

# Transition State Features in the Hepatitis Delta Virus Ribozyme Reaction Revealed by Atomic Perturbations

Selene C. Koo,<sup>†</sup> Jun Lu,<sup>†,#</sup> Nan-Sheng Li,<sup>†</sup> Edward Leung,<sup>†</sup> Subha R. Das,<sup>‡</sup> Michael E. Harris,<sup>§</sup> and Joseph A. Piccirilli<sup>\*,†</sup>

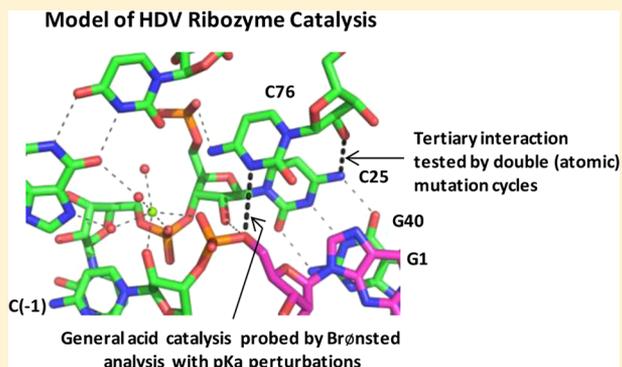
<sup>†</sup>Department of Biochemistry & Molecular Biology and Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, United States

<sup>‡</sup>Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

<sup>§</sup>Department of Biochemistry and Center for Proteomics, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, United States

## S Supporting Information

**ABSTRACT:** Endonucleolytic ribozymes constitute a class of non-coding RNAs that catalyze single-strand RNA scission. With crystal structures available for all of the known ribozymes, a major challenge involves relating functional data to the physically observed RNA architecture. In the case of the hepatitis delta virus (HDV) ribozyme, there are three high-resolution crystal structures, the product state of the reaction and two precursor variants, with distinct mechanistic implications. Here, we develop new strategies to probe the structure and catalytic mechanism of a ribozyme. First, we use double-mutant cycles to distinguish differences in functional group proximity implicated by the crystal structures. Second, we use a corrected form of the Brønsted equation to assess the functional significance of general acid catalysis in the system. Our results delineate the functional relevance of atomic interactions inferred from structure, and suggest that the HDV ribozyme transition state resembles the cleavage product in the degree of proton transfer to the leaving group.



## INTRODUCTION

Natural ribozymes constitute a category of non-coding RNAs that mediate diverse biological functions including RNA splicing (group I and group II self-splicing introns),<sup>1–3</sup> replication of pathogenic RNA genomes including hammerhead- and hepatitis delta virus (HDV)-like ribozymes,<sup>4,5</sup> tRNA processing (RNase P),<sup>6</sup> gene expression (glmS),<sup>7,8</sup> and protein synthesis (ribosome).<sup>9</sup> With the exception of the ribosome, all known naturally occurring ribozymes catalyze phosphotransfer reactions of two distinct mechanistic categories: small ribozymes that use nucleobases and organic or metal ion cofactors to mediate endonucleolytic RNA cleavage, and large ribozymes that use multiple divalent metal ions to mediate nucleotidyl transfer.<sup>10–12</sup> High-resolution crystal structures now exist for all of the known small endonucleolytic ribozymes,<sup>8,13–21</sup> providing powerful foundations for defining catalytic functions in these systems. A grand challenge now lies in the integration of structural and functional data to understand how the RNA architecture facilitates the formation of specific interactions that stabilize the chemical transition state. Meeting this challenge entails defining the bonding changes between the ground state and the transition state, identifying and quantifying the energetic contributions of the

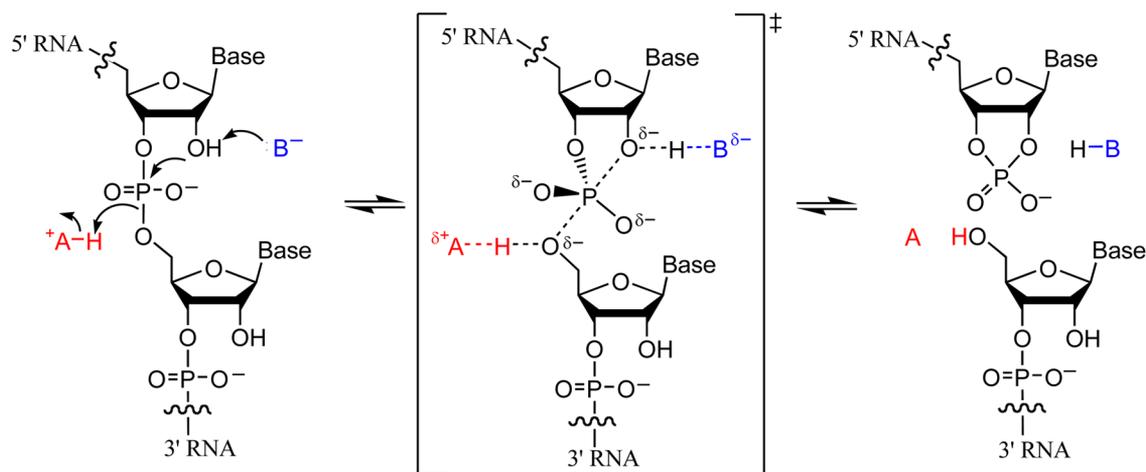
catalytic interactions, and establishing the functional relevance of the structural data.

The RNA genome of the hepatitis delta virus encodes two forms of the HDV ribozyme, genomic and antigenomic, which mediate essential steps in viral replication.<sup>22</sup> Both forms of the ribozyme adopt an overall similar secondary structure and likely employ similar catalytic mechanisms. Like the other small endonucleolytic ribozymes, the HDV ribozyme catalyzes RNA cleavage resulting in the formation of 2',3'-cyclic phosphate and 5'-hydroxyl termini (Figure 1). In vitro selection and bioinformatic analysis of genomic databases have revealed widespread distribution of HDV ribozyme-like motifs in organisms such as humans, unicellular ciliates, insects, fish, fungi, and plants.<sup>5,23</sup> The locations of these HDV-like motifs within specific genomes may implicate a variety of uncharacterized biological roles for these RNAs, some of which may include retrotransposition and RNA processing.<sup>24</sup>

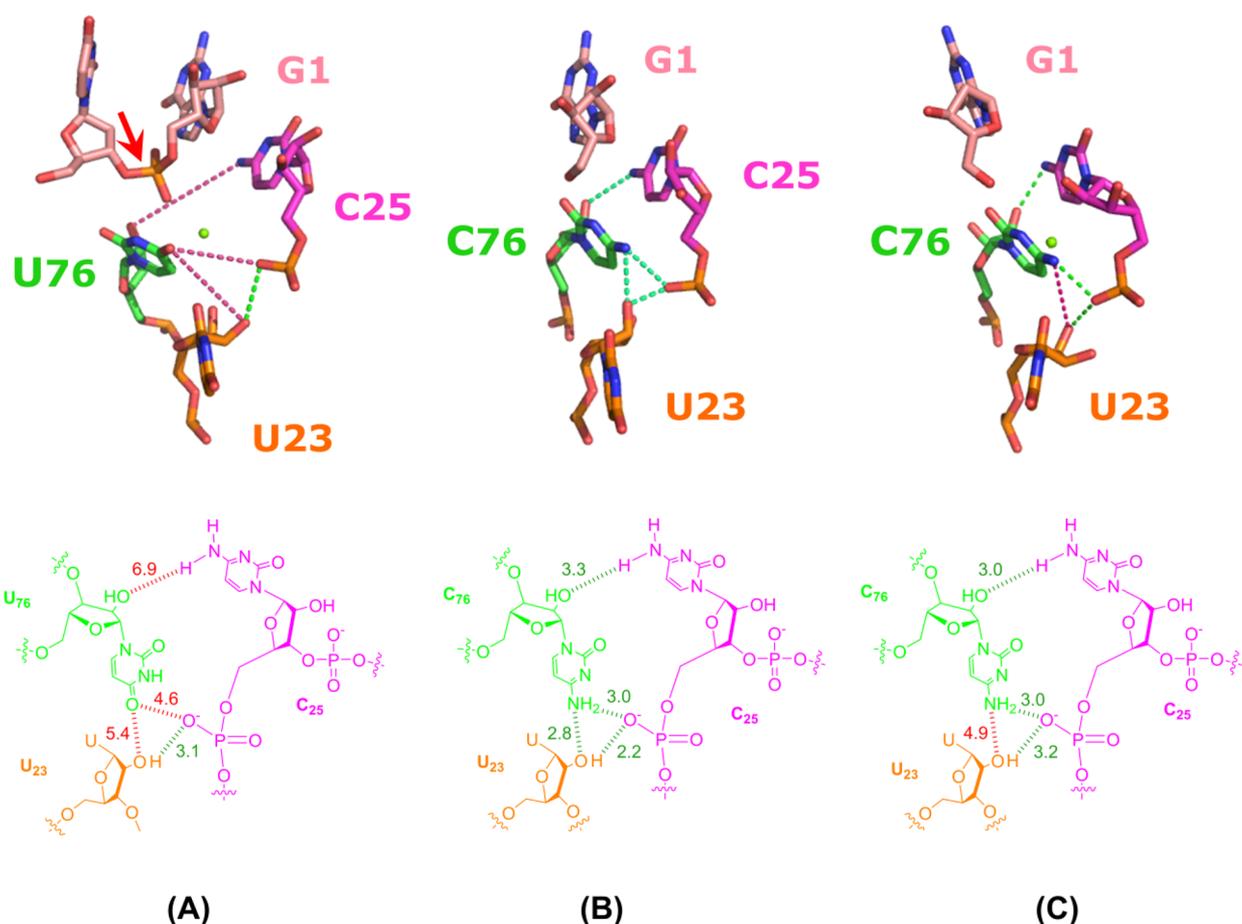
For the past 15 years, efforts to understand the structural basis of HDV ribozyme catalysis have relied on high-resolution crystal structures for two genomic HDV ribozyme constructs—

