

Charged Nucleobases and Their Potential for RNA Catalysis

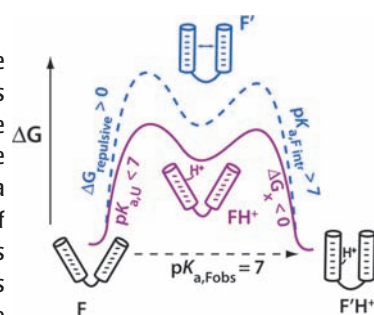
JENNIFER L. WILCOX, AMARPREET K. AHLUWALIA, AND
PHILIP C. BEVILACQUA*

*Department of Chemistry, The Pennsylvania State University, University Park,
Pennsylvania 16802, United States*

RECEIVED ON FEBRUARY 22, 2011

CONSPECTUS

Catalysis in living cells is carried out by both proteins and RNA. Protein enzymes have been known for over 200 years, but RNA enzymes, or “ribozymes”, were discovered only 30 years ago. Developing insight into RNA enzyme mechanisms is invaluable for better understanding both extant biological catalysis as well as the primitive catalysis envisioned in an early RNA-catalyzed life. Natural ribozymes include large RNAs such as the group I and II introns; small RNAs such as the hepatitis delta virus and the hairpin, hammerhead, VS, and glmS ribozymes; and the RNA portion of the ribosome and spliceosome. RNA enzymes use many of the same catalytic strategies as protein enzymes, but do so with much simpler side chains. Among these strategies are metal ion, general acid–base, and electrostatic catalysis. In this Account, we examine evidence for participation of charged nucleobases in RNA catalysis. Our overall approach is to integrate direct measurements on catalytic RNAs with thermodynamic studies on oligonucleotide model systems.



The charged amino acids make critical contributions to the mechanisms of nearly all protein enzymes. Ionized nucleobases should be critical for RNA catalysis as well. Indeed, charged nucleobases have been implicated in RNA catalysis as general acid–bases and oxyanion holes. We provide an overview of ribozyme studies involving nucleobase catalysis and the complications involved in developing these mechanisms. We also consider driving forces for perturbation of the pK_a values of the bases. Mechanisms for pK_a values shifting toward neutrality involve electrostatic stabilization and the addition of hydrogen bonding. Both mechanisms couple protonation with RNA folding, which we treat with a thermodynamic formalism and conceptual models. Furthermore, ribozyme reaction mechanisms can be multichannel, which demonstrates the versatility of ribozymes but makes analysis of experimental data challenging.

We examine advances in measuring and analyzing perturbed pK_a values in RNA. Raman crystallography and fluorescence spectroscopy have been especially important for pK_a measurement. These methods reveal pK_a values for the nucleobases A or C equal to or greater than neutrality, conferring potential histidine- and lysine/arginine-like behavior on them. Structural support for ionization of the nucleobases also exists: an analysis of RNA structures in the databases conducted herein suggests that charging of the bases is neither especially uncommon nor difficult to achieve under cellular conditions. Our major conclusions are that cationic and anionic charge states of the nucleobases occur in RNA enzymes and that these states make important catalytic contributions to ribozyme activity. We conclude by considering outstanding questions and possible experimental and theoretical approaches for further advances.

Introduction

A paradigm shift in biological catalysis occurred in the 1980s when Cech and Altman discovered that certain RNAs are enzymes.¹ This was remarkable for elucidating new mechanisms of enzyme catalysis, and for insights into possible mechanisms of primitive catalysis in an RNA World. Since these landmark discoveries, much has been learned about ribozymes: additional classes have been discovered, numerous ribozymes and DNA enzymes

(“DNAzymes”) have been reported, high-resolution structures of ribozymes have been solved, and theoretical approaches to understanding ribozyme reactions have made headway.²

The focus of this Account is the direct participation of nucleobases in RNA catalysis. We begin by considering roles for ionized bases in RNA catalysis. Structural support for ionization is provided via analysis of databases. This is followed by discussion of experimental and conceptual

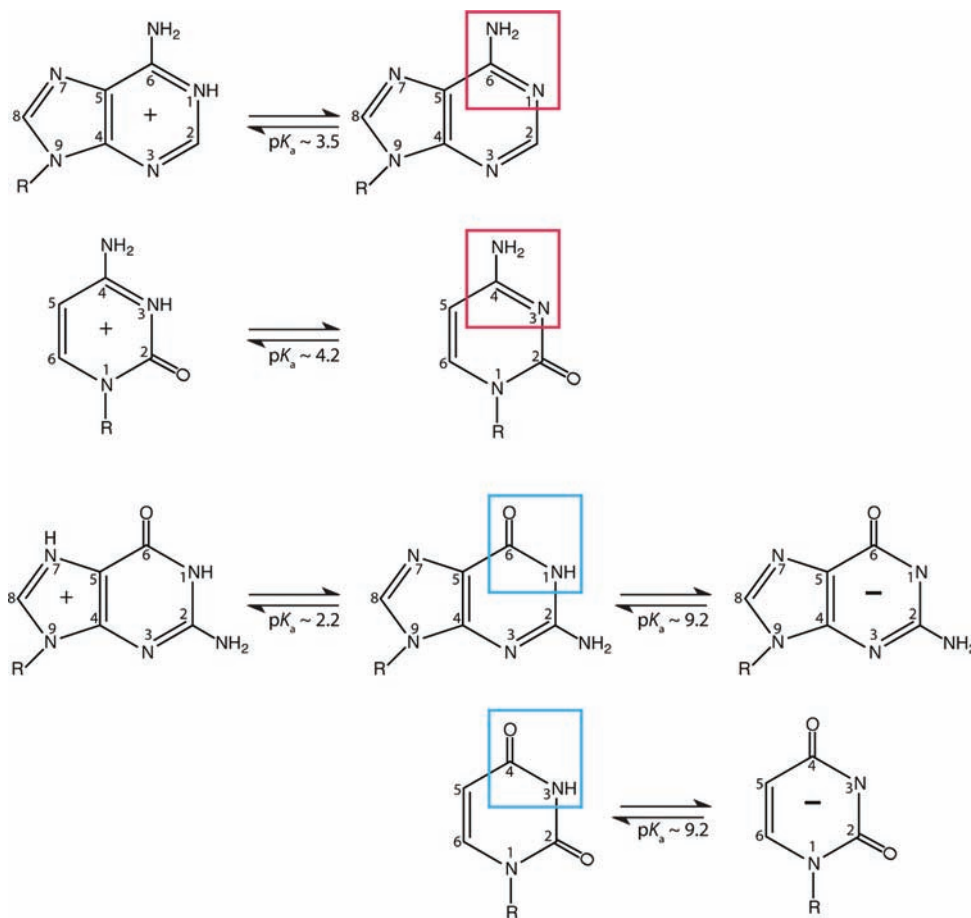


FIGURE 1. Charged nucleobases in RNA. Adenosine and cytidine can protonate on N1 and N3, with pK_a 's near 3.5 and 4.2. Guanosine and uridine can deprotonate on N1 and N3 with pK_a 's near 9.2; guanine has protonation possible at N7. Boxes indicate functional groups responsible for these groupings. These pK_a values are often perturbed in natural RNAs.

advances for measuring and interpreting pK_a values. Next, we examine select cases wherein ionized bases have been implicated in RNA catalysis. Lastly, driving forces for shifting pK_a values are considered. Our approach is to bridge functional study on ribozymes with thermodynamic measurements on model oligonucleotides.

