



## Determination of hepatitis delta virus ribozyme N(–1) nucleobase and functional group specificity using internal competition kinetics



Daniel L. Kellerman<sup>a</sup>, Kandice S. Simmons<sup>a</sup>, Mayra Pedraza<sup>a</sup>, Joseph A. Piccirilli<sup>b</sup>, Darrin M. York<sup>c</sup>, Michael E. Harris<sup>a,\*</sup>

<sup>a</sup> Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

<sup>b</sup> Department of Chemistry and Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA

<sup>c</sup> Center for Integrative Proteomics Research, BioMaPS Institute for Quantitative Biology and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA

### ARTICLE INFO

#### Article history:

Received 18 February 2015

Received in revised form 17 April 2015

Accepted 23 April 2015

Available online 1 May 2015

#### Keywords:

Enzyme kinetics

Ribozyme

Internal competition

### ABSTRACT

Biological catalysis involves interactions distant from the site of chemistry that can position the substrate for reaction. Catalysis of RNA 2'-O-transphosphorylation by the hepatitis delta virus (HDV) ribozyme is sensitive to the identity of the N(–1) nucleotide flanking the reactive phosphoryl group. However, the interactions that affect the conformation of this position, and in turn the 2'O nucleophile, are unclear. Here, we describe the application of multiple substrate internal competition kinetic analyses to understand how the N(–1) nucleobase contributes to HDV catalysis and test the utility of this approach for RNA structure–function studies. Internal competition reactions containing all four substrate sequence variants at the N(–1) position in reactions using ribozyme active site mutations at A77 and A78 were used to test a proposed base-pairing interaction. Mutants A78U, A78G, and A79G retain significant catalytic activity but do not alter the specificity for the N(–1) nucleobase. Effects of nucleobase analog substitutions at N(–1) indicate that U is preferred due to the ability to donate an H-bond in the Watson–Crick face and avoid minor groove steric clash. The results provide information essential for evaluating models of the HDV active site and illustrate multiple substrate kinetic analyses as a practical approach for characterizing structure–function relationships in RNA reactions.

Published by Elsevier Inc.

Enzyme catalysis is due to selective stabilization of the transition state and can result from interactions that are both proximal and distal from the site of bond making and breaking [1]. Understanding how substrates are positioned for reaction on an enzyme, therefore, is essential for understanding the origin of its catalytic power. The hepatitis delta virus (HDV)<sup>1</sup> ribozyme catalyzes self-cleavage via RNA 2'-O-transphosphorylation and is an important experimental system for understanding mechanisms of RNA catalysis [2]. In addition, because the enzymatic activity of the HDV ribozyme is required for the replication of the human pathogen, the hepatitis delta virus, this catalytic RNA is a potential target for the development of antiviral therapies. Current models of the HDV catalytic mechanism involve nucleobase-mediated leaving group stabilization by general acid catalysis and metal ion

catalysis of nucleophile activation [3]. Information on interactions that position the nucleotides flanking the reactive phosphoryl group comes from high-resolution structures and biochemical data. The conserved G residue at N(1) that contains the 5'O leaving group is base-paired to U in the ribozyme active site [4,5]; however, interactions with the N(–1) nucleobase that contains the 2'O nucleophile are less clear. This site is absent from product structures of the HDV ribozyme, the resolution of precursor models at this position are low, and current models of the active site are based on the structure of the hammerhead ribozyme [3,5]. Understanding whether, or how, the HDV ribozyme positions the N(–1) nucleobase is important because phosphoryl transfer reactions require in-line geometry of nucleophile and leaving group [6].

Although most small catalytic RNAs, including the hairpin, hammerhead, and Varkut satellite (VS) ribozymes, use base-pairing interactions upstream and downstream of the cleavage site in order to orient the substrate for in-line attack, in the HDV ribozyme the cleavage site is flanked immediately upstream by a single-stranded 5' sequence. Only a single nucleotide upstream of the cleavage site, N(–1), is required for catalysis, and

\* Corresponding author. Fax: +1 216 368 3419.

E-mail address: [meh2@cwru.edu](mailto:meh2@cwru.edu) (M.E. Harris).

<sup>1</sup> Abbreviations used: HDV, hepatitis delta virus; PAGE, polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; 2AP, 2-aminopurine; DAP, 2,6-diaminopurine.

this 5' sequence has previously been shown to lower the energetic barrier for HDV ribozyme catalysis and drive the reaction forward via ground state destabilization [7]. However, the precise molecular interactions that contribute to this effect are unknown, and current crystal structures show significant disorder in this region. Direct metal ion coordination to the N(-1) nucleobase is not likely because the divalent metal ion concentration dependence is identical for wild-type substrates and substrates with an abasic residue at N(-1) [7]. Therefore, it is plausible that interactions between the N(-1) nucleobase and additional nucleotides of the ribozyme backbone may be necessary for positioning the nucleophile for catalysis.

Dissecting the structure–function relationships important for RNA catalysis depends, in large part, on comparing the effects of changes in nucleobase and backbone structure on reaction rate and equilibrium constants [8–11]. For RNA processing and RNA catalysis reactions *in vitro*, activity is typically measured by scoring the change in the concentration of a single substrate or product over time and fitting these data to the appropriate rate equation [12]. A powerful example is the use of double mutant cycles, in which the interdependence of the effects of mutation or modification at sites on the substrate and enzyme is analyzed [13,14]. Although direct fitting of time courses is usually preferable, its application can be limited due to differences in enzyme activity and reaction conditions between assays and when absolute concentrations of substrate or product are difficult to quantify.

Internal competition is an alternative method for quantifying relative enzyme rate constants that involves analyzing the change in the alternative substrates or product ratios as a function of reaction progress in reactions containing multiple substrates [15–18]. In some cases, this approach can complement analyses by direct fitting of kinetic data by providing higher precision, increased throughput, and less sensitivity to variation in enzyme specific activity [15,16,19]. Internal competition kinetics has been used extensively to measure kinetic isotope effects [18,20], and substrate and product ratios have been measured using a wide range of analytical methods, including mass spectrometry [21,22], nuclear magnetic resonance [23,24], and radioactive remote labeling [25–27]. Recently, we described the application of alternative substrate kinetics to simultaneously determine the relative rate constants for multiple pre-tRNA (transfer RNA) substrates for RNase P using both polyacrylamide gel electrophoresis (PAGE) and Illumina sequencing analyses to quantify substrate ratios [28].

Here, we describe the application of multiple substrate internal competition kinetics to determine relative rate constants for HDV ribozyme reactions containing multiple alternative oligonucleotide substrates in order to understand the nucleobase specificity at the N(-1) position. Analysis of precursor ratios for internal competition kinetics using standard molecular biology methods provides accurate rate constants relative to analyses of individual time courses. Importantly, the results show that accuracy depends on accounting for large differences in intrinsic rate constants for different substrate variants, differences in reaction rate constants due to sequence length, and background levels of unreacted substrates. General strategies for accounting for and minimizing these effects are discussed.

This method was used to analyze reactions containing substrate sequence variants at the N(-1) position together with ribozyme active site mutations at A77 and A78 in order to test a proposed base-pairing interaction between these residues and N(-1). The results from experiments comparing the effects on nucleotide analog substitutions at N(-1) indicate that both steric clash in the minor groove and H-bonding in the Watson–Crick binding face are responsible for the observed specificity for uridine residues at N(-1). These results validate multiple substrate internal competition as a general approach for analyzing RNA structure–function

relationships and provide benchmarks for evaluating models of the contribution of the N(-1) nucleobase to catalysis.

