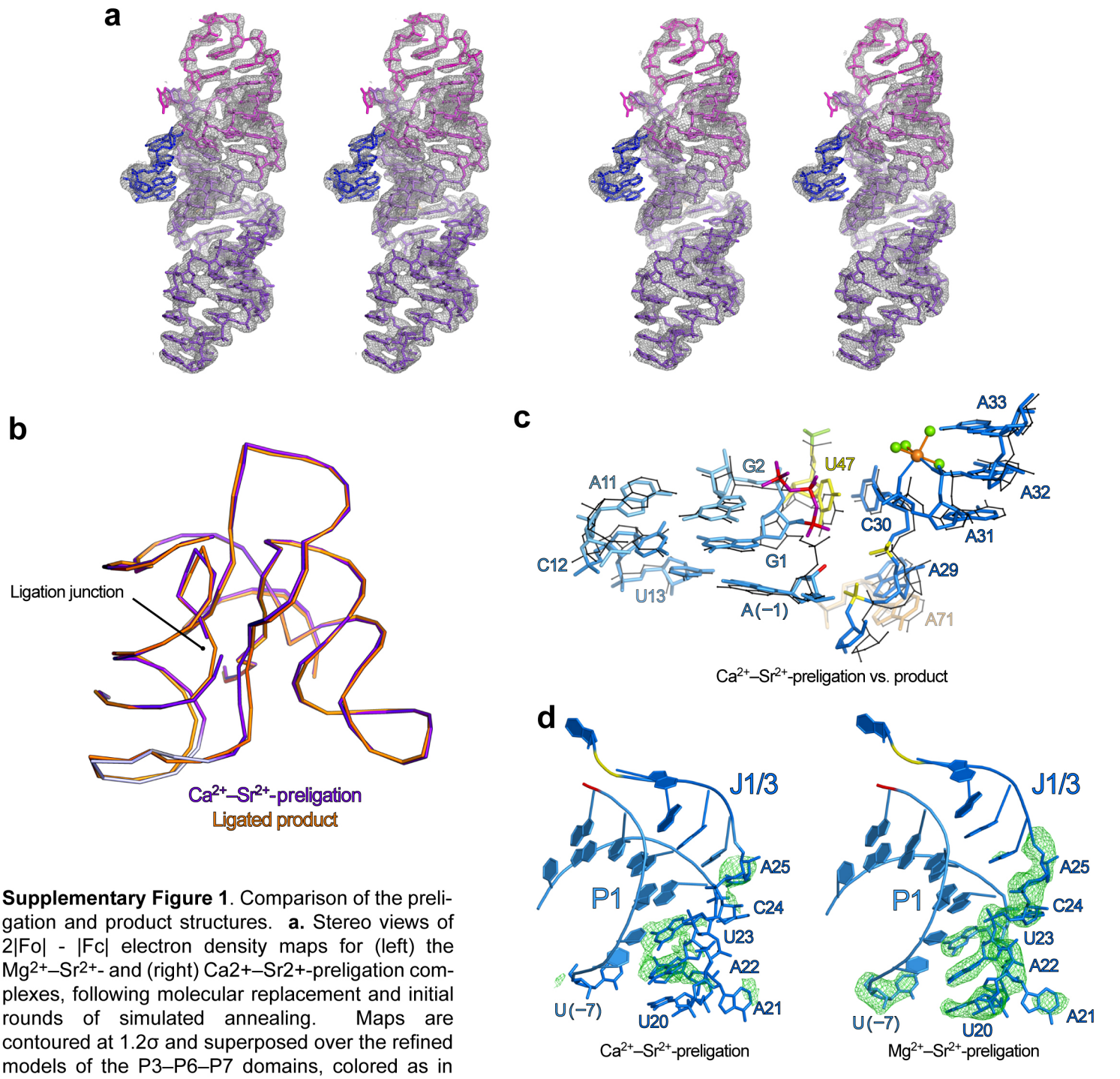
Supplementary information for

The structural basis of RNA-catalyzed RNA polymerization

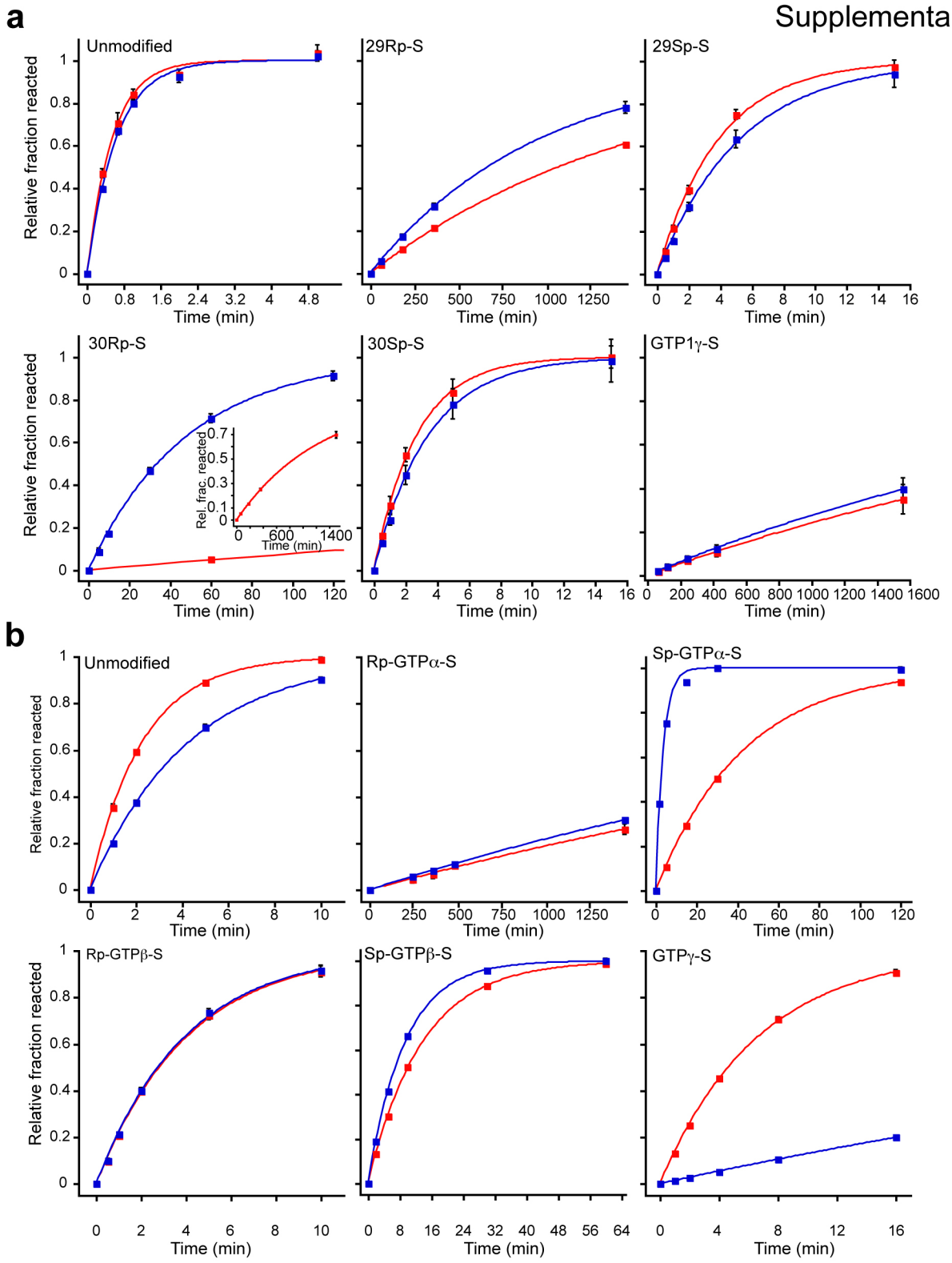
David M. Shechner and David P. Bartel

This document includes:

