Supporting Information

for

Non-canonical 3'-5' Extension of RNA with Prebiotically Plausible Ribonucleoside 2', 3'-Cyclic Phosphates

Hannes Mutschler and Philipp Holliger*

MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH , UK

*To whom correspondence should be addressed (ph1@mrc-lmb.cam.ac.uk)

Experimental Section

Materials

HPLC grade 2‘, 3‘-cyclic nucleoside mono-phosphates were purchased from Sigma-Aldrich (A>p), Santa Cruz Biotechnology, Inc. (U>p, C>p), and Biolog Life Science Institute (G>p, NAD>p). The trinucleotide ACA>p was purchased from ChemGenes. 3’-6-carboxyfluorescein labeled 5NTz 5’-acceptor (5’-GUCCUGCCC-3’), cleavage substrate (5’-AGUCCUGCCC-3’), and 5’ acceptor for pentanucleotide additions (5’-GUCCAGUCCA-3’) were purchased from IDT and further purified using denaturing gel electrophoresis. Primers for the generation of double stranded DNA templates were purchased in gel-purified form from IDT. The pentanucleotide GUCCA>p was generated by incubating 120 µM GUCCA-3P with 20 mM N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (Sigma) in 300 mM MES buffer (pH 5.5) for 30 min at 37°C similar as described previously.[1] The product was precipitated using isopropanol and washed extensively with 80% EtOH. The pellet was air-dried and resuspended in H_2O. All oligonucleotide concentrations were determined spectroscopically using the extinction coefficients at 260 nm calculated by OligoCalc.[2]

Ribozyme synthesis and purification

Ribozymes were prepared by in vitro transcription using linearized double-stranded DNA templates, which were generated by fill-in PCR using 5T7 as forward and TX_5NTz (5NTz), TX_5NTzD2 (5NTzΔ2), or TX_pentaHPz (multiple pentamer transferase) as reverse primer (Table S1). RNA constructs shown in Figure S2 were generated using TX_loopATempH2 (lane 2), TX_loopAH2 (lane 4, 6), and TX_HPz57 (lane 5). For constructs with only a single-nucleotide gap between 5’ and 3’ end, two 2’-O-methyl nucleotides (mGmU) at the 5’-ends of the fill-in ensured homogeneous 3’-ends during transcription (Table S1).[3] The PCR products were purified using QIAquick columns (Qiagen) and 400 ng used in 20 µL transcription reactions using the MEGAshortscript T7 Kit (Invitrogen). Transcription products were purified using the RNeasy mini kit (Qiagen). RNA homogeneity was confirmed by denaturing gel electrophoresis (15%) followed by SYBR Gold staining (Invitrogen). Ribozyme concentrations were determined spectroscopically using the extinction coefficients at 260 nm calculated by OligoCalc.[2]

Nucleotidyl transfer reactions

A typical sample (5 µL) of an equilibrium reaction contained 2 µM 5NTz (or 5NTzΔ2 for ACA>p addition), 1 µM acceptor RNA and varying concentrations of N>p mononucleotide, NAD>p dinucleotide, or ACA>p in reaction buffer (1 mM Tris pH 8.3 (at 22°C), 25 mM NaCl and 10 mM $MgCl_2$).[4] For 5’-transfer under eutectic conditions, all sample were incubated for 10 min at 37°C followed by 5 min in-

cubation on ice. To create eutectic conditions, samples were frozen on dry ice and incubated at -7°C in a TC120-R4 refrigerated water bath (Grant Instruments). The freezing step was omitted for aqueous reactions at ambient or supercooled conditions. Aqueous 5' adenylation at 0°C (Figure S4) was performed in presence of 100 mM $MgCl_2$ to provide similar magnesium levels as under eutectic conditions. Multiple addition experiments of the pentamer GUUCA>p were carried out in presence of 2.5 µM HPz variant (Figure S7b), 0.5 µM acceptor RNA, and 25 µM GUCCA>p substrate. All reactions were stopped at the time points indicated by re-suspension of the ice pellet in 4 volumes loading buffer (95% formamide, 10 mM EDTA pH 8, 0.25% bromphenol blue).