## Materials and methods

### Preparation of HDV ribozyme and substrate RNA

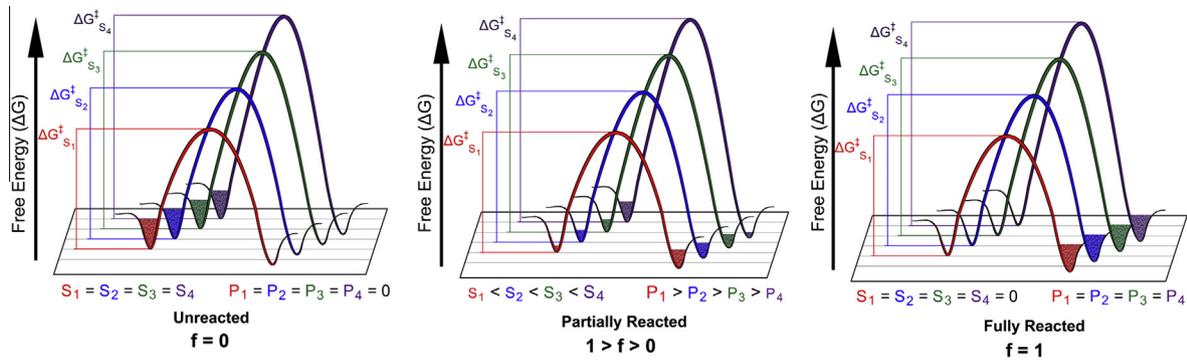
The wild-type HDV ribozyme and site-specific mutant ribozymes were prepared by T7 RNA polymerase transcription from cloned plasmid DNA templates (see Table S1 in online Supplementary material) containing the required promoter sequence using standard procedures [29]. The trans-cleaving version of the HDV ribozyme used in these studies is derived from the antigenomic sequence and contains a shortened P4 stem and a discontinuous J1/2 region (see Fig. 2A in Results and Discussion) [30]. Briefly, 96-bp DNA duplexes containing a 5' EcoRI overhang, a 3' BamHI overhang, the antigenomic HDV ribozyme sequence driven by a T7 promoter, and a BbsI site positioned to cut 3' of the last nucleotide of the HDV ribozyme sequence on the template strand were designed and commercially synthesized (IDT). This construct was inserted into the pUC19 plasmid and cloned into NEB 5-alpha competent *Escherichia coli* cells. HDV ribozyme–pUC19 plasmid was purified using the Qiagen Plasmid Plus Maxi Kit. Mutations at A78 and A78 were introduced by polymerase chain reaction from purified HDV ribozyme–pUC19 plasmid using reverse primers designed to introduce site-specific mutations [31]. These mutant constructs were inserted into the pUC19 plasmid, cloned into NEB 5-alpha competent *E. coli* cells, and purified using the Qiagen Plasmid Plus Maxi Kit.

Purified plasmid was prepared for runoff *in vitro* transcription by restriction digest with BbsI to generate a linear DNA template. Transcription reactions contained 40 mM Tris–HCl (pH 8.0), 20 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 2 mM spermidine, 4 mM each ribonucleoside triphosphate (rNTP), 20 μl inorganic pyrophosphatase (0.01 U/μl), and 10 μl T7 RNA polymerase (10 U/μl) and were incubated at 37 °C for 8 h. The enzyme was removed by phenol–chloroform extraction, and full-length ribozyme was purified by 6% denaturing PAGE and subsequent ultraviolet (UV) shadowing. Excised gel slices were eluted in TE buffer containing 0.1% sodium dodecyl sulfate, and RNA was recovered by ethanol precipitation following phenol–chloroform extraction using standard protocols.

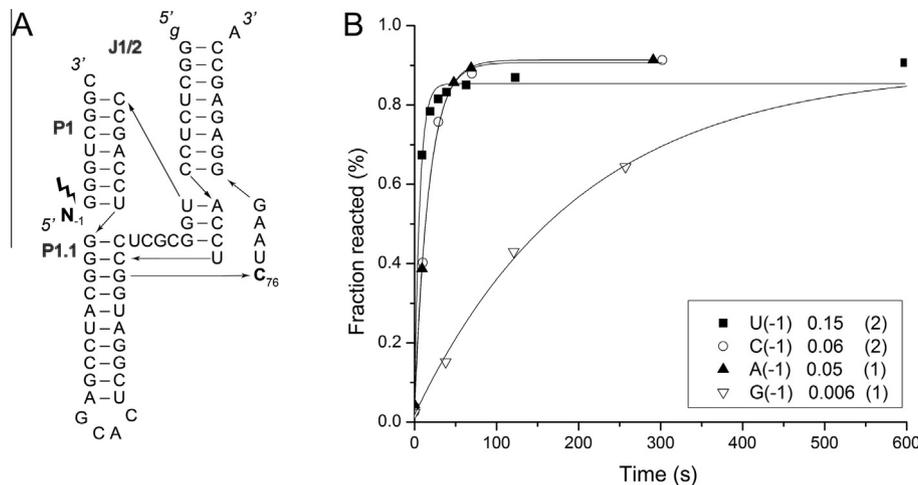
Nine nucleotide synthetic RNA substrates were designed that recognize the seven nucleotides of the P1 stem by base pairing and contain a one-nucleotide overhang at the 5' and 3' ends. Specific substrate variants were designed that contain all four nucleotide possibilities at position N(-1), 5' to the reactive phosphoryl group. As discussed below, additional variants were designed that incorporate one to three additional uridine residues at the 3' end to facilitate separation of substrate variants by PAGE (Table S1). For example, compared with the nine-nucleotide reference substrate with a U(-1), the substrate C12 contains a C(-1) at the 5' end and an additional three uridine nucleotides at the 3' end. Substrates containing nucleobase analogs at the N(-1) position were also designed to include 3' U residues permitting resolution of multiple substrates in the same reaction by PAGE. RNA substrates were 5' end-labeled using  $\gamma$ -<sup>32</sup>P ATP and OptiKinase and purified by 20% PAGE, elution of gel slices into TE, phenol–chloroform extraction, and ethanol precipitation.

### HDV ribozyme multiple and single substrate reactions

To facilitate proper folding prior to reaction, ribozyme RNA was heated at 60 °C for 1 min in 40 mM Tris–HCl (pH 8.0) and 1 mM ethylenediaminetetraacetic acid (EDTA), followed by an additional



**Fig. 1.** The theoretical framework that defines internal competition kinetics can be described by two-dimensional free energy landscapes for the competitive reactions of four substrates at the start of the reaction (left,  $f = 0$ ), at an intermediate time during the reaction (middle,  $1 > f > 0$ ), and after completion (right,  $f = 1$ ). In each panel, the relative concentrations of each competitive substrate and product are shown on the left and right of the free energy diagram, respectively. Due to differences in activation energy ( $\Delta G^\ddagger$ ) for the reactions of the competitive substrates, at intermediate times during the reaction the ratio of the substrates and products will be offset from their initial and final values, respectively. The faster reacting substrates become enriched in the product population ( $P_2/P_1 > 1$ ), whereas the slower reacting species are enriched in the residual unreacted substrate population ( $S_2/S_1 < 1$ ). These ratios change as a function of  $f$  according to the relative rate constant  $k_{rel}$ , which is defined by the differences in activation energy ( $\Delta G^\ddagger$ ).



