ChemBioChem

Supporting Information

Sequence-Dependent Acylation of Peptide Lysine Residues by DNAzymes

Prakriti K. Das and Scott K. Silverman*

Sequence-Dependent Acylation of Peptide Lysine Residues by DNAzymes

Prakriti K. Das and Scott K. Silverman*

Department of Chemistry, University of Illinois Urbana-Champaign,
600 South Mathews Avenue, Urbana, Illiniois 61801 (USA)

Table of Contents

Uncatalyzed background reactions .... page S3

Nucleotide details of in vitro selection and DNAzyme assays .... page S4

Procedures for in vitro selection and cloning .... page S4

Yield data for DNAzymes .... page S7

Metal ion dependence of DNAzymes .... page S8

Mass spectrometry of EDC adducts and DNAzyme products .... page S9

Peptide concentration dependence of DNAzymes .... page S9

Evaluating DNAzymes with the tethered peptide substrate .... page S10

DNAzyme secondary structure predictions using mfold .... page S11

Assays of DNAzymes with mutated peptide substrates .... page S25

References for Supporting Information .... page S25

Uncatalyzed background reactions

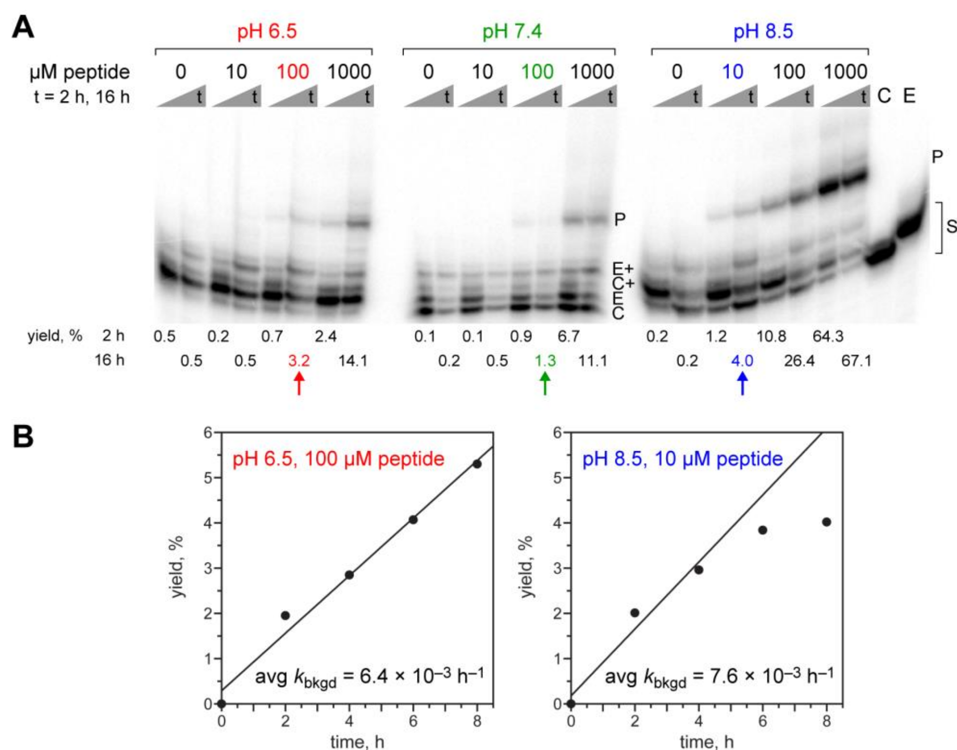


Figure S1. Assays of the uncatalyzed background reactions with the N_3-ASQSWK hexapeptide. All background assays used the N_{40} pool in place of an individual DNAzyme. (A) Assays at 2 and 16 h for each combination of pH (6.5, 7.4, and 8.5) and peptide concentration (0, 10, 100, and 1000 µM), to determine suitable incubation times for the enrichment step of selection. The observed yield for each combination is shown below each lane. The colored arrows and labels mark the chosen conditions, in which the uncatalyzed background yield is <5%. S = acyl donor substrate, P = background product. E (ester) = 5′-TFPE oligonucleotide acyl donor, C (carboxyl) = 5′-CO_2H oligonucleotide acyl donor. + denotes nonspecific EDC adduct on a nucleobase of E or C. (B) Single-turnover assays using initial-rate kinetics to quantitatively determine the k_{bkgd} values, which are as follows (mean ± sd). pH 6.5, 100 µM peptide (6.4 ± 2.8) × 10^{-3} h^{-1} (n = 6); pH 8.5, 10 µM peptide (7.6 ± 0.7) × 10^{-3} h^{-1} (n = 3).

Nucleotide details of in vitro selection and DNAzyme assays

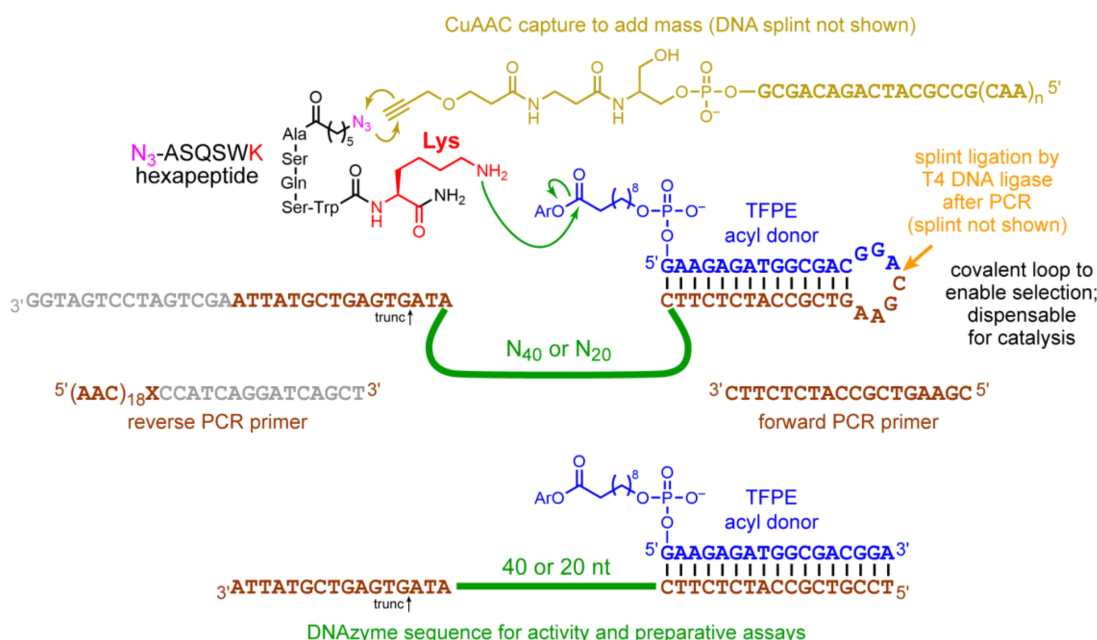


Figure S2. Nucleotide details of the in vitro selection experiments and DNAzyme assays with 5′-TFPE oligonucleotide as acyl donor (Ar = 2,3,5,6-tetrafluorophenyl). In the reverse primer, X = hexa(ethylene glycol), HEG, spacer to stop Taq polymerase. The splint for ligation of the TFPE acyl donor oligonucleotide to the pool was 5′-AGAG-ATGGCGACTTCGTCCGTCGCCATCTCTTC-3′. The splint for the capture reaction was 5′-TTCGTCCGTCGCCATCTC-TTCACGCTGTCTGATGