



Supporting Information

DNAzyme-Catalyzed Site-Specific N-Acylation of DNA Oligonucleotide Nucleobases

*M. M. Kennebeck, C. K. Kaminsky, M. A. Massa, P. K. Das, R. D. Boyd, M. Bishka,
J. T. Tricarico, S. K. Silverman**

Table of Contents

Oligonucleotide preparative procedures.....	page S3
Preliminary in vitro selection experiment	page S3
Activation of 5'-CO ₂ H to form 5'-TFPE oligonucleotide	page S4
Common procedure for DNAzyme-catalyzed reactions	page S5
Uncatalyzed splinted background reactions	page S5
Procedures for in vitro selection and cloning.....	page S7
Sequences of individual DNAzymes.....	page S8
Single-turnover DNAzyme assay procedure.....	page S9
Yield data for DNAzymes.....	page S10
Product hydrolysis under alkaline conditions to establish amide formation	page S11
Mass spectrometry of DNAzyme products	page S12
Assays with CCCT ₆ ddC substrate of DNAzymes that N-acylate C nucleobases	page S12
Base treatment assays to identify nucleotide attachment sites	page S13
DNAzyme secondary structure predictions using mfold.....	page S15
References for Supporting Information.....	page S25

Oligonucleotide preparative procedures

DNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) or prepared by solid-phase synthesis on an ABI 394 instrument using reagents from Glen Research, including the 5'-CO₂H modifier (5'-carboxy-modifier C10, 10-1935) and the ddC solid support (2',3'-ddC-CPG, 20-2017). Oligonucleotides were purified by 7 M urea denaturing 20% or 8% PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3), extracted from the polyacrylamide with TEN buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 300 mM NaCl), and precipitated with ethanol. The 18 nt 5'-CO₂H oligonucleotide was purified by HPLC, using a Shimadzu Prominence instrument with a Phenomenex Gemini-NX C₁₈ column (5 μm, 10 × 250 mm), solvent A (20 mM triethylammonium acetate in 50% acetonitrile/50% water, pH 7.0), solvent B (20 mM triethylammonium acetate in water, pH 7.0), gradient of 15% A/85% B at 0 min to 30% A/70% B at 45 min, and flow rate of 3.5 mL/min, with retention time of 23.7 min.

Preliminary in vitro selection experiment

We performed a preliminary in vitro selection experiment, alphanumerically designated JB2, to identify DNAzymes that use the 5'-TFPE oligonucleotide of Figure 1A for C nucleobase N-acylation. We found that unlike PE and 4FPE, each of which must be formed preparatively from a 5'-CO₂H oligonucleotide and then HPLC-purified, TFPE has the practical advantage that it can be formed in situ from the corresponding phenol and a 5'-CO₂H oligonucleotide. In vitro selection was performed using the CCC 3'-overhang substrate (Figure 2A and Figure S4). In each selection round, one of the two 5'-CO₂H oligonucleotides of Figure S4 was used to form the TFPE (short n=0, 18 nt, in even-numbered rounds; long n=1, 55 nt, in odd-numbered rounds). The selection procedure was performed as described in the section "Procedures for in vitro selection and cloning", except the enrichment step incubation had final conditions of 70 mM HEPES, pH 7.5, 40 mM MgCl₂, 20 mM MnCl₂, 1 mM ZnCl₂, 150 mM NaCl, 50 mM each EDC and TFP at 37 °C for 2 h. The 5'-CO₂H oligonucleotide (250 pmol in round 1; 50 pmol in subsequent rounds) was activated in 12 μL (round 1) or 6 μL (subsequent rounds) of 167 mM each EDC and TFP at room temperature for 2 h, then added to the other reaction components for a final volume of 40 μL (round 1) or 20 μL (subsequent rounds). The selection experiment showed 37% yield at round 8, and a substantial 19% yield was maintained when the DNA substrate was not connected to the pool (in trans). Cloning revealed five DNAzymes, with initial assays shown in Figure S1.

Only after this preliminary TFPE selection was completed did we realize that we had not properly controlled the pH during the enrichment step, because insufficient HEPES buffer concentration was included to overwhelm the EDC and TFP used to activate the 5'-CO₂H. Empirically we found that the actual pH value in the enrichment step was 5.9 rather than 7.5. Because of this pH discrepancy, we included the combination of the CCC 3'-overhang substrate and three properly controlled pH values in our subsequent systematic in vitro selection experiments. Fortunately, as shown in Figure S1, we found that the five DNAzymes from this preliminary selection experiment each had approximately equivalent activities under the conditions of their selection (pH 5.9) and under the properly pH-controlled conditions (pH 6.5), where these conditions were subsequently developed as described in the section "Activation of 5'-CO₂H to form 5'-TFPE oligonucleotide".

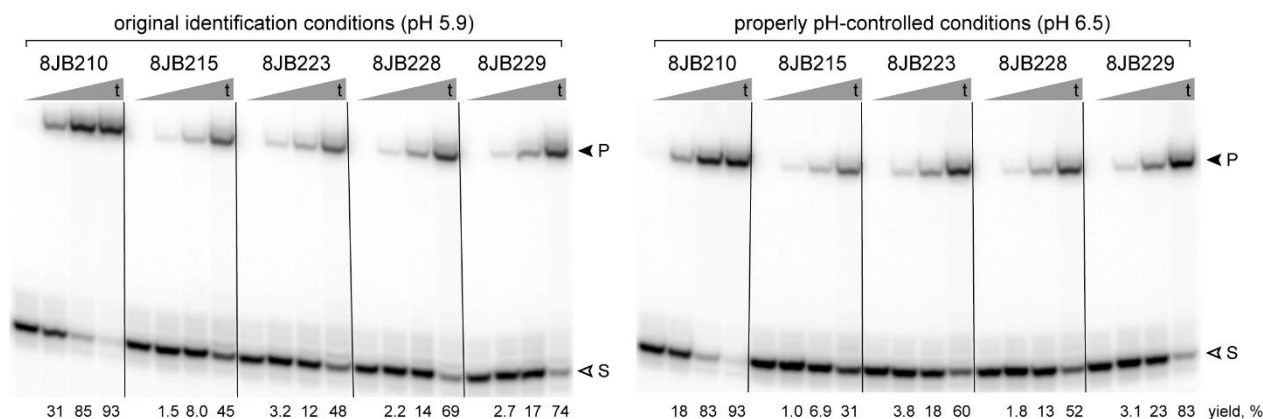


Figure S1. Assays of the five 8JB2 DNAzymes found in the preliminary *in vitro* selection experiment, using the pH 5.9 conditions of their identification as described in the text (left) and using the properly controlled pH 6.5 conditions (right). $t = 0.5$ min, 0.5 h, 2 h, 16 h. S = substrate, P = product. Sequences of these five DNAzymes are in Figure S5.

Activation of 5'-CO₂H to form 5'-TFPE oligonucleotide

To form the 5'-TFPE oligonucleotide for enrichment steps of selection, single-turnover DNAzyme assays, and mass spectrometry assays, a 6, 6, 6, 12, or 12 μ L sample tube containing 10, 40, 50, 250, or 360 pmol of 5'-CO₂H oligonucleotide, 100 mM EDC (added from a 600 mM stock in water), and 100 mM TFP (added from a 600 mM stock in DMF) in unbuffered water was placed at room temperature for 1 h (up to 16 h is permissible). For enrichment steps of selection, single-turnover DNAzyme assays, and making DNAzyme products for subsequent mass spectrometry analysis, the sample of 5'-TFPE oligonucleotide was added to the other components according to the common procedure (see page S4). For MALDI-TOF mass spectrometry analysis of 5'-TFPE, the 6 μ L sample containing 40 pmol of 5'-TFPE oligonucleotide was diluted with 4 μ L of water, desalted by Millipore C₁₈ ZipTip, and analyzed 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. The mass spectrum shows $\geq 80\%$ TFPE (Figure S2A).

To achieve the intended incubation conditions in selection for the three properly controlled pH values of 6.5, 7.4, and 8.5 in the presence of 30 mM each EDC and TFP, we manually varied the concentration and pH of buffer components while monitoring with a pH probe. The three final sample compositions were as follows, each also including 30 mM each EDC and TFP as well as 150 mM NaCl. For pH 6.5, we used a mixture of 105 mM HEPES, pH 7.5, 40 mM MgCl₂, 40 mM MnCl₂, and 1 mM ZnCl₂, where 20 mM of the HEPES is contributed by the ZnCl₂ stock solution. For pH 7.4, we used a mixture of 200 mM HEPES, pH 8.0, 20 mM HEPES, pH 7.5, 40 mM MgCl₂, 40 mM MnCl₂, and 1 mM ZnCl₂, where the 20 mM HEPES, pH 7.5, is contributed by the ZnCl₂ stock solution. For pH 8.5, we used a mixture of 240 mM CHES, pH 9.0, and 40 mM MgCl₂. The metal ion stock samples are described in the common procedure.

MALDI-TOF mass spectrometry was used to assess the presence of the 5'-TFPE oligonucleotide as a function of incubation time under each set of conditions (Figure S2B). A 6 μ L sample tube containing 40 pmol of 5'-CO₂H oligonucleotide, 100 mM EDC (added from a 600 mM stock in water), and 100 mM TFP (added from a 600 mM stock in DMF) was placed at room temperature for 1 h, to form the 5'-TFPE oligonucleotide. Separately, a second tube was prepared with the necessary volume of 1 M HEPES or 500 mM CHES buffer to achieve the intended final incubation conditions as described above, then dried fully in a SpeedVac, which is required in most cases to avoid exceeding the final intended volume. To this second tube was introduced the divalent metal ion(s), NaCl, and water, followed by the 5'-TFPE oligonucleotide sample from the first tube, for a final volume of 20 μ L. This sample was incubated at 37 $^{\circ}$ C for 10 min (pH 7.4 only), 2 h, 6 h, or 16 h, followed immediately by ZipTip and mass spectrometry analysis. Notably, the TFPE decay was much more rapid at pH 7.4 than at pH 6.5 or 8.5.