Received: February 3, 2015

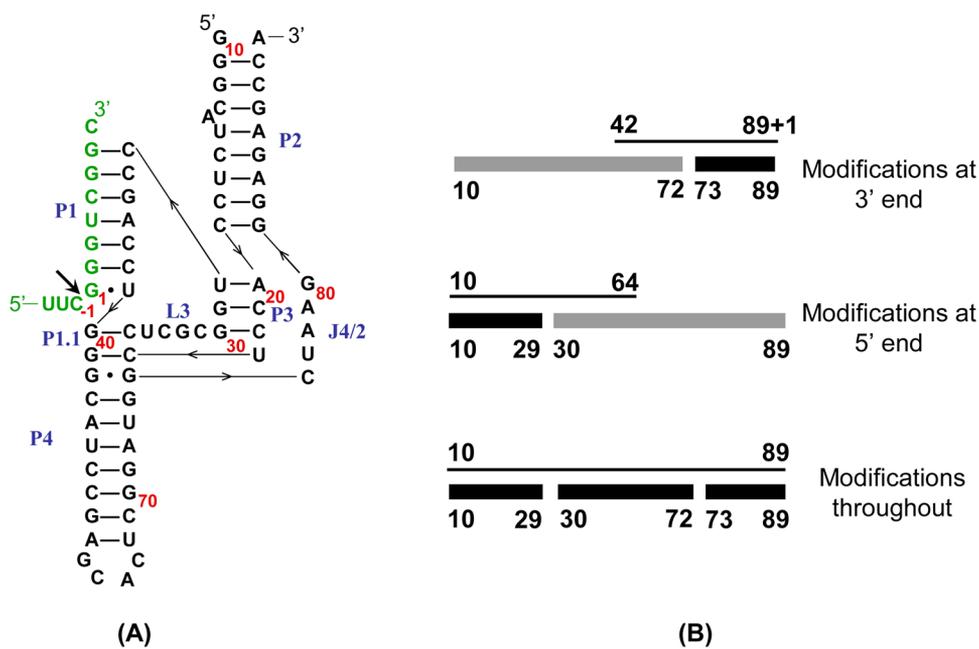
Published: June 30, 2015



**Figure 1.** RNA cleavage reaction catalyzed by the HDV ribozyme. Cleavage occurs with attack of a 2'-hydroxyl group on the adjacent phosphate to give 2',3'-cyclic phosphate and 5'-hydroxyl-containing products. A-H<sup>+</sup> represents a putative general acid, which protonates the 5'-leaving group; B<sup>-</sup> represents a general base, which activates the 2'-nucleophile. The scheme assumes that bond formation and bond cleavage occur in a single step.



**Figure 2.** Interactions with the active site nucleobase C76 differ in the product and precursor crystal structures. Nucleotide numbering is as in the antigenomic ribozyme. Green dashes connect atoms that reside within hydrogen-bonding distance ( $<3.5$  Å) of each other, and red dashes connect atoms that are not within hydrogen-bonding distance of each other ( $>3.5$  Å). Line drawings of the active-site interactions, including distances between groups in Å, are shown below each crystal structure. (A) In the U76 mutant precursor structure, the active-site U76 nucleobase interacts with a divalent metal ion (green sphere). (B) In the product structure, the N3 imino nitrogen of C76 is within hydrogen-bonding distance of the 2'-hydroxyl leaving group and engages in a network of hydrogen bonds. (C) In the ribozyme-1dC inhibitor pre-cleavage complex structure, the N3 imino nitrogen of C76 is within hydrogen-bonding distance of the 2' hydroxyl leaving group; in addition, C76 engages in a similar network of hydrogen bonds as in the product structure. There is a divalent metal ion (green sphere) in close proximity to C76. Nucleotide numbering is as in the antigenomic ribozyme. Green dashes connect atoms that reside within hydrogen-bonding distance ( $<3.5$  Å) of each other, and red dashes connect atoms that are not within hydrogen-bonding distance of each other ( $>3.5$  Å).



**Figure 3.** Construction of ribozymes bearing atomic mutations by DNA splint-mediated ligation. (A) Secondary structure of the *trans*-acting antigenomic ribozyme used in this work. The ribozyme contains a discontinuous J1/2, dividing the RNA into substrate and ribozyme portions, and a shortened P4. The substrate sequence is shown in green, and the ribozyme is in black; the cleavage site is indicated by a black arrow. Numbering maintains that of the *cis*-acting antigenomic ribozyme with a wild-type length P4 stem. (B) Components used to construct mutant ribozymes by splint-mediated ligation. Black rectangles represent synthetic RNA, gray rectangles represent transcribed RNA, and black lines represent complementary DNA splints. Numbering follows the secondary structure in A.

an inactive precursor construct with the catalytically critical C75 mutated to a uridine (Supporting Information, Figure S1B),<sup>15</sup> and the self-cleaved product (Figure S1C).<sup>13</sup> These two structures share an overall similar global fold but exhibit several differences in the active site, leading to two distinct mechanistic models for catalysis involving a catalytic nucleobase and a metal ion. Since C75 in the genomic ribozyme is equivalent to C76 in the antigenomic ribozyme, all subsequent nucleotide numbers refer to that of the antigenomic ribozyme, in which the current work was done (Figure S1A). In the C76U mutant precursor structure (Figure 2A and Figure S1B), a metal ion sits in the active site poised to coordinate the O4 keto group of U76 directly and the 5'-O leaving group indirectly via a metal-bound water molecule. Based on this structure, a simple bond rotation about the scissile phosphate would bring the nucleophilic 2'-hydroxyl group within hydrogen-bonding distance of the U76 imino nitrogen, leading the authors to propose that in the wild-type ribozyme, the C76 nucleobase acts as a general base that activates the nucleophile. In contrast, in the product structure (Figure 2B and Figure S1C) the C76 imino nitrogen is within hydrogen-bonding distance of the 5'-O, implicating the C76 nucleobase as a general acid that stabilizes the leaving group. No divalent metal ion was visualized in the active site (Figure 2B).

In 2010, a third crystal structure of the genomic HDV ribozyme provided additional insight into the structural basis of HDV ribozyme catalysis.<sup>18</sup> This structure is based on a fast-folding, *trans*-acting construct with the 2'-hydroxyl nucleophile removed by deoxynucleotide substitution, referred to here as the ribozyme:-1dC inhibitor pre-cleavage complex (Figure 2C and Figure S1D). In this structure, as in the product structure, the C76 imino nitrogen is positioned near the 5'-O leaving group (Figure 2C). In addition, there is a divalent metal ion proximal to C76 within the active site. A model of the electron

density around this metal ion places the scissile phosphate and a phosphate within the L3 loop as inner-sphere contacts and the G28 nucleobase as an outer-sphere contact.

Functional and spectroscopic data favor the mechanistic model implied by both the ribozyme:-1dC inhibitor pre-cleavage complex structure and the product structure, supporting the role of C76 as a general acid.<sup>25–29</sup> Despite key biochemical results favoring this mechanistic model, the crystal structures that serve as the basis for this model each lack critical elements of the reaction coordinate. The product structure lacks a metal ion, the scissile phosphate, and the nucleophile. The C76U precursor structure contains all three elements lacking in the product structure, but the crystallized construct has no catalytic activity in the absence of exogenous imidazole. The ribozyme:-1dC inhibitor pre-cleavage complex structure includes a cleavage site metal ion and scissile phosphate but lacks the nucleophile. Additionally, the electron density in the vicinity of the scissile phosphate is disordered, requiring model building based on the crystal structure of another ribozyme. Thus, limitations associated with the product and both precursor constructs make it difficult to unambiguously relate the reaction pathway to the physical structure of the RNA.

As a strategy to reduce the ambiguity in relating the reaction mechanism to the physical structure of the RNA, we describe here a series of new functional approaches that exploit principles of steric and electronic complementarity to test for spatial proximity between groups not directly involved in catalysis. We also describe a strategy that combines synthesis of ribozymes containing site-specific nucleotide analogue substitutions with use of a hyperactivated substrate to assess the extent of nucleobase-mediated proton transfer. The results show that the catalytic C76 residue in the HDV ribozyme transition state mirrors the “product-like” mechanistic model

with respect to both spatial position and degree of proton transfer to the leaving group.

