Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

DNA-Catalyzed Lysine Side Chain Modification**

*Benjamin M. Brandsen, Tania E. Velez, Amit Sachdeva, Nora A. Ibrahim, and Scott K. Silverman**

anie_201404622_sm_miscellaneous_information.pdf

Table of Contents

Procedures for in vitro selection and cloning

The key selection step of each round is shown in Fig. 2b, and full nucleotide details are depicted in Fig. S1. The DNA anchor oligonucleotide sequence was 5′-GGATAATACGACTCACTAT-3′. The 5′-ImpDNA sequence was 5′-TATGTTCTAGCGCTTCG-3′, prepared as described below. The random deoxyribozyme pool was 5′-CGAAGCGCTAGAACAT-N_{40}-ATAGTGAGTCGTATTAAGCTGATCCTGATGG-3′. PCR primers were 5′-CGAAGCGCTAGAACAT-3′ (forward primer) and 5′-$(AAC)_4$XCCATCAGGATCAGCT-3′, where X is the HEG spacer to stop Taq polymerase (reverse primer). In each round, the ligation step to attach the deoxyribozyme pool at its 3′-end with the 5′-end of the amine-containing substrate was performed using a DNA splint and T4 DNA ligase. The splint sequence was 5′-ATAGTGAGTCGTATTATCC<u>T</u>CCATCAGGATCA-GCTTAATACGACTCACTAT-3′, where the underlined <u>T</u> is included to account for the untemplated A nucleotide that is added at the 3′-end of each PCR product by Taq polymerase. This <u>T</u> nucleotide was omitted from the splint used for ligation of the initially random N_{40} pool, which was prepared by solid-phase synthesis without the untemplated A.

Procedure for preparation of 5′-ImpDNA. A 60 µL sample containing 1.0 nmol of 5′-phosphorylated DNA, 100 mM 1-ethyl 3-(3-dimethylaminopropyl) carbodiimide (EDC), and 100 mM imidazole (pH 6.0 with HCl) was incubated at room temperature for 2 h. A Micro Bio-Spin P-6 desalting column (Bio-Rad) was prepared by centrifuging at 1000 *g* for 1 min and rinsing 4× by adding 500 µL of water followed by centrifuging at 1000 *g* for 1 min. The 60 µL sample was applied to the column and eluted by centrifuging at 1000 *g* for 4 min. The resulting sample was quantified by UV absorbance (A_{260}).

Procedure for ligation step in round 1. A 25 µL sample containing 600 pmol of DNA pool, 750 pmol of DNA splint, and 900 pmol of DNA-C_3-NH_2 or DNA-HEG-CKA substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 3 µL of 10× T4 DNA ligase buffer (Fermentas) and 2 µL of 5 U/µL T4 DNA ligase (Fermentas). 10× T4 DNA ligase buffer that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) was used with disulfide-linked oligonucleotide-peptide conjugates. The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

Procedure for ligation step in subsequent rounds. A 17 µL sample containing the PCR-amplified DNA pool (~5–10 pmol), 30 pmol of DNA splint, and 50 pmol of DNA-C_3-NH_2 or DNA-HEG-CKA substrate was annealed in 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. To this solution was added 2 µL of 10× T4 DNA ligase buffer (Fermentas) and 1 µL of 1 U/µL T4 DNA ligase (Fermentas). 10× T4 DNA ligase buffer that lacks DTT (400 mM Tris, pH 7.8, 100 mM $MgCl_2$, and 5 mM ATP) was used with disulfide-linked oligonucleotide-peptide conjugates. The sample was incubated at 37 °C for 12 h and purified by 8% PAGE.

Procedure for selection step in round 1. Each selection experiment was initiated with 200 pmol of the ligated pool. A 20 µL sample containing 200 pmol of ligated pool and 300 pmol of 5′-ImpDNA was annealed in (conditions *A*) 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA or (conditions *B*) 5 mM CHES, pH 9.0, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 40 µL total volume containing (conditions *A*) 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, and 150 mM NaCl or (conditions *B*) 50 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl. The Mn^{2+} was added from a 10× stock solution containing 200 mM $MnCl_2$. The Zn^{2+} was added from a 10× stock solution containing 10 mM $ZnCl_2$, 20 mM HNO_3, and 200 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM $ZnCl_2$ in 200 mM HNO_3. The metal ion stocks were added last to the final sample. The sample was incubated at 37 °C for 14 h.

Procedure for selection step in subsequent rounds. A 10 µL sample containing the ligated pool and 30 pmol of 5′-ImpDNA was annealed in (conditions *A*) 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA or (conditions *B*) 5 mM CHES, pH 9.0, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The selection reaction was initiated by bringing the sample to 20 µL total volume containing (conditions *A*) 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, and 150 mM NaCl or (conditions *B*) 50 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 14 h.

Procedure for PCR. In each selection round, two PCR reactions were performed, 10-cycle PCR followed by 30-cycle PCR. First, a 100 µL sample was prepared containing the PAGE-purified selection product, 200 pmol of forward primer, 50 pmol of reverse primer, 20 nmol of each dNTP, and 10 µL of 10× Taq polymerase buffer [1× = 20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100]. This sample was cycled 10 times according to the following PCR program: 94 °C for 2 min, 10× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. Taq polymerase was removed by phenol/chloroform extraction. Second, a 50 µL sample was prepared containing 1 µL of the 10-cycle PCR product, 100 pmol of forward primer, 25 pmol of reverse primer, 10 nmol of each dNTP, 20 µCi of α-^{32}P-dCTP (800 Ci/mmol), and 5 µL of 10× Taq polymerase buffer. This sample was cycled 30 times according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. Samples were separated by 8% PAGE.

Cloning and screening of individual deoxyribozymes. The PCR primers used for cloning were 5′-CGAAGCGCTAGAACAT-3′ (forward primer; same as in selection) and 5′-TAATTAATTAATTACCCATCAGG-ATCAGCT-3′ (reverse primer). The 10-cycle PCR product from the appropriate selection round was diluted 10^3-fold. A 50 µL sample was prepared containing 1 µL of the diluted 10-cycle PCR product from the appropriate selection round, 100 pmol of forward cloning primer, 25 pmol of reverse cloning primer, 10 nmol of each dNTP, and 5 µL of 10× Taq polymerase buffer. This sample was cycled 30 times according to the following PCR program: 94 °C for 2 min, 30× (94 °C for 30 s, 47 °C for 30 s, 72 °C for 30 s), 72 °C for 5 min. The sample was separated by 1% agarose gel and extracted using a GeneJET Gel Extraction Kit (Fermentas). The extracted product was quantified by absorbance (A_{260}) and diluted to 4–8 ng/µL. A 1 µL portion of the diluted PCR product was inserted into the pCR2.1-TOPO vector using a TOPO TA cloning kit (Life Technologies). Individual *E. coli* colonies harboring plasmids with inserts were identified by blue-white screening and grown in LB/amp media. Miniprep DNA was prepared using a GeneJET Plasmid Miniprep Kit (Fermentas) and screened for properly sized inserts by EcoRI digestion and agarose gel analysis. Before sequencing, assays of individual deoxyribozyme clones were performed with PAGE-purified DNA strands prepared by PCR from the miniprep DNA, using the single-turnover assay procedure described in the main text.

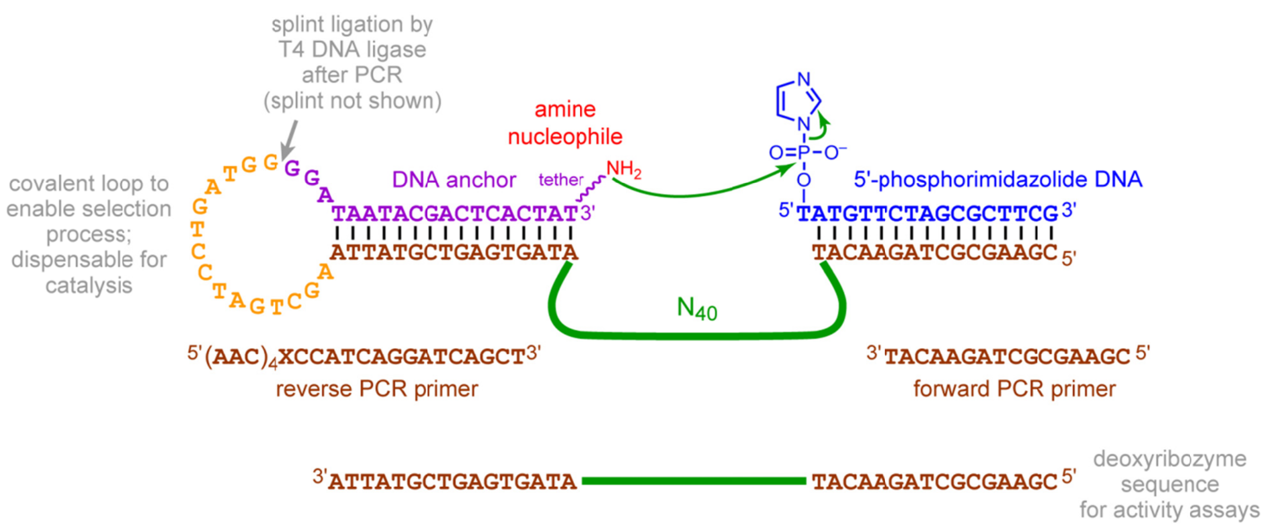

Figure S1. Nucleotide details of the in vitro selection experiments. The tether and amine nucleophile structures are drawn in Fig. 2a.