Pseudo first-order kinetics

Time-dependency of 5'-extension under pseudo-first order conditions was measured using a 5-fold excess of ribozyme (2.5 µM) over 3'-substrate (0.5 µM) and varying concentrations of excess A>p nucleotide (0.1 mM–12.8 mM). Pre-annealed 5NTz/substrate complex (15 min at 37°C) was mixed with nucleotide followed by immediate quick-freezing on dry ice. Reactions were started by incubating the samples at -7°C. Reactions were stopped at different time points by re-suspension of the ice pellet in 4 volumes loading buffer (95% formamide, 10 mM EDTA pH 8, 0.25% bromphenol blue). The amount of 5'-extension was determined by resolving the samples using 20% urea PAGE followed by analysis using a Typhoon Trio scanner (GE Healthcare). Kinetics of 5'-adenylation was fit with GraphPad Prism using the following equation:

$$E(t) = E_{ss}\,(1 - e^{-k_{obs}\cdot t}) + C$$

E(t) is the time-dependent fractional increase in 5'-extension, k_{obs} the observed rate constant, E_{ss} the apparent reaction endpoint, and *C* the residual densitometric offset after background subtraction. The apparent K_d for A>p was obtained by fitting E_{ss} as a function of the nucleotide concentration *X* using the equation:

$$E_{ss}(X) = \frac{E_{max}\cdot X}{K_d + X}$$

where E_{max} is the maximal level of 5'-adenylation under saturating nucleotide levels.

Reverse reaction

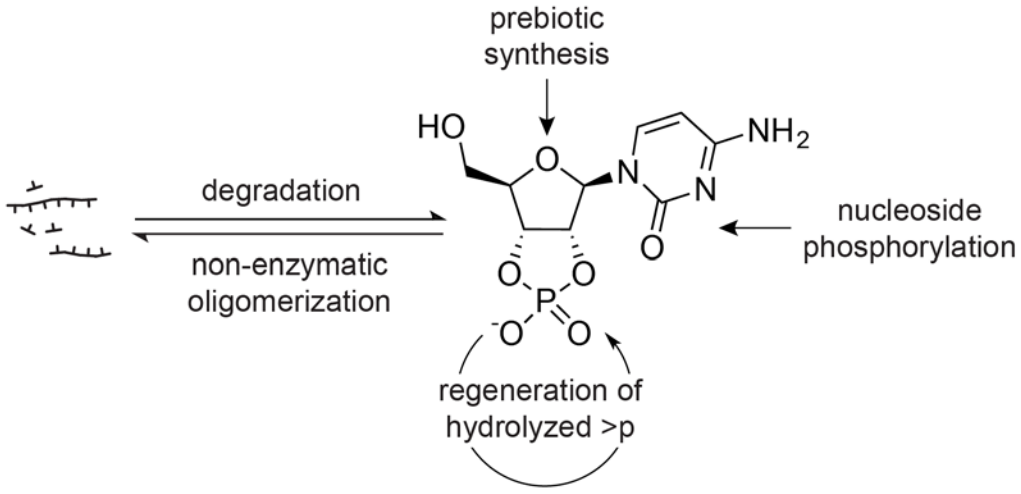
To estimate the fraction of active ribozymes 0.5 µM fully adenylated 3'-substrate was mixed with 2.5 µM pre-annealed 5NTz (15 min at 37°C) and immediately frozen on dry ice. The reaction was carried out in absence and presence of 3 mM A>p. In absence of A>p, reverse 5'-adenylation is negligible and the amount of cleaved product was therefore expected to represent the fraction of substrate strands

that are in complex with active ribozymes. Reactions were started by incubating the samples at -7°C. Reactions were stopped at different time points by re-suspension of the ice pellet in 4 volumes loading buffer (95% formamide, 10 mM EDTA pH 8, 0.25% bromphenol blue) and analyzed as described above. Kinetics of 5'-deadenylation was fit using GraphPad Prism using a mono-exponential decay model.

$$E(t) = e^{-k_{obs} \cdot t} + E_{\infty}$$

where *E(t)* is the remaining amount of 5'-adenylation at time *t* and E_{∞} is the final amount of substrate cleavage.