## EXPERIMENTAL SECTION

**Chemical Synthesis.** The 5'-S substrate ( $S_{5'-S}$ ; wild-type substrate will be referred to as  $S_{5'-O}$ ) was synthesized by solid-phase synthesis using 2'-O-TBS-3'-phosphoramidites, 2'-O-*o*-nitrobenzyl-3'-cytidine phosphoramidite<sup>30</sup> and 5'-S-DMTr-3'-guanosine phosphoramidite as described previously;<sup>27</sup> the substrate used here (UU-C<sub>2'-OonBn</sub>G<sub>5'-S</sub>GGUCGGC, where OonBn represents the photolabile protecting group 2'-O-*o*-nitrobenzyl) contains a guanosine rather than a 2'-deoxyguanosine immediately 3' of the 5'-sulfur linkage. MALDI-TOF mass spectrometry confirmed the structure of the 5'-S RNA substrate (calcd for MH<sup>+</sup>, 3647.5; found, 3647.3). Oligonucleotides containing modified bases or sugars were either purchased from Dharmacon or synthesized from the corresponding phosphoramidites; the synthesis of several of these phosphoramidites has been described.<sup>31</sup>

**Ribozyme Preparation.** The HDV ribozyme used is a well-characterized *trans*-acting antigenomic ribozyme with a discontinuous J1/2 and a shortened P4 helix (Figure 3A),<sup>32</sup> which facilitates analysis by allowing the use of modified substrates as separate entities. The wild-type HDV ribozyme and C25U:G40A ribozymes were synthesized by T7 RNA polymerase transcription from a single-stranded DNA template (IDT). All ribozymes containing site-specific modifications were synthesized via DNA splint-mediated ligation from oligonucleotides made by chemical synthesis or by T7 RNA polymerase transcription as previously described (Figure 3B).<sup>27,33</sup> Without changing the ribozyme sequence, ligation junctions were selected to allow for ligation with transcribed RNAs initiating with an efficient GG sequence.<sup>34</sup> Ribozymes containing modifications in the 5' portion were synthesized from GGGCAUCUCCACCUCCUCGC (underlining indicates position 25) ligated using T4 DNA ligase in the presence of a DNA splint to the 3' sequence transcribed in the presence of guanosine 5'-monophosphate. Ribozymes containing modifications at both positions 25 and 76 were synthesized from one-step DNA splint-mediated ligation of three synthetic oligonucleotides: GGGCAUCUCCACCUCCUCGC (20 nucleotides, C25 is underlined), GGUCGACCUGGGCAUCCGAGCACUCGGGA (29 nucleotides), and UGGCUAAGGGAGAGCCA (17 nucleotides; C76 is in bold). The oligonucleotides used to generate modified ribozymes are summarized in Table S2. The ligated wild-type ribozyme compared favorably to the transcribed ribozyme in both purity and activity (Figure S2). Because modified ribozymes could not be directly made by transcription, the identity of the isolated ligation product was verified by running the purified product on a denaturing gel and comparing with the size of the wild-type transcribed ribozyme; in all cases only one product band, of the same size as the wild-type transcribed ribozyme, was observed.

**Ribozyme Reactions and Kinetic Analysis.** Ribozyme-catalyzed substrate cleavage reactions were conducted as previously described.<sup>27</sup> Buffers used contained 25 mM acetic acid, 25 mM MES, and 50 mM Tris (pH 4.0–8.0), or 50 mM MES, 25 mM Tris, and 25 mM 2-amino-2-methyl-1-propanol (pH 7.5–10.0). Ribozyme (final concentration 1  $\mu$ M) was pre-incubated with 10 mM MgCl<sub>2</sub> at 70 °C for 2 min, then 25 °C for 14 min; buffer was added, and reaction was initiated by the addition of a trace amount of 5'-radiolabeled substrate. The 5'-S substrate was prepared for use by irradiation with a 100-W UV lamp (365 nm) for 4 min immediately prior to addition to the reaction mixture. Reaction aliquots were removed at appropriate times and quenched by the addition of a 90% formamide, 10 mM EDTA solution and placed on ice. Cleaved products were separated from uncleaved substrate on a denaturing 20% polyacrylamide/7 M urea gel and quantified using a Typhoon Trio PhosphorImager and ImageQuant TL software (Molecular Dynamics). Data for cleavage reactions were fitted to the equation:  $y = 1 - y_0 - A e^{-k_{\text{cleavage}}t}$  using Origin 7.0 (OriginLab), where  $y$  is the product fraction and  $k_{\text{cleavage}}$  is the first-order rate constant for substrate cleavage.  $y_0$  and  $A$  are constant parameters that are used to fit the equation; at the reaction end-point,

$y_0$  corresponds to the fraction of unreacted substrate and  $A$  corresponds to the fraction of reacted substrate. In all cases where the reaction reached an end-point,  $y_0$  was less than 30%. For reactions with slow cleavage rates ( $<1 \times 10^{-4} \text{ min}^{-1}$ ), data were fitted to a single exponential assuming a maximum cleavage end-point of 80–90%. To calculate reaction  $pK_a$ 's from reaction rates obtained at different pH values, rate constants were fitted to the following equation:  $k_{\text{cleavage}} = k_{\text{max}}/(1 + 10^{\text{pH}-\text{p}K_a})$ .

$\Delta\Delta G_{Z \rightarrow Z1}^\ddagger$ , the difference in transition state energies for a ribozyme–substrate complex containing functional group  $Z$  and one replacing  $Z$  with  $Z1$ , was calculated as  $\Delta\Delta G_{Z \rightarrow Z1}^\ddagger = -RT \ln(k_{Z1}/k_Z)$ , where  $R$  is the universal gas constant,  $T$  is the reaction temperature in kelvin,  $k_Z$  is the rate of substrate cleavage by ribozyme–substrate complex containing functional group  $Z$ , and  $k_{Z1}$  is the rate of substrate cleavage by ribozyme–substrate complex containing functional group  $Z1$ .

**Derivation of a Modified Brønsted Equation To Estimate the Extent of Proton Transfer in the Transition State of the HDV Ribozyme-Catalyzed Reaction.** A Brønsted relationship for general acid catalysis reflects a linear free energy relationship between the reaction rate constant ( $k_{\text{obs}}$ ) and the  $pK_a$  values for a series of acids and is given by the following equation:

$$\log(k_{\text{obs}}) = -\alpha pK_{a,\text{HA}^+} + \log(G_A) \quad (1)$$

In our study, we treat the wild-type (wt) and mutant (mut) HDV ribozymes as a series of general acids.

$$\log(k_{\text{obs}}^{\text{mut}}) = -\alpha pK_{a,\text{HA}^+} + \log(G_A') \quad (2)$$

However, mutations within an enzyme likely would cause a Brønsted relationship to deviate from linearity due to the spatial and electrostatic constraints imposed by the active site. Thus, estimating the effects of a mutation specific to general acid catalysis for use in a Brønsted analysis requires that the rate constants of the mutant be corrected by a factor  $\gamma$  that accounts for the active site constraints.

$$(k_{\text{obs}}^{\text{mut}})_{\text{cor}} = k_{\text{obs}}^{\text{mut}} \gamma \quad (3)$$

Since general acid catalysis does not facilitate cleavage of the 5'-S substrate, we hypothesize that rate of cleavage of 5'-S substrate by the wild-type ribozyme relative to the rate of cleavage by a mutant ribozyme provides an operational measure of  $\gamma$  according to eq 4,

$$\gamma = \frac{(k_{\text{obs}}^{\text{wt}})_{5'-S}}{(k_{\text{obs}}^{\text{mut}})_{5'-S}} \quad (4)$$

$$(k_{\text{obs}}^{\text{mut}})_{\text{cor}} = k_{\text{obs}}^{\text{mut}} \gamma = (k_{\text{obs}}^{\text{mut}})_{5'-O} \frac{(k_{\text{obs}}^{\text{wt}})_{5'-S}}{(k_{\text{obs}}^{\text{mut}})_{5'-S}} \quad (5)$$

Substituting eq 5 into eq 2 gives

$$\begin{aligned} \log(k_{\text{obs}}^{\text{mut}})_{\text{cor}} &= \log\left(\frac{(k_{\text{obs}}^{\text{wt}})_{5'-S}}{(k_{\text{obs}}^{\text{mut}})_{5'-S}} (k_{\text{obs}}^{\text{mut}})_{5'-O}\right) \\ &= -\alpha pK_{a,\text{HA}^+}^{\text{mut}} + \log(G_A') \end{aligned} \quad (6)$$