In vitro selection progressions

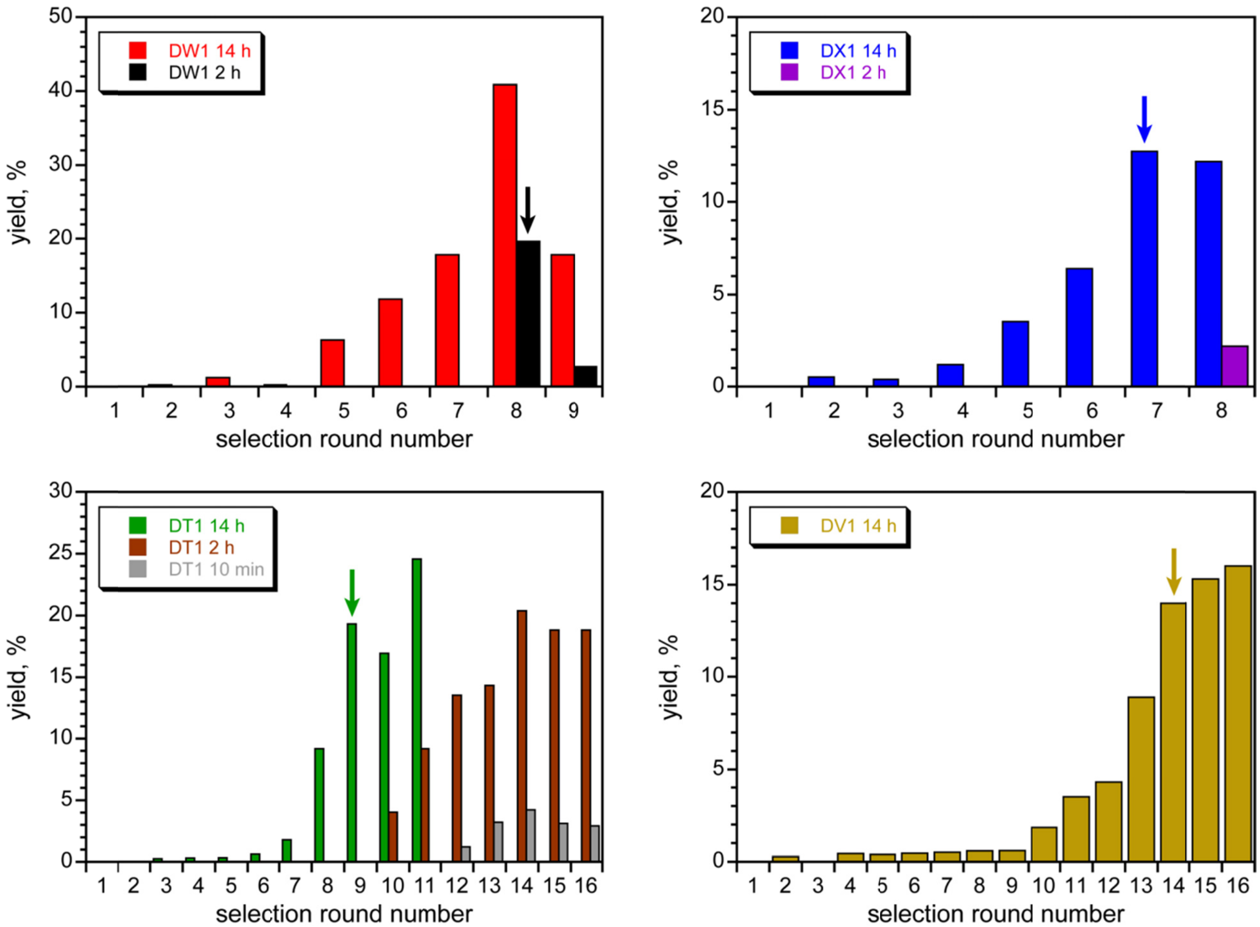

Figure S2. Progressions of the in vitro selection experiments. Arrows mark the cloned rounds. The DW1 and DX1 selections were for reaction of the DNA-C_3-NH_2 substrate under conditions *A* and *B*, respectively, using an N_{40} random region. The DT1 and DV1 selections were for reaction of the DNA-HEG-CKA substrate under conditions *A* and *B*, respectively, using an N_{40} random region. Conditions *A* were 50 mM HEPES, pH 7.5 with 40 mM Mg^{2+}, 20 mM Mn^{2+}, 1 mM Zn^{2+}, and 150 mM Na^{+} at 37 °C for 14 h or 2 h as indicated. Conditions *B* were 50 mM CHES, pH 9.0 with 40 mM Mg^{2+} and 150 mM Na^{+} at 37 °C for 14 h, 2 h, or 10 min as indicated. All metal ions were provided as chloride salts. Arrows mark the cloned rounds.

Sequences of individual deoxyribozymes

	1–10	11–20	21–30	31–40	
8DW103	ACGGGGAACG	GTATTAGGGA	GTACTTGGAG	AGCGCGTAGC	40 (1)
8DW112	CAA...G.AC	.AGG.T.T..	.A.G.AA.G.	C.T.A.....	40 (2)
8DW113	C.A.ACCG.C	AAT...AAAG	.GG...CTG.	T..T..GG.G	40 (1)
8DW115	C...C...AT	C..C..C..G	TATAAC.AGA	T...GAC.CG	40 (1)
8DW118	....	TATG..TA.G	AAGAAGCAGA	C.........	40 (2)
8DW120	.G...A.CG.	AG.GA.AC..	A..G.AC.C.	..T.......	40 (1)
8DW134	....	AA.GGT.TAC	T..AGAT.TA	C.T.......	40 (1)

	1–10	11–20	21–30	31–40	
7DX107	AGCAGCAATA	TGTCGCGGTA	TAGAGAGGAT	TGACTTGCGT	40 (1)
7DX108	..GG..GGA.	ACAGATA.A.	CGAGAGA.TG	C.T.CCCGTA	40 (1)
7DX110	..TG.TGCGC	GTCAC.CAAC	GCC.A.A.CG	CTGTAACATA	40 (1)
7DX111	.TAGTG.G.C	GTGT.T.ACG	C.A.A.TTCG	GATTCGAAAG	40 (1)
7DX112	..A.A.TGGG	ACA.CGACGC	CGA...TCCA	CT.GAACATA	40 (1)
7DX114	..GGCAGTG.	A.GACAA..G	.CA....CC.	G.TGGCC...	40 (1)
7DX124	.CGG.A.GCG	ACGAAG.C.T	C.AGA...GA	CT.GGACATA	40 (1)

	1–10	11–20	21–30	31–40	
9DT105	GTGCAGGATA	GCACTGGGAT	CCGGTACGCC	TCGGAGGGTT	40 (6)
9DT114	TA..T.TC..	.G..CTATCG	TACTG.GAAT	CATTCCAAAG	40 (3)

	1–10	11–20	21–30	31–40	
14DV103	AGCGAGCGGC	ATGATGTCGA	ACATCGATCA	TTATGTGTGA	40 (6)
14DV104	........A.	G.AGC..GTT	.A.CAC.A.C	C..C......	40 (2)
14DV108	....	GGAG.TGGAT	GAT.TACCA.	.A.C.C....	40 (3)
14DV111	........A.	.CCGGCCA.C	.ATCG..AG.	.GG.......	40 (1)
14DV117	........A.	GG..GA..AT	TTTA.ATGAT	AA.C......	40 (1)
14DV131	.........A	.C.TGCCA.C	CGTCGT.GGG	ACG.TC....	40 (1)

Figure S3. Sequences of the deoxyribozymes described in this study. Only the initially random (N_{40}) sequences are shown. All deoxyribozymes were used as 5′-CGAAGCGCTAGAACAT-N_{40}-ATAG<u>T</u>GAGTCGTATTA-3′. For 9DT114, the underlined T in the 3′-binding arm was deleted; its presence led to no catalytic activity (data not shown). In each alignment, a dot denotes conservation, i.e., the same nucleotide as in the uppermost sequence; a dash denotes a gap. On the far right is shown the sequence length and (in parentheses) the number of times that the particular sequence was found during cloning.

