Supplementary Information

Nucleobase Modification by an RNA Enzyme

Raghav R. Poudyal,[1,2] Phuong D. M. Nguyen,[1,2] Melissa P. Lokugamage,[2,3] Mackenzie K. Callaway,[2,3] Jesse V. Gavette,[4] Ramanarayanan Krishnamurthy,[4] Donald H. Burke[1,2,3,5§]

[1]Dept. of Biochemistry
[2]Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA
[3]Dept. of Biological Engineering
[4]Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA
[5]Dept. of Molecular Microbiology and Immunology

§ Corresponding author: email BurkeDH@Missouri.edu

Table of Contents

Supplementary Text:

Prior work from our group that assigned G1 as the acceptor nucleotide assumed a 2′OH phosphorylation product (1) (which we now know to have been an incorrect assumption, based on the reactions reported here with all-DNA acceptor substrates). Briefly, the previous experiments allowed 5′-radiolabeled ribozyme to react with GTPγS, then subjected it to alkaline cleavage. All 5′-radiolabeled fragments representing cleavage after G2 carried the adduct and were visible on the gel. Because 2′-*O*-thiophosphorylation blocks alkaline digestion, these observations were consistent with 2′-*O*- thiophosphorylation of G1. However, the results in the present study establish that adduct formation does not require 2′OH in the acceptor strand, and thiophosphorylation of the G2 nucleobase would not have prevented alkaline digestion at G2. That same earlier study also observed termination of reverse transcription after inserting a base across from A3 in an all-RNA, cis-acting ribozyme, as observed here for two different all-DNA substrates. Prior work had shown that 2′-*O*-thiophosphorylation normally causes RT to pause one position prior to the modification (2), which suggested G2 as the modification site, seemingly in conflict with the alkaline digestion pattern noted above. To resolve this conflict, we noted that it was also known that RT binds only very weakly to template-primer complexes with short 5′ overhangs, such as those encountered near the 5′ end of an RNA template. Given the compelling observation of the adduct on the GG dinucleotide and the assumption of 2′-*O*-phosphorylation, the failure to insert a nucleotide across from G2 was interpreted as being due to a combination of the covalent modification on G1 and reduced affinity near 5′ ends. The present study resolves each of those conundrums, identifying the nucleobase (not the 2′O) of G2 (not G1) as the acceptor.

Supplementary Methods

Thiophosphorylation reactions with Inosine, 2-Aminopurine and 7-deazaGuanosine substitutions at position G2: 10,000-20,000 cpm of [5′ ^{32}P] end-labeled acceptor substrate strand (50-166 nM) was added to 2 μM Ribozyme strand. Ribozyme:Substrate mixture was heated at 85 °C for 5 min. 5X Thiophosphorylation buffer (1X = 25 mM Tris pH 8.0, 30 mM $MgCl_2$, 10 μM $CuCl_2$, 200 mM KCl and 15 mM NaCl) was added and the tubes were incubated at room temperature for 5 min and moved to ice to form the Ribozyme:Substrate complex. Reactions were initiated by adding 1 mM GTPγS (final concentration). Tubes were then incubated for the 10 °C for indicated times.

Time course for mass spectrometry : Five identical 20 μL reactions were assembled, each consisting of 110 μM ribozyme, 100 μM substrate, 1X T8M30 buffer (1X=25 mM Tris pH 8.0, 30 mM MgCl2, 200 mM KCl, 15 mM NaCl, and 10 μM CuCl2). Ribozyme and substrate were suspended in water, heated at 85 °C for 3 min and 5X T8M30 buffer was added. Reactions were left at room temperature for 5 min, followed by addition of GTPγS to a final concentration of 200 μM. Reactions were then incubated at 10°C, and quickly frozen in dry ice ethanol bath at indicated time points. 80 μL of water was added to each tube, followed by 25 μL of 7.5 M NH_4OAc. 300 μL of 100% ethanol was added and samples were moved to -80°C for 5 min. Samples were then centrifuged for 30 min and supernatants were decanted, followed by resuspension in 1 M NH_4OAc and precipitation with 100% ethanol. This was done to remove excess sodium. Samples were finally resuspended in 20 μL of water, followed by MALDI TOF. Percent product yield for 7321 or 7225 product was calculated as Peak height for 7321 or 7225/ Sum of peak heights for unreacted substrate, 7321 product and 7225 product. A 12379.1 kDa standard was included in all of the data acquisition to monitor precision of experiments, for which the SEM was found to be 0.02.

Generation of 5' trithiophosphorylated RNAs (**designated as sPPP-RNA**): 50 µL transcription reactions were assembled with 0.5 mM GTP and 2 mM GTPγS and ~100 µCi of α-^{32}PCTP. Transcription reactions were ethanol precipitated, and resuspended in 50 µL of water. Samples were then separated by 10% denaturing PAGE containing mercury layer. 5' trithiophosphorylated RNAs were gel extracted and resuspended in 30 µL of water.

Generation of internally labeled DNA substrate: An internally labeled DNA substrate 5' GGACGGATCGGATAGCTCAGCC 3' was generated by 5' end-labeling ~ 20 pmol 5' AGCTCAGCC 3' using [γ^{32}P]-ATP and T4 Polynucleotide Kinase (NEB). The labeled oligo was gel extracted and ligated with 5' GGACGGATCGGAT 3' using T4 DNA ligase. A splint 5' GGCTGAGCTATCCGATCCG 3' was used to base pair with both the 5' and 3' fragments of d3.1.

Phosphatase reactions with CIP: 20 µL of thiophosphorylation reactions were carried out using K28min ribozyme and an internally labeled DNA substrate. After 6 hr, 80 µL of water was added, followed by 10 µL of 3M NaOAc and 300 µL of 100% ethanol. Samples were precipitated and resuspended in water (**designated as d3.1 sP**). Approximately 5000-10000 cpm of d3.1 sP and 5' PPPs samples were resuspended in water and 1X Cutsmart buffer (NEB). ~10 U of the Calf Intestinal Phosphatase was added to experimental samples, whereas water was added for no treatment control. Samples were then incubated at 10 °C to prevent temperature induced thiophosphoryl hydrolysis of sPPP-RNA and d3.1 sP products. Aliquots were taken out after 30 min and 12 hr and separated on APM gels to calculate the % products that still contained a thiophosphoryl group.