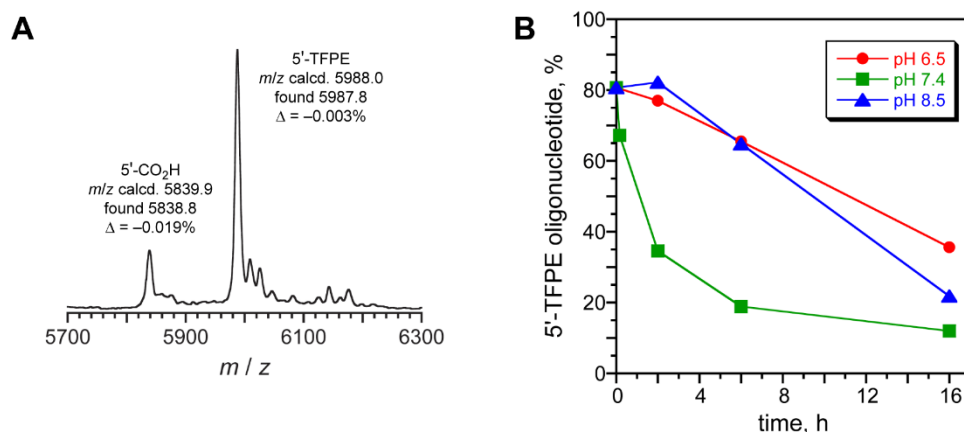


Figure S2. Formation of 5'-TFPE from 5'-CO₂H oligonucleotide, and monitoring 5'-TFPE decay under selection conditions. (A) MALDI-TOF mass spectrum of 18 nt 5'-TFPE oligonucleotide formed under unbuffered conditions (100 mM each EDC and TFP, room temperature, 1 h). (B) Using MALDI-TOF mass spectrometry to monitor the relative amounts of the 5'-TFPE and 5'-CO₂H oligonucleotides upon incubation in the pH 6.5, 7.4, and 8.5 selection conditions. The first nonzero timepoint at pH 7.4 is at 10 min. When a nonspecific EDC nucleobase adduct of the 5'-CO₂H or 5'-TFPE oligonucleotide was apparent in the mass spectrum, this EDC adduct was included as part of the combined peak area for the 5'-CO₂H or 5'-TFPE oligonucleotide.

Common procedure for DNAzyme-catalyzed reactions

This manuscript describes many different purposes and scales of DNAzyme-catalyzed reactions, such as for enrichment steps of selection, for single-turnover DNAzyme assays, and for mass spectrometry assays. In all cases, the common procedural elements are described here, and specific volume and mole amounts are given with the individual procedures. The Mg²⁺ was added from a 10× stock solution containing 400 mM MgCl₂. The Mn²⁺ was added from a 10× stock solution containing 200 mM MnCl₂. The Zn²⁺ was added from a 10× stock solution containing 10 mM ZnCl₂, 20 mM HNO₃, and 200 mM HEPES at pH 7.5; this stock solution was freshly prepared from a 100× stock of 100 mM ZnCl₂ in 200 mM HNO₃. The metal ion stocks were added last to the final sample. The final incubation conditions were as follows, including 30 mM each EDC and TFP, 150 mM NaCl, and 5% (v/v) DMF for the solubility of TFP. pH 6.5: 105 mM HEPES, 40 mM MgCl₂, 20 mM MnCl₂, and 1 mM ZnCl₂. pH 7.4: 220 mM HEPES, 40 mM MgCl₂, 20 mM MnCl₂, and 1 mM ZnCl₂. pH 8.5: 240 mM CHES and 40 mM MgCl₂.

Three tubes were prepared. To tube 1 was added the PCR product (for enrichment steps) or DNAzyme (for assays), the 10× annealing buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, and 1 mM EDTA), and water to reach the intended annealing volume. In tube 2, the 5'-CO₂H oligonucleotide was activated to 5'-TFPE by placing the 5'-CO₂H oligonucleotide, 100 mM EDC, and 100 mM TFP in unbuffered water at room temperature for 1 h. To tube 3 was added the required volume of 1 M HEPES or 500 mM CHES buffer to achieve the intended final incubation conditions (see the section “Activation of 5'-CO₂H to form 5'-TFPE oligonucleotide”), and the contents were evaporated to dryness in a SpeedVac, which is necessary in most cases to avoid exceeding the final intended volume. Tube 1 was annealed by heating at 95 °C for 1 min and cooling on ice for 5 min. The contents of tubes 1 and 2 were added to tube 3, followed by water, NaCl, and metal ions as required to achieve the intended final concentrations. The final sample, which has 30 mM EDC and 30 mM TFP, was incubated at 37 °C.

Uncatalyzed splinted background reactions

Before we began *in vitro* selection, to determine suitable incubation conditions for the enrichment step with each combination of nucleobase and pH, we performed assays of the uncatalyzed splinted background reaction (Figure 2B). The goal was to identify conditions in which the background product is observable but its yield does not exceed 5%, where 5% background limits the enrichment factor per selection round to 20-fold. From the background yields at 2 and 16 h for each combination of nucleobase and pH, we found that one of 2 or 16 h incubation was suitable for each combination, as shown in Figure 3 and marked in

Figure S3A. Separately, for each combination of nucleobase and pH that led to DNazymes (the five selections with red arrows in Figure 3), we determined k_{bkgd} (Figure S3B). These k_{bkgd} values were used to compute rate enhancements $k_{\text{obs}}/k_{\text{bkgd}}$ from single-turnover k_{obs} values such as those in Figure 6.

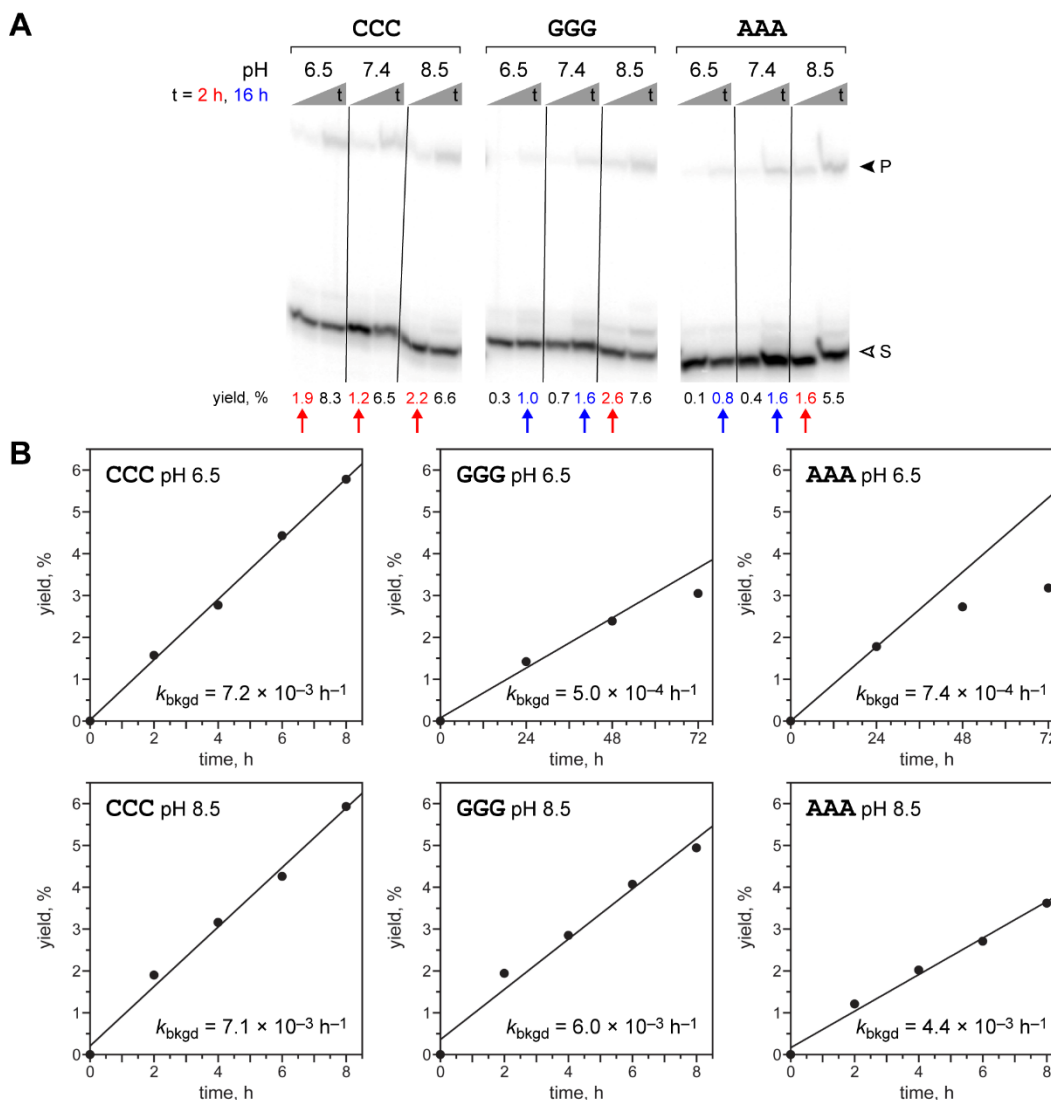


Figure S3. Assays of the uncatalyzed splinted background reactions. (A) Assays at 2 and 16 h for each combination of trinucleotide 3'-overhang and pH, to determine suitable incubation times for the enrichment step of selection. The observed yield for each combination is shown below each lane. The red (2 h) and blue (16 h) arrows mark the chosen conditions, in which the uncatalyzed background yield is <5%. S = substrate, P = background product. (B) Single-turnover assays (initial-rate kinetics) to quantitatively determine the k_{bkgd} values in each plot. No background product was observed (<0.2%) when the splint was omitted, except 0.4% unsplinted background product in 8 h was observed for AAA at pH 8.5. When the trinucleotide 3'-overhang was CCddC (i.e., 2',3'-dideoxy terminus) instead of CCC, the k_{bkgd} values were very close to those observed for CCC: $5.5 \times 10^{-3} \text{ h}^{-1}$ and $6.8 \times 10^{-3} \text{ h}^{-1}$ at pH 6.5 and 8.5, respectively. When the trinucleotide 3'-overhang was TTT, the k_{bkgd} values were very close to those observed for GGG and AAA: $6.3 \times 10^{-4} \text{ h}^{-1}$ and $4.2 \times 10^{-3} \text{ h}^{-1}$ at pH 6.5 and 8.5, respectively. From the combined data, we conclude that at pH 6.5, the CCC background product arises primarily from reaction at one or more of the three C nucleobase N^4 -amino groups. In contrast, for GGG and AAA at pH 6.5, or for CCC, GGG, and AAA at pH 8.5, the data do not allow determination of the balance between nucleobase amine and 3'-OH background reactivity. When the trinucleotide 3'-overhang was GGrG, i.e., a 3'-terminal ribonucleotide, the yield at pH 6.5 and 24 h was 16%, which is 10-fold higher than with GGG. We therefore assign the GGrG background product as an ester, likely from the 2'-OH but possibly with some contribution from the 3'-OH of the 2',3'-diol. This ester-linked product was used as the ester standard in Figure 5. The aliphatic amide standard in Figure 5 was made by splinted background reaction of a substrate in which the trinucleotide 3'-overhang was replaced with an aliphatic $(\text{CH}_2)_3\text{-NH}_2$ linker (MOPS pH 7.0, room temperature, 24 h).^[1]