Assays of the 8DW1 deoxyribozymes

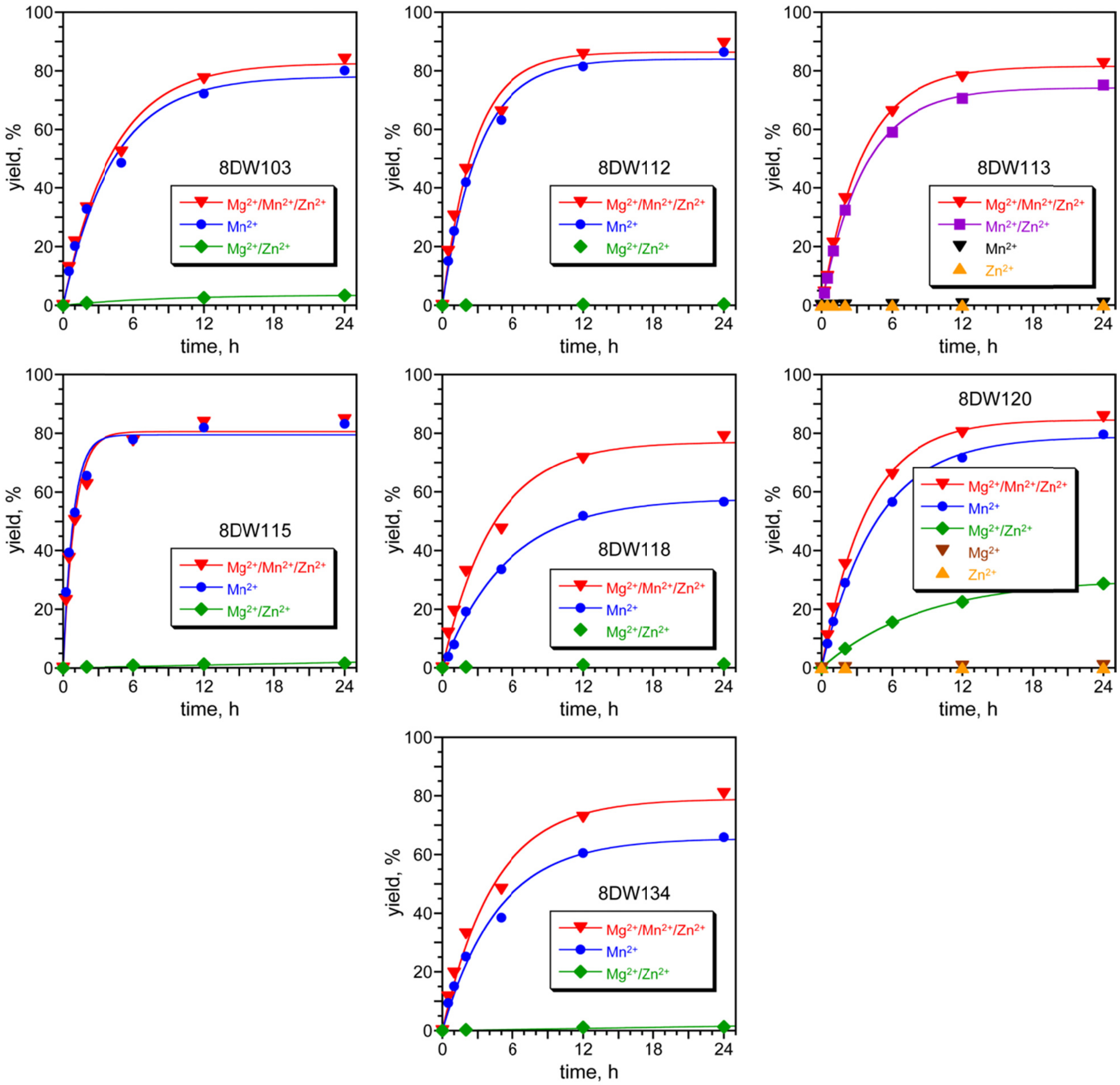

Figure S4. Kinetic plots with various metal ions for all seven 8DW1 deoxyribozymes. Incubation conditions: 50 mM (–Zn^{2+}) or 70 mM (+Zn^{2+}) HEPES, pH 7.5 with 150 mM NaCl and the indicated combinations of 40 mM $MgCl_2$, 20 mM $MnCl_2$, and 1 mM $ZnCl_2$ at 37 °C. These data sets are representative; each illustrated data set was acquired at least two times independently. k_{obs} values (h^{-1}): 8DW103 Mg/Mn/Zn 0.24, Mn 0.23, Mg/Zn 0.0022; 8DW112 Mg/Mn/Zn 0.37, Mn 0.32; 8DW113 Mg/Mn/Zn 0.28, Mn/Zn 0.27; 8DW115 Mg/Mn/Zn 1.02, Mn 1.20, Mg/Zn 0.0010; 8DW118 Mg/Mn/Zn 0.23, Mn 0.18; 8DW120 Mg/Mn/Zn 0.26, Mn 0.22, Mg/Zn 0.12; 8DW134 Mg/Mn/Zn 0.23, Mn 0.21; Mg/Zn 0.0009.

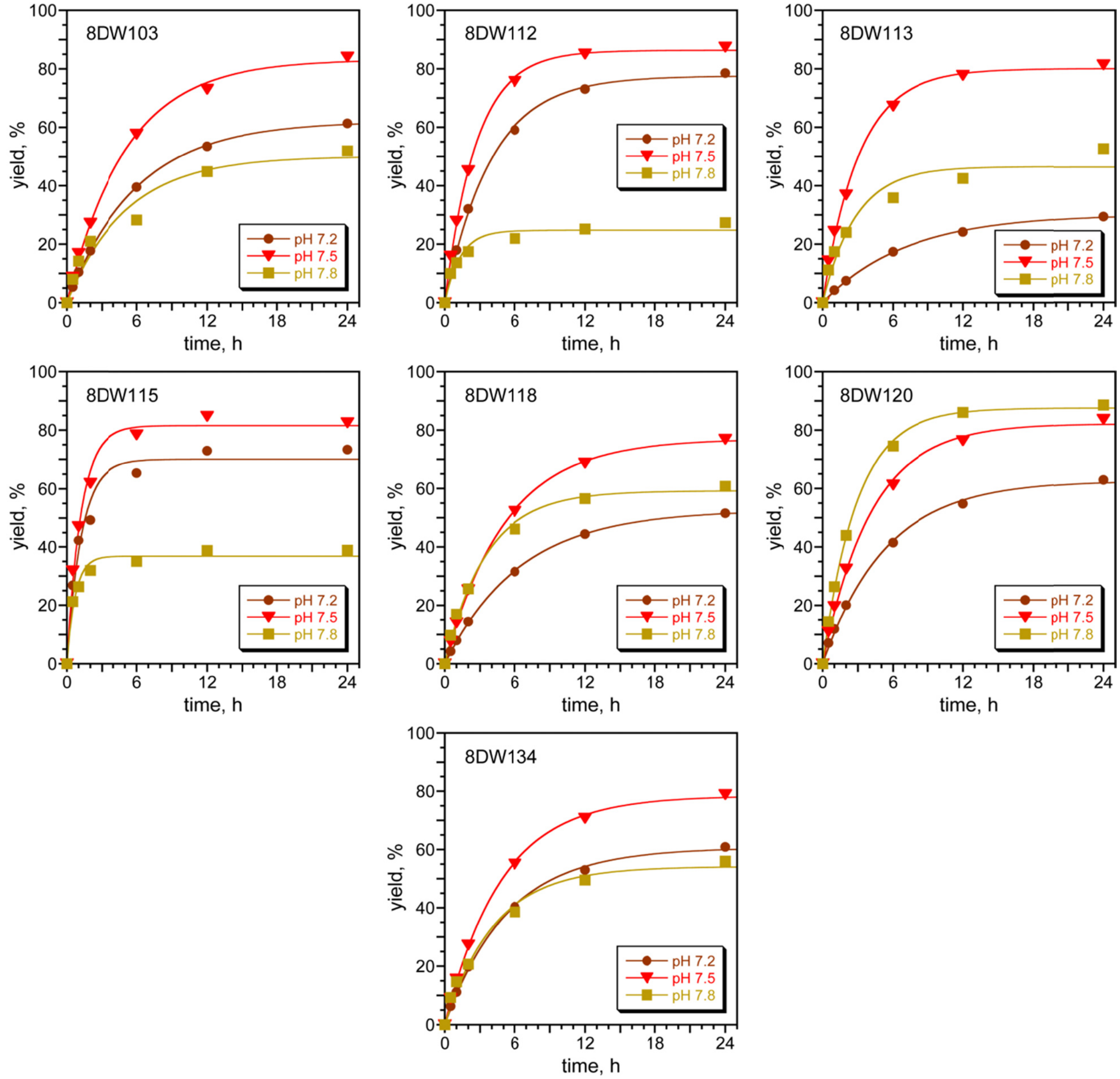

Figure S5. Kinetic plots at various pH values for all seven 8DW deoxyribozymes. Incubation conditions: 70 mM HEPES, pH 7.2 or 7.5 or 7.8, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, and 150 mM NaCl at 37 °C.

Assays of the 7DX1 deoxyribozymes

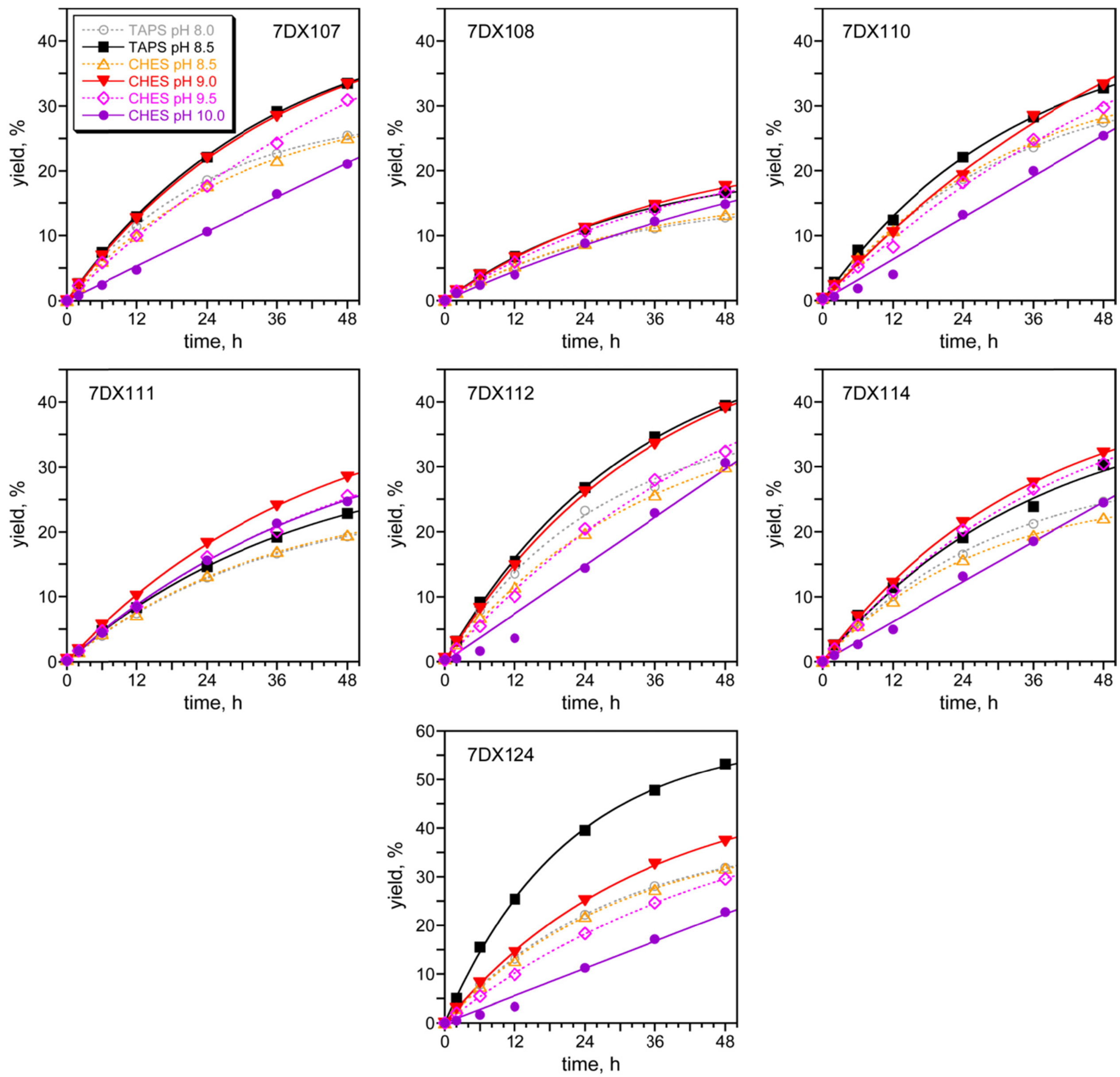

Figure S6. Kinetic plots at various pH values for all seven 7DX1 deoxyribozymes. Incubation conditions: 50 mM buffer, pH 8.0–10.0 with 40 mM $MgCl_2$ and 150 mM NaCl at 37 °C.

