Direct $pK_a$ Measurement of the Active-Site Cytosine in a Genomic Hepatitis Delta Virus Ribozyme

Andrej Lupták, † Adrian R. Ferré-D’Amaré, § Kaihong Zhou, ‡ Kurt W. Zilm, † and Jennifer A. Doudna* ,§, ‡

Contribution from the Departments of Chemistry and Molecular Biophysics and Biochemistry, and Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520

Received April 26, 2001. Revised Manuscript Received June 19, 2001

Abstract: Hepatitis delta virus ribozymes have been proposed to perform self-cleavage via a general acid/base mechanism involving an active-site cytosine, based on evidence from both a crystal structure of the cleavage product and kinetic measurements. To determine whether this cytosine (C75) in the genomic ribozyme has an altered $pK_a$ consistent with its role as a general acid or base, we used $^{13}$C NMR to determine its microscopic $pK_a$ in the product form of the ribozyme. The measured $pK_a$ is moderately shifted from that of a free nucleoside or a base-paired cytosine and has the same divalent metal ion dependence as the apparent reaction $pK_a$’s measured kinetically. However, under all conditions tested, the microscopic $pK_a$ is lower than the apparent reaction $pK_a$, supporting a model in which C75 is deprotonated in the product form of the ribozyme at physiological pH. While additional results suggest that the $pK_a$ is not shifted in the reactant state of the ribozyme, these data cannot rule out elevation of the C75 $pK_a$ in an intermediate state of the transesterification reaction.

Introduction

During rolling-circle replication of the hepatitis delta virus (HDV) RNA, genome-encoded self-cleaving ribozymes produce unit length genomic and anti-genomic RNA products (reviewed by Lai1). Two structurally homologous ribozymes, one in the positive strand and one in the negative strand of the virus, carry out site-specific strand scission with a minimal requirement for divalent metal ions.2,3 A crystal structure of the self-cleaved form of the genomic ribozyme revealed a cytosine (C75) residue proximal to the 5′ hydroxyl leaving group of the transesterification reaction.4 This led to a hypothesis that C75 acts as a general acid or base in the reaction,5-6 having a role similar to that of the catalytic histidines in RNase A (see Figure 1 for a general mechanism).

Recent investigation of the reaction kinetics of both the genomic and anti-genomic HDV ribozymes has provided strong evidence that C75 (C76 in the antigenomic ribozyme) is directly involved in the transesterification reaction. While mutation of this cytosine to G or U inactivates both ribozymes, substitution


10.1021/ja016091x CCC: $20.00 © 2001 American Chemical Society
Published on Web 08/07/2001
Figure 1. Proposed mechanism of general acid/base-catalyzed cleavage of a phosphodiester bond in RNA. B denotes general base which deprotonates the nucleophile. In a related proposal this base is specific. A denotes the general acid which protonates the 5′ leaving group. The transition state is assumed, and no hypothesis about the protonation of the dianion or any intermediates is made here.

of an A at this site supports self-cleavage, albeit at a rate 2–3 orders of magnitude slower than that of the wild-type ribozymes. Comparison of the kinetic pH profiles of the wild-type ribozymes and the C75A/C76A point mutants showed a difference in the apparent pKₐ of the reaction by approximately the same amount as in free C versus A nucleosides (~0.7 pH unit). In addition, the activity of C76U and C76 deletion forms of the T7 RNA polymerase center of the ribosome. 20,24

To measure the pKₐ of C75 in the active site of the genomic HDV ribozyme directly, we designed a trans-acting construct analogous to one previously described (Figure 2a). This allowed site-specific ¹³C,¹⁵N-labeling of individual residues. We show that the in situ pKₐ of C75 in the product form of the HDV ribozyme, as calculated from ¹³C chemical shifts of the base carbons, is only modestly shifted from that of free cytidine. Mechanistic implications of this finding are discussed.

