Quantum Mechanical/Molecular Mechanical Simulation Study of the Mechanism of Hairpin Ribozyme Catalysis

Kwangho Nam,[†,‡,§] Jiali Gao,[†,‡] and Darrin M. York*,[†]

Department of Chemistry and Supercomputing Institute and Digital Technology Center, University of Minnesota, Minneapolis, Minnesota 55455−0431

Received August 7, 2007; E-mail: york@chem.umn.edu

Abstract: The molecular mechanism of hairpin ribozyme catalysis is studied with molecular dynamics simulations using a combined quantum mechanical and molecular mechanical (QM/MM) potential with a recently developed semiempirical AM1/d-PhoT model for phosphoryl transfer reactions. Simulations are used to derive one- and two-dimensional potentials of mean force to examine specific reaction paths and assess the feasibility of proposed general acid and base mechanisms. Density-functional calculations of truncated active site models provide complementary insight to the simulation results. Key factors utilized by the hairpin ribozyme to enhance the rate of transphosphorylation are presented, and the roles of A38 and G8 as general acid and base catalysts are discussed. The computational results are consistent with available experimental data, provide support for a general acid/base mechanism played by functional groups on the nucleobases, and offer important insight into the ability of RNA to act as a catalyst without explicit participation by divalent metal ions.

1. Introduction

The hairpin ribozyme is a small RNA enzyme that catalyzes the reversible, site-specific cleavage of the phosphodiester bond of the substrate RNA strand.[1−5] The chemical mechanism of the cleavage reaction involves an associative in-line attack of an activated ribose 2′-OH nucleophile on an adjacent phosphorus atom to create a 2′,3′-cyclic phosphate and a 5′-OH terminus.[1−5] The ribozyme-catalyzed phosphate cleavage has a rate of ~0.3 min[−1],[1,6,7] which represents a rate enhancement of approximately 10[6]-fold relative to the nonenzymatic process in aqueous solution.[8] The hairpin ribozyme catalyzes the reverse ligation reaction at a rate up to an order of magnitude faster than the forward cleavage reaction.[6,7,9−11] Despite an abundance of experimental data, a detailed picture of the reaction mechanism remains unclear. Of central importance is to characterize the specific role of conserved nucleotides in the active site and assess the contributions of solvent and the electrostatic environment of the ribozyme on the reaction rate.

The hairpin ribozyme is distinct in its catalytic mechanism[2,7,12−15] in comparison with other small self-cleaving ribozymes such as hammerhead[16,17] and hepatitis delta virus[18,19] ribozymes. For minimal hairpin ribozymes, the apparent lack of an explicit metal ion requirement for catalysis[12,14,15] and a shallow, bell-shaped pH-rate profile[7,15] are consistent with a general acid−base mechanism with the nucleobases playing a key role in activation of the 2′ nucleophile and stabilization and protonation of the 5′ leaving group. In this scenario, the ribozyme environment would considerably perturb the pK_a of one or more nucleobases. However, the nature of the ionization events of those pK_a-perturbed bases that affect activity cannot be unambiguously identified from kinetic data alone. As has been noted elsewhere,[9] different hairpin ribozyme variants exhibit different pH-rate behavior. For the most stable variants, catalytic activity increases linearly with pH, consistent with a kinetic model where protonation of a functional group with apparent pK_a of approximately 6 was correlated with activity.[9]

Recent high-resolution X-ray crystallographic structures provide valuable insight into the active site conformation in the reactant state,[20−22] transition state, and product state mimic structures.[4] Comparison of these structures shows that the active

---

† Department of Chemistry and Supercomputing Institute.
‡ Digital Technology Center.
§ Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138.

(1) Walter, N. G.; Burke, J. M. *Curr. Opin. Chem. Biol.* **1998**, *2*, 24−30.
(2) Fedor, M. J. *J. Mol. Biol.* **2000**, *297*(2), 269−291.
(3) Doherty, E. A.; Doudna, J. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 457−475.
(4) Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferré-D'Amaré, A. R. *Science* **2002**, *298*, 1421−1424.
(5) Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2259−2265.
(6) Hegg, L. A.; Fedor, M. J. *Biochemistry* **1995**, *34*, 15813−15828.
(7) Nesbitt, S. M.; Erlacher, H. A.; Fedor, M. J. *J. Mol. Biol.* **1999**, *289*, 1009−1024.
(8) Li, Y.; Breaker, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 5364−5372.
(9) Kuzmin, Y. I.; Costa, C. P. D.; Cottrell, J. W.; Fedor, M. J. *J. Mol. Biol.* **2005**, *349*, 989−1010.
(10) Nahas, M. K.; Wilson, T. J.; Hohng, S.; Jarvie, K.; Lilley, D. M. J.; Ha, T. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1107−1113.
(11) Liu, S.; Bokinsky, G.; Walter, N. G.; Zhuang, X. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*(31), 12634−12639.

(12) Hampel, A.; Cowan, J. A. *Chem. Biol.* **1997**, *4*, 513−517.
(13) Chowrira, B. M.; Burke, J. M. *Biochemistry* **1991**, *30*, 8518−8522.
(14) Young, K. J.; Gill, F.; Grasby, J. A. *Nucleic Acids Res.* **1997**, *25*, 3760−3766.
(15) Nesbitt, S.; Hegg, L. A.; Fedor, M. J. *Chem. Biol.* **1997**, *4*, 619−630.
(16) Scott, W. G.; Murray, J. B.; Arnold, J. R. P.; Stoddard, B. L.; Klug, A. *Science* **1996**, *274*, 2065−2069.
(17) Scott, W. G. *Q. Rev. Biophys.* **1999**, *32*, 241−294.
(18) Shih, I.-h.; Been, M. *Biochemistry* **2000**, *39*, 9055−9066.
(19) Shih, I.-h.; Been, M. D. *Annu. Rev. Biochem.* **2002**, *71*, 887−917.

site bases surrounding the cleavage site do not undergo large variation and remain poised for the in-line nucleophilic attack. The cleavage site, especially the scissile phosphate itself, undergoes greater variation as the reaction proceeds. In close proximity to the scissile phosphodiester bond in the active site, there are two key conserved nucleotides, G8 and A38,[23,24] that interact (mainly through hydrogen bonds) with the nucleophile, the leaving group, and the nonbridging phosphoryl oxygens. On the basis of the transition state mimic structure with a pentavalent vanadate ion at the scissile phosphate position, it has been suggested that A38, in protonated form, acts as a general acid in the cleavage reaction.[4,24,25] The G8 nucleotide base has been implicated in the stabilization of the transition state.[4,25] This is consistent with the results from nucleotide analog interference mapping experiments.[26,27] G8 and A38 are critical in catalysis: the deletion of the G8 base decreases the reaction rate by a factor of 350, and the deletion of the A38 base decreases the rate more than 10 000-fold relative to the wild-type control.[9,24,28,29]

One important factor is the protonation states of the scissile phosphate and other nucleotides present in the active site and their relationship to the pH-rate profiles. Although G8, A38, and other nucleobases have been found to significantly affect the catalytic efficiency,[26,27,30,31] their specific roles in catalysis have not yet been clearly defined. Although alternative interpretations are possible,[5,32] the proposal that a protonated A38 acts as a general acid in the cleavage reaction would require the pK_a of A38 to be considerably shifted from a value of ~4.3 in solution[33] or correlated with a functional group with the apparent pK_a value of ~6.1 in the hairpin ribozyme.[9,34,35] If the A38 pK_a remained at its value in solution, the pH-rate profile suggests that A38 would be deprotonated over the catalytic pH range.[7,9,15] Moreover, the protonation state and the role of G8 in the catalysis are also unclear. The protonation state of the reactive phosphorane intermediate and/or transition state is elusive due to the transient nature of these species in solution and in the ribozyme environment. Theoretical and reaction kinetic studies on small model compounds suggest that reactive phosphorane intermediates are most likely monoprotic[36–40] in

accord with their expected protonation state over the catalytic pH range based on estimates of oxyphosphorane pK_a values in solution.[41,42]

In the present computational study, the main goal is to decipher the specific roles of residues G8, A38, and other nucleotide bases in hairpin ribozyme catalysis. Free energy simulations employing a combined quantum mechanical and molecular mechanical (QM/MM) potential have been applied successfully to many chemical reactions in solution[43–48] and in enzymes[49,50] and are used in the present study. The high concentration of charged residues and conformational variability of RNA that can accompany the chemical catalytic steps in ribozyme catalysis present special challenges to QM/MM simulations. These conditions necessitate the use of an accurate quantum model that is sufficiently fast to allow extensive sampling of the solvent degrees of freedom and anharmonic macromolecular motions[47,49,51–53] to derive reliable, converged thermodynamic quantities, including often significant entropic contributions to the activation free energy barriers.[54,55] Toward this end, we employ a recently developed semiempirical AM1/d-PhoT Hamiltonian model specifically designed to study phosphoryl transfer reactions in enzymes and ribozymes.[56] The model has been tested and demonstrated to be robust for phosphoryl transfer reactions[48,56] and sufficiently fast so as to allow long-time molecular dynamics simulations to be performed. We further examine truncated reaction model systems with density-functional theory in order to assess the reliability of the AM1/d-PhoT model and provide additional insight into mechanism.

