Supporting Information

Table S1: The angles and distances for the in-line S_N2 attack.

	Reactant	Reactant*	2GOZ†	Full Length†	1HMH	1MME	299D	300D	301D	359D
R(C17:$O_{2'}$,C1.1:P)	3.61 ± 0.23	3.83 ± 0.19	3.18	3.27	4.36	4.32	3.46	3.51	4.08	3.54
Angle(C17:$O_{2'}$,C1.1:P,C1.1:$O_{5'}$)	141.9 ± 11.3	138.0 ± 10.3	162.0	150.7	127.0	144.9	56.9	60.0	74.9	63.4
In-line fitness (F)	0.43 ± 0.13	0.34 ± 0.08	0.73	0.61	0.20	0.25	0.06	0.07	0.09	0.08
Percentage in NAC‡	5.58	0.12								

*Reactant state simulation without Mg^{2+} ion. †The full length hammerhead RNA crystallographic structure of Scott and Martick[1] that was also used in the paper as the starting structure. †An unpublished full length hammerhead RNA crystallographic structure with Mn^{2+} ions at 2.0Å resolution. ‡ Near in-line Attack Conformation (NAC) region by Torres and Bruice.[2]

Simulation Setup

Initial structures used in the simulations were based on the crystallographic structure of the full length hammerhead ribozyme of Martick and Scott.[1] The positions of hydrogen atoms are determined using HBUILD facility in the program CHARMM[3] (version c32a2). Four simulations were performed: one simulation of the reactant state in the absence of Mg^{2+} ions, and three simulations of the reactant state, early and late transition state (TS) mimics all in the presence of a bridging Mg^{2+} ion. The Mg^{2+} ion was placed between the A9 and scissile phosphate coordinating the non-bridging O_{P2} atoms that were 4.3 Å apart in the crystal structure. This type of coordination was implicated based on thio/rescue effect experiments that suggest a single divalent metal might bridge these positions in the transition state, despite having never been observed crystallographically.[4] The ribozyme was then immersed in a rhombic dodecahedral cell of 10,062 pre-equilibrated TIP3P[5] waters centered about the active site, and pruned such that any water molecule within 2.8 Å from the solute was removed. The ion atmosphere consisted of Na^+ and Cl^- ions that were added at random positions to neutralize the system and reach the physiologic concentration of 0.14 M. The ions positions were kept initially at least 4.7 Å away from any solute atoms. The resulting system (the reactant state) contained 9,053 water molecules, 82 Na^+, 23 Cl^-, and 2,021 RNA atoms.

Simulations were performed with CHARMM[3] (version c32a1) using the all-atom CHARMM27 nucleic acid force field[6,7] with extension to reactive intermediate models (e.g., transition state mimics)[8] and TIP3P water model.[5] The early TS mimic is close to an ideal phosphorane in that the bond lengths to the endocyclic 2'O and exocyclic 5'O are roughly equal. Density-functional studies of model transesterification reactions suggest that the rate-controlling transition state for a dianionic mechanism are "late" in that the bond to the nucleophile is almost fully formed and shorter whereas the bond to the leaving group is almost completely cleaved and significantly longer with more accumulated charge on the oxygen. The parameters for the late transition state mimic involved a minor modification of the early transition state mimic. The late TS mimic has the nucleophilic P-$O_{2'}$ and leaving group P-$O_{5'}$ bond lengths of 1.856 and 2.382 Å, respectively, and the charge of the leaving $O_{5'}$ atom was increased from -0.68e to -0.90e while the charge of $C_{5'}$ was adjusted in compensation to ensure overall charge conservation. The modifications for the late transition state mimic were based on density-functional calculations of the rate-controlling transition state for the transesterification reaction of a sugar-phosphate with methoxide leaving group reported previously[9] and available in the QCRNA database.[10]

Periodic boundary conditions were used along with the isothermal-isobaric ensemble (NPT) at 1 atm and 298 K using extended system pressure algorithm[11] with effective mass of 500.0 amu and Nosé-Hoover thermostat[12,13] with effective mass of 1000.0 kcal/mol-ps^2, respectively. The smooth particle mesh Ewald (PME) method[14,15] was employed with a κ value of 0.35 Å$^-$1, 80 FFT grid points for each of the lattice directions, and a B-spline interpolation order of 6. Non-bonded interactions were treated using an atom-based cutoff of 10 Å with shifted van der Waals potential. Numerical integration was performed using the leap-frog Verlet algorithm with 1 fs time step.[16] Covalent bond lengths involving hydrogen were constrained using the SHAKE algorithm.[17]

Equilibration Protocol

The following equilibration procedures (total 1 ns) were applied to the system prior the production simulations. The positions of the solute atoms, including the Mg^{2+} ion, were first kept in the equilibration stages.