Materials and Methods

Preparation of Site-Specifically ¹³C,¹⁵N-Labeled HDV Ribozymes. The two RNAs that comprise the NMR construct were transcribed in vitro. The 5′ RNA was transcribed from a Bsal-linearized pUC19-derived plasmid containing the T7 promoter and the ribozyme coding sequence, using a standard protocol for in vitro transcription;26 the 3′ RNA was transcribed separately from a synthetic DNA duplex in which the last two coding nucleotides were 2′-methoxylated to prevent addition of untemplated nucleotides to the transcript. 27 The 3′ RNA was transcribed with a limiting ¹³C,¹⁵N-labeled CTP (Mertak Biosciences, Cambridge Isotope Laboratories). The initial CTP concentration was 250 μM, while the concentration of ATP, GTP, and UTP was 5 mM each. The incorporation of the labeled CTP during transcription was monitored using anion-exchange chromatography (Dionex NucleoPac PA-100), and the reaction was terminated when no free CTP could be detected. Both RNAs were purified using preparative-scale denaturing polyacrylamide gel electrophoresis. After electrophoresis and concentration, the RNAs were assembled by heating to 65 °C for 5 min and annealing at 37 °C for 1 h in 10 mM MgCl₂, 100 mM LiCl, and 10 mM potassium cacodylate, pH 7. Following self-cleavage, the product complex was purified away from the precursor and the individual RNAs using preparative size-exclusion chromatography (Superdex 75; Pharmacia Biotech, Uppsala, Sweden) in the presence of 100 mM LiCl, 1 mM MgCl₂, and 10 mM KH₂PO₄/K₂HPO₄, pH 7. The purity of the complex was confirmed using analytical denaturing gel electrophoresis (Figure 2b). The sample was concentrated and exchanged into the final


The 1H-decoupled 13C spectra were recorded on a Varian Inova 500 spectrometer at 125.885 MHz using a 13C direct-detect probe with 2.5 s 1H NOE presaturation. The spectra were recorded at 37 °C with 7168 complex points with a sweep width of 18 002 Hz and the carrier at 109.3 ppm. Typically, from 500 to several thousand scans were recorded to obtain an adequate signal-to-noise ratio.

NMR Spectroscopy. The 1H-decoupled 13C spectra were recorded on a Varian Inova 500 spectrometer at 125.885 MHz using a 13C direct-detect probe with 2.5 s 1H NOE presaturation. The spectra were recorded at 37 °C with 7168 complex points with a sweep width of 18 002 Hz and the carrier at 109.3 ppm. Typically, from 500 to several thousand scans were recorded to obtain an adequate signal-to-noise ratio.

Determination of pKₐ of C75. The pKₐ's were determined from 13C chemical shifts using the equation

\[
\delta_{\text{observed}} = \frac{\delta_{\text{protonated}} + \delta_{\text{deprotonated}} \times 10^{(pK_a - \delta_{\text{observed}})}}{1 + 10^{(pK_a - \delta_{\text{observed}})}}
\]

where \(\delta_{\text{protonated}}\) and \(\delta_{\text{deprotonated}}\) are the chemical shifts of protonated and deprotonated states, respectively. A more convenient form of the equation for fitting our experimental data is obtained by essentially referencing the chemical shifts of each observed carbon to their high pH value. The equation then becomes

\[
\Delta \delta = \delta_{\text{observed}} - \delta_{\text{deprotonated}} = \frac{\delta_{\text{protonated}} - \delta_{\text{deprotonated}}}{1 + 10^{(pK_a - \delta_{\text{observed}})}} = \frac{\delta_{\text{max}}}{1 + 10^{(pK_a - \delta_{\text{observed}})}}
\]

where \(\delta_{\text{max}}\) is the full change in chemical shift of the carbon resonance upon protonation. When full protonation of the titrated residue is not observed, both pKₐ and \(\delta_{\text{max}}\) have to be estimated from the available data. In most of our titrations, unambiguous 13C spectra could only be observed up to about the midpoint of the titration. In the cases where only partial protonation of the base was observed, estimation of both pKₐ and \(\delta_{\text{max}}\) resulted in unrealistically high values of \(\delta_{\text{max}}\). In such cases the pKₐ was calculated while assuming the value of \(\delta_{\text{max}}\) to be the same as that in free nucleoside monophosphate, as was observed in a lead-dependent ribozyme. A similar \(\delta_{\text{max}}\) has been reported for cytosine in a CACACA hexanucleotide in D₂O, although an about 20% lower \(\delta_{\text{max}}\) is observed with the same molecule in dimethyl sulfoxide. Somewhat lower \(\delta_{\text{max}}\) values have been reported for adenosines with unusually shifted pKₐ's in the hairpin ribozyme, as compared with a lead-dependent ribozyme, although full deprotonation was not observed. Lowering \(\delta_{\text{max}}\) increases the estimated pKₐ of the titrated base. The effect of decreased \(\delta_{\text{max}}\) on the pKₐ of C75 is reflected in Table 1.