Acid lability of sPPP-RNA and d3.1 sP products: Approximately 5000-10000 cpm of d3.1 sP and sPPP-RNA samples were resuspended in water. Samples were then incubated in 100 mM Acetate pH 4.0 or water at 4°C for 9 hr.

Periodate oxidation and dye labeling: 100 µL of thiophosphorylation reactions were carried out with 1.5 µM K28min ribozyme and 1 µM unlabeled d3.1 DNA substrate and ~5000 c.p.m. of trace 5' end labeled DNA substrate for 6 hr (~30% yield). Reactions were ethanol precipitated and resuspended in 20 µl 10 mM $NaIO_4$, followed by 2 hr incubation at 4 °C in the dark. 2 µL of 1 M HEPES pH 8.5 was added after 2 hr to a final concentration ~100 mM. 3 µl (15 ng) of Cy-3 hydrazide (Lumiprobe, in 100% DMSO) was added to the sample. Reactions were incubated for 2 hr at room temperature in dark, followed by ethanol precipitation. Samples were washed with 70% ethanol (vortex, and centrifuged) to remove unincorporated dye. Samples were resuspended in 95% formamide and 50 mM EDTA, and aliquots were separated by 20% denaturing PAGE.

26. Biondi E, Maxwell A, & Burke DH (2012) A small ribozyme with dual-site kinase activity. *Nucleic Acids Res*. 40(15):7528-7540.
19. Lorsch JR, Bartel D, & Szostak JW (1995) Reverse transcriptase reads through a 2'-5' linkage and a 2'-thiophosphate in a template. *Nucleic Acids Res*. 23(15):2811-2814.

Supplementary Figure Legends:

Supp Fig S1. **Reengineered Stem I alleviates misfolding at high concentration**. **A.** Self-thiophosphorylation reactions were performed using ribozyme 1.130 (dark bars) or K28(1-77)C (lightly shaded bars) at the indicated concentrations. Plotted values are the averages of two measurements for ribozyme 1.130 (error bars indicate observed range) and one measurement for ribozyme K28 (1-77)C. Reactions were incubated at 32 °C for 18 hr. **B.** Ribozyme K28 (1-77)C includes a palindromic sequence CCCUAGGG that may contribute to non-productive intermolecular interactions or internal misfolding at high concentration and that may have precluded early efforts to make 2-strand versions. Those potential interactions are weakened in ribozymes 1.130 and 1.140.

Supp Fig S2. **Multiple 2′F substitutions and copper dependence of 2′F substituted ribozymes.** **A.** Thiophosphorylation reactions were performed using ribozyme 1.140 transcripts that carried 2'F substitutions in various combinations of nucleotides. Plotted values are the averages of three iterations; error bars represent the standard error of mean. **B.** Thiophosphorylation reactions were performed in the presence or absence of Cu^{2+} for ribozyme 1.140 and in the absence of Cu^{2+} for 2′F substituted ribozymes. All reactions contained 30 mM $MgCl_2$ and were incubated at 32 °C for 6 hr.

Supp Fig S3. **Specificity of acceptor site**. **A.** Only the major site near 5′ end is reactive in the two-stranded ribozyme. Thiophosphorylation reactions were assembled for parental K28(1-77)C or with a trans-acting version of ribozyme 1.130 with radiolabel (asterisk) on either the acceptor or ribozyme strand. No signal was retained at the mercury layer when only the catalytic RNA strand carried the radiolabel, indicating that the internal site is inactivate in these two-strand constructs. **B.** Thiophosphorylation reactions using acceptor strands that carried the indicated

substitutions for nucleotide dG2 were performed at 10 °C for 0, 1, 3, 6, 9 and 18 hr, and aliquots were separated on a denaturing APM gel.

Supp Fig S4. **Sequence requirements for nucleotide that flank the modification site**. **A. – B.** Thiophosphorylation reactions were performed at 32 °C for 18 hr with DNA substrates carrying canonical **(A)** or purine analog **(B)** base substitutions at 5′GGA. **C.** Thiophosphorylation reactions for the G1I variant substrate were performed at 10 °C.

Supp. Fig S5. **Temperature dependence of quasi-diffusible acceptor substrates indicate weak interactions between the ribozyme and GGA trinucleotide**. **A.** Structures of hexaethylene-glycol (HEG) linked substrates. **B.** Thiophosphorylation reactions were performed at indicated temperatures using the acceptor analogs in (A). Product yields after 18 hours within each set were normalized to the yield at 4 °C for that acceptor. Plotted values are the average of two iterations and error bars represent the range.

Supp FigS6. **Thiophosphorylated product forms much faster than it degrades at low and moderate temperatures**. **A.** Temperature dependence of product formation. Thiophosphorylation reactions were performed at the indicated temperatures and stopped as a function of time. Plotted values are averages of three measurements, error bars represent the standard error of mean. Observed rate constants (k_{obs}) are 0.0062±0.001 min^{-1}, 0.0089±0.001min^{-1}, 0.0064±0.003 min^{-1}, 0.0073±0.009, and 0.0021±0.001min^{-1} for reactions performed at 10 °C, 20 °C, 32 °C, 40 °C and 50 °C; differences in overall activity arise from differences in f_{max}, which often reflects the fraction of the RNA that folds productively. **B.** Temperature dependence of product degradation. Thiophosphorylated product was ethanol precipitated and incubated in 25 mM Tris pH 8.0 at indicated temperatures and times and then analyzed for percent retained at the APM layer.