The present paper is organized as follows. Section 2 describes the computational details of the QM/MM simulations, generation of the free energy profiles, and active site model calculations. Section 3 presents the results for the dianionic and in-line monoanionic reaction mechanisms. The remainder of the section discusses the possible roles of A38 and G8 nucleobases and the details of the mechanism that are consistent with the

(20) Rupert, P. B.; Ferré-D'Amaré, A. R. *Nature* **2001**, *410*, 780–786.
(21) Salter, J.; Krucinska, J.; Alam, S.; Grum-Tokars, V.; Wedekind, J. E. *Biochemistry* **2006**, *45*, 686–700.
(22) Torelli, A. T.; Krucinska, J.; Wedekind, J. E. *RNA* **2007**, *13*, 1052–1070.
(23) Butcher, S. E.; Burke, J. M. *Biochemistry* **1994**, *33*, 992–999.
(24) Cottrell, J. W.; Kuzmin, Y. I.; Fedor, M. J. *J. Biol. Chem.* **2007**, *282*, 13498–13507.
(25) Ferré-D'Amaré, A. R. *Biopolymers* **2004**, *73*, 71–78.
(26) Ryder, S. P.; Strobel, S. A. *Nucleic Acids Res.* **2002**, *30*, 1287–1291.
(27) Ryder, S. P.; Oyelere, A. K.; Padilla, J. L.; Klostermeier, D.; Millar, D. P.; Strobel, S. A. *RNA* **2001**, *7*, 1454–1463.
(28) Lebruska, L. L.; Kuzmine, I. I.; Fedor, M. J. *Chem. Biol.* **2002**, *9*, 465–473.
(29) Kuzmin, Y. I.; Da Costa, C. P.; Fedor, M. J. *J. Mol. Biol.* **2004**, *340*, 233–251.
(30) Pinard, R.; Hampel, K. J.; Heckman, J. E.; Lambert, D.; Chan, P. A.; Major, F.; Burke, J. M. *EMBO J.* **2001**, *20*, 6434–6442.
(31) Thomas, J. M.; Perrin, D. M. *J. Am. Chem. Soc.* **2006**, *128*, 16540–16545.
(32) Wilson, T. J.; Ouellet, J.; Zhao, Z.-y.; Harusawa, S.; Araki, L.; Kurihara, T.; Lille, D. M. *RNA* **2006**, *12*, 980–987.
(33) Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics*, 83rd ed.; CRC Press LLC: Boca Raton, FL, 2003.
(34) Tang, C. L.; Alexov, E.; Pyle, A. M.; Honig, B. *J. Mol. Biol.* **2007**, *366*, 1475–1496.
(35) Wilson, T. J.; Nahas, M.; Araki, L.; Harusawa, S.; Ha, T.; Lilley, D. M. *Blood Cell. Mol. Dis.* **2007**, *38*, 8–14.
(36) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **1997**, *36*, 432–450.
(37) Liu, Y.; Gregersen, B. A.; Lopez, X.; York, D. M. *J. Phys. Chem. B* **2005**, *109*, 19987–20003.
(38) Lopez, X.; Dejaegere, A.; Leclerc, F.; York, D. M.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 11525–11539.
(39) Liu, Y.; Gregersen, B. A.; Hengge, A.; York, D. M. *Biochemistry* **2006**, *45*, 10043–10053.
(40) 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.
(41) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*(18), 5010–5018.
(42) Davies, J.; Doltsinis, N.; Kirby, A.; Roussev, C.; Sprik, M. *J. Am. Chem. Soc.* **2002**, *124*, 6594–6599.
(43) Gao, J. *Acc. Chem. Res.* **1996**, *29*, 298–305.
(44) Gao, J.; Garcia-Viloca, M.; Poulsen, T. D.; Mo, Y. *Adv. Phys. Org. Chem.* **2003**, *38*, 161–181.
(45) Gregersen, B. A.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2003**, *125*, 7178–7179.
(46) Gregersen, B. A.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 7504–7513.
(47) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2005**, *1*(1), 2–13.
(48) Nam, K.; Gao, J.; York, D. M. in *Large Scale MD, Meso, and Nanoscale Modeling*; Ross, R., Ed.; ACS Symposium Series; Oxford University Press: New York, 2006.
(49) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186–195.
(50) Gao, J.; Ma, S.; Major, D.; Nam, K.; Pu, J.; Truhlar, D. *Chem. Rev.* **2006**, *106*(8), 3188–3209.
(51) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
(52) Gao, J. *Rev. Comput. Chem.* **1995**, *7*, 119–185.
(53) Gao, J. *Curr. Opin. Struc. Biol.* **2003**, *13*, 184–192.
(54) Villà, J.; Štrbj, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*(22), 11899–11904.
(55) Štrajbl, M.; Sham, Y. Y.; Villà, J.; Chu, Z.-T.; Warshel, A. *J. Phys. Chem. B* **2000**, *104*, 4578–4584.
(56) Nam, K.; Cui, Q.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 486–504.

simulations and experiment. Section 4 concludes with a summary of the key points and identifies future research directions.

2. Computational Details

2.1. Model of Hairpin Ribozyme in Solution. The crystal structure of a transition state analog with the hairpin ribozyme (PDB code 1M5O) contains two independent (nonsymmetry-related) ribozyme structures.[4] The second ribozyme structure was chosen in the present simulation study, which consists of a 92-nucleotide ribozyme strand, a 21-nucleotide substrate strand, 11 Ca^{2+} ions, and 34 crystallographic water molecules. Three of the Ca^{2+} ions and the 18 crystal waters closest to the active site were considered in the simulations. The vanadate ion between the A−1 and G+1 scissile nucleotides were replaced by a phosphate group to form a trigonal bipyramidal phosphorane geometry. Nucleotides 25 Å away from the scissile phosphate were removed, leaving a total of 77 nucleotides. Both the 5′-end and the 3′-end were terminated by a hydroxyl group.

Fully solvated MM simulations with explicit solvent under periodic boundary conditions were initially performed for the reactant state using the smooth particle mesh Ewald method[57,58] at 1 atm and 298 K to equilibrate the water molecules and ions (0.14 M NaCl). It became evident, however, that the computational effort required to treat the fully solvated system using the linear-scaling QM/MM-Ewald method[47] in conjunction with the extensive QM/MM umbrella sampling that was needed to construct the 1-D and 2-D PMF profiles was not feasible with the available resources. Consequently, it was necessary to construct a considerably smaller model system for activated dynamics simulations with stochastic boundaries.

The reduced model system was constructed as follows. All nucleotide bases were modeled in their normal protonation states, except in the simulations that considered A38 to be protonated at N_1 position. The positions of hydrogen atoms were determined using the HBUILD facility in the program CHARMM[59] (version c32a2). The resulting system was further solvated out to 30 Å with a sphere of pre-equilibrated water molecules centered at the active site. Water molecules located beyond 25 Å or within 2.5 Å from any non-hydrogen atoms of the ribozyme and ions were removed. This procedure was repeated seven times by randomly rotating the water spheres to fill the potential cavities for water occupations. A total of 41 Na^+ counterions were included. The number of Na^+ ions added was determined from analysis of the molecular dynamics (MD) trajectories of the fully solvated system under periodic boundary conditions, from which the number of ions within the 25 Å sphere from the scissile phosphate was averaged. The final system contained 2484 RNA atoms, 56 ions, and 1625 water molecules for the system containing a neutral A38 base and monoprotic phosphorane in the active site. For completeness, we set up a total of four systems with different protonation states of A38 (neutral or protonated) and the phosphorane intermediate (monoprotic/monoanionic or unprotonated/dianionic) and two systems in which either G8 or A38 base was removed.

2.2. Combined QM/MM Potential. To model the phosphate cleavage reaction, a combined QM/MM potential[51,52,60] was used in all calculations. The system was partitioned into a QM region, constituting the active site and modeled with the AM1/d-PhoT Hamiltonian,[56] and a MM region, containing the rest of the system modeled with the all-atom CHARMM27 nucleic acid force field[61,62] and TIP3P water model.[63] The AM1/d-PhoT model[56] is a reparametrized AM1/d model

for phosphoryl transfer reactions in enzymes and ribozymes based on model reactions in a new database of quantum calculations for RNA catalysis (*QCRNA*)[40,64] and has recently been applied to simulations of the full length hammerhead ribozyme.[65] Simulations were performed using the CHARMM software package[59] (version c32a2) interfaced with the MNDO97 program.[66] The QM subsystem consists of the scissile phosphodiester group and parts of the A−1 and G+1 ribose rings. The bonded QM/MM interaction (to treat division of the QM and MM subsystem across a covalent bond) was modeled using the generalized hybrid orbital (GHO) method[67] that was used at the following bond positions (asterisk indicates the GHO atom): *C1′−C2′ and *C4′−C3′ bonds in A−1, *C4′−C5′ bond in G+1, and in simulations with A38 protonated, the *C1′−N9 bond in A38. When the G8 base was included in the QM subsystem, the QM/MM boundary was set at the *C1′−N9 bond in G8. The nonbonded QM/MM van der Waals interaction was modeled with a Lennard-Jones potential with parameters for the QM subsystem optimized such that the QM/MM interactions in TIP3P water clusters reproduce corresponding density-functional results (see Supporting Information).

2.3. Molecular Dynamics Simulations. The starting systems were initially relaxed by restrained minimizations to alleviate high-energy contacts without significantly altering the active site structure. Then, stochastic boundary molecular dynamics simulations (SBMD)[68,69] were carried out to further equilibrate the system and to sample the configuration space in the free energy simulations. In SBMD, the aqueous environment was represented by a 25 Å sphere of water molecules centered at the geometric center of the active site, defined as the center of QM subsystem. A solvent boundary potential[68] was applied to mimic the effect of external solvent and to act as a confining potential to prevent diffusion of water outside the sphere. Nucleotides more than 25 Å away from the center of the sphere were held fixed during the simulations. The nucleotides in the region from 21 to 25 Å from the center were harmonically restrained to the initial X-ray crystal positions with force constants determined from the crystallographic B-factor values according to the relation

$$f_i = \frac{3k_\mathrm{B}T}{\langle \Delta r_i^2 \rangle} = 3k_\mathrm{B}T \frac{8\pi^2}{3B_i} \tag{1}$$

where f_i is the restoring harmonic force constant on atom i, k_B is Boltzmann's constant, T is temperature, and B_i is the isotropic crystallographic B-factor for the heavy (non-hydrogen) atom i. The force constants thus derived were multiplied by a scaling factor that varies linearly from 0 to 1 between 21 and 25 Å.[69] In the MD simulations, the atoms within a sphere of 21 Å from the center were treated by Newtonian dynamics, and the atoms in the radial range between 21 and 25 Å were treated by Langevin dynamics, which also serves as a thermal bath to maintain constant temperature during the simulation. Atomic positions were propagated using the leapfrog Verlet algorithm with 1 fs integration time.[70] Initially, the systems were slowly heated from 8 to 298 K over 11.6 ps and subsequently equilibrated at 298 K for 100 ps. During the equilibration, harmonic restraints (force constant of 500.0 kcal/mol-Å^2) were used for the nucleophilic $O_{2'}$−P and leaving group P−$O_{5'}$ bonds to ensure the integrity of the transition state phosphorane geometry. Throughout MD simulations, the nonbonded electrostatic and van der Waals interactions were evaluated

(57) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing, L.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*(19), 8577−8593.
(58) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155−179.
(59) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(60) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(61) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86−104.
(62) MacKerell, A. D., Jr.; Banavali, N. K. *J. Comput. Chem.* **2000**, *21*, 105−120.

