Supporting Information

Thiolated uridine substrates and templates improve the rate and fidelity of ribozyme-catalyzed RNA copying

Noam Prywes[ab†], Yale S. Michaels[ac†], Ayan Pal[ad], Seung Soo Oh[ad] and Jack W. Szostak[abd*]

a) Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA

E-mail: szostak@molbio.mgh.harvard.edu

b) Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge, Massachusetts 02138, USA.

c) Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DS, UK

d) Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA

† These authors contributed equally to this work.

METHODS

In-vitro transcription

The b1-233t ribozyme (Supplemental table 1) was *in vitro* transcribed using T7-MEGAshort script™ (Invitrogen, Carlsbad CA). Template DNA (Integrated DNA Technologies, Coralville IA) was diluted in water to a concentration of 10µM. 2µl of template DNA was added to an 18µl reaction mixture containing 2µl of each rNTP (75mM) as well 2µl 10x T7-MEGAshort script™ reaction buffer, 2µl T7 enzyme mix and 6µl water as per the manufacturer's instructions. Transcription reactions were carried out overnight at 37˚C.

Transcript purification

The transcripts generated by *in vitro* transcription (IVT) were purified by denaturing polyacrylamide gel electrophoresis (PAGE). An equal volume of 8M urea (Sigma-Aldrich, St. Louis MO) was added to the crude transcription product. The crude product with urea was heated at 95˚C for 2 minutes and run at 55W on a 20% PAGE gel. Desired gel bands were excised from the gel and crushed. RNA was extracted from the gel sections using 500mM NH_4OAc (Ambion®- Thermo Fisher Scientific, Waltham MA). After rocking at 4˚ overnight, samples were centrifugally purified using a 200 micron filtration column (Merck, Kenilworth NJ) at 13000 rpm for 90 seconds. The flow-through was diluted in 2 volumes of EtOH and precipitated on dry ice (-78˚C) for 30 minutes. RNA was separated at 13000 rpm for 30 minutes at 4˚C. The supernatant was then aspirated and RNA pellets were re-suspended in 10µl of purified water. RNA concentration was determined by spectrophotometry using a NanoDrop™ 2000 instrument. Purified RNA was diluted in distilled water to a working concentration of 10µM.

Ribozyme Reaction Conditions

The b1-233t ribozyme polymerase and thiolated-uracil variants, were generated by IVT and used to assay the rate of primer extension. Primer extension reactions were carried out at a final volume of 20µl in 200µl PCR tubes. Reactions were conducted in a final concentration of 300µM Tris (pH = 8), 600µM $MgCl_2$ and 2M KCL. Ribozyme RNA was heated at 95˚C for 2 minutes and cooled to room temperature gradually to allow secondary structure formation. 2 µl of ribozyme RNA was then added to the reaction for a final concentration of 1 µM. The desired template (Supplemental figure 1) as well as the desired monomer (either ATP, GTP,UTP, s^2UTP or s^2rTTP) were both added to the reaction mix (1 µM final) as well as a 5′ fluorescein labeled primer (2 µM final) (See Supplemental table 1 for sequences). Reactions proceeded at room temperature. At the designated time points, 5µl of the reaction was removed and stored at -80˚C. The class 1 ligase ribozyme and thiolated-uracil variants were generated by IVT. Ligation reactions were carried out at room temperature by combining each ligase variant (1 µM final) with the 5′ fluorescein labeled primer (0.5 µM) in ligation reaction buffer (30mM TRIS pH = 8, 60mM $MgCl_2$, 200mM KCl, 600uM EDTA) unless otherwise specified. (See Supplemental table 1 for sequences).

Quantifying Primer Extension

Ribozyme reactions were quenched in 100mM EDTA and 8M urea. Samples were subjected to denaturing PAGE at 32W. Gels were scanned with a Typhoon 9410 scanner (GE Healthcare, Little Chalfont UK) using a fluorescein filter (λ_{nm} = 526-532) at a resolution of 100 microns. Primer extension was quantified using ImageQuant 5.2 software (GE Healthcare). Each gel was

background corrected and the extension ratio was calculated as the fraction of the plus one product divided by the un-extended primer.