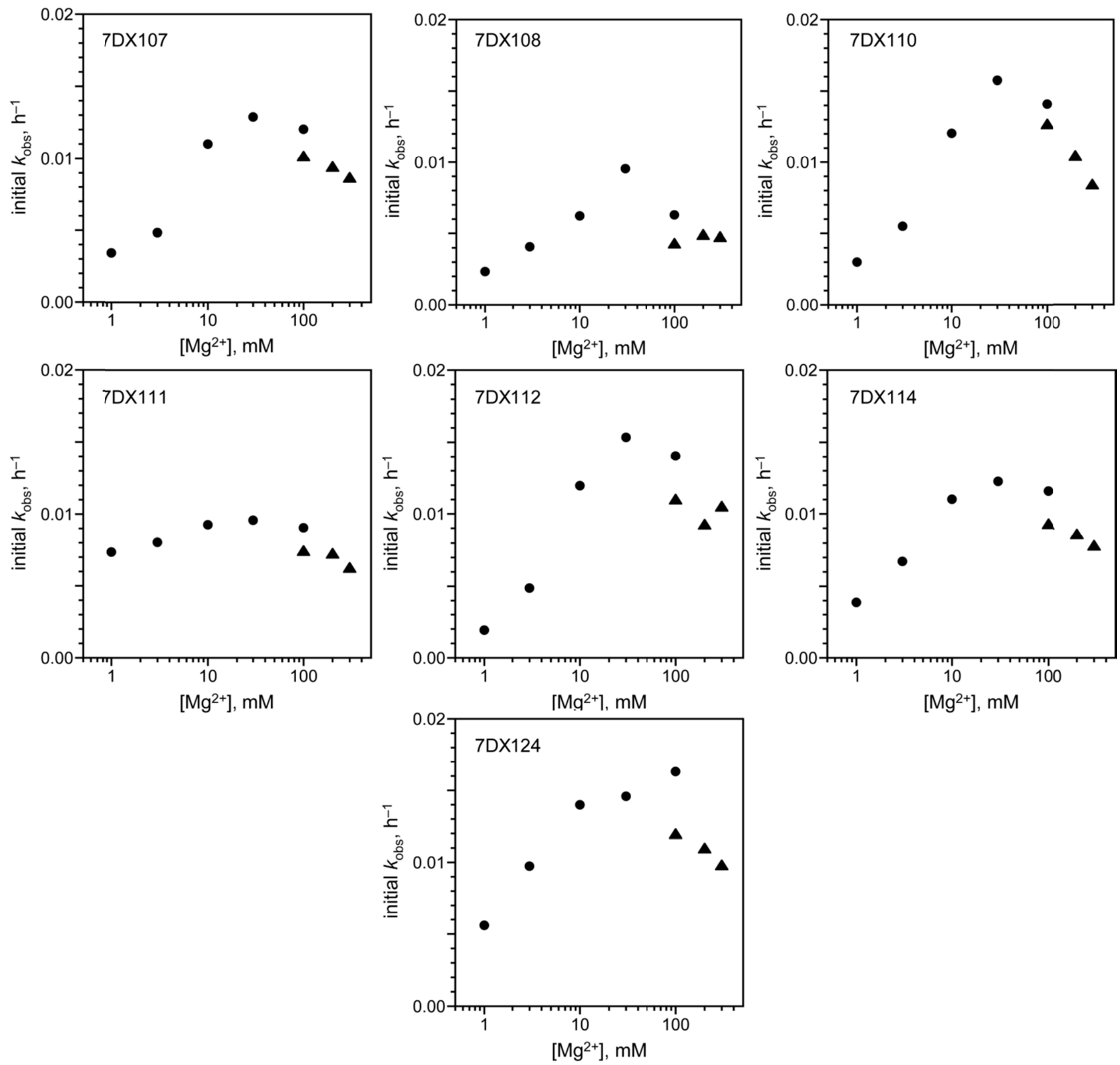

Figure S7. Mg^{2+} concentration dependence of the 7DX1 deoxyribozymes. Incubation conditions: 50 mM CHES, pH 9.0, 0–300 mM $MgCl_2$, and 150 mM NaCl at 37 °C. k_{obs} values were determined from initial-rate kinetics (linear fits to plots of yield versus time from 0 to 12 h). The two symbols represent data sets collected at different times.

pH dependence of the 9DT105 deoxyribozyme

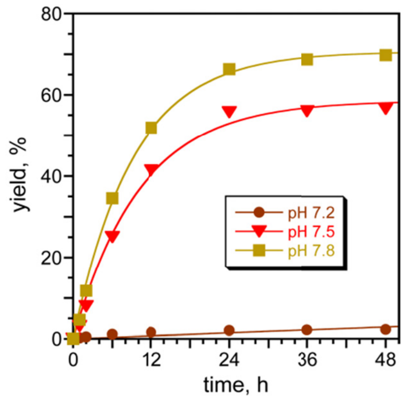

Figure S8. pH dependence of the 9DT105 deoxyribozyme. Incubation conditions: 70 mM HEPES, pH 7.2 or 7.5 or 7.8, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, and 150 mM NaCl at 37 °C.

Determination of reaction product of the 9DT114 deoxyribozyme

The 9DT114 deoxyribozyme forms a reaction product with k_{obs} 0.17 h^{-1}, requiring both Mn^{2+} and Zn^{2+} for optimum yield (Fig. S9a). A DNA substrate with 3′-OH was used, because the 3′-HEG-CKA segment was found to be dispensable for catalysis. The DNA substrate was 5′-GGATAATACGACTCACTAT-3′, where the 3′-terminal T is designated as T(–1), the adjacent A is A(–2), and so on. The T nucleotide in the deoxyribozyme 3′-binding arm located across from A(–5) in the substrate was found to be deleted; restoring this T into the binding arm led to complete loss of catalytic activity (data not shown).

Experiments were performed to determine the connectivity of the product formed by 9DT114 (Fig. S9b). Three 5′-^{32}P-radiolabeled DNA substrates with 3′-OH were prepared, each with a single ribonucleotide substitution at one of C(–4), A(–5), or C(–6). All three ribo-substituted substrates were accepted by 9DT114. The product from each substrate was isolated by PAGE and cleaved under alkaline conditions (100 mM NaOH, 75 °C, 2 h), which removes several nucleotides from the 3′-end of the product and leaves behind a 2′,3′-cyclic phosphate. The 20% PAGE migration rates of the three products clearly reveal that C(–4) is the site of nucleophilic reactivity on the DNA substrate. The rC(–4) product, after alkaline cleavage, migrates considerably more slowly than does the analogous rC(–4) substrate after alkaline cleavage (green arrow in the figure). In contrast, the rA(–5) and rC(–6) products both migrate at the same position as their analogous ribo-modified substrates after alkaline cleavage.

The only reasonable nucleophile on nucleotide C(–4) is its C^4-NH_2 group, which would form a phosphoramidate (P–N) bond upon reaction with 5′-ImpDNA. To support the assignment of the C^4-NH_2 functional group of nucleotide C(–4) as the 9DT114 modification site, a functional group deletion experiment was performed (Fig. S9c). Two new 5′-^{32}P-radiolabeled DNA substrates with 3′-OH were prepared, using Glen Research phosphoramidites for solid-phase synthesis. Each substrate was substituted at C(–4) with either 5-methyl-2′-deoxyzebularine (where zebularine is the cytidine analogue that lacks the C^4-NH_2 group) or 5-methyl-2′-deoxycytidine as a control, noting that for zebularine, only the 5-methyl-2′-deoxy phosphoramidite is commercially available. After 48 h incubation, the parent substrate with cytidine provided 51% yield; the substrate with 5-methyl-2′-deoxycytidine provided 7.2% yield; and the substrate with 5-methyl-2′-deoxyzebularine provided 0.5% yield. Therefore, the 5-methyl modification by itself is deleterious but still allows activity, whereas removal of the C^4-NH_2 group almost completely abolishes activity. This finding is consistent with the conclusion that 9DT114 uses the C^4-NH_2 group of nucleotide C(–4) of the DNA substrate as the nucleophile. The identity of the trace amount of product observed with the 5-methyl-2′-deoxyzebularine substrate is unknown; this product likely arises from secondary usage by 9DT114 of an alternate nucleophile in the substrate.