Chart S1. Plausible prebiotic sources for N>p formation: RNA degradation by transesterification,[5] *de novo* nucleotide synthesis,[6] nucleoside phosphorylation,[7-10] and re-cyclization of 2'/3' nucleoside monophosphates[11,12]. Non-enzymatically, N>p's can form short RNA oligomers under concentrating conditions in presence of organic catalysts.[13-15]



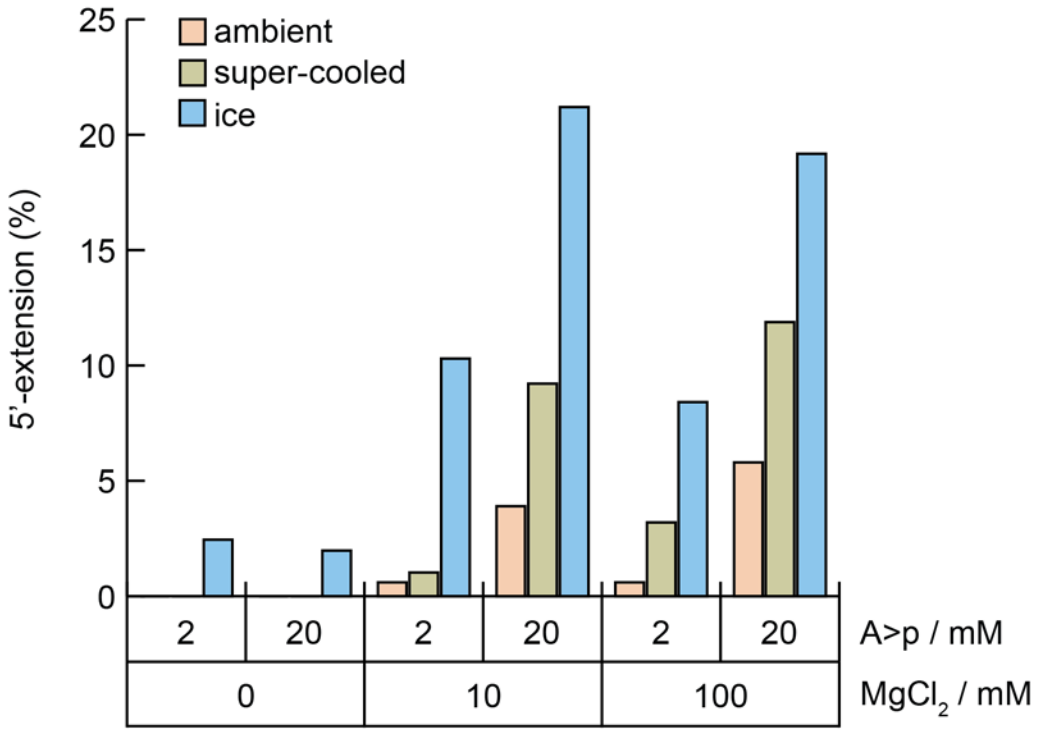


Figure S1. Ribozyme catalyzed 5'-adenylation at various reaction conditions. A) Samples containing 2 µM 5NTz and 1 µM 5'-acceptor strand and varying A>p and $MgCl_2$ concentrations were incubated for 96 h under ambient (17°C), super-cooled (-7°C but unfrozen), or ice (-7°C, frozen). The amount of 5'-extension was quantified after gel-electrophoresis.