Determining Kinetic Parameters

Individual timepoints were plotted in Excel™ (Microsoft, Redmond WA). The natural log was determined for each timecourse and the slope of a linear fit was taken as the rate. The rates for each concentration series were then plotted in Prism 6™ (GraphPad, San Diego CA) and standard Michaelis-Menten parameters k_{cat} and K_m, along with standard errors, were calculated

Liquid Chromatography-Mass Spectrometry

IVT was performed in accordance with the above method, replacing U with an equal concentration of s^2U. Yield was compared by 20% denaturing PAGE. Gels were scanned and quantified in accordance with the previously described procedure. s^2U incorporation was verified by Liquid Chromatography-Mass Spectrometry (LCMS).

LCMS was performed with a 6520 Q-TOF mass analyzer (Agilent Technologies, Santa Clara CA) and 1200 series High Performance Liquid Chromatograph (HPLC) with an XBridge C18 column (3.5 μm, 1x100 mm, Waters, Milford MA). Mobile phase A was aqueous 200 mM hexafluoroisopropanol and 3 mM Tris/Borate/EDTA at pH = 7, and mobile phase B was methanol. HPLC was performed for 10 μL of a 200 μM solution at a linear increase of 5% to 20% B over 30 min at 0.1 mL/min at 60 ˚C. A_{260nm} was used to monitor sample elution and the eluate was passed directly to an electrospray ionization source with 325 °C drying nitrogen gas flowing at 8.0 L/min, a nebulizer pressure of 30 psig and a capillary voltage of 3500 V. Agilent MassHunter Qualitative Analysis software was used for Q-TOF derived MS data and the spectra were deconvoluted using the maximum entropy method.

Table S1: RNA Sequences and Sources

RNA name	Description	Sequence	Source
b1-233t	Ribozyme polymerase	5′-GGAAAAAGACAAAUCUGCCCUCAGAGUUGAGAACAUC UUCGGAUGCAGAGGGAUGCAGAGGAGGCAGCCUC CGGUGG CUUUAACGCCAACGUUCUCAACAAUAGCCA-3′	*in vitro* transcription
b1-233t_G	Ribozyme template (unpaired 5′ G)	5′-GGCUAUAGGACUGGAACCA-3′	IDT
b1-233t_U	Ribozyme template (unpaired 5′ U)	5′-GGCUAUAUGACUGGAACCA-3′	IDT
Class 1 ligase	Ligase ribozyme	5'- GGAACACUAUACGACUGGUACCGUAAAAGAC AAAUCUGCCCUCAGAGCUUGAGAACAUCUUCG GAUGCAGGGGAGGCAGCCCCCGGUGGCUUUAA CGCCAACGUUCUCAACAAUAGUGA- 3'	*in vitro* transcription
b1-233t_P	Fluorescien labelled b1-233t primer	5′-Fluor-dAdAdAdACCAGUC-3′	IDT