Charged Nucleobases and Their Potential for Catalysis

A major difference between RNA and protein enzymes is the diversity of the side chains. Proteins are composed of 20 chemically diverse amino acids, whereas RNA has just four similar heterocycles.³ It is well established that the charged amino acids play key roles in protein enzyme catalysis: lysine and arginine act as oxyanion holes to stabilize charge development in the transition state, whereas histidine cycles between neutral and positive when it transfers protons. Thus, one might inquire whether the nucleobases play similar roles.

In contrast to the amino acids, the unmodified nucleosides in RNA are typically neutral at biological pH.³ The pK_a values closest to neutrality on adenosine and cytidine are ~ 3.5 and ~ 4.2 , respectively, while those on guanosine and uridine are ~ 9.2 (Figure 1).³ A and C become cationic when protonated, while G and U become anionic when deprotonated. However, lacking pK_a perturbation, the bases are only ~ 0.1 – 1% in their charged forms at neutrality. On the one hand, these pK_a 's help nucleic acids store genetic information by virtue of complementary hydrogen bonding. However, these pK_a 's limit potential contributions to catalysis by electrostatics and proton transfer.

The intrinsic ability of charged and near-neutral functionalities to enhance nucleic acid enzyme activity has been elegantly demonstrated by Perrin and co-workers through studies on chimeric DNAzymes containing amine and imidazole functionalities.⁴ We have asserted that poor population of the functional charged forms of the bases is one of the greatest limitations to RNA catalysis, because once

properly charged RNA catalyzes reactions with rates estimated to equal those of protein enzymes.⁵ Acidic and basic amino acids often have pK_a values shifted 2 or more units in proteins. The pK_a 's of RNA side chains shift as well. Indeed, every time a Watson–Crick base pair forms, linkage to folding causes the pK_a to shift further away from neutrality by a few units.⁶

Interestingly, examples are also known where A and C shift toward neutrality. We divided such pK_a shifts into “Class I” and “Class II”, depending on whether the loaded proton is sequestered in base pairing (Figure 2).⁷ Both classes have potential to participate in catalysis: Class I primarily by serving as an oxyanion hole (Figure 3, Interaction “3”), and Class II by participating primarily in proton transfer (Figure 3, Interactions “1” and “2”), although these roles are not exclusive owing to acidification of external amino protons upon protonation of an imino proton.⁸ As described below, pK_a values for Class I motifs from 7 to 8 have been reported in model systems, indicating potential for populating the protonated state at neutrality, and lysine/arginine-like behavior. Also, pK_a values for Class II motifs near 7 have been reported for A and C, indicating potential for histidine-like behavior.

Structural Support for Ionization of the Nucleobases

Given that enzymes need a compact tertiary structure with a cleft for conducting catalysis, a major issue in RNA folding is neutralization of the backbone. Catalytic RNAs achieve compact structures primarily through counterion condensation: cations neutralize the backbone, which allows compaction.⁹ Another strategy for attaining electrostatic stabilization is protonation of A or C, which affords a two-hydrogen bond cationic base. We focus on $A^+\bullet C$ wobbles because these are common and compatible with an A-form helix, making them amenable to experiments.¹⁰

In an effort to assess structural support for protonated bases in RNA, we surveyed databases for Class I protonated base pairs. We inspected the Fox lab RNA database as well as a database from the Major lab.^{11,12} A base motif was inferred as protonated if two hydrogen-bonding heteroatoms were properly positioned (2.5–3.5 Å between heteroatoms) and maintained an angle ($X-H\cdots Y$) of 140–180°. These criteria were verified for both of the hydrogen bonds in each $A^+\bullet C$ wobble. We found 57 unique occurrences in 37 crystal or NMR structures consistent with these criteria for protonation of the A. Of these, 54 motifs were isolated $A^+\bullet C$ pairs and three were base triples involving an $A^+\bullet C$ base pair.

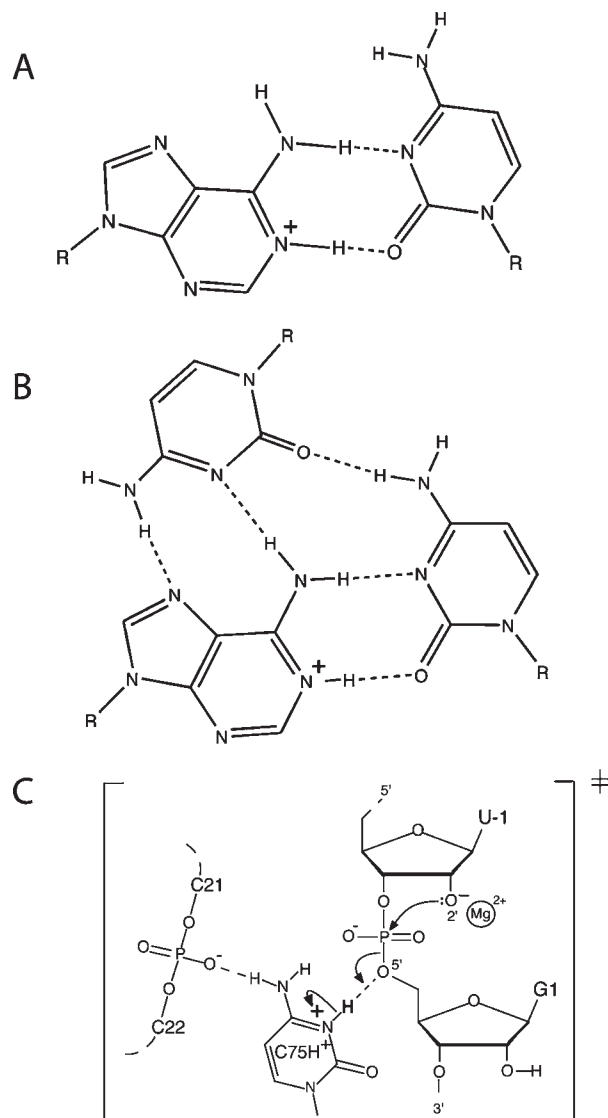


FIGURE 2. Examples of Class I and II protonated base interactions. (A) Structure of a two-hydrogen bond $A^+\bullet C$ base pair. (Class I base pair.) (B) Structure of a base triple involving a two-hydrogen bond $A^+\bullet C$ base pair and a third base (Class I base pair). (C) Model for active site of the HDV ribozyme (Class II interaction).²⁴

For those structures that reported experimental pH, half had a pH in the range of 6.5–7.5, suggesting a pK_a near biological pH. This analysis suggests that charging of the bases into a cationic state is not uncommon.