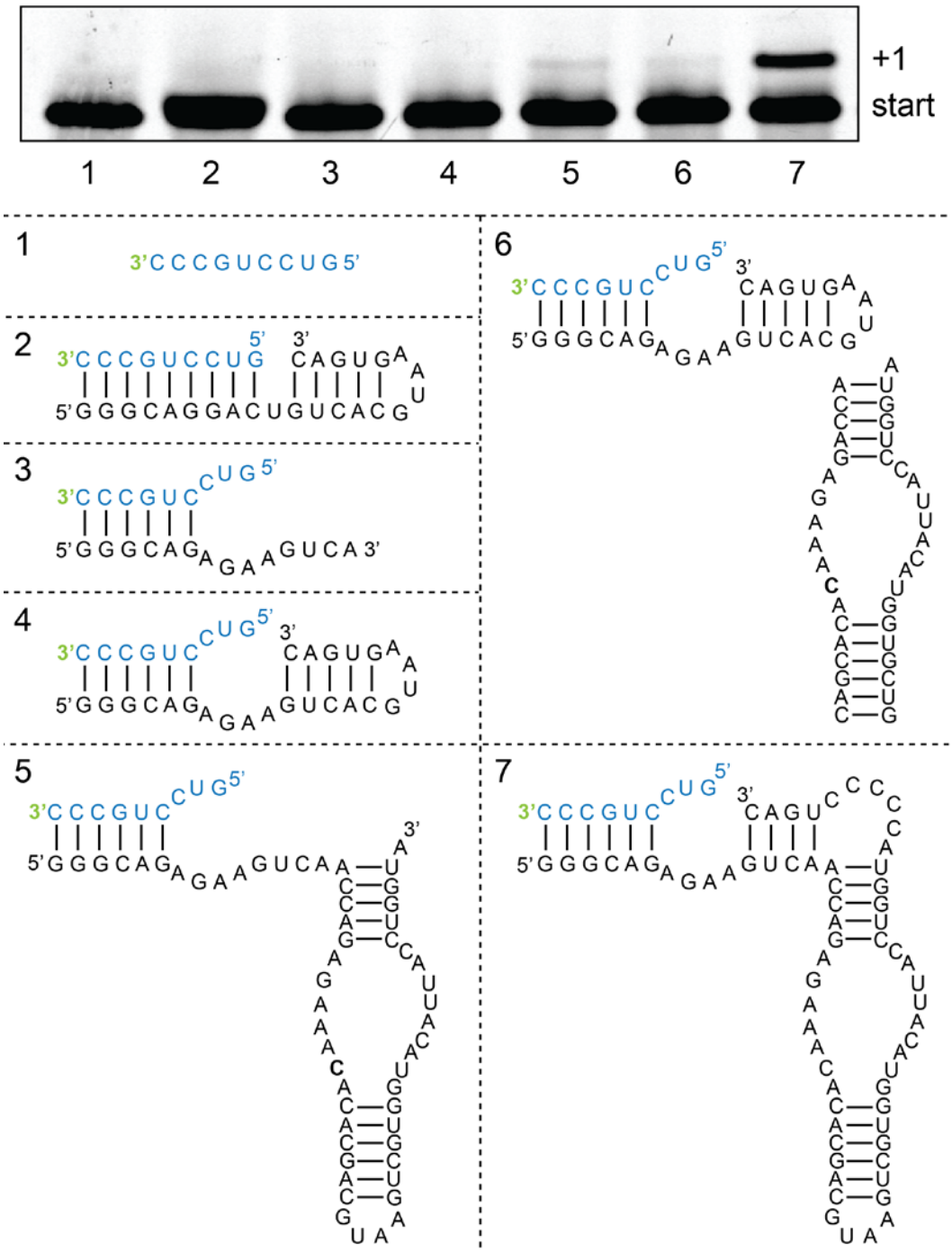


Figure S2. 5’ adenylation of fluorescently labeled acceptor strand (blue, 0.5 µM) in presence of various RNA constructs (black, 2.5 µM construct) and 3 mM A>p under eutectic conditions (-7°C, ice). Detectable amounts if 5’-extension after 72h occur in presence of (5) parental HPz, (6) a construct where the catalytic site has to assemble in trans, and (7) the full-length NTz construct used in this study.