**Fig. 2.** (A) Secondary structure of the *trans* antigenomic ribozyme used in this kinetic analysis, where the cleavage site is indicated by a bolt. This cleavage reaction requires the presence of at least a single nucleotide 5' of the scissile phosphate, and the identity of the upstream sequence at N(-1) affects the catalytic efficiency of the ribozyme. (B) Single turnover kinetic analysis of N(-1) substrate variation using individual substrate reactions and direct data fitting. U(-1) results in fast cleavage by the HDV ribozyme, whereas a G(-1) mutation reduces catalytic activity by 25-fold. The single turnover rate constants ( $s^{-1}$ ) for the N(-1) substrate variants are presented in the inset, and the error in the final decimal place from at least three independent reactions is presented in parentheses.

5 min of preincubation at 37 °C.  $MgCl_2$  was added to 6 mM, and the RNA was preincubated for 5 min at 37 °C. Cleavage reactions were initiated by the addition of preincubated RNA substrate in reaction buffer. Single turnover kinetics were monitored by removing aliquots at specific times and quenching the reaction in an equal volume of 90% formamide and 100 mM EDTA. Separation of radiolabeled substrate variants and products was achieved by heating to 90 °C for 1 min, followed by resolving the precursor and product species using 20% denaturing PAGE. Substrate and product bands were quantified using a Phosphorimager Storm 8 and ImageQuant software (Molecular Dynamics).

For individual substrate reactions, time courses were fit using a single exponential function. Experimental errors were obtained from the standard deviation of at least three independent measurements. For internal competition reactions containing four substrates, the product bands and individual substrate bands ranging from 9 to 12 nucleotides were quantified as described above. The observed rate constant for the entire population time course was fit using a single exponential first-order rate equation in Origin.

Relative rate constants from internal competition assays were determined using the internal competition kinetic analysis described below.

#### Internal competition kinetics of reactions containing multiple substrates

It is well established that under steady-state conditions alternative substrates act as competitive inhibitors, and their relative rate constants are the ratio of their respective  $V/K$  values multiplied by the ratio of their concentrations, where  $V/K$  is the second-order rate constant ( $M^{-1} s^{-1}$ ) at limiting  $S$ . The ratio of  $V/K$  values for alternative substrates is in fact the quantitative definition of “enzyme specificity” [12,15]. Differences in specificity are due to differences in the activation energies that are measured by  $V/K$  for alternative substrates for the same enzyme. As a consequence, faster reacting substrates will be depleted more quickly from the substrate population, resulting in differences in the ratios of individual

substrates relative to each other or to a specific reference substrate (Fig. 1).

Differences in rate constants between alternative substrates will also result in a transient accumulation of the products for the substrate with the largest  $V/K$  ratio at partial fractions of reaction. By quantitatively analyzing the change in the ratio of the individual substrate and product concentrations as a function of the reaction progress, the rate constants for individual substrates relative to a reference substrate can be measured.

As illustrated in Scheme 1, a single population of enzyme (E) can combine with multiple substrates ( $S_1, S_2, S_3, \dots, S_i$ ) to form individual ES complexes ( $ES_1, ES_2, ES_3, \dots, ES_i$ ) that react with rate constants ( $V_1, V_2, V_3, \dots, V_i$ ) to form the corresponding products ( $P_1, P_2, P_3, \dots, P_i$ ). The rate of product formation of any individual substrate ( $v_{obs1}$ ) is proportional to the fraction of total enzyme in the  $ES_1$  form [32],

$$v_{obs1} = V_1 E f_{ES1}, \quad (1)$$

where  $v_{obs1}$  is the observed reaction rate (M/s) for depletion of substrate  $S_1$ ,  $V_1$  is the first-order rate constant ( $s^{-1}$ ) for the reaction of  $ES_1$  to yield free product ( $P_1$ ) and free enzyme, and  $f_{ES1}$  is the fraction of total enzyme that is in the form that reacts with rate constant  $V_1$  to form free product.

Alternative substrates deplete  $ES_1$  and, consequently, the rate of formation of  $P_1$ . For two substrates, the multiple turnover rate equation is essentially that for competitive inhibition, and the ratio of the two observed rates simplifies to [12,15,17,33].

$$\frac{v_{obs2}}{v_{obs1}} = \frac{(V/K)_2}{(V/K)_1} \left( \frac{S_2}{S_1} \right), \quad (2)$$

Thus, the relative rate constant, or the ratio of the two individual rate constants for the two competing substrates, is the ratio of their respective  $V/K$  values multiplied by the ratio of their concentrations. Here, we define this ratio as  $k_{rel} = (V/K)_2/(V/K)_1$ , where the substrate considered in the denominator is the reference so that a substrate with a larger  $V/K$  relative to the reference substrate will have an  $k_{rel} > 1$ , whereas a  $k_{rel} < 1$  indicates a lower  $V/K$ . As illustrated below, the same relationship holds for any competitive pseudo-first-order reaction and, therefore, is applicable to both multiple turnover and single turnover reactions. For single turnover reactions, the  $k_{rel}$  will reflect the ratio of the rate constants for the experimental and reference substrate.

Integration of Eq. (2) describes how the ratio of substrates and products will change over the time course for first-order and pseudo-first-order reactions [32,34]:

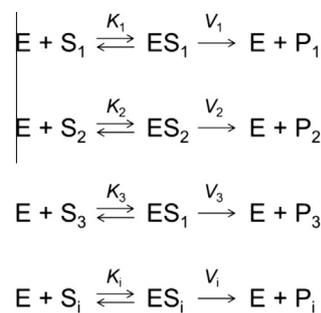
$$k_{rel} = \frac{\ln\left(\frac{S_2}{S_{2,0}}\right)}{\ln\left(\frac{S_1}{S_{1,0}}\right)}, \quad (3)$$

where  $S_{1,0}$  and  $S_{2,0}$  are the initial concentrations of the two substrates and  $S_1$  and  $S_2$  are their concentrations after a specific time interval.

Given the following definitions for the ratios of  $S_1$  and  $S_2$ , an expression for  $k_{rel}$  can be obtained in terms of these ratios for the two substrates:

$$R_2 = S_2/S_1 \\ R_0 = S_{2,0}/S_{1,0},$$

where  $S_1$  and  $S_2$  are the concentrations of these two substrates at a fraction of reaction  $f$  and  $S_{1,0}$  and  $S_{2,0}$  are their initial concentrations ( $f=0$ ). Note that in this case the reference substrate,  $S_1$ , is in the denominator and so the meaning of  $k_{rel}$  is that defined above. Also note that the parameter  $f$ , referred to as the fraction of reaction of the  $S_1$  substrate, is defined as follows:



Scheme 1. Multiple turnover kinetic mechanism.

$$f = (S_{1,0} - S_1)/S_{1,0} = P_1/S_{1,0} \\ f = 1 - (S_1/S_{1,0}) \\ (1 - f) = (S_1/S_{1,0}).$$

Integration of Eq. (3) expanded to multiple alternative substrates can be used to give the following expression for  $k_{rel}$  in terms of substrate ratios  $R_i$  and  $R_{i,0}$  and the fraction of reaction ( $f$ ) and the mole fractions for each species ( $X_i$ ) [17]:

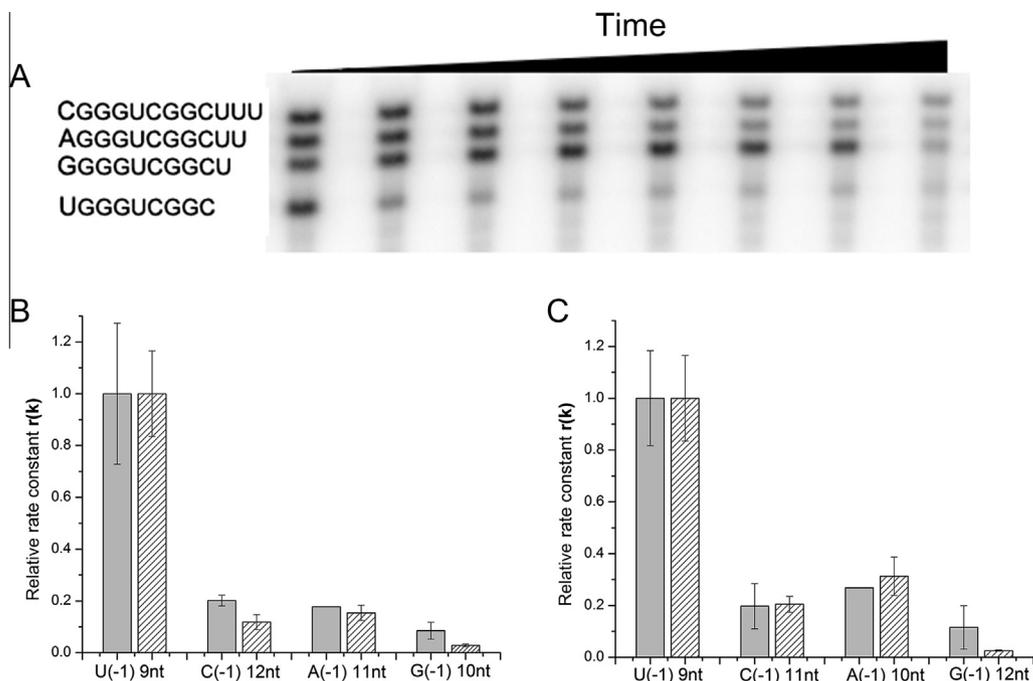
$$k_{rel} = \frac{\ln\left(\frac{1-f}{\frac{R_{2,0}}{R_2} X_1 + \sum\left(\frac{R_i}{R_{i,0}} X_i\right) + X_2}\right)}{\ln\left(\frac{1-f}{X_1 + \left(\frac{R_2}{R_{2,0}}\right) X_2 + \sum\left(\frac{R_i}{R_{i,0}}\right) X_i}\right)} \quad (4)$$

A more complete derivation of Eq. (4) is provided in Guenther and coworkers [28]. Importantly, the  $k_{rel}$  for any two substrates will be independent of the presence of all other competing substrates, allowing for the analysis of much larger substrate populations. Using this expression, the experimental data needed to determine the relative rate constants for all substrate species are the fraction of reaction of the population  $f$  and the  $R_i$  ratio for each individual substrate relative to the specific reference ( $S_1$ ). As described in the text, this information is readily obtained using standard molecular biology methods, allowing application of internal competition to a wide range of RNA processing reactions, including RNA catalysis.

## Results and discussion

### Internal competition kinetic analysis of N(-1) nucleobase specificity by the antigenomic HDV ribozyme

We used the specificity of the antigenomic HDV ribozyme for the N(-1) nucleotide of a substrate RNA oligonucleotide added in *trans* as a model system to define the utility and limitations of multiple alternative substrate kinetics. The rate constants for HDV ribozyme cleavage of substrates with different nucleobases at N(-1) have previously been shown to vary in decreasing order,  $U > C > A > G$ , with the first-order rate constant for the U(-1) substrate being approximately 18-fold higher than that for G(-1) [7,35–37]. Recent results from a minimal HDV ribozyme identified in the human microbiome (47 nt) show similar specificity at N(-1) (e.g.,  $U > C > A > G$ ); however, neither the structural nor mechanistic basis for these effects is known [2]. Fig. 2 shows that the *trans* cleavage reaction of the wild-type, antigenomic HDV ribozyme accurately reproduces this feature of HDV ribozyme catalytic specificity. Direct fitting of the individual time courses to an exponential function shows that changing U(-1) to an A or C results in a 5- to 8-fold decrease in the rate constant for catalysis ( $k_{obs}$ ), whereas a G(-1) substrate reacts with an approximately 30-fold slower  $k_{obs}$ .



**Fig. 3.** (A) PAGE analysis of internal competition reactions containing four alternative HDV ribozyme substrates varying in nucleobase composition at the N(-1) position. The length of these substrates was altered by the addition of one to three uridine residues to the 3' end of the RNA in order to resolve these species by PAGE. The U(-1) substrate is depleted early in the time course, whereas the G(-1) substrate that reacts with the slowest rate constant accumulates at intermediate times relative to the other substrates. The A(-1) and C(-1) substrates are depleted from the residual substrate population with similar apparent kinetics. Thus, the qualitative changes in the band intensities from PAGE analysis correspond to expectations from measurement of individual reaction time courses. (B and C) To control for the effect of the addition of 3' uridine residues, two substrate populations were examined. The first population consisted of U9, G10, A11, and C12, the second one consisted of U9, A10, C11, and G12, where the N(-1) base identity is indicated with its associated substrate sequence length. The  $k_{rel}$  values measured by internal competition for both populations containing mixtures of four substrates (gray bars) correspond with  $k_{rel}$  ratios ( $k_{rel} = k_{(obs)exp}/k_{(obs)ref}$ ) calculated from reactions containing individual substrates (shaded bars).

Application of internal competition kinetics requires the precise measurement of the ratio of the concentrations of the competing substrates and the fraction of reaction [17] (see Fig. 1). To individually quantify different sequence variants at the N(-1) position, we designed a series of substrates in which additional uridine residues were added to the 3' end of the RNA to facilitate separation by PAGE. As shown in Fig. 3A, four alternative substrates ranging in length from 9 to 12 nucleotides are easily resolved and can be quantified separately using phosphorimager analysis. The additional uridine residues essentially serve as a remote label for the identity of the nucleobase at the N(-1) position. These values can be used to compute substrate ratios ( $R_i$ ) and reaction progress ( $f$ ) as described above. Note that the PAGE results by inspection show that the relative concentration of the U(-1) substrate is depleted early in the time course, whereas the relative concentration of the G(-1) substrate that reacts with the slowest rate constant is enriched at intermediate times. The A(-1) and C(-1) substrates are depleted from the residual substrate population with similar apparent kinetics. Thus, the qualitative changes in the band intensities from PAGE analysis correspond to expectations from measurements of individual reaction time courses.