Supp.Fig. S7. **Nucleotidyl transfer. A.** Competition between GTP and [α-^{32}P]GTP. Incorporation of α-^{32}P label into an unlabeled d3.1 acceptor substrate was demonstrated as in Fig 4B, but with the opposite order of addition ([α-^{32}P]GTP before GTP). Controls in first two lanes show 5′-^{32}P-labeled d3.1 acceptor substrate alone (NR) or incubated with ribozyme K28min and GTPγS. For remaining lanes, non-radiolabeled d3.1 substrate was incubated with ribozyme K28min and unlabeled GTP at indicated concentration (μM) for 5 min, followed by addition of [α-^{32}P]GTP. Reactions were performed at 10 °C for 18 hr. **B-D.** Dye labeling of 5' radiolabeled acceptor DNA substrate as in fig 4C, but with dual labels. Thiophosphorylation reactions containing the indicated combinations of ribozyme, radiolabeled acceptor substrate and donor ethanol precipitated and resuspended in sodium periodate, followed by reductive amination with Cy3 hydrazide dye. For the sample in Lane 4, the product of the ribozyme-catalyzed reaction was purified from the organomercury layer of an APM gel prior to the dye labeling reaction. **B.** Gel was scanned for Cy3 dye fluorescence. The faint (X) marker on the left was drawn manually with a radioactive spot at the center of the (X) to assist in overlaying the phosphor and Cy3 scanned images. **C**. Phosphor image of same experiment from panel B, scanned for ^{32}P. The product of the ribozyme reaction appears as a doublet on the phosphor image because the presence or absence of the Cy3-dye (~543 da) affects mobility, but it is a single band in the fluorescent image because only the dye-labeled material is visible (Compare lanes 4 in panels B and C). **D.** Overlay of images from panels B and C, showing that the newly-acquired dye label co-migrates with the ^{32}P-labeled product. Red indicates Cy3 signal, green indicates ^{32}P signal. Thus, the DNA substrate acquires a ribonucleotide with an intact ribose during the ribozyme-catalyzed reaction, and both hydroxyls remain available for further reactivity.

Supp Fig S8 **Reaction kinetics under reduced GTPγS concentrations analyzed by mass spectrometry. A.** Reactions were performed at 0.1 mM ribozyme and substrate and 0.2 mM GTPγS. They were stopped at indicated times and subjected to MALDI-TOF mass spectrometry.

Samples from a given time point were measured in triplicate. Yields for 7321 m/z (full product, red trace) and 7225 m/z (lacking the thiophosphoryl moiety, blue trace) at each time point are calculated as follows: % Product = X/ (Product 6784 + Product 7321 + Product 7325), where X is the peak intensity of either Product 7321 or Product 7225, and Product 6784 is the peak intensity of unreacted DNA substrate. Precision among measurements (n = 3) was evaluated by monitoring the peak height of an unrelated internal loading control standard at 12379.1 Da, for which the standard error of mean was ± 0.2 m/z. Observed first-order rate constants for the two traces are indistinguishable (0.028 min^{-1} for 7321 m/z and 0.024 min^{-1} for 7225 m/z). **B.** Representative mass spectra from indicated time points used for quantification in panel (A). Peaks used for those quantifications are indicated with red and blue arrows. Background peaks at 7166.72, 7207.44 and 7372.18 m/z (small cyan, orange and yellow arrows) are present at t = 0 hr and likely reflect partial degradation of the ribozyme.

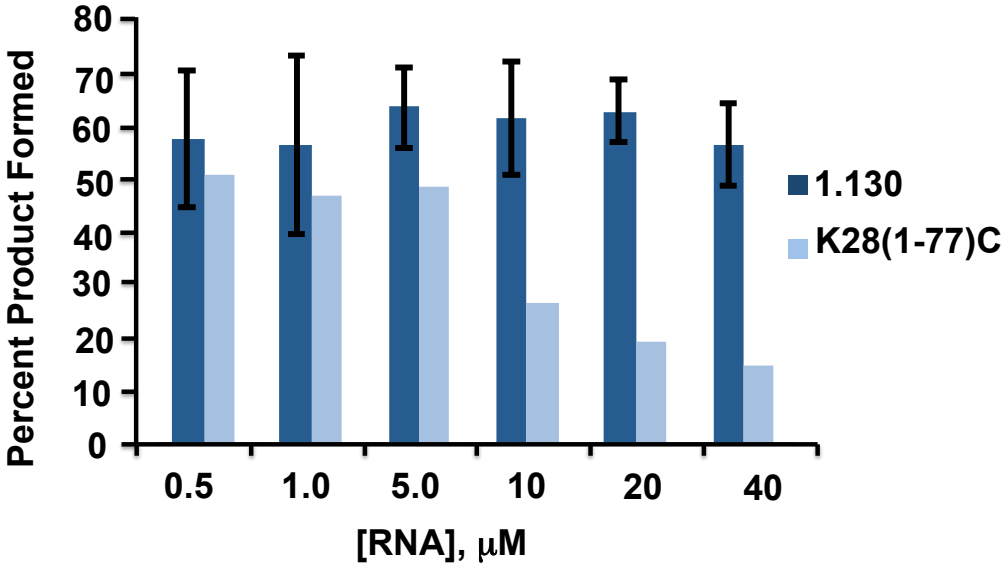
Supp. Fig S9. **Acid-induced decomposition of ribozyme products. A.** Self-thiophosphorylation reactions were performed for ribozyme K28min (nucleobase-modifying) and for ribozyme K37 (2′OH-modifying). Both ribozymes use GTPγS as the thiophosphoryl donor. Products were precipitated and resuspended either in water (sP) or in 30 mM $MgCl_2$, 200 mM KCl, 15 mM NaCl and 25 mM Acetate pH 4.5 (H^+) for 5 hr at 32 °C, then analyzed for percent retained at the APM interface. Values for each ribozyme are normalized to percent retained for the sP material. **B.** Following mild acid treatment as in panel (A), thiophosphorylation reaction products were no longer retained at the APM interface, but instead migrate as a +1 upshifted band (see inset).