(63) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(64) QCRNA; http://theory.chem.umn.edu/Database/QCRNA.
(65) Lee, T.-S.; Silva-Lopez, C.; Martick, M.; Scott, W. G.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 325−327.
(66) Thiel, W. *MNDO97, version 5.0*; University of Zurich, 1998.
(67) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.
(68) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(69) Brooks, C. L., III; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843−865.
(70) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.

Scheme 1. Scheme of the Dianionic Reaction Mechanism for the Self-Cleavage Catalyzed by Hairpin Ribozyme
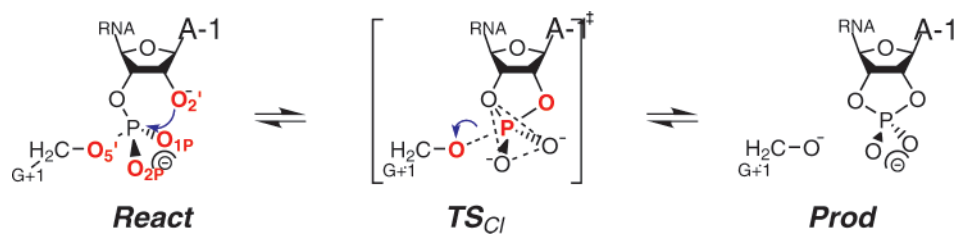
without truncation (no-cutoff), and all bond lengths involving hydrogen atoms in the MM subsystem were constrained using the SHAKE algorithm.[71] During the umbrella sampling MD simulations, harmonic restraining potentials with a force constant of 12.0 kcal/mol-Å^2 were applied to the Na^+ ions in the radial range between 21 and 25 Å at their equilibrated positions in accord with the restraints applied to other nucleotide atoms in that range.

2.4. Free Energy Simulations. The potentials of mean force (PMFs) for the ribozyme catalyzed cleavage reaction were determined using umbrella sampling[72] SBMD simulations and the QM/MM potential without electrostatic or van der Waals cutoff as described above. In the present study, the PMF is represented as a function of one or two reaction coordinates. In the former, the reaction coordinate is defined as

$$\zeta_1 = R_{P-O_{5'}} - R_{P-O_{2'}} \tag{2}$$

where $R_{P-O_{5'}}$ and $R_{P-O_{2'}}$ are the distances of the ribosyl $O_{5'}$ and $O_{2'}$ atoms from the P atom, respectively. When two reaction coordinates are considered, the first reaction coordinate ζ_1 is the same as that of eq 2, and the second reaction coordinate ζ_2 is defined as

$$\zeta_2 = R_{Donor-H} - R_{Acceptor-H} \tag{3}$$

where $R_{Donor-H}$ and $R_{Acceptor-H}$ are the distances of the proton to be transferred from the donor and to the acceptor atom, respectively. The one-dimensional PMF models the free energy change for the nucleophilic substitution reaction itself, whereas the two-dimensional PMF addresses the question of concerted versus stepwise mechanism involving both substitution and proton transfer in a general acid−base catalysis. A series of umbrella sampling MD simulations were carried out to span the entire range of the reaction coordinate(s) from the reactant to the products. To enhance sampling efficiency, a harmonic restraining potential was applied with a force constant ranging between 30.0 and 110.0 kcal/mol-Å^2 based on the shape of the PMF profile for each system. This ensures sufficient overlap of the probability distribution with neighboring windows. Each umbrella sampling window was equilibrated for 17 ps followed by 30−50 ps of configuration sampling with data collected at every dynamics integration step. The weighted histogram analysis method (WHAM)[73,74] was used to combine the sampled configurations to compute the unbiased PMFs as a function of the reaction coordinate(s). Further details for individual simulations are provided in the Supporting Information.

We have carried out 8 independent 2-D PMF simulations for the nucleophilic substitution step and the exocyclic cleavage step with either the nonbridging oxygen (O_{1P} or O_{2P}) acting as a proton acceptor from $O_{2'}$ nucleophile and with protonated and unprotonated A38 base, respectively. However, the free energy values obtained from the 2-D profiles can only be considered as an estimation for the barrier height, and a more rigorous evaluation of the contribution from the two reaction

coordinates is required.[75,76] A 1-D reaction coordinate was constructed as a linear combination of the two reaction coordinates in the 2-D PMF, which corresponds to the minimum free energy path on the 2-D surface. Free energy simulations along this 1-D reaction coordinate were then performed for 100 ps to construct the 1-D PMF profile.[77] To estimate the errors of computed PMFs, the 100 ps of sampled data were partitioned into 20 ps segments and PMFs were separately computed for each 20 ps data set. From the resulting PMF profiles, the standard deviation was obtained and used as a statistical error estimate. We have also tested the effects of G8 and A38 base deletion by repeating 1-D PMF simulations with a hydrogen atom replacing G8 and A38 base, respectively.

2.5. Active Site Model Calculations. To further validate the results using the AM1/d-PhoT model at the rate-determining exocyclic cleavage transition state, a series of density functional calculations was carried out on truncated active site models. The active site model for the dianionic reaction (24 atoms) consists of a nucleophilic ribose ring with $O_{2'}$ deprotonated, a scissile phosphate, and a leaving ethyl group; the activation energy at the transition state was taken from the QCRNA database.[40] The active site model for the in-line monoanionic reaction (49 atoms) consists of two ribose rings for A−1 and G+1 but with the bases removed, the scissile phosphate, and the G8 base with N9 atom capped by a hydrogen atom. For the A38 mediated monoanionic reaction, the model (65 atoms) includes an additional protonated A38 base. Except for the dianionic reaction that was taken directly from the QCRNA database, the starting geometry for each model reaction was the equilibrated geometry of the intermediate state followed by energy minimizations at the semiempirical AM1/d-PhoT QM/MM level. The geometry optimizations were carried out at the reactant state and the exocyclic cleavage transition state. All structures were optimized in the gas phase with Kohn−Sham density functional theory (DFT) methods using the hybrid exchange functional of Becke[78,79] and the Lee−Yang−Parr correlation functional[80] (B3LYP) at the 6-31+G(d) basis set level. Vibrational frequency calculations were performed to confirm the correct topological nature of the minima and transition states. To mimic the restraints applied by the hairpin ribozyme, several atoms were kept frozen through all geometry optimizations; the frozen atoms were $O_{4'}$ of A−1, $C_{3'}$ of G+1, N9 of A38 and G8 base, respectively, which are reasonably separated from the scissile phosphate. Electronic energies were further refined by single-point energy calculations using 6-31++G(d,p) basis set. The same procedure was followed at the semiempirical AM1/d-PhoT level to assess the reliability of the model. All DFT calculations were performed with the GAUSSIAN03 suite of programs,[81] and AM1/d-PhoT model calculations were carried

(71) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(72) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(73) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(74) Rajamani, R.; Naidoo, K. J.; Gao, J. *J. Comput. Chem.* **2003**, *24*, 1775−1781.
(75) Truhlar, D. G.; Garrett, B. C. *Acc. Chem. Res.* **1980**, *13*, 440−448.
(76) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*(12), 5828−5833.
(77) To test the effects of electrostatic solvation outside of 25 Å water sphere on the computed barrier heights, we have repeated QM/MM 1-D umbrella sampling simulations along with COSMO-reaction field (see Gregersen, B. A.; York, D. M. *J. Phys. Chem. B* **2005**, *109*, 536) for the in-line monoanionic O_{2P} pathway and A38(+) mediated O_{1P} pathway, respectively. The computed barrier heights agree with the results without reaction field within 1−2 kcal/mol.
(78) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(79) Becke, A. D. *J. Chem. Phys.* **1993**, *98*(7), 5648−5652.
(80) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
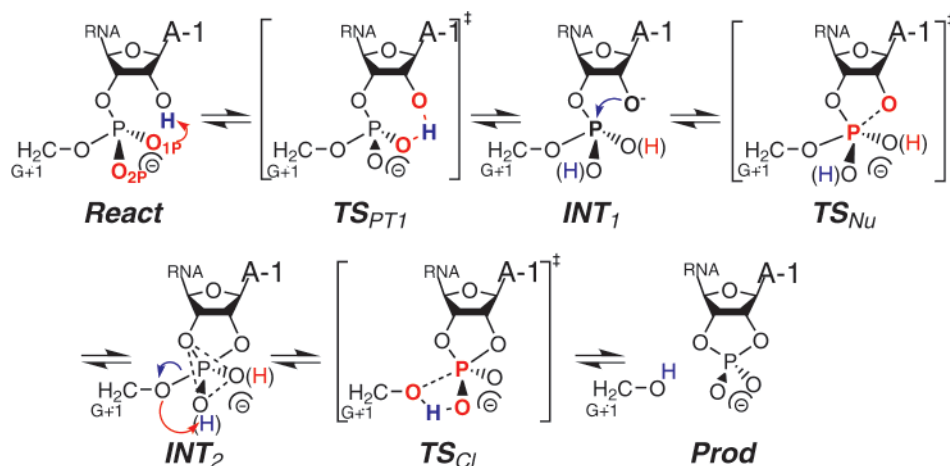(81) Frisch, M. J.; et al. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

Scheme 2. Scheme of the In-Line Monoanionic Reaction Mechanisms for the Self-Cleavage Catalyzed by Hairpin Ribozyme[a]

React   **TS$_{PT1}$**   **INT$_1$**   **TS$_{Nu}$**

INT$_2$ **TS$_{Cl}$**   **Prod**

[a] O$_{1P}$ and O$_{2P}$ pathways are denoted as blue and red colored hydrogens in phosphorane intermediate, respectively.