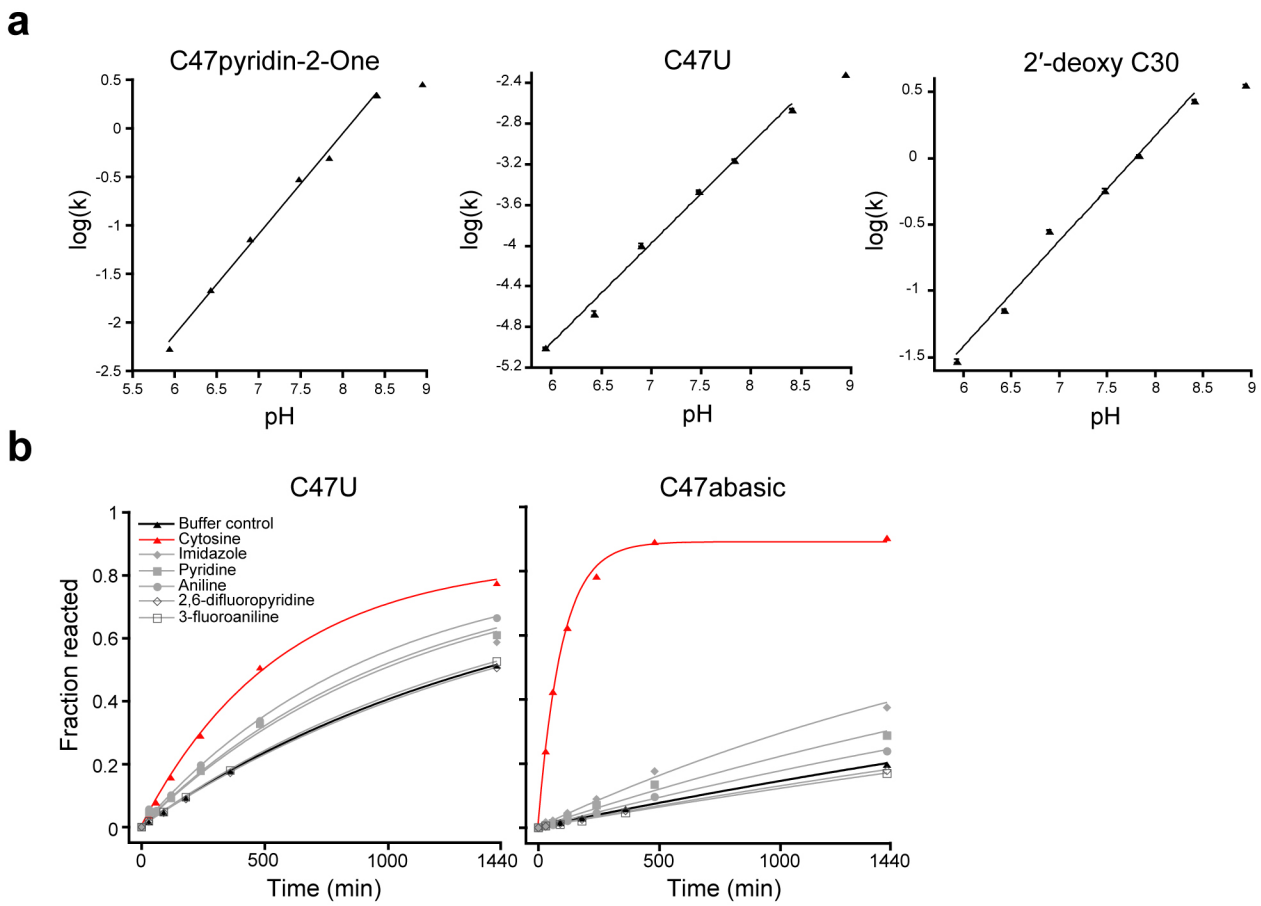
Supplementary Figures 1–5
Supplementary Discussion
Supplementary Methods



Supplementary Figure 1. Comparison of the preligation and product structures. **a.** Stereo views of 2|Fo| - |Fc| electron density maps for (left) the Mg^{2+}–Sr^{2+}- and (right) Ca2+–Sr2+-preligation complexes, following molecular replacement and initial rounds of simulated annealing. Maps are contoured at 1.2σ and superposed over the refined models of the P3–P6–P7 domains, colored as in **Figure 1b**. **b.** Superposition of the Ca2+–Sr2+-preligation structure and the product structure (PDB ID: 3HHN)[28]. Shown are backbone traces of all-atom superpositions, excluding the U1A protein and its cognate loop. Residues near the 5′-end of J1/3, colored light gray, were poorly ordered. **c.** Superposition of the Ca^{2+}–Sr^{2+}-preligation complex and the product structure, as viewed from the active site. The product structure is depicted as black lines. The preligation structure is depicted as sticks colored as in **Figure 1d.** In the Ca^{2+}–Sr^{2+}-complex, widening of the P2 helix weakens the base pair between G1 and the template residue C12, increasing the distance between their C1′-carbons by 0.3 Å and rotating the C12 base 24° away from ideal pairing geometry. The G1 ribose is concomitantly shifted ~1 Å away from its position in the product structure, which, owing to the steric effects of the 5′-triphosphate, pushes the U47 base out of the active site by the same distance. **d.** Crystallographic disorder in the Ca^{2+}–Sr^{2+} structure. The 5′-end of P1–P2 domain, comprising helix P1 and the first nine residues of J1/3, is shown for each preligation structure. Meshes are simulated-annealing |Fo| – |Fc| OMIT maps, contoured at 3.5σ, in which the residues depicted as sticks were excluded from the map calculations. These residues were colored in gray in **b.**

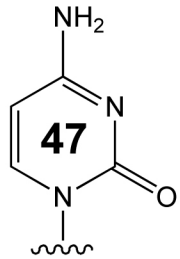


Supplementary Figure 2. Phosphorothioate interference and rescue data. **a.** Self-ligation timecourses. Ligation timecourses for the unmodified ligase, and for each of the phosphorothioate constructs tested, in the absence (red curves, 50 mM $MgCl_2$ alone) and presence (blue curves, 50 mM $MgCl_2$ + 2 mM $CdCl_2$) of rescuing metal. Data have been renormalized to the calculated theoretical maximum fraction reacted, F_M, and plotted on the same scale. All data are the mean values of three independent experiments; error bars correspond to the standard deviation. The inset for the 30*Rp*-S experiment shows the complete timecourse in absence of Cd^{2+}. **b.** Primer-extension timecourses. Otherwise, as in **a.** Apparent rescue of the *Sp*-GTPα phosphorothioate substrate (itself not a dramatically inhibitory substitution[6] nor a position observed making structural contacts in our crystal structures) might be explained if inner-sphere coordination of Cd^{2+} induced a tautomeric shift in the α-phosphate, pulling the *Sp*-sulfur into an apparent thioether and ablating the inhibitory negative charge on the *Rp*-oxygen.