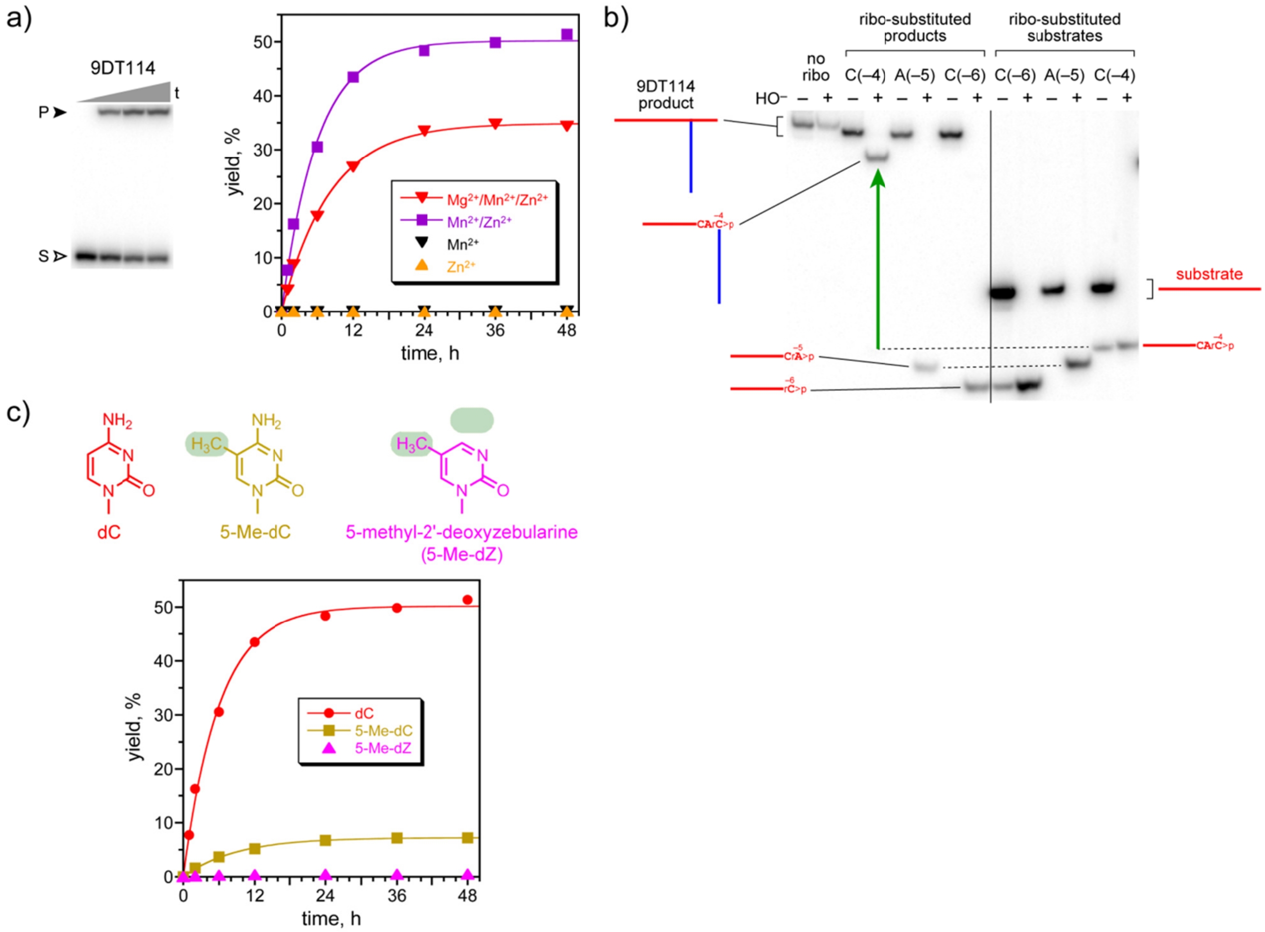

Figure S9. 9DT114 deoxyribozyme and determination of its reaction product connectivity. a) Kinetic assay and metal dependence of 9DT114, using the DNA-3′-OH substrate. Incubation conditions: 50 mM (–Zn^{2+}) or 70 mM (+Zn^{2+}) HEPES, pH 7.5 with 150 mM NaCl and the indicated combinations of 40 mM $MgCl_2$, 20 mM $MnCl_2$, and 1 mM $ZnCl_2$ at 37 °C (t = 30 s, 6 h, 24 h, 48 h). b) Using ribo-substituted DNA substrates to ascertain that 9DT114 modifies the DNA substrate at nucleotide C(–4). c) Functional group deletion experiments to support the assignment of the C^4-NH_2 of nucleotide C(–4) as the nucleophile (conditions of panel a, with Mn^{2+} and Zn^{2+}).

Assays of the 14DV1 deoxyribozymes

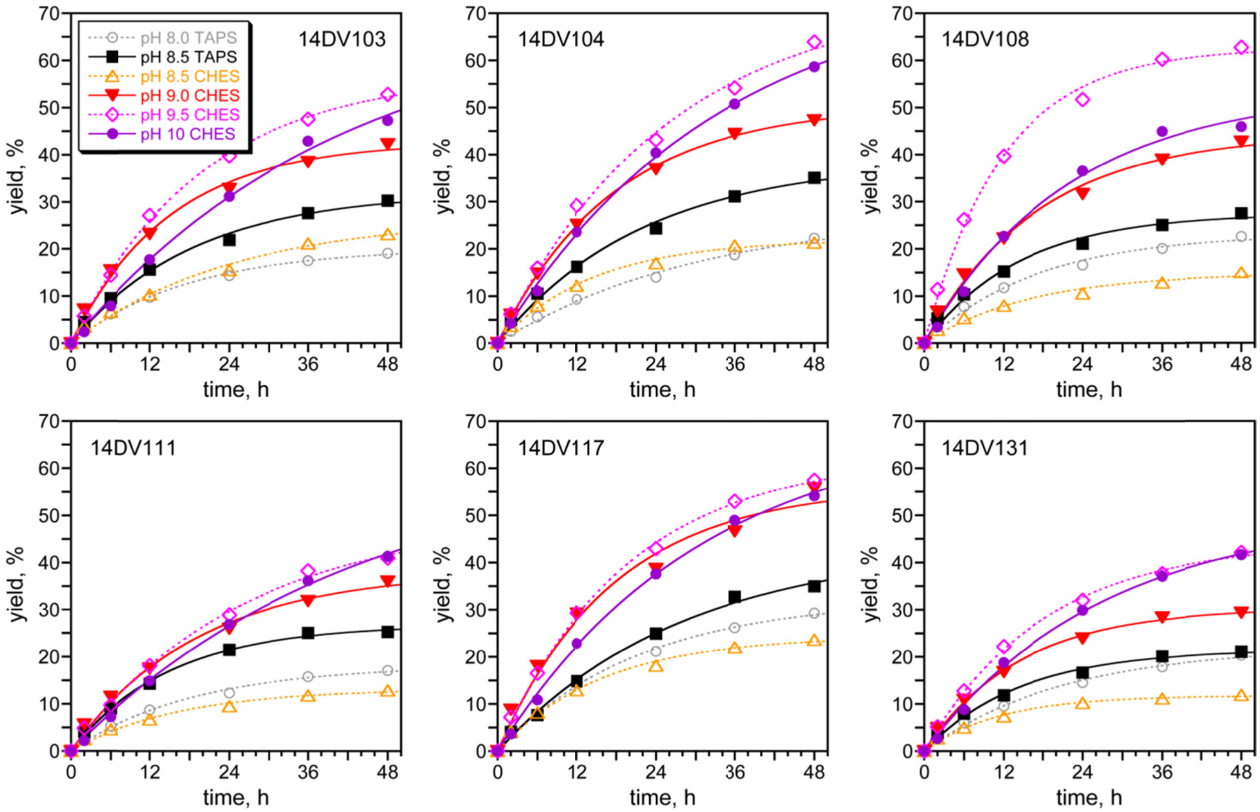

Figure S10. Kinetic plots at various pH values for all six 14DV1 deoxyribozymes. Incubation conditions: 50 mM buffer, pH 8.0–10.0 with 40 mM $MgCl_2$ and 150 mM NaCl at 37 °C.

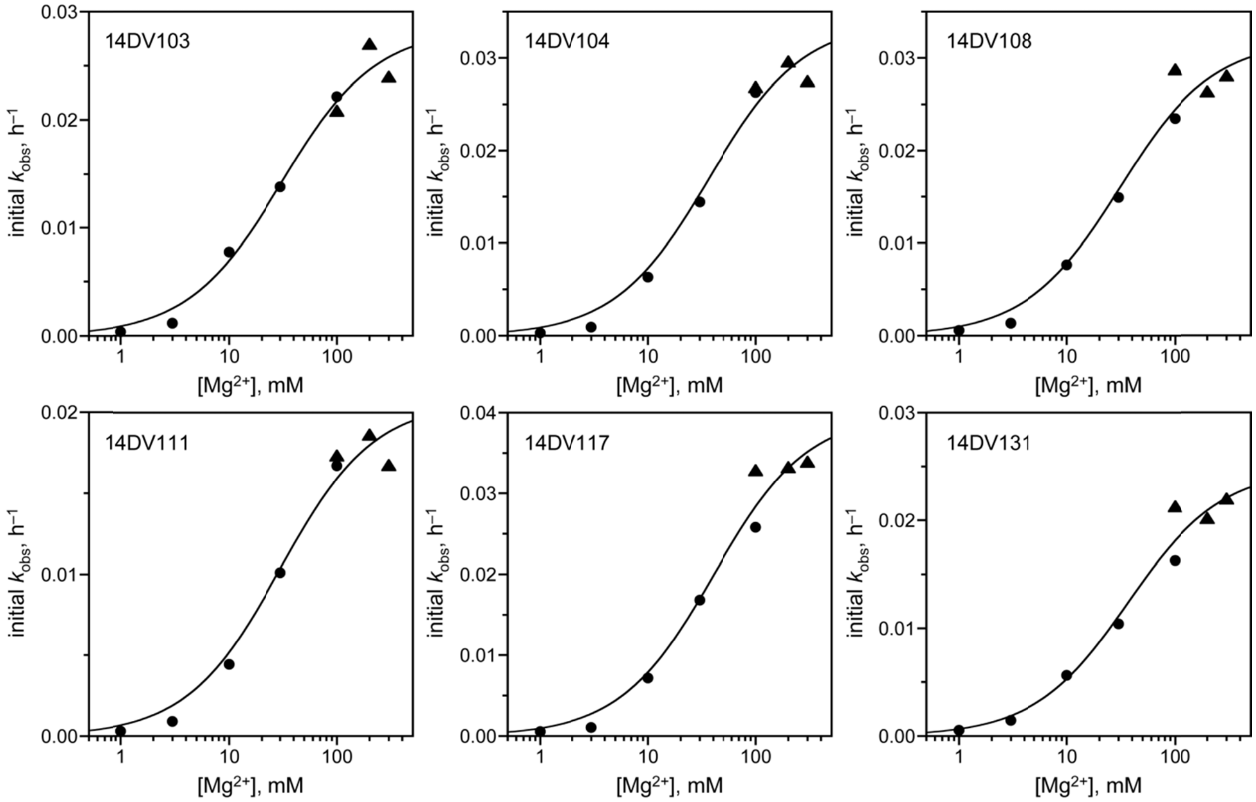

Figure S11. Mg^{2+} concentration dependence of the 14DV1 deoxyribozymes. Incubation conditions: 50 mM CHES, pH 9.0, 0–300 mM $MgCl_2$, and 150 mM NaCl at 37 °C. k_{obs} values were determined from initial-rate kinetics (linear fits to plots of yield versus time from 0 to 6 h). The two symbols represent data sets collected at different times. Each illustrated curve fit is to the standard titration equation $k_{obs} = k_{max}\bullet([Mg^{2+}]/(K_d+[Mg^{2+}])$, assuming a Hill coefficient of 1. Each combined data set is fit well with apparent K_d of 30 to 41 mM. However, because the data do not turn over at high $[Mg^{2+}]$ (up to 300 mM), we more conservatively estimate that each apparent K_d is >50 mM. Curve fit values of apparent K_d (mM): 14DV103 31 ± 7; 14DV104 37 ± 8; 14DV108 31 ± 8; 14DV111 30 ± 7; 14DV117 41 ± 10; 14DV131 37 ± 10.

<u>MALDI mass spectrometry data for the deoxyribozyme products</u>

The products of representative individual deoxyribozymes were analyzed by MALDI mass spectrometry (Fig. S12). Each product was prepared from a 50 µL sample containing 200 pmol of DNA-C_3-NH_2 or DNA-HEG-CKA substrate, 300 pmol of deoxyribozyme, and 400 pmol of 5′-ImpDNA. The sample was annealed in (for conditions *A*) 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA or (for conditions *B*) 5 mM CHES, pH 9.0, 15 mM NaCl, and 0.1 M EDTA by heating at 95 °C for 3 min and cooling on ice for 5 min. The DNA-catalyzed reaction was initiated by bringing the sample to 100 µL total volume containing (conditions *A*) 70 mM HEPES, pH 7.5, 1 mM $ZnCl_2$, 20 mM $MnCl_2$, 40 mM $MgCl_2$, and 150 mM NaCl or (conditions *B*) 70 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl. The sample was incubated at 37 °C for 24 h, precipitated with ethanol, purified by 20% PAGE, and analyzed by MALDI mass spectrometry. Data were acquired on a Bruker UltrafleXtreme MALDI-TOF mass spectrometer with matrix 3-hydroxypicolinic acid in positive ion mode at the UIUC School of Chemical Sciences Mass Spectrometry Laboratory (instrument purchased with support from NIH S10RR027109A).