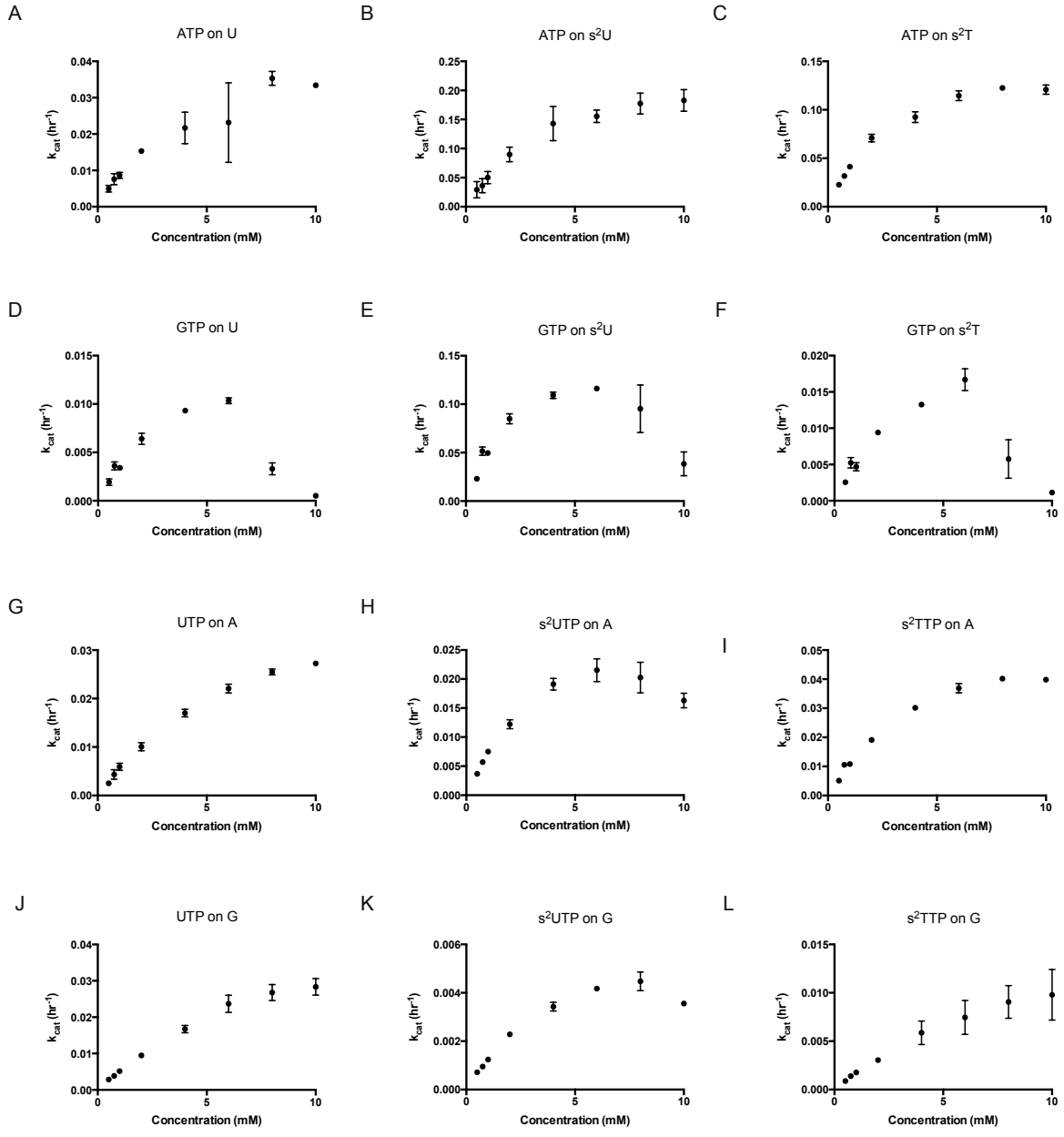

Fig. S1:

Saturation curves for NTP substrates. For each concentration of NTP a polymerization rate was calculated. These curves were fit to Michaelis-Menten parameters and k_{cat} and K_M values were extracted (Supplemental Table 2).

Supplemental Table 2:

		Template				
	Monomer	**A**	**G**	**U**	**s^2U**	**s^2T**
K_M (mM)	**A**	-	-	5.4±1.9	3.8±0.7	2.8±0.2
	G	-	-	3.1±0.4	1.9±0.2	4.2±0.7
	U	7.3±0.6	9.3±1.5	-	-	-
	s^2U	2.8±0.5	4.4±0.4	-	-	-
	s^2T	3.9±0.3	10.2±4.1	-	-	-
k_{cat} (hr^{-1})	**A**	-	-	0.05±0.009	0.26±0.02	0.16±0.004
	G	-	-	0.016±0.0009	0.16±0.008	0.028±0.002
	U	0.048±0.002	0.057±0.005	-	-	-
	s^2U	0.030±0.002	0.0071±0.0003	-	-	-
	s^2T	0.058±0.002	0.020±0.005	-	-	-

Kinetic parameters. k_{cat} and K_M values and standard errors calculated from the curves in figure S1 using Prism 6™.