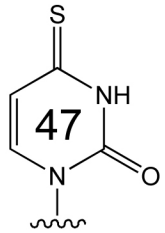


Supplementary Figure 3. Additional biochemical assays. **a.** pH dependencies for self-ligation of the C47P2o (left), C47U (middle) and 2′-deoxy C30 (right) variants. For each, the fit is log-linear with pH between 5.8 and 8.5, as was previously observed for the parental ligase[7]. Values are the mean of three independent experiments; error bars are standard deviations. **b.** Exogenous base-rescue experiments[36-38]. Self-ligation kinetics for the C47U (left) and C47abasic (right) variants, supplemented with exogenous small molecules. The axes are the same for both plots. Rescuing compounds were added at 35 mM, which approached the solubility limit for cytosine; kinetics at the solubility limits of the other rescuing compounds, or at 400 mM imidazole, were indistinguishable from those shown here (data not shown). Exogenous cytosine (red) rescued activity of the C47abasic construct (starting k_{rel} <10^5 below that of the parent) 33-fold, while only rescuing the C47U mutant ~2.2 fold. Therefore, of the rescue observed with the C47abasic construct, ~15-fold can be attributed to the exogenous base functionally mimicking C47 (ref. 36). Because the cytosine N4 has an unmeasurable pKa, additional candidate rescuing compounds (gray) were selected based on their presentation of amino or imino groups known pKas. However, neither these analogs nor the general acid-base catalyst imidazole could specifically rescue the C47abasic ligase.

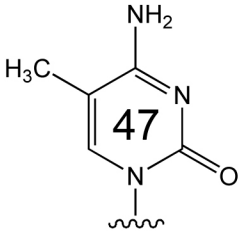
a



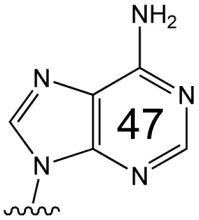
Cytosine (unmodified)
$k = 0.93 \pm 0.0091\ min^{-1}$



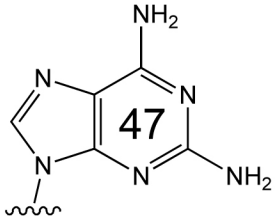
4-Thiouridine
$k_{rel_1} = 0.23 \pm 0.040$
$k_{rel_2} = 7.5\times10^{-5} \pm 1.7\times10^{-5}$



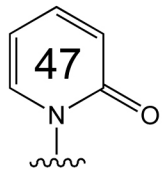
5-Methylcytosine
$k_{rel} = 0.091 \pm 0.0024$



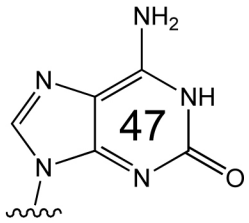
Adenosine
$k_{rel} = 0.033 \pm 8.7\times10^{-4}$



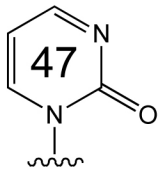
2,6-Diaminopurine
$k_{rel} = 0.011 \pm 5.1\times10^{-4}$



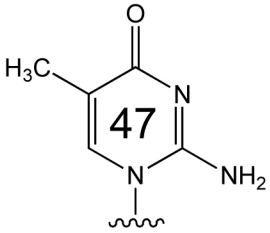
Pyridin-2-one
$k_{rel} = 5.6\times10^{-3} \pm 5.1\times10^{-4}$



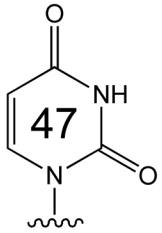
Isoguanosine
$k_{rel} = 2.4\times10^{-3} \pm 1.1\times10^{-4}$



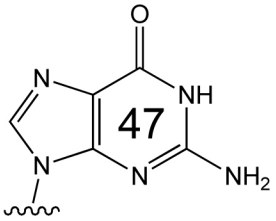
Zebularine
$k_{rel} = 1.6\times10^{-3} \pm 5.4\times10^{-5}$