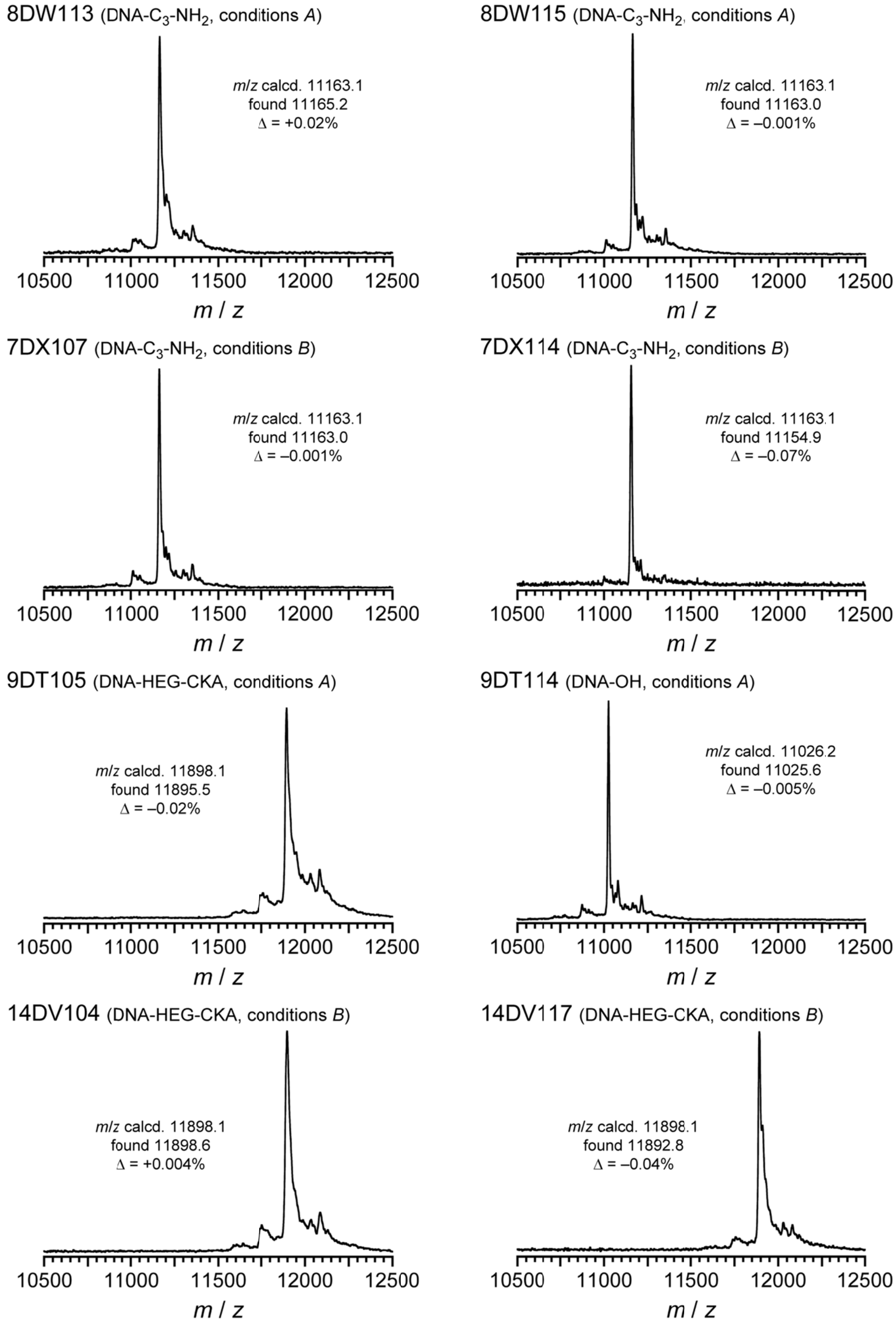

Figure S12. MALDI mass spectrometry to establish identities of products from representative deoxyribozymes. All *m*/*z* values are $[M+H]^+$. Although 9DT114 was identified by selection with the DNA-HEG-CKA substrate, it functions well with the unmodified DNA substrate (3′-OH), which was therefore used for the mass spectrometry experiment. See main text for experimental procedure. For each of 9DT105, 14DV104, and 14DV117, the product was additionally treated with 50 mM DTT (pH 7.5, 37 °C, 2 h), precipitated with ethanol, and again analyzed by MALDI mass spectrometry. In all three cases (not shown), we observed the two peaks expected upon disulfide cleavage between the HEG and CKA components of the product (Fig. 2a).

Acid sensitivity of the phosphoramidate products

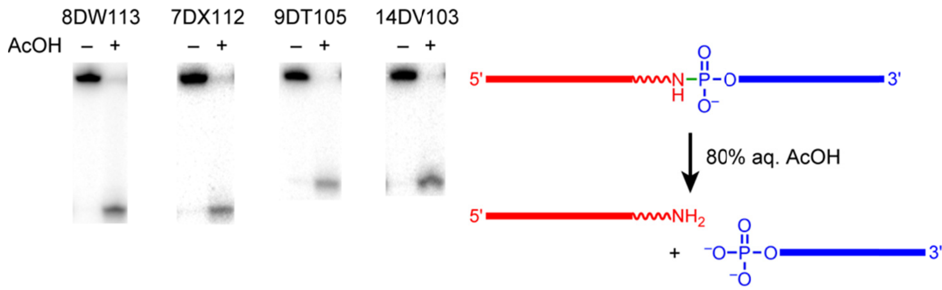

Figure S13. Treatment with 80% aqueous acetic acid to verify the identities of the newly formed phosphoramidate (P–N) linkages. A phosphoramidate linkage involving an aliphatic amine is sensitive to such treatment.[1] Products from the 21 deoxyribozymes collectively obtained from the four different selection experiments (excluding 9DT114) were each assayed. One representative data image is shown for the deoxyribozymes from each selection experiment (DW1, DX1, and DV1). The 5′-^{32}P-radiolabeled product from each deoxyribozyme was incubated in 80% aqueous acetic acid at 37 °C for 5 h.

Negative controls with substrates lacking the nucleophilic amine group or 5′-Imp

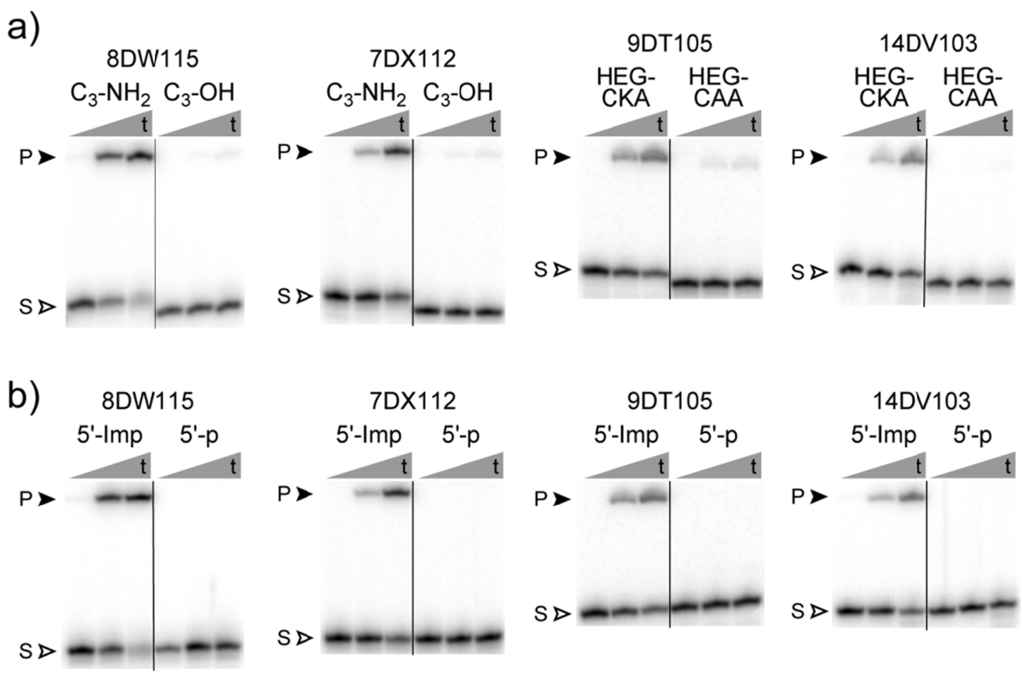

Figure S14. Negative control experiments to assess participation of the nucleophilic amine group or electrophilic 5′-Imp group. The 21 deoxyribozymes collectively obtained from the four different selection experiments (excluding 9DT114) were each assayed. One representative data image is shown for each selection experiment. Incubation conditions as in Fig. 3 and Fig. 4. a) Negative control experiments for nucleophilic amine group. Each deoxyribozyme was assayed with both the DNA-C_3-NH_2 substrate and its DNA-C_3-OH counterpart in which NH_2 is replaced by OH (DW1, DX1 deoxyribozymes) or both the DNA-HEG-CKA substrate and its DNA-HEG-CAA counterpart in which Lys is replaced by Ala (9DT105 and 14DV1 deoxyribozymes). For the 8DW1 and 7DX1 deoxyribozymes, trace activity is observed with C_3-OH rather than C_3-NH_2 (e.g., 1.3% in 24 h for 8DW115; 1.3% in 48 h for 7DX112). This is unsurprising because the hydroxyl group is still a nucleophile. For 9DT105 and the 14DV1 deoxyribozymes, trace activity is observed with HEG-CAA rather than HEG-CKA (e.g., 2.8% in 48 h for 9DT105; 0.5% in 48 h for 14DV103). The trace product is attributed to a small amount of disulfide reduction of the HEG-CAA substrate, which forms a nucleophilic HEG-thiol that is directly analogous to HEG-NH_2 as tested in Fig. 5. b) Negative control experiments for 5′-Imp group. Each deoxyribozyme was assayed with both the 5′-ImpDNA substrate and its unactivated 5′-phosphorylated DNA precursor.