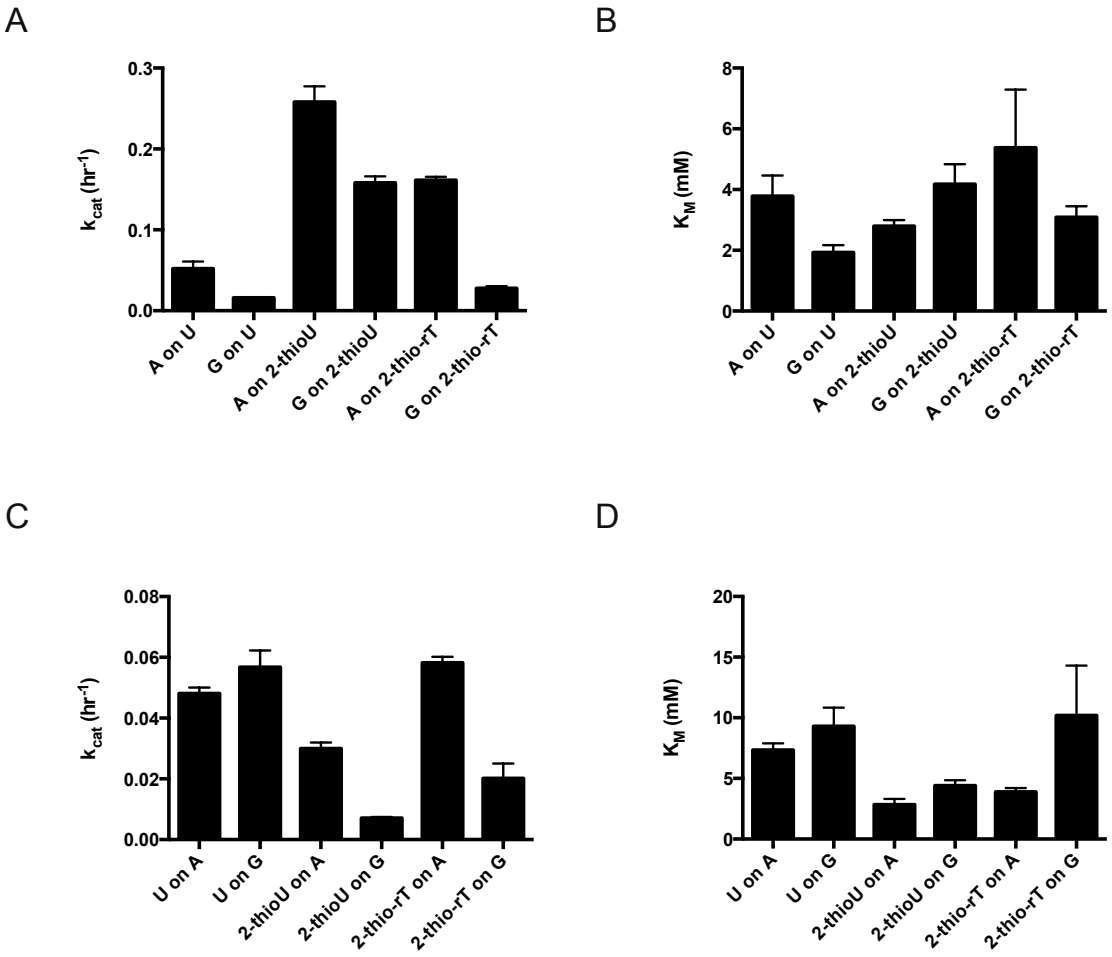

Figure S2

Kinetic parameters for ribozyme catalyzed primer extension. Values from Supplemental Table 2 plotted as bar graphs. **a** and **b** display k_{cat} and K_M values when uridine and uridine analogs are in the template. **c** and **d** display k_{cat} and K_M values for uridine and uridine analogs as NTP substrates. The convention used is N on N′, where N is the NTP and N′ is in the template.