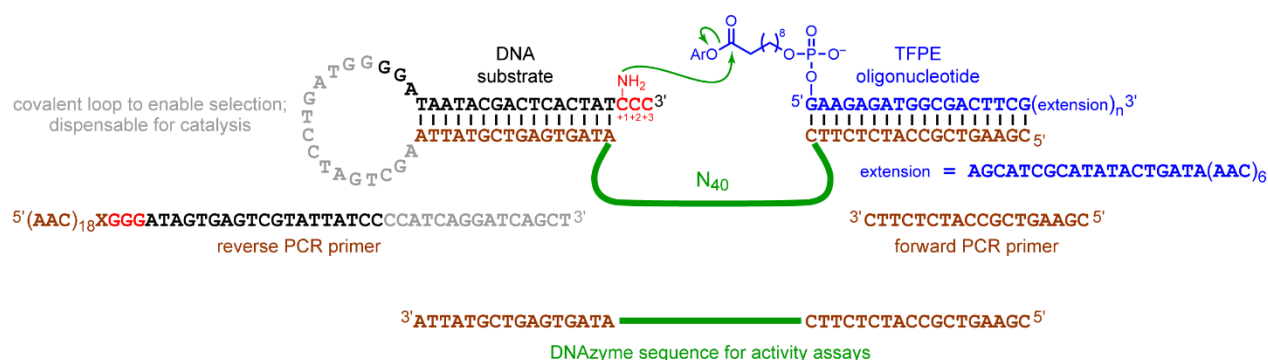
Procedures for in vitro selection and cloning

Figure S4. Nucleotide details of the in vitro selection experiments with 5'-TFPE oligonucleotide as acyl donor (Ar = 2,3,5,6-tetrafluorophenyl). The DNA substrate is shown with CCC 3'-overhang, and the reverse primer has GGG to match; the appropriate changes in the reverse primer were made for GGG and AAA 3'-overhangs. In the reverse primer, X = hexa(ethylene glycol), HEG, spacer to stop Taq polymerase. The short (n=0; 18 nt) version of the 5'-TFPE oligonucleotide was used in even-numbered selection rounds, and the long (n=1; 55 nt) version in odd-numbered rounds, to avoid enriching noncatalytic DNA sequences that have abnormal PAGE migration positions. The 55-mer was prepared from the 18-mer by splint ligation with the 37 nt extension oligonucleotide, 37 nt splint 5'-TATCAGTAT-ATGCGATGCTCGAAGTCGCCATCTCTTC-3', and T4 DNA ligase. Although only one of the three CCC 3'-overhang nucleotides (C+1) is shown as providing the nucleophile for amine acylation, for any individual DNzyme this role could be played by any of these three C nucleotides as well as other nucleotides of the DNA substrate, or the 3'-OH could be the nucleophile for unwanted O-acylation. The product migration standard during each selection round was the uncatalyzed pH 6.5 background product formed from the CCC N₄₀ pool, the 5'-TFPE oligonucleotide, and splint 5'-CGAAGTCGCCATCTCTTCATAGTGAGTCGTATTATCCC-CATCAGGATCAGCT-3' (this is the splint used for the background reaction illustrated in Figure 2B), incubated at 37 °C for 16 h.

Procedure for enrichment step in round 1. Following the common procedure, each selection experiment was initiated with 200 pmol of the pool and 250 pmol of 5'-TFPE oligonucleotide, with annealing volume of 16 μL, activation volume of 12 μL, and final incubation volume of 40 μL. The sample was incubated for 2 or 16 h and separated by 8% PAGE.

Procedure for enrichment step in subsequent rounds. Following the common procedure, each selection round was performed with the 30-cycle PCR product and 50 pmol of 5'-TFPE oligonucleotide, with annealing volume of 8 μL, activation volume of 6 μL, and final incubation volume of 20 μL. The sample was incubated for 2 or 16 h and separated by 8% PAGE.

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, 10 μL of 10× Taq polymerase buffer [1× = 20 mM Tris-HCl, pH 8.8, 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, and 0.1% Triton X-100], and 1 μL of Taq polymerase (expressed as reported^[2]). 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 2 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-³²P-dCTP (800 Ci/mmol), 5 μL of 10× Taq polymerase buffer, and 0.5 μL of Taq polymerase. 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 DNzymes. The PCR primers used for cloning were 5'-CGAAGTCGCCATCTCTTC-3' (forward primer; same as in selection) and 5'-TAATTAATTAATTACCCATCAGGATCAGCT-3' (reverse primer). The 10-cycle PCR product from the appropriate selection round was diluted 10³-fold. A 50 μL sample was prepared containing 1 μL of the diluted 10-cycle PCR product from the appropriate selection round, 25 pmol of forward cloning primer, 25 pmol of reverse cloning primer, 10 nmol of each dNTP, 5 μL of 10× Taq polymerase buffer, and 0.5 μL of Taq polymerase. 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 (Thermo Fisher). The extracted product was quantified by absorbance (A_{260}) and diluted to 5–10 ng/ μ L. A 4 μ L portion of the diluted PCR product was inserted into the pCR2.1-TOPO vector using a TOPO TA cloning kit (Thermo Fisher). 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 (Thermo Fisher) and screened for properly sized inserts by EcoRI digestion and agarose gel analysis. Before sequencing, assays of individual DNAzyme clones were performed with PAGE-purified DNA strands prepared by PCR from the miniprep DNA, using the procedure described in the section “Single-turnover DNAzyme assay procedure”.

Sequences of individual DNAzymes

	1	10	20	30	40	
*8JB210	TAGGCGTACA	TGGCCTGGT	TGGGGGGTA	ACGATGGTCA	40	(1)
8JB215	C...A.AGG	G...CGA.TAC	G...ACAACAC	G.TG.T...G	40	(2)
8JB223	CCA.GACTA	ATA..A.AT	.AC.A..A.C	TGATGT-C.G	39	(1)
8JB228	C...A.GGG	A..CATC.T	GA.T..T.C	CGTCGC.CAC	40	(1)
*8JB229	GG..G..GAT	A.CATT...G	C.....ACAC	.G.GCCAGTC	40	(1)
	1	10	20	30	40	
*6MK208	AGCGAGCATG	AATGGACCGT	GGATGGGGAC	CCTCTAGGCT	40	(3)
6MK205	CCACCAGGGA	..GT.GAACG	.TC...ATCG	TG.GAG...A	40	(3)
6MK215	C.A.GAA.CT	..GCAGG.T	.CTGT...G	AGGGATACAC	40	(1)
6MK232	GAG...ACAT	T.GAA.TG..	..G.....GG	A..AAC.A.C	40	(1)
	1	10	20	30	40	
*7MN217	GAA CGCGGGG	AGAAGTGACG	CGTCCCTTTC	GGACTAGGGC	40	(1)
	1	10	20	30	40	
*8MQ213	AACCAGGATG	GGCGCAAGTG	TATGAATAAT	GCACTACGCG	40	(1)
8MQ209C...	..TAAGC.A	CCATGGA.TG	.T.TCGATTA	40	(1)
8MQ214	GGG.GCCAC	CA.ATCTTAA	CGGC..G.GG	..GGG.ACAC	40	(1)
8MQ215G.A.CC.CA	ATGC.CATCC	T.TAA.G...40	(1)	
8MQ216A...GG...	CGAAGGAGGC	.G...CGT..40	(1)	
8MQ230A...	..T...GA.T	.CGATC.CGG	.GGGCCT..A40	(1)	
8MQ236ACGCA	A.TAGGGACC	GG..CGGT..	...TG.A.G.40	(1)	
8MQ237C...	.TT.A.T.A	CGAATG.GGC	AGCTC.GT.A40	(1)	
8MQ241	GG...AA.CC	TTA.GCTTAA	ACC.G.GG..	TG.AAGT.GA40	(1)	
	1	10	20	30	40	
*7MR202	CGTGGTTCGA	ATGCGGGCAG	AGAACTTGAG	CGGCGTCAGT	40	(1)
7MR205	T.CA.AGGC	G.CGTACACA	...T.GA.GT	GTT..CA..A40	(1)	
7MR206	GCGT.G...C	T.A.CAT.G	G.T.A.G..A	..AGTCTGAC40	(1)	
7MR218	.TGAACGGAG	TCA.A.CGGA	GAGTAAA.G	A.CT.CAGA.40	(1)	
	1	10	20	30	40	
*6MT241	CGGGGGCGAG	AAAAGAACT	ATTATCGAGG	AGAGTTTGCA	40	(1)
6MT201	GCACT.G.TA	G.TGAC.GAA	C.GGA.ATC	.C...GGAGT40	(1)	
6MT204AGA..	GCT.ATGGGA	..GCA.TGAA-.	39	(6)

Figure S5. Sequences of the DNAzymes identified in this study. See the section “Procedures for in vitro selection and cloning” for details of the selection process. Only the initially random (N_{40}) sequences are shown. All DNAzymes were assayed as 5'-CGAAGTCGCCATCTCTTC- N_{40} -ATAGTGAGTCGTATTA-3' (see Figure S4). In each alignment, a dot denotes conservation, i.e., the same nucleotide as in the uppermost sequence within the family; 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. Two sequences are 39 nt, each presumably due to a Taq polymerase deletion in an unknown selection round. The seven representative DNAzymes of Figure 6 are each denoted here with an asterisk in front of their name. In most cases, no common or conserved motifs were apparent either within or between sequence families, although several sequence families have some internal similarities. See Table S4 and Figure S9 for predicted secondary structures of the DNAzymes.