Anionic base formation has less direct structural support. Given that RNA is a polyanion, electrostatic stabilization of a negative charge might seem improbable. Supporting this notion, molecular dynamics (MD) trajectories on the hairpin and *glmS* ribozymes suggest that an anionic G near the active site is disfavored due to electrostatic repulsion.^{13,14} On the other hand, experimental data show that RNA is fully capable of binding anions. Anion sites were first described

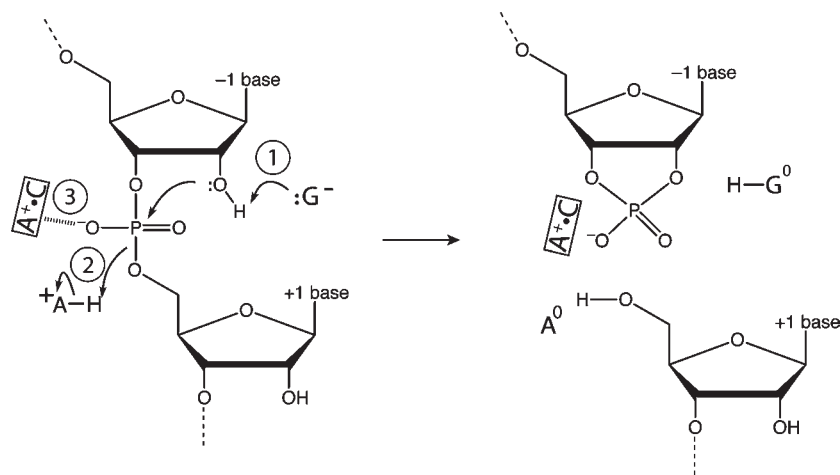


FIGURE 3. Involvement of charged bases in RNA catalysis. Shown are three potential interactions of charged nucleobases in the mechanism of small ribozymes. Interaction 1 involves an anionic G acting as a general base. Interaction 2 involves a cationic A acting as a general acid. Interaction 3 involves a cationic A⁺•C base pair acting as an oxyanion hole.

by Auffinger and co-workers,¹⁵ and Golden et al. recently described sites for sulfate and selenate in RNA.¹⁶ They proposed that anion binding can be achieved through two motifs: consecutive helical C's, whose major groove amines provide a positive potential, or positioning of one or two metal ions near the anion. A third possible motif is through-space electrostatic interaction with a cationic base, while a fourth is hydrogen bonding of a deprotonated functionality, in analogy to protonated functionalities. For example, an anionic guanosine could hydrogen bond with a 2'-OH nucleophile (Figure 3), which is supported structurally from hairpin ribozyme precleaved, product, and transition state structures.^{17,18}

Measuring and Interpreting pK_a Values in RNA

There has been growing interest in measuring pK_a values of RNA. The two basic approaches involve monitoring the fraction of RNA in the protonated state as a function of pH as reported by (1) spectroscopic or (2) kinetics measurements. First, we consider spectroscopic approaches. Importantly, pK_a shifts have to occur in the ground state, rather than the transition state, to affect reactive populations;^{7,19} this means such shifts should be measurable by nonkinetics means such as spectroscopy. A two-state model of protonated-to-unprotonated states is typically invoked to interpret these data. A spectroscopic signal (e.g., UV, fluorescence, Raman, or NMR) is fit to a two-state Henderson–Hasselbalch equation to yield a pK_a and a Hill coefficient, which allows for multiple linked protonations.

The pK_a values of larger RNAs have been determined by NMR spectroscopy, especially by changes in ¹³C chemical

shift. Legault and Pardi showed, using the lead-dependent ribozyme, that nucleobase pK_a's cluster into three classes: unstructured loops have pK_a's that are *unshifted*, Watson–Crick base pairs have pK_a's perturbed *away* from neutrality, and wobble A⁺•C pairs have pK_a's perturbed *toward* neutrality.⁶

NMR was also applied to the HDV ribozyme (structure and mechanism provided in Figure 4) to measure the pK_a of the catalytic nucleobase C75 using ¹³C spectroscopy.²⁰ In the cleaved form, the pK_a was found to be largely unperturbed, while line broadening and splitting precluded pK_a determination in the precleaved state. More recently, we collaborated with the Carey and Golden laboratories to measure a pK_a for the precleaved state of the HDV ribozyme using Raman crystallography.²¹ The interested reader is directed to a review article on the technique.²² We designed non-cleavable forms of the HDV ribozyme in which the 2'-hydroxyl of U-1 was modified. A pK_a of 6.15 was found in 20 mM Mg²⁺, which shifted to 6.4 in 2 mM Mg²⁺ (Figure 5). These pK_a's and their anticooperative coupling to Mg²⁺ are in agreement with intrinsic pK_a's from kinetics of 5.9 and 7.25 for Mg²⁺-bound and -free states, respectively.²³ This anticooperative coupling is likely due to through-space repulsion between protonated C75 and bound Mg²⁺, as suggested by a crystal structure of a C75-containing precleaved HDV ribozyme.²⁴ Remarkably, these near-neutral pK_a values confer a histidine-like pK_a on cytosine, which is optimal for general acid–base chemistry.¹⁹ The ability of Raman crystallography of RNA, which is conducted at room temperature,²² to measure these pK_a's may be because crystallography selects for one state of the ribozyme, thus

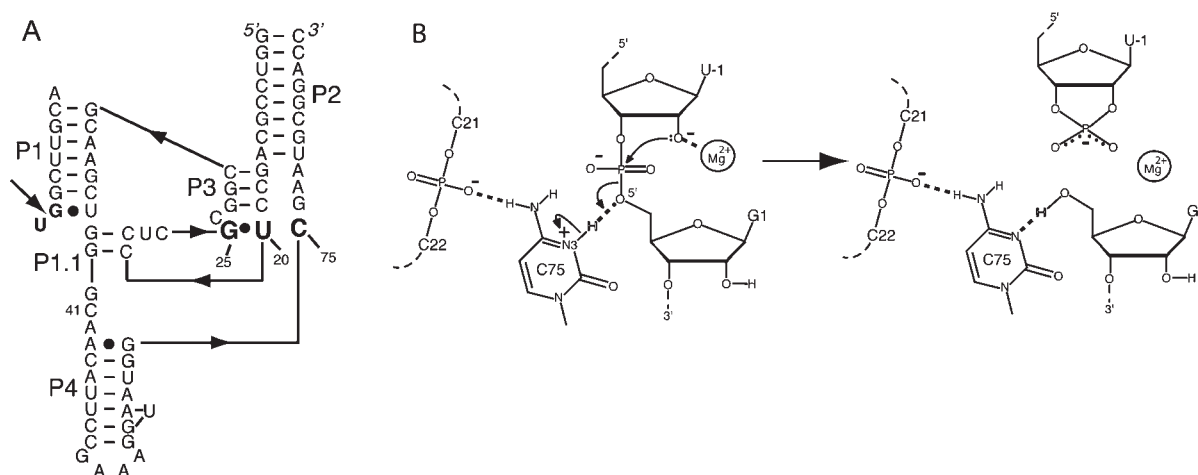


FIGURE 4. Structure and proposed mechanism of the HDV ribozyme. (A) Secondary structure of two-piece ribozyme. Arrow depicts cleavage site. (B) Proposed mechanism with Mg^{2+} acting as a Lewis acid and protonated C75 as a general acid. Reproduced with permission from ref 49. Copyright 2011 American Chemical Society.