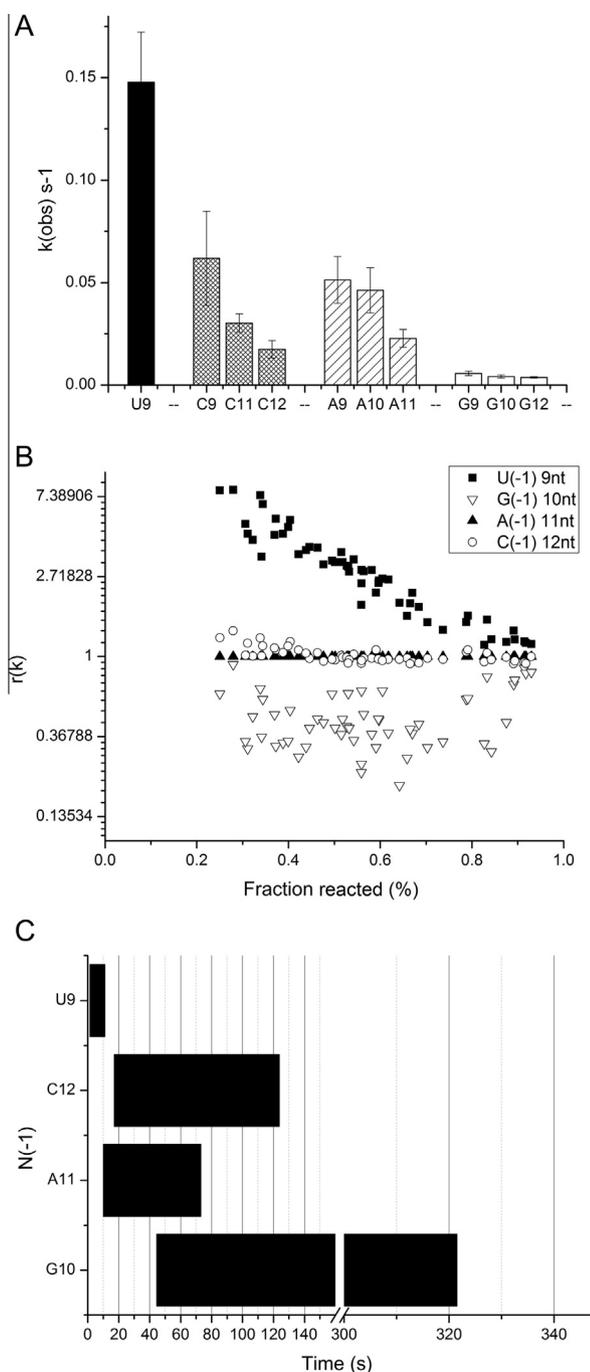
#### Optimization of substrate sequence length and reference substrate selection

A potential source of inaccuracy in determination of relative rate constants by internal competition could be interfering or additive contribution to individual rate constants from the presence of additional 3' uridine residues used to identify different substrate variants. Therefore, we compared the  $k_{rel}$  values obtained from reactions containing two different substrate populations containing four substrates each in which different N(-1) mutations were

identified with different substrate lengths. The first population consisted of U9, G10, A11, and C12, and the second one consisted of U9, A10, C11, and G12, where the N(-1) base identity is indicated with its associated substrate sequence length from the native 9 nucleotides up to 12 with the addition of 3 extra uridine residues (Fig. 3A). The  $k_{rel}$  values for mutations at N(-1) were determined by internal competition using Eq. (4) for both populations and were also calculated from rate constants determined by direct fitting data from reactions containing single substrates. In single turnover reactions, all substrates were observed to react to more than 85% completion and fit well to a single exponential function characteristic of a pseudo-first-order reaction. As shown in Fig. 3B and C, the  $k_{rel}$  values measured by internal competition for both populations containing mixtures of four substrates correspond with  $k_{rel}$  ratios ( $k_{rel} = k_{(obs)exp}/k_{(obs)ref}$ ) calculated from reactions containing individual substrates.

Nonetheless, comparison of individual time course data revealed that the sequence length difference used to permit resolution by PAGE does introduce differences in rate constants that must be corrected to precisely measure the intrinsic difference in  $k_{rel}$  due to N(-1) mutation. Although the effects of additional 3' U residues on  $k_{obs}$  are minimal for most substrates, a 2- to 3-fold decrease was observed for A11 and C11 compared with A9 and C9 (Fig. 4A). The largest effect (~3.5-fold) was observed for C12 relative to the C9 reference substrate. Thus, for most substrates the observed  $k_{rel}$  is accurate; however, the structural changes introduced to allow resolution by PAGE can introduce inaccuracies that require correction. The  $k_{rel}$  values determined for the HDV ribozyme catalyzed cleavage of N(-1) substrate variants reported below have been corrected for the effect of 5' U addition.

A second potential source of inaccuracy can arise due to the reduced signal-to-noise ratio at higher fractions of reactions. As



**Fig. 4.** (A) Effects of additional 3' U residues on  $k_{\text{obs}}$  were calculated by direct data fitting of individual substrate, single turnover kinetic time courses. These effects are minimal for most substrates (>2-fold). The largest effect (~3.5-fold) was observed for C12 relative to the C9 reference substrate. Thus, for most substrates the observed  $k_{\text{rel}}$  is accurate; however, the structural changes introduced to allow resolution by PAGE can introduce inaccuracies that require correction. (B) The observed  $k_{\text{rel}}$  values decline or increase with increasing  $f$  before converging at a  $k_{\text{rel}}$  of 1. This effect results from the fact that at higher fractions of reaction the substrate ratio measurement becomes dominated by the contributions from residual unreacted substrate. Restricting the analysis to fractions of reaction progress where all substrates are undergoing reaction can avoid potential inaccuracies from this effect. For this analysis, population reaction progress from  $f = 0.2$  to  $0.4$  was analyzed. (C) Range of time for which the experimental substrate population is undergoing reaction ( $f = 0.2$ – $0.8$ ). Large differences in observed rate constants between individual substrate variants and the reference substrate can introduce inaccuracies due to signal-to-noise limitations resulting from the relative abundance of the reference substrate. The A(-1) substrate undergoes reaction with kinetics such that its concentration changes over a range that overlaps most with the other substrates in the population and is the optimal reference for this set of substrates.

shown in Fig. 3A, the relative abundances of the different substrate species becomes altered due to differences in rate constants at intermediate time points. However, at higher fractions of reaction, the substrate ratio measurement becomes dominated by the contributions from residual unreacted substrate. This is evident in Fig. 4B, where the observed  $k_{\text{rel}}$  values converge at 1 at high fractions of reaction. The reason for this effect is that at later time points the  $R_i$  values for both the experimental and reference substrates will be similar due to the residual signal from unreacted substrate RNA. Restricting the analysis to low  $f$  values where the signal is dominated by the reactive RNA population circumvents this complication. For subsequent analysis, population reaction progress from  $f = 0.2$  to  $0.4$  was analyzed to minimize these effects.

A third possible source of inaccuracy apparent from these analyses arises from large differences between individual substrate variants and the reference substrate (Fig. 4C). That is, when the slow reacting G(-1) substrate reaches an  $f$  of 0.2, the U(-1) substrate is at an  $f$  of more than 0.9 and measurement of its concentration will be subject to inaccuracy due to signal-to-noise limitations. Thus, the measurement of the  $k_{\text{rel}}$  for G(-1) is less precise due to error in the measurement of the relative abundance of the reference substrate. This effect is illustrated in Fig. 4C, where the range of time where the  $f$  value is 0.2 to 0.8 is shown for the U(-1), C(-1), A(-1), and G(-1) substrates. These data show that the fastest reacting U(-1) substrate is essentially completely reacted by the time the concentrations of A(-1) and C(-1) have begun to change appreciably. The G(-1) substrate as the slowest reacting substrate nonetheless changes over a range of time that overlaps with the A(-1) and C(-1) substrates. The A(-1) substrate undergoes reaction with kinetics such that its concentration changes over a range that overlaps most with the other substrates in the population. For these reasons, the A(-1) substrate is chosen as an optimal reference for this set of substrates because its rate constant is intermediate between the slowest and fastest reacting species. Importantly, the collection and processing of data for the experiment do not depend a priori on the choice of the reference substrate, which can be varied in order to cover the range of relative rate constants occurring in the experiment.

