Role of the Active Site Guanine in the *glmS* Ribozyme Self-Cleavage Mechanism: Quantum Mechanical/Molecular Mechanical Free Energy Simulations

Sixue Zhang,[†] Abir Ganguly,[†] Puja Goyal,[†] Jamie L. Bingaman,[‡] Philip C. Bevilacqua,*[,‡] and Sharon Hammes-Schiffer*[,†]

[†]Department of Chemistry, University of Illinois at Urbana—Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, United States

[‡]Department of Chemistry and Center for RNA Molecular Biology, Pennsylvania State University, 104 Chemistry Building, University Park, Pennsylvania 16802, United States

ⓢ *Supporting Information*

ABSTRACT: The *glmS* ribozyme catalyzes a self-cleavage reaction at the phosphodiester bond between residues A-1 and G1. This reaction is thought to occur by an acid—base mechanism involving the glucosamine-6-phosphate cofactor and G40 residue. Herein quantum mechanical/molecular mechanical free energy simulations and pK_a calculations, as well as experimental measurements of the rate constant for self-cleavage, are utilized to elucidate the mechanism, particularly the role of G40. Our calculations suggest that an external base deprotonates either G40(N1) or possibly A-1(O2′), which would be followed by proton transfer from G40(N1) to A-1(O2′). After this initial deprotonation, A-1(O2′) starts attacking the phosphate as a hydroxyl group, which is hydrogen-bonded to deprotonated G40, concurrent with G40(N1) moving closer to the hydroxyl group and directing the in-line attack. Proton transfer from A-1(O2′) to G40 is concomitant with attack of the scissile phosphate, followed by the remainder of the cleavage reaction. A mechanism in which an external base does not participate, but rather the proton transfers from A-1(O2′) to a nonbridging oxygen during nucleophilic attack, was also considered but deemed to be less likely due to its higher effective free energy barrier. The calculated rate constant for the favored mechanism is in agreement with the experimental rate constant measured at biological Mg^{2+} ion concentration. According to these calculations, catalysis is optimal when G40 has an elevated pK_a rather than a pK_a shifted toward neutrality, although a balance among the pK_a's of A-1, G40, and the nonbridging oxygen is essential. These results have general implications, as the hammerhead, hairpin, and twister ribozymes have guanines at a similar position as G40.
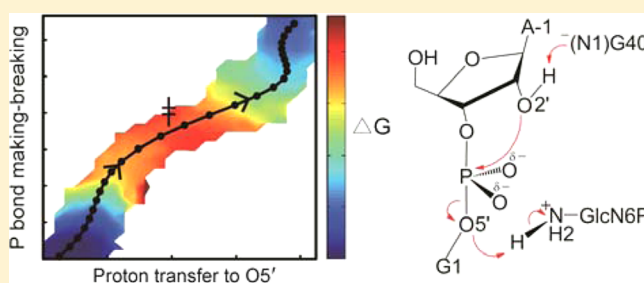
■ INTRODUCTION

Ribozymes catalyze essential reactions for RNA processing and protein synthesis. One such reaction is sequence-specific cleavage and ligation of a phosphodiester bond, which is catalyzed by the hammerhead,[1] hairpin,[2] hepatitis delta virus (HDV),[3] Varkud satellite,[4] twister,[5] and *glmS*[6] ribozymes. First reported in 2004, there are now over 450 identified *glmS* ribozyme representatives.[7] The ribozyme resides in the 5′ untranslated region of an mRNA upstream from the coding region for L-glutamine/D-fructose-6-phosphate aminotransfer-ase, which catalyzes the synthesis of glucosamine-6-phosphate (GlcN6P).[8] In addition to being the product of L-glutamine/D-fructose-6-phosphate aminotransferase, GlcN6P serves as a cofactor for the *glmS* ribozyme.[6] Thus, the ribozyme is involved in a negative feedback mechanism, where expression of the *glmS* gene leads to higher GlcN6P production, which in turn results in *glmS* ribozyme activation and mRNA cleavage and subsequent down-regulation of the *glmS* gene. GlcN6P eventually forms uridine 5′-diphospho-*N*-acetyl-D-glucosamine,

which participates in the biosynthesis of amino sugar-containing macromolecules.[9]

The tertiary structure of the *glmS* ribozyme features a doubly pseudoknotted core, as depicted in Figure 1A.[10] The ribozyme undergoes self-cleavage at the phosphodiester bond between A-1 and G1.[8] The active site consists of A-1, G1, another guanine residue positioned near the nucleophilic 2′-OH (G40 in *Thermoanaerobacter tengcongensis*[11] and G33 in *Bacillus anthracis*[12]), and the GlcN6P cofactor positioned near the leaving group 5′oxygen (Figure 1B). We use the "G40" notation in this article. The ribozyme's self-cleavage reaction is thought to employ a general acid—base mechanism, in which a general base whose identity is unclear deprotonates A-1(O2′); A-1(O2′) makes an in-line attack on the scissile phosphate at the backbone between A-1 and G1; and the cofactor serves as a general acid to protonate the G1(O5′). At the end of the
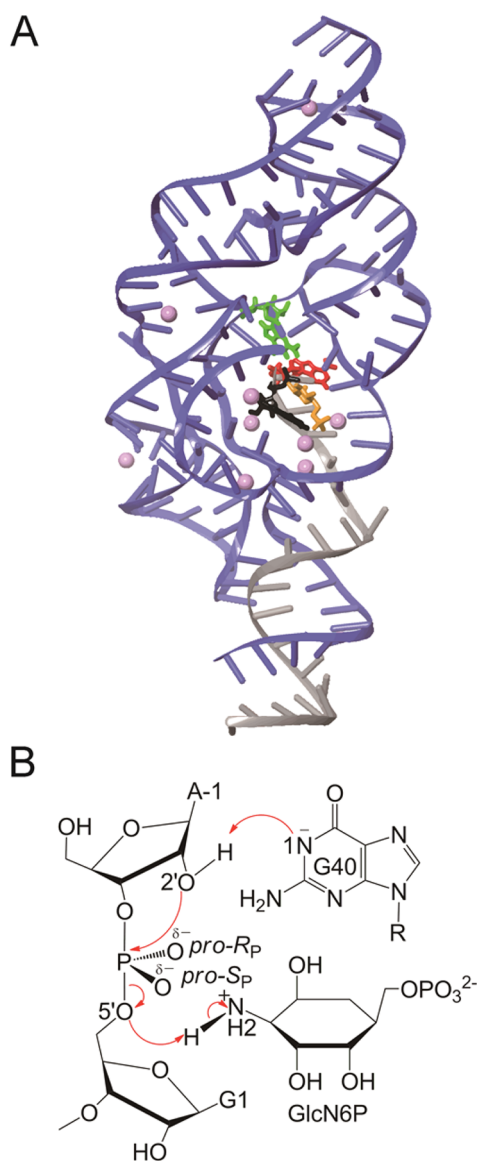
A

B

Figure 1. (A) Structure of the *glmS* ribozyme from *Thermoanaerobacter tengcongensis* (PDB ID 2Z75). Chain A, which is an oligomer substrate, is shown in gray, and Chain B, which is a motif from the *glmS* ribozyme RNA, is shown in blue. The active site consists of A-1 (shown in red), G1 (shown in black), G40 (shown in green), and the cofactor (shown in orange). The Mg^{2+} ions in the crystal structure are represented by pink spheres. (B) Schematic picture of the active site. For simplification, some parts of the residues are not shown. In this mechanism, A-1(O2′) is deprotonated by G40 during the nucleophilic attack of the phosphate.

reaction, the P−O5′ bond is broken and the P−O2′ bond is formed to generate 2′,3′-cyclic phosphodiester and 5′-hydroxyl termini products.[12−15]

Experimental studies involving truncation and gene deletion[6,8,11,16] have shown that a small region around the active site, which has high sequence conservation, is crucial to the catalytic activity. Mutation of A-1G1 to CC or mutation of G1 to A inhibits catalytic activity, while mutation of A-1 to G leads to a more modest rate reduction.[6] Studies on the *glmS* ribozyme bound to GlcN6P and its analogues illustrate the catalytic importance of the amine group of GlcN6P.[10,15] According to Raman crystallography, the pK_a of the amine group of bound GlcN6P is shifted from a solution pK_a of 8.06

± 0.05 to 7.26 ± 0.09,[17] close to neutrality. This shift, as well as structural studies that show the amine group of GlcN6P positioned near G1(O5′),[17,18] support the cofactor's potential role as a general acid. Additionally, classical molecular dynamics (MD) simulations have predicted that the cofactor will be protonated when bound to the active site.[19,20]

The possible roles of G40[21,22] and metal ions[16,23] have also been explored. Studies involving G40 mutations have shown that substitution of a different base for guanine at this position results in a significant reduction of activity, where the greatest effect is seen with the G40A mutation with at least a 10,000-fold decrease in rate.[12] G40 has been proposed to be the general base in the cleavage mechanism, but evidence challenging this proposition exists. Studies with the fluorescence analogue with an 8-azaguanine at this position have suggested that G40 has a basic-shifted pK_a compared to the pK_a of 9.2 for free guanine in solution.[21,24] In addition, metal ion studies on the *glmS* ribozyme have suggested that divalent metal ions may have a nonspecific role in the cleavage mechanism, as numerous divalent metal ions and monovalent ions at high concentrations support cleavage activity.[6,25] Despite these previous studies on the *glmS* ribozyme, several crucial aspects of the mechanism remain unknown: (1) the identity of the base that deprotonates A-1(O2′); (2) the role of G40, which appears to be essential based on mutation studies;[22] and (3) the fundamental mechanism of self-cleavage (i.e., concerted or sequential chemical steps).

A variety of theoretical methods have been used to study the mechanisms of small ribozymes.[26] Molecular dynamics simulations using classical force fields provide useful information about hydrogen-bonding interactions and structural motifs; however, they cannot describe the making and breaking of chemical bonds in the cleavage reaction.[27] Traditional quantum mechanical/molecular mechanical (QM/MM) geometry optimizations are helpful in probing intermediates along possible reaction pathways, but they do not include conformational sampling and entropic contributions, which are essential for obtaining free energy barriers for comparison to experimentally measured rate constants.[28,29] In contrast, QM/MM free energy simulations that combine umbrella sampling[30] and a finite temperature string method[28] can be used to generate the multidimensional free energy surface and to identify possible reaction paths, denoted minimum free energy paths (MFEPs), for complex biological reactions.[31−37] This approach is computationally tractable because only the relevant portions of the free energy surface are sampled. We recently used this approach to study the self-cleavage reaction catalyzed by the HDV ribozyme.[38]

Herein, we apply this QM/MM free energy approach to the *glmS* ribozyme in an effort to address the mechanistic issues discussed above. Prior to the QM/MM free energy simulations, we performed classical MD simulations to elucidate key hydrogen-bonding interactions for the various protonation states in the proposed mechanisms. We also performed QM/MM geometry optimizations to investigate the structures of the reactant, product, and intermediate states associated with the various reaction pathways and to obtain initial structures for the QM/MM free energy simulations. The subsequent free energy simulations were used to generate the MFEPs for the proposed mechanisms and to calculate the relative free energies of key states in the reaction pathways. To gain further insight into the relative ease of deprotonating either A-1(O2′) or G40(N1), we also conducted pK_a calculations using the Poisson−Boltzmann

with Linear Response Approximation (PB/LRA) approach. To provide a degree of validation for the QM/MM free energy simulations, we experimentally measured the rate constant for the self-cleavage reaction catalyzed by the *glmS* ribozyme under various conditions, including those related to the *in silico* conditions. The theoretical studies, supported by the experimental measurements, provide new insights into the mechanism of this self-cleavage reaction.