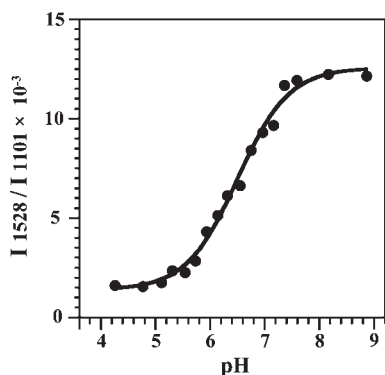


FIGURE 5. pH titration of HDV ribozyme crystals, using construct shown in Figure 4A. The y-axis is the relative intensity ratio of the 1528 cm^{-1} band (neutral cytosine) to the intensity of the band at 1101 cm^{-1} (PO_2^- , internal standard). A pK_a of 6.40 ± 0.05 at 2 mM Mg^{2+} was obtained. Data were fit to a Henderson–Hasselbalch equation with a Hill constant of unity. Adapted with permission from ref 21. Copyright 2007 American Chemical Society.

preventing misfolding. Indeed, we have shown that the HDV ribozyme is especially prone to misfolding.²⁵ Also, our crystallographic constructs were based upon sequences engineered to have fast and monophasic folding.²⁵

Measurements of pK_a 's of the general acid and base of the hairpin ribozyme have been made recently. Wedekind, Carey, and co-workers used Raman crystallography to determine a pK_a of 5.5 for A38, the putative general acid in the hairpin ribozyme,²⁶ and a similar value of 5.9 was extrapolated by Fedor and colleagues using fluorescence spectroscopy;²⁷ moreover, these pK_a values agree with analyses of rate–pH studies. In addition, Fedor and colleagues used fluorescence spectroscopy to obtain a pK_a of 10.6 for G8, the putative general base, which was extrapolated from a value of 9.5

measured with the fluorescent base 8-azaguanine.²⁸ There appears to be agreement that A38 serves as the general acid, but the role of G8 is unsettled: rate–pH profiles can be interpreted with G8 as the general base in the reaction,^{19,29} but hydroxide could deprotonate the 2'OH while G8 could simply hydrogen bond in the transition state.^{28,30}

We have also measured pK_a values for model oligonucleotides using ^{31}P NMR and UV–vis spectroscopy.^{10,31,32} For Class I oligonucleotides with an $\text{A}^+\cdot\text{C}$ base pair, ionization was measured by placing a thiophosphate near the site of protonation.³¹ The resulting mixture of diastereomers leads to two peaks in the ^{31}P NMR that are separated from other phosphorus resonances and can be tracked as a function of pH. We measured effects of position, temperature, and ionic strength for an $\text{A}^+\cdot\text{C}$ wobble in DNA using this method.³³ Under the most favorable conditions (low temperature, low ionic strength, and central positioning of the wobble), a pK_a of 8.0 was found, suggesting that nucleobases can be cationic under biological conditions.

In addition, pK_a values of unstructured oligonucleotides with a single ionizable group, such as UUCUU and UUAUU, have been determined using UV-detected pH titrations.^{10,32} These measurements provide unfolded state pK_a estimates (see below).¹⁰ Also, the ionic strength dependence of these systems was used to interpret effects of ionic strength on the HDV ribozyme as lack of ion penetration.³²

Lastly, chemical modification, such as methylation, as a function of pH has been used to determine pK_a 's in large RNAs, such as the ribosome. These have led to measurement of a pK_a near 7.³⁴ One problem with this method is potential for nonequivalence in sites of deprotonation and

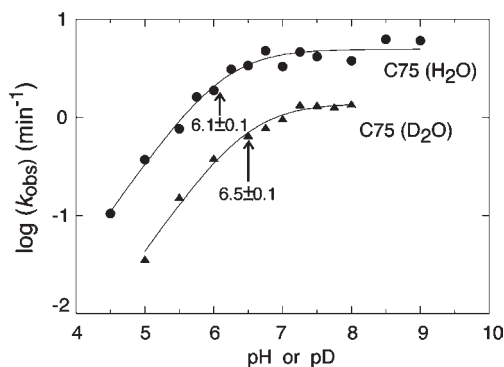


FIGURE 6. Rate–pH profile for the HDV ribozyme in H₂O or D₂O and 10 mM Mg²⁺. Note that the pK_a is near the rounding or flex point of the curve and that data can be obtained over a wide range of rates. The slope of the log–linear region is ~1, consistent with one proton ionizing in this region. The offset in the pK_a and the difference in the plateau values between H₂O and D₂O are consistent with chemistry as the rate-limiting step. Adapted with permission from ref 5. Copyright 2000 American Association for the Advancement of Science.

methylation, which is because complex RNA structures can refold in a way that can expose distal, nonionizing sites to methylation.

Next we consider kinetics approaches of attaining pK_a's. Most often kinetics data are plotted linear in free energy, that is, log *k*_{obs} versus pH (Figure 6). This is doable because decreasingly small amounts of RNA in the functional protonation state remain detectable by gathering longer time points. Kinetic assays thus offer the ability to detect poorly populated (e.g., 1% or less) general acids and bases; typically this is not afforded by spectroscopy, which often has detection limits of ~5% bound. By analyzing such “log–linear” regions of rate–pH profiles, one can attain additional information, such as reaction channels^{35,36} and nonproductive roles for protons, such as misfolding the RNA.^{10,37} The slope affords the net number of protons ionizing in the step.