Supp Fig S10 **Product formation under reducing or non-oxidizing conditions. A.** Structures of the proposed product from ribozyme-catalyzed reaction and a genuine 5′ γ-thiotriphosphorylated RNA. **B. – C.** Calf Intestinal Phosphatase treatment of product formed by ribozyme K28min with acceptor substrate d3.1 and of 5′ γ-thiotriphosphorylated RNA. Treatments were for 30 mins **(B)** or for 12 hr **(C)**. **D. – E.** Thiophosphorylation reactions were

performed as described in materials and methods with indicated concentrations of Mg^{2+} and either in the presence or absence of 0.5 mM DTT and 0.01 mM Cu^{2+}, as indicated. Samples were separated by 20% denaturing PAGE with an APM layer (**D**) or without an APM layer (**E**).

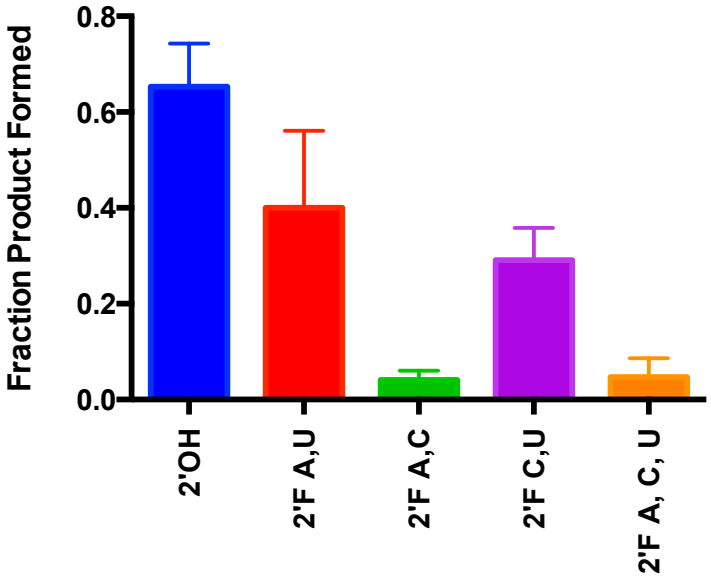
Supp Fig S11 **ITPγS does not serve as donor or compete with GTPγS. A.** Thiophosphorylation reactions were assembled by adding buffer, ITPγS, GTPγS and supplemental copper in the combinations indicated. Reactions proceeded for 18 hr at 10 °C and were then analyzed on a denaturing APM gel. **B.** Quantification of two experiments where 1 mM ITPγS or 1 mM GTPγS was used as donor with the indicated ribozyme and substrate concentrations. For relative product yields (right panel), data are normalized to product yield obtained at 200 nM ribozyme and substrate. Only trace signal was observed at the APM layer when K28min was annealed with acceptor strand d3.1 and incubated with ITPγS donor, and this signal largely disappeared upon increasing the concentration of ribozyme and substrate. **C.** Thiophosphorylation reactions were assembled as described in (A) and analyzed by non-APM denaturing PAGE. The data in these three panels indicate that ITPγS is not readily used as a donor substrate, nor does it act as a competitor to interfere with product formation by GTPγS.

Supplementary Figure S1

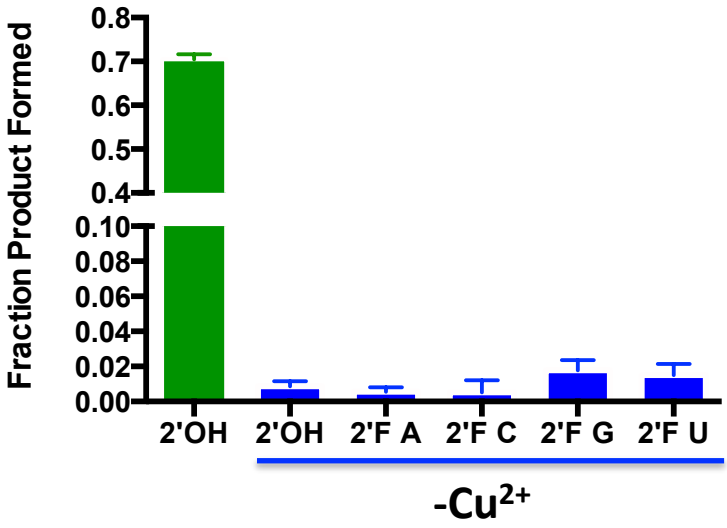


Supplementary Figure S2

A.

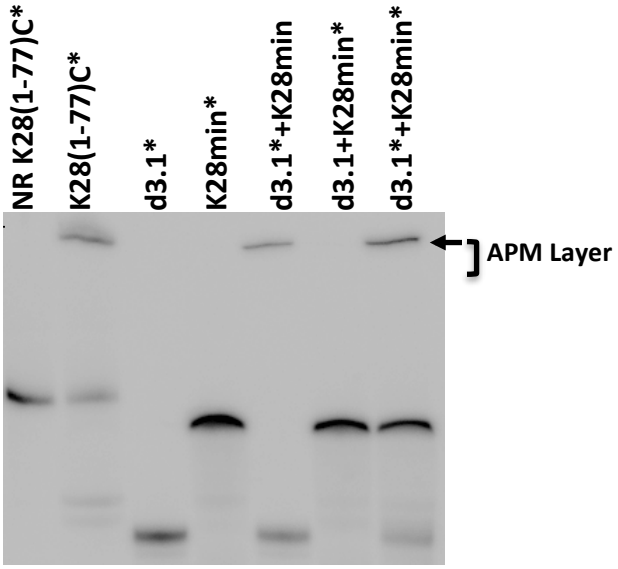


B.