Single-turnover DNAzyme assay procedure

Each single-turnover assay for an individual DNAzyme or background splint followed the common procedure, using 0.5 pmol of DNA or RNA substrate that was 5'-³²P-radiolabeled using γ -³²P-ATP and T4 polynucleotide kinase, 5 pmol of DNAzyme, and 10 pmol of 5'-TFPE acyl donor oligonucleotide, with annealing volume of 8 μ L, activation volume of 6 μ L, and final incubation volume of 20 μ L. At each time point, a 2 μ L aliquot was quenched with 7 μ L of stop solution (80% formamide, 1 \times TBE [89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3], 50 mM EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol). Before PAGE for most assays, to each quenched sample was added 100 pmol of a “decoy oligonucleotide”, which was the 37-mer background splint that is complementary to both DNAzyme binding arms (Figure 2B). This decoy oligonucleotide was added to displace the DNAzyme from the product. In these cases when the decoy was omitted, the product bands were noticeably smeared, which inhibited proper quantification. Quenched samples were separated by 20% PAGE and quantified using a Phosphorimager. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., $\text{yield} = Y \cdot (1 - e^{-kt})$, where $k = k_{\text{obs}}$ and Y is the final yield. Where appropriate, each k_{obs} value is reported with error calculated as the standard deviation from the indicated number of independent determinations. For the background reactions (Figure S3), the initial points were fit to a straight line, and k_{bkgd} was taken as the slope of the line.

Yield data for DNAzymes

DNAzyme	0.5 h yield, %	2 h yield, %	16 h yield, %	48 h yield, %
8JB210	34	87	92	n.d.
8JB215	1	7	31	n.d.
8JB223	4	18	60	n.d.
8JB228	2	13	52	n.d.
8JB229	3	23	83	n.d.
6MK205	8	29	61	n.d.
6MK208	31	77	91	n.d.
6MK215	13	37	81	n.d.
6MK232	24	57	81	n.d.
7MN217	n.d.	15	48	64
8MQ209	25	56	91	n.d.
8MQ213	20	56	91	n.d.
8MQ214	7	23	52	n.d.
8MQ215	11	41	87	n.d.
8MQ216	28	65	92	n.d.
8MQ230	5	10	28	n.d.
8MQ236	8	33	71	n.d.
8MQ237	6	24	69	n.d.
8MQ241	3	8	23	n.d.
7MR202	n.d.	47	94	97
7MR205	n.d.	5	14	24
7MR206	n.d.	6	25	40
7MR218	n.d.	11	45	64
6MT201	3	12	38	n.d.
6MT204	11	29	45	n.d.
6MT241	32	55	70	n.d.

Table S1. Initial yield data for DNAzymes with the corresponding DNA substrate.

Product hydrolysis under alkaline conditions to establish amide formation

DNAzyme	trial 1	trial 2	trial 3	trial 4	trial 5	mean $t_{1/2}$	error
8JB210	0.67	0.94	0.90	0.82	0.64	0.79	0.13
8JB215					0.40		
8JB223					0.43		
8JB228					0.43		
8JB229	0.33	0.61	0.53		0.49	0.49	0.12
6MK205					0.38		
6MK208					0.61		
6MK215					0.46		
6MK232					0.53		
7MN217				0.74	0.67	0.70	0.04
8MQ209					0.55		
8MQ213				0.86	0.54	0.70	0.16
8MQ214					0.54		
8MQ215				0.80	0.36	0.58	0.22
8MQ216					0.62		
8MQ230					0.65		
8MQ236					0.71		
8MQ237					0.71		
8MQ241					0.56		
7MR202				0.56	0.50	0.53	0.03
7MR205					0.44		
7MR206				0.43	0.50	0.47	0.03
7MR218					0.47		
6MT201				0.48	0.47	0.48	0.005
6MT204					0.52		
6MT241					0.40		

Table S2. Hydrolysis half-lives (h) for the acylation products formed by all 26 DNAzymes (see Figure 5). “Error” is listed for the eight DNAzymes with $n = 2$ or more, as standard deviation ($n = 4$ or 5) or half of range ($n = 2$). Mean \pm standard deviation by category: all 26 DNAzymes, 0.54 ± 0.11 ; 9 C DNAzymes (8JB2, 6MK2), 0.50 ± 0.13 ; 10 G DNAzymes (7MN2, 8MQ2), 0.59 ± 0.11 ; 7 A DNAzymes (7MR2, 6MT2), 0.47 ± 0.04 .

Mass spectrometry of DNAzyme products

The products from the seven individual DNAzymes of Figure 6A were analyzed by MALDI-TOF mass spectrometry. Following the common procedure, each reaction was initiated with 300 pmol of DNA substrate, 330 pmol of DNAzyme, and 360 pmol of 5'-CO₂H oligonucleotide, with annealing volume of 16 μ L, activation volume of 12 μ L, and final incubation volume of 40 μ L. The sample was incubated for 24 h, separated by 20% PAGE, and desalted by Millipore C₁₈ ZipTip. 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. All m/z values are for [M+H]⁺.

DNAzyme	m/z calcd.	m/z found	Δ
8JB210	12485.2	12481.1	-0.03%
8JB229	12485.2	12481.9	-0.03%
6MK208	12485.2	12482.1	-0.03%
7MN217	12605.3	12602.3	-0.02%
8MQ213	12605.3	12599.7	-0.04%
7MR202	12557.3	12557.0	-0.002%
6MT241	12557.3	12554.3	-0.02%

Table S3. MALDI-TOF mass spectrometry data of representative DNAzyme products.

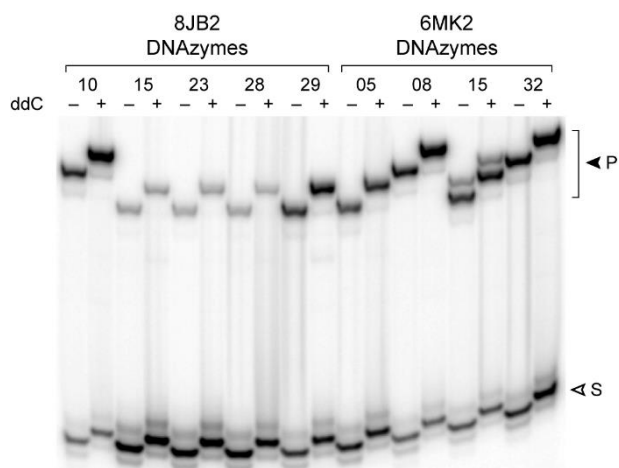
Assays with CCCT₆ddC substrate of DNAzymes that N-acylate C nucleobases

Figure S6. Assay results using DNA substrates with CCCT₆ and CCCT₆ddC 3'-overhangs, analogous to the results in Figure 4 for the same DNAzymes using DNA substrates with CCC and CCddC 3'-overhangs.

Base treatment assays to identify nucleotide attachment sites

For the assay of Figure 7B, the 33 nt RNA substrate was the 22 nt 3'-CCC substrate ligated with the arbitrary 11 nt sequence 5'-UGUGCACGCCU-3', using 51 nt splint 5'- AACAACAACAACAACAACAG-GCGTGCACAGGGATAGTGAGTCGTATTATCC-3' (ligation junction underlined) and T4 DNA ligase. We followed the common procedure for making the 8JB210 product using 3 pmol of 5'-³²P-radiolabeled RNA substrate, 10 pmol of DNAzyme, and 20 pmol of 5'-TFPE acyl donor oligonucleotide, with annealing volume of 8 μ L, activation volume of 6 μ L, and final incubation volume of 20 μ L. The sample was incubated for 24 h, separated by 12% PAGE (on 20% PAGE, the DNAzyme itself overlaps with the product), and redissolved in 20 μ L of water. For the base treatment assay, 4 μ L of the product (and a parallel sample of 33 nt RNA substrate) was brought to 6 μ L total volume containing 50 mM Tris, pH 8.0, and 200 mM MgCl₂, incubated at 90 °C for 2 min, and quenched with 10 μ L of stop solution. These assay conditions were optimized with regard to pH, Mg²⁺ concentration, temperature, and time, to minimize amide hydrolysis (acyl cleavage) while ensuring substantial backbone cleavage by 2'-OH transesterification, as shown in Figure S7. The RNase T1 standard sample was made using 0.5 pmol of 5'-³²P-radiolabeled RNA substrate and 1 U of RNase T1 (Thermo Scientific) in 5 μ L total volume containing 50 mM Tris, pH 8.0 and 1 mM EDTA, incubated at room temperature for 5 min, and quenched with 10 μ L of stop solution and 1.25 μ L of 1 M MgCl₂ (to match the Mg²⁺ in the 8JB210 product assay sample). The samples were separated by 20% PAGE.