Extraction of pK_a values from kinetics can be very complex, however. First, data can fit well to the Henderson–Hasselbalch equation but have nothing to do with charging of bases; for example, apparent pK_a's can be due to a change in the rate-limiting step from chemistry to a conformational change, or to another reaction channel. Second, because there is typically ionization of both a general acid *and* a general base (along with any other ionization related to structure), the shape of the rate–pH profile is not a reliable guide as to whether the observed pK_a comes from the general acid or base. We have developed a partition function analysis to aid interpreting rate–pH profiles, which can provide insight, but not proof of mechanism.^{19,29} Third, extremes of pH chemically denature RNA and DNA by rupturing

hydrogen bonding. Moreover, because ionization of a number of bases can unfold the RNA, an apparent pK_a for loss of cleavage activity shifted *toward neutrality* can occur owing to statistical effects, as first described by Knitt and Herschlag.^{10,37}

Because of the above hazards in interpreting rate–pH profiles, it is critical to have complementary measurements of the pK_a, such as the aforementioned spectroscopic approaches. Calculations can help as well, and we recently used MD simulations to gain insight into how protonation of a noncatalytic base triple affects the active site.³⁸ These simulations were consistent with complex kinetic rate–pH profiles.²³ In addition, deconvoluting pK_a's from rate–pH profiles can be aided by studies with nucleobase analogues with altered pK_a's.^{27,28,39,40}

Involvement of Charged Bases in RNA Catalysis

The reaction pathways of RNA enzymes vary in terms of contributions of metal ion, general acid–base, and electrostatic catalysis. In general, large ribozymes prefer metal ion catalysis, while smaller ribozymes prefer general acid–base catalysis. Large ribozymes have been extensively studied biochemically and structurally.¹ With the exception of the spliceosome, high-resolution structures of these RNAs have been solved, often in several catalytic states. A major contribution to the catalytic mechanism is stabilization of charge development by divalent ions, which act as Lewis acids.⁴¹ Smaller ribozymes such as the HDV, hairpin, hammerhead, VS, and *glm S* have also been intensively studied by biochemical and structural approaches.¹ With the exception of the VS ribozyme, there are high-resolution structures of these RNAs as well, including reactant states, product states, and transition state mimics. A major catalytic theme for these reactions has been the direct involvement of nucleobases in self-cleavage.⁴²

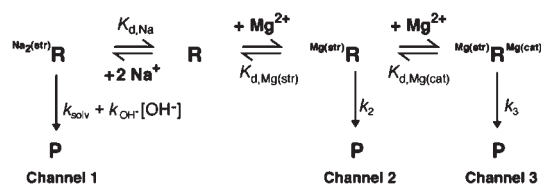
One can associate the termini produced in RNA catalysis with the chemical roles that catalytic nucleobases play. Small ribozymes perform phosphodiester bond cleavage in which a nucleophilic 2'–OH attacks the adjoining phosphorus, with the bridging 5'–oxygen as the leaving group (Figure 3). This leaving group has no vicinal source of labile protons, which might serve as a general acid. In contrast, in large ribozymes the bridging 3'–oxygen leaving group has the 2'–OH to assist in its departure. In addition, metal ion and nucleobase catalysis are not exclusive within these ribozyme classes: there is strong evidence that the HDV ribozyme employs a Mg²⁺ ion as a Lewis acid (Figure 4B),²⁴ and it is possible that large ribozymes use nucleobases in catalysis.

We favor general acid–base chemistry as a predominant mechanism for small ribozymes for three reasons: (1) rate–pH profiles are consistent with general acid–base chemistry,^{5,19,29,32,35} (2) hyperactivated leaving group studies implicate nucleobases as general acids,^{43,44} and (3) Brønsted plots based on small molecule rescue show nucleobase acidity contributes to catalysis.⁴⁵ Participation in general acid–base chemistry does not eliminate structural or electrostatic roles for the bases; in fact, these are likely given that the general acid and base are charged for part of the reaction cycle. Indeed, RNase A, which catalyzes the same reaction, uses a lysine to stabilize charge development.⁴⁶ Kinetically equivalent modes of catalysis in which nucleobases hydrogen bond to the active site and solvent species transfer protons remain possibilities in ribozyme reactions where Brønsted and activated leaving group studies have not been performed.

An emerging theme in the general acid–base mechanisms for most small ribozymes is a reactant state in which the general base is anionic and general acid is cationic (Figure 3). Anionic guanine has been implicated as a general base for the hairpin, VS, *glm S*, and hammerhead ribozymes. At the general acid position, cationic species are generally found: cytosine for the HDV ribozyme, adenine for the hairpin and VS ribozymes, and GlcN6P for the *glm S* ribozyme. All of these species cycle from charged to neutral as they accept and donate their protons (Figure 3). The observation that the general acid and base are charged together in the reactant state of many ribozymes suggests the possibility for through-space electrostatic stabilization.

Mechanisms used by small ribozymes can be remarkably complex. We described the reaction of the HDV ribozymes with a multichannel mechanism, in which each channel leads to the same products (Scheme 1).^{32,35,36} Channel 1 operates without participation of Mg^{2+} , Channel 2 with structural Mg^{2+} , and Channel 3 with structural and catalytic Mg^{2+} , and thus, the dominant reaction mechanism is a function of Mg^{2+} concentration. Ions were classified as structural or catalytic on the basis of effects of ion size and affinity.^{35,36} The presence of multiple channels was inferred from the complexity of rate–pH and rate– Mg^{2+} profiles over wide pH and Mg^{2+} ranges.^{32,35,36} Channels 1 and 2 have rate–pH profiles inverted from Channel 3. Moreover, Channels 1 and 2 have subchannels, in which water or hydroxide participates as the base, depending on pH.³² One theme to emerge is that C75 is cationic and serves as the general acid in all three channels,

SCHEME 1



indicating robustness and importance of its shifted pK_a to function.³²

Other ribozymes likely have multiple reaction channels as well.⁴⁷ On the one hand, multichannel mechanisms have the potential to confound model-driven interpretation of rate–pH profiles; indeed, careful tuning of experimental conditions is needed to isolate a single channel. On the other hand, these mechanisms reveal the versatility of ribozymes: that they can employ metal ions, nucleobases, and solvent to catalyze the same reaction in multiple ways.

Driving Forces for Ionizing the Bases

The Thermodynamic driving forces for pK_a perturbation have been studied.^{10,48} Two types of molecular mechanisms for these driving forces are (1) gain of hydrogen bonding and (2) electrostatic stabilization, both of which involve coupling of protonation and RNA folding.

Coupling between protonation of A or C and gain of hydrogen bonding can lead to non-Watson-Crick base pairs (Figure 2).¹⁰ We previously developed a formalism in which the driving force for nucleobase protonation in the folded state is the stabilization in RNA folding offered by protonation, ΔG_x (Figure 7A, right side green box).

$$\Delta G_x = -1.42(\text{pK}_{a,\text{Fobs}} - \text{pK}_{a,\text{U}}), \quad \text{in kcal/mol (at 37 }^\circ\text{C)} \quad (1)$$

Given that pK_a shifts of 3 or 4 units are known, ΔG_x can be a remarkably large, -4 to -6 kcal/mol. This arises from strengthened stacking, hydrogen bonding, and electrostatics.

The large value of ΔG_x is generally not observed experimentally, however. This is because directly measuring it would require starting in the protonated unfolded state (Figure 7A, UH^+ state), but this is impossible because lowering the pH enough to fully populate it leads to protonation of other sites in the unfolded state (Figure 7A, pink box), which drives acid denaturation of the helix. Practically speaking, the minimum in folding free energy for most RNAs occurs around pH 5.5. We thus typically calculate ΔG_x , using the thermodynamic cycle in Figure 7A (green box) and the experimental values of $\text{pK}_{a,\text{Fobs}}$ and $\text{pK}_{a,\text{U}}$.