5-Methylisocytosine
$k_{rel} = 1.6\times10^{-5} \pm 8.6\times10^{-8}$

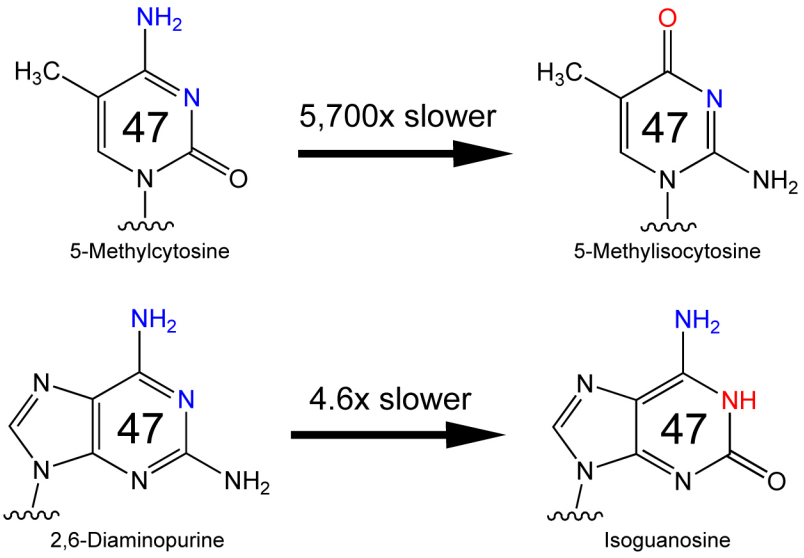


Uridine
$k_{rel} = 1.1\times10^{-5} \pm 1.3\times10^{-7}$

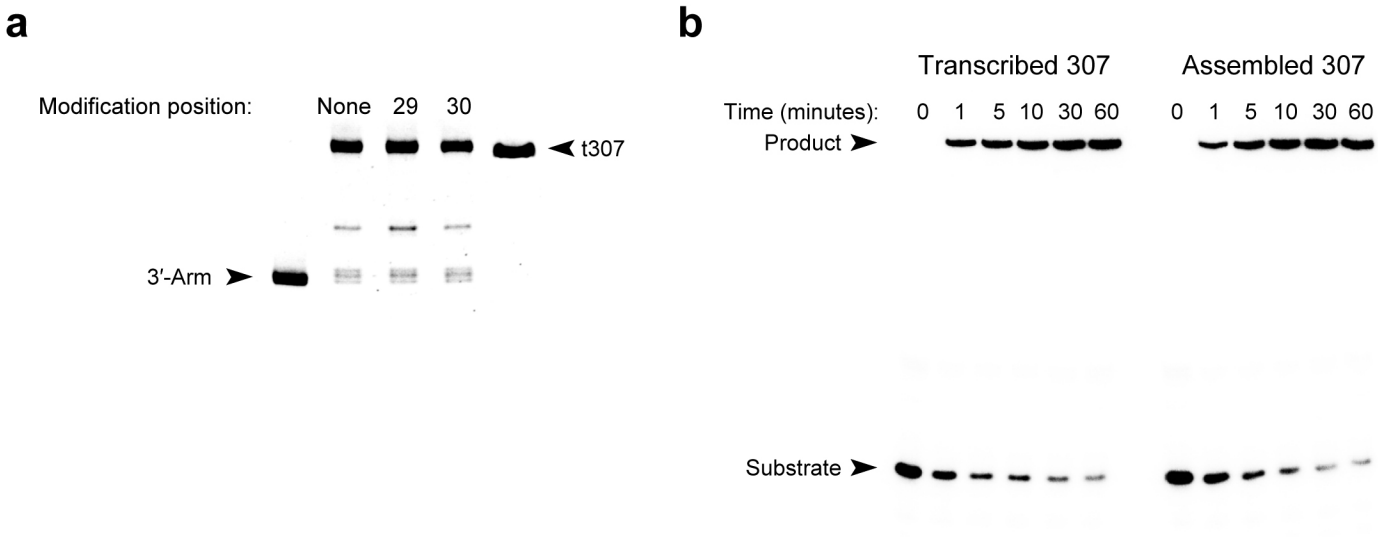


Guanosine
$k_{rel} = 2.3\times10^{-6} \pm 4.4\times10^{-8}$

b



Supplementary Figure 4. Variants used in the biochemical dissection of C47. **a.** Chemical structures and activities (scaled relative to the parent, top) of variants shown in **Figure 3b**, sorted in order of descending self-ligation rate. The 4-thiouridine variant ("^{4S}U") exhibited biphasic kinetics wherein ~10% reacted with a rate 4.4-fold slower than that of the parent, and the remainder reacted > 1000-fold slower. **b.** Evidence that the C47 N4 amine—and not the N3 imine—contributes to catalysis. Shown are comparisons between two pairs of isosteric variants that either maintain the N3 imine (top), but ablate the N4 amine, or maintain the exocyclic amine but protonate the imine (bottom). In each, numbers above the arrows indicate the fold reduction in first-order rate constants incurred by introducing the given modification (highlighted in red). Although replacement of the amine with a carbonyl group is extremely deleterious, protonation of the imine was comparably innocuous. This interpretation of these results assumes that the purines can be accommodated within the active site such that the N6 exocyclic amine can play the same role as a pyrimidine N4 exocyclic amine. Analogous accommodation has been proposed in the active sites of the HDV and VS ribozymes[37,39,40]. Previous results from a modification interference study provide further evidence that the C47 N3 lacks a direct catalytic role, because methylation of C47 N3 did not perturb activity[28,41].



Supplementary Figure 5. Assembly of active ligases by DNA-splinted ligation. **a.** Efficient DNA-splinted assembly of modified t307 species, as analyzed on a denaturing polyacrylamide gel stained with ethidium bromide. Leftmost lane: the 3′-Arm oligonucleotide, the longest of the RNAs used in assembly. Middle three lanes: end points (4 h) of DNA-splinted assembly reactions; products are t307 derivatives bearing phosphorothioate modifications at the indicated positions. Rightmost lane: *in vitro*-transcribed, unmodified t307. This material runs slightly faster than the assembled t307 products because it terminates with a 2′-3′ cyclic phosphate and lacks 3′- transcription heterogeneities. **b.** Activity of transcribed and assembled ligases. Shown are self-ligation timecourses using a 5′-^{32}P-labeled oligonucleotide substrate. Left lanes: T7 *in vitro*-transcribed ligase construct 307. Right lanes: the equivalent species generated by DNA-splinted assembly. The observed rates of these samples differed by approximately two-fold.

Supplementary Discussion

Interpretation of the kinetic isotope results. Although the linear proton inventories shown in **Figure 4a** are consistent with our model in which a single proton is transferred during catalysis, these results have alternative explanations and thus do not conclusively determine the number of protons in transit. Nonetheless, we argue that because the apparent number of protons in transit is unchanged among each of the ligase variants tested, neither the C47 N4 amine nor the C30 2′-hydroxyl functions by proton transfer. The discussion below supports this assertion in the contexts of alternative explanations for the solvent kinetic isotope results.

Similar kinetic isotope results have been explained by either **1**) an unanticipated rate-determining conformation change dependent on proton transfer[1], **2**) differential enzyme protonation (deuteration) states–and not catalytic rates–in H_2O and D_2O[2-4], or **3**) interactions between the isotopic waters and catalytic cofactors that account for the apparent SDKIE of a catalytic reaction devoid of proton transfer[2]. For the ligase, model **1** can be ruled out because catalysis, and not a conformation change, is known to be rate limiting under the conditions used here[5-7]. However, models **2** or **3** cannot be conclusively ruled out.

Model **2** stipulates that, because pKa values differ between D_2O and H_2O[3,8], measured kinetic isotope effects inevitably comprise contributions both from the true kinetic isotope effects (that is, the change in catalytic rate incurred by the change in isotope) and from the effects incurred by protonating (deterating) different proportions of the enzyme population in each solvent[3,9]. This issue can be circumvented by performing experiments at a pH-rate "plateau," such that the protonation (deuteration) state does not contribute to the catalytic rate[2-4,10-13], or can be partially alleviated by the method of velocity-derived "equivalent pL's" (where L is either H or D) described in Ref. 3. However, because the ligase rate increases over the entire pH range in which it is expected to properly fold[7], neither method is

applicable. Thus, we can neither compensate for nor discretely measure the degree to which the observed kinetic isotope effects are determined by a difference in enzyme pKa between the two isotopic waters. Nonetheless, this ΔpKa effect could not explain our results if the ligase ΔpKa was near to those of the buffer systems employed (ΔpKa ~ 0.6 for both MES and Tris buffers[8]), as is often the case in other systems[3]. Moreover, the atomic mutagenesis data support the notion that our measured proton inventories were not purely the products of enzyme protonation (deuteration) effects, since the removal or replacement of active-site functional groups did not alter the ribozyme's pH-dependence (and hence the apparent pKa, **Supplementary Fig. 3a**), but significantly altered the SDKIE. If the apparent isotope effect parameters were entirely artifactual consequences of the ligase ΔpKa, we would have expected these parameters, like the pH-dependence, to be unchanged among the variants surveyed. We therefore infer that the ΔpKa effects proposed by model **2** are not the principle determinants of the observed proton inventories, thus supporting the notion that these experiments provide direct readout of ribozyme catalytic function. But even if this were not the case, and the apparent isotope effect parameters were entirely artifactual consequences of the ligase ΔpKa, then there would be no protons in transit during the transition states of our wild-type and mutant constructs, and thus neither the C47 N4 amine nor the C30 2′-hydroxyl could function by proton transfer.