For the wild-type ribozyme,

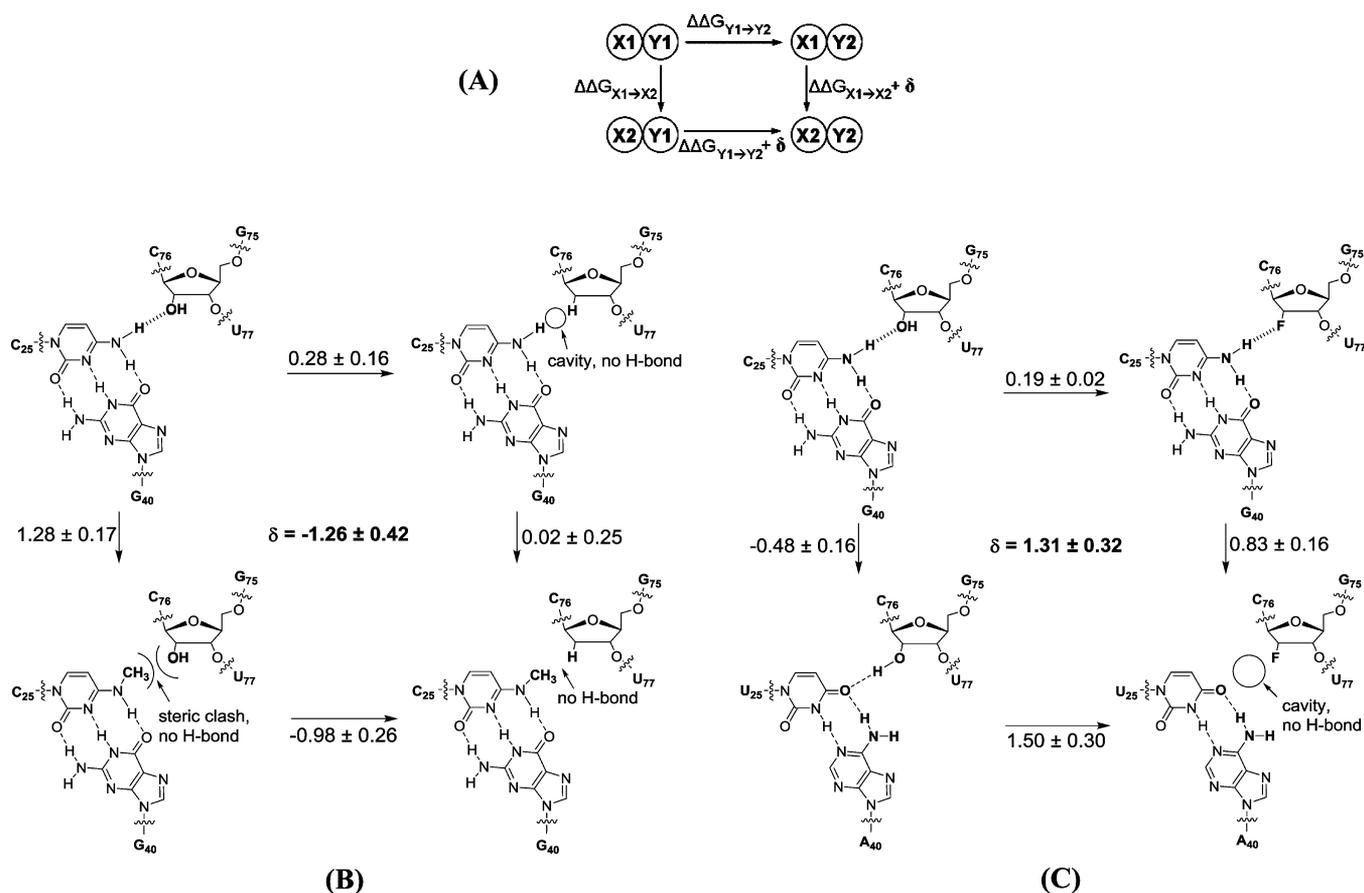
$$\log(k_{\text{obs}}^{\text{wt}})_{5'-O} = -\alpha pK_{a,\text{HA}^+}^{\text{wt}} + \log(G_A') \quad (7)$$

$$\log(G_A') = \log(k_{\text{obs}}^{\text{wt}})_{5'-O} + \alpha pK_{a,\text{HA}^+}^{\text{wt}} \quad (8)$$

Substituting eq 8 into eq 6 gives

$$\log\left(\frac{(k_{\text{obs}}^{\text{wt}})_{5'-S}}{(k_{\text{obs}}^{\text{mut}})_{5'-S}} k_{\text{obs}}^{\text{mut}}\right) = -\alpha pK_{a,\text{HA}^+}^{\text{mut}} + \log(k_{\text{obs}}^{\text{wt}})_{5'-O} + \alpha pK_{a,\text{HA}^+}^{\text{wt}} \quad (9)$$

Rearranging terms gives



**Figure 4.** Atomic mutation cycle analysis. (A) Setting up an atomic mutation cycle. Each partner of the interaction is first mutated separately (single mutants) and then together (double mutants). A mutation at one site, X (i.e., X1 → X2), that disrupts the interaction would be expected to attenuate the energetic effects of mutation at the second site, Y (i.e., Y1 → Y2; relative to the mutation in the wild-type ribozyme), because the interaction has already been disrupted by the first mutation. In essence, the second mutation “suppresses” the energetic effects of the first mutation. The second mutation is made both at the putative interacting partner and elsewhere, to distinguish suppression effects specific to the partners from more general suppression effects. The non-additive effect of combining two mutations on the catalytic activity of the HDV ribozyme is quantified in terms of the free energy term  $\delta$ , calculated as follows:  $\delta = \Delta G_{(X1Y1 \rightarrow X2Y2)}^{\ddagger} - \Delta G_{(X1 \rightarrow X2)}^{\ddagger} - \Delta G_{(Y1 \rightarrow Y2)}^{\ddagger} + \Delta G_{\text{wild-type}}^{\ddagger}$ . Positive values of  $\delta$  suggest that site X and site Y interact synergistically. Mutations at non-interacting sites are expected to yield additive effects, i.e.,  $\delta = 0$ . Non-additive effects (negative  $\delta$ ) may implicate cooperative interaction in cases where the mutations act in a mutually repulsive manner. Alternatively, negative  $\delta$  values could emerge when two sites do not interact, if the double mutant undergoes local rearrangements that form favorable, non-native interactions that interfere with catalysis. (B) Atomic mutation cycle for establishing steric proximity of C25 N4 exocyclic amine and C76 2'-hydroxyl group. Based on the terminology established for part A, X1 = C25 N4 NH<sub>2</sub>; X2 = C25 N4 methyl; Y1 = C76 2'-OH; Y2 = C76 2'-H. (C) Atomic mutation cycle for establishing electronic proximity of C25 N4 exocyclic amine and C76 2'-hydroxyl group. Based on the terminology established for part A, X1 = C25 N4 NH<sub>2</sub>; X2 = C25 O4 (i.e., C25U:G40A); Y1 = C76 2'-OH; Y2 = C76 2'-F.

$$\log(k_{\text{obs}}^{\text{mut}})_{S'-O} - \log((k_{\text{obs}}^{\text{mut}})_{S'-S}) + \log((k_{\text{obs}}^{\text{wt}})_{S'-S}) = -\alpha(\text{p}K_{\text{a,HA}^+}^{\text{mut}} - \text{p}K_{\text{a,HA}^+}^{\text{wt}}) + \log(k_{\text{obs}}^{\text{wt}})_{S'-O} \quad (10)$$

Grouping terms gives

$$\log\left(\frac{(k_{\text{obs}}^{\text{mut}})_{S'-O}}{(k_{\text{obs}}^{\text{mut}})_{S'-S}}\right) - \log\left(\frac{(k_{\text{obs}}^{\text{wt}})_{S'-O}}{(k_{\text{obs}}^{\text{wt}})_{S'-S}}\right) = \alpha(\text{p}K_{\text{a,HA}^+}^{\text{wt}} - \text{p}K_{\text{a,HA}^+}^{\text{mut}}) \quad (11)$$

Thus, a plot of  $\log(k_{\text{obs}}^{\text{mut}}/(k_{\text{obs}}^{\text{mut}})_{S'-S}) - \log((k_{\text{obs}}^{\text{wt}})_{S'-O}/(k_{\text{obs}}^{\text{wt}})_{S'-S})$  vs  $(\text{p}K_{\text{a,HA}^+}^{\text{mut}} - \text{p}K_{\text{a,HA}^+}^{\text{wt}})$  should give a line with slope  $\alpha$ . An assumption in this formulation is that the relationship between  $\text{p}K_{\text{a}}$  and general acid catalysis is not coupled to the correction term,  $\gamma$ . This assumption likely represents an oversimplification, considering that the HDV ribozyme has evolved to function in the context of a cytidine nucleobase as the general acid. Nevertheless, we expect that application of the correction term renders the data more meaningful than they would be otherwise.

Relative rate or binding constants have been used previously to investigate ribozyme interactions. For example, in studies of the group

I intron,<sup>35,36</sup> removal of the 2'-hydroxyl group from the 3-position of the substrate or product causes a significant reduction in catalysis or binding, respectively. To identify the residue in the ribozyme that makes this interaction, mutations were tested with both ribo- and 3-deoxynucleotide substrates. The 3-deoxynucleotide substrate/product lacks the 2'-hydroxyl group with which to make the interaction, so it is used as a correction factor to assess the effects of a mutation independent of loss of the interaction with the 2'-OH group. Thus, a residue that upon mutation has a greater deleterious effect on the ribo substrate relative to the deoxyribo substrate would be a viable candidate for the interacting nucleotide. Our use of the  $\gamma$  correction factor parallels this analysis. In our study, we are testing an interaction in the transition state of the HDV ribozyme. We use a modified substrate that gains no catalytic advantage from the interaction (the 5'-S substrate). In the same manner as the deoxynucleotide substrate, the 5'-S substrate is used as a measure of the effects of the mutation independent of the interaction being investigated.

## RESULTS

**Atomic Mutation Cycles Establish Spatial Proximity between C25 N4 and C76 2'-Hydroxyl Group.** Atomic mutation cycles (or double-mutant cycles) assess the energetic coupling between two positions within a macromolecule by delineating the energetic cost of a mutation at one position by itself and in the context of a mutation at the second position (Figure 4A). The approach provides a biochemical signature to distinguish functional groups that can “sense” one another from those that cannot, allowing for functional tests of atomic proximity implicated by crystal structures.<sup>37–40</sup> We compared the crystal structures of the genomic product (PDB: 1CX0, 2.3 Å resolution), ribozyme:-1dC inhibitor pre-cleavage complex (PDB: 3NKB, 1.9 Å resolution), and C76U precursor (PDB: 1SJ3, 2.2 Å resolution) and observed significant differences in the distances between two functional groups near the active site—the 2'-hydroxyl group of C76 and the exocyclic N4 amino group of C25 (Figure 2). These groups reside within hydrogen-bonding distance in the crystal structures of the ribozyme in product form (3.3 Å) and the ribozyme:-1dC inhibitor pre-cleavage complex (3.0 Å) but not in the C76U precursor crystal structure (6.9 Å). In the C76U precursor structure, these two functional groups have different nearest neighbors. The C25 N4 amino group lies within 3.5 Å of the exocyclic amino group of C24 due to stacking between C24 and C25 in the P1.1 helix; the 2'-hydroxyl group of U76 sits within 3 Å of both non-bridging oxygens of U77.