Results

Design and Preparation of the Construct. In the two-piece RNA construct used in this study, the 5' RNA contains a leader sequence followed by the first 52 nucleotides of the ribozyme in which the P4 helix is shortened and the loop L4 left out in order to allow for a break in the backbone (black letters in Figure 2a). The 3' RNA is a 23-nucleotide oligomer that base-pairs with the 5' strand in regions P2 and P4 to form an active ribozyme (gray in Figure 2a). The G10-C85 base pair (wild-type genomic ribozyme numbering) is inverted, leaving the 3' RNA with only two cytosines. This allows for selective incorporation of 13C,15N-labeled cytosines at the active site.
could the sigmoidal shape of the titration curve be observed. In 2001, Biol.
556.
samples were prepared in 1 mM MgCl2, 100 mM LiCl, and the self-cleaved complex was purified as described above. The samples were prepared in 1 mM MgCl2, 100 mM LiCl, and 10% D2O. The HDV ribozyme requires the presence of divalent metal ions to maintain its tertiary fold.30–32 lithium was chosen to prevent formation of alternative structures by the G-rich 3′ RNA strand.

**NMR Spectra of the Self-Cleaved Form of the Trans HDV Ribozyme.** Initial 1H NMR characterization of the self-cleaved form of the ribozyme revealed a well-dispersed imino region of the spectrum with about 20 distinguishable peaks (Figure 1S, Supporting Information). No proton resonance in the 15 ppm region was observed at any pH or temperature, indicating that solvent exchange with any potential imino protons of unusually basic cytosines was fast on the NMR time scale. However, two amino resonances with a pH- and temperature-dependent chemical shift were observed (pKₐ ~7–7.5). Specific isotopic labeling of the C75 confirmed that neither of these 1H resonances belongs to the amino group of C75. In the absence of full assignment of the molecule, we hypothesize that the observed pH- and temperature-dependent spectral change resulted from a structural rearrangement, most likely in the base quadruple involving protonated C41,33,34

We were able to determine the pKₐ of C75 using indirect detection of the base protonation through base carbon chemical shift measurements. The downfield 13C spectral region of the constructs used in this study is presented in Figure 3. This region contains resonances corresponding to carbons 2, 4, and 6 on the cytosine ring. Positions 2 and 4 are not protonated and exhibit the highest chemical shift change upon protonation of the N3 imino nitrogen.28,29 Figure 3a shows the 13C spectra of cytosines 75 and 84 at acidic to neutral pH’s. The spectra contain two peaks for each carbon type, one of which shows a more dramatic chemical shift change as a function of pH (indicated by the solid line in Figure 3a). Comparison with the spectrum of the C84U mutant (Figure 3b) demonstrates that these shifting peaks correspond to the C75 carbon resonances. The resonances of C84 (indicated by the dotted line in Figure 3a) exhibit only a slight change of chemical shift. This nucleotide is involved in a Watson–Crick base-pair interaction,4 which stabilizes the deprotonated state of the N3 of cytosine.

**Measurement of the pKₐ of the Active-Site Cytosine.** Figure 4 shows the titration curves of the C2 carbons on the 13C,15N-labeled cytosines under various salt conditions. The pKₐ values were calculated by fitting the chemical shift data to eq 2, with pKₐ as a variable. In only one case, in the presence of calcium, could the sigmoidal shape of the titration curve be observed. In that case, both δmax and pKₐ were used as variables to fit the

![Figure 3. Carbon-13 NMR spectra of the selectively labeled HDV ribozymes. (a) C75- and C84-labeled complex at pH 4.70, 5.25, 5.65, 6.58, and 7.08. (b) Complex with C84U mutation showing a single 13C resonance for each carbon type at pH 7.44. Smaller peaks correspond to the natural abundance 13C spectrum of the 74 unlabeled nucleotides. Dotted lines and solid lines connect resonances of carbons at positions 2, 4, and 6 of C84 and C75, respectively. All spectra were collected at 37 °C and processed with 20 Hz line broadening.](image)

![Figure 4. Protonation of C75. The chemical shifts of the C2 carbons were referenced to the highest pH measured and to the maximum chemical shift, δmax, used to fit the data (see Table 1). Sample conditions were as follows: 0.7 mM MgCl2, 70 mM LiCl (circles); 3 mM MgCl2, 40 mM LiCl (triangles); 0.7 mM CaCl2, 70 mM LiCl (squares). Lines represent the best-fit curves generated using eq 2. Dashed line is a theoretical curve based on kinetic data representing the protonation of C75 in the wild-type genomic ribozyme in 0.87 mM MgCl2 at 37 °C.](image)
Discussion and Conclusions

In this study, we used a 13C-labeled form of the genomic HDV ribozyme to determine the pKₐ of an active-site cytosine proposed to be involved in general acid/base catalysis during ribozyme self-cleavage. In order for C75 to be an efficient general acid/base catalyst in the phosphodiester cleavage reaction, its pKₐ should be near 7 at some stage during ribozyme self-cleavage to enable it to accept or donate a proton.