Model **3** has been best illustrated in minimal truncation systems derived from the hammerhead ribozyme[2,14,15]. Like the ligase, these constructs also exhibit a log-linear rate-pH profile with a slope of unity, and display an apparent SDKIE of 4.4 in a simple equivalent pL setting (that is, pH = pD) in the presence of Mg^{2+} (Ref. 2). These results were interpreted as supporting a model in which differential protonation (deuteration) of a catalytic Mg^{2+}-hydroxide complex, and not catalytic proton transfer, was responsible for the observed isotope effect[2,16]. The authors furthermore estimated that, for reactions requiring the activity of a $Mg^{2+}OH$ complex, an apparent SDKIE of ~4.5 would be observed simply due

to the differential population of active (metal-hydroxide) versus inactive (metal-hydrate) complexes in different isotopic waters[2]. Although subsequent structural and mechanistic work has proven this model to be incorrect for the hammerhead ribozyme (in that restoration of a distal tertiary interaction has a dramatic effect on the pH and metal dependencies[17], establishing an active site that employs general acid-base catalysts[18,19] but only ancillary use of metal ions[20]), the possibility merits consideration for the ligase ribozyme. Nonetheless, we disfavor model **3** for the ligase. Model **3** attributes an SDKIE of ~4.5 to the function of $Mg^{2+}OH$ alone[2], and although this effect matches the effect observed for the wild type ligase (~5.0 ± 0.3), it exceeds those of the C47U, C47P2o and C30Deoxy variants (~2.9 ± 0.1, 2.4± 0.1 and 3.1± 0.2, respectively). Since each of these variants exhibits the same pH-dependence, monophasic kinetics and fractions of active molecules as the unmodified ribozyme, we infer that all of the constructs surveyed react through the same reaction pathway and presumably make equivalent use of the proposed catalytic Mg^{2+} ion. Thus, were the observed SDKI effects and proton inventories entirely the result of a differential metal-hydroxide/hydrate balance, the measured SDKIE would be equal for all species. Because these values differ, we conclude that the effects predicted by model **3** are insufficient to describe the kinetic isotope data. But even if this were not the case, and the apparent isotope effect parameters were entirely consequences of the differential population of active (metal-hydroxide) versus inactive (metal-hydrate) complexes, then during the transition state no protons in transit would be attributable to either the C47 N4 amine or the C30 2′-hydroxyl. Hence, neither group functions by proton transfer.

Supplementary Methods

Crystallographic model-building and refinement procedures. All refinement steps followed the "individual ADP" strategy in PHENIX[21], employing sequential steps of automated bulk-solvent correction, positional and individual atomic *B*-factor refinement during each round of refinement. Target geometric weights (the "wxc_scale" in PHENIX[21]) were held at relatively restrictive values of 0.05–0.01; the "wxu_scale" was held at 1.0 throughout all refinement steps. To remove model bias from the molecular replacement solutions, the first three rounds of refinement also employed rounds of simulated annealing (5000K to 300K, in 100K steps) and rigid-body refinement, defining the individual ligase helices (P1, P2, etc…), the U1A protein, its cognate loop, and 3–6-nt subdivisions of the unpaired joining regions (J1/2, J1/3 and J3/4) as independent rigid bodies. Neither NCS-averaging nor TLS was used during either refinement.

The $|F_O| - |F_C|$ difference fourier maps resulting from initial refinement rounds exhibited prominent (7–10σ) peaks in the vicinity of nucleotide G1 (**Figs. 1b** and **2a,b**), corresponding to the 5′-triphosphate. This 5′-GTP was built by hand in COOT[22], and the resulting model was subjected to another round of simulated annealing in addition to the ADP regimen outlined above. Fully- or partially-hydrated metal ions outside the active site were built next, their degree of hydration assigned by first placing dehydrated metals into strong (>3.5σ) $|F_O| - |F_C|$ difference peaks, and inspecting the resultant temperature factors and difference peaks after a subsequent round of refinement[23,24]. This process was repeated iteratively, until no strong peaks remained. Active-site metal ions were placed last, using this same strategy. Since crystal nucleation required the presence of Sr^{2+} ions, during model building, conspicuous Fourier difference peaks were tested as potential strontium sites. In each case, however, inspection of the coordination geometry, temperature factors, R_{free} values and difference fourier peaks in subsequent rounds of refinement eliminated this possibility. Hence, neither final model

contains a single well-ordered Sr^{2+}-ion, which is consistent with the observation that strontium is a poor competitive inhibitor with magnesium[7].

Owing to the resolution limits, metal ions and their bound waters could not be individually refined in a chemically sensible manner. So, partially- or fully-hydrated metal clusters were treated as monomers during refinement, having defined model residues with idealized bond lengths and geometries using the SKETCHER program the CCP4 program suite[25]. Magnesium–water bond distances were defined as 2.07 Å; hexacoordinate geometry was used in for all magnesium monomers[26,27]. Calcium–water bond distances were similarly fixed at 2.43 Å. However, owing to the variable coordination geometries of calcium hydrate clusters[26,27], bound waters were not built unless their geometry with respect to the calcium ion was clearly indicated in difference Fourier maps. For the metal-hydrate clusters bound to the P6 A-minor triad[28] (residues A31 and A32), bond distances were restrained to their ideal values, and the angle was fixed at 90°; none of the other metal hydrates were fixed with regard to the neighboring macromolecule.

Each monomer in the final Mg^{2+}-bound structure contains all 130 nucleotides of the ligase core, all seven nucleotides of the substrate, and all but the first 6–7 amino acids of the U1A protein. The occupancies have been set to zero for protein side chains lacking electron density. Despite high temperature factors in the Ca^{2+}–Sr^{2+}-complex structure (**Supplementary Fig. 1d**), refining the occupancies of, or altogether removing problematic nucleotides resulted in an increase in the R_{free} value. Hence, the final structure consists of all nucleotides of the ligase core and substrate.

Structural figures were made in PyMol (http://www.pymol.org, Schrödinger LLC).. Morphing movies were created in MORPH2 (Martz, E. http://www.umass.edu/microbio/rasmol/pdbtools.htm #martz) and PyMol.

Synthetic RNA and DNA oligonucleotides. Most of the RNA oligonucleotides used for the splinted assembly of modified ligases (see below) were purchased from Dharmacon. Oligonucleotides bearing

the pyridin-2-one and abasic modifications were purchased from Trilink Biotechnologies. Those bearing zebularine, 2′-deoxy isoguanosine and 2′-deoxy 5-methylisocytosine modifications were purchased from the W.M. Keck Foundation Biotechnology Resource Laboratory at Yale University (hereafter "Yale Keck"), using phosphoramidites from Glen Research. DNA transcription template oligonucleotides bearing 5′-terminal 2′-O-methyl modifications (see below) were also synthesized by the Yale Keck facility. All other DNA oligonucleotides were purchased from Integrated DNA Technologies (IDT).