#### Use of internal competition to test a potential pairing interaction involving N(-1)

One likely mechanism for positioning the 2'-hydroxyl nucleophile is formation of a pairing interaction between the N(-1) nucleobase and another nucleotide in the catalytic core of the ribozyme. Recent molecular dynamic simulations comparing different points along the reaction pathway suggested that U(-1) may form interactions A78 and/or A79 [38]. The in-line conformation of U(-1) in an activated precursor state where the 2'O nucleophile is deprotonated was observed to form an H-bond with the N6 of A78, consistent with the experimentally observed importance of the exocyclic NH<sub>2</sub> group of this residue [39–41]. In addition, a Watson–Crick base pair was observed to form between A79 and U(-1) when the simulations were performed with an early transition state mimic consistent with the importance of the identity of A78 [39–41] as well as the nucleobase preference of the -1 position [36,37] for reactions in vitro. Alternatively, the N(-1) nucleobase could form some variation on these interactions within the catalytic core resulting in the observed specificity. The simple base-pairing model predicts that the detrimental effect of mutation at A78 would be ameliorated by a corresponding change in the substrate N(-1) nucleobase that preserves Watson–Crick base-pairing. More generally, if there are nucleobase-specific interactions in the region of these two nucleobases, the specificity of mutant ribozymes for sequence variation at the N(-1) position could still be affected.

Individual point mutations at A78 and A79 caused significant defects in HDV ribozyme catalysis, as observed previously [39–41]. To analyze the effects of mutations at these positions on N(-1) nucleobase specificity, it is necessary to employ HDV ribozyme mutants that retain sufficient activity to allow product formation to be scored. Of all eight possible individual mutations at A78 and A79, only three ribozymes retained significant activity. The A78G, A78U, and A79G mutants retained the greatest activity, with  $k_{\text{obs}}$  values in single turnover experiments of 0.03, 0.0017, and  $0.0015 \text{ s}^{-1}$ , respectively (see Fig. 2), corresponding to decreases of 4- to 100-fold compared with the native HDV ribozyme. As shown in Fig. 5, the specificity of the HDV ribozyme cleavage is insensitive to mutation of A78 to either U or G and, likewise, does not change for the A79G mutant. For all three mutant ribozymes, there are significant and quantitatively similar decreases in  $k_{\text{obs}}$  when U(-1) is mutated to A, C, or G (Table S2). The simplest interpretation is that there is no nucleobase-specific interaction between N(-1) and A78 or A79. Alternatively, these interactions may occur at a point along the reaction coordinate that is not rate limiting and, therefore, not reflected in the  $k_{\text{obs}}$  measured under the current conditions.

#### Comparative analysis of the effects of chemical mutagenesis of N(-1)

To gain additional information on the chemical basis for the N(-1) nucleobase specificity, we used internal competition to determine the effects of nucleotide analog substitutions at the N(-1) position. Previous analysis showed that whereas purines at N(-1) bind tightly, this stability correlates with lower catalytic efficiency [36,37]. The preference for pyrimidine nucleobases at the N(-1) position of the HDV ribozyme substrate has also been attributed to ground state destabilization [36]. An abasic residue at N(-1) binds with a dissociation constant similar to a pyrimidine at this position yet reacts 10-fold slower, which has been interpreted as reflecting greater disorder and, consequently, deviation from optimal inline geometry [7]. In solution, different nucleobases flanking the reactive phosphoryl group result in different rates due to the intrinsic effect on  $\text{pK}_a$  of 2'-OH and, presumably, could reflect differences in the HDV active site as well. Thus, no clear role has emerged for interactions involving the nucleobase 5' to the HDV reactive phosphoryl group despite evidence that aspects of its chemical structure contribute to catalysis.

Accordingly, the effects of a series of nucleotide analog substitutions at N(-1) were tested (Fig. 6A). Individual substrates were modified to contain additional 5'-U residues to provide resolution by PAGE. Two substrate populations consisting of four individual substrates were analyzed. Due to the low rate constants, for cleavage of the nucleotide analog substrates, we used the G(-1) substrate variant as the reference substrate in order to determine the relative rate constants for the remaining three substrates in the reaction. The two substrate populations consisted of substrates in which the N(-1) position was changed from guanosine (G9) to inosine (I10), 2-aminopurine (2AP11), or purine (Pu12), and additional uridine residues were added to achieve substrate oligonucleotide lengths of 9, 10, 11, or 12 nucleotides, respectively. Similarly, the substrates in the second population were modified at the N(-1) position to contain 2,6-diaminopurine (DAP10), 2,2-dimethyladenosine (DMA11), or 3-methyluridine (3MU12) with similar additional sequence lengths to allow separate quantification by PAGE. As illustrated above for N(-1) mutations, the  $k_{\text{rel}}$  value for each substrate was determined by fitting  $R_i$  and  $f$  data using Eq. (4) at  $f$  values of 0.2 to 0.4. To provide further validation for the analytical method, individual rate constants were determined by direct fitting data from reactions containing each substrate individually. The values obtained using the two different methods are compared in Fig. 6B.

An increase in  $k_{\text{rel}}$  is observed for inosine at N(-1) compared with the reference G9 substrate, consistent with interference due to the presence of an N2 amine group in the minor groove. This interpretation is consistent with the effects of 2,6-diaminopurine (DAP) and purine relative to the A(-1) substrate variant. The DAP and 2AP substrates differ from adenosine by the addition of an N2 amine in the minor groove, and both result in a decrease in  $k_{\text{rel}}$  compared with adenosine at N(-1). Interestingly, substitution with purine at this position, which shares the sterically unrestricted minor groove of adenosine but lacks a hydrogen bond acceptor in the major groove, also shows a decrease in  $k_{\text{rel}}$  compared with adenosine. This chemical modification suggests that a hydrogen bond donor in the major groove may be a positive determinant of catalysis. The substrate variant containing N6-dimethyladenosine further explored this possibility. The relative rate constant for catalysis of the N6-dimethyladenosine substrate is equivalent to that observed for purine ribonucleoside, which is consistent with a hydrogen bond donor in the major groove as a positive determinant for catalytic specificity. The results for 3-methyluridine show a significant decrease in  $k_{\text{rel}}$  compared with a U at the N(-1) position. Although aromatic methylation leads to increased hydrophobicity and stacking ability, in light of the detrimental effect of N6 methylation of A, the effect of methylation of the N3 of U(-1) may arise due to decreased ability to hydrogen bond in the major groove.

High-resolution crystal structures in the pre-cleavage and post-cleavage configurations exist for the HDV ribozyme and provide additional insight into the role of this position. Although the determination of the structure for the self-cleaved form of the HDV ribozyme revealed the nested double pseudo-knot structure of the ribozyme and the general positioning of the substrate, the absence of the 3' product led to minimal insight regarding the specificity at the N(-1) position [20]. In the pre-cleavage structure, the U(-1) nucleotide is sharply bent approximately  $180^\circ$  away from the downstream substrate strand, positioning this residue in the catalytic core [5]. However, this structure contains an inactivating mutation at C75 that is predicted to destabilize the conformation of the active site. In addition, in the pre-cleavage structure most likely to reflect the active conformation of the HDV ribozyme catalytic core, poor electron density was observed for U(-1) [3,42]. Orientation of this critical region of the substrate strand within the