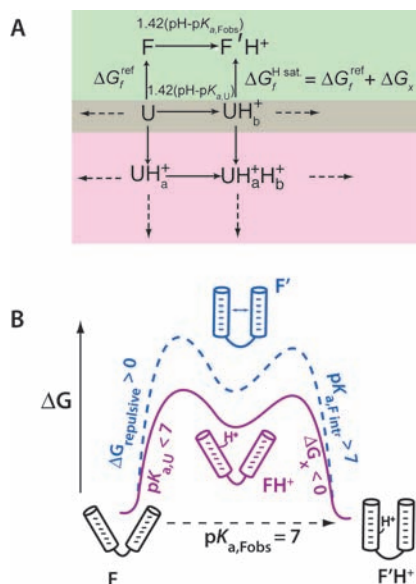


FIGURE 7. Linkage between folding and $\text{p}K_{\text{a}}$ shifting.¹⁰ (A) Thermodynamic model linking RNA folding (green box, vertical lines) and protonation of the folded (F) and unfolded (U) states (green box, horizontal lines; pink boxes, all lines). This model assumes a single shifted $\text{p}K_{\text{a}}$ in the folded state. Values are free energies in kcal/mol. We measure the folded and unfolded state $\text{p}K_{\text{a}}$'s, as well as the reference free energy $\Delta G_{\text{f}}^{\text{ref}}$, and calculate ΔG_{x} . (B) Free energy diagram showing model-driven interpretation of $\text{p}K_{\text{a}}$ values for electrostatic stabilization at pH 7. The model is conceptual, and the difference in the free energy of the two paths is unknown; the two paths are offset only to allow adequate visualization. Protonation is three-state, with either path involving a high-energy intermediate. Panel (A) is adapted with permission from ref 10. Copyright 2005 RNA Society.

Regarding electrostatic stabilization as a driving force for ionization, Pyle, Honig, and co-workers have reported nonlinear Poisson–Boltzmann (NLPB) electrostatics calculations in which regions of high negative electrostatic potential drive protonation of A and C,⁴⁸ and we have provided qualitatively similar NLPB arguments on the HDV ribozyme for upward shifting of the $\text{p}K_{\text{a}}$ of C75 (Figure 8).^{35,48,49} These types of calculations confirm which regions of the RNA contribute to shifting of ionization potentials, but are presently limited in predicting precise $\text{p}K_{\text{a}}$ values.

As mentioned, mechanisms for $\text{p}K_{\text{a}}$ shifting invoke coupling between proton loading and RNA folding. This coupling can be understood in terms of the simplified conceptual free energy diagram in Figure 7B. Coupling arises because the proton-binding site has too much negative repulsion to be stable without a cation (state F') and too little affinity to bind proton without folding (state FH⁺). A minimal model for coupling between folding and protonation is thus three-state. Because the intermediate is high in energy, it is generally hidden from experimental observation, and the

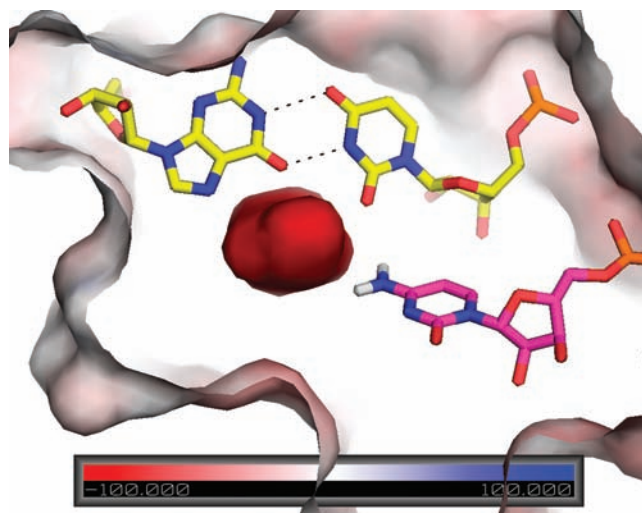


FIGURE 8. Surface electrostatic potential of precleaved HDV ribozyme. Reverse G•U wobble is in yellow and C75 is in magenta. Note the extreme negative potential near the N4 of C75. This potential helps stabilize the positive charge of protonated C75, which delocalizes to N4 according to RESP calculations, and orient the N3 for proton transfer to the leaving group. Reproduced with permission from ref 49. Copyright 2011 American Chemical Society.

system behaves cooperatively. It is therefore customary to treat the system as two-state, with a single $\text{p}K_{\text{a}}$ in the folded state (Figure 7A). Nonetheless, it is conceptually useful to consider $\text{p}K_{a,\text{Fobs}}$ as arising from two terms.

Along path 1 (Figure 7B, solid line), we consider the system protonating with unperturbed $\text{p}K_{a,\text{U}}$ to a high energy state followed by a folding bonus (i.e., negative) of ΔG_{x}

$$\text{p}K_{a,\text{Fobs}} = \text{p}K_{a,\text{U}} - \Delta G_{\text{x}}/1.42, \quad \text{in kcal/mol (at } 37^\circ\text{C)} \quad (2)$$

Thus, $\text{p}K_{a,\text{Fobs}}$ could be raised by interactions upon folding of the protonated state. Along path 2 (Figure 7B, dashed line), we consider the system folding to a high energy state with a repulsive (i.e., positive) free energy ΔG_{rep} followed by protonating with an intrinsic $\text{p}K_{a,\text{Fintr}}$ given by

$$\text{p}K_{a,\text{Fintr}} = \text{p}K_{a,\text{Fobs}} + \Delta G_{\text{rep}}/1.42, \quad \text{in kcal/mol (at } 37^\circ\text{C)} \quad (3)$$

where $\text{p}K_{a,\text{Fintr}}$ is even more basic than $\text{p}K_{\text{Fobs}}$. If there are extensive interactions in the RNA that help it remain folded despite repulsive interactions, then the repulsive state might become the reference state (i.e., F' and F'H⁺ both lowered in energy), and the observed $\text{p}K_{\text{a}}$ could become $\text{p}K_{a,\text{Fintr}}$ and be exceptionally basic. One might envision such a mechanism in a large RNA with extensive

preorganization, such as the ribosome. Coupling of folding to molecular crowding could further favor protonation, as supported by recent studies on the i-motif.⁵⁰

At present, almost nothing is known about coupling between folding and deprotonation of the nucleobases. However, if new interactions are gained upon deprotonation, there will also be a driving force. New interactions could be offered by metal ions or protonated bases, the partially cationic functionalities on the bases, or hydrogen bonding to the 2'OH. Moreover, we note that the unperturbed pK_a 's of G and U start closer to biological pH than do the unperturbed pK_a 's of A and C, meaning less driving force is needed to shift to neutrality.