The reagents from Trilink and IDT did not require deprotection or further purification before use. DNA oligonucleotides from Yale Keck were purified from denaturing 10% polyacrylamide gels as described for the RNA substrate oligonucleotide. RNA oligonucleotides from Yale Keck were deprotected with 55.6% (v/v) triethylamine trihydrofluoride (TEA.3HF) in anhydrous DMSO for 2.5 hr at 65ºC. Reaction mixtures were then briefly chilled, brought to 300 mM sodium acetate and precipitated by addition of five volumes of isopropanol and overnight incubation at –20ºC. Following centrifugation, pellets were washed twice with absolute ethanol, then dried and resuspended in water. Oligonucleotides from Dharmacon were deprotected according to that manufacturer's instructions, and in biochemistry assays were used without further purification. The purity and mass of all modified RNA oligonucleotides was confirmed by MALDI mass spectrometry. All reagents were stored at –20ºC until use.

Rp- and *Sp*-isomers of phosphorothioate-modified oligonucleotides were purified by reversed-phase HPLC on a C18 column, using conditions in which the *Rp*-isomer is known to elute first[29,30]. The purity of isolates was confirmed by RP-HPLC. Pure samples were pooled, and freed of excess salt by passing through several cycles of evaporation (speed-vac) and resuspension in pure water, and stored at –20ºC until use.

Phosphorothioate GTP analogs. Stereopure *Rp*- and *Sp*-GTPαS were purchased from Axxora and used without further purification. GTPγS was purchased from Jena Biosciences and resuspended in 50 mM Tris, pH 7.5, before use. Stereopure *Rp*- and *Sp*-GTPβS were the generous gifts of Fritz Eckstein. Since approximately 10% of these samples had hydrolyzed to GDPβS, the intact GTP analogs were purified by ion-exchange HPLC using the same separation system described for their original synthesis[31]. The purity of each sample was confirmed by analytical ion-exchange HPLC; pure samples were pooled, evaporated to dryness in a speed-vac (using methanol as a co-solvent) and resuspended in 50 mM Tris, pH 7.5 before use. All GTP analogs were stored at -20°C

Transcription templates for RNAs used in biochemical assays. The plasmid template for the primer-extending variant ("t307") was generated by QuickChange mutagenesis from pH307HP (ref. 28), which brackets the ligase core with 5′-hammerhead (HH) and 3′-hepatitus delta virus (HDV) self-cleaving ribozymes. Excision of these ribozymes ensured that the final ligase species has homogeneous termini[32]. The relevant sequence of the insert was

GCGTAATACGACTCACTATA*GGAGAGTAGTATAGTGCTGATGAGTCCGTGAGGACGAAACGGTACCCG*
*GTACCGTC***C**ACTATACTACTGGATAATCAAAGACAAATCTGCCCGAAGGGCTTGAGAACATACCCATTG
CACTCCGGGTATGCAGAGGTGGCAGCCTCCGGTGGGTTAAAACCCAACGTTCTCAACAATAGTG**A***GGCCC*
GGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATTCCGAGGGGACCGTCCCCTCGGTAATG
GCGAATGGGACCCAC

where highlighted letters indicate the T7 promoter, bold nucleotides denote the 5′ and 3′ ends of the mature ligase species, italicized nucleotides denote the 5′-HH and 3′-HDV ribozymes and underlined nucleotides denote the U1A modification in stem P5. The plasmid was linearized by digestion with EarI endonuclease before use.

Templates for the C47A, C47G and C47Δ variants of the self-ligating construct, and the "3′-Arm" species used in assembling modified ligases by splinted ligation (see below) were PCR products generated from overlapping DNA oligonucleotides[33]. Templates for the active-site mutants did not include self-cleaving ribozymes; that for the "3′-Arm" construct consisted of a 5′-HH ribozyme followed by nucleotides 52–121 of the ligase core, appended with the P5 U1A-loop extension. The full sequence of this construct was

GCGTAATACGACTCACTATA*GGAGATGGGTATGTTCCTGATGAGTCCGTGAGGACGAAACGGTACCCG*
*GTACCGTC***G**AACATACCCATTGCACTCCGGGTATGCAGAGGTGGCAGCCTCCGGTGGGTTAAAACCCAA
CGTTCTCAACAATAGTG**A**

annotated as above.

The template for the "5′-Arm" species used in assembling modified *cis*-ligases by splinted ligation (see below) was a DNA oligonucleotide modified with two 2′-O-methyl nucleotides at its 5′-terminus. These modifications have been shown to limit the 3′-heterogeneity of transcription products[34]. The template oligonucleotide was annealed to a 21-nt T7 promoter top strand DNA oligonucleotide and used as previously described[6,34].

Transcription and purification of RNAs used in biochemical assays. T7 *in vitro* transcription reactions were performed as described[33], in volumes of 0.05–1 mL. Templates were included at concentrations of 20 μg/mL in reactions using linearized plasmids, 6 μg/mL for those using PCR products, and 6 μg/mL for those using DNA oligonucleotides. Transcription reactions were quenched with the addition of EDTA, extracted with phenol and chloroform, ethanol precipitated and resuspended in water. Constructs employing self-cleaving ribozymes were subsequently desalted and refolded to allow full processing of the HH or HDV ribozymes, as previously described[28,32]. The terminal 2′-3′

cyclic phosphate was not removed from these species. All RNAs were purified from denaturing 6% or 20% 0.5X TBE polyacrylamide gels, passively eluted into 300 mM NaCl at 4ºC overnight, ethanol precipitated, resuspended in water and stored at –20ºC before use.

Assembly and purification of modified ligases. Equimolar concentrations of all five RNA oligonucleotides and the DNA splint were combined in water to give a final concentration of 12 μM. The reaction mixture was annealed by heating and cooling (85ºC and 37ºC, for five minutes each), supplemented with T4 RNA Ligase 2 (RNAl2, New England Biolabs) reaction buffer and incubated at 37ºC for 15 minutes before addition of RNAl2 (final concentration of 1 U/μL). After four hours at 37ºC, the reaction was quenched by the addition of EDTA to 60 mM and NaCl to 300 mM, extracted with phenol and ethanol precipitated. Pellets were resuspended in water, brought to 5 M urea and purified on denaturing 6% 0.5X TBE polyacrylamide gels. RNA was passively eluted into 300 mM NaCl overnight at 4ºC, ethanol precipitated, resuspended in water and stored at –20ºC before use.

Under these reaction conditions, the assembly reaction appears nearly quantitative, as monitored by ethidium bromide-stained gels (**Supplementary Fig. 5a**). However, due to the inefficiency of elution, reactions starting with 2 nmol of oligonucleotide substrates typically yielded 0.5–0.8 nmol fully assembled product. Unmodified ligase species assembled in this fashion exhibited activities indistinguishable from those that had been transcribed whole *in vitro* (**Supplementary Fig. 5b**).