■ MATERIALS AND METHODS

1. Classical MD Simulations. The system is based on a precleaved crystal structure (PDB ID 2Z75)[11] of wild-type *glmS* ribozyme from *Thermoanaerobacter tengcongensis*. The ribozyme consists of a substrate oligonucleotide (chain A) and a longer RNA enzyme strand (chain B). In this crystal structure, A-1 and G1 have the rare C2′-endo sugar pucker, the cofactor is the α-anomer, and A-1 has a 2′-deoxyribose to prevent the cleavage reaction.[39] For our simulations, the hydroxyl group on the C2′ of A-1 was incorporated by superimposing it with another crystal structure of the *glmS* ribozyme (PDB ID 2HO7), which utilized substitution of the amine group in the cofactor with a hydroxyl group to prevent the cleavage reaction.[10] Three exterior uridine bases, U6, U49, and U103, which were not resolved in the crystal structure, were added to the system using the Maestro program,[40] followed by energy minimization of only these bases. These bases are distal from the active site and unlikely to affect catalysis. The nine Mg^{2+} ions that were resolved in the crystal structure[11] were included in the system, although none of these is near the site of cleavage. The hydrogens were added with the Accelrys Discover Studio Visualizer 2.0 program. The system was solvated in an orthorhombic box containing TIP3P water[41] and 0.15 M NaCl, which is near the physiological concentration of monovalent ions.[42] Additional Na$^+$ ions were added to neutralize the phosphate backbone.

The classical MD simulations utilized the AMBER99 force field,[43] periodic boundary conditions, and the Ewald treatment for long-range electrostatics.[44] The charges for the cofactor, the N1-deprotonated G40, and the O2′-deprotonated A-1 were calculated using the RESP procedure[45] (for details see Supporting Information, p S3, and Tables S1−S3 for parameters). The MD simulations were performed using the DESMOND program[46] with the same protocol previously used for the HDV ribozyme.[47,48] The temperature of all classical MD and QM/MM free energy simulations was 300 K. The structures obtained after more than 25 ns of classical MD were found to be similar to the crystal structure except for changes expected for different protonation states and in flexible regions (Table S4).

2. QM/MM Free Energy Simulations. The QM/MM free energy simulations combine umbrella sampling and a finite temperature string method. The QM region is depicted in Figure 2 and was treated with density functional theory (DFT) with the B3LYP functional and the 6-31G** basis set. This region included the cofactor and the critical residues G40, A-1, and G1, as well as the scissile phosphate. Test calculations with the 6-31+G** basis set, which includes diffuse basis functions, were in qualitative agreement with the results obtained with the 6-31G** basis set (Table S5). The MM region was described by the AMBER99 force field,[43] and the standard single-link atom approach, in which the atoms at the boundary are capped by hydrogen atoms, was used to describe the QM/MM interface.[49] An interface between Q-Chem[50] and CHARMM[51] was used to perform these QM/MM free energy simulations, following the same protocol used previously to study the HDV ribozyme.[38] Here we briefly summarize this general approach. Additional technical details are provided elsewhere.[38,52]

In this approach, *M* reaction coordinates are used to describe the reaction of interest. An initial string connecting the reactant and product structures is constructed in the *M*-dimensional reaction coordinate space. This initial string is divided into *N* images, where each image corresponds to specified values of the *M* reaction coordinates. Umbrella sampling is performed for each image to sample that region of phase space. In particular, harmonic potentials with a force constant of 100 kcal mol^{-1} Å^{-2} centered at the specified
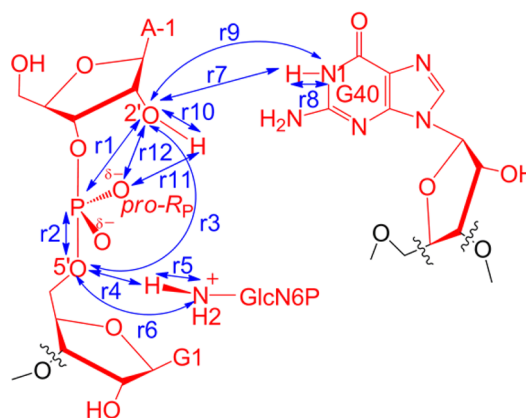

Figure 2. Definition of the QM region and the reaction coordinates. The QM region, which includes the entire GlcN6P cofactor, is shown in red, and the wavy lines indicate the boundaries between the QM and MM regions. The reaction coordinates used in the QM/MM free energy simulations are shown in blue.

values of the *M* reaction coordinates are applied during an MD trajectory for each image. After 100 fs of MD, the *N* images are redistributed along the string, and the centers of the harmonic restraints are updated on the basis of the average reaction coordinates of each image. This procedure is repeated until the string is determined to be converged on the basis of several criteria described in the Supporting Information (p S11 and Figures S3−S5). The final converged string is the MFEP. The weighted histogram analysis method (WHAM)[53] is used to unbias the data from all iterations to obtain the multidimensional free energy surface in the region of the MFEP.

We carried out QM/MM free energy simulations to study two parts of the reaction: phosphate bond cleavage, which also includes proton transfer from the cofactor to G1(O5′), and proton transfer between G40(N1) and the nucleophilic A-1(O2′). For phosphate bond cleavage, we consider mechanisms that are initiated by an external base as well as mechanisms that do not require an external base. The proposed mechanisms for these two cases, denoted "cleavage with external base" and "cleavage without external base", are depicted in Figure 3. The overall reaction evolves from the initial state, termed "State A" to the post-cleavage state, termed "State P". When the reaction is initiated by an external base, the external base deprotonates either A-1(O2′), producing State B, or G40(N1), producing State C1. The identity of the external base was not considered here; it could be a metal-bound hydroxide ion, a buffer molecule, or another atom on the ribozyme itself, and will require additional study. As will be shown below, A-1(O2′) was assumed to be deprotonated at the start of the *initial* string, but G40(N1) ended up deprotonated at the start of the *converged* string, and the proton subsequently transfers from A-1(O2′) to G40(N1) in the early stages of the MFEP. This change in the mechanism of the converged MFEP compared to the initial string illustrates that the methodology has the flexibility to alter the mechanism. When the reaction does not require an external base, both A-1(O2′) and G40(N1) are protonated (State A) at the beginning of the reaction pathway studied, and the proton is found to transfer to the *pro-R*$_P$ oxygen of the scissile phosphate in the early stages of the MFEP.

These QM/MM free energy simulations of the phosphate bond cleavage included 9 and 12 reaction coordinates for "cleavage with external base" and "cleavage without external base" simulations, respectively. These reaction coordinates are defined in Figure 2, where reaction coordinates r1−r9 are the same for the two types of mechanisms. The initial string for the "cleavage with external base" simulations was comprised of 11 images generated by a linear interpolation connecting the pre-cleavage and post-cleavage structures obtained from the QM/MM geometry optimizations described in the Supporting Information. This type of an initial string mimics a

A: Cleavage with external base
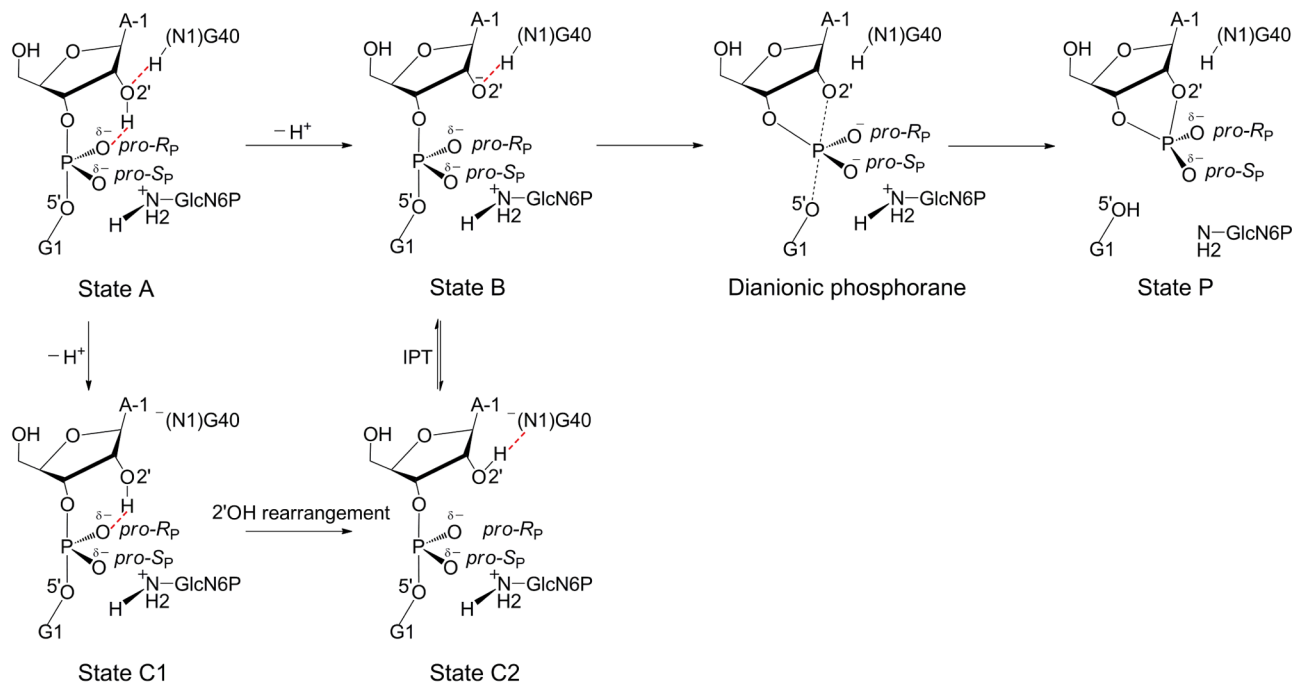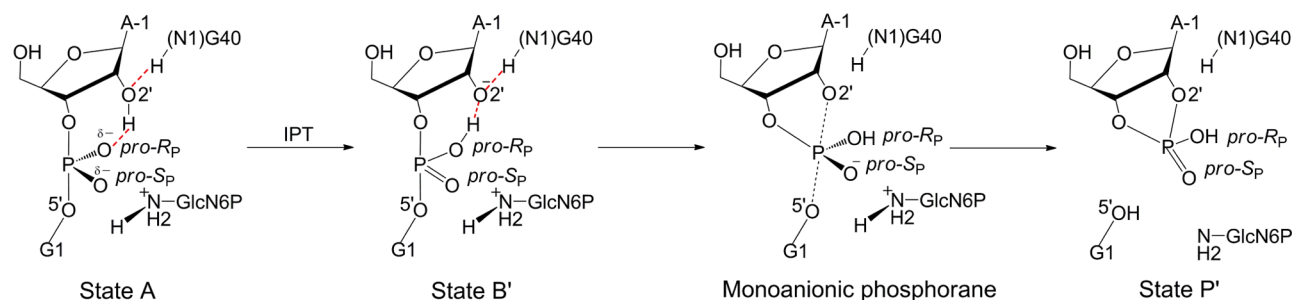

B: Cleavage without external base

Figure 3. Proposed mechanisms for the self-cleavage reaction (A) when an external base activates the reaction, denoted "cleavage with external base" and (B) when an external base does not activate the reaction, denoted "cleavage without external base". Red dashed lines indicate hydrogen bonds related to the initial proton transfer. The overall reaction evolves from State A to the post-cleavage State P. When the reaction is activated by an external base, the external base deprotonates either A-1(O2′), directly producing State B, or G40(N1), producing State C1, which can undergo rearrangement of the O2′ hydrogen-bonding interaction to produce State C2, followed by intramolecular proton transfer (IPT) from A-1(O2′) to G40(N1) to produce State B. The reaction proceeds from State B to a dianionic phosphorane transition state or intermediate, depending on whether the reaction is concerted or sequential, ultimately leading to the post-cleavage product State P. When the reaction is not activated by an external base, IPT from A-1(O2′) to the pro-R_p oxygen produces State B′, which proceeds to a monoanionic phosphorane transition state or intermediate, ultimately leading to the post-cleavage product State P′. State P′ is expected to rapidly evolve to State P because the pK_a on a cyclic phosphate is very low.

concerted mechanism. After seven iterations, the number of images was increased to 28 to obtain a higher degree of resolution for the MFEP, and a total of 21 iterations were performed, giving a total simulation time of 46.9 ps. The initial string for the "cleavage without external base" was also generated from a linear interpolation connecting the pre-cleavage and post-cleavage structures and so too represents a concerted mechanism. This initial string was comprised of 28 images, and 27 iterations were performed, resulting in a total simulation time of 75.6 ps. Note that these total simulation times include all images for all iterations.