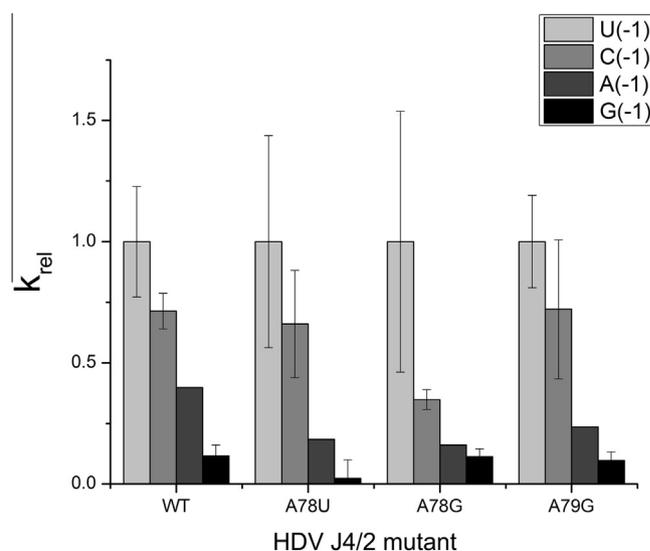
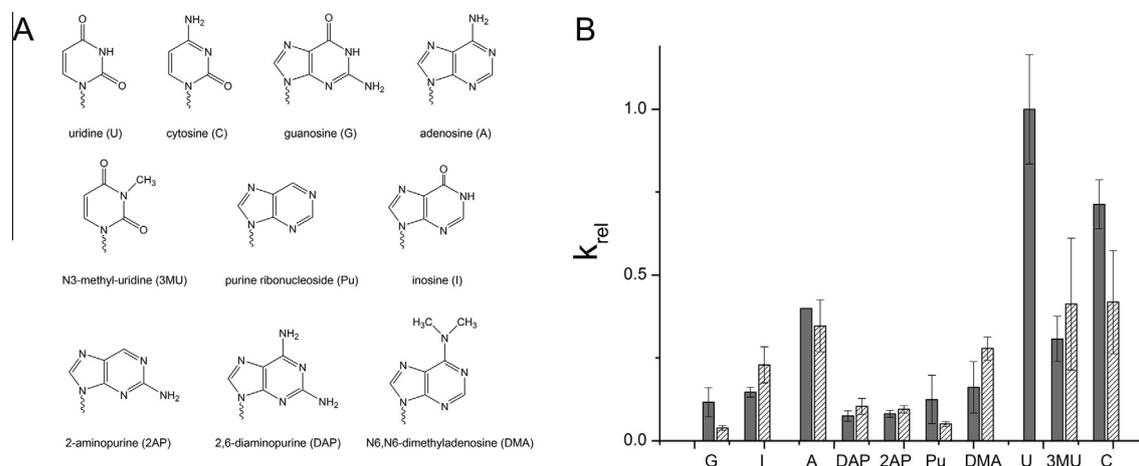


Fig. 5.  $k_{\text{rel,obs}}$  values measured by internal competition for substrate populations containing all four natural nucleotides at the N(-1) position corrected for 3' U addition and normalized to U9.



**Fig. 6.** (A) Analysis of nucleobase analogs designed to probe the functional importance of hydrogen bond donors and acceptors in the major and minor grooves at N(-1) in the HDV ribozyme substrate. (B)  $k_{rel}$  values for N(-1) substrate variants determined by internal competition have been corrected for 3' U addition and normalized to U9. The  $k_{rel}$  values measured by internal competition for both populations containing mixtures of four substrates (gray bars) correspond with  $k_{rel}$  ratios ( $k_{rel}$  ratio =  $k_{(obs)exp}/k_{(obs)ref}$ ) calculated from reactions containing individual substrates (shaded bars).

active site of this structure was achieved by superpositioning the corresponding residue downstream of the scissile phosphate in the hammerhead ribozyme structure. However, biochemical probes of ribozyme structure reveal that global and local structures are influenced by the identity of the nucleobase at the N(-1) position.

Although mutations at N(-1) show only subtle changes in terbium cleavage patterns, fluorescence resonance energy transfer (FRET) fluorophores attached to the termini of P2 and P4 show significant changes, suggesting substantial fluctuations in the architecture of the catalytic core [35]. Molecular mechanic simulations observe significant changes in the angle of the kink around the cleavage site, which varies from 178° for U(-1) to 86° for A(-1) [31]. These simulations support a model in which the relative lack of H-bond interactions for U(-1) results in the sharpest kink around the scissile phosphate, forming a U-turn motif to expose the cleavage site to the active site general acid (C75). Although attractive, this model requires rigorous kinetic examination in order to understand the precise interactions underlying the structural changes due to variation at N(-1).

## Summary

The results presented here illustrate the use of multiple substrate reactions and the application of internal competition kinetics to interrogate in vitro RNA processing reactions. The analyses reveals important aspects of experimental design and interpretation to consider in order to maximize accuracy of the resulting relative rate constants. However, this general approach can be used to rapidly test the effects of substrate mutations as well as chemical modification on reaction rate constants. This method should be widely applicable to analyze RNA processing reactions in vitro because it takes advantage of substrate labeling, separation, and detection methods that are standard in the field. The results obtained for the HDV ribozyme provide important new mechanistic details regarding the chemical specificity of the RNA active site for the nucleobase at N(-1) that is highly valuable for evaluating structural and mechanistic models of HDV ribozyme catalysis.

## Acknowledgments

This work was supported by R01GM096000 and R01GM056740 to M.E.H., R01AI081987 to J.A.P., and R01GM062248 to

D.M.Y. In addition, D.L.K. was supported by National Institute of General Medical Sciences (NIGMS) training Grant T32GM008056. K.S.S. and M.P. received support from NIGMS training Grant R25GM049010.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ab.2015.04.024>.

## References

- [1] W.P. Jencks, Binding energy, specificity, and enzymic catalysis: the circe effect, *Adv. Enzymol. Relat. Areas Mol. Biol.* 43 (1975) 219–410.
- [2] N. Riccitelli, A. Lupták, HDV family of self-cleaving ribozymes, in: A.S. Garrett (Ed.), *Progress in Molecular Biology and Translational Science*, Academic Press, San Diego, 2013, pp. 123–171.
- [3] B.L. Golden, Two distinct catalytic strategies in the hepatitis delta virus ribozyme cleavage reaction, *Biochemistry* 50 (2011) 9424–9433.
- [4] A.L. Cerrone-Szkal, D.M. Chadalavada, B.L. Golden, P.C. Bevilacqua, Mechanistic characterization of the HDV genomic ribozyme: the cleavage site base pair plays a structural role in facilitating catalysis, *RNA* 14 (2008) 1746–1760.
- [5] A.R. Ferre-D'Amare, K. Zhou, J.A. Doudna, Crystal structure of a hepatitis delta virus ribozyme, *Nature* 395 (1998) 567–574.
- [6] W.W. Cleland, A.C. Hengge, Mechanisms of phosphoryl and acyl transfer, *FASEB J.* 9 (1995) 1585–1594.
- [7] S. Jeong, J. Sefcikova, R.A. Tinsley, D. Rueda, N.G. Walter, Trans-acting hepatitis delta virus ribozyme: catalytic core and global structure are dependent on the 5' substrate sequence, *Biochemistry* 42 (2003) 7727–7740.
- [8] J.K. Frederiksen, J.A. Piccirilli, Identification of catalytic metal ion ligands in ribozymes, *Methods* 49 (2009) 148–166.
- [9] J.L. Wilcox, A.K. Ahluwalia, P.C. Bevilacqua, Charged nucleobases and their potential for RNA catalysis, *Acc. Chem. Res.* 44 (2011) 1270–1279.
- [10] J.A. Nelson, O.C. Uhlenbeck, When to believe what you see, *Mol. Cell* 23 (2006) 447–450.
- [11] T.K. Stage-Zimmermann, O.C. Uhlenbeck, Hammerhead ribozyme kinetics, *RNA* 4 (1998) 875–889.
- [12] A. Fersht, *Structure and Mechanism in Protein Science*, W. H. Freeman, San Francisco, 1998.
- [13] M. Forconi, R.N. Sengupta, J.A. Piccirilli, D. Herschlag, A rearrangement of the guanosine-binding site establishes an extended network of functional interactions in the Tetrahymena group I ribozyme active site, *Biochemistry* 49 (2010) 2753–2762.
- [14] M. Forconi, R.H. Porecha, J.A. Piccirilli, D. Herschlag, Tightening of active site interactions en route to the transition state revealed by single-atom substitution in the guanosine-binding site of the Tetrahymena group I ribozyme, *J. Am. Chem. Soc.* 133 (2011) 7791–7800.
- [15] S. Cha, Kinetics of enzyme reactions with competing alternative substrates, *Mol. Pharmacol.* 4 (1968) 621–629.
- [16] V. Schellenberger, R.A. Siegel, W.J. Rutter, Analysis of enzyme specificity by multiple substrate kinetics, *Biochemistry* 32 (1993) 4344–4348.
- [17] W.W. Cleland, The use of isotope effects to determine enzyme mechanisms, *Arch. Biochem. Biophys.* 433 (2005) 2–12.