Scheme 3. Scheme of the A38 Mediated Monoanionic Reaction Mechanisms for the Self-Cleavage Catalyzed by Hairpin Ribozyme[a]

React   **TS$_{PT1}$**   **INT$_1$**   **TS$_{Nu}$**

INT$_2$ **TS$_{Cl}$**   **Prod**
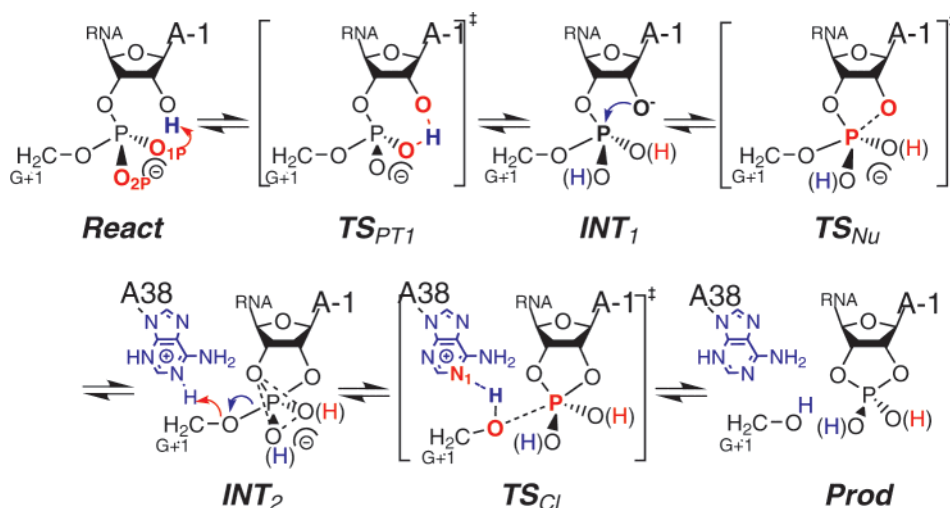
[a] O$_{1P}$ and O$_{2P}$ pathways are denoted as blue and red colored hydrogens in phosphorane intermediate, respectively.

out using the MNDO97 program.[66] The DFT-optimized geometries and energies are provided in the Supporting Information.

3. Results and Discussion

This section presents the results from free energy simulations and discusses their implications into the mechanism of the hairpin ribozyme catalyzed self-cleaving reaction. Several mechanisms are explored to provide insight into the role of different conserved active site residues, in particular, G8 and A38, which have been suggested to be involved as general acid–base catalysts. Three general mechanisms are considered: (1) a dianionic mechanism whereby it is assumed that the nucleophile has been previously activated and that the leaving group is unprotonated, (2) a monoanionic mechanism whereby a nonbridging phosphoryl oxygen acts as the general base and acid, and (3) an A38(+) mediated mechanism where the cleavage step is facilitated by proton donation from a nearby adenine (A38) residue protonated at the N1 position. These mechanisms are illustrated in Scheme 1–3. Thermodynamic and kinetic results from the free energy simulations are summarized in Table 1, and the averages for key structural parameters are listed in Tables 2 and 3. A more detailed decomposition of the nucleobases and solvent contributions in stabilizing the transition

states and reactive intermediates are provided in Table 4 and in the Supporting Information. The 2-D PMF profiles for the in-line monoanionic mechanism are shown in Figure 1. The structures of the phosphorane intermediate and transition state in the active site are illustrated in Figures 2 and 3. The 1-D PMF profiles are compared in Figure 4.

3.1. Dianionic Reaction. The dianionic mechanism (Scheme 1) considered here assumes that the deprotonation of nucleophilic O$_{2'}$ and the protonation of leaving group O$_{5'}$ occur sequentially and are not rate-controlling and therefore can be treated independently from the transphosphorylation reaction step. We consider here only the transphosphorylation step where the 2′ hydroxyl nucleophile has already been activated (for example by a hydroxide ion coordinated to an active site base) and the leaving 5′ alkoxide group has not yet been protonated (for example from the solvent or a nucleobase) after the exocyclic cleavage. As shown by recent biochemical studies that explore the pH-rate profiles for unmodified and chemically modified hairpin ribozyme at different temperatures,[9] the dianionic mechanism is not likely to be the most physiologically relevant. Nevertheless, characterization of this mechanism provides useful reference for comparison and elucidation of the key factors that govern catalytic proficiency.

Table 1. Computed Reaction Free Energies and Barrier Heights for the Dianionic, the In-Line Monoanionic, and the A38(+) Mediated Monoanionic Reaction Mechanisms[a]

		TS_{PT1}	INT_1	TS_{Nu}	INT_2	TS_{Cl}	$TS_{Cl}^{DFT,f}$	prod
dianionic[b]						15.6 ± 2.0 (29.3)	14.9[g] (28.6)	
in-line[c]	O_{1P}	15.0 ± 0.6 (15)	13.5 ± 0.5 (14)	13.9 ± 0.8 (15)	11.8 ± 0.5 (13)	22.7 ± 0.8 (25)	25.0 (27)	−1.9 ± 1.5 (−7)
	O_{2P}	10.5 ± 0.7 (12)	9.2 ± 0.7 (11)	12.3 ± 1.3 (14)	10.6 ± 1.6 (13)	18.0 ± 1.5 (21)	20.5 (24)	−3.9 ± 1.2 (−5)
	O_{2P}+dG8	10.7 ± 0.5	9.4 ± 0.6	14.1 ± 0.8	13.6 ± 0.6	21.6 ± 0.9	24.1	−1.7 ± 1.4
	O_{2P}+dA38	10.3 ± 0.2	9.0 ± 0.1	11.0 ± 0.6	9.6 ± 0.6	18.6 ± 1.4	21.1	−5.8 ± 1.1
A38(+)[d]	O_{1P}	17.2 ± 0.7 (20.9)	15.6 ± 0.8 (19.3)	17.1 ± 1.2 (20.8)	14.5 ± 1.4 (18.2)	24.4 ± 1.6 (28.1)	16.4 (20.1)	−9.3 ± 2.5 (−5.6)
	O_{2P}	10.0 ± 0.4 (13.7)	9.4 ± 0.5 (13.1)	16.6 ± 1.3 (20.3)	12.0 ± 2.0 (15.7)	22.2 ± 2.1 (25.9)	14.4 (18.1)	−4.5 ± 2.1 (−0.8)
expt[e]							20−21	~1

[a] All free energies and standard deviations are given in kcal/mol. [b] The dianionic mechanism assumes that the 2′ OH has been previously deprotonated, possibly by solvent, in a step that is not rate-controlling. An estimate of the free energy to deprotonate the 2′ OH at neutral pH is made based on the pK_a value of 17.1 for 2-propanol[33] (13.7 kcal/mol), and the shifted free energy values based on this estimate are shown in parenthesis. [c] The values are those derived from the 1-D PMF profiles described in the text, whereas the values in parenthesis are estimated directly from the 2-D PMF profiles given in Figure 1, and the abasic reactions of G8 and A38 are denoted as dG8 and dA38, respectively. [d] The A38(+) mediated mechanism assumes that the A38 has been previously protonated in a step that is not rate-controlling. If we make an estimate of the free energy to protonate A38, assuming a pK_a value of 4.3 for a free adenine base[33] (3.7 kcal/mol), the shifted free energy values based on this estimate are shown in parenthesis. [e] Experimental values are taken from references. [f] Free energy barrier is for the TS_{Cl} state after applying DFT correction based on active site models. [g] Density functional results on the dianionic reaction are taken from QCRNA database.[40]

Table 2. Key Average Geometric Quantities and Standard Deviations Obtained from QM/MM Molecular Dynamics Simulations at Different Stages along the Reaction Path: Proton-Transfer Transition State(TS_{PT1}), Phosphorane Intermediate(INT_2), and Exocyclic Cleavage Transition State(TS_{Cl})[a]

path		$O_{2'}-P$	$P-O_{5'}$	$O_{2'}-P-O_{5'}$	$O_2-H_{2'}$	$O_{XP}-H_{2'}$[b]	$O_{5'}-H_{2'}$
In-line monoanionic mechanism							
O_{1P}	react	3.05 ± 0.13	1.67 ± 0.03	154 ± 6	1.01 ± 0.02	2.42 ± 0.08	
	TS_{PT1}	3.05 ± 0.07	1.63 ± 0.03	141 ± 11	1.28 ± 0.05	1.22 ± 0.04	
	INT_2	1.85 ± 0.05	1.71 ± 0.03	166 ± 6		1.01 ± 0.03	
	TS_{Cl}	1.68 ± 0.03	2.32 ± 0.12	163 ± 5		1.02 ± 0.03	1.99 ± 0.13
O_{2P}	react	2.94 ± 0.11	1.64 ± 0.03	160 ± 6	1.01 ± 0.04	2.11 ± 0.08	
	TS_{PT1}	3.10 ± 0.08	1.63 ± 0.03	141 ± 7	1.25 ± 0.06	1.23 ± 0.06	
	INT_2	1.83 ± 0.04	1.74 ± 0.03	161 ± 5		1.00 ± 0.03	
	TS_{Cl}	1.69 ± 0.03	2.36 ± 0.19	162 ± 5		1.02 ± 0.03	2.03 ± 0.17
A38(+) mediated monoanionic mechanism							
O_{1P}	react	2.87 ± 0.11	1.65 ± 0.03	160 ± 5	1.01 ± 0.03	2.15 ± 0.08	
	TS_{PT1}	3.09 ± 0.08	1.64 ± 0.03	120 ± 8	1.27 ± 0.05	1.23 ± 0.04	
	INT_2	1.87 ± 0.05	1.70 ± 0.03	166 ± 4		1.01 ± 0.03	
	TS_{Cl}	1.71 ± 0.04	2.09 ± 0.10	167 ± 4		1.02 ± 0.04	2.43 ± 0.25
O_{2P}	react	3.08 ± 0.21	1.64 ± 0.03	149 ± 10	1.01 ± 0.03	2.11 ± 0.09	
	TS_{PT1}	3.11 ± 0.08	1.64 ± 0.03	114 ± 8	1.29 ± 0.05	1.19 ± 0.04	
	INT_2	1.80 ± 0.05	1.66 ± 0.03	134 ± 6		1.01 ± 0.03	
	TS_{Cl}	1.70 ± 0.04	2.04 ± 0.10	159 ± 6		1.02 ± 0.03	2.48 ± 0.20

[a] Units for distance and angle are Å and degree, respectively. [b] O_{XP} is the nonbridging oxygen acting as a general base in the nucleophilic attack step (i.e., the O_{1P} atom in the O_{1P} pathways and the O_{2P} atom in the O_{2P} pathways).