To examine whether the MFEP depends on the choice of initial string, we also performed "cleavage with external base" simulations with an initial string associated with the sequential rather than the concerted mechanism. The details of these calculations are provided in the Supporting Information (p S11 and Figure S6), and the outcome illustrates that the results are not sensitive to the choice of the initial

string. In addition, we performed a statistical error analysis to estimate the error in the calculated free energy barriers due to statistical fluctuations. As shown in the Supporting Information, Table S6, we found the error due to statistical fluctuations to be less than 1 kcal/mol. Note that this analysis reflects only the statistical error and does not account for systematic error due to limitations in the density functional, basis set, and classical force field.

We also performed QM/MM free energy simulations to investigate the proton transfer reaction from A-1(O2′) to G40(N1) (i.e., the reaction from State C2 to State B in Figure 3). These simulations of the initial proton transfer reaction are denoted "O2′/N1 PT" simulations. In this case, the initial string was generated from a linear interpolation connecting States C2 and B obtained from QM/MM geometry optimizations. The QM region consisted of 30 atoms and included the sugar of A-1 and the base of G40. Because the QM region did not include the scissile phosphate, the self-cleavage reaction was

not able to proceed in these simulations. This simulation included only three reaction coordinates, defined as *r7*, *r8*, and *r9* in Figure 2, and the string was represented by 18 images. Note that the r10 coordinate was not included in these simulations because this proton was not present. We performed 100 fs of MD for each image per iteration. The simulations were terminated after 13 iterations because the system began to evolve toward a nonphysical geometry. The total simulation time was 23.4 ps.

In addition, we used this approach to calculate the free energy difference between States C1 and C2 (Figure 3) to determine the relative populations of these two states. These simulations are denoted "2′OH rearrangement" simulations. Because no chemical bonds are broken or formed, the entire system was described by the AMBER99 classical force field.[43] The initial string was generated from a linear interpolation connecting the structures of States C1 and C2 obtained from QM/MM geometry optimizations. The reaction coordinates considered in these simulations were the angles A-1(O2′):A-1(O2′H):G1(O2P) and A-1(O2′):A-1(O2′H):G40(N1), as defined in Figure S11, and the string was represented by 30 images. We performed 50 ps of MD for each image per iteration, and the simulations were converged after 15 iterations, corresponding to a total simulation time of 22.5 ns.

3. pK_a Calculations. The PB/LRA approach enables the calculation of the shift in pK_a of a residue or base in a protein or nucleic acid environment with respect to its pK_a in solution. The contribution of electrostatics, which is the dominant factor, to the pK_a shift of residue AH can be calculated as

$$\Delta p K_a = \frac{(G_{A^-,env} - G_{AH,env}) - (G_{A^-,aq} - G_{AH,aq})}{2.303RT} \quad (1)$$

where R is the gas constant, "env" denotes a particular microenvironment, and "aq" denotes aqueous solution.[54−56] Knowledge of the experimentally measured pK_a in solution can then provide an estimate of the absolute pK_a in the environment of interest. Studies in the literature[57,58] have suggested that the value of the dielectric constant for the RNA environment, ε_{env}, should be based on the type of structures for which the PB calculations are carried out. If only the crystal structure is used, the value of ε_{env} may need to be quite high because the nuclear relaxation of the environment in response to the change in the protonation state of the residue is not included. Alternative approaches have been proposed to include these effects, such as methods that account for side-chain rotamer sampling (i.e., the Multi-Conformer Continuum Electrostatics method[59,60]) or methods that account for relaxation of the environment by utilizing snapshots from MD trajectories.[61]

In the work described in this paper, we performed the PB calculations for snapshots sampled from MD trajectories for both the protonated and deprotonated states of the residue in the environment of interest. The pK_a shifts were then averaged over both kinds of snapshots in an LRA framework to obtain the final pK_a shift. This method was shown to be effective in previous continuum electrostatics studies aimed at calculating pK_a shifts in proteins.[61] The choice of dielectric constant for the environment has been discussed extensively in the literature. In the limit of infinite sampling, it has been suggested that $\varepsilon_{env} = 2$ is appropriate for use with PB/LRA calculations to account for deficiencies of standard nonpolarizable force fields in describing electronic polarization.[62] In this study, the snapshots used for the PB/LRA calculations were sampled from relatively long 25 ns MD trajectories. To account for modest deficiencies in the sampling, we used $\varepsilon_{env} = 4$ for the results reported in the main paper. To test the sensitivity of the results to the choice of dielectric constant, we also performed the same calculations with $\varepsilon_{env} = 8$ and found that the results remain qualitatively similar (Table S7).

In addition, we found that the ions in the bulk must be treated explicitly rather than implicitly, as is typical within the PB approach for proteins. The explicit treatment of ions is necessary for a proper description of neutralization of the backbone charges of the ribozyme. Hence all explicit ions in the primary box used in the MD simulations were retained in the PB calculations, and the positions of these ions

were different for each snapshot sampled from the MD trajectories. The atomic radii and partial charges were defined using the customized AMBER99 parameter set. We used a solvent dielectric constant of 80.0 to describe the aqueous environment. For the PB calculations of the reference species in aqueous solution, the residue of interest was extracted from the biomolecular environment and placed in a dielectric continuum with $\varepsilon = 80$. In these calculations, the monovalent ions in the bulk were treated implicitly, with concentrations of 150 mM and radii of 2 Å for both negatively and positively charged ions. For the benchmarking studies on adenine guanine dinucleotide (ApG), with a reference of ApEt in which the guanine is substituted by an ethyl group,[63] the explicit ions were also retained for the reference calculations. The pK_a calculations for A-1(O2′) also used a reference of ApEt and explicit ions. The pK_a calculations on G40 used guanosine as the reference.

The three-dimensional grid for the PB calculations was chosen such that its boundaries extend ~15−20 Å beyond all explicit atoms and ions. A grid spacing of 0.4 Å and a solvent probe radius of 1.4 Å were used for all calculations. For each PB/LRA calculation, approximately 200 snapshots separated by at least 10 ps were obtained from the classical MD trajectories.

4. Reaction Rate Measurement. The crystal structure used for the calculations (PDB ID 2Z75) was solved for RNA crystals stabilized in 1.7 M LiCl, 30 mM MgCl$_2$, 100 mM Tris-HCl (pH 8.5), and 15 mM GlcN6P.[11] As mentioned above, in the free energy simulations, 150 mM NaCl was used, which mimics physiological ionic strength, and the temperature was 25 °C. For the experiments, conditions of 25 °C and 150 mM NaCl were also used. In addition, 100 mM HEPES (pH 8.5) and 10 mM GlcN6P were used in the experiments; the concentration of GlcN6P is more than 3 times the K_d for GlcN6P binding to *Bacillus anthracis glmS* ribozyme of 1.4 mM, and should therefore be saturating.[18] We performed experiments at 3 mM Mg^{2+}, which closely matches those of the calculations, as well as 30 mM Mg^{2+}.

For experiments at 3 mM Mg^{2+}, manual mixing was used. Time points were collected, quenched with 20 mM EDTA, and were also immediately placed on dry ice. For experiments at 30 mM Mg^{2+}, the KinTek RQF-3 Rapid Chemical Quench-Flow instrument was used. Here, time points were collected, quenched with 150 mM EDTA, and were also immediately placed on dry ice. For both Mg^{2+} conditions, reaction aliquots were combined with formamide loading buffer as well as 1 mM heparin, which minimized retention of RNA in the wells, and fractionated on denaturing polyacrylamide gels containing 8.3 M urea. Gels were dried and visualized using a PhosphorImager and analyzed using ImageQuant. Data were fit to the following single-exponential equation:

$$f_c = A + B\, e^{-k_{obs}t} \quad (2)$$

■ RESULTS AND DISCUSSION

1. Classical MD Simulations. We performed classical MD simulations for each of three different protonation states depicted in Figure 3: State A, where both A-1(O2′) and G40(N1) are protonated; State B, where A-1(O2′) is deprotonated; and States C1/C2 (collectively denoted State C), where G40(N1) is deprotonated and two hydrogen-bonding arrangements are possible. In all cases, the cofactor has a protonated amine group and a doubly deprotonated phosphate tail, as depicted in Figure 1B. For each state, we propagated at least two 25 ns MD trajectories starting with different initial conditions. The preferred hydrogen-bonding network in the active site for each state during the MD trajectories is depicted in Figure S7, and the average donor–acceptor distances for the three key hydrogen-bonding interactions are given in Table 1. The results for States A and C are qualitatively consistent with previously published MD simulation results on the *glmS* ribozyme.[19] In State A, the

Table 1. Average Hydrogen-Bonding Distances (Å) in Different Protonation States from Classical MD Simulations[a]

state	A-1(O2′)-G40(N1)	A-1(O2′)-G1(O2P)	GlcN6P(N2)-G1(O5′)
A	3.20 (0.28)	3.20 (0.21)	2.95 (0.20)
C1[b]	4.16 (0.36)	3.33 (0.24)	3.09 (0.53)
C2[c]	2.88 (0.11)	4.56 (0.28)	3.09 (0.53)
B	2.93 (0.19)	4.45 (0.22)	3.52 (0.49)
crystal[d]	N/A	N/A	3.06

[a]States A, B, C1, and C2 are defined in Figure 3. Standard deviations are given in parentheses. [b]Configurations where A-1(O2′) forms a hydrogen bond with the pro-R_P oxygen on the scissile phosphate. [c]Configurations where A-1(O2′) forms a hydrogen bond with G40(N1). [d]The crystal structure used herein (ref 11) does not contain A-1(O2′).

hydrogen atom on G40(N1) always forms a hydrogen bond with A-1(O2′), and the hydrogen atom on A-1(O2′) always forms a hydrogen bond with the pro-R_p oxygen. In State B, the hydrogen atom on G40(N1) always forms a hydrogen bond with the deprotonated A-1(O2′).