Pre-annealing Stage

Water and ion molecules were first energy-optimized then underwent a constant volume simulation annealing for 50 ps. The temperature was increased from 0 K to 298 K in a 7.5 ps period then was kept at 298 K. The annealing simulation were repeated twice with temperature increased from 298 K to 498 K and then back to 298 K.

Annealing Stage

Four steps of constant volume simulations (50 ps each) were performed in the this stage.

First step: The temperature increased from 298 K to 498 K in 7.5 ps then was kept at 498 K.

Second step: The temperature increased from 498 K to 698 K in 7.5 ps then was kept at 698 K.

Third step: The temperature decreased from 698 K to 498 K in 7.5 ps then was kept at 498 K.

Fourth step: The temperature decreased from 498 K to 298 K in 7.5 ps then was kept at 298 K.

The whole annealing stage was repeated three times before the post-annealing stage.

Post-annealing Stage

Three steps of constant volume simulations were performed in the this stage.

First step (50 ps) : The temperature increased from 298 K to 498 K in 7.5 ps then was kept at 498 K.

Second step (50 ps) : The temperature decreased from 498 K to 298 K in 7.5 ps then was kept at 298 K.

Third step (150 ps): The temperature was kept at 298 K for 150 ps.

Solute Relaxation Stage

The solute atoms were energy-optimized and then were allowed to move under harmonic restraints over a 50 ps simulation at 298 K under constant pressure of 1 atm. The harmonic force constant (in kcal mol^{-1} Å^{-2}) on each heavy atom was obtained from the empirical formula $k_i = 25 + 2{\times}10^3/B_i$ where k_i is the force constant for atom i and B_i is the corresponding crystallographic B-value. The restraints were exponentially released over 50 ps with a half-life decay parameter of 10 ps. At the end of the 50 ps simulation, the restraints were reduced to about 3 percents of the initial restraints. Three harmonic restraints of 20 kcal mol^{-1} Å^{-2} were added to keep the Mg^{2+} ion in the middle of the C1.1:O_{P2} and A9:O_{P2} positions. Another harmonic restraint of 20 kcal mol^{-1} Å^{-2} was used to force the distance between G8:H_{OP2} and C1.1:O_{P5} to be around 1.8 Å, which is to ensure that the H_{OP2} of G8 is initially hydrogen bonded. All restraints were then released prior to the production simulation.

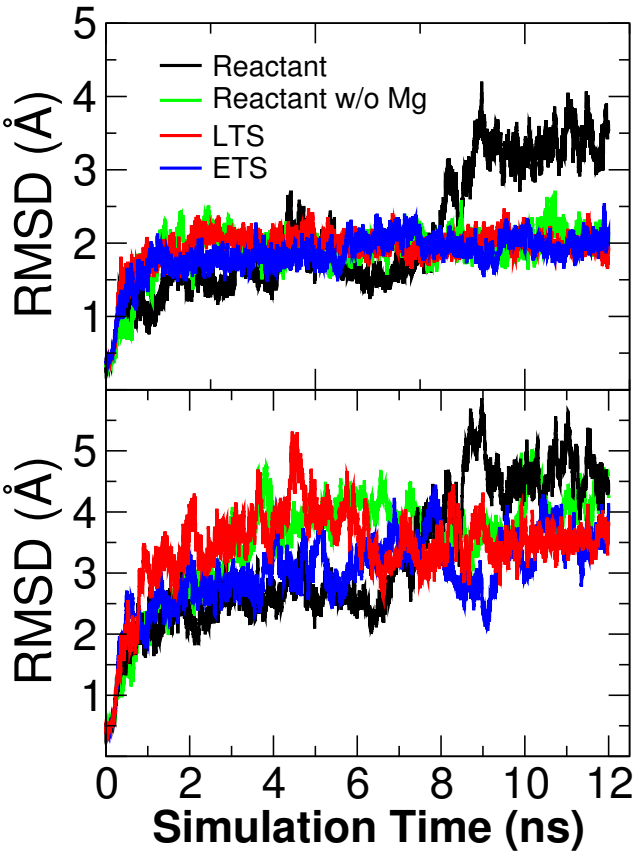

Figure S1: The heavy-atom root-mean-square deviation (RMSD) plots of the whole full length hammerhead system (bottom) and the active site residues (top) with respect to the crystallographic structure.[1] The active site was defined as all residues within 10 Å from the scissile phosphate.