13C NMR is perhaps the most sensitive method to measure the protonation state of specific functional groups in macromolecules. While 1H NMR can be used to detect a titrating proton directly, usually in combination with indirect 15N chemical shift observation, this is only possible under a slow-exchange regime and in cases where the 15N relaxation is slow enough to allow for efficient magnetization transfer. Direct detection of the heteronuclei on the titratable nucleobase circumvents the problem of rapid relaxation in large molecules and rapid solvent exchange. For this purpose, 13C detection is superior to 15N observation, given the higher receptivity of 13C and the large chemical shift range for the base carbons that are directly bonded to the titrated imino groups.

The sensitivity of these 13C chemical shifts to pH makes this method an especially informative probe of the state of protonation of the site of interest.

13C spectra of the site-specifically labeled HDV RNA show that, in the product state of the reaction, the pKₐ of C75 is only moderately shifted from that of the free nucleoside. The pKₐ estimated from the 13C chemical shifts of the C2 and C4 carbons of the C75 base is about 4.8 when measured in 0.7 mM Mg²⁺ and 70 mM Li⁺, 0.6 unit above the pKₐ of free cytidine. The pKₐ of the same group measured at 3 mM Mg²⁺ is not shifted from that of free nucleoside at all, while in the presence of 0.7 mM Ca²⁺ it is slightly higher (5.4).

As a control experiment, titration of 1 mM cytidine nucleoside was carried out in samples having between 0 and 3 mM MgCl₂ in the background of 70 mM LiCl. Using the chemical shifts of the H5 and H6 protons, the pKₐ of the base was found to be independent of the divalent metal ion concentration under these conditions. Our results show that, in the product state of the ribozyme, C75 is fully deprotonated close to neutral pH.

The general trend of the microscopic pKₐ shifts of C75 follows the same trend in apparent reaction pKₐ's: the pKₐ shifts downward with an increasing concentration of Mg²⁺ and upward in the presence of Ca²⁺ (Table 1). However, the microscopic pKₐ's measured in the self-cleaved form of our construct are about 2 units lower than the apparent pKₐ's calculated from the kinetic pH profiles of the cis-acting wild-type genomic ribozyme under similar conditions (Figure 4).

Our data indicate that if C75 is involved in general acid/base catalysis, its pKₐ must be higher in the reactant or intermediate state of the cleavage reaction. We have attempted to measure the pKₐ of the C75 in the uncleaved form of the ribozyme. While we were able to purify the assembled trans complex in the presence of EDTA and cobalt hexammine (an inhibitor of the cleavage reaction), the lack of well-dispersed 1H spectra indicated that the complex was not in a unique conformation and, close to neutral pH, the C2 and C4 resonances of the specifically labeled C75 had chemical shifts similar to those of the deprotonated cytosine in the self-cleaved ribozyme (Figure 3S, Supporting Information). The lack of unique resonances corresponding to the base carbons precluded a pKₐ analysis of this form of the ribozyme, although we estimate that the pKₐ is less than 5 (in 1 mM EDTA, 1.5 mM cobalt hexammine, and 150 mM LiCl). A related construct of the ribozyme, which contained a 5’ monophosphate at the cleavage site, exhibited a well-dispersed 1H spectrum similar to the 1H spectrum of the cleaved form of the ribozyme. The 13C spectra, on the other hand, showed multiple peaks for each type of carbon on the cytosine ring, barring pKₐ analysis in this construct as well. As in the case of the uncleaved form of the ribozyme, near neutral pH the 13C chemical shifts were similar to those of a deprotonated cytidine (in 0.9 mM MgCl₂ and 90 mM LiCl).

This suggests that the C75 may not have an unusually shifted pKₐ in the precursor state of the ribozyme.

To understand these results in light of the kinetic data published by other groups requires consideration of the general mechanism of RNA cleavage (Figure 1). Breaker and co-workers have investigated the influence of the 2’ nucleophile on the rate of phosphodiester cleavage and have concluded that the maximum rate of cleavage with the fully deprotonated 2’ hydroxyl group is about 0.01 min⁻¹. In addition, when the conformation of the 2’ hydroxyl and the reactive phosphate group is close to the theoretical geometry of an in-line attack, the rate of cleavage is accelerated by about 10-fold. The observed maximum rate of cleavage of the HDV ribozymes is about 10⁻¹⁻¹, about 100-fold higher than the maximum theoretical rate resulting from the nucleophilic attack by the 2’ oxyanion group. This means that further acceleration of transesterification must come from stabilization of the transition state and/or stabilization of the leaving group. The general acid/base mechanisms proposed previously have suggested stabilization by protonation of the 5’ oxyanion leaving group as one of the steps involved in the catalysis, although protonation of an anionic phosphorane intermediate related to the transition state cannot be ruled out.