- [18] L. Melander, W.H. Saunders, *Reaction Rates of Isotopic Molecules*, John Wiley, New York, 1980.
- [19] H.-C. Lin, L.E. Yandek, I. Gjermeni, M.E. Harris, Determination of relative rate constants for in vitro RNA processing reactions by internal competition, *Anal. Biochem.* 467 (2014) 54–61.
- [20] A. Kohen, H.H. Limbach, *Isotope Effects in Chemistry and Biology*, Taylor & Francis, Boca Raton, FL, 2005.
- [21] N. Pi, J.A. Leary, Determination of enzyme/substrate specificity constants using a multiple substrate ESI–MS assay, *J. Am. Soc. Mass Spectrom.* 15 (2004) 233–243.
- [22] M.E. Harris, Q. Dai, H. Gu, D.L. Kellerman, J.A. Piccirilli, V.E. Anderson, Kinetic isotope effects for RNA cleavage by 2'-O-transphosphorylation: nucleophilic activation by specific base, *J. Am. Chem. Soc.* 132 (2010) 11613–11621.
- [23] K.A. Manning, B. Sathyamoorthy, A. Eletsky, T. Szyperki, A.S. Murkin, Highly precise measurement of kinetic isotope effects using  $^1\text{H}$ -detected 2D [ $^{13}\text{C}$ ,  $^1\text{H}$ ]-HSQC NMR spectroscopy, *J. Am. Chem. Soc.* 134 (2012) 20589–20592.
- [24] J. Chan, A.R. Lewis, M. Gilbert, M.F. Karwaski, A.J. Bennet, A direct NMR method for the measurement of competitive kinetic isotope effects, *Nat. Chem. Biol.* 6 (2010) 405–407.
- [25] D.A. Hiller, M. Zhong, V. Singh, S.A. Strobel, Transition states of uncatalyzed hydrolysis and aminolysis reactions of a ribosomal P-site substrate determined by kinetic isotope effects, *Biochemistry* 49 (2010) 3868–3878.
- [26] D.W. Parkin, H.B. Leung, V.L. Schramm, Synthesis of nucleotides with specific radiolabels in ribose: Primary  $^{14}\text{C}$  and secondary  $^3\text{H}$  kinetic isotope effects on acid-catalyzed glycosidic bond hydrolysis of AMP, dAMP, and inosine, *J. Biol. Chem.* 259 (1984) 9411–9417.
- [27] J.G. Bertram, K. Oertell, J. Petruska, M.F. Goodman, DNA polymerase fidelity: Comparing direct competition of right and wrong dNTP substrates with steady state and pre-steady state kinetics, *Biochemistry* 49 (2010) 20–28.
- [28] U.P. Guenther, L.E. Yandek, C.N. Niland, F.E. Campbell, D. Anderson, V.E. Anderson, M.E. Harris, E. Jankowsky, Hidden specificity in an apparently nonspecific RNA-binding protein, *Nature* 502 (2013) 385–388.
- [29] D. Chadalavada, A. Cerrone-Szakal, J. Wilcox, N. Siegfried, P. Bevilacqua, Mechanistic analysis of the hepatitis delta virus (HDV) ribozyme: methods for RNA preparation, structure mapping, solvent isotope effects, and co-transcriptional cleavage, in: J.S. Hartig (Ed.), *Ribozymes*, Humana, Totowa, NJ, 2012, pp. 21–40.
- [30] S.R. Das, J.A. Piccirilli, General acid catalysis by the hepatitis delta virus ribozyme, *Nat. Chem. Biol.* 1 (2005) 45–52.
- [31] H. Kadowaki, T. Kadowaki, F.E. Wondisford, S.I. Taylor, Use of polymerase chain reaction catalyzed by Taq DNA polymerase for site-specific mutagenesis, *Gene* 76 (1989) 161–166.
- [32] W.W. Cleland, P.F. Cook, *Enzyme Kinetics and Mechanism*, Garland, New York, 2007.
- [33] D. Herschlag, The role of induced fit and conformational changes of enzymes in specificity and catalysis, *Bioorg. Chem.* 16 (1988) 62–96.
- [34] W.W. Cleland, *Enzyme mechanisms from isotope effects*, in: A. Kohnen (Ed.), *Isotope Effects in Chemistry and Biology*, Taylor & Francis, Boca Raton, FL, 2006.
- [35] P. Deschenes, D.A. Lafontaine, S. Charland, J.P. Perreault, Nucleotides-1 to -4 of hepatitis delta ribozyme substrate increase the specificity of ribozyme cleavage, *Antisense Nucleic Acid Drug Dev.* 10 (2000) 53–61.
- [36] I. Shih, M.D. Been, Energetic contribution of non-essential 5' sequence to catalysis in a hepatitis delta virus ribozyme, *EMBO J.* 20 (2001) 4884–4891.
- [37] J. Sefcikova, M.V. Krasovska, J. Sponer, N.G. Walter, The genomic HDV ribozyme utilizes a previously unnoticed U-turn motif to accomplish fast site-specific catalysis, *Nucleic Acids Res.* 35 (2007) 1933–1946.
- [38] T.S. Lee, G. Giambasu, M.E. Harris, D.M. York, Characterization of the structure and dynamics of the HDV ribozyme at different stages along the reaction path, *J. Phys. Chem. Lett.* 2 (2011) 2538–2543.
- [39] P.K. Kumar, Y.A. Suh, H. Miyashiro, F. Nishikawa, J. Kawakami, K. Taira, S. Nishikawa, Random mutations to evaluate the role of bases at two important single-stranded regions of genomic HDV ribozyme, *Nucleic Acids Res.* 20 (1992) 3919–3924.
- [40] F. Nishikawa, M. Shirai, S. Nishikawa, Site-specific modification of functional groups in genomic hepatitis delta virus (HDV) ribozyme, *Eur. J. Biochem.* 269 (2002) 5792–5803.
- [41] H.-N. Wu, Z.-S. Huang, Mutagenesis analysis of the self-cleavage domain of hepatitis delta virus antigenomic RNA, *Nucleic Acids Res.* 20 (1992) 5937–5941.
- [42] J.H. Chen, R. Yajima, D.M. Chadalavada, E. Chase, P.C. Bevilacqua, B.L. Golden, A 1.9 Å crystal structure of the HDV ribozyme precleavage suggests both Lewis acid and general acid mechanisms contribute to phosphodiester cleavage, *Biochemistry* 49 (2010) 6508–6518.