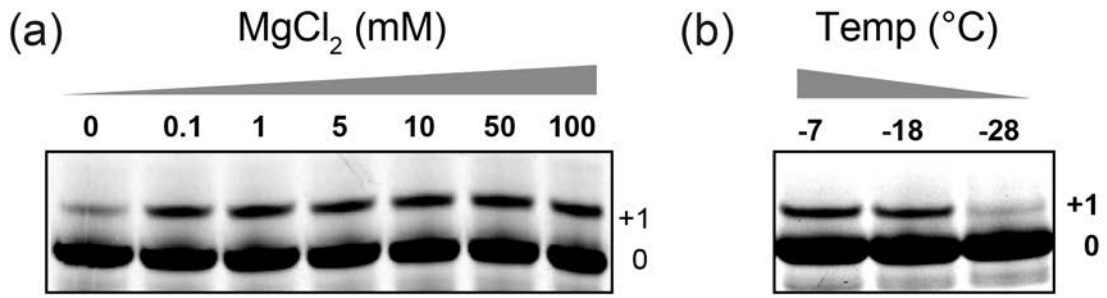


Figure S3. Eutectic nucleotidyl transfer (A>p), is (a) largely Mg^{2+} and (b), above the eutectic point, temperature independent.

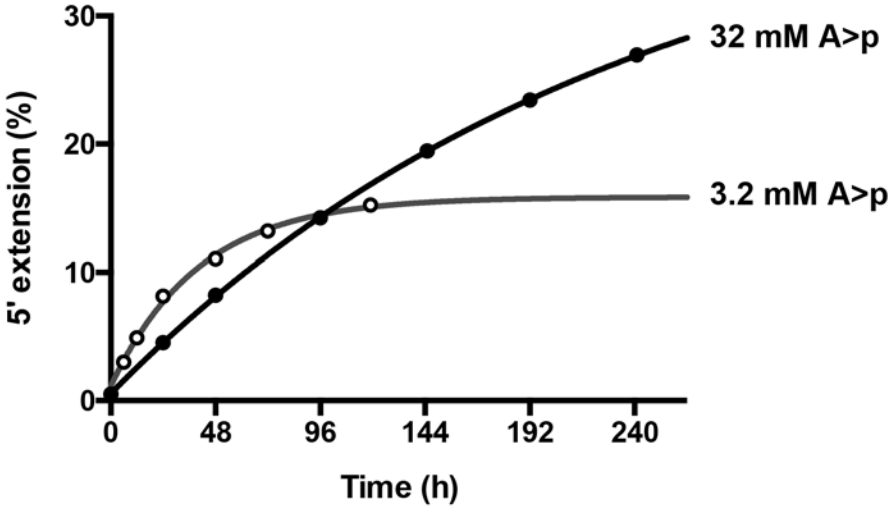


Figure S4. Substrate inhibition of 5NTz by N>p's affects the rate of 5' extension of the substrate strand. Data sets including mono-exponential fits are shown for 3.2 mM A>p and 32 mM A>p.