There are three basic possibilities that would explain the equilibrium data presented in this paper in view of the available kinetic data for the HDV ribozymes: (1) if the observed reaction pKₐ's reflect a rate-limiting step other than chemistry, C75/C76 may not be a general acid or base in the transesterification reaction; (2) if none of our precursor and precursor analogue samples had properly preorganized active sites, the pKₐ of C75 immediately before the reaction may in fact be significantly shifted, reflecting the apparent reaction pKₐ; (3) if our precursor NMR samples had active sites resembling the correct precleaved state, the apparent reaction pKₐ may reflect transient protonation of C75 during an intermediate step of the reaction. We now consider the evidence for or against each of these scenarios.

The apparent reaction pKₐ's reflect the pKₐ of an active-site group only if the rate-limiting step of the reaction represents a chemical step in the transesterification. The wealth of HDV ribozyme kinetic data indicates that during the observed step of catalysis, there is at least one, and in most cases only one, proton transferred, in contrast to the six proton transfers hypothesized for an analogous reaction catalyzed by RNase A. Proton inventory and kinetic solvent isotope effects on the cleavage reaction support a single-proton transfer in the ratelimiting step of the reaction. The ΔpKₐ for the self-cleavage reaction in D₂O versus H₂O was about 0.4 and 0.7 for the genomic and anti-genomic ribozymes, respectively, while the ΔpKₐ of the free cytidine is 0.53 (Figure 2S, Supporting Information). While these effects demonstrate a general proton...
transfer, they could reflect a structural rearrangement involving formation or breakage of a hydrogen bond during the reaction, rather than reaction chemistry. However, imidazole rescue experiments performed with a variant of the antigenomic ribozyme lacking C76 retained the kinetic solvent isotope effect observed with the wild-type ribozyme, making a structural rearrangement involving C76 unlikely. Furthermore, the apparent pK_a's for the reaction have been shown to be the same for trans cleavage of either a 3'-5' or 2'-5' phosphodiester in the anti-genomic ribozyme, even though the observed rates of cleavage were substantially different. These pK_a's are near the pK_a's measured in the wild-type cis ribozymes. These results, and the results obtained with the C75A/C76A mutants described above, strongly suggest that C75/C76 is directly involved in a rate-limiting chemical step in the transesterification, as opposed to a folding transition.

Since our 13C NMR measurements demonstrate that the pK_a of C75 is only moderately shifted in the product state, the apparent reaction pK_a must correspond to a microscopic pK_a either in the precursor active site or in a transition-state-related intermediate. In the reactant state, we detected no protonation of C75 near neutral pH in three different constructs with active sites containing a 5' phosphate, a 5' phosphate with an upstream deoxynucleotide, or a 5' phosphate with an upstream ribonucleotide (inhibited from self-cleaving by cobalt hexammine). The lack of protonation might result from these constructs being improperly folded. However, this is unlikely, since they were assembled under conditions similar to those in which the all RNA construct assembles and is catalytically active; and the imino regions of the proton spectra of the 5'phosphate- and 5'-deoxyribonucleotide-containing constructs were well dispersed throughout the pH titration. Thus, it appears that the pK_a of C75 in the reactant state is not significantly shifted. Therefore, the unusual shift in the pK_a of C75 obtained from kinetic pH profiles of the transesterification reaction likely demonstrates a rate-limiting, short-lived pK_a shift more closely associated with the phosphorane intermediate, related to the trigonal bipyramidal transition state of the reaction (Figure 1).

Acknowledgment. We thank Donald Crothers, Ronald Breaker, Robert Batey, Jeffrey Kieft, Jeremy Murray, Scott Strobel, and Adegboyega Oyelere for useful discussions; Melanie Cocco, R. Batey, and J. Kieft for help with NMR; and A. Oyelere for synthesis of dArG dinucleotide. This work was supported by a grant from the NIH (J.A.D.). The authors acknowledge generous support of the high-field NMR facilities by the W. M. Keck Foundation.

Supporting Information Available: Figure 1S, imino region of the 1H spectrum of the self-cleaved ribozyme; Figure 2S, deuterium isotope effect on pK_a of cytidine; Figure 3S, down-field region of the 13C spectrum of the uncleaved form of the ribozyme at pH 6.2 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.