Kinetic Assays. Kinetic assays used substrate oligonucleotides that had been 5′-radiolabeled using PNK and [^{32}P]γ-ATP. Unless otherwise noted, all kinetic assays were performed in 50 mM buffer (sodium cacodylate for pH 6.0–6.5; Tris for pH 7.0–9.0), 10 mM $MgCl_2$, 200 mM KCl and 600 μM EDTA. Self-ligation assays were performed as follows[7]: the ribozyme was heated (5 minutes, 80ºC) and

cooled (5 minutes, 22ºC) in water, and reactions were initiated at 22ºC by the simultaneous addition of buffer (pH 6.0, unless otherwise indicated), salts and substrate oligonucleotide. The final ligase concentration was 1 μM; substrate was added to a final concentration of 0.5 μM. Primer-extension assays were performed using a nearly identical protocol[6], except that the reaction was initiated by the simultaneous addition of buffer (pH 7.5), salts, substrate oligonucleotide and 1 mM GTP (or analog). Furthermore, because NTPs chelate Mg^{2+}, the reactions were supplemented with additional $MgCl_2$, such that the supplemental Mg^{2+} concentration equaled that of the NTP[6].

For all reactions, aliquots were taken at the specified time points and quenched by mixing with an equal volume of gel-loading buffer (8 M urea, 120 mM EDTA, trace bromphenol blue and xylene cyanol). Samples were separated on denaturing 0.5X TBE 20% polyacylamide gels. Gels were visualized by phosphorimaging (Fujifilm BAS-2500). For each time point the fraction product was measured as

$$F_P(t) = \frac{P(t)}{(P(t) + R(t))}$$

where $P(t)$ and $R(t)$ are the product and reactant at a given time, respectively, and fit to the equation

$$F_P(t) = F_M(1 - e^{-k_{OBS}t})$$

were t equals time, treating F_M (the maximum fraction reacted) and k_{OBS} (the observed rate constant) as unknowns.

Compounds used in exogenous base rescue experiments (**Supplementary Fig. 3b**) were prepared in 50 mM Tris•HCl buffer, pH 8.0. Rescuing compounds were added to ligation reactions concomitantly with the substrate, buffer and salts. In every other regard, these timecourses were implemented and analyzed as for the standard self-ligation assays described above. Base rescue was calculated as follows

$$Rescue = \frac{\left(\frac{^{+}k_{abas}}{^{-}k_{abas}}\right)}{\left(\frac{^{+}k_{C47U}}{^{-}k_{C47U}}\right)}$$

where $^{+}k_{abas}$ and $^{-}k_{abas}$ are the apparent first-order self-ligation rate constants for the C47abasic construct in the presence and absence of rescuing base, respectively, and $^{+}k_{C47U}$ and $^{-}k_{C47U}$ are the comparable rate constants for the C47U construct.

Phosphorothioate interference and metal rescue assays. Phosphorothioate interference was measured by the standard self-ligation (for modifications at positions 29, 30 or the GTP1γ phosphate) or primer-extension (for modifications at all GTP phosphate positions) reactions described above. However, the concentration of $MgCl_2$ was 50 or 51 mM for self-ligation and primer-extension experiments, respectively. Interference values were calculated as follows:

$$Interference = \frac{^{Mg}k_S}{^{Mg}k_{oxy}}$$

where $^{Mg}k_s$ and $^{Mg}k_{oxy}$ are the rate constants for the thio-modified and unmodified constructs, respectively, in magnesium. Rescue values were calculated as follows:

$$Rescue = \frac{\left(\frac{^{X}k_S}{^{Mg}k_S}\right)}{\left(\frac{^{X}k_{oxy}}{^{Mg}k_{oxy}}\right)}$$

where $^{X}k_s$ and $^{X}k_{oxy}$ are the rate constants for thio-modified and unmodified constructs, respectively, in magnesium supplemented with a rescuing metal, X. Although potential rescue conditions were screened at 10, 50 and 100 mM $MgCl_2$, in the presence of 0.1–25mM $CdCl_2$, $CoCl_2$, $MnCl_2$ or $ZnCl_2$, the most significant rescue observed was at 50 mM $MgCl_2$, supplemented with 2 mM $CdCl_2$. All values quoted in **Figure 2e** were collected under these conditions.

Kinetic Isotope Experiments. To determine solvent isotope effects, all reaction components (ribozymes, substrates, buffers, $MgCl_2$, KCl, EDTA) were reformulated in parallel in H_2O and D_2O. Aliquots of ribozyme and radiolabeled substrate in H_2O were dried overnight in a speed-vac, resuspended in D_2O, quantified by UV-vis spectroscopy and stored at –80ºC until use. Buffers were brought to pL (pH or pD) by titration with the conjugate acid and base of the buffer system. MES (free acid, sodium salt) was used for experiments at pL 6.0 and Tris (hydrochloride, free base) for experiments at pL 8.0. In all cases, the pL was measured using a glass electrode; for buffers in D_2O, the apparent pD was corrected by adding a value of 0.4 to the instrument readings[35].

Other than the experimental variation of H_2O and D_2O, proton inventory experiments were performed under standard self-ligation conditions: 50 mM Buffer (pL 6.0 or 8.0), 10 mM $MgCl_2$, 200 mM KCl, 600 μM EDTA, 1 μM ribozyme and 0.5 μM substrate. Ribozyme samples in D_2O and H_2O were combined in appropriate ratios to produce mixtures in 100%, 80%, 60%, 40%, 20% and 0% D_2O. Substrate oligonucleotides, buffers and salts were similarly combined to obtain a parallel series of reaction start mixes at the same fractions of D_2O. Self-ligation reactions were then performed, and at each molar fraction of D_2O (termed "*n*") the apparent rate constant, k_n, was calculated. Proton inventories were calculated by plotting the ratio of k_n to the apparent rate constant in pure water, k_{H2O}, as a function of *n*, and fitting the data to the modified Gross-Butler equation[9] for either a two-proton-transfer model,

$$\frac{k_n}{k_{H2O}} = (1 - n + n\cdot \Phi_1)(1 - n + n\cdot \Phi_2)$$

or for a one-proton-transfer model,

$$\frac{k_n}{k_{H2O}} = (1 - n + n\cdot \Phi)$$

where each Φ is the inverse of the of the SKIE for an individual ionizable group.

Supplementary References

1. Tinsley, R.A., Harris, D.A. & Walter, N.G. Significant kinetic solvent isotope effects in folding of the catalytic RNA from the hepatitis delta virus. *J Am Chem Soc* **125**, 13972-3 (2003).

2. Sawata, S., Komiyama, M. & Taira, K. Kinetic evidence based on solvent isotope effects for nonexistence of proton-transfer process in reactions catalyzed by a hammerhead ribozyme: implication to the double-metal-ion mechanism of catalysis. *J Am Chem Soc* **117**, 2357-2358 (1995).