Production Simulation

After the 1 ns of solvent equilibration, the whole system was energy-optimized and unconstrained dynamics simulation began from 0 K under constant pressure of 1 atm. The temperature was increased to 298 K at the rate of 1 K/ps and then kept fixed at 298 K. The same equilibration process was applied for each simulation. A total of 12 ns of unconstrained dynamics was performed for each of the four simulations (reactant with and without Mg^{2+}, early TS mimic and late TS mimic), the last 10 ns of which were used for analysis. The motions and relaxation of solvent and counterions are notoriously slow to converge in nucleic acid simulations,[18] and careful equilibration is critical for reliable simulations. In summary, for each simulation, a total of 3 ns of equilibration (1 ns of solvent relaxation and 2 ns of solvent and structure relaxation) has been carried out before 10 ns of data sampling. This represents to date, the longest reported simulations of the hammerhead system.

RMSD results

The heavy-atom root-mean-square deviation (RMSD) plots of the whole full length hammerhead system and the active site residues with respect to the crystallographic structure[1] are shown in Figure S1.

Preliminary QM/MM results

We have performed preliminary QM/MM simulations on the early and late transition state mimics. Initial structures were taken from snapshots of the purely classical molecular dynamics simulations after 2 ns. The system is partitioned into a quantum mechanical (QM) region constituting the active site that is represented by the AM1/d-PhoT Hamiltonian[19] and the modified AM1 magnesium parameters of Hutter and co-workers.[20] The total number of solute and solvent atoms, setup of periodic boundary conditions, etc., was identical to the classical simulations. The QM subsystem was defined as the 43 atoms around the active site, and included the scissile and A9 phosphates, parts of the nucleophilic and leaving ribose rings, and Mg^{2+} ion and coordinated waters. The generalized hybrid orbital (GHO) method[21] is used to cut a covalent bond to divide the system into QM and MM region. Full electrostatic interactions were calculated using a recently introduced linear-scaling QM/MM-Ewald method.[22] The results are shown in Table S2. The roles inferred from purely classical MD simulations are confirmed by the QM/MM results: The Mg^{2+} ion is bonded to G8:$O_{2'}$ in the early transition state mimic and switches to C1.1:$O_{5'}$ of the leaving group in the late transition state mimic. In the late transition state mimic, the G8:$H_{O_{2'}}$ is strongly hydrogen-bonded to C1.1:$O_{5'}$. In fact, the proton transfer from G8:$O_{2'}$ to C1.1:$O_{5'}$ is observed and confirms that Mg^{2+} may significantly lower the pK_a of G8:$O_{2'}$ such that it may act as a general acid at the catalytic pH.

Table S2: Key distances (Å) in the hammerhead active site from early transition state and late transition state mimics from molecular dynamics (MD) and combined quantum mechanics and molecular mechanics (QM/MM) simulations. ‡

	Early-TS (MD)	Early-TS (QM/MM)	Late-TS (MD)	Late-TS (QM/MM)
Active site RMSD	1.92(16)	1.61 (10)	2.00(12)	1.99(10)
Mg ↔ G8:$O_{2'}$	2.24(13)	2.14 (33)	3.21(23)	2.42(33)
Mg ↔ C1.1:$O_{5'}$	3.68(35)	3.95 (10)	2.09(05)	2.13(10)
G8:$H_{O_{2'}}$ ↔ C1.1:$O_{5'}$	5.09(74)	3.19 (39)	2.36(42)	1.79(39)

‡The active site is defined as all residues within 10 Å from the scissile phosphate. ¶The results for MD simulations were calculated over the last 10 ns with data collected every 1 ps. The QM/MM results were calculated over 250 ps with data collected every 0.5 ps. Entries shown are average values and standard deviations in parenthesis (only in the last two or three digits).