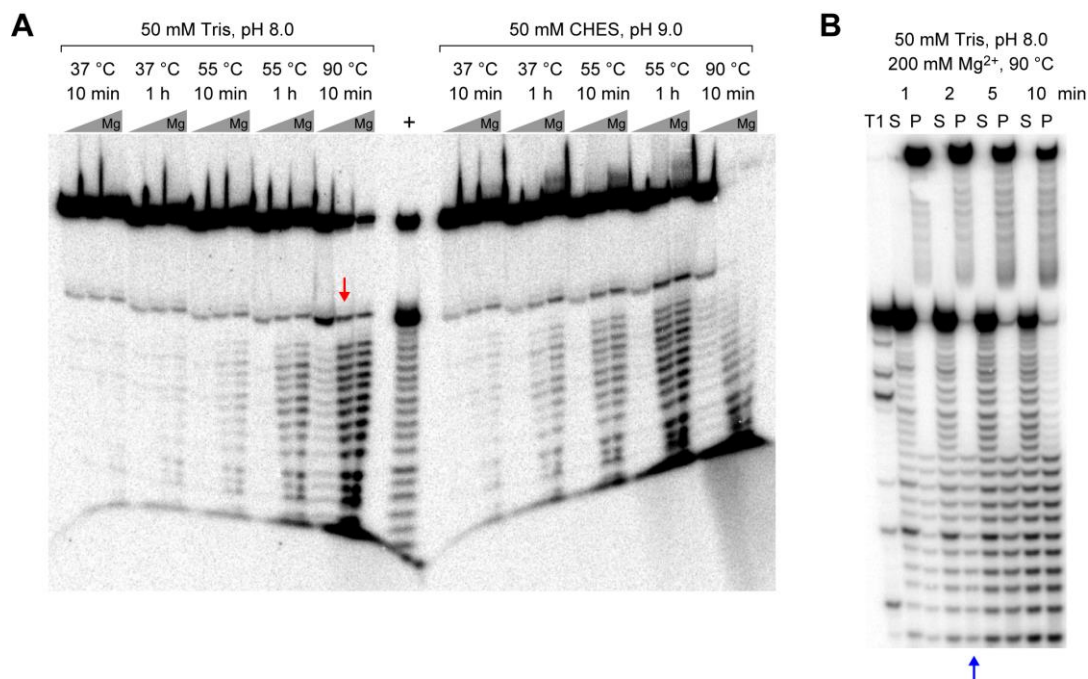


Figure S7. Optimization of base treatment conditions for the assay of Figure 7B. (A) Initial optimization experiments to screen a variety of pH values, Mg²⁺ concentrations, temperatures, and times. For each set of three lanes, the Mg²⁺ concentrations were 0, 167, and 500 mM. + = positive control (50 mM Na₂CO₃, pH 10.0, 90 °C, 10 min). The substrate for these base treatment experiments was the 8JB210 product made using the 22 nt RNA substrate with CCC 3'-overhang. A key objective to enable the Figure 7B assay was to identify incubation conditions that favor RNA phosphodiester cleavage by 2'-OH attack while disfavoring amide hydrolysis. Based on these data, we chose 50 mM Tris, pH 8.0, 200 mM Mg²⁺, 90 °C, and 10 min (red arrow) as our starting point for further optimization. The presence of high Mg²⁺ concentration in many lanes leads to unusual gel appearance at the top and bottom of the image. (B) Follow-up experiment to optimize the incubation time. The substrate for these experiments was the 8JB210 product made using the 33 nt RNA substrate with 14 nt CCCUGUGCACGCCU 3'-overhang. To suppress the noted Mg²⁺ effects on gel lanes evident in panel A, two blank samples each without nucleic acid but with buffer and 200 mM Mg²⁺ were loaded in lanes on either side of the gel, both here and in Figure 7B itself. Based on these data, we chose 50 mM Tris, pH 8.0, 200 mM Mg²⁺, 90 °C, and 2 min (blue arrow) as the conditions for the Figure 7B assay. T1, S, and P are marked as in Figure 7B. The older sample of RNase T1 used in this experiment led to some stray digestion bands that were absent in the Figure 7B experiment that used fresh RNase T1.

For each assay as in Figure 7C and Figure S8, we 5'-³²P-radiolabeled each member of a set of five DNA substrates, each with a single ribonucleotide at an internal position (U-1, C+1, C+2, C+3, or U+4) of the CCCT₆ 3'-overhang DNA substrate. We followed the common procedure for making each DNAzyme product, using 3 pmol of radiolabeled substrate, 10 pmol of DNAzyme, and 20 pmol of 5'-TFPE acyl donor oligonucleotide, with annealing volume of 8 μL, activation volume of 6 μL, and final incubation volume of 20 μL. The sample was incubated for 24 h and quenched with 20 μL of stop solution. The DNAzyme product was separated by 12% PAGE (on 20% PAGE, the DNAzyme itself overlaps with the product), precipitated with ethanol, and dissolved in 10 μL of water. For the base treatment assay under optimized conditions, 2 μL of product and 10 pmol of “blocking oligonucleotide” [6-mer (AAC)₂₀, to prevent nonspecific binding] was brought to 10 μL total volume containing 50 mM CHES, pH 9.0, and 600 mM MgCl₂, incubated at 90 °C for 10 min, quenched with 1 μL of 3 M NaOAc, pH 5.2, and 11 μL of stop solution, and separated by 20% PAGE.

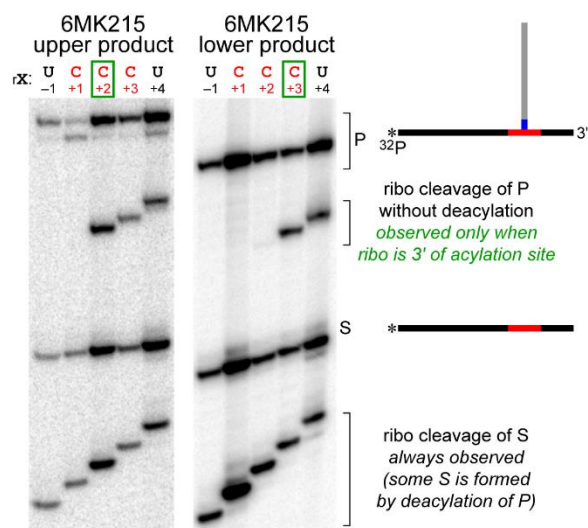


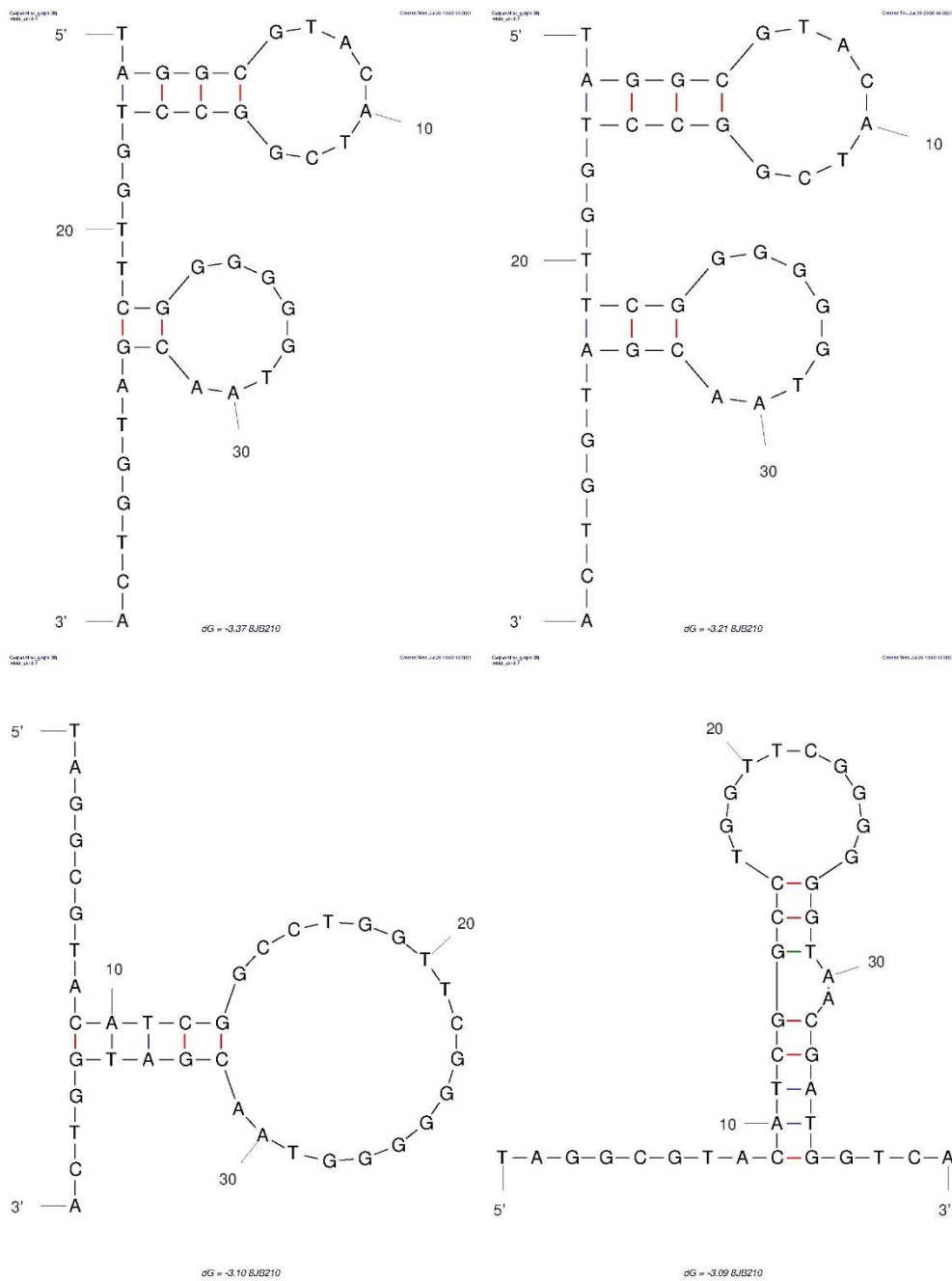
Figure S8. Assays similar to those in Figure 7C for the two products of the 6MK215 DNAzyme.