3. Schowen, K.B. & Schowen, R.L. Solvent isotope effects of enzyme systems. *Methods Enzymol.* **87**, 551-606 (1982).

4. Jencks, W.P. *Catalysis in chemistry and enzymology*, xvi, 644 p. (McGraw-Hill, New York,, 1969).

5. Bergman, N.H., Johnston, W.K. & Bartel, D.P. Kinetic framework for ligation by an efficient RNA ligase ribozyme. *Biochemistry* **39**, 3115-23 (2000).

6. Glasner, M.E., Yen, C.C., Ekland, E.H. & Bartel, D.P. Recognition of nucleoside triphosphates during RNA-catalyzed primer extension. *Biochemistry* **39**, 15556-62 (2000).

7. Glasner, M.E., Bergman, N.H. & Bartel, D.P. Metal ion requirements for structure and catalysis of an RNA ligase ribozyme. *Biochemistry* **41**, 8103-12 (2002).

8. Prini, R.F., De Pattin, C.R., Tanak, K. & Bates, R.G. Acidic dissociation constants of weak bases in D2O from measurements with ion-selective electrodes. *J. Electroanal. Chem.* **144**, 415-422 (1982).

9. Venkatasubban, K.S. & Schowen, R.L. The proton inventory technique. *CRC Crit Rev Biochem* **17**, 1-44 (1984).

10. 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* **130**, 14504-20 (2008).

11. Nakano, S. & Bevilacqua, P.C. Proton inventory of the genomic HDV ribozyme in Mg(2+)-containing solutions. *J Am Chem Soc* **123**, 11333-4 (2001).

12. Shih, I.H. & Been, M.D. Involvement of a cytosine side chain in proton transfer in the rate-determining step of ribozyme self-cleavage. *Proc Natl Acad Sci U S A* **98**, 1489-94 (2001).

13. Smith, M.D. & Collins, R.A. Evidence for proton transfer in the rate-limiting step of a fast-cleaving Varkud satellite ribozyme. *Proc Natl Acad Sci U S A* **104**, 5818-23 (2007).

14. Takagi, Y. & Taira, K. Analyses of kinetic solvent isotope effects of a hammerhead ribozyme reaction in NH4+ and Li+ ions. *Nucleic Acids Res Suppl*, 273-4 (2002).

15. Takagi, Y. & Taira, K. Detection of a proton-transfer process by kinetic solvent isotope effects in NH(4)(+)-mediated reactions catalyzed by a hammerhead ribozyme. *J Am Chem Soc* **124**, 3850-2 (2002).

16. Takagi, Y., Warashina, M., Stec, W.J., Yoshinari, K. & Taira, K. Recent advances in the elucidation of the mechanisms of action of ribozymes. *Nucleic Acids Res* **29**, 1815-34 (2001).

17. Roychowdhury-Saha, M. & Burke, D.H. Extraordinary rates of transition metal ion-mediated ribozyme catalysis. *RNA* **12**, 1846-52 (2006).

18. Thomas, J.M. & Perrin, D.M. Probing general acid catalysis in the hammerhead ribozyme. *J Am Chem Soc* **131**, 1135-43 (2009).

19. Thomas, J.M. & Perrin, D.M. Probing general base catalysis in the hammerhead ribozyme. *J Am Chem Soc* **130**, 15467-75 (2008).

20. Martick, M., Lee, T.S., York, D.M. & Scott, W.G. Solvent structure and hammerhead ribozyme catalysis. *Chem Biol* **15**, 332-42 (2008).

21. Adams, P.D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-54 (2002).

22. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32 (2004).

23. Stahley, M.R., Adams, P.L., Wang, J. & Strobel, S.A. Structural metals in the group I intron: a ribozyme with a multiple metal ion core. *J Mol Biol* **372**, 89-102 (2007).

24. Klein, D.J., Moore, P.B. & Steitz, T.A. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **10**, 1366-79 (2004).

25. The CCP4 suite: programs for protein crystallography. *Acta Crysallogr D Biol Crystallogr* **50**, 760-3 (1994).

26. Harding, M.M. The geometry of metal-ligand interactions relevant to proteins. II. Angles at the metal atom, additional weak metal-donor interactions. *Acta Crystallogr D Biol Crystallogr* **56**, 857-67 (2000).

27. Harding, M.M. Geometry of metal-ligand interactions in proteins. *Acta Crystallogr D Biol Crystallogr* **57**, 401-11 (2001).

28. Shechner, D.M. *et al.* Crystal structure of the catalytic core of an RNA-polymerase ribozyme. *Science* **326**, 1271-5 (2009).

29. Slim, G. & Gait, M.J. Configurationally defined phosphorothioate-containing oligoribonucleotides in the study of the mechanism of cleavage of hammerhead ribozymes. *Nucleic Acids Res* **19**, 1183-8 (1991).

30. Frederiksen, J.K. & Piccirilli, J.A. Separation of RNA Phosphorothioate Oligonucleotides by HPLC. *Methods Enzymol.* **468**, 289-309 (2009).

31. Connolly, B.A., Romaniuk, P.J. & Eckstein, F. Synthesis and characterization of diastereomers of guanosine 5'-O-(1-thiotriphosphate) and guanosine 5'-O-(2-thiotriphosphate). *Biochemistry* **21**, 1983-9 (1982).

32. Price, S.R., Ito, N., Oubridge, C., Avis, J.M. & Nagai, K. Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *J Mol Biol* **249**, 398-408 (1995).

33. Ekland, E.H. & Bartel, D.P. The secondary structure and sequence optimization of an RNA ligase ribozyme. *Nucleic Acids Res* **23**, 3231-8 (1995).

34. Kao, C., Zheng, M. & Rudisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5**, 1268-72 (1999).

35. Glasoe, P.K. & Long, F.A. Use of glass electrodes to measure acidities in deuterium oxide. *J Phys Chem* **64**, 188-190 (1960).

36. Peracchi, A., Beigelman, L., Usman, N. & Herschlag, D. Rescue of abasic hammerhead ribozymes by exogenous addition of specific bases. *Proc Natl Acad Sci U S A* **93**, 11522-7 (1996).

37. Perrotta, A.T., Shih, I. & Been, M.D. Imidazole rescue of a cytosine mutation in a self-cleaving ribozyme. *Science* **286**, 123-6 (1999).

38. Lebruska, L.L., Kuzmine, II & Fedor, M.J. Rescue of an abasic hairpin ribozyme by cationic nucleobases: evidence for a novel mechanism of RNA catalysis. *Chem Biol* **9**, 465-73 (2002).

39. Nakano, S., Chadalavada, D.M. & Bevilacqua, P.C. General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* **287**, 1493-7 (2000).

40. Zhao, Z.Y. *et al.* Nucleobase participation in ribozyme catalysis. *J Am Chem Soc* **127**, 5026-7 (2005).

41. Bagby, S.C., Bergman, N.H., Shechner, D.M., Yen, C. & Bartel, D.P. A class I ligase ribozyme with reduced Mg2+ dependence: Selection, sequence analysis, and identification of functional tertiary interactions. *RNA* **15**, 2129-46 (2009).