In State C, two different hydrogen-bonding interactions were observed, as depicted schematically in Figure 3. In State C1, the hydrogen atom on A-1(O2′) still forms a hydrogen bond with the pro-R_p oxygen, as in State A. In State C2, the hydrogen atom on A-1(O2′) forms a hydrogen bond with the deprotonated G40(N1) instead. As shown in Figure 4, the system fluctuates between States C1 and C2 during the 25 ns MD trajectory. Because there is no hydrogen-bonding interaction between the A-1 sugar and G40 base in State C1, the average A-1(O2′)-G40(N1) distance is noticeably larger
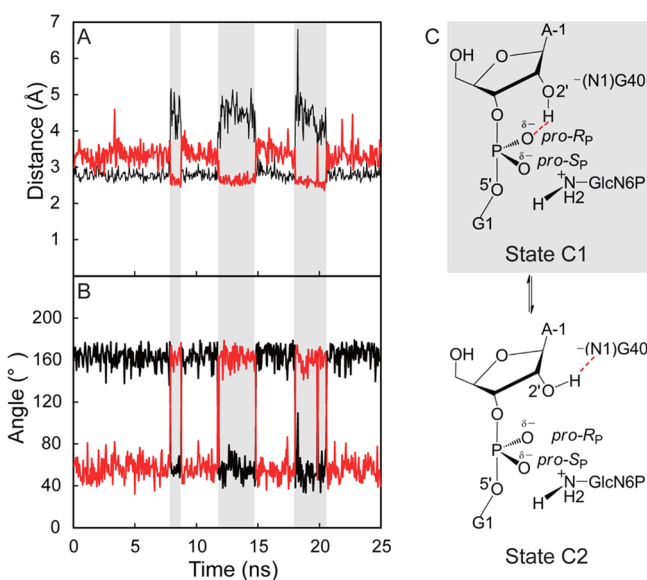

Figure 4. Key distances and angles along a 25 ns trajectory of States C1/C2. (A) Distance between A-1(O2′) and G40(N1) is shown in black, and distance between A-1(O2′) and pro-R_p is shown in red. Larger distances between A-1(O2′) and G40(N1) indicate State C1 (shaded area), and smaller distances between A-1(O2′) and G40(N1) indicate State C2. (B) Hydrogen-bonding angle A-1(O2′):H:G40(N1) is shown in black, and hydrogen-bonding angle A-1(O2′):H:G1(O2P) is shown in red. The changes in angles are correlated with the changes in distances. (C) Schematic pictures of State C1 and State C2.

than that in the other states. Specifically, this average distance is 4.16 Å for State C1 and is 2.88−3.20 Å for States A, B, and C2, as given in Table 1.

In all of these MD trajectories, the cofactor remained in the active site, and one of the three hydrogen atoms covalently bonded to GlcN6P(N2) was hydrogen-bonded to G1(O5′). In addition, one of the hydroxyl groups of GlcN6P was hydrogen-bonded to the pro-R_p oxygen of the scissile phosphate. Although nine Mg^{2+} ions were included in the simulations, these ions, which started at their crystallographic sites, remained localized at these sites throughout the simulations. These stable interactions include the two Mg^{2+} ions in contact with the phosphate tail of the cofactor, as previously observed by experiments[23] and simulations.[19] These particular Mg^{2+} ions are distal from the active site and are not thought to play a direct role in the self-cleavage reaction mechanism of the glmS ribozyme. In support of this notion, the glmS ribozyme functions in the absence of divalent ions when high concentrations of monovalent ions are present, albeit at a substantially slower rate, and in the presence of the exchange inert cobalt hexammine ion.[16]

2. QM/MM Free Energy Simulations. We performed QM/MM free energy simulations to generate the multidimensional free energy surfaces and MFEPs for the self-cleavage reaction catalyzed by the glmS ribozyme. As discussed above, we considered two different types of phosphate bond cleavage mechanisms. In the "cleavage with external base" simulations, either the A-1(O2′) or the G40(N1) was deprotonated prior to the start of the reaction pathway studied, whereas in the "cleavage without external base" simulations, both A-1(O2′) and G40(N1) were protonated at the beginning of the reaction pathway studied. In both types of simulations, we considered three possible reaction pathways for the self-cleavage reaction, as depicted in Figure 5. In this figure, all of the pathways are shown to start with State B in the "cleavage with external base" simulation. In this case, an external base is assumed to deprotonate either A-1(O2′), producing State B, or G40(N1), producing State C1, which can transform to State B via the pathway shown in Figure 3. An analogous figure for the "cleavage without external base" simulation starts with State B′, in which the pro-R_p oxygen is protonated, and is provided in Figure S8.

The three pathways depicted in Figure 5 correspond to one concerted and two sequential mechanisms. In the concerted pathway, denoted by the middle arrow, the P−O bonds are formed and broken simultaneously along with the proton transfer from the cofactor to G1(O5′), passing through a single phosphorane-like transition state. In the upper sequential mechanism, the nucleophilic attack of A-1(O2′) on the phosphate occurs first, generating a phosphorane intermediate, followed by protonation of G1(O5′). In the lower sequential mechanism, proton transfer from the cofactor occurs first, generating an intermediate with protonated G1(O5′), followed by the nucleophilic attack of A-1(O2′) on the phosphate. From the QM/MM optimizations described below, we found a stable phosphorane intermediate corresponding to the upper sequential pathway but were unable to find a stable intermediate corresponding to the lower sequential pathway. As discussed above, the initial string in both types of cleavage free energy simulations corresponded to the concerted mechanism. To test the sensitivity of the approach to the initial string, we also performed the "cleavage with external base" simulation with an initial string corresponding to the
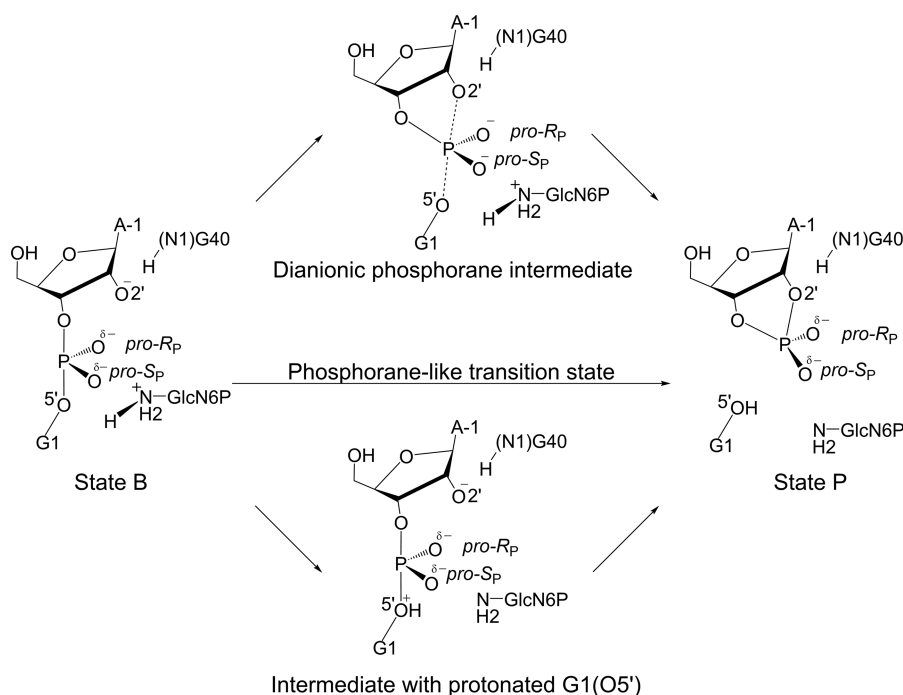
Figure 5. Three possible pathways for "cleavage with external base" subsequent to the deprotonation of O2′. The upper and lower pathways are sequential, and the middle pathway is concerted. For the mechanism that does not require an external base (shown in Supporting Information, Figure S8), the *pro-R*$_P$ oxygen will be protonated in all of these structures.

Table 2. Key Distances (Å) in Structures Optimized with QM/MM Method^a

state	r1	r2	r3	r4	r5	r6	r7	r8	r9
C2	3.00	1.69	4.67	1.90	1.04	2.93	1.00	1.90	2.88
B	2.97	1.70	4.65	1.87	1.04	2.90	1.45	1.15	2.60
Int	1.94	1.88	3.80	1.62	1.08	2.70	1.98	1.00	3.01
P	1.69	3.04	4.70	1.01	1.71	2.71	2.98	1.03	3.89
crystal	N/A	1.60	N/A	N/A	N/A	3.06	N/A	N/A	3.65

^{a}Reaction coordinates $r1-r9$ are defined in Figure 2. The two pre-cleavage states, C2 and B, the dianionic phosphorane intermediate state, Int, and the post-cleavage state, P, are depicted in Figure 3. The crystal structure used herein (ref 11) does not contain A-1(O2′).

upper sequential mechanism and found that both initial strings lead to the same mechanism.

QM/MM Geometry Optimizations. The initial strings for the QM/MM free energy simulations were generated from structures obtained by QM/MM geometry optimizations, which provided minimum-energy structures corresponding to State C2 and State B defined in Figure 3, as well as a phosphorane-like intermediate and a post-cleavage state. The key distances $r1-r9$ for these four states are provided in Table 2. In the two pre-cleavage states, the P−O5′ bond is fully formed, and the amine group of the cofactor is protonated, as indicated by the values of $r2$ and $r5$, respectively. Note that $r5$ is chosen to correspond to the specific amine hydrogen that is hydrogen-bonded to O5′. The C2 and B pre-cleavage states differ in the protonation site for the hydrogen-bonding interaction between A-1 and G40. In State C2, G40(N1) is deprotonated and A-1(O2′) is protonated, whereas in State B, G40(N1) is protonated and A-1(O2′) is deprotonated, as indicated by a comparison of $r7$ and $r8$ in Table 2. In the phosphorane-like intermediate (Int), denoted dianionic phosphorane in Figure 3, the P−O2′ and P−O5′ bond lengths (i.e., $r1$ and $r2$) are nearly equal, and the G1(O5′) is not yet protonated (i.e., the proton remains predominantly on the amine group of the cofactor, as indicated by $r4$ and $r5$).[64] In the

post-cleavage state, the P−O2′ bond has formed, the P−O5′ bond with G1 has broken, and the proton has been transferred from the amine group of the cofactor to G1(O5′), as indicated by $r1$, $r2$, and $r4$, respectively, and depicted in Figure 3.

Phosphate Bond Cleavage with External Base Activation. First we discuss the results from the "cleavage with external base" simulations. The initial string was associated with the concerted mechanism and was generated from a linear interpolation connecting the QM/MM optimized structures for the pre-cleavage State B and the post-cleavage product State P. Figure 6A depicts the initial (dashed) and converged (solid) strings, as well as the two-dimensional (2D) free energy surface, projected along the collective reaction coordinates ($r2−r1$) and ($r5−r4$). As shown in Figure 2, ($r2−r1$) is associated with the P−O2′/P−O5′ bond making/breaking, and ($r5−r4$) is associated with proton transfer from the cofactor amine group to G1(O5′). The one-dimensional (1D) free energy profile along the converged string is depicted in Figure 6B, and the changes in key reaction coordinates along this MFEP are shown in Figure 6C.