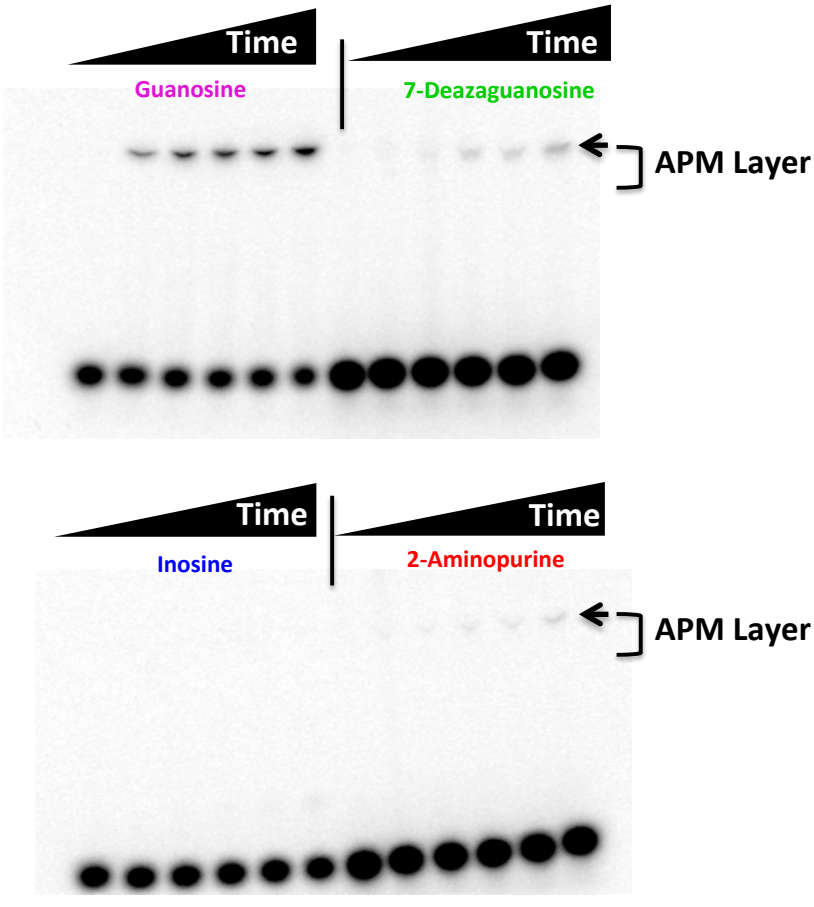


Supplementary Figure S3

A.

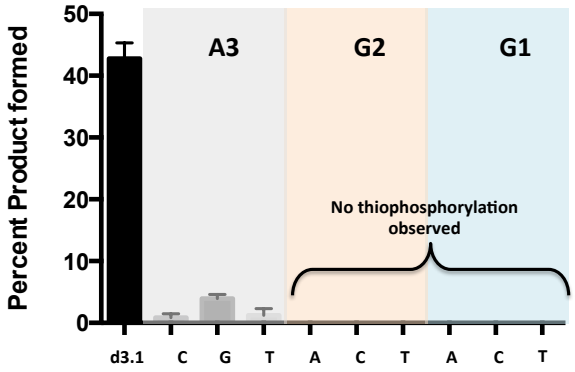


B.

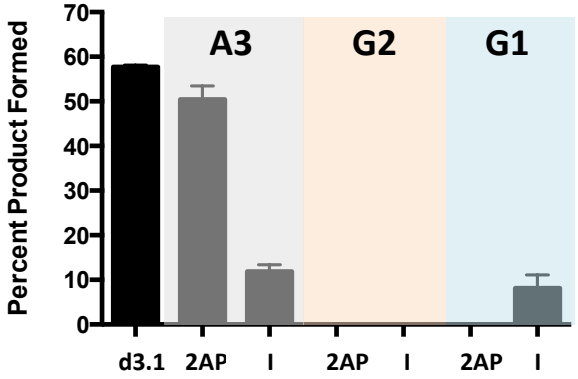


Supplementary Figure S4

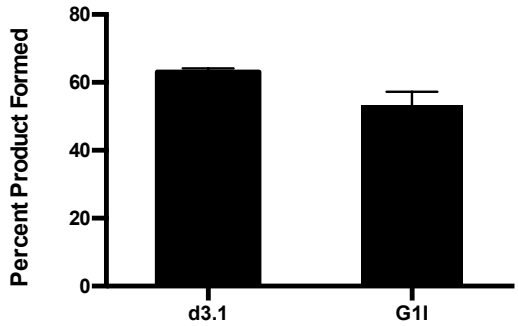
A.



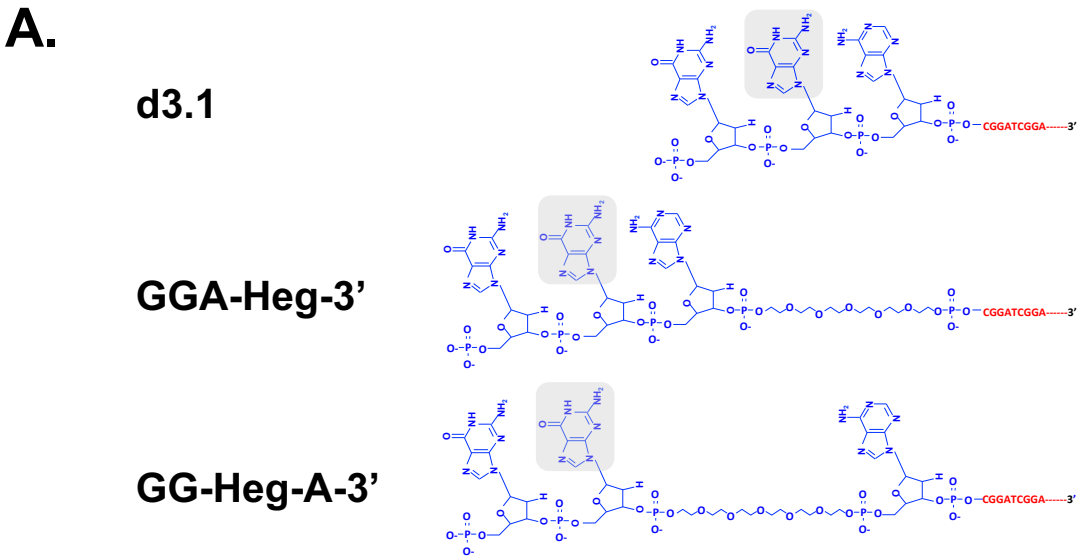
B.