To probe the functional relevance of the proximity between the exocyclic N4 amino group of C25 and the 2'-hydroxyl group of C76 observed in two of the structures, we created atomic mutants of each and analyzed them with double-mutant cycles (Figure 4A). Specifically, we asked whether a deleterious effect caused by a steric or electronic alteration of the functional group at one position could be suppressed by a complementary steric or electronic alteration of the functional group at the other position. We conducted assays in the log-linear region of the pH-rate profile<sup>27,41,42</sup> (pH 4.9).

We first tested the functional importance of the C76 2'-hydroxyl group by incorporating 2'-deoxycytidine or 2'-fluorocytidine into position 76 of the ribozyme. Neither of these modifications affected the reaction rate significantly (at pH 4.9, 25 °C, under single-turnover conditions) (Table 1), suggesting this 2'-hydroxyl group contributes minimally to the

overall HDV ribozyme reaction. The absence of a significant functional contribution from the C76 2'-hydroxyl group precludes the use of strategies that probe linkage to the exocyclic N4 amino group of C25 by swapping donor-acceptor properties of putative hydrogen-bonding partners. Indeed, the C25U:G40A double mutant, which effectively replaces the C25 N4 amino group with a U25 O4 keto group while maintaining Watson-Crick complementarity with its P1.1 pairing partner (residue 40), reacts with near-wild-type activity (Table 1), further suggesting that a C76 2'-OH:C25 N4 amino interaction contributes little to HDV ribozyme catalysis. Consequently, we designed alternative strategies to probe proximity between the two functional groups. Namely, we installed mutations that induce steric or electronic interference at one position and then tested whether the compensatory mutation of the putative neighbor suppressed the interference.

To introduce steric interference (Figure 4B, Figure S4), we installed N4-methylcytidine at residue 25, replacing a hydrogen atom with a more bulky hydrophobic methyl group. The C25N4-methyl ribozyme variant catalyzes substrate cleavage from the enzyme-substrate (ES) complex 9-fold less efficiently than does the wild-type ribozyme, corresponding to an energetic cost,  $\Delta\Delta G^\ddagger = +1.3 \pm 0.2$  kcal/mol (Table 1). Presumably the microenvironment surrounding the N4 amino group lacks sufficient space to accommodate the methyl group as the enzyme-substrate complex traverses the reaction coordinate landscape, consistent with the product and ribozyme:-1dC inhibitor pre-cleavage complex structures. In contrast, the C76U precursor structure suggests that little or no steric clash would occur. Strikingly, a 4-NMe-C25/dC76 double mutant reacted faster than the 4-NMe single mutant, meaning that in the context of the C76dC ribozyme, the introduced N4-methyl group at C25 has only a minimal effect on catalysis ( $\Delta\Delta G^\ddagger = +0.02 \pm 0.25$  kcal/mol). These results suggest that the absence of the 2'-hydroxyl group at C76 allows the ribozyme to accommodate the methyl group better than does the wild-type ribozyme.

The accommodation of the methyl group on C25 by removal of the 2'-hydroxyl group at C76 can also be assessed by constructing an atomic mutation cycle (AMC; Figure 4A,B); the calculated negative value for the cooperativity term  $\delta$  ( $-1.3 \pm 0.4$  kcal/mol) implies that removing the C76 2'-hydroxyl group suppresses the energetic cost of introducing the methyl group on C25. In contrast, the C25 N4-methyl mutation was not suppressed by 2'-deoxy mutations at U77 in the ribozyme (data not included), suggesting that the observed suppression occurs specifically. Removal of the 2'-hydroxyl group from C76 could eliminate potential van der Waals overlap with the methyl group and the unfavorable cost of juxtaposing a polar hydroxyl group with a non-polar methyl group, or engender a favorable hydrophobic interaction between the C25 N4-methyl and the 2'-methylene group of the dC76 residue. Attenuation of the effect from methyl installation upon removal of the 2'-hydroxyl group implies that the two functional groups reside close to one another in the transition state for the reaction. These findings support the functional relevance of the ribozyme product and ribozyme:-1dC inhibitor pre-cleavage complex.

As a second strategy to probe 4-NC25/2'-OHC76 proximity, we altered the electronic character of the same functional groups to induce electrostatic repulsion (Figure 4C). The physical basis for this experimental design comes from NAIM experiments in the context of the  $\Delta$ C209 P4-P6 independently folding RNA domain and functional experiments in the context

**Table 1. Reaction Rates for Mutant Ribozymes Used in Atomic Mutation Cycle Analysis<sup>a</sup>**

ribozyme	$k$ (min <sup>-1</sup> ) <sup>b</sup>	$\Delta\Delta G_{wt \rightarrow mut}^\ddagger$ (kcal/mol) <sup>c</sup>
wt	0.055 ± 0.009	N/A
4-NMe-C25	0.0063 ± 0.0008	1.28 ± 0.17
dC76	0.034 ± 0.004	0.28 ± 0.16
4-NMe-C25/dC76	0.033 ± 0.002	0.30 ± 0.09
C25U:G40A	0.12 ± 0.02	-0.48 ± 0.16
2'-F-C76	0.040 ± 0.005	0.19 ± 0.02
4-NMe-C25/2'-F-C76	0.00205 ± 0.00008	1.94 ± 0.12
C25U:G40A/dC76	0.097 ± 0.005	-0.34 ± 0.09
C25U:G40A/2'-F-C76	0.0098 ± 0.0008	1.02 ± 0.14

<sup>a</sup>Included rates were measured at pH 4.9, within the log-linear region of the pH-rate profiles for all ribozymes tested. Data are the average of three independent measurements. <sup>b</sup>Cleavage of wild-type substrate at pH 4.9, 10 mM MgCl<sub>2</sub>, 25 °C. <sup>c</sup> $\Delta\Delta G_{wt \rightarrow mut}^\ddagger$  is defined as in the Experimental Section.

of the group I ribozyme.<sup>43</sup> Those analyses revealed that 2'-F substitution caused deleterious effects when the 2'-OH group resides within hydrogen-bonding distance of an oxygen atom, but not when the 2'-OH groups resides within hydrogen-bonding distance of a nucleobase exocyclic amine.

To explore the proximity between the C25 and C76 functional groups, we examined the effect of introduction of 2'-fluorocytidine at residue 76 in the context of the wild-type and C25U:G40A mutant ribozyme. Again, the C25U:G40A double mutant, which reacts with near-wild-type activity (Table 1),<sup>44</sup> effectively replaces the C25 N4 amino group with a U25 O4 keto group while maintaining Watson–Crick complementarity with its P1.1 pairing partner. In this P1.1-modified ribozyme, introduction of 2'-fluorocytidine at residue 76 causes the enzyme–substrate (ES) complex to react 12-fold less efficiently than the wild-type ribozyme (Table 1), corresponding to an energetic cost,  $\Delta\Delta G^\ddagger = +1.5 \pm 0.3$  kcal/mol. In contrast, in the wild-type ribozyme, introduction of 2'-fluorocytidine at residue 76 has minimal effect on catalysis ( $\Delta\Delta G^\ddagger = +0.19 \pm 0.02$  kcal/mol) (Table 1). These results suggest that replacing the O4 keto group at residue 25 with an amino group suppresses the deleterious effect of a 2'-fluorine atom of residue 76, consistent with the NAIM signature that 2'-fluorine atoms may reside close to amino groups without deleterious consequences. The significant effect from fluorine substitution in the C25U:G40A ribozyme, but not in the wild-type ribozyme corroborates the findings from the steric interference suppression analysis described above that C76 2'-hydroxyl and C25 N4 lie in close proximity in the transition state. Considering these data in the context of an atomic mutation cycle, the positive value for  $\delta$  ( $+1.3 \pm 0.3$  kcal/mol) suggests that reversion of C25U:G40A to C25:G40 provides better accommodation of a 2'-F atom at C76.

As controls, we also constructed double-mutant ribozymes containing one each of a space-filling mutant (dC76, N4-methyl-C25) and an electron-density-increasing mutant (2'-F-C76, C25U:G40A) and calculated the resulting  $\delta$  values (Figure 4A). For these double mutants (C25 N4-methyl/2'-F-C76, C25U:G40A/dC76), the  $\delta$  values obtained have a magnitude less than 0.5 kcal/mol (C25 N4-methyl/2'-F-C76,  $\delta = +0.5 \pm 0.3$  kcal/mol; C25U:G40A/dC76,  $\delta = -0.1 \pm 0.4$  kcal/mol), which supports the specificity of the observed effects probed by our site-specific mutants.