These simulations assume that an external base deprotonated either A-1(O2′) or G40(N1) prior to the start of the reaction pathway studied. On the basis of Figure 6A,B, the subsequent self-cleavage mechanism is concerted because there is no stable
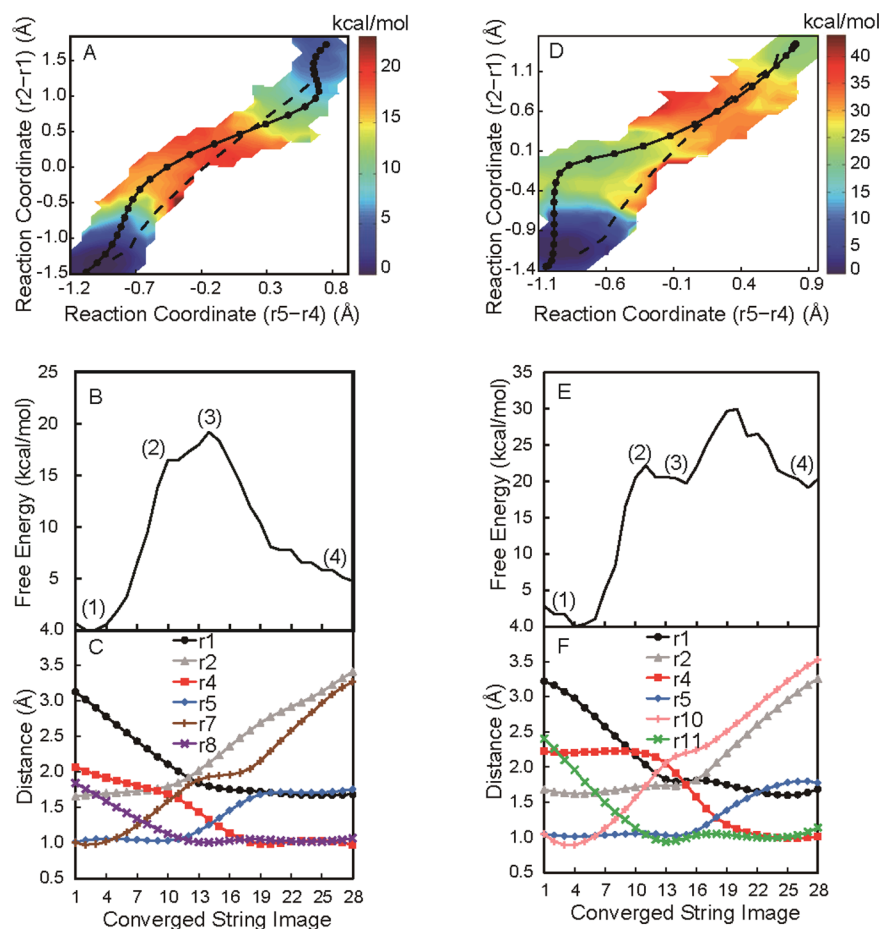
Figure 6. QM/MM free energy simulation results for (A−C) "cleavage with external base" and (D−F) "cleavage without external base" simulations. Panels (A) and (D) are 2D free energy surfaces projected in the ($r5−r4$) and ($r2−r1$) space, corresponding to proton transfer from the cofactor to G1(O5′) and P−O bond-making/breaking, respectively. The initial string corresponding to a concerted pathway is depicted as a dashed black line, and the converged MFEP is depicted as a solid black line. Panels (B) and (E) are 1D free energy profiles along the MFEPs shown in panels (A) and (D), respectively. Panels (C) and (F) depict the values of the most important reaction coordinates along the MFEPs. Each symbol corresponds to an image along the string, and the images have been rearranged to be equally spaced along the MFEP. Numbers in panels (B) and (E) indicate important states along the reaction pathway, which are illustrated in Figure 7

intermediate (i.e., there is only a single transition state). In the 2D reaction coordinate system shown in Figure 6A, the diagonal nature of the converged MFEP line suggests a concerted mechanism because the P−O bonds are forming and breaking simultaneously with the proton transfer reaction. The free energy surface corresponds to a pathway with a single free energy barrier that is ∼19 kcal/mol. As will be discussed below, this free energy barrier is consistent with experimentally measured rate constants.

Analysis of the changes of key reaction coordinates along the MFEP in Figure 6C provides insights into the details of the mechanism. As mentioned above, at the beginning of the initial string, A-1(O2′) was deprotonated and G40(N1) was protonated, as in State B. In the converged string, however, the proton was found to be bound to A-1(O2′) instead of G40(N1) at the beginning of the string (State C2) and then to transfer to G40(N1) (State B) during the self-cleavage reaction. The proton transfer reaction from A-1(O2′) to G40(N1) (i.e., from State C2 to State B) is depicted by the curves associated with reaction coordinates $r7$ and $r8$ in Figure 6C. Note that $r8$ decreases while $r7$ remains relatively constant over the first five images, indicating that G40(N1) is moving closer to the hydroxyl group on A-1(O2′). This movement occurs as O2′

begins its attack on the phosphorus, indicated by accompanying decreases in $r1$ and $r8$. In other words, G40 "follows" the O2′ as it attacks the phosphorus and therefore plays a role in directing this in-line attack. The proton is midway between A-1(O2′) and G40(N1) when the curves associated with reaction coordinates $r7$ and $r8$ cross, which occurs between images 7 and 8. At this point A-1(O2′) is ∼2.3 Å from the scissile phosphate, as indicated by the curve associated with $r1$. In other words, this proton transfer does not occur until the attacking A-1(O2′) is close enough to the scissile phosphate, thereby lowering the pK_a of A-1(O2′). The observation that the mechanism changed during the iterative procedure for the QM/MM free energy simulations (i.e., the converged string begins in State C2 even though the initial string began in State B) indicates that this approach does not rely strongly on the initial string but rather is flexible enough to evolve toward the MFEP even when the mechanism of the MFEP differs significantly from that of the initial string.

After this proton transfer from A-1(O2′) to G40(N1), the A-1(O2′) is 2.01 Å from the scissile phosphate ($r1$ in image 11 in Figure 6C). As indicated by reaction coordinates $r1$ and $r2$ in Figure 6C, the A-1(O2′)-P distance decreases only slightly as the G1(O5′)-P distance increases. At image 12, the A-1(O2′)-P

A: Cleavage with External Base
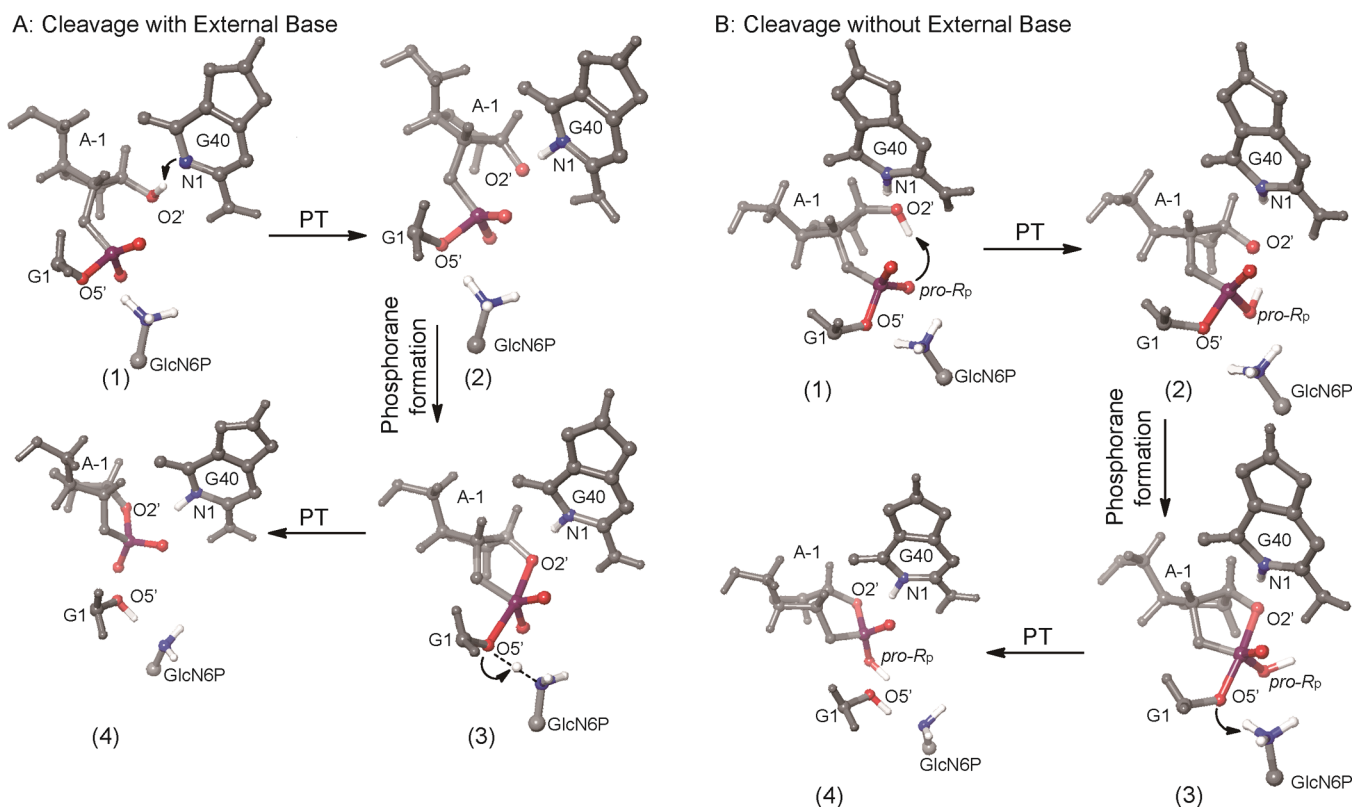
B: Cleavage without External Base

Figure 7. Representative structures along the MFEPs obtained from the QM/MM free energy simulations. The structures are numbered according to the locations along the MFEPs shown in Figure 6B and Figure 6E for the "cleavage with external base" and "cleavage without external base" simulations, respectively. The steps for "cleavage with external base" are as follows: (1) to (2) corresponds to PT from A-1(O2′) to G40(N1) accompanied by a decrease in the P−O2′ distance; (2) to (3) corresponds to the formation of a dianionic phosphorane transition state; (3) to (4) corresponds to PT from the cofactor N to G1(O5′). The mechanism associated with (1) to (4) is concerted yet asynchronous. The initial deprotonation of G40 is not shown here but is assumed to precede the concerted reaction shown here. The steps for "cleavage without external base" are as follows: (1) to (2) corresponds to PT from A-1(O2′) to the $pro\text{-}R_p$ oxygen; (2) to (3) corresponds to the formation of a monoanionic phosphorane intermediate; (3) to (4) corresponds to PT from the cofactor N to G1(O5′). This mechanism is sequential, where the first step is the concerted yet asynchronous process (1) to (3), producing a stable monoanionic phosphorane intermediate, and the second step is (3) to (4). Note that these structures do not represent stationary points but rather represent selected structures along the MFEP, and the charges are not shown.

and G1(O5′)-P distances are both 1.92 Å, resembling a phosphorane-like structure that is not a minimum on the free energy surface. When the G1(O5′)-P distance is slightly longer than the A-1(O2′)-P distance, the proton begins to transfer from the amine group of the cofactor to G1(O5′), as indicated by reaction coordinates $r4$ and $r5$ in Figure 6C. The A-1(O2′)-P bond forms and the G1(O5′)-P bond breaks during this proton transfer reaction. The proton is midway between the amine group and G1(O5′) between images 14 and 15, followed by completion of this proton transfer reaction by image 18. At the end of the reaction pathway, the distance between A-1 and both G40 and G1 increases because of the cleavage reaction, as revealed by large increases in $r7$ and $r2$, respectively.

The transition state, which is associated with the point of highest free energy, occurs at image 14. In this structure, the A-1(O2′)-P distance is 1.80 Å, the G1(O5′)-P distance is 2.13 Å, the $pro\text{-}R_p$-P distance is 1.50 Å, the $pro\text{-}S_p$-P distance is 1.56 Å, and the A-1(O3′)-P distance is 1.72 Å. Note that the last three bond lengths were not reaction coordinates but were obtained by averaging over the final iteration of the free energy simulation. This type of a product-like, "late" transition state, where the O5′-P distance is longer than the O2′-P distance, is typical for reactions in which the protonation of the leaving group is delayed.[65]

This analysis suggests that the reaction is concerted but asynchronous, where the sequence of events is as follows: (1) proton transfer from A-1(O2′) to G40(N1), (2) O−P bond making/breaking, and (3) proton transfer from the cofactor amine group to G1(O5′). This sequence of events is illustrated schematically in Figure 3A, which shows the progression from State C2 to State B to the dianionic phosphorane-like transition state to the post-cleavage product. In addition, Figure 7A depicts structures along the MFEP denoted numerically in Figure 6B. Note that these structures are not stable intermediates (i.e., they are not minima on the free energy surface) but rather simply represent selected structures along the MFEP to illustrate the asynchronous mechanism. Moreover, these steps are not completely independent and overlap somewhat during the reaction pathway, as illustrated by Figure 6C. The mechanism is still defined to be concerted in the sense that the reactant and product are connected by a single transition state without any stable intermediates. To test the dependence of this result on the initial string, we constructed an initial string associated with the sequential mechanism using the QM/MM optimized phosphorane intermediate. As shown in Figure S6, this sequential mechanism evolved toward a concerted pathway during the iterative procedure. These simulations suggest that the "cleavage with external base"

simulations occur via a concerted yet asynchronous mechanism, without a strong dependence on the initial string.