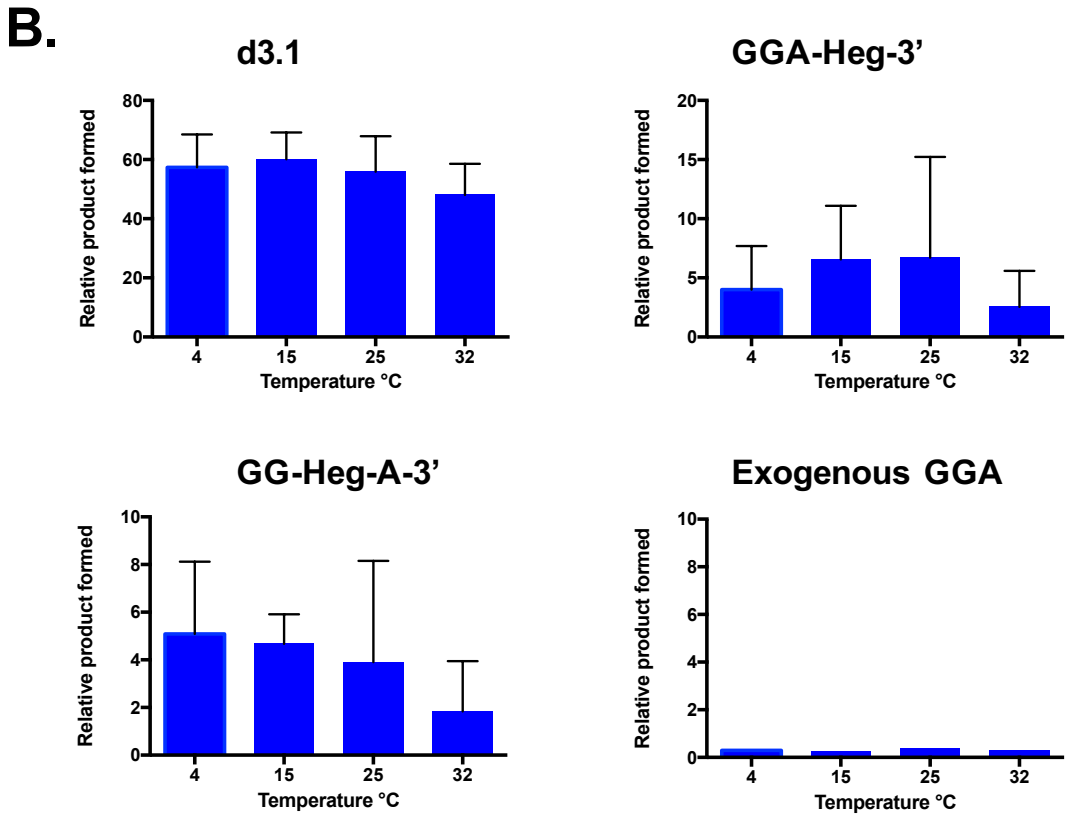


C.