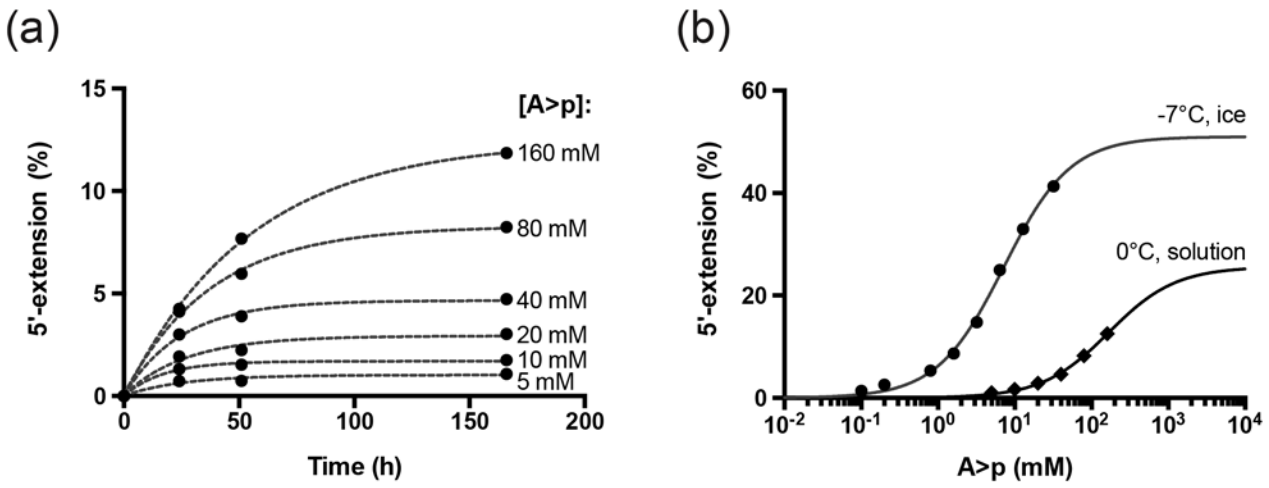


Figure S5. Extrapolation of 5' adenylation endpoints at different nucleotide concentrations in solution (0°C). (a) 5' adenylation levels were measured at three different time-points after on-ice incubation and the approximate endpoints of the reaction extrapolated using mono-exponentials. (b) Re-plot of the approximated reaction endpoints from (a) vs A>p concentration. Amplitudes were fitted using a one-site binding model resulting in a K_d^{app} of 170+/-22mM and a maximal level of 5' adenylation of 26+/-2%. The analogous plot from the in-ice experiments shown in the main text (Figure 3c) is shown as a comparison.

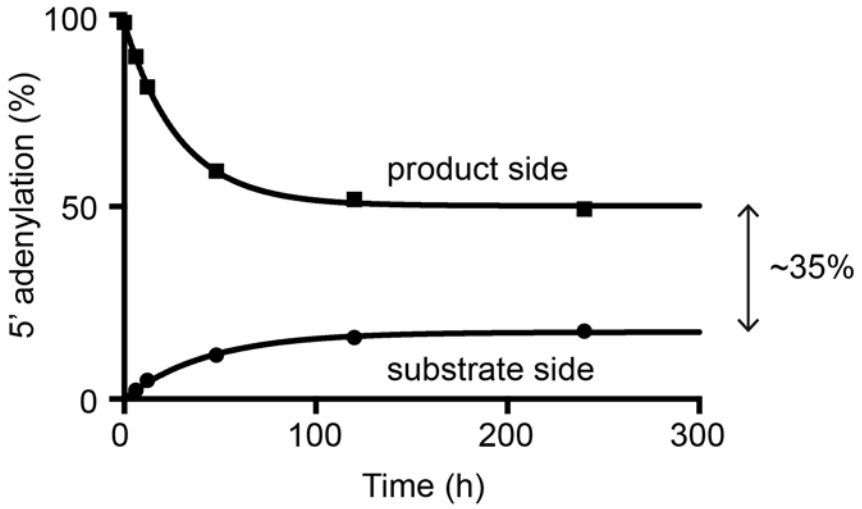


Figure S6. Reaction course of 5' adenylation (substrate side) and product deadenylation (product side) under the same reaction conditions (3 mM A>p, 2 µM 5NTz, and 1 µM 5'-acceptor strand). Both reactions were fit using a mono-exponential. The diverging reaction endpoints indicated by the fit imply that a considerable fraction of the AS is trapped by binding to inactive ribozymes and that no significant substrate exchange between the active and inactive ribozyme populations occurs during the time of incubation (10 days).