1-D PMF Profiles. Free energy simulations have been carried out on the transphosphorylation step using a total of 27 separate simulations that span the entire reaction. The reaction coordinate is defined by eq 2. The reaction proceeds via a single transition state corresponding to the exocyclic cleavage (Scheme 1) and the computed barrier height is 15.6 kcal/mol (Table 1). This can be compared with a barrier of 30.1 kcal/mol computed for the uncatalyzed reaction estimated from a nonenzymatic model for transesterification proceeding through a dianionic methyl ethylene phosphorane transition state in water.[48] The ribozyme environment decreases the barrier by 14.5 kcal/mol for the dianionic mechanism.

Density-Functional Calculations. The density-functional result for the dianionic in-line mechanism is in close agreement with that of the AM1/d-PhoT model. The adiabatic DFT barrier for the dianionic model reaction is 28.8 kcal/mol in the gas phase, whereas the corresponding AM1/d-PhoT result is 29.5 kcal/mol. If this 0.7 kcal/mol difference is applied as a correction to the AM1/d-PhoT barrier from the 1-D PMF profile, the barrier decreases from 15.6 to 14.9 kcal/mol (Table 1). For the dianionic reaction mechanisms, the AM1/d-PhoT model is in good accord with the DFT results.[56]

pK_a of the 2′ OH. It remains to obtain an estimate of the free energy required to activate the nucleophile, for example, from a pK_a calculation using a QM/MM free energy perturbation approach[82] or quantum/continuum calculations.[41,83] A theoretical prediction of the pK_a of the 2′ OH requires careful calibration of the theoretical model and quantum corrections, and is a topic of future work. Nonetheless, a rough estimate of the free energy required to activate 2′ OH can be obtained from the pK_a of a secondary alcohol such as 2-propanol (pK_a 17.1[33]).

(82) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715−17733.

(83) Lyne, P. D.; Karplus, M. *J. Am. Chem. Soc.* **2000**, *122*, 166−167.

Table 3. Average Distances and Standard Deviations Involving G8, A9, A10, and A38 Obtained from QM/MM Molecular Dynamics Simulations at Different Stages along the Reaction Path: Proton-Transfer Transition State (TS_{PT1}), Phosphorane Intermediate (INT_2), and Exocyclic Cleavage Transition State (TS_{Cl})[a]

		G8		A9	A10	A38	
path		$O_{2'}-N_1$	$O_{2P}-N_2$	$O_{1P}-N_6$	$O_{1P}-N_6$	$O_{5'}-N_1$	$O_{1P}-N_6$
				In-line monoanionic mechanism			
	react	3.98 ± 0.29	3.68 ± 0.35	3.21 ± 0.23	4.55 ± 0.29	6.61 ± 0.30	7.87 ± 0.46
O_{1P}	TS_{PT1}	4.14 ± 0.32	3.94 ± 0.63	3.89 ± 0.24	5.41 ± 0.68	7.55 ± 0.44	8.58 ± 0.75
	INT_2	3.10 ± 0.22	3.31 ± 0.29	3.52 ± 0.16	4.56 ± 0.52	7.04 ± 0.27	7.66 ± 0.38
	TS_{Cl}	2.70 ± 0.12	3.86 ± 0.32	3.28 ± 0.22	5.02 ± 0.41	7.42 ± 0.27	8.85 ± 0.41
	react	2.48 ± 0.08	3.38 ± 0.23	3.13 ± 0.17	5.37 ± 0.35	5.09 ± 0.29	5.56 ± 0.42
O_{2P}	TS_{PT1}	2.59 ± 0.12	3.02 ± 0.13	3.12 ± 0.15	4.91 ± 0.32	4.43 ± 0.34	4.07 ± 0.57
	INT_2	2.70 ± 0.12	3.20 ± 0.21	3.10 ± 0.14	5.27 ± 0.36	5.00 ± 0.31	5.54 ± 0.48
	TS_{Cl}	2.64 ± 0.11	3.56 ± 0.27	3.17 ± 0.19	5.39 ± 0.32	6.54 ± 0.23	7.19 ± 0.33
				A38(+) mediated monoanionic mechanism			
	react	3.76 ± 0.24	3.30 ± 0.24	3.47 ± 0.40	4.90 ± 0.29	2.94 ± 0.11	2.99 ± 0.21
O_{1P}	TS_{PT1}	4.65 ± 0.34	5.99 ± 0.53	4.91 ± 0.42	6.48 ± 0.54	2.95 ± 0.14	3.93 ± 0.40
	INT_2	3.01 ± 0.18	3.35 ± 0.24	4.30 ± 0.28	6.13 ± 0.25	2.88 ± 0.09	2.97 ± 0.14
	TS_{Cl}	2.79 ± 0.19	3.52 ± 0.28	3.57 ± 0.26	6.06 ± 0.42	2.61 ± 0.06	3.06 ± 0.22
	react	3.39 ± 0.83	4.36 ± 0.51	5.27 ± 0.56	5.01 ± 0.35	3.17 ± 0.43	3.13 ± 0.28
O_{2P}	TS_{PT1}	2.59 ± 0.11	3.98 ± 0.25	5.66 ± 0.43	5.18 ± 0.38	4.11 ± 0.22	2.98 ± 0.30
	INT_2	3.02 ± 0.17	3.22 ± 0.20	3.80 ± 0.24	4.36 ± 0.29	3.19 ± 0.22	2.98 ± 0.15
	TS_{Cl}	3.85 ± 0.49	4.02 ± 0.40	4.33 ± 0.31	4.55 ± 0.31	2.64 ± 0.06	2.87 ± 0.09

[a] Unit for distance is Å.

Table 4. Contributions of Key Nucleobases and Water Molecules in Stabilizing the Intermediate and Transition States from In-Line Monoanionic and A38(+) Mediated Mechanisms[a]

	in-line O_{1P}			in-line O_{2P}			A38(+)+O_{1P}[b]			A38(+)+O_{2P}[b]		
range (Å)[c]	TS_{PT1}	INT_2	TS_{Cl}	TS_{PT1}	INT_2	TS_{Cl}	TS_{PT1}	INT_2	TS_{Cl}	TS_{PT1}	INT_2	TS_{Cl}
						Ribozyme strand						
G8	4.9	−2.9	2.5	1.2	−0.8	5.3	17.0	−1.7	1.7	−9.3	−4.9	7.8
A9	2.1	0.1	6.8	1.0	−1.1	3.4	4.7	5.1	4.6	−1.0	−1.7	−0.8
A10	−2.8	0.7	2.0	0.0	0.2	0.5	−2.3	−1.5	−0.8	−0.3	0.0	0.4
C25	0.9	−1.2	−1.5	0.2	0.6	0.0	−1.9	−2.4	−2.9	−2.8	−1.1	−1.1
A38	−0.2	0.1	0.3	−2.2	1.0	2.0	−0.3	0.2	−16.8	−5.5	−18.3	−41.5
						Substrate strand						
A−1	−0.7	−1.0	0.2	−2.1	−0.3	−0.1	−1.5	0.5	0.5	0.2	0.9	1.0
G+1	−5.0	2.0	3.2	−1.1	0.5	2.8	−1.5	1.6	0.3	7.7	4.6	8.4
U+2	0.1	−1.4	−3.0	2.1	0.6	−1.6	1.2	1.3	−0.1	1.2	1.8	1.6
C+3	5.6	−2.9	−5.7	3.6	5.2	3.9	−0.3	−1.5	−2.3	−1.2	1.3	1.2
						Water molecules						
water (1−3)	0.4	0.2	0.5	0.5	−0.3	−4.3	0.5	0.4	0.5	−1.8	−1.1	1.5
(3−5)	8.9	2.3	−18.9	−0.2	−8.1	−21.6	−14.6	−12.6	−1.1	14.9	16.2	13.3

[a] Energy values are computed as $\Delta E_X = \langle E_X \rangle - \langle E_{React} \rangle$, where X denotes intermediate and transition states, E_X is the interaction energy in kcal/mol of the specified base or water molecules with the scissile phosphate group, and react is reactant state, in the QM/MM molecular dynamics simulations. [b] In the A38(+) mediated mechanism, A38 base, which is protonated at the N_1 site, is excluded from the QM subsystem for the consistent comparison with other mechanisms and represented by point charges. [c] The distance range in Å of the water molecules from the center of quantum region.

If the pK_a shift due to the ribozyme environment is neglected (a considerable approximation), this results in a 13.7 kcal/mol free energy requirement at 300 K to activate the nucleophile at neutral pH, shifting the rate-controlling barrier from 15.6 to 29.3 kcal/mol. This estimated barrier is too high in comparison to the experimental estimate of 20−21 kcal/mol,[2,7,9,84] which is to be expected because RNA cleavage is thought to proceed via a phosphorane monoanion.[36,42,85] Nonetheless, it is of interest to note that the ribozyme environment considerably lowers the activation barrier for the dianionic mechanism relative to the uncatalyzed reaction and likely contributes to the catalytic proficiency.

3.2. In-Line Monoanionic Reaction. In the in-line monoanionic mechanism (Scheme 2), nonbridging oxygens play a dual role as a proton acceptor in the nucleophilic attack by the $O_{2'}$ atom, forming a monoprotic phosphorane intermediate, and a

(84) Fedor, M. J. *Biochemistry* **1999**, *38*, 11040−11050.
(85) Breslow, R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1208−1211.