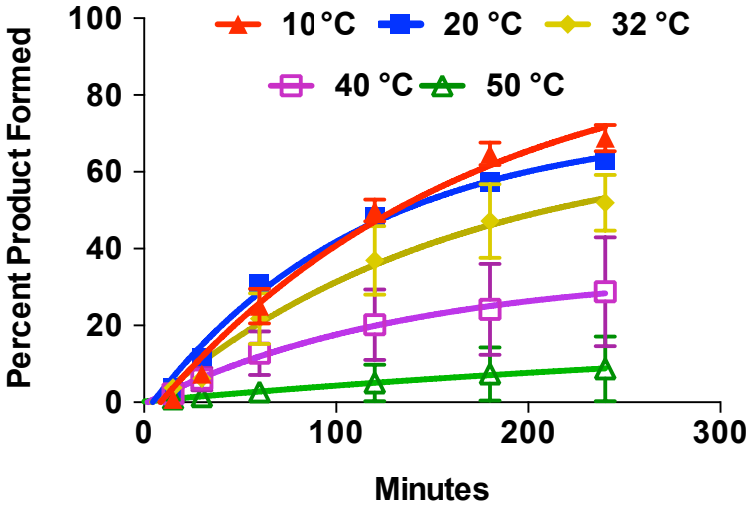
Supplementary Figure S5



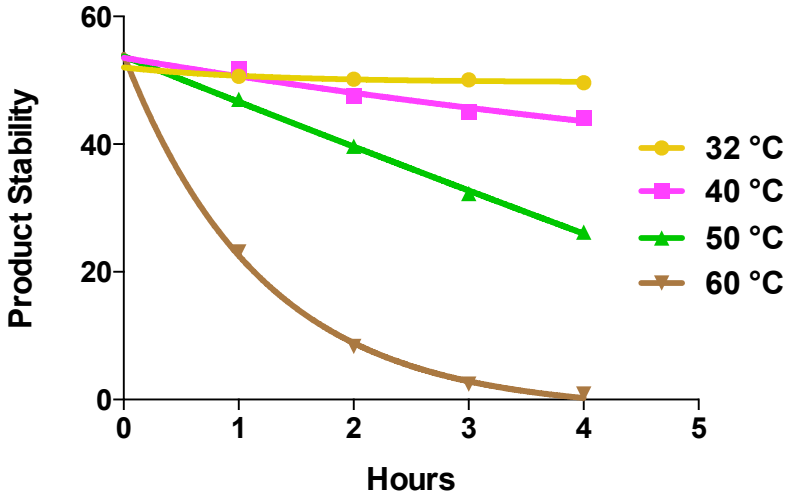


Supplementary Figure S6

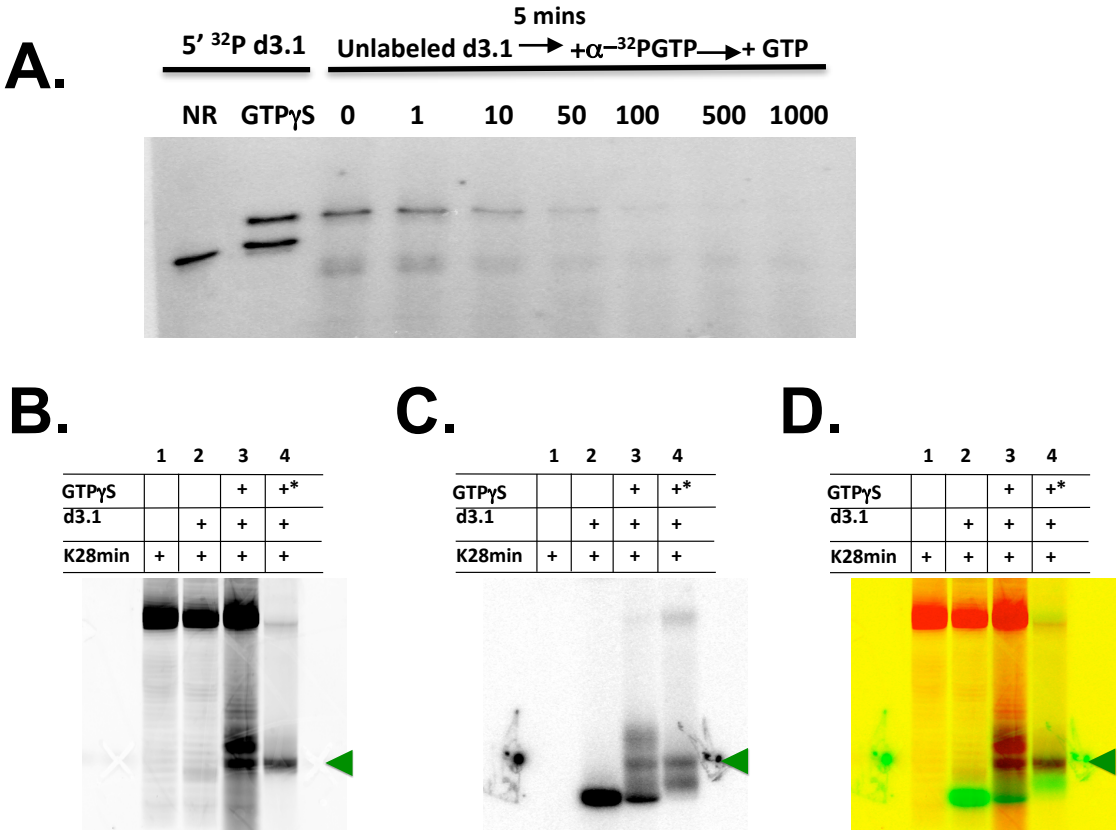
A.



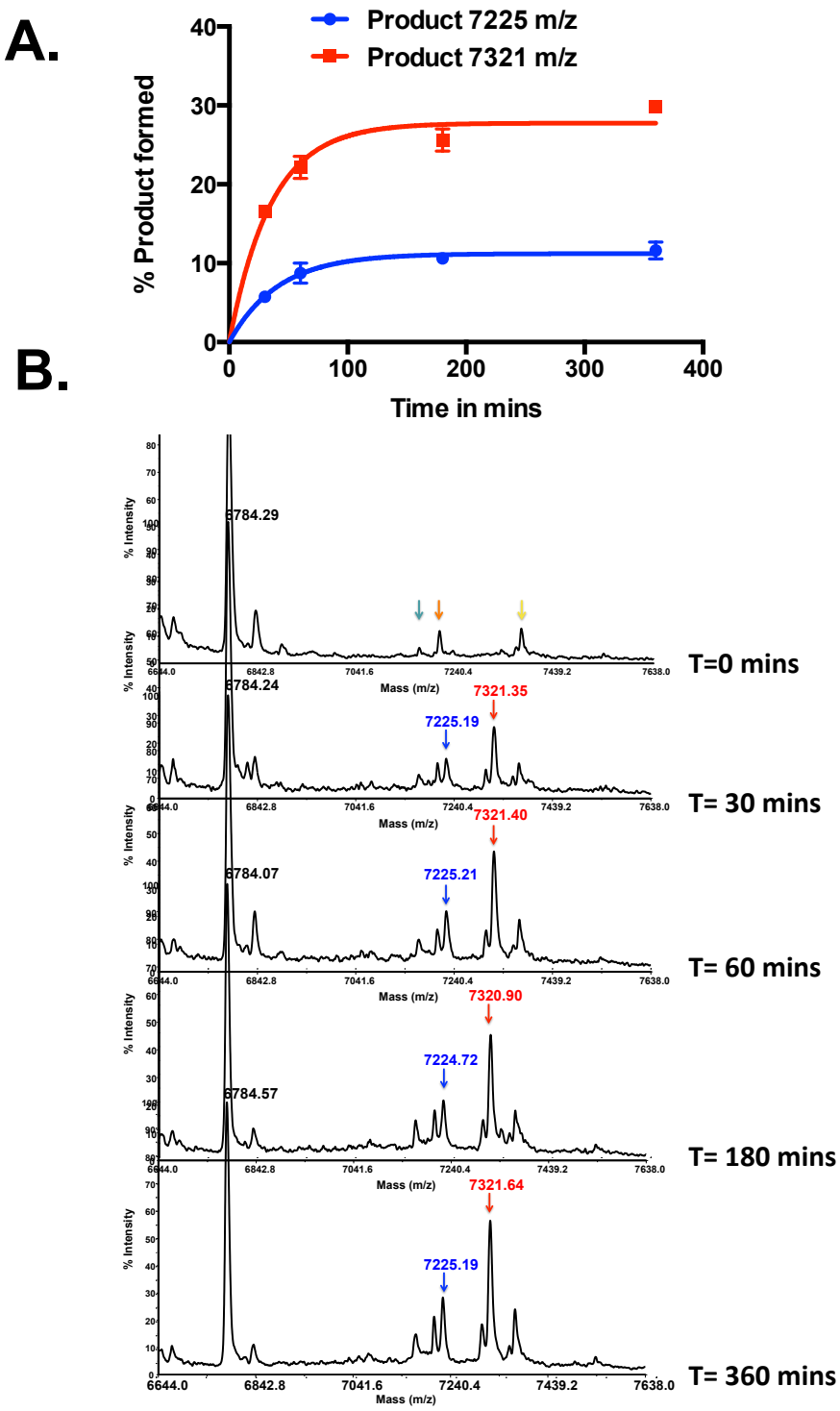
B.