In-line Fitness Analysis

The in-line fitness (F) has been used as a measurement for the likelihood of activity of an RNA,[23] and it has been shown that when an RNA would be likely active when $F >= 0.4$, where F is defined by

$$F = \frac{\tau - 45}{180 - 45} \times \frac{3^3}{d^3_{O2'-P}} \quad (1)$$

where τ is the $O_{2'}$-P-$O_{5'}$ angle in degrees and $d^3_{O2'-P}$ is the $O_{2'}$-P distance in Å. Most of the currently available Hammerhead RNA structures are obtained from the minimal sequence hammerhead RNA motif.[24–26] We calculated the in-line fitness (F) for selected minimal length Hammerhead RNA's, two full length Hammerhead RNA's, and from our reactant state simulation results, with and without Mg ion. The results are listed in Table S1. The in-line fitness value for the two full length structures are 0.73 and 0.61 while our simulations results of the reaction state with an Mg^{2+} ion have an average of 0.42 with fluctuations that have considerable population that encompass the full length crystal values. The minimal length structures have much lower in-line fitness values, ranged from 0.07 to 0.25. The values of the full length structures are close to the F value for the hairpin RNA with in-line confirmation, which is 0.83 (considering only the angle term since the distance term exceeds the limit).[27] A plot of the angle versus distance approach of the 2′-hydroxyl of residue C17 to the phosphate of residue C1.1 is shown in Figures S2 and S3. It is evident from the figure that two full length crystal structures and the reactant simulation are in positions that are in near attack conformations ready to react while minimal length structures are quite far from the ideal attack conformation. The reaction state with a catalytic Mg^{2+} ion has 5.56% population in the Near in-line Attack Conformation (NAC) region defined by Torres and Bruice[2] while the population is reduced to 0.12% for the reactant state simulation without Mg^{2+}.

Whether the placement of the Mg^{2+} ion between the A9 and scissile phosphates in the reactant simulation is highly occupied in the ground state has yet to be established. It is possible that the Mg^{2+} ion is mainly coordinated elsewhere in the reactant ground state, such as between the A9 phosphate and the N7 position of G10.1,[28,29] and that a pH-dependent transition to a metal bridge between the A9 and scissile phosphates either precedes or occurs synchronously with movement to the transition state.[4] The results of the present simulations suggest that Mg^{2+} coordination between the A9 and scissile phosphates is at least transiently stable in the reactant and stable in the early and late TS mimic simulations, and that in this position the Mg^{2+} is well poised to play an active chemical role in catalysis by shifting the pK_a of the 2'OH of G8 to act as a general acid, and/or by providing electrostatic stabilization of the accumulating charge of the 5' leaving group through either direct or indirect coordination.

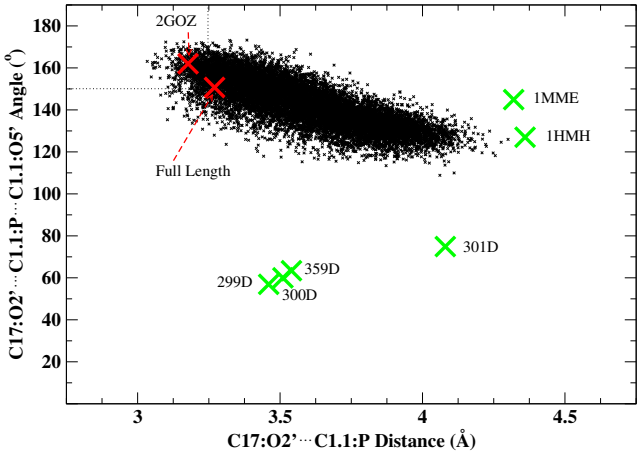

Figure S2: Plot of the angle versus distance approach of the 2′-hydroxyl of residue C17 to the phosphate of residue C1.1 for the reactant state simulation with a bridging Mg^{2+} ion (small black crosses), previously reported Hammerhead RNA X-ray structures (large green crosses, listed in Table S1), and the two full length hammerhead structures (large red crosses). Dashed lines at 3.25 Å and 150 degrees are the Near in-line Attack Conformation (NAC) region defined by Torres and Bruice.[2]