Dependence of catalysis on substrate structure

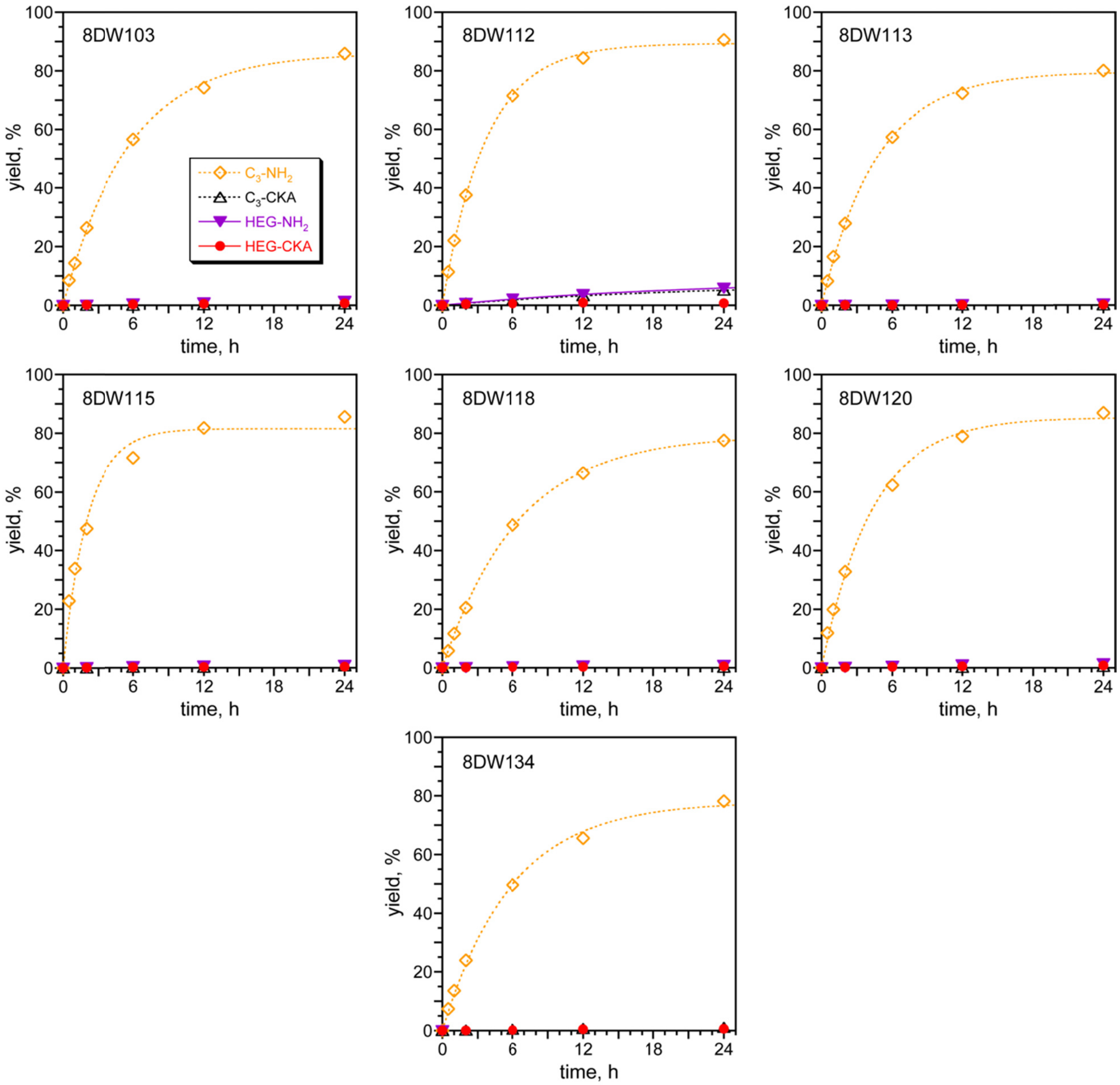

Figure S15. Comprehensive assays of the 8DW1 deoxyribozymes with substrates that have different tether lengths and amine contexts. The 8DW115 data is shown in Fig. 5 and reproduced here to facilitate comparisons. Incubation conditions: 70 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$, and 150 mM NaCl at 37 °C. A small amount (ca. 2%) of a dimeric species is observed when using the C_3-CKA or HEG-CKA substrate, in which the peptide is conjugated to its DNA anchor via a disulfide linkage. This dimer is attributed to a small amount of disulfide reduction followed by thiol-disulfide interchange. For the HEG-CKA substrate, the dimer band migrates slower than the product band on PAGE, and product quantification is not impacted; no correction is needed. For the C_3-CKA substrate, the dimer band exactly comigrates with the product band. As a correction, the average product yield of the 30 s timepoint (ca. 2%) was subtracted from the product yield at each timepoint.

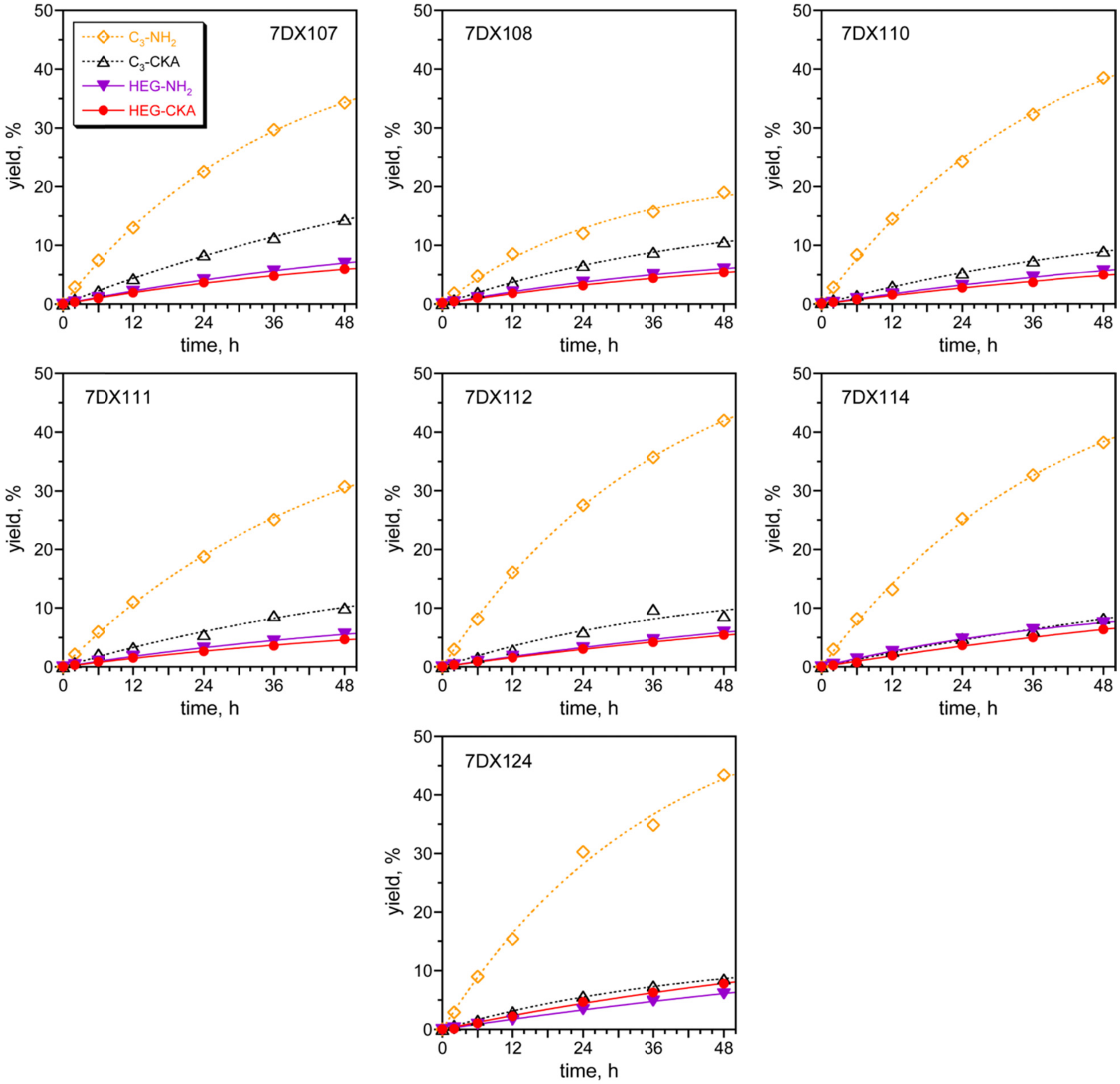

Figure S16. Comprehensive assays of the 7DX1 deoxyribozymes with substrates that have different tether lengths and amine contexts. The 7DX107 data is shown in Fig. 5 and reproduced here to facilitate comparisons. Incubation conditions: 50 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. See Fig. S15 for C_3-CKA correction due to small amount of dimer formation.