proton donor to the leaving group ($O_{5'}$) in the exocyclic cleavage step. Because the two nonbridging oxygens, O_{1P} and O_{2P} (*proR* and *proS* O, respectively), experience different electrostatic environments, we have tested both as potential acid/base catalysts. The free energy simulations have been divided into two separate steps: (1) the nucleophilic attack along with the proton transfer from $O_{2'}$ onto the nonbridging oxygen and (2) the exocyclic cleavage of the P−$O_{5'}$ bond with the transfer of proton on the nonbridging oxygen to the $O_{5'}$ atom. To provide insight into the extent to which the nucleophilic attack/cleavage and proton transfers are concerted or stepwise, 2-D umbrella sampling calculations were performed. Each step (nucleophilic attack or exocyclic cleavage) requires two umbrella sampling reaction coordinates defined by eqs 2 and 3. For the O_{1P} pathway, for example, a total of 212 and 183 separate simulations (windows) have been carried out for the nucleophilic attack and cleavage steps, respectively. Each simulation window was performed for 17 ps of equilibration, followed by 50 ps of
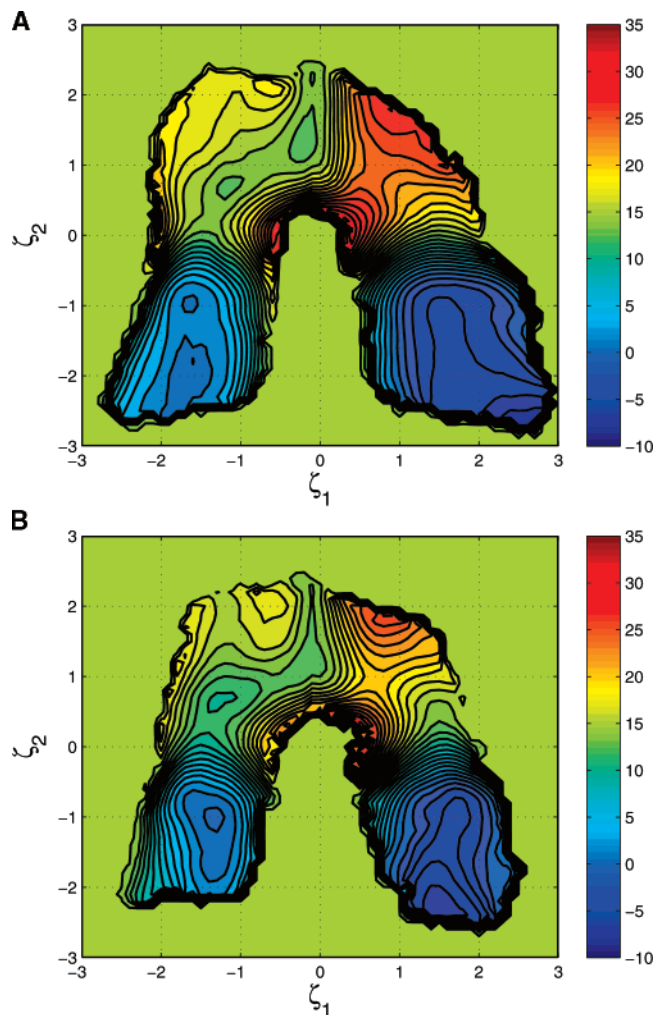
Figure 1. Computed two-dimensional potentials of mean force for the in-line monoanionic mechanism, accompanied by (A) the O_{1P} proton-transfer pathway, and (B) the O_{2P} proton-transfer pathway. The abscissa defines the reaction coordinate for the nucleophilic substitution, $\zeta_1 = R(P-O_{5'}) - R(P-O_{2'})$, and the ordinate depicts the reaction coordinate for the proton-transfer process, ζ_2. The left-hand side region, where $\zeta_1 < 0.0$ Å, corresponds to the coordinate $\zeta_2 = R(O_{2'} - H_{2'}) - R(O_{NB} - H_{2'})$, whereas in the right-hand side, where $\zeta_1 > 0.0$ Å, the ordinate is defined by $R(O_{5'} - H_{2'}) - R(O_{NB} - H_{2'})$. O_{NB} is for the O_{1P} proton transfer in (A), and for the O_{2P} proton transfer in (B). Free energies are given in kcal/mol.

sampling, resulting in a total of 26 ns of MD simulation. The same procedure was repeated for the O_{2P} pathway.

2-D PMF Profile for the O_{1P} Pathway. Figure 1A presents the two-dimensional PMF profile for the hairpin ribozyme transesterification reaction, in which the nonbridging O_{1P} atom plays the role of both general base and general acid. The first reaction step of the O_{1P} pathway (the left half of the 2-D PMF in Figure 1A) is almost stepwise: the proton on the $O_{2'}$ hydroxyl group is first transferred to O_{1P} followed by the nucleophilic attack of $O_{2'}$ at phosphorus. The computed free energy barrier from Figure 1A is about 15 kcal/mol for the O'_2 proton-transfer step, which is slightly higher than the barrier for the nucleophilic attack (Table 1). The free energy of the phosphorane intermediate relative to the reactant state is about 13 kcal/mol. The second reaction step (the right half of the 2-D PMF in Figure 1A) exhibits only one transition state, corresponding to the exocyclic cleavage of the $P-O_{5'}$ bond followed by barrierless proton transfer from O_{1P} to the $O_{5'}$ leaving group. The free energy

barrier for the exocyclic cleavage is about 12 kcal/mol from the intermediate state (Figure 1A). The overall free energy barrier from the 2-D profile for the cleavage reaction is 25 kcal/mol (Table 1).

2-D PMF Profile for the O_{2P} Pathway. The reaction path via the nonbridging O_{2P} oxygen shows a similar profile (Figure 1B). The first step of the O_{2P} pathway (the left half of the 2-D PMF in Figure 1B) indicates the proton transfer has a lower barrier than that of the O_{1P} pathway by about 3 kcal/mol. The barrier for the nucleophilic attack to the phosphorus center is about 14 kcal/mol relative to the reactant state (Table 1). The intermediate state is also about 13 kcal/mol higher than the reactant state. The results suggest that the proton transfer to the O_{2P} position is favorable relative to the proton transfer to the O_{1P} oxygen by $\sim$3 kcal/mol. This arises from the differential stabilization of accumulated charge on the nonbridging oxygens at the scissile phosphate, mainly through specific hydrogen bonds to conserved residues such as G8, A9, and A38, and water solvation in the asymmetric ribozyme active site (Figure 2A and B).[21,22,86,87] These interactions facilitate proton transfer to O_{2P}. Table 2 compares several key average distances and angles in the reactions, and Table 3 shows that the G8 base makes close contact with $O_{2'}$ at the monoprotic phosphorane intermediate in the O_{2P} pathway.

The barrier heights for the exocyclic cleavage step differ between the O_{1P} and O_{2P} reaction pathways, despite the overall similar shape of their PMF profiles (Figure 1). The barrier height from the intermediate to the transition state is about 8 kcal/mol in the O_{2P} pathway, which is about 4 kcal/mol lower than that for the O_{1P} pathway. The overall barrier for the entire cleavage step is approximately 21 kcal/mol based on the 2-D PMF profile (Table 1). The difference in the barrier heights can be rationalized by the corresponding transition state geometries (Figure 3A and B). From Figure 3B, the $H-O_{2P}$ bond is aligned adjacent and antiparallel to the elongated $P-O_{5'}$ bond. These dipoles interact favorably to provide additional stabilization on the polarization of $P-O_{5'}$ bond. The 2-D PMF plot for the O_{2P} pathway (Figure 1B) exhibits a more gradual rise in free energy around 1.0 Å along the proton-transfer reaction coordinate (ζ_2) relative to that of the O_{1P} pathway.

The transition state for the in-line monoanionic O_{2P} pathway is preferentially stabilized by solvent compared with the O_{1P} pathway (Table 4). This is due to the orientations of the $O_{5'}$ leaving group and unprotonated nonbridging oxygens that share a formal -1 charge in the rate-controlling transition state. In the O_{2P} pathway, solvation is more effective due to the relative orientations of these groups, which leave them more solvent exposed (see Figure 3B): in the first solvation sphere (out to 5.0 Å), water stabilizes the O_{1P} and O_{2P} pathways by -18.4 and -25.9 kcal/mol, respectively. The preferential solvent stabilization of the rate-controlling TS_{Cl} transition state for the O_{2P} pathway relative to that of the O_{1P} pathway is also observed for the INT_2 phosphorane intermediate, and underscores the importance of solvation in discerning specific reaction pathways.

1-D PMF Profiles. It is worthwhile to briefly clarify the differences in the free energy values obtained from the 2-D PMF