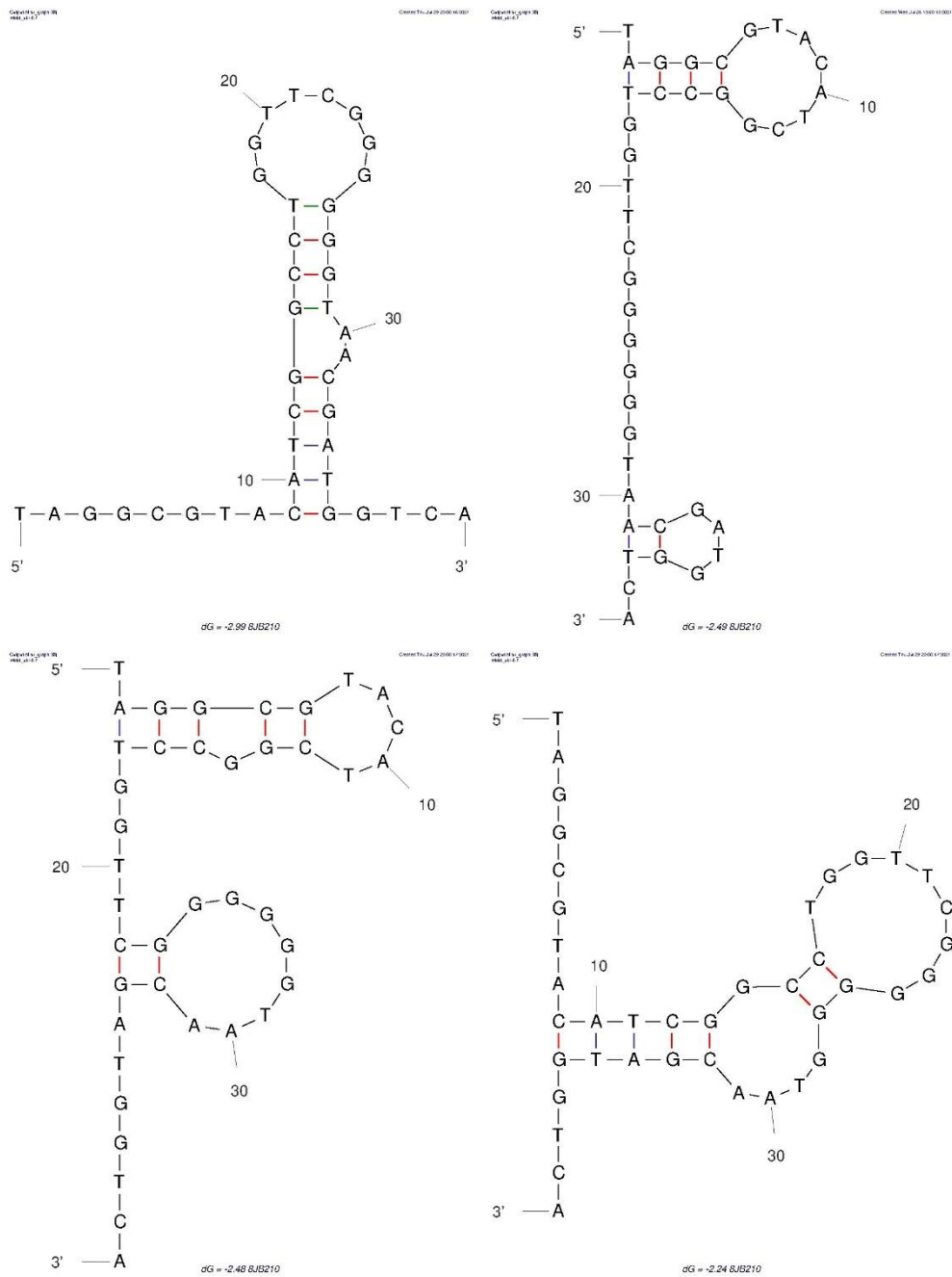
DNAzyme secondary structure predictions using mfold

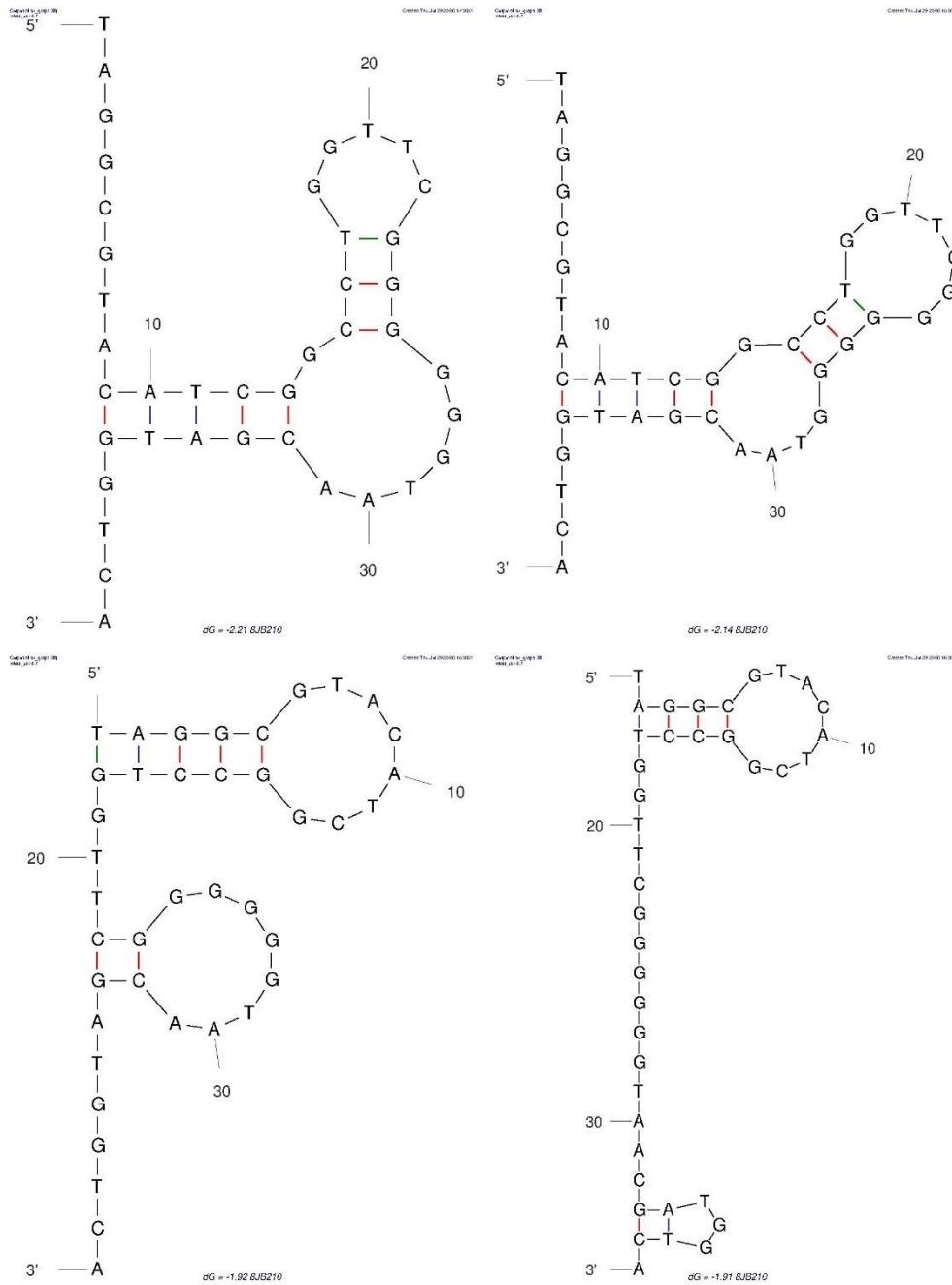
DNAzyme	number of structures	lowest ΔG , kcal/mol	DNAzyme	number of structures	lowest ΔG , kcal/mol
8JB210	13	-3.4	8MQ215	3	-6.2
8JB215	17	-8.0	8MQ216	4	-6.2
8JB223	3	-2.4	8MQ230	2	-4.3
8JB228	6	-4.2	8MQ236	3	-5.0
8JB229	10	-5.0	8MQ237	1	-7.9
6MK205	3	-3.4	8MQ241	2	-4.0
6MK208	9	-1.3	7MR202	5	-1.4
6MK215	4	-3.3	7MR205	1	-3.7
6MK232	4	-1.2	7MR206	5	-4.5
7MN217	2	-6.1	7MR218	1	-3.0
8MQ209	1	-7.2	6MT201	1	-5.0
8MQ213	1	-7.3	6MT204	5	-0.6
8MQ214	1	-5.4	6MT241	3	-2.0

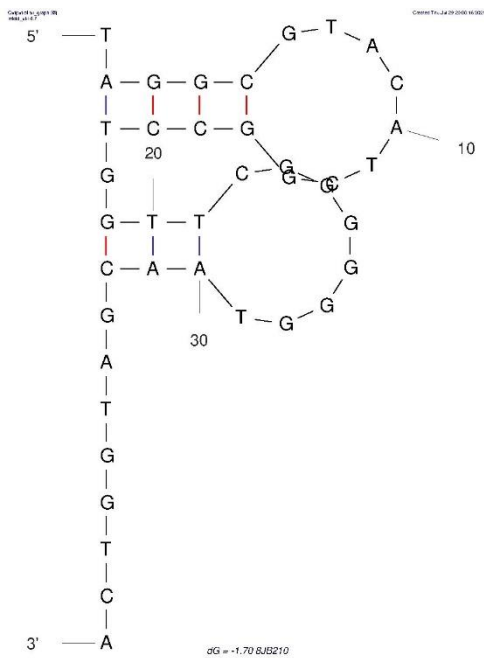
Table S4. Summary of mfold-predicted^[3] secondary structures of all 26 new DNAzymes reported in this study. The default settings were used for the sequences of the initially random N₄₀ regions with the DNA Folding Form at <http://unafold.rna.albany.edu/?q=mfold>, adjusted to 150 mM Na⁺ and 40 mM Mg²⁺. The boldface entries, corresponding to seven of the DNAzymes of Figure 6, are those for which all predicted secondary structures are shown in Figure S9.