Concluding Remarks

Nucleobases can have pK_a 's near neutrality that contribute to catalysis. This functional diversity is remarkable when one considers that it arises solely from reversible, noncovalent interactions of functionally bereft monomers. Roles of cationic bases in RNA catalysis seem well-established. At the same time, much remains controversial or unknown in the field: Do negatively charged bases populate and do they contribute to catalysis? Do the bases tautomerize? Are current ribozyme structures catalytically competent? In RNAs with measurable pK_a 's, which bases are ionizing? Are the ionizing bases communicating with each other? Do nucleobases contribute to catalysis in large ribozymes? Do in vivo conditions alter pK_a shifting? Answering these questions requires collaborations between chemists, biologists, crystallographers, spectroscopists, and theorists. Theory will be critical because charged bases often populate poorly and thus remain hidden from many structural and spectroscopic studies.

BIOGRAPHICAL INFORMATION

Jennifer L. Wilcox received her B.S. in chemistry from The University of Scranton in 2007 where she completed undergraduate research in organic geochemistry. She is now a graduate student at The Pennsylvania State University pursuing her Ph.D. in chemistry. Her research focuses on determining the driving forces for pK_a shifting in RNA.

Amarpreet Ahluwalia is an undergraduate student pursuing a B.S. in biochemistry and molecular biology at The Pennsylvania State University in the Schreyer Honors College. Her research interests include identifying protonated bases in RNA through computational and experimental approaches.

Philip C. Bevilacqua received his B.S. in chemistry and physics from John Carroll University in 1987. He obtained his Ph.D. from The University of Rochester in 1993 from Douglas Turner, and was a postdoctoral fellow at the University of Colorado, Boulder with Thomas Cech. He is currently Professor of Chemistry at The

Pennsylvania State University. His research interests include the structure and function of catalytic and regulatory RNA.

This work was supported by NIH Grant GM095923. We thank our collaborators Paul Carey, Barbara Golden, and Sharon Hammes-Schiffer for their helpful comments.

FOOTNOTES

*To whom correspondence should be addressed. Telephone: (814) 863-3812. Fax: (814) 865-2927. E-mail: pcb@chem.psu.edu.

REFERENCES

- 1 *Ribozymes and RNA Catalysis*; Lilley, D. M., Eckstein, F., Eds.; Royal Society of Chemistry: Cambridge, 2008; pp 1–318.
- 2 Ferre-D'Amare, A. R.; Scott, W. G. Small self-cleaving ribozymes. *Cold Spring Harbor Perspect. Biol.* **2010**, a003574.
- 3 Saenger, W. *Principles of Nucleic Acid Structure*; Cantor, C. R., Ed.; Springer-Verlag: New York, 1984; pp 1–556.
- 4 Thomas, J. M.; Yoon, J. K.; Perrin, D. M. Investigation of the catalytic mechanism of a synthetic DNAzyme with protein-like functionality: an RNaseA mimic? *J. Am. Chem. Soc.* **2009**, *131*, 5648–5658.
- 5 Nakano, S.; Chadalavada, D. M.; Bevilacqua, P. C. General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* **2000**, *287*, 1493–1497.
- 6 Legault, P.; Pardi, A. Unusual dynamics and pK_a shift at the active site of a lead-dependent ribozyme. *J. Am. Chem. Soc.* **1997**, *119*, 6621–6628.
- 7 Bevilacqua, P. C.; Brown, T. S.; Nakano, S.; Yajima, R. Catalytic roles for proton transfer and protonation in ribozymes. *Biopolymers* **2004**, *73*, 90–109.
- 8 Gueron, M.; Leroy, J. L. Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.* **1995**, *261*, 383–413.
- 9 Draper, D. E. A guide to ions and RNA structure. *RNA* **2004**, *10*, 335–343.
- 10 Moody, E. M.; Lecomte, J. T.; Bevilacqua, P. C. Linkage between proton binding and folding in RNA: A thermodynamic framework and its experimental application for investigating pK_a shifting. *RNA* **2005**, *11*, 157–172.
- 11 Nagaswamy, U.; Larios-Sanz, M.; Hury, J.; Collins, S.; Zhang, Z.; Zhao, Q.; Fox, G. E. NCIR: a database of non-canonical interactions in known RNA structures. *Nucleic Acids Res.* **2002**, *30*, 395–397.
- 12 Lemieux, S.; Major, F. Automated extraction and classification of RNA tertiary structure cyclic motifs. *Nucleic Acids Res.* **2006**, *34*, 2340–2346.
- 13 Banas, P.; Walter, N. G.; Sponer, J.; Otyepka, M. Protonation states of the key active site residues and structural dynamics of glmS riboswitch as revealed by molecular dynamics. *J. Phys. Chem. B* **2010**, *114*, 8701–8712.
- 14 Mlynsky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H+ forms are most consistent with crystal structures of hairpin ribozyme. *J. Phys. Chem. B* **2010**, *114*, 6642–6652.
- 15 Auffinger, P.; Bielecki, L.; Westhof, E. Anion binding to nucleic acids. *Structure* **2004**, *12*, 379–388.
- 16 Kieft, J. S.; Chase, E.; Costantino, D. A.; Golden, B. L. Identification and characterization of anion binding sites in RNA. *RNA* **2010**, *16*, 1118–1123.
- 17 Rupert, P. B.; Ferre-D'Amare, A. R. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature* **2001**, *410*, 780–786.
- 18 Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferre-D'Amare, A. R. Transition state stabilization by a catalytic RNA. *Science* **2002**, *298*, 1421–1424.
- 19 Bevilacqua, P. C. Mechanistic considerations for general acid-base catalysis by RNA: Revisiting the mechanism of the hairpin ribozyme. *Biochemistry* **2003**, *42*, 2259–2265.
- 20 Luptak, A.; Ferre-D'Amare, A. R.; Zhou, K.; Zilm, K. W.; Doudna, J. A. Direct $pK(a)$ measurement of the active-site cytosine in a genomic hepatitis delta virus ribozyme. *J. Am. Chem. Soc.* **2001**, *123*, 8447–8452.
- 21 Gong, B.; Chen, J. H.; Chase, E.; Chadalavada, D. M.; Yajima, R.; Golden, B. L.; Bevilacqua, P. C.; Carey, P. R. Direct measurement of a $pK(a)$ near neutrality for the catalytic cytosine in the genomic HDV ribozyme using Raman crystallography. *J. Am. Chem. Soc.* **2007**, *129*, 13335–13342.
- 22 Gong, B.; Chen, J. H.; Yajima, R.; Chen, Y.; Chase, E.; Chadalavada, D. M.; Golden, B. L.; Carey, P. R.; Bevilacqua, P. C. Raman crystallography of RNA. *Methods* **2009**, *49*, 101–111.
- 23 Nakano, S.; Bevilacqua, P. C. Mechanistic characterization of the HDV genomic ribozyme: a mutant of the C41 motif provides insight into the positioning and thermodynamic linkage of metal ions and protons. *Biochemistry* **2007**, *46*, 3001–3012.