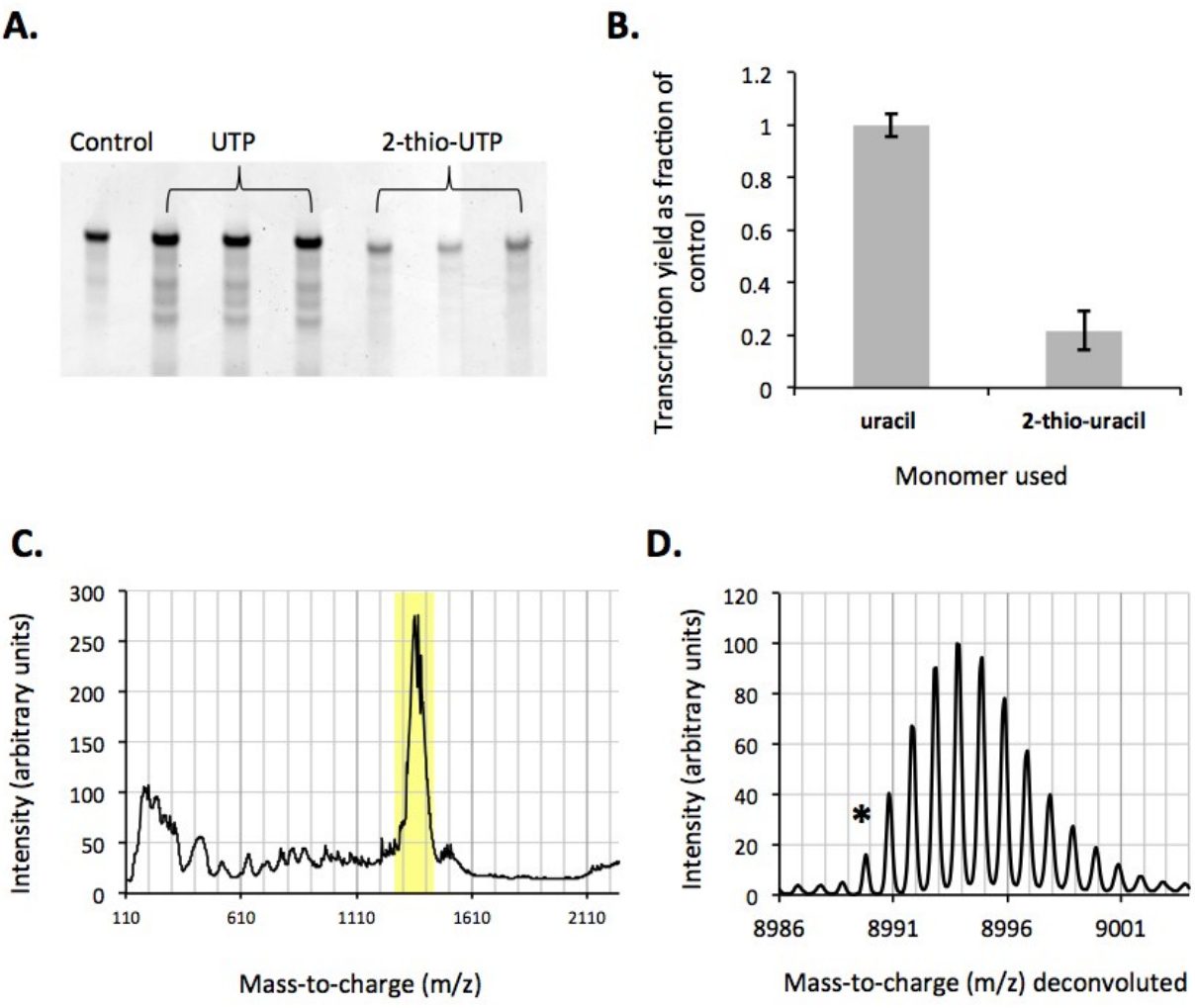

Figure S3

T7 RNA Polymerase Transcribes RNAs with 2-thio-Uracil.