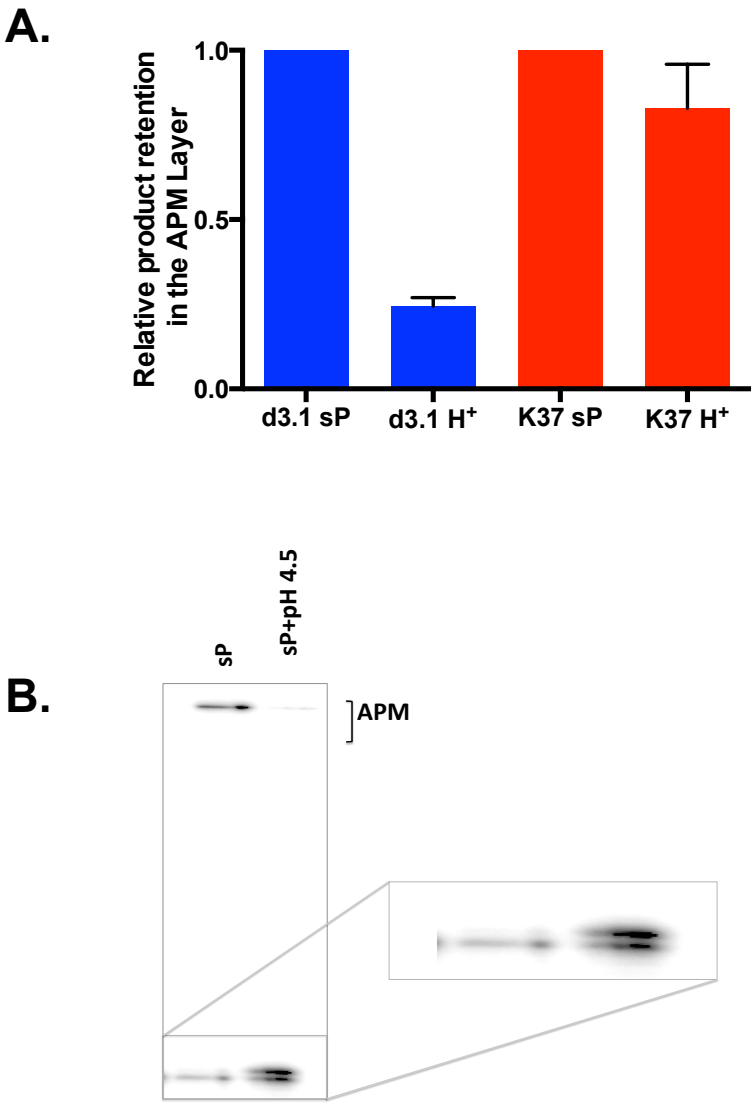
Supplementary Figure S7



Supplementary Figure S8

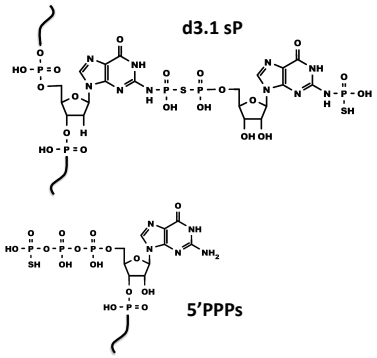


Supplementary Figure S9

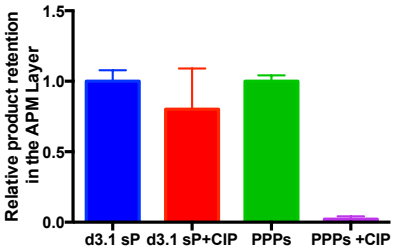


Supplementary Figure S10

A.

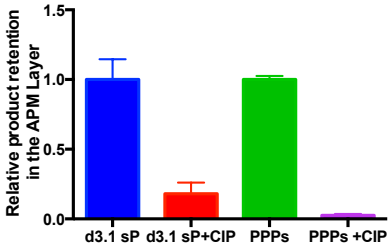


B.



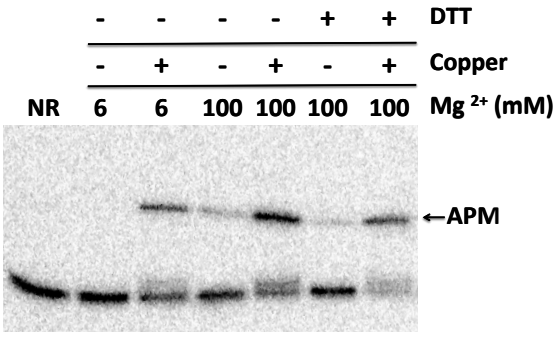
CIP Reaction= 0.5 hours

C.

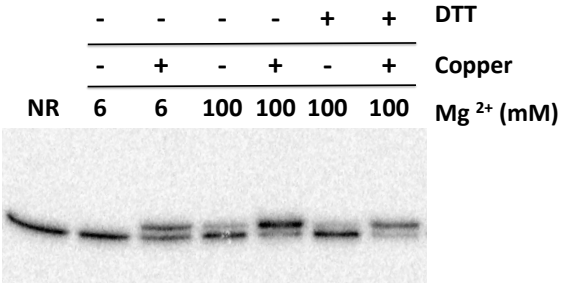


CIP Reaction=12 hours

D.



E.



Supplementary Figure S11