- 24 Chen, J. H.; Yajima, R.; Chadalavada, D. M.; Chase, E.; Bevilacqua, P. C.; Golden, B. L. A 1.9 Å crystal structure of the HDV ribozyme pre-cleavage suggests both Lewis acid and general acid mechanisms contribute to phosphodiester bond cleavage. *Biochemistry* **2010**, *49*, 6508–6518.
- 25 Brown, T. S.; Chadalavada, D. M.; Bevilacqua, P. C. Design of a highly reactive HDV ribozyme sequence uncovers facilitation of RNA folding by alternative pairings and physiological ionic strength. *J. Mol. Biol.* **2004**, *341*, 695–712.
- 26 Guo, M.; Spitale, R. C.; Volpini, R.; Krucinska, J.; Cristalli, G.; Carey, P. R.; Wedekind, J. E. Direct Raman measurement of an elevated base pK(a) in the active site of a small ribozyme in a precatalytic conformation. *J. Am. Chem. Soc.* **2009**, *131*, 12908–12909.
- 27 Cottrell, J. W.; Scott, L. G.; Fedor, M. J. The pH dependence of hairpin ribozyme catalysis reflects ionization of an active site adenine. *J. Biol. Chem.* **2011**, *286*, 17658–17664.
- 28 Liu, L.; Cottrell, J. W.; Scott, L. G.; Fedor, M. J. Direct measurement of the ionization state of an essential guanine in the hairpin ribozyme. *Nat. Chem. Biol.* **2009**, *5*, 351–357.
- 29 Wilson, T. J.; Lilley, D. M. Do the hairpin and VS ribozymes share a common catalytic mechanism based on general acid-base catalysis? A critical assessment of available experimental data. *RNA* **2011**, *17*, 213–221.
- 30 Kuzmin, Y. I.; Da Costa, C. P.; Fedor, M. J. Role of an active site guanine in hairpin ribozyme catalysis probed by exogenous nucleobase rescue. *J. Mol. Biol.* **2004**, *340*, 233–251.
- 31 Moody, E. M.; Brown, T. S.; Bevilacqua, P. C. Simple method for determining nucleobase pKa values by indirect labeling and demonstration of a pKa of neutrality in dsDNA. *J. Am. Chem. Soc.* **2004**, *126*, 10200–10201.
- 32 Cerrone-Szakal, A. L.; Siegfried, N. A.; Bevilacqua, P. C. Mechanistic characterization of the HDV genomic ribozyme: Solvent isotope effects and proton inventories in the absence of divalent metal ions support C75 as the general acid. *J. Am. Chem. Soc.* **2008**, *130*, 14504–14520.
- 33 Siegfried, N. A.; O'Hare, B.; Bevilacqua, P. C. Driving forces for nucleic acid pK(a) shifting in an A(+)C wobble: effects of helix position, temperature, and ionic strength. *Biochemistry* **2010**, *49*, 3225–3236.
- 34 Muth, G. W.; Ortoleva-Donnelly, L.; Strobel, S. A. A single adenosine with a neutral pKa in the ribosomal peptidyl transferase center. *Science* **2000**, *289*, 947–950.
- 35 Nakano, S.; Proctor, D. J.; Bevilacqua, P. C. Mechanistic characterization of the HDV genomic ribozyme: assessing the catalytic and structural contributions of divalent metal ions within a multichannel reaction mechanism. *Biochemistry* **2001**, *40*, 12022–12038.
- 36 Nakano, S.; Cerrone, A. L.; Bevilacqua, P. C. Mechanistic characterization of the HDV genomic ribozyme: classifying the catalytic and structural metal ion sites within a multichannel reaction mechanism. *Biochemistry* **2003**, *42*, 2982–2994.
- 37 Knitt, D. S.; Herschlag, D. pH dependencies of the Tetrahymena ribozyme reveal an unconventional origin of an apparent pKa. *Biochemistry* **1996**, *35*, 1560–1570.
- 38 Veeraraghavan, N.; Bevilacqua, P. C.; Hammes-Schiffer, S. Long distance communication in the HDV ribozyme: Insights from molecular dynamics and experiments. *J. Mol. Biol.* **2010**, *402*, 278–291.
- 39 Das, S. R.; Fong, R.; Piccirilli, J. A. Nucleotide analogues to investigate RNA structure and function. *Curr. Opin. Chem. Biol.* **2005**, *9*, 585–593.
- 40 Suydam, I. T.; Levandoski, S. D.; Strobel, S. A. Catalytic importance of a protonated adenosine in the hairpin ribozyme active site. *Biochemistry* **2010**, *49*, 3723–3732.
- 41 Frederiksen, J. K.; Piccirilli, J. A. Identification of catalytic metal ion ligands in ribozymes. *Methods* **2009**, *49*, 148–166.
- 42 Bevilacqua, P. C.; Yajima, R. Nucleobase catalysis in ribozyme mechanism. *Curr. Opin. Chem. Biol.* **2006**, *10*, 455–464.
- 43 Das, S. R.; Piccirilli, J. A. General acid catalysis by the hepatitis delta virus ribozyme. *Nat. Chem. Biol.* **2005**, *1*, 45–52.
- 44 Wilson, T. J.; Li, N. S.; Lu, J.; Frederiksen, J. K.; Piccirilli, J. A.; Lilley, D. M. Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11751–11756.
- 45 Perrotta, A. T.; Wadkins, T. S.; Been, M. D. Chemical rescue, multiple ionizable groups, and general acid–base catalysis in the HDV genomic ribozyme. *RNA* **2006**, *12*, 1282–1291.
- 46 Raines, R. T.; Ribonuclease, A. *Chem. Rev.* **1998**, *98*, 1045–1065.
- 47 Zhou, J. M.; Zhou, D. M.; Takagi, Y.; Kasai, Y.; Inoue, A.; Baba, T.; Taira, K. Existence of efficient divalent metal ion-catalyzed and inefficient divalent metal ion-independent channels in reactions catalyzed by a hammerhead ribozyme. *Nucleic Acids Res.* **2002**, *30*, 2374–2382.
- 48 Tang, C. L.; Alexov, E.; Pyle, A. M.; Honig, B. Calculation of pK(a)s in RNA: on the structural origins and functional roles of protonated nucleotides. *J. Mol. Biol.* **2007**, *366*, 1475–1496.
- 49 Veeraraghavan, N.; Ganguly, A.; Chen, J. H.; Bevilacqua, P. C.; Hammes-Schiffer, S.; Golden, B. L. Metal binding motif in the active site of the HDV ribozyme binds divalent and monovalent ions. *Biochemistry* **2011**, *50*, 2672–2682.
- 50 Rajendran, A.; Nakano, S.; Sugimoto, N. Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* **2010**, *46*, 1299–1301.