(**a**) Polyacrylamide gel electrophoresis showing the transcription yield for T7 RNA polymerase with GTP, CTP, ATP and either UTP or s^2UTP (75mM each). The yields were quantified in (**b**) (error bars are standard deviation amongst three replicates). (**c**) The Mass-to-Charge profile for LCMS of a short RNA oligonucleotide (5′-GGA ACA Cs^2UA s^2UAC GAC s^2UG Gs^2U s^2Us^2Us^2U s^2UA AA-3′) transcribed by T7 with s^2UTP**.** (**d**) is a zoomed and deconvoluted plot of the region of interest (yellow). The major product (8989.8828 m/z) is denoted with an asterisk. The predicted molecular weight for the s^2U-containing transcript is 8989.8827 m/z.

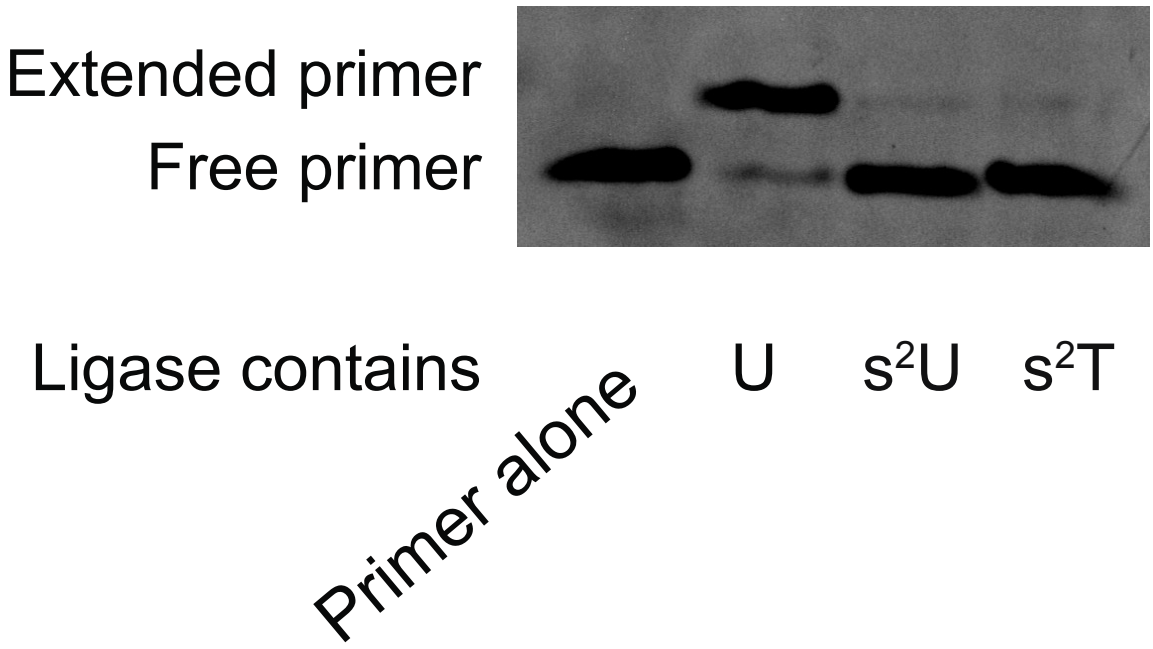

Figure S4

Modified ribozyme polymerase activity: Gel showing ribozyme polymerase activity for WT and thiolated ribozyme varients after 4 days.