In these simulations, we assumed that the scissile phosphate is not protonated. At image 12 of the converged string, the O2′-P and O5′-P distances are both 1.92 Å, resembling the phosphorane intermediate obtained from the QM/MM geometry optimizations, which exhibited distances of 1.94 and 1.88 Å, respectively. If this reaction occurs at relatively low pH, the dianionic phosphorane may become singly or doubly protonated, potentially favoring a stable phosphorane intermediate and therefore a sequential mechanism. Previous experiments and calculations on model systems suggest that the first protonation of a dianionic phosphorane is associated with a pK_a of ~14.[66,67] Although adding a proton to the phosphorane intermediate was not allowed for these simulations, we considered this possibility in the "cleavage without external base" simulations described in the next subsection.

In addition, the dianionic phosphorane may be stabilized by an electrostatic interaction between the positive charge of the cofactor amine group and the negative charges of the pro-R_P and pro-S_P oxygens.[68] From the reactant to the transition state along the MFEP, we found that the pro-S_P and pro-R_P oxygen atoms are 3.0−3.5 and 3.5−4.0 Å, respectively, from the nitrogen atom on the cofactor amine group (Figure S9). Natural Bond Orbital (NBO) analyses[69,70] on the QM atoms of selected snapshots along the MFEP indicate substantial negative charges on the two nonbridging oxygen atoms and a positive charge on the cofactor amine group up to the transition state (Figure S9). The relatively short distances and opposite charges support electrostatic stabilization effects in this region. In addition, we observed stable hydrogen-bonding interactions between the pro-S_P oxygen and G39(N1) and G39(N2) and between the pro-R_P oxygen and G65(N2) and O1 of GlcN6P, providing further stabilization of the negative charges on these oxygen atoms (Figure S13).

Phosphate Bond Cleavage without External Base. The "cleavage without external base" simulations assume that both A-1(O2′) and G40(N1) are protonated prior to the start of the reaction pathway studied, as in State A (Figure 3B). The initial string was associated with the concerted mechanism and was generated from a linear interpolation connecting the QM/MM optimized structures for the pre-cleavage State A, in which both A-1(O2′) and G40(N1) are protonated, and the post-cleavage product state in which the proton has transferred from A-1(O2′) to the pro-R_P oxygen (State P′). Note the similarity of the initial (dashed) strings for the two types of simulations in Figure 6A,D. Figure 6D depicts the initial (dashed) and converged (solid) strings, as well as the 2D free energy surface, projected along the collective reaction coordinates $(r2−r1)$ and $(r5−r4)$. Figure 6E depicts the 1D free energy profile along the converged string, and Figure 6F depicts the changes in key reaction coordinates along this MFEP.

The mechanism obtained from these simulations differs from that obtained with an external base in that the initial proton transfer reaction is from A-1(O2′) to the pro-R_P oxygen instead of to G40(N1). In addition, the phosphorane-like structure that occurs along the reaction pathway is monoanionic instead of dianionic and represents a stable intermediate (i.e., a minimum on the free energy surface). In the simulations without an external base, the proton on A-1(O2′) transfers to the pro-R_P oxygen in the early stage of the MFEP, concurrent with the attack of A-1(O2′) on the scissile phosphate. This initial proton transfer is indicated by the intersection of the curves

corresponding to $r10$ and $r11$ in Figure 6F, while the concurrent attack of A-1(O2′) on the scissile phosphate is indicated by the decrease of reaction coordinate $r1$. A minimum on the free energy surface occurs subsequent to this initial proton transfer reaction, corresponding to a monoanionic phosphorane intermediate in which the pro-R_P oxygen is protonated (image 15). This phosphorane intermediate is a minimum for these simulations because it is stabilized by proton transfer to the pro-R_P oxygen to generate the monoanionic phosphorane, in contrast to the dianionic phosphorane-like structure, which is not a stable intermediate in the "cleavage with external base" simulations.

Because a minimum is observed on the free energy surface, the reaction in the "cleavage without external base" simulations is determined to be sequential with a phosphorane intermediate. The first step is a concerted yet asynchronous process involving the initial proton transfer from A-1(O2′) to the pro-R_P oxygen followed by the formation of a phosphorane intermediate. This first step is concerted in that the reactant and the phosphorane intermediate are connected by a single transition state, but the proton transfer and phosphorane intermediate formation occur in an asynchronous yet coupled fashion. In State B′, the hydrogen-bonding interactions of A-1(O2′) with both the pro-R_P oxygen ($r10$ and $r11$ in Figure 6F) and G40(N1) ($r9$ in Figure S14) restrict the in-line attack of A-1(O2′) on the phosphate. The second step in these simulations involves completion of the P−O bond making/breaking and proton transfer from the cofactor to G1(O5′). This second step is also concerted in that the phosphorane intermediate and the product are connected by a single transition state. G40 remains neutral during this entire mechanism and is not involved in any proton transfer reaction, although it could assist in orienting O2′ for in-line attack of the phosphorus. This sequence of events is illustrated schematically in Figure 3B, which shows the progression from State A to State B′ to the monoanionic phosphorane intermediate to the post-cleavage product. In addition, Figure 7B depicts structures along the MFEP denoted numerically in Figure 6E. Some of these structures, such as the phosphorane intermediate, are stable intermediates, but others are simply structures occurring along the MFEP to illustrate the asynchronous mechanism.

The overall effective free energy barrier along this MFEP is ~30 kcal/mol. The free energy barrier for the first step is ~22 kcal/mol, and the phosphorane intermediate is ~20 kcal/mol higher than the reactant. This free energy difference between the phosphorane intermediate and the reactant includes the free energy associated with proton transfer from A-1(O2′) to the pro-R_P oxygen and the formation of the phosphorane intermediate. In principle, the free energy associated with the initial proton transfer can be estimated from the pK_a differences between the proton donor and acceptor, but this initial proton transfer cannot be rigorously separated from the formation of the phosphorane intermediate, which also occurs in the first step. The free energy barrier for the second step is ~10 kcal/mol and corresponds mainly to proton transfer from the cofactor to G1(O5′), as well as completion of the P−O bond making/breaking.

In comparing the mechanisms and free energy barriers obtained from the two types of simulations, it is important to keep in mind that the "cleavage with external base" mechanism requires an initial deprotonation of the G40(N1), which would diminish the observed rate constant in the form of a pre-equilibrium constant. This contribution to the free energy can

be determined from the pK_a of G40(N1), which is estimated experimentally to be around 10.[21] Using the Henderson–Hasselbalch equation for a pH of 7, the reaction free energy for the deprotonation of G40(N1) (i.e., State A to State C1) is ~4 kcal/mol. Thus, the overall effective free energy barrier associated with the "cleavage with external base" simulations is the sum of 19 kcal/mol shown in Figure 6B and 4 kcal/mol associated with the pre-equilibrium constant for initial deprotonation of G40(N1), leading to an overall effective free energy barrier of ~23 kcal/mol. Although this analysis provides only qualitative estimates, comparison of the overall effective free energy barriers of ~23 kcal/mol and ~30 kcal/mol for the simulations with and without external base, respectively, suggests that the mechanism using an external base may be favorable over the mechanism that does not invoke an external base.

Proton Transfer between A-1(O2′) and G40(N1): The O2′/N1 PT Simulations. The initial proton transfer reaction between A-1(O2′) and G40(N1) was also examined independently with this QM/MM free energy simulation approach. This proton transfer corresponds to conversion between States C2 and B in Figure 3. In these simulations, denoted "O2′/N1 PT", only the sugar ring in A-1 and the G40 base were included in the QM region, and one of these residues was assumed to have been deprotonated by an external base prior to the reaction studied. The self-cleavage reaction was inhibited because the scissile phosphate moiety was not included in the QM region. The MFEP generated by these simulations indicates a minimum for the system with the proton covalently bonded to A-1(O2′) but not with the proton covalently bonded to G40(N1) (Figure S10). Moreover, the distance between A-1(O2′) and G40N1 ($r9$) is shorter when the proton is bonded to A-1(O2′), providing a more stable hydrogen bond (Figure S10). This result implies that the proton will remain on A-1(O2′) prior to the self-cleavage reaction.

The results from these simulations are consistent with the results from the "cleavage with external base" simulations, where the proton is covalently bonded to A-1(O2′) in the early stages of the MFEP and is not transferred to G40(N1) until the attacking A-1(O2′) is ~2.3 Å from the phosphorus. These phosphate bond cleavage simulations suggest that A-1(O2′) starts attacking the phosphate as a hydroxyl group, concurrent with G40(N1) moving closer to the H of the hydroxyl group, and then transfers its proton to G40(N1) midway through this attack. When the O2′ of A-1 gets close enough to the phosphate, its pK_a decreases to an extent that it can be deprotonated by G40(N1), which has moved closer and thus facilitates the proton transfer reaction. In the "O2′/N1 PT" simulations, A-1(O2′) does not get close enough to the phosphate to enable this proton transfer reaction because self-cleavage is inhibited by omitting the scissile phosphate from the QM region. In this somewhat artificial situation, the pK_a of A-1(O2′) remains much higher than that of G40(N1), and proton transfer to G40(N1) does not occur.

Transformation between States C1 and C2: The 2′OH Rearrangement Simulations. In Figure 3A, G40(N1) is deprotonated by an external base, resulting in State C1, followed by a hydrogen-bonding rearrangement to produce State C2, which undergoes a proton transfer reaction to produce State B. The transformation between States C1 and C2 involves the rotation of the 2′OH to move the hydrogen bond from the *pro-R*$_p$ oxygen to G40(N1). Because it does not

require the making or breaking of chemical bonds, this rearrangement can be described by a classical force field. In principle, the relative free energies of States C1 and C2 could be determined from extensive classical MD via the relative populations of these two states. As shown in Figure 4, we observed conversions between States C1 and C2 during a 25 ns classical MD trajectory, which illustrates that this transformation is not sterically forbidden. Although we also propagated another independent 25 ns trajectory, 50 ns of classical MD is not a sufficient amount of sampling to allow a quantitative calculation of the relative free energies of States C1 and C2.

Instead, we used umbrella sampling with the finite temperature string method to calculate the relative free energies of States C1 and C2 in the "2′OH rearrangement" simulations. Figure S11 depicts the two angular coordinates used to describe this hydrogen-bonding rearrangement, and Figure S12 presents the 2D free energy surface and MFEP. The free energy barrier for this hydrogen-bonding rearrangement is ~5 kcal/mol, indicating that it would be reasonably facile at room temperature. The free energy of state C1 is ~1 kcal/mol lower than that of state C2, suggesting that they are almost equally thermodynamically favorable. These results are used in the pK_a calculations described in the next subsection.

3. pK_a Calculations. We used the PB/LRA method to calculate the relative pK_a of G40(N1) and A-1(O2′) in an effort to identify the thermodynamically favorable deprotonation site, thereby determining whether State A will be more likely to evolve to State B or State C1 (Figure 3). Prior to these calculations, we benchmarked this methodology for these types of systems by calculating the pK_a of A(O2′) in ApG in aqueous solution. The details of these benchmarking calculations are provided in the Supporting Information (pp S23 and S24). We calculated the pK_a of the 2′-OH group of adenine in ApG in 1 M NaCl to be ~13.3, which is ~1 pK_a unit higher than the experimentally measured value of 12.31 (Table 3).[63]

Table 3. pK_a Values from PB/LRA Calculations with ε_{env} = 4^a

system	calculated	experimental
ApG (A(O2′))	13.3 (0.7)	12.31 ± 0.02^c
G40 (State A to C1/C2)^b	12.8 (1.6)	~10.0^d
A-1 (State A to B)	19.6 (0.8)	N/A

^{a}The numbers in parentheses are standard deviations. The calculated value for ApG corresponds to 1 M NaCl, and the calculated values for G40 and A-1 correspond to 150 mM NaCl. ^{b}The pK_a of G40 is a weighted average of the value obtained from MD configurations in State C1, 13.0 (1.6), and State C2, 11.6 (1.5). ^{c}Experimental data from ref 63. The experimental pK_a for ApG (deprotonation of O2′ of A) was measured in 1 M NaCl. ^{d}Obtained by extrapolation of data in ref 21. Note that this value could be shifted upward to at least ~10.5 based on the possibility that the titration may be leveling off due to alkaline denaturation.

