within the title section of the XYZ file format.

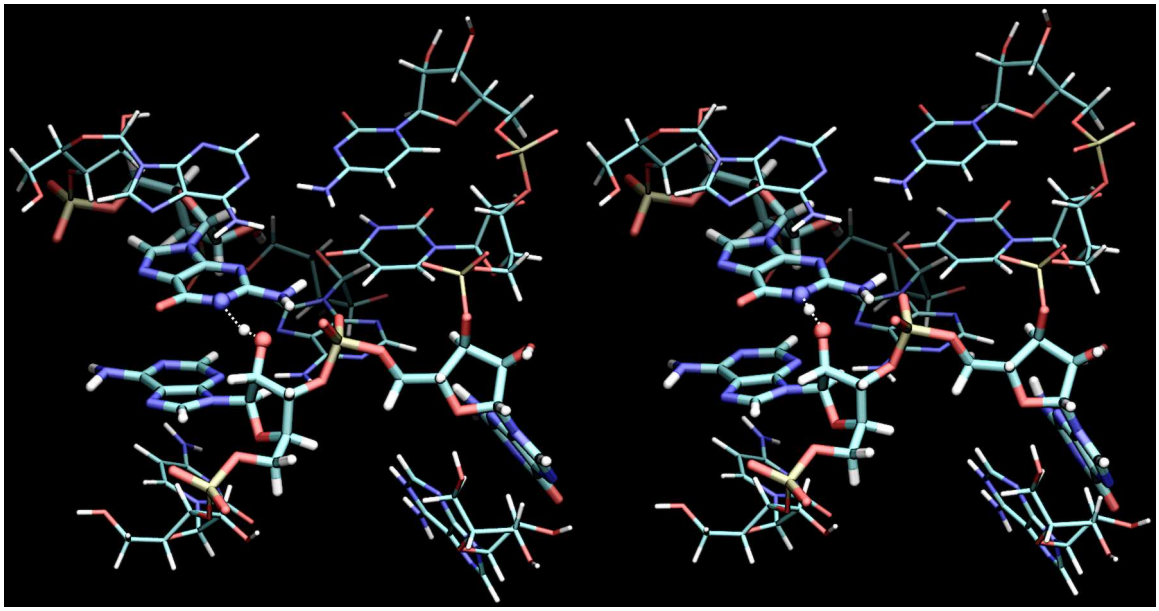

Figure S9: The "RNA model" system used to construct potential energy surfaces corresponding to the proton transfer between O2' and the deprotonated Guanine N1. The left structure shows the proton bonded to the O2' and the right structure shows it bonded to N1.

References

(1) Salter, J.; Krucinska, J.; Alam, S.; Grum-Tokars, V.; Wedekind, J. E. *Biochemistry* **2006**, *45*, 686–700.

(2) Torelli, A. T.; Krucinska, J.; Wedekind, J. E. *RNA* **2007**, *13*, 1052–1070.

(3) Torelli, A. T.; Spitale, R. C.; Krucinska, J.; Wedekind, J. E. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 154–158.

(4) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kaleé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(5) Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, C., R. Sagui; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. AMBER 12. University of California, San Francisco: San Francisco, CA, 2012.

(6) Pérez, A.; Marchán, I.; Svozil, D.; Šponer, J.; Cheatham III, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817–3829.

(7) Joung, I. S.; Cheatham III, T. E. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

(8) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(9) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(10) Janowski, P. A.; Cerutti, D. S.; Holton, J.; Case, D. A. *J. Am. Chem. Soc.* **2013**, *135*, 7938–7948.

(11) Lee, T.-S.; Giambaşu, G. M.; Sosa, C. P.; Martick, M.; Scott, W. G.; York, D. M. *J Mol Biol* **2009**, *388*, 195–206.

(12) York, D. M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.

(13) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1177–1152.