(A) 8JB210

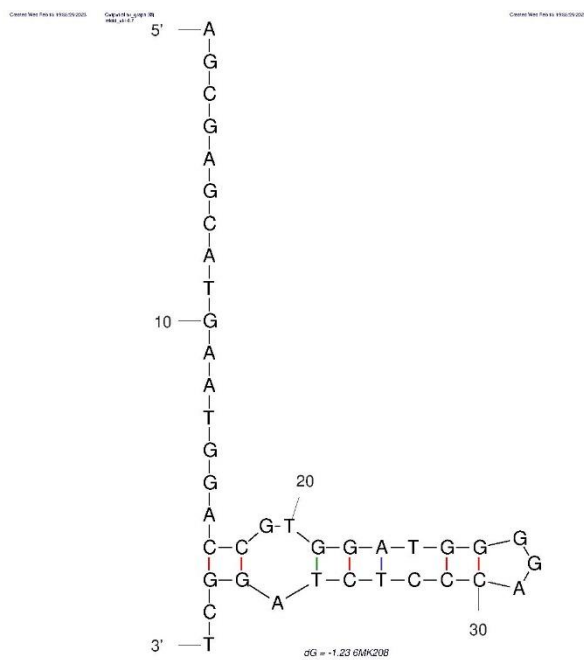
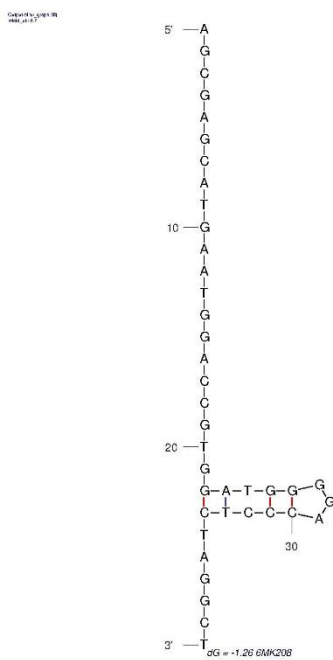




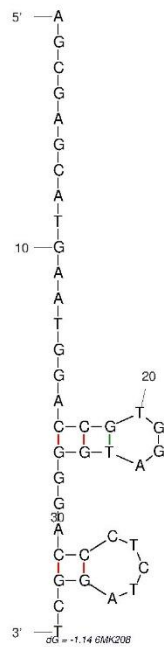




(B) 6MK208



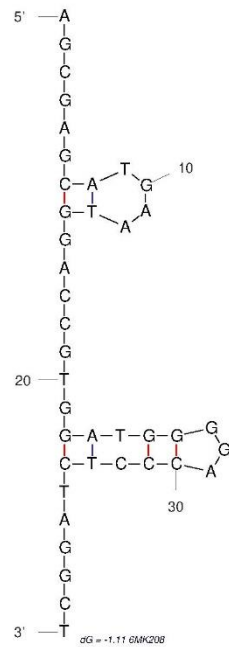
Supporting Figure S1
15/09/2019



$\Delta G = -1.14 \text{ GMK208}$

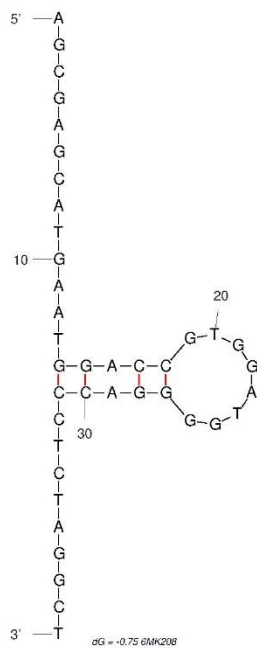
Supporting Figure S2
15/09/2019

Supporting Figure S3
15/09/2019



$\Delta G = -1.11 \text{ GMK208}$

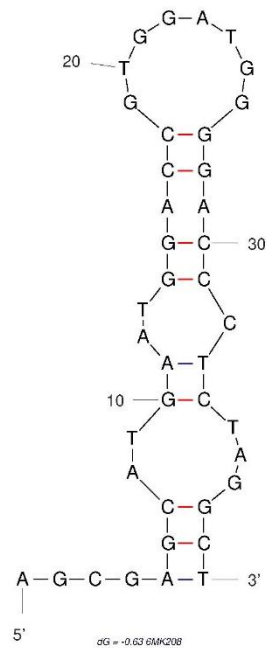
Supporting Figure S4
15/09/2019



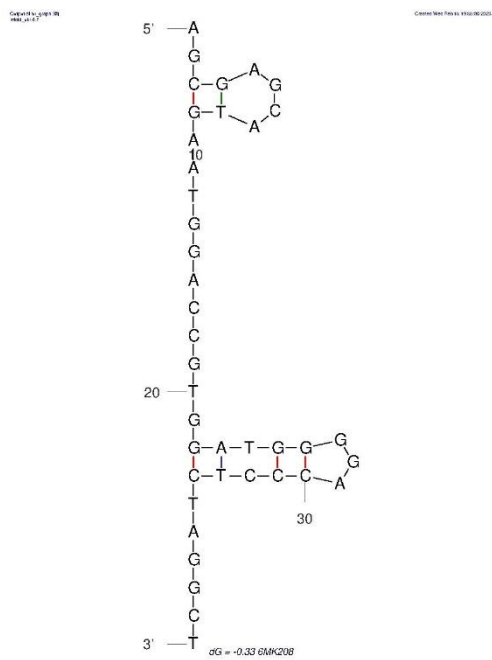
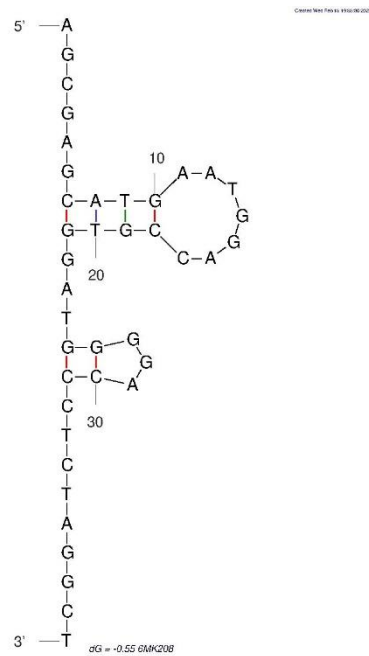
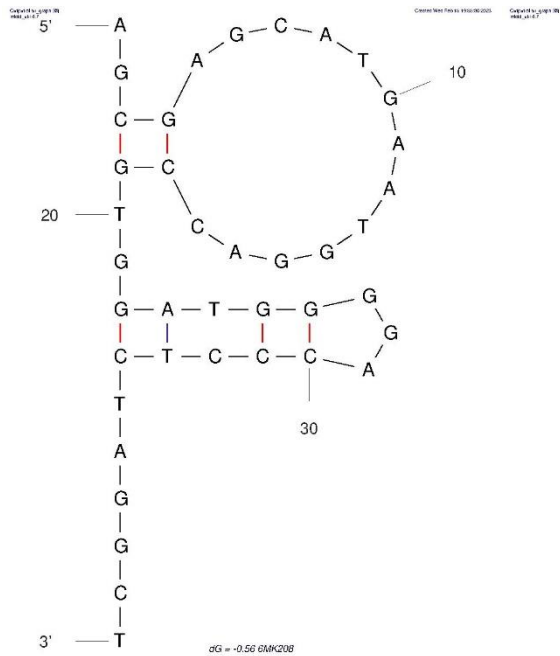
$\Delta G = -0.75 \text{ GMK208}$