To provide additional benchmarking, we calculated the pK_a of G40(N1) and compared it to an experimentally measured value. When G40(N1) is deprotonated and A-1(O2′) is protonated, the system samples two states, denoted State C1 and State C2 in Figure 3. Thus, the pK_a of G40(N1) corresponds to the deprotonation of State A to form a mixture of States C1 and C2. We used the free energy difference calculated from the "2′OH rearrangement" simulations

described in the previous subsection to weight the pK_a values associated with these two states. Upon comparing relative free energies, C1 is lower than C2 by ~1 kcal/mol on the basis of the 2′OH rearrangement simulations, but C2 is lower than C1 by ~2 kcal/mol on the basis of the PB/LRA pK_a calculations; this difference is within the error bars of both types of calculations and does not impact the qualitative conclusion that States C1 and C2 are approximately equally thermodynamically favorable. Again, the details of these benchmarking calculations are provided in the Supporting Information (pp S23 and S24). We calculated the pK_a of the G40(N1) to be ~12.8, which is ~2.8 units higher than the value of ~10.0 determined from experimental measurements. Overall, these benchmarking studies suggest that the calculated pK_a's in the ribozyme may be overestimated by ~1−3 pK_a units.

Following these benchmarking studies, we calculated the pK_a of A-1(O2′), which has not been experimentally measured in the *glmS* ribozyme. When A-1(O2′) is deprotonated and G40(N1) is protonated, the system samples a single state, denoted State B in Figure 3. Thus, the pK_a of A-1(O2′) corresponds to the deprotonation of State A to form State B. As with ApG, we used ApEt in solution as the reference,[63] and the experimentally measured[63] pK_a of 12.51 ± 0.05 for ApEt at 1 M NaCl was extrapolated to 150 mM NaCl using the approach of Li and Breaker[71] to obtain a pK_a of 12.7 for ApEt at 150 mM NaCl. Using this reference, the pK_a value for A-1(O2′) was calculated as ~19.6 at 150 mM NaCl.

Thus, the calculations indicate that the pK_a of A-1(O2′) is ~7 units higher than that of G40(N1). Analyzing such a *difference* in pK_a avoids systematic errors observed in the benchmarking studies because systematic errors cancel out. The unusually high pK_a for A-1(O2′) may be due to the strong hydrogen-bonding interaction with the *pro-R*$_P$ oxygen of the scissile phosphate, as depicted in Figure 3 for State A. This pK_a difference supports the formation of State C1 rather than State B from State A because deprotonation of G40(N1) is ~9 kcal/mol more thermodynamically favorable than deprotonation of A-1(O2′).

In addition, these pK_a calculations imply that the free energy of State B is ~11 kcal/mol higher than that of State C2. This result is consistent with the "O2′/N1 PT" free energy simulations that revealed a significantly lower free energy for State C2 than for State B. This result is also consistent with the "cleavage with external base" MFEP, where the proton starts on A-1(O2′) and is only transferred to G40(N1) as the nucleophilic attack of A-1(O2′) on the phosphate group lowers its pK_a value. Thus, the pK_a calculations are consistent with the free energy simulations described in the previous subsection.

4. Experimentally Measured Rate Constant. Experiments were carried out to determine whether the calculated free energy barrier for the self-cleavage reaction is reasonable. We initially chose a Mg^{2+} concentration of 3 mM as it is close to biological for bacteria where the *glmS* ribozyme is found, as well as physiological ionic strength of 150 mM NaCl.[72,73] Moreover, these Mg^{2+} and Na^+ conditions are close to those used in the calculations. Plots of fraction cleaved versus time for the 3 mM Mg^{2+} conditions are shown in Figure 8A, where the data from the average of three trials were fit to the single exponential expression given in eq 2. The corresponding rate constant at 3 mM $MgCl_2$ is 0.013 ± 0.003 min^{-1} (Figure 8A). The data show only one phase with an amplitude of ~0.6.
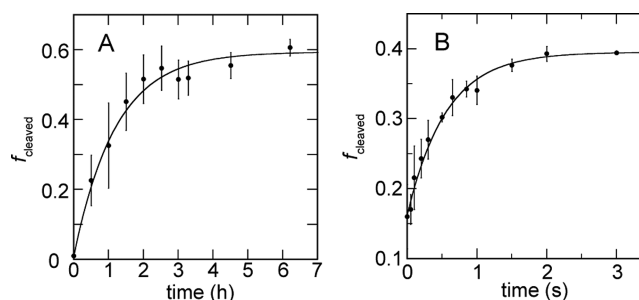

Figure 8. Plot of fraction cleaved versus time for experiments. (A) 3 mM Mg^{2+} conditions. Average rate constant is 0.013 ± 0.003 min^{-1}. (B) 30 mM Mg^{2+} conditions. Average rate constant is 110 ± 10 min^{-1}. For each Mg^{2+} condition, three trials were performed and data points were averaged, yielding an average rate constant.

Substantial degradation of the ribozyme was observed at time points longer than those used in the fits.

From the theoretical simulations, an overall effective free energy barrier of ~23 kcal/mol was calculated for the mechanism that was activated by an external base. We converted this effective free energy barrier to a rate constant using the transition state theory expression

$$k = \frac{k_B T}{h} \, e^{-\Delta G^{\ddagger}/RT} \tag{3}$$

where k_B is Boltzmann's constant, h is Planck's constant, and $\Delta G^{\ddagger}$ is the effective free energy barrier. At 300 K, the prefactor is 6.25 ps^{-1}, leading to a rate constant of 0.0066 min^{-1}. This value is comparable to the experimentally measured rate constant of 0.013 ± 0.003 min^{-1} at 3 mM $MgCl_2$. On the other hand, if the experimental rate constant is converted to a free energy barrier using eq 3, the experimentally obtained free energy barrier is 22.6 ± 0.2 kcal/mol, which is in excellent agreement with the calculated free energy barrier of 23 kcal/mol, which has an error of ~1 kcal/mol based on the statistical error analysis in the Supporting Information and additional error introduced by the estimate of the free energy of the initial deprotonation.

We also measured the rate constant for *glmS* self-cleavage at the higher Mg^{2+} concentration of 30 mM. The rate under these conditions was considerably faster, with an observed rate constant of 110 ± 10 min^{-1} (Figure 8B). The data collected at 30 mM $MgCl_2$ show only one phase with an amplitude of ~0.3. This rate constant agrees with other measurements under similar temperature and Mg^{2+} conditions.[15] Given that our *in silico* conditions are more similar to the 3 mM Mg^{2+} experimental conditions, we compare our calculations to those results. Future experimental and theoretical studies will be required to address the possible roles of Mg^{2+} in the reaction.

■ CONCLUSIONS

QM/MM free energy simulations and pK_a calculations were performed to investigate the self-cleavage reaction catalyzed by the *glmS* ribozyme. In one proposed mechanism, the reaction is initiated by deprotonation of G40(N1) by an external base, presumably by a metal-bound hydroxide ion, a buffer molecule, or another atom on the ribozyme itself. G40(N1) rather than A-1(O2′) is deprotonated because the pK_a of A-1(O2′) is significantly higher than that of G40(N1) in this initial state, most likely due to the hydrogen-bonding interaction between

A-1(2′OH) and the *pro-R*$_p$ oxygen. Following deprotonation of G40(N1), hydrogen-bonding between A-1(2′OH) and the negatively charged G40(N1) is also possible. Both of these hydrogen-bonding configurations are thermodynamically accessible at room temperature, but only the A-1(2′OH):G40(N1) hydrogen-bonding configuration can initiate the self-cleavage reaction. In this catalytically active hydrogen-bonding configuration, the proton remains bound to A-1(O2′) because the pK_a of A-1(O2′) is significantly higher than that of G40(N1). As A-1(O2′) attacks the phosphorus to form an O−P bond, the pK_a of A-1(O2′) is lowered and the G40(N1)-H distance shortens, allowing the proton to transfer from A-1(O2′) to G40(N1). The subsequently formed dianionic phosphorane-like structure is not a stable intermediate, but rather represents a transition state on the free energy surface. The self-cleavage reaction continues as a proton transfers from the cofactor amine group to G1(O5′) and the P−O bond making/breaking is completed, generating the cleaved product state. In this mechanism, after the initial deprotonation of G40(N1), the self-cleavage reaction is concerted yet asynchronous.

The pK_a of G40(N1) is still quite high, as determined experimentally[74] and from the present calculations. The even higher pK_a of A-1(O2′) calculated herein is most likely due to the hydrogen-bonding interaction of A-1(2′OH) with the *pro-R*$_p$ oxygen, which ties up the proton and keeps it from transferring. Another possibility is that a solvent fluctuation breaks this hydrogen-bonding interaction, thereby significantly lowering the pK_a of A-1(O2′) and producing State B directly from State A. On the basis of our calculated free energies, State B would probably evolve to State C2, which is significantly lower in free energy, via proton transfer from G40(N1) to A-1(O2′). When A-1(O2′) attacks the phosphorus, the proton would transfer back to G40(N1), as observed in the QM/MM free energy simulations. According to the mechanism requiring initial deprotonation of G40(N1) or A-1(O2′), the shift of the high pK_a of G40(N1) even *further* from neutrality is important for assisting in deprotonation of A-1(O2′) during the self-cleavage reaction. This observation is quite interesting, as nucleobase catalysis typically assumes the pK_a of the nucleobase should shift *toward* neutrality; however, this is only valid if the nucleobase is a classical general acid−base catalyst.

A second mechanism that does not require initial deprotonation by an external base was also investigated. This mechanism was found to be sequential in that a stable phosphorane intermediate representing a minimum on the free energy surface was observed. As in the mechanism activated by an external base, the initial state has a hydrogen-bonding interaction between A-1(2′OH) and the nonbridging oxygen. The proton remains bound to A-1(O2′) because the pK_a of A-1(O2′) is much higher than that of the nonbridging oxygen. As A-1(O2′) attacks the phosphorus to form an O−P bond, the pK_a of A-1(O2′) is lowered, allowing the proton to transfer from A-1(O2′) to the *pro-R*$_p$ oxygen during formation of the stable monoanionic phosphorane intermediate. Thus, the first step of this sequential mechanism involves proton transfer from A-1(O2′) to the *pro-R*$_p$ oxygen and formation of a monoanionic phosphorane intermediate. The second step involves proton transfer from the cofactor to G1(O5′) as well as completion of the P−O bond making/breaking. In this case, the self-cleavage process is sequential because the phosphorane intermediate is stabilized by initial proton transfer from A-1(O2′) to the *pro-R*$_p$ oxygen, leading to a monoanionic rather than a dianionic phosphorane.

The rate constant was measured experimentally to be 0.013 ± 0.003 min^{-1} for 3 mM MgCl$_2$. Estimating the preequilibrium constant for the initial deprotonation, the overall effective free energy barrier for the mechanism activated by an external base was calculated to be ∼23 kcal/mol. The calculated rate constant of 0.0066 min^{-1} is in qualitative agreement with the experimentally measured rate constant at low Mg^{2+} ion concentration. The overall free energy barrier for the second proposed mechanism, which does not require an external base, was calculated to be ∼30 kcal/mol, implying that this mechanism is less favorable than the first mechanism and has a rate constant that is significantly lower than the experimentally measured value. Therefore, the second proposed mechanism is determined to be less likely, although it cannot be ruled out completely, particularly given the potential role of Mg^{2+} ions in providing electrostatic stabilization.