To validate the approach further, we also constructed single- and double-mutant ribozymes that perturb a hydrogen bond found in all three crystal structures—the 2'-hydroxyl group of U23 and the *pro-R<sub>p</sub>* oxygen of C25 (Figure 2). We constructed an atomic mutation cycle with the following modifications: 2'-deoxy-U23 and/or a *pro-R<sub>p</sub>*-non-bridging phosphorothioate at C25 (Table S3 and Figure S5). Taken together the kinetic data revealed a  $\delta$  value of  $-1.2$  kcal/mol. The relatively large negative  $\delta$  value suggests that removal of the 2'-hydroxyl group from U23 at least partially suppresses the deleterious effect from a bulky phosphorothioate substitution at C25. This observation is consistent with an interaction between the 2'-hydroxyl group of U23 and the *pro-R<sub>p</sub>* oxygen of C25 in the transition state of the ribozyme-catalyzed cleavage reaction. Thus, these findings support the functional relevance of the interaction found all three structural forms of the ribozyme.

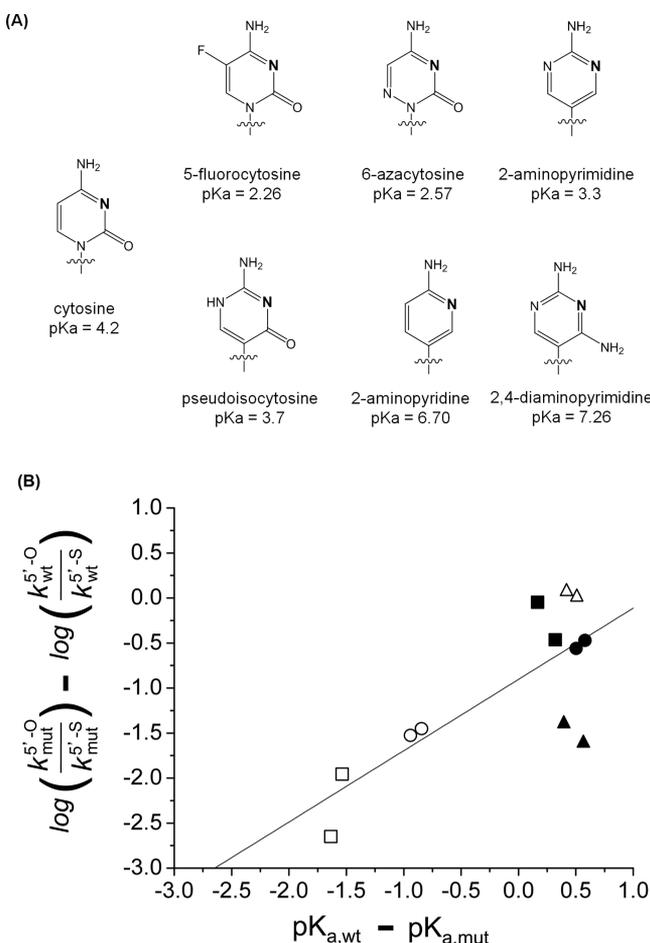
**Brønsted Analysis Implicates a Late Transition State for the HDV Ribozyme-Catalyzed Reaction.** In addition to defining the architectural features of the transition state, we investigated the general acid function of C76 using Brønsted

analysis. Brønsted analyses monitor the reaction rate versus  $\text{p}K_{\text{a}}$  of the general acid or base and provide information about the degree of proton transfer in the transition state. In practice,  $\text{p}K_{\text{a}}$  variations are obtained using a series of analogues that, ideally, have minimal structural variation and impart  $\text{p}K_{\text{a}}$  changes electronically. However, in the context of enzyme and ribozyme active sites, the inability to account for effects independent of proton transfer frequently confounds such analyses.<sup>45</sup>

Previously, we showed that a hyperactivated substrate bearing a 5'-sulfur leaving group at the scissile site ( $S_{5'S}$ ) reacts with minimal dependence on the ribozyme's capacity for general acid catalysis, but retains sensitivity to other features of catalytic mechanism,<sup>27</sup> including inhibition by cobalt hexamine and pH–rate dependence at reaction pH less than 7. Thus, the reaction rate constant of the wild-type substrate,  $S_{5'O}$ , relative to the 5'-sulfur containing substrate,  $S_{5'S}$ , allows the steric effects of analogue substitution at C76 to be estimated. These effects can thus provide a more sensitive measure of the electronic effects on proton transfer than analysis of the effects of analogue substitution on the reaction of  $S_{5'O}$  alone.

Accordingly, we synthesized a series of cytosine analogues that span a range of solution  $\text{p}K_{\text{a}}$ 's (Figure 5A) and installed them site-specifically at C76.<sup>30</sup> To accomplish this, individual cytosine analogue phosphoramidite derivatives were incorporated into oligonucleotides by solid-phase synthesis. These oligonucleotides were subsequently incorporated into the HDV ribozyme via DNA splint-mediated ligation, yielding ribozymes bearing cytosine analogues with altered  $\text{p}K_{\text{a}}$  at C76. We determined the pH–rate profiles for the mutant  $ES_{5'O}$  complexes to obtain apparent  $\text{p}K_{\text{a}}$  values corresponding to titration of the general acid.<sup>27,28</sup> To conduct the Brønsted analysis, we plotted the  $\text{p}K_{\text{a}}$  values for the six different atomically perturbed ribozymes and the wild-type ribozyme against the reaction rate of  $ES_{5'O}$  relative to that for  $ES_{5'S}$ . The data give a linear correlation ( $n = 12$ ,  $R = 0.77$ ) with slope  $\alpha = 0.8 \pm 0.2$  (Figure 5B). The apparent Brønsted coefficient ( $\alpha_{\text{app}}$ ) could provide a rough measure of the change in effective charge on the general acid en route to the transition state; thus, the effective charge on C76 decreases from +1 in the ground state to +0.2 in the transition state, implying a significant degree of proton transfer to the leaving group. In contrast, Brønsted analysis of the  $S_{5'O}$  reaction without the correction does not show an apparent correlation between reaction rate and  $\text{p}K_{\text{a}}$  (Figure S6).

Several factors could obscure  $\alpha_{\text{app}}$  as a true measure of the effective charge change that C76 experiences as the reaction goes from the ground state to the transition state.<sup>46</sup> First, Brønsted analyses generally attribute changes in reactivity directly to the changes in  $\text{p}K_{\text{a}}$  of the functional group undergoing proton transfer, thereby assuming that the modifications installed to perturb the  $\text{p}K_{\text{a}}$ 's have no indirect effects on the reaction. In the HDV ribozyme active site, the modifications used to achieve  $\text{p}K_{\text{a}}$  variation can influence the reaction in ways other than  $\text{p}K_{\text{a}}$  shifts. We used the  $S_{5'S}$  substrate to account for these indirect effects, but as described in the Experimental Section, this correction assumes these effects to be separable from the contribution of general acid catalysis eq 5. In other words, the assumption implies that if a hypothetical series of  $\text{p}K_{\text{a}}$ -varied ribozymes could be constructed in the context of each C76 analogue, all would give the same  $\alpha_{\text{app}}$  value. Although we cannot test this assumption directly, we found no evidence that  $\alpha_{\text{app}}$  reflects changes in molecular volumes when we plotted reaction rates of  $S_{5'O}$ ,  $S_{5'S}$ ,



**Figure 5.** Brønsted analysis to assess the degree of proton transfer in the transition state of the HDV ribozyme-catalyzed reaction. (A) Cytidine analogues used for Brønsted analysis. Solution  $pK_a$ 's for N3 (indicated in bold) of each analogue are included.<sup>57–63</sup> (B) Brønsted analysis of the dependence of reaction rate on the acidity of the N3 imino nitrogen of C76. The  $x$ -axis represents the reaction  $pK_a$  for each mutant ribozyme, and the  $y$ -axis represents proton-transfer-specific electronic effects of each mutation, as described in the Experimental Section. Data for each point on the plot are taken from Table 2. Reaction rates and  $pK_a$  values obtained for the same ribozyme in different experiments differed by less than 30%. Each modification is represented on the figure as follows: (●) 5-fluorocytosine, (■) 6-azacytosine, (▲) 2-aminopyrimidine, (○) 2-aminopyridine, (□) 2,4-diaminopyrimidine, and (△) pseudoisocytosine. The error-weighted best fit to the data is shown; the equation of the best-fit line is  $y = (0.79 \pm 0.21)x - (0.90 \pm 0.17)$  ( $R = 0.77$ ), where the slope of the line is equal to  $\alpha_{app}$  and the included error is the error of the fit.

and the corresponding ratios against molecular volumes of the C76 analogues (calculated using solvent-excluded molecular volume and with group contributions; Table S4 and Figures S7–S9). Second, if the C76H<sup>+</sup> ribozyme populates a non-productive conformation in a manner that depends on the nucleobase  $pK_a$ , the Brønsted coefficient for the conformational equilibrium will affect the value of  $\alpha_{app}$ . Third, catalytic interactions involving the C76H<sup>+</sup> exocyclic substituents will influence  $\alpha_{app}$  if their catalytic contribution varies with nucleobase  $pK_a$ . Correction of the  $S_{5'-O}$  reaction rates using  $S_{5'-S}$  minimizes the potential influence of the latter two factors, but we note that the stated factors may affect the reactivity of the two substrates unequally. In the future, it will be important to benchmark our result against independent measures of the degree of proton transfer (e.g., kinetic isotope effects on the leaving group or imino group of C76H<sup>+</sup>). Despite these caveats, the data suggest an increase in reaction rate with increasing nucleobase acidity, consistent with a significant catalytic contribution from proton transfer by C76H<sup>+</sup> to the leaving group. Moreover, the value of  $\alpha_{app}$  provides a biochemical signature that may be used in future studies to determine how structural features of the HDV ribozyme influence general acid catalysis (i.e., by measuring  $\alpha_{app}$  in the context of other mutations).