Supporting Figure S5
15/09/2019

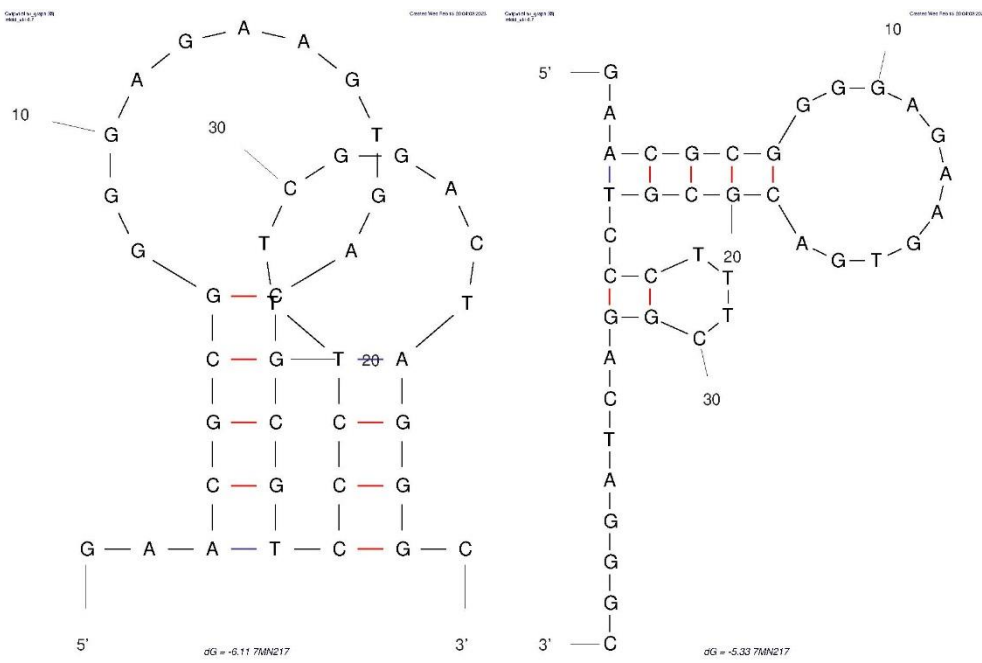
Supporting Figure S6
15/09/2019



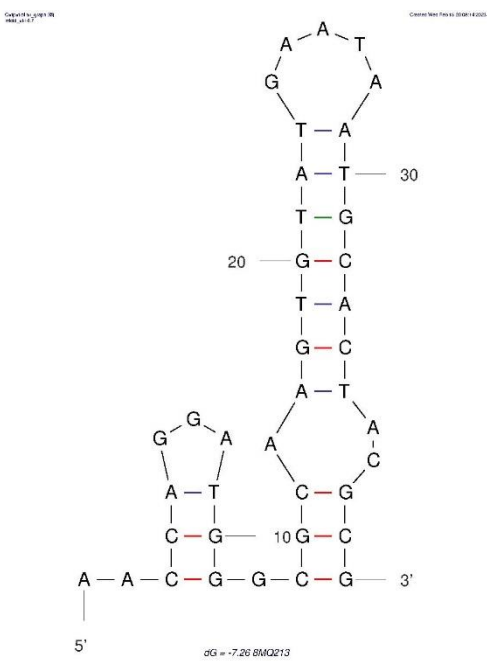
$\Delta G = -0.63 \text{ GMK208}$



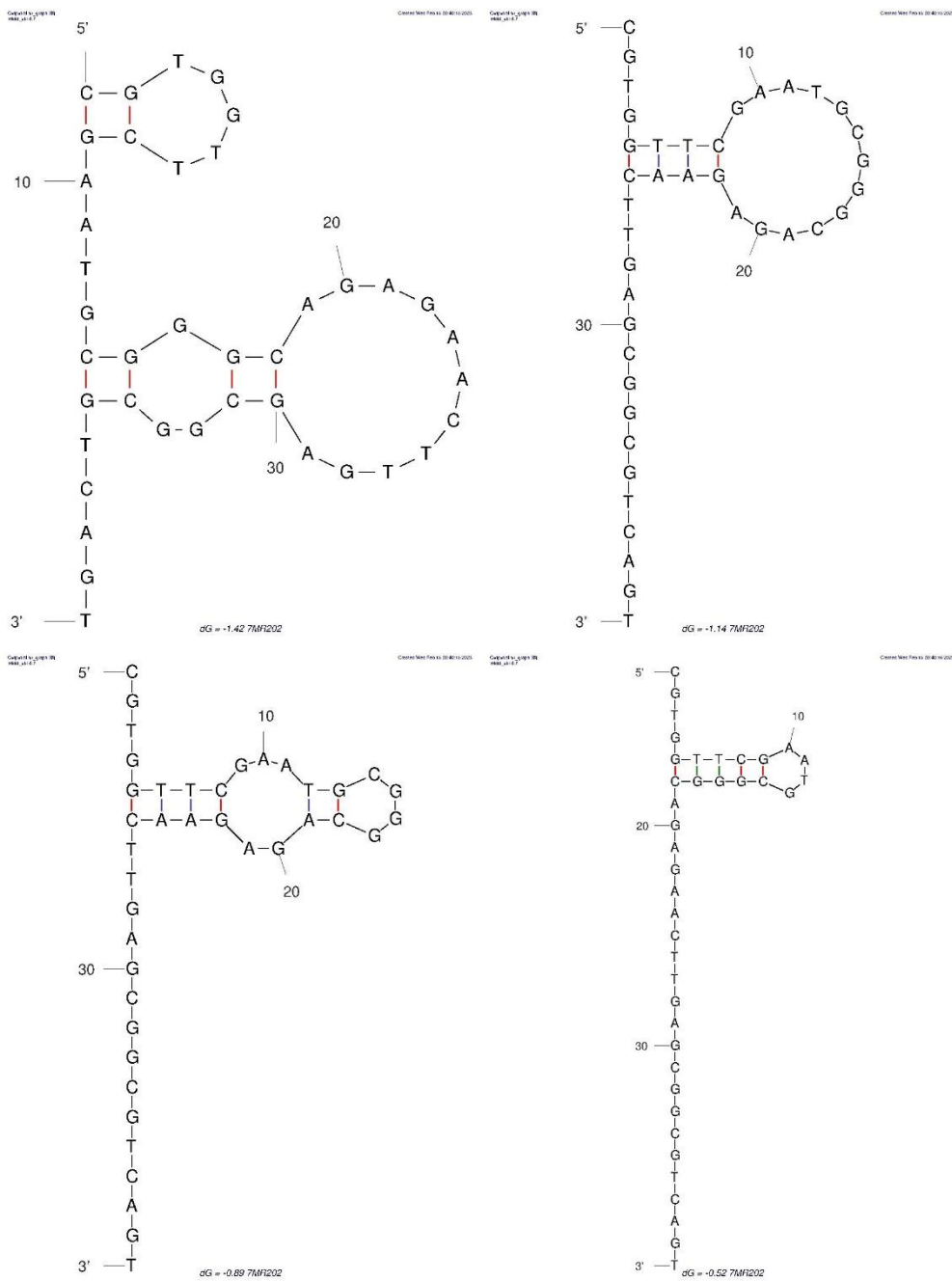
(C) 7MN217

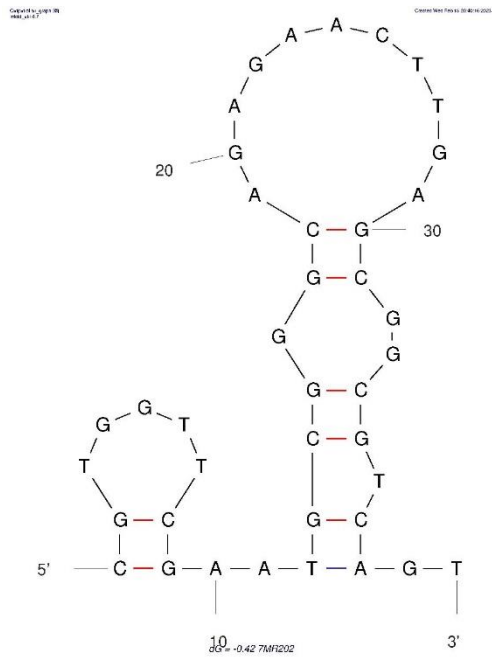


(D) 8MQ213

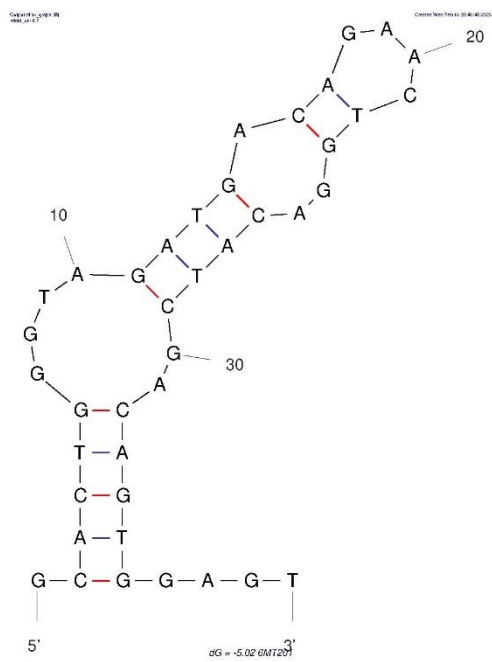


(E) 7MR202





(F) 6MT201



(G) 6MT241

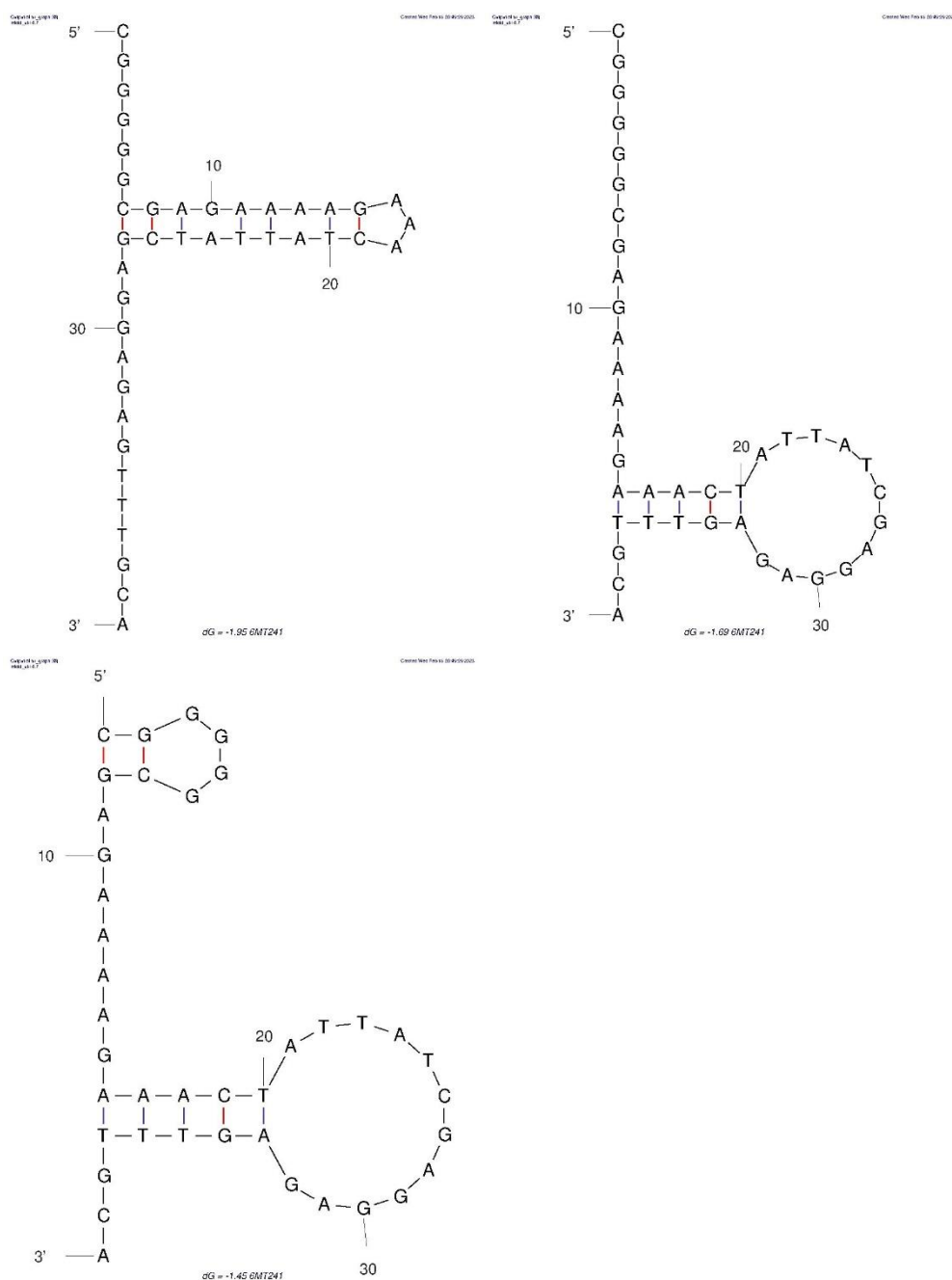


Figure S9. Mfold-predicted^[3] secondary structures for representative DNAzymes reported in this study. The default settings were used for the initially random N_{40} regions with the DNA Folding Form at <http://unafold.rna.albany.edu/?q=mfold>, adjusted to 150 mM Na^+ and 40 mM Mg^{2+} . Shown are structures for the seven of the DNAzymes of Figure 6. Where multiple structures are shown for an individual DNAzyme, the lowest-energy structure (with most negative ΔG value) is shown first, followed by the remaining structure(s) in order of increasing energy. See Table S4 for full tabulation of number of structures and lowest ΔG value for all 26 DNAzymes identified in this study.

References for Supporting Information

- [1] C. Zhou, J. L. Avins, P. C. Klauser, B. M. Brandsen, Y. Lee, S. K. Silverman, *J. Am. Chem. Soc.* **2016**, *138*, 2106-2109.
- [2] U. J. Desai, P. K. Pfaffle, *BioTechniques* **1995**, *19*, 780-784.
- [3] M. Zuker, *Nucleic Acids Res.* **2003**, *31*, 3406-3415.