The catalytic role of G40 is of interest for understanding not only the mechanism of *glmS* cleavage, but also the mechanism of cleavage of other small ribozymes. A counterpart to G40 is found in most of the other small ribozymes (G8 in the hairpin,[75] G12 in the hammerhead,[76,77] G638 in the VS,[78] and G45 in the twister[79] ribozyme), with the HDV ribozyme being an exception to this trend. Future calculations and experiments will be needed to test the generality of our findings among the G40-counterpart ribozymes as well as any relationship to the HDV ribozyme. On the basis of classical MD simulations, the analogous guanine, G8, in the hairpin ribozyme was also proposed to assist the in-line attack through hydrogen-bonding interactions.[80] Because these classical MD simulations did not allow chemical bonds to break and form, however, they did not provide information about the potential role of this guanine in acid/base chemistry, although earlier QM/MM geometry optimizations[81] suggested that a deprotonated G8$^-$, together with a protonated A38H$^+$, could yield a reasonable activation barrier for self-cleavage for the hairpin ribozyme. The present QM/MM free energy simulations of the *glmS* ribozyme have provided insights into the catalytic role of the active site guanine in this ribozyme using QM/MM methodology that includes conformational sampling. As the role of these active site guanines may be universal, deciphering the role of G40 in *glmS* cleavage may provide insight into the catalytic roles of active site guanines in other small ribozymes.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

RESP procedure and charges; additional analysis of classical MD simulations; details about QM/MM geometry optimizations and further analysis; additional information about free energy simulations, including computational details, convergence criteria, statistical error analysis, comparison of two simulations with different initial strings, and free energy surfaces; NBO analysis; PB/LRA benchmarking, computational details, and results with a higher dielectric constant; additional figures describing hydrogen-bonding interactions and mechanisms; coordinates and energies of QM/MM-optimized geometries. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

pcb5@psu.edu
shs3@illinois.edu

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Prody, G. A.; Bakos, J. T.; Buzayan, J. M.; Schneider, I. R.; Bruening, G. Science 1986, 231, 1577.

(2) Buzayan, J. M.; Gerlach, W. L.; Bruening, G. Nature 1986, 323, 349.

(3) Sharmeen, L.; Kuo, M. Y. P.; Dintergottlieb, G.; Taylor, J. J. Virol. 1988, 62, 2674.

(4) Saville, B. J.; Collins, R. A. Cell 1990, 61, 685.

(5) Roth, A.; Weinberg, Z.; Chen, A. G. Y.; Kim, P. B.; Ames, T. D.; Breaker, R. R. Nat. Chem. Biol. 2014, 10, 56.

(6) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. Nature 2004, 428, 281.

(7) McCown, P. J.; Roth, A.; Breaker, R. R. RNA 2011, 17, 728.

(8) Barrick, J. E.; Corbino, K. A.; Winkler, W. C.; Nahvi, A.; Mandal, M.; Collins, J.; Lee, M.; Roth, A.; Sudarsan, N.; Jona, I.; Wickiser, J. K.; Breaker, R. R. Proc. Natl. Acad. Sci. U.S.A. 2004, 101, 6421.

(9) Milewski, S. Biochim. Biophys. Acta: Protein Struct. Mol. Enzymol. 2002, 1597, 173.

(10) Klein, D. J.; Ferre-D'Amare, A. R. Science 2006, 313, 1752.

(11) Klein, D. J.; Wilkinson, S. R.; Been, M. D.; Ferre-D'Amare, A. R. J. Mol. Biol. 2007, 373, 178.

(12) Cochrane, J. C.; Lipchock, S. V.; Strobel, S. A. Chem. Biol. 2007, 14, 97.

(13) Cochrane, J. C.; Lipchock, S. V.; Smith, K. D.; Strobel, S. A. Biochemistry 2009, 48, 3239.

(14) McCarthy, T. J.; Plog, M. A.; Floy, S. A.; Jansen, J. A.; Soukup, J. K.; Soukup, G. A. Chem. Biol. 2005, 12, 1221.

(15) Viladoms, J.; Fedor, M. J. J. Am. Chem. Soc. 2012, 134, 19043.

(16) Roth, A.; Nahvi, A.; Lee, M.; Jona, I.; Breaker, R. R. RNA 2006, 12, 607.

(17) Gong, B.; Klein, D. J.; Ferre-D'Amare, A. R.; Carey, P. R. J. Am. Chem. Soc. 2011, 133, 14188.

(18) Davis, J. H.; Dunican, B. F.; Strobel, S. A. Biochemistry 2011, 50, 7236.

(19) Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. J. Phys. Chem. B 2010, 114, 8701.

(20) Xin, Y.; Hamelberg, D. RNA 2010, 16, 2455.

(21) Viladoms, J.; Scott, L. G.; Fedor, M. J. J. Am. Chem. Soc. 2011, 133, 18388.

(22) Klein, D. J.; Been, M. D.; Ferre-D'Amare, A. R. J. Am. Chem. Soc. 2007, 129, 14858.

(23) Brooks, K. M.; Hampel, K. J. Biochemistry 2011, 50, 2424.

(24) Izatt, R. M.; Christen, Jj; Rytting, J. H. Chem. Rev. 1971, 71, 439.

(25) Klawuhn, K.; Jansen, J. A.; Souchek, J.; Soukup, G. A.; Soukup, J. K. ChemBioChem 2010, 11, 2567.

(26) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. Acc. Chem. Res. 2010, 43, 40.

(27) Banas, P.; Jurecka, P.; Walter, N. G.; Sponer, J.; Otyepka, M. Methods 2009, 49, 202.

(28) Weinan, E.; Ren, W. Q.; Vanden-Eijnden, E. J. Phys. Chem. B 2005, 109, 6688.

(29) Radhakrishnan, R. Biophys. J. 2007, 93, 2391.

(30) Torrie, G. M.; Valleau, J. P. J. Comput. Phys. 1977, 23, 187.

(31) Wojtas-Niziurski, W.; Meng, Y. L.; Roux, B.; Berneche, S. J. Chem. Theory Comput. 2013, 9, 1885.

(32) Lans, I.; Medina, M.; Rosta, E.; Hummer, G.; Garcia-Viloca, M.; Lluch, J. M.; Gonzalez-Lafont, A. J. Am. Chem. Soc. 2012, 134, 20544.

(33) Zhu, F. Q.; Hummer, G. Proc. Natl. Acad. Sci. U.S.A. 2010, 107, 19814.

(34) Pan, A. C.; Sezer, D.; Roux, B. J. Phys. Chem. B 2008, 112, 3432.

(35) Zheng, H.; Zhang, Y. K. J. Chem. Phys. 2008, 128.

(36) Faradjian, A. K.; Elber, R. J. Chem. Phys. 2004, 120, 10880.

(37) Cui, Q.; Karplus, M. J. Phys. Chem. B 2000, 104, 3721.

(38) Ganguly, A.; Thaplyal, P.; Rosta, E.; Bevilacqua, P. C.; Hammes-Schiffer, S. J. Am. Chem. Soc. 2014, 136, 1483.

(39) Hampel, K. J.; Tinsley, M. M. Biochemistry 2006, 45, 7861.

(40) Maestro, version 9.3; Schrödinger, LLC: New York, 2012.

(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J. Chem. Phys. 1983, 79, 926.

(42) Maguire, M. E.; Cowan, J. A. Biometals 2002, 15, 203.

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

(44) Darden, T.; York, D.; Pedersen, L. J. Chem. Phys. 1993, 98, 10089.

(45) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. J. Phys. Chem. 1993, 97, 10269.

(46) Desmond Molecular Dynamics System, version 3.1; D.E. Shaw Research: New York, 2012.

(47) Veeraraghavan, N.; Ganguly, A.; Golden, B. L.; Bevilacqua, P. C.; Hammes-Schiffer, S. J. Phys. Chem. B 2011, 115, 8346.

(48) Veeraraghavan, N.; Ganguly, A.; Chen, J. H.; Bevilacqua, P. C.; Hammes-Schiffer, S.; Golden, B. L. Biochemistry 2011, 50, 2672.

(49) Vasilyev, V. V. J. Mol. Struct.: THEOCHEM 1994, 110, 129.

(50) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., III; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Phys. Chem. Chem. Phys. 2006, 8, 3172.

(51) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J. J. Comput. Chem. 1983, 4, 187.

(52) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. J. Am. Chem. Soc. 2011, 133, 8934.

(53) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. J. Comput. Chem. 1992, 13, 1011.

(54) Warwicker, J.; Watson, H. C. J. Mol. Biol. 1982, 157, 671.

(55) Davis, M. E.; Mccammon, J. A. Chem. Rev. 1990, 90, 509.

(56) Honig, B.; Nicholls, A. Science 1995, 268, 1144.

(57) Tang, C. L.; Alexov, E.; Pyle, A. M.; Honig, B. J. Mol. Biol. 2007, 366, 1475.

(58) Goh, G. B.; Knight, J. L.; Brooks, C. L. J. Chem. Theory Comput. 2013, 9, 935.

(59) Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. Biophys. J. 2002, 83, 1731.

(60) Alexov, E. G.; Gunner, M. R. Biophys. J. 1997, 72, 2075.

(61) Archontis, G.; Simonson, T. Biophys. J. 2005, 88, 3888.

(62) Gilson, M. K.; Honig, B. H. Biopolymers 1986, 25, 2097.

(63) Acharya, S.; Foldesi, A.; Chattopadhyaya, J. J. Org. Chem. 2003, 68, 1906.

(64) Perreault, D. M.; Anslyn, E. V. Angew. Chem., Int. Ed. 1997, 36, 432.

(65) Kellerman, D. L.; York, D. M.; Piccirilli, J. A.; Harris, M. E. Curr. Opin. Chem. Biol. 2014, 21, 96.

(66) Davies, J. E.; Doltsinis, N. L.; Kirby, A. J.; Roussev, C. D.; Sprik, M. *J. Am. Chem. Soc.* **2002**, *124*, 6594.

(67) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 5010.

(68) Ferre-D'Amare, A. R.; Scott, W. G. *Cold Spring Harbor Perspect. Biol.* **2010**, *2*, No. a003574.

(69) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.

(70) Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066.

(71) Li, Y. F.; Breaker, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 5364.

(72) Lusk, J. E.; Williams, R. J.; Kennedy, E. P. *J. Biol. Chem.* **1968**, *243*, 2618.

(73) Truong, D. M.; Sidote, D. J.; Russell, R.; Lambowitz, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 3800.

(74) Fedor, M. J.; Williamson, J. R. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 399.

(75) Kath-Schorr, S.; Wilson, T. J.; Li, N. S.; Lu, J.; Piccirilli, J. A.; Lilley, D. M. J. *J. Am. Chem. Soc.* **2012**, *134*, 16717.

(76) Thomas, J. M.; Perrin, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 15467.

(77) Han, J.; Burke, J. M. *Biochemistry* **2005**, *44*, 7864.

(78) Wilson, T. J.; Li, N. S.; Lu, J.; Frederiksen, J. K.; Piccirilli, J. A.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11751.

(79) Liu, Y. W.; Wilson, T. J.; McPhee, S. A.; Lilley, D. M. J. *Nat. Chem. Biol.* **2014**, *10*, 739.

(80) Heldenbrand, H.; Janowski, P. A.; Giambasu, G.; Giese, T. J.; Wedekind, J. E.; York, D. M. *J. Am. Chem. Soc.* **2014**, *136*, 7789.

(81) Mlynsky, V.; Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2011**, *115*, 13911.