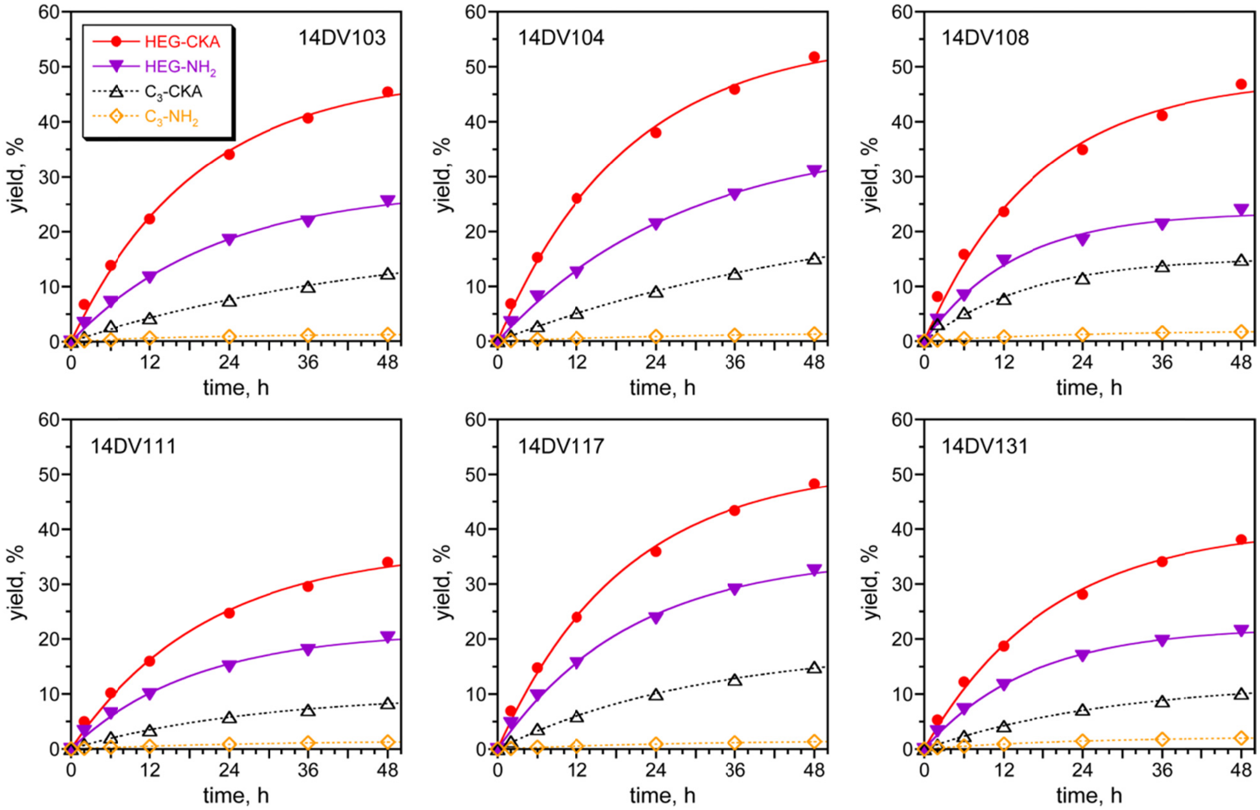

Figure S17. Comprehensive assays of the 14DV1 deoxyribozymes with substrates that have different tether lengths and amine contexts. The 14DV103 data is shown in Fig. 5 and reproduced here to facilitate comparisons. Incubation conditions: 50 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl at 37 °C. See Fig. S15 for C_3-CKA correction due to small amount of dimer formation.

Background reaction rates and rate enhancement calculations

As shown in Fig. S18, observed rate constants, k_{bkgd}, for various background (uncatalyzed) reactions were determined for 5′-ImpDNA reacting with either DNA-C_3-NH_2 or DNA-HEG-CKA under either conditions *A* (pH 7.5, Mg^{2+}/Mn^{2+}/Zn^{2+}) or *B* (pH 9.0, Mg^{2+}). The Imp functional group is relatively unstable under conditions *A*, and the background reaction yields increased very little after 24 h of incubation. Imp is more stable under conditions *B*, and yields continued to increase through 96 h.

For both substrates under conditions *A*, the deoxyribozyme was replaced with either an exactly complementary DNA splint (equivalent to the deoxyribozyme but with deletion of the entire 40-nucleotide catalytic region; "$n = 0$"), a complementary DNA splint that includes a single unpaired T nucleotide in place of the 40-nucleotide catalytic region ("$n = 1$"), or the random N_{40} pool ("$n = 40$"). The $n = 0$ and $n = 1$ assays provide an overestimate of k_{bkgd} for computing the rate enhancement (k_{obs}/k_{bkgd}), because the deoxyribozyme is structurally more complex than a simple $n = 0$ or 1 splint. In contrast, the $n = 40$ assays provide a fairer assessment of the catalytic roles of the particular 40-nucleotide regions.

For both substrates under conditions *B*, a more expansive set of experiments was performed. The deoxyribozyme was replaced with $n = 0$, 1, 2, 3, and 4 splints (with 0 to 4 unpaired T nucleotides in place of the 40-nucleotide catalytic region), the $n = 40$ random pool as splint, a "hairpin" splint in which the 40-nucleotide catalytic region was replaced with a well-defined 6 bp stem-loop structure, or a "scrambled hairpin" splint in which the stem-loop sequence was scrambled and therefore lacked any secondary structure. The hairpin splint has the effect of merely holding the two substrate-containing duplexes close together, which would be achievable by a 40-nucleotide catalytic region that is selected for simple hairpin formation rather than formation of a more sophisticated catalytic structure. The scrambled hairpin splint is functionally equivalent to a random pool splint that has $n = 16$ nucleotides.

For the DNA-C_3-NH_2 and DNA-HEG-CKA substrates under conditions *A*, the $n = 40$ yields at 24 h were 0.22% and 0.20%, respectively. From these values, we estimate $k_{bkgd} \approx 10^{-4}$ h^{-1} in both cases. Noting k_{obs} as high as 1.2 h^{-1} for DNA-C_3-NH_2 (Fig. 3) and 0.08 h^{-1} for DNA-HEG-CKA (Fig. 4), the corresponding rate enhancements (k_{obs}/k_{bkgd}) are ~10^4 and ~10^3.

For the DNA-C_3-NH_2 substrate under conditions *B*, the yields at 96 h for the $n = 40$ and scrambled hairpin splint were 2.2% and 2.5%, respectively, corresponding to k_{bkgd} of ~2×10^{-4} h^{-1}. With k_{obs} of ~0.03 h^{-1} for the 7DX1 deoxyribozymes (Fig. S6), the rate enhancements (k_{obs}/k_{bkgd}) are 150 (~10^2).

For the DNA-HEG-CKA substrate under conditions *B*, the yields at 96 h for the $n = 40$ and scrambled hairpin splint were 1.9% and 1.6%, respectively, corresponding to k_{bkgd} of ~2×10^{-4} h^{-1}. With k_{obs} of ~0.05 h^{-1} for the 14DV1 deoxyribozymes (Fig. S10), the rate enhancements (k_{obs}/k_{bkgd}) are 250 (~10^2).

Note that for the DNA-C_3-NH_2 background assay under conditions *B*, the hairpin splint gave much lower yield than did the $n = 0$ splint (7% versus 60% at 96 h). This finding emphasizes the very high degree of preorganization provided by the $n = 0$ splint and supports the use of the $n = 40$ splint to compute k_{bkgd}. In this assay, also note the steady drop in yield from $n = 0$ to $n = 4$, the equivalent yields for $n = 4$ and the hairpin splint, and the equivalent yields for $n = 40$ and the scrambled hairpin splint.

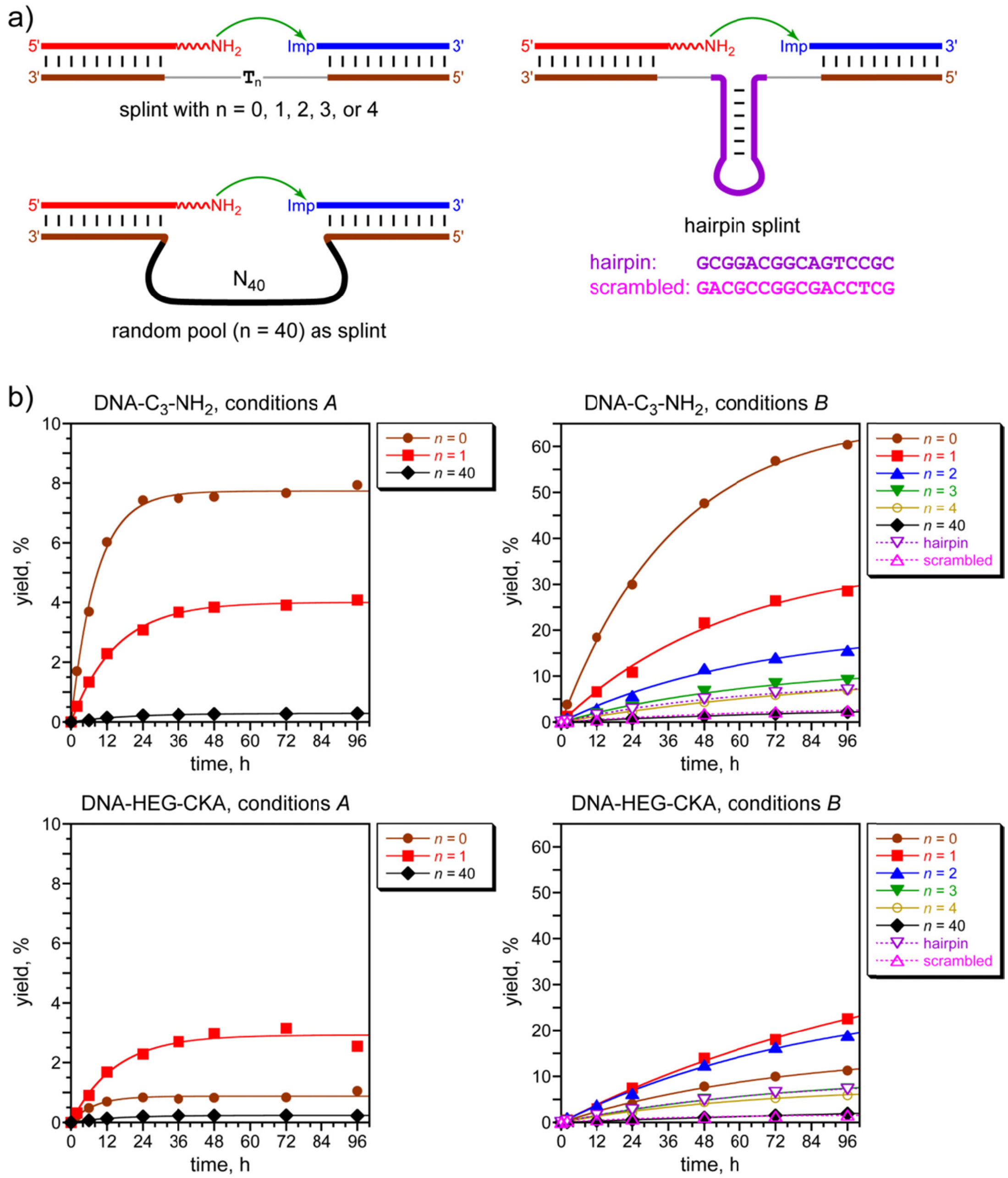

Figure S18. Determining observed rate constants, k_{bkgd}, for various background (uncatalyzed) reactions. a) Depiction of the background assays. b) Kinetic plots.

<u>References for Supporting Information</u>

[1] a) T. Wada, T. Moriguchi, M. Sekine, *J. Am. Chem. Soc.* **1994**, *116*, 9901-9911; b) M. Chandra, A. Sachdeva, S. K. Silverman, *Nat. Chem. Biol.* **2009**, *5*, 718-720; c) A. Sachdeva, S. K. Silverman, *Chem. Commun.* **2010**, *46*, 2215-2217; d) A. Sachdeva, S. K. Silverman, *Org. Biomol. Chem.* **2012**, *10*, 122-125.