## DISCUSSION

Here, we used two distinct approaches to study the transition state of the HDV ribozyme-catalyzed cleavage reaction: atomic mutation cycle analysis to identify proximal functional groups in the transition state, and Brønsted analysis to infer the degree of proton transfer from the general acid to the 5' leaving group. Both lines of inquiry suggest the ribozyme:product complex and the ribozyme:-1dC inhibitor pre-cleavage complex crystal structures as functionally relevant to the transition state and reinforce the role of C76 as the general acid in the HDV ribozyme-catalyzed cleavage reaction.<sup>25,27</sup>

**Double-Mutant Cycles Establish Functional Relevance of an Interaction Inferred from the Product Structure and the Deoxynucleotide Inhibitor Complex.** The available crystal structures for the HDV ribozyme show excellent agreement with respect to the global architecture. Differences reside at the active site, especially with respect to the position and inferred catalytic role of the active site nucleobase C76. Beyond differences in nucleobase positioning, the 2'-hydroxyl group of residue 76 experiences different microenvironments in these structures. Specifically, in both the product complex and ribozyme:-1dC inhibitor complex, the 2'-hydroxyl group of residue 76 resides within hydrogen-bonding distance of the N4 amino group of the C25 nucleobase on the

**Table 2. Reaction Rates and  $pK_a$ 's Determined for C76 Mutants for Brønsted Analysis<sup>a</sup>**

modification at C76	reaction $pK_a$	$pK_{a,wt} - pK_{a,mut}$	$k^{5'-O}/k^{5'-S}$	$\log(k_{mut}^{5'-O}/k_{mut}^{5'-S}) - \log(k_{wt}^{5'-O}/k_{wt}^{5'-S})$
wt	$5.99 \pm 0.07$	0	$2.97 \pm 0.79$	0
5-fluorocytosine	$5.36 \pm 0.08$	$0.54 \pm 0.06$	$0.80 \pm 0.20$	$-0.51 \pm 0.07$
6-azacytosine	$5.74 \pm 0.07$	$0.25 \pm 0.11$	$1.83 \pm 0.44$	$-0.25 \pm 0.29$
2-aminopyrimidine	$5.51 \pm 0.05$	$0.48 \pm 0.12$	$0.10 \pm 0.04$	$-1.50 \pm 0.17$
pseudoisocytosine	$5.65 \pm 0.19$	$0.37 \pm 0.19$	$3.03 \pm 0.04$	$0.02 \pm 0.06$
2-aminopyridine	$6.86 \pm 0.10$	$-0.90 \pm 0.06$	$0.087 \pm 0.003$	$-1.45 \pm 0.06$
2,4-diaminopyrimidine	$7.58 \pm 0.08$	$-1.59 \pm 0.07$	$0.0082 \pm 0.0023$	$-2.31 \pm 0.49$

<sup>a</sup>Tabulated reaction  $pK_a$  values for each ribozyme are the average of at least two independent measurements; relative rates are the average of at least four independent measurements.

5' side of P1.1, whereas in the C76U precursor structure the 2'-hydroxyl group of residue 76 has no apparent interacting partners.

We used double-mutant cycles comprised of atomic mutations to explore linkage between the 2'-hydroxyl group of C76 and the N4 exocyclic amino group of C25 (Figure 4). We show that although the 2'-hydroxyl group appears to make no energetic contribution to catalysis, the two sites exhibit mutual sensitivity both to volume and electrostatic alterations, suggesting that the two groups lie proximal to one another in the transition state.

Although the product structure fails to capture several key elements that comprise the reaction coordinate (e.g., the scissile phosphate, the metal ion, and the 2'-hydroxyl nucleophile), it preserves C76 and appears to contain catalytically relevant interactions. The ribozyme:-1dC inhibitor complex also lacks the nucleophilic hydroxyl group but contains the metal ion and the scissile phosphate and thus provides a reasonable mimic of the precursor form of the ribozyme. The structural model derived from X-ray crystallographic analysis of this complex shows the C76 2'-OH and C25 4-NH<sub>2</sub> groups in proximity, suggesting that these two groups achieve their relative orientation in the precursor ground state. We conclude that the ribozyme maintains this functional group juxtaposition in the transition state through to the product state.

**Brønsted Analysis Demonstrates the Influence of C76 Acidity on HDV Ribozyme Catalysis.** The conceptual underpinnings for general acid/base catalysis by protein and RNA enzymes derive predominantly from structure–reactivity analyses of non-enzymatic reactions.<sup>47,48</sup> Nevertheless, defining the energetic roles of general acid/base catalysis for enzymatic reactions presents greater challenges because the constraints imposed by the enzyme active site frequently restrict use of substrate or side-chain analogues that allow systematic variation of pK<sub>a</sub> via inductive effects. Consequently, more often than not, for a given enzyme we extrapolate mechanism from structure, mutagenesis, and computational analysis using chemical intuition derived from non-enzymatic reactions. Here, we probed general acid catalysis more directly using a series of ribozymes bearing modifications at the catalytic nucleobase. To allow correlation of the reaction rate versus the apparent pK<sub>a</sub>, we used the reactivity of the hyperactivated substrate S<sub>5',5'</sub> which gains no advantage from general acid catalysis, to correct for effects not specifically associated with pK<sub>a</sub> variation.

This approach implicates a significant dependence of the reaction rate on nucleobase acidity and yields an apparent Brønsted coefficient of  $\alpha_{\text{app}} = 0.8 \pm 0.2$  (Figure 5B), implying a product-like transition state, in which proton transfer from the catalytic nucleobase to the leaving group has advanced significantly. Been and co-workers reported a Brønsted coefficient of  $\alpha_{\text{app}} = 0.5$  for rescue of the C76U or C76Δ ribozymes by a series of exogenously added imidazole analogues.<sup>42</sup> This value implies a lesser degree of proton transfer mediated by imidazole than by C76, but includes the effects on imidazole binding.<sup>42,49,50</sup> In addition, whether imidazole facilitates the reaction of these mutant ribozymes by acting as a general acid or by acting as a general base remains ambiguous.<sup>50</sup> If imidazole does act as a general acid, the transition states for the imidazole-catalyzed and natural reactions need not have the same degree of proton transfer to the leaving group.

In non-enzymatic reactions, RNA undergoes base-catalyzed endonucleolytic cleavage with  $\beta_{\text{lg}} = -1.28$ ,<sup>51</sup> and a large <sup>18</sup>O

KIE on the 5'-O leaving group is observed,<sup>52</sup> reflecting the change in effective charge on the leaving group as the reaction proceeds from the ground state to the transition state in which 5'O–P bond cleavage is advanced.<sup>53</sup> Crudely estimating the effective charge on the leaving group in the ground state as +0.5 gives an effective charge equal to  $\sim -0.8$  (i.e., 0.5–1.28) in the transition state in the absence of a catalyst.<sup>54</sup> If the transition state in the HDV ribozyme active site resembles the transition state for the non-enzymatic reaction, the relatively large value of  $\alpha_{\text{app}}$ , aforementioned caveats notwithstanding, implies that the negative charge on the leaving group would be neutralized entirely. General acid catalysis appears to impart a significant fraction of the rate acceleration provided by the HDV ribozyme, similar to classic models of ribonuclease A catalysis,<sup>48,55,56</sup> in which Brønsted and KIE analysis of the leaving group reveal significant decreases in  $\beta_{\text{lg}}$ <sup>55</sup> and <sup>18</sup>O<sub>lg</sub><sup>56</sup> respectively, compared to corresponding non-enzymatic reactions. When  $\alpha$  exceeds 0.85 for reactions at pH 7 (as appears to be nearly the case for the HDV ribozyme), a general acid with pK<sub>a</sub> = 5 provides greater catalysis than a general acid of pK<sub>a</sub> = 7 (Note S1), despite the smaller fraction being in the kinetically competent form.<sup>37,41</sup> Perhaps ribozymes that use nucleobases to mediate proton transfer generally will have large Brønsted coefficients to compensate for their lack of functional groups with pK<sub>a</sub>'s near 7.

## CONCLUSION

Despite available precursor and product crystal structures of the HDV ribozyme, the relationship between the reaction pathway and structure has remained ambiguous. The product structure lacks three critical features that comprise the reaction coordinate: (i) a metal ion, the scissile phosphate, and the nucleophile; (ii) the C76U mutant used to determine the precursor structure has no catalytic activity in the absence of exogenous imidazole; and (iii) the ribozyme:-1dC inhibitor pre-cleavage complex structure lacks the nucleophile and required modeling to resolve ambiguous electron density surrounding the scissile phosphate. Here we have developed new approaches to investigate the spatial proximity between functional groups and to evaluate the efficiency of general acid catalysis within enzyme active sites. By using steric and electronic double-mutation cycles and Brønsted analysis with cytosine analogues at the catalytic residue 76, we demonstrate that C76 in the HDV ribozyme transition state resembles the reaction product structure with respect to both spatial position and degree of proton transfer to the leaving group. This work reduces some ambiguities between the reaction pathway and the physical RNA structure and supports the catalytic relevance of the most recently reported structure of the pre-cleavage complex of the ribozyme bound to a substrate lacking the nucleophilic 2'-hydroxyl group.

## ASSOCIATED CONTENT

### Supporting Information

Figure S1–S9, Note S1, and Tables S1–S4. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b01189.