(86) Park, H.; Lee, S. *J. Chem. Theory Comput.* **2006**, *2*, 858−862.
(87) Rhodes, M. M.; Réblová, K.; Sponer, J.; Walter, N. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13380−13385.
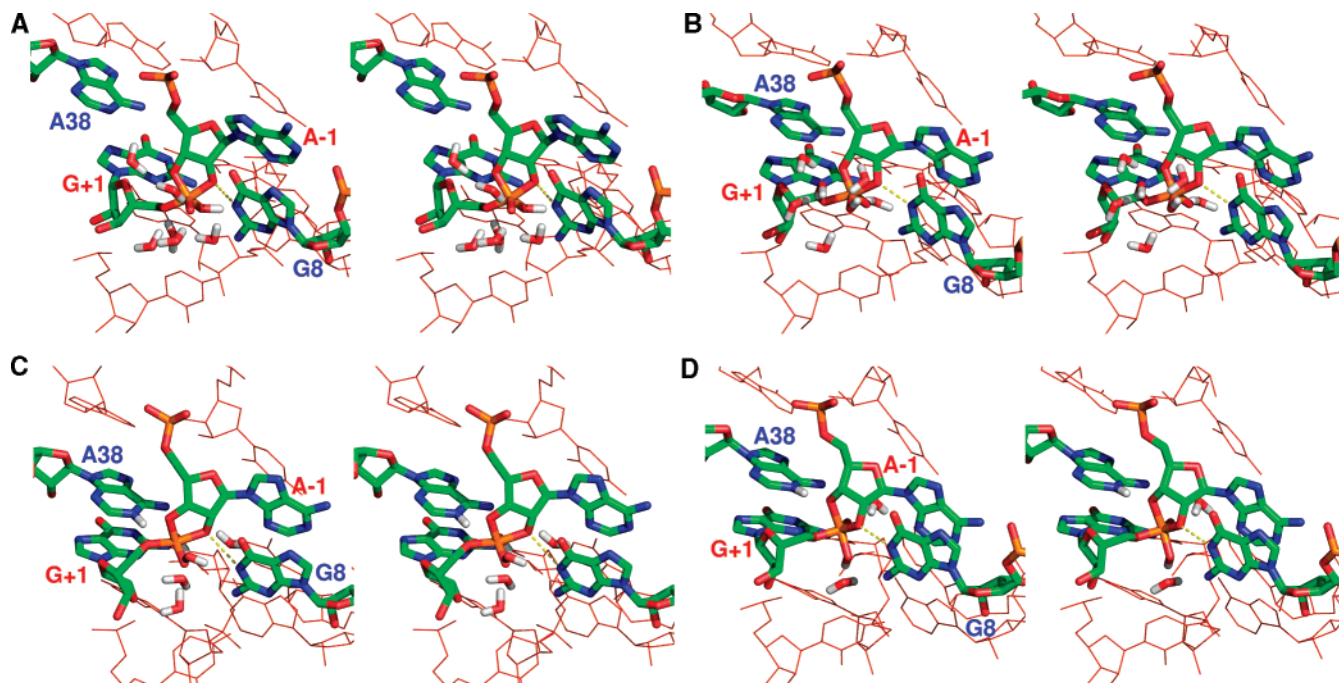
Figure 2. Stereoview of active site snapshot of the monoprotic phosphorane intermediate from the in-line monoanionic pathway via (A) O_{1P} atom and (B) O_{2P} atom, and the A38(+) mediated monoanionic pathways via (C) O_{1P} atom and (D) O_{2P} atom. Water molecules that interact with nonbridging oxygens and the leaving group $O_{5'}$ are also displayed.
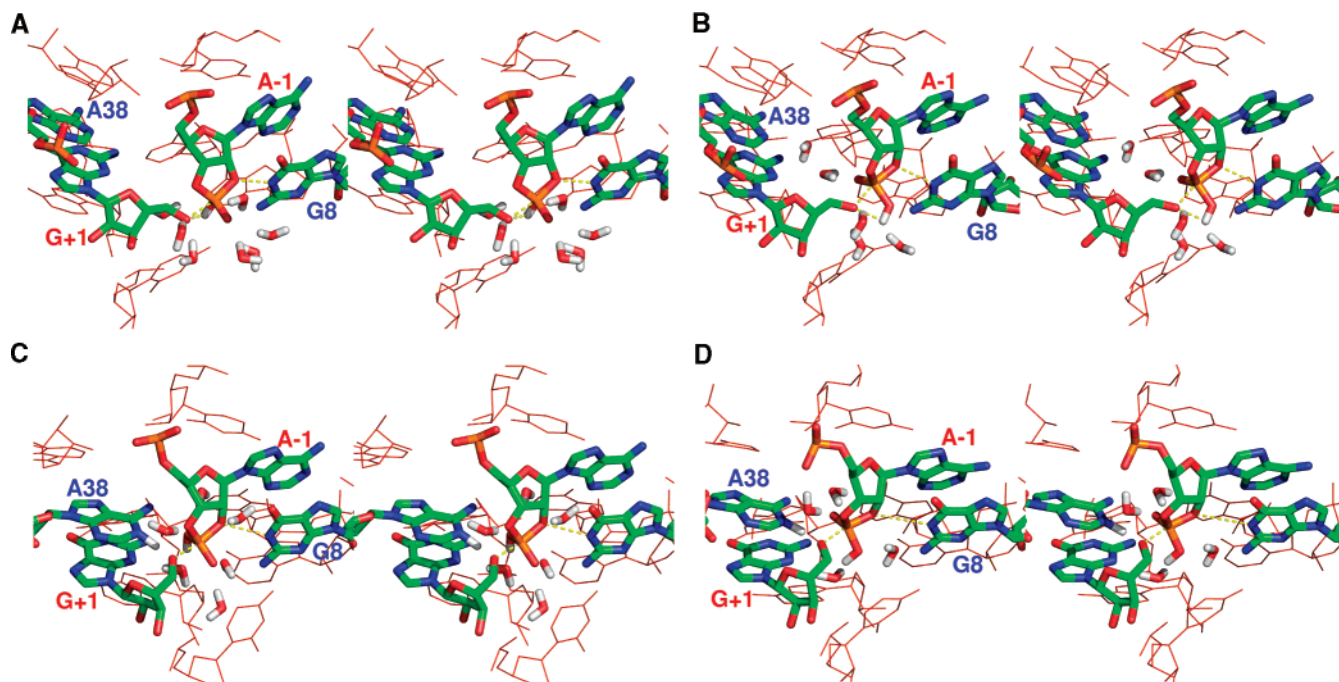

Figure 3. Stereoview of active site snapshot at the transition state of the exocyclic cleavage step from the in-line monoanionic pathway via (A) O_{1P} atom and (B) O_{2P} atom, and the A38(+) mediated monoanionic pathways via (C) O_{1P} atom and (D) O_{2P} atom. Water molecules that interact with nonbridging oxygens and the leaving group $O_{5'}$ are also displayed.

and 1-D PMF profiles on the in-line monoanionic mechanism. The 2-D PMF is a function of two degrees of freedom, and the 1-D PMF is a function along the minimum free energy path in the 2-D surface. Separate free energy simulations were performed to sample all other degrees of freedom except that corresponding to this minimum path, resulting in a variationally optimized transition state along the reaction coordinate. The effects from the second degree of freedom in the 2-D PMF cannot be assumed negligible, especially when there is a shallow

and broad basin in the 2-D PMF such as that in the reactant and product states, which makes significant entropic contribution. In the in-line monoanionic O_{1P} pathway, the effects are not large in the reactant state but are considerable in the product state (Figure 1A), whereas in the O_{2P} pathway the effects are qualitatively reversed (Table 1). The differences between the free energy values derived from the 1-D PMF quantitatively refine, but do not alter, the general conclusions about the hairpin mechanism based on the 2-D PMF profiles.
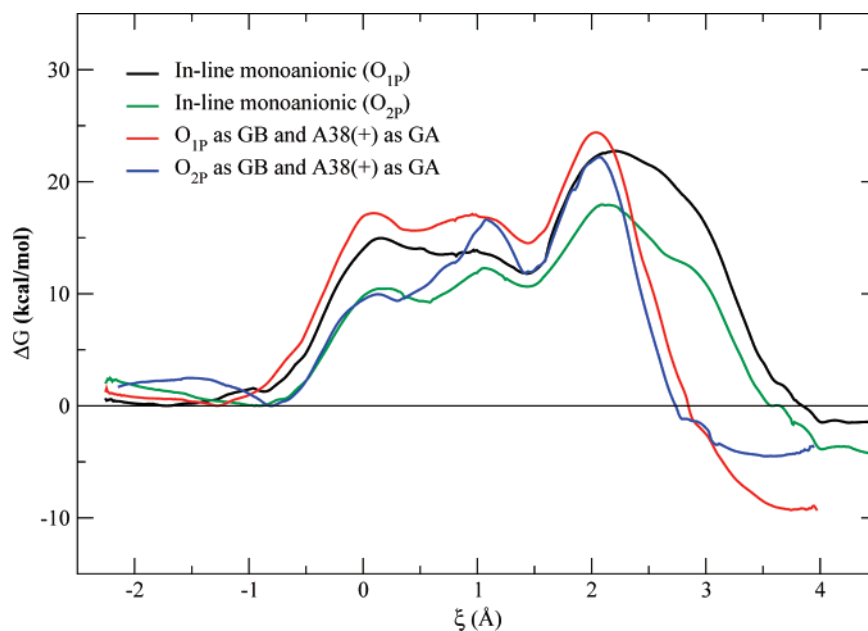
Figure 4. Comparisons of 1-dimensional potentials of mean forces for the in-line monoanionic (black and green colored solid line) and A38(+) mediated reaction mechanisms (red and blue colored solid line). The reaction coordinate is defined in eqs 2 and 3 by following minimum PMF pathways for each separate 2-D PMF profiles. GB and GA are the general base and acid, respectively.

Density-Functional Calculations. For the in-line monoanionic mechanism, the DFT results indicate that the AM1/d-PhoT model has similar differences for the O_{1P} and O_{2P} pathways (2.3 and 2.5 kcal/mol of error, respectively). The DFT-corrected barrier for the exocyclic cleavage step is 25.0 kcal/mol in the O_{1P} pathway and 20.5 kcal/mol in the O_{2P} pathway, respectively (Table 1). The hairpin ribozyme stabilizes the intermediate state by more than 15 kcal/mol compared to the transphosphorylation reaction of methyl ethylene phosphate (MEP) in water,[56] a simple model for the uncatalyzed reaction. There are several hydrogen bonds that play important roles in the stabilization of the intermediate and transition states, in particular, functional groups on G8 and A9 bases, as well as solvent molecules (Table 3 and also see the Supporting Information).

3.3. Role of A38 Nucleobase. Ferré-D'Amaré and co-workers proposed on the basis of crystallographic analyses that the protonated A38 nucleobase may act as a general acid to donate a proton to the $O_{5'}$ leaving group.[4,25] The involvement of A38 nucleotide base on catalysis is further supported by Fedor and co-workers, who found that the reaction rate decreases more than 10^4 -fold after A38 base is removed.[9] To address the proposal that A38 acts as a general acid in the exocyclic cleavage, two different simulations are carried out with A38 protonated at the N_1 position, henceforth designated A38(+).

1-D PMF Profiles. The first set of simulations follows analogously the reaction coordinate for the in-line monoanionic mechanism, but with a protonated A38 base (Scheme 3). The 1-D PMF profiles are compared in Figure 4. The initial step is the proton transfer from $O_{2'}$ to the nonbridging oxygen, followed by nucleophilic attack of $O_{2'}$ at the P atom, forming the monoprotic phosphorane intermediate as described in Section 3.2. The A38(+) base forms hydrogen bonds with the non-bridging O_{1P} oxygen and the $O_{5'}$ leaving group. In the second step, the proton on N_1 of A38(+) is transferred to the leaving group along with the exocyclic cleavage of the $P-O_{5'}$ bond, whereas the proton from the nucleophile remains on the nonbridging oxygen.

For the nucleophilic attack step in the A38(+) mediated mechanism, the proton transfer from the nucleophile favors the O_{2P} by 7.2 kcal/mol over the O_{1P} position, and the rate-controlling cleavage step by 2.2 kcal/mol from the AM1/d-PhoT model (Table 1). The presence of proton on O_{1P} hinders the hydrogen bond with the exocyclic amine group on A38, whereas this hydrogen bond is strengthened in the O_{2P} pathway. In the O_{2P} pathway, the accumulation of negative charge on O_{1P} after the protonation of O_{2P} further strengthens the interaction with protonated A38 base (Table 4).