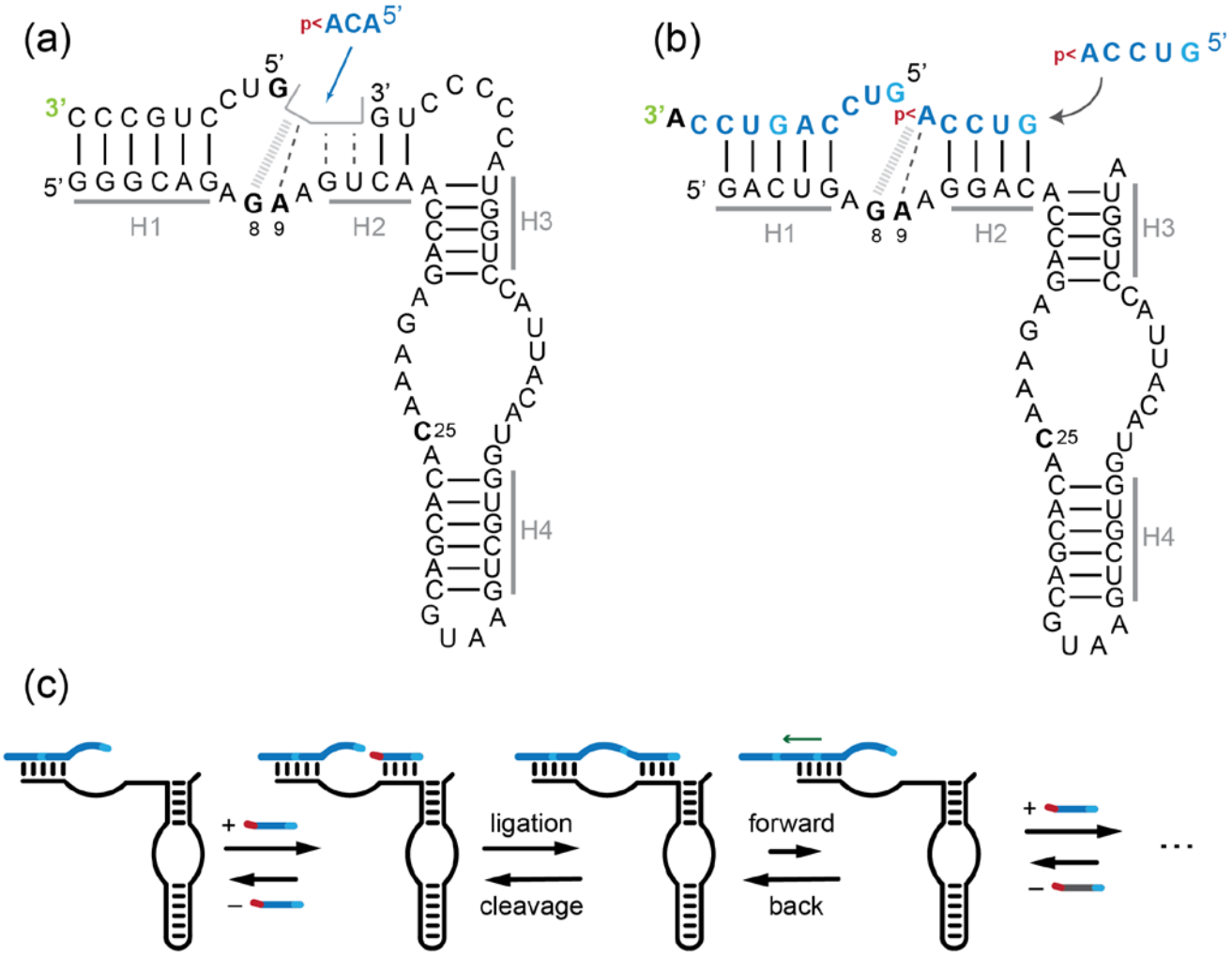


Figure S7. (a) Design of the 5NTzΔ2 ribozyme, which catalyzes 5' transfer of the ACA>p trinucleotide. (b) Design of a HPz variant that catalyzes multiple 5' additions of GUCCA>p pentanucleotide to a decamer AS (GUCCAGUCCA-6FAM). (c) Presumed mode of GUCCA>p concatenation by HPz.

Table S1. Nucleic acids. Those sequences not listed with previous descriptions of protocols are listed below. RNA residues are listed in black, and DNA in gray; IVT is in vitro transcription.