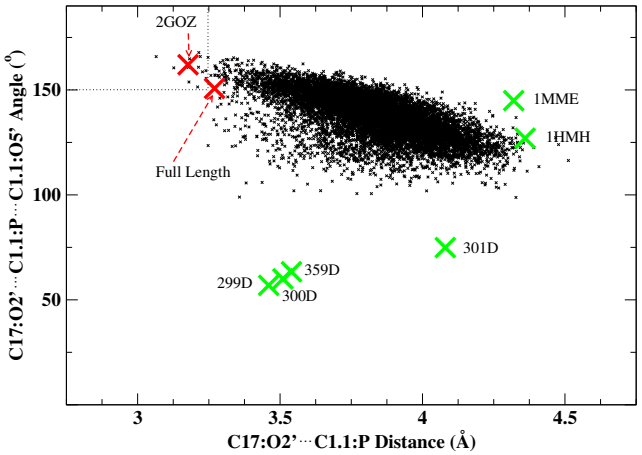

Figure S3: Plot of the angle versus distance approach of the 2′-hydroxyl of residue C17 to the phosphate of residue C1.1 for the reactant state simulation in the absence of Mg^{2+} (small black crosses), previously reported Hammerhead RNA X-ray structures (large green crosses, listed in Table S1), and the two full length hammerhead structures (large red crosses). Dashed lines at 3.25 Å and 150 degrees are the Near in-line Attack Conformation (NAC) region defined by Torres and Bruice.[2]

(1) Martick, M.; Scott, W. G. *Cell* **2006,** *126*, 309–320.
(2) Torres, R. A.; Bruice, T. C. *J. Am. Chem. Soc.* **2000,** *122*, 781–791.
(3) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983,** *4*, 187–217.
(4) Wang, S.; Karbstein, K.; Peracchi, A.; Beigelman, L.; Herschlag, D. *Biochemistry* **1999,** *38*, 14363–14378.
(5) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983,** *79*, 926–935.
(6) Foloppe, N.; MacKerell, Jr., A. D. *J. Comput. Chem.* **2000,** *21*, 86–104.
(7) MacKerell, Jr., A. D.; Banavali, N. K. *J. Comput. Chem.* **2000,** *21*, 105–120.

(8) Mayaan, E.; Moser, A.; Mackerell, Jr., A. D.; York, D. M. *J. Comput. Chem., in press*.
(9) Liu, Y.; Gregersen, B. A.; Lopez, X.; York, D. M. *J. Phys. Chem. B* **2005,** *109*, 19987–20003.
(10) Giese, T. J.; Gregersen, B. A.; Liu, Y.; Nam, K.; Mayaan, E.; Moser, A.; Range, K.; Nieto Faza, O.; Silva Lopez, C.; Rodriguez de Lera, A.; Schaftenaar, G.; Lopez, X.; Lee, T.; Karypis, G.; York, D. M. *J. Mol. Graph. Model.* **2006,** *25*, 423–433.
(11) Andersen, H. C. *J. Chem. Phys.* **1980,** *72*, 2384–2393.
(12) Nosé, S.; Klein, M. L. *Mol. Phys.* **1983,** *50*, 1055–1076.
(13) Hoover, W. G. *Phys. Rev. A* **1985,** *31*, 1695–1697.
(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing, L.; Pedersen, L. G. *J. Chem. Phys.* **1995,** *103*, 8577–8593.
(15) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999,** *28*, 155–179.
(16) Allen, M.; Tildesley, D. *Computer Simulation of Liquids;* Oxford University Press: Oxford, 1987.
(17) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977,** *23*, 327–341.
(18) Ponomarev, S. Y.; Thayer, K. M.; Beveridge, D. L. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 14771-14775.
(19) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory. Comput., in press*.
(20) Hutter, M. C.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1998,** *102*, 8080–8090.
(21) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998,** *102*, 4714–4721.
(22) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2005,** *1*, 2-13.
(23) Soukup, G. A.; Breaker, R. R. *RNA* **1999,** *5*, 1308–1325.
(24) Pley, H. W.; Flaherty, K. M.; McKay, D. B. *Nature* **1994,** *372*, 111–113.
(25) Scott, W. G.; Finch, J. T.; Klug, A. *Cell* **1995,** *81*, 991–1002.
(26) Scott, W. G.; Murray, J. B.; Arnold, J. R. P.; Stoddard, B. L.; Klug, A. *Science* **1996,** *274*, 2065–2069.
(27) Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferré-D'Amaré, A. R. *Science* **2002,** *298*, 1421–1424.
(28) Peracchi, A.; Beigelman, L.; Scott, E. C.; Uhlenbeck, O. C.; Herschlag, D. *J. Biol. Chem.* **1997,** *272*, 26822–26826.
(29) Peracchi, A.; Beigelman, L.; Usman, N.; Herschlag, D. *Proc. Natl. Acad. Sci. USA* **1996,** *93*, 11522–11527.