The formation of a strong hydrogen bond with the A38(+) weakens the interaction of the phosphate group with G8 (the average interaction energies of selected nucleotide bases are provided in the Supporting Information). Comparison of inter-mediate geometries shows that the A38(+) base pulls the phosphate group closer and its distance to G8 concomitantly lengthens (Figure 2C and D). This distorts the orientation of phosphate toward the A38(+) base and hinders nucleophilic attack of $O_{2'}$ on the P atom. In particular, the phosphorane geometry from the A38(+) mediated O_{2P} pathway is partially pseudo-rotated to maintain the interaction with G8 nucleobase, having an $O_{2'}-P-O_{5'}$ angle of 134 degrees (Table 2 and Figure 2D). The protonated A38 base is a better proton donor than the O_{2P} nonbridging oxygen of the phosphorane intermediate, and the A38(+) mediated mechanism has an earlier transition state than that of the in-line monoanionic mechanism. The computed barrier heights from the two A38(+) mediated pathways are 24.4 and 22.2 kcal/mol for O_{1P} and O_{2P} atoms acting as the general base from the AM1/d-PhoT model, respectively.

Density-Functional Calculations. The AM1/d-PhoT trun-cated active site model has an activation energy for the rate-controlling exocyclic cleavage step that is 8 kcal/mol higher than the corresponding density-function result for both the O_{1P} and O_{2P} pathways. If this difference is applied as a DFT correction, the barrier height for the TS_{Cl} step becomes 16.4 and 14.4 kcal/mol for the O_{1P} and O_{2P} pathways, respectively (Table 1). However, consideration must be taken into account

for the free energy required to protonate the A38 nucleobase to quantify the true activation barrier. A rough estimate of the free energy required to protonate A38 at neutral pH can be made based on the pK_a of 4.3 of a free adenine base,[33] leading to an increase in barrier height by 3.7 kcal/mol at 300 K. The shifted and DFT corrected barriers become 20.1 and 18.1 kcal/mol for the O_{1P} and O_{2P} pathways, respectively (Table 1). In each pathway, the proton from A38(+) is transferred without barrier immediately following the transition state of the exocyclic cleavage step.

3.4. Role of G8 Nucleobase. Until now, the focus is on the nucleophilic step along with the in-line proton transfer from $O_{2'}$ onto the nonbridging oxygen. Several experiments implicate the involvement of G8 in the catalytic step,[20,21,28−30,32] although it has not yet been possible to establish unambiguously its precise role as a general base. The energetic contribution of the G8 base from the current simulations indicates that G8 contributes to the stabilization of the transition states and the phosphorane intermediate by providing stabilizing hydrogen bond interactions. However, it is possible that G8 acts as a general base to activate the $O_{2'}$ nucleophile. This section aims at providing insight into the role of the G8 nucleobase, under the assumption that the active form of G8 is neutral. This is reasonable because the negatively charged active site would not prefer to have an extra negative charge near the scissile phosphate. It is plausible that tautomerization between N_1 and O_6 on the guanine base[30] is important for activity. In this case, N_1 accepts a proton from $O_{2'}$ prior to the nucleophilic attack and becomes positively charged, which can further stabilize the negative charges on the scissile phosphate. The energetic difference between these two forms in the gas phase is small; the computed adiabatic energy difference between the two tautomeric forms is 1.1 kcal/mol from B3LYP/6-311++G(3df,-2p) and 0.6 kcal/mol from the AM1/d-PhoT model, where the tautomer with proton on N_1 has lower energy. Therefore, the proton affinities of two functional groups are similar in the gas phase.[88]

To test this possibility in the ribozyme, we have carried out free energy simulations to compute the free energy difference between the two tautomeric forms. The resulting free energy difference is 19.1 kcal/mol in the active site of the hairpin ribozyme. Although it is possible that the calculations are biased by the starting X-ray structure, the large free energy difference essentially rules out the possibility that the tautomeric form of G8 acts as the general base in the nucleophilic attack step. The present result supports the importance of a proton located on N_1 of G8 as well as the exocyclic amine group, as proposed by Pinard et al.[30] In addition, it is consistent with the recent X-ray crystallographic structures with several G8 variants suggesting that the base G8 primarily plays a structural role in the ribozyme reaction.[21]

3.5. Comparison with Experiment. A main mechanistic question is to identify the general acid and base and the associated protonation state of the reactive intermediates. The present work provides important insight into these questions. Although we cannot rule out the possibility of a water acting as a specific base, the nonbridging oxygens are a plausible candidate as a general base to activate the $O_{2'}$ nucleophile. The

G8 base, on the other hand, while providing important hydrogen bond stabilization in the reaction, is not a viable candidate as a general base for the mechanisms that were explored in the present study. The present results show that O_{2P} is the preferred general base that accepts a proton from $O_{2'}$ with a computed barrier of 10.5 kcal/mol for the proton-transfer process in the in-line monoanionic mechanism (Figure 1B). Kinetic experiments revealed that RNA cleavage reactions are most preferred through a phosphorane monoanion charge state,[36,42,85] but direct measurement of pK_a values for phosphoranes is difficult due to their transient nature. Proposed values from experiment and theory are in the range of 6.5−11.0 and of 11.3−15.0 for the first and second ionizations, respectively.[36,41,42] As the reaction proceeds to the phosphorane intermediate, the pK_a is shifted to a higher value, and the nonbridging oxygens increase their proton affinity. If the reaction is characterized by a dianionic mechanism, in which a base other than the nonbridging oxygens extracts the proton from $O_{2'}$ prior to the nucleophilic attack,[36] the lifetime of the dianionic phosphorane (likely a true transition state) may be insufficient to undergo protonation whereby the reaction would proceed directly to the product. Alternative mechanisms are also possible. One aspect that has not yet been thoroughly addressed is the possible chemical role played by solvent networks in the active site, for example in mediating proton transfers.[89] Trapped water molecules have been observed to interact with nonbridging oxygens in the active site of hairpin ribozyme,[87] but these waters are readily exchangeable with bulk solvent and exhibit considerable configurational variation. A detailed study of the role of solvent networks in the chemical steps of acid/base catalysis is nontrivial and will be a focus of future work.

The in-line monoanionic mechanism explains the mechanism for protonation of the phosphorane intermediate, and identifies the nonbridging phosphoryl oxygen as a plausible candidate to act as the general base. This mechanism implies a relatively minor chemical role of the A38 nucleobase in the catalysis and is different from the mechanisms proposed from structural[4,25] and kinetic analysis.[9] One finding from molecular dynamics free energy simulations in the O_{1P} pathway is that the neutral A38 nucleobase prefers to base stack with A40 and retracts slightly from the scissile phosphate and the leaving group. The retraction of A38 is less pronounced for the O_{2P} pathway where a weak hydrogen bond with the O_{1P} atom is retained (Table 3).

The reduction of reaction rates by the removal of G8 and A38 base[9,28,29] suggests the involvement of those bases in the catalysis. On the other hand, the current simulation study supports the mechanism of G8 playing structural role by providing hydrogen bonds to the scissile phosphate, and A38 being a general acid in the exocyclic cleavage step. Although this does not imply a disagreement, to verify the agreement of the simulations with the experiments, two free energy simulations were carried out, in particular, along the in-line monoanionic O_{2P} pathway, each of which has either removed G8 or A38 base, respectively. The results are presented in Table 1. After the DFT-correction, the A38 base deletion raises the barrier by 6.7 kcal/mol (or 3.0 kcal/mol with pK_a correction for A38) with respect to the A38(+) mediated O_{2P} pathway. The increase in barrier height is consistent with the reduction

(88) Gorb, L.; Podolyan, Y.; Leszczynski, J.; Siebrand, W.; Fernández-Ramos, A.; Smedarchina, Z. *Biopolymers* **2002**, *61*, 77−83.

(89) Gerratana, B.; Sowa, G. A.; Cleland, W. W. *J. Am. Chem. Soc.* **2000**, *122*(51), 12615−12621.

of reaction rate more than 10^4-fold. Similarly, the deletion of G8 base raises the reaction barrier by 3.6 kcal/mol in the in-line monoanionic O_{2P} pathway. It is likely that the strong interaction of G8 base with the scissile phosphate is not fully compensated by water molecules occupying the space of removed G8 base (see the Supporting Information for the snapshots at the INT_2 state).

The present simulations reveal a possible role of the non-bridging oxygen as a general acid catalyst in the absence of the A38 base.[9] On the other hand, when A38 is protonated, the current results support the proposal that it is the general acid in catalysis, in addition to its role in hydrogen bonding and base stacking interactions.[4,9,24,25] The computed activation barriers for the A38(+) mediated mechanism do not include the free energy contribution required to protonate A38, whereas an estimate of its contribution is provided in Table 1 based on the pK_a of a free adenine base. Determination of the pK_a shift of A38 in the ribozyme environment via quantum chemical calculation or NMR experiments can be valuable to aid in the resolution of the kinetic ambiguity inherent in the analysis of pH-rate profiles. Additional calculations to reproduce kinetic and thermodynamic properties of mutations that designed to probe the active site architecture of the hairpin ribozyme[24] can further contribute to formulating a consistent mechanism based on theory and experiment.

4. Conclusion

This work presents a theoretical study of the hairpin ribozyme self-cleavage mechanism using molecular simulations and a combined QM/MM potential, making use of the recently developed AM1/d-PhoT Hamiltonian. The latter has been developed against high level theoretical and experimental results. The present study further demonstrates that the AM1/d-PhoT potential is a reliable Hamiltonian in comparison with density functional results for the hairpin ribozyme reactions. The computed barrier height relative to that in solution is consistent with experimental results. The enhancement of the reaction rate by the ribozyme is due to the stabilization of the monoprotic phosphorane intermediate and the transition state by the ribozyme. The simulation results offer a plausible alternative to the notion that catalysis is by a general acid/base mechanism with the direct involvement of active site nucleotide bases. The experimentally identified nucleobases, including G8, A9, and A38, are intimately involved in the stabilization of phosphorane intermediates. The present results provide evidence that the role of the G8 nucleobase in the catalysis is mainly structural, which is unlikely to play the role of a general base to activate $O_{2'}$ nucleophile. The role of the A38 nucleobase as general acid has not been fully resolved, although the nonbridging phosphoryl oxygen is shown to be a feasible candidate to play the role of general base in abstracting a proton prior to or concerted with nucleophilic attack. The nucleotide bases, phosphate, and water molecules in the active site provide an extensive hydrogen bond network, stabilizing the transition state and increasing the reaction rate by more than a million-fold relative to the uncatalyzed reaction in solution.

Supporting Information Available: Details of umbrella sampling MD simulations, average interaction energy of key nucleotide bases, DFT-optimized geometries and energies of active site models, snapshots of intermediate states from removed G8 and A38 base simulations, and the complete citation for ref 81. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0759141