## AUTHOR INFORMATION

### Corresponding Author

\*jpicciri@uchicago.edu

## Present Address

#J.L.: SynChem Inc., Elk Grove Village, IL 60007

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank members of the Piccirilli laboratory, Raghuvir Sengupta, and Dr. Jason Schwans for their comments and suggestions on the manuscript. We also thank Barbara Golden for kindly sharing the coordinates of her HDV ribozyme model, which we used to create the TOC figure. This work was funded by NIH Grant R01AI081987 to J.A.P. and R01 GM096000 to M.E.H. S.C.K. was partially supported by a National Institute of General Medical Sciences Medical Scientist National Research Service Award (5T32GM07281).

## REFERENCES

- (1) Hougland, J. L.; Piccirilli, J. A.; Forconi, M.; Lee, J.; Herschlag, D. In *The RNA World*, 3rd ed.; Gesteland, R. F., Cech, T. R., Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2006.
- (2) Fedorova, O.; Zingler, N. *Biol. Chem.* **2007**, *388*, 665.
- (3) Toor, N.; Keating, K. S.; Pyle, A. M. *Curr. Opin. Struct. Biol.* **2009**, *19*, 260.
- (4) Fedor, M. J. *Annu. Rev. Biophys.* **2009**, *38*, 271.
- (5) Webb, C.-H. T.; Riccitelli, N. J.; Ruminski, D. J.; Luptak, A. *Science* **2009**, *326*, 953.
- (6) Altman, S. *Mol. BioSyst.* **2007**, *3*, 604.
- (7) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. *Nature* **2004**, *428*, 281.
- (8) Cochrane, J. C.; Lipchock, S. V.; Strobel, S. A. *Chem. Biol.* **2007**, *14*, 97.
- (9) Schmeing, T. M.; Ramakrishnan, V. *Nature* **2009**, *461*, 1234.
- (10) Lilley, D. M. J.; Eckstein, F., Eds. *Ribozymes and RNA Catalysis*; Royal Society of Chemistry: Cambridge, 2008.
- (11) Frederiksen, J. K.; Fong, R.; Piccirilli, J. A. In *Nucleic Acid-Metal Ion Interactions*; Hud, N. V., Ed.; Royal Society of Chemistry: Cambridge, 2009; pp 260–306.
- (12) Roth, A.; Weinberg, Z.; Chen, A. G. Y.; Kim, P. B.; Ames, T. D.; Breaker, R. R. *Nat. Chem. Biol.* **2013**, *10*, 56.
- (13) Ferré-D'Amaré, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567.
- (14) Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferré-D'Amaré, A. R. *Science* **2002**, *298*, 1421.
- (15) Ke, A. L.; Zhou, K. H.; Ding, F.; Cate, J.; Doudna, J. A. *Nature* **2004**, *429*, 201.
- (16) Klein, D. J.; Ferré-D'Amaré, A. R. *Science* **2006**, *313*, 1752.
- (17) Martick, M.; Scott, W. G. *Cell* **2006**, *126*, 309.
- (18) Chen, J.-H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. *Biochemistry* **2010**, *49*, 6508.
- (19) Suslov, N. B. The crystal structure of the Varkud satellite ribozyme: Implications for RNA catalysis, biology and evolution. Ph.D. Thesis, The University of Chicago, 2012.
- (20) Eiler, D.; Wang, J.; Steitz, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 13028.
- (21) Liu, Y.; Wilson, T. J.; McPhee, S. A.; Lilley, D. M. J. *Nat. Chem. Biol.* **2014**, *10*, 739.
- (22) Shih, I.-H.; Been, M. D. *Annu. Rev. Biochem.* **2002**, *71*, 887.
- (23) Salehi-Ashtiani, K.; Lupták, A.; Litovchick, A.; Szostak, J. W. *Science* **2006**, *313*, 1788.
- (24) Eickbush, D. G.; Eickbush, T. H. *Mol. Cell. Biol.* **2010**, *30*, 3142.
- (25) Nakano, S.; Chadalavada, D. M.; Bevilacqua, P. C. *Science* **2000**, *287*, 1493.
- (26) Nakano, S.; Proctor, D. J.; Bevilacqua, P. C. *Biochemistry* **2001**, *40*, 12022.
- (27) Das, S. R.; Piccirilli, J. A. *Nat. Chem. Biol.* **2005**, *1*, 45.
- (28) Gong, B.; Chen, J.-H.; Chase, E.; Chadalavada, D. M.; Yajima, R.; Golden, B. L.; Bevilacqua, P. C.; Carey, P. R. *J. Am. Chem. Soc.* **2007**, *129*, 13335.
- (29) Cerrone-Szakal, A. L.; Siegfried, N. A.; Bevilacqua, P. C. *J. Am. Chem. Soc.* **2008**, *130*, 14504.
- (30) Lu, J.; Koo, S. C.; Li, N.-S.; Piccirilli, J. A. *Nucleosides, Nucleotides Nucleic Acids* **2015**, *34*, 114.
- (31) Lu, J.; Li, N.-S.; Koo, S. C.; Piccirilli, J. A. *J. Org. Chem.* **2009**, *74*, 8021.
- (32) Shih, I.-H.; Been, M. D. *Biochemistry* **2000**, *39*, 9055.
- (33) Kao, C.; Zheng, M.; Rudisser, S. *RNA* **1999**, *5*, 1268.
- (34) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. *Nucleic Acids Res.* **1987**, *15*, 8783.
- (35) Pyle, A. M.; Murphy, F. L.; Cech, T. R. *Nature* **1992**, *358*, 123.
- (36) Herschlag, D.; Eckstein, F.; Cech, T. R. *Biochemistry* **1993**, *32*, 8299.
- (37) Carter, P. J.; Winter, G.; Wilkinson, A. J.; Fersht, A. R. *Cell* **1984**, *38*, 835.
- (38) Fersht, A. *Structure and Mechanism in Protein Science*; W.H. Freeman and Co.: New York, 1999.
- (39) Horovitz, A. *Folding Des.* **1996**, *1*, R121.
- (40) Hougland, J. L.; Deb, S. K.; Maric, D.; Piccirilli, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 13578.
- (41) Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2259.
- (42) Perrotta, A. T.; Wadkins, T. S.; Been, M. D. *RNA* **2006**, *12*, 1282.
- (43) Forconi, M.; Schwans, J. P.; Porecha, R. H.; Sengupta, R. N.; Piccirilli, J. A.; Herschlag, D. *Chem. Biol.* **2011**, *18*, 949.
- (44) Wadkins, T. S.; Perrotta, A. T.; Ferré-D'Amaré, A. R.; Doudna, J. A.; Been, M. D. *RNA* **1999**, *5*, 720.
- (45) Kirsch, J. F.; Toney, M. D. *Ann. N.Y. Acad. Sci.* **1990**, *585*, 48.
- (46) Hollfelder, F.; Herschlag, D. *Biochemistry* **1995**, *34*, 12255.
- (47) Oivanen, M.; Kausela, S.; Lönnberg, H. *Chem. Rev.* **1998**, *98*, 961.
- (48) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 432.
- (49) Perrotta, A. T.; Shih, I.; Been, M. D. *Science* **1999**, *286*, 123.
- (50) Shih, I. H.; Been, M. D. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1489.
- (51) Kosonen, M.; Youseti-Salakdeh, E.; Stromberg, R.; Lonnberg, H. *J. Chem. Soc., Perkin Trans. 2* **1997**, 2661.
- (52) Harris, M. E.; Dai, Q.; Gu, H.; Kellerman, D. L.; Piccirilli, J. A.; Anderson, V. E. *J. Am. Chem. Soc.* **2010**, *132*, 11613.
- (53) Wong, K.-Y.; Gu, H.; Zhang, S.; Piccirilli, J. A.; Harris, M. E.; York, D. M. *Angew. Chem.* **2012**, *124*, 671.
- (54) Ye, J.-D.; Li, N.-S.; Dai, Q.; Piccirilli, J. A. *Angew. Chem., Int. Ed.* **2007**, *46*, 3714.
- (55) Raines, R. T. *Chem. Rev.* **1998**, *98*, 1045.
- (56) Gu, H.; Zhang, S.; Wong, K.-Y.; Radak, B. K.; Dissanayake, T.; Kellerman, D. L.; Dai, Q.; Miyagi, M.; Anderson, V. E.; York, D. M.; Piccirilli, J. A.; Harris, M. E. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 13002.
- (57) Albert, A.; Goldacre, R.; Phillips, J. J. *Chem. Soc.* **1948**, 2240.
- (58) Wempen, I.; Duschinsky, R.; Kaplan, L.; Fox, J. J. *J. Am. Chem. Soc.* **1961**, *83*, 4755.
- (59) Barlin, G.; Pfeleiderer, W. *J. Chem. Soc. B* **1971**, 1425.
- (60) Premecz, J. E.; Ford, M. E. *J. Chromatogr.* **1987**, *388*, 23.
- (61) Kan, L.-S.; Lin, W.-C.; Yadav, R. D.; Shih, J. H.; Chao, I. *Nucleosides Nucleotides* **1999**, *18*, 1091.
- (62) Chen, D. L.; McLaughlin, L. W. *J. Org. Chem.* **2000**, *65*, 7468.
- (63) Miller, J. M.; Blackburn, A. C.; Shi, Y.; Melzak, A. J.; Ando, O. Y. *Electrophoresis* **2002**, *23*, 2833.