Application	Name, Source	Sequence (5' -> 3')
ribozymes	5NTz, IVT	GGGCAGAGAAGUCAACCAGAGAAACACACGACGUAAGUCGUGG UACAUUACCUGGUACCCCCUGAC
	5NTzΔ2, IVT	GGGCAGAGAAGUCAACCAGAGAAACACACGACGUAAGUCGUGG UACAUUACCUGGUACCCCCUG
	pentaHPz, IVT	GACUGAGAAGGACACCAGAGAAACACACGACGUAAGUCGUGGU ACAUUACCUGGUA
	loopATempH2, IVT (Figure S2, 2)	GGGCAGGACUGUCACGUAAGUGAC
	loopA, IDT (Figure S2, 3)	GGGCAGAGAAGUCAC
	loopAH2, IVT (Figure S2, 4)	GGGCAGAGAAGUCACGUAAGUGAC
	HPz57, IVT (Figure S2, 5)	GGGCAGAGAAGUCAACCAGAGAAACACACGACGUAAGUCGUGG UACAUUACCUGGUA
	loopB1, IDT (Figure S2, 6)	ACCAGAGAAACACACGAC
	loopB2, IDT (Figure S2, 6)	GUCGUGGUACAUUACCUGGUA
Fill-in primer	5T7	GATCGATCTCGCCCGCGAAATTAATACGACTCACTATA
	TX_5NTz, IDT	mGmUCAGGGGGTACCAGGTAATGTACCACGACTTAC-GTCGTGTGTTTCTCTGGTTGACTTCTCTGCCCTATAGTGAGTCGTA TTAATTTC
	TX_5NTzD2, IDT	CAGGGGGTACCAGGTAATGTACCACGACTTACGTCGTGTTTCTCT GGTTGACTTCTCTGCCCTATAGTGAGTCGTATTAATTTC
	TX_loopATempH2	mGmUCACTTACGTGACAGTCCTGCCCTATAGTGAGTCG-TATTAATTTC
	TX_loopAH2	mGmUCACTTACGTGACTTCTCTGCCCTATAGTGAGTCG-TATTAATTTC
	TX_HPz57	TACCAGGTAATGTACCACGACTTACGTCGTGTGTTTCTCTGGTTG ACTTCTCTGCCCTATAGTGAGTCGTATTAATTTC

Supplemental References

(1) Hertel, K. J.; Herschlag, D.; Uhlenbeck, O. C. *Biochemistry* **1994**, *33*, 3374.
(2) Kibbe, W. A. *Nucleic Acids Res* **2007**, *35*, W43.
(3) Kao, C.; Zheng, M.; Rudisser, S. *Rna* **1999**, *5*, 1268.
(4) Vlassov, A. V.; Johnston, B. H.; Landweber, L. F.; Kazakov, S. A. *Nucleic Acids Res* **2004**, *32*, 2966.
(5) Li, Y.; Breaker, R. R. *J Am Chem Soc* **1999**, *121*, 5364.
(6) Powner, M. W.; Gerland, B.; Sutherland, J. D. *Nature* **2009**, *459*, 239.
(7) Lohrmann, R.; Orgel, L. E. *Science* **1968**, *161*, 64.
(8) Lohrmann, R.; Orgel, L. E. *Science* **1971**, *171*, 490.
(9) Schoffstall, A. M. *Orig Life* **1976**, *7*, 399.
(10) Yamagata, Y.; Inoue, H.; Inomata, K. *Orig Life Evol Biosph* **1995**, *25*, 47.
(11) Westheimer, F. H. *Acc Chem Res* **1968**, *1*, 70.
(12) Crowe, M. A.; Sutherland, J. D. *Chembiochem* **2006**, *7*, 951.
(13) Renz, M.; Lohrmann, R.; Orgel, L. E. *Biochim Biophys Acta* **1971**, *240*, 463.
(14) Verlander, M. S.; Lohrmann, R.; Orgel, L. E. *J Mol Evol* **1973**, *2*, 303.
(15) Verlander, M. S.; Orgel, L. E. *J Mol Evol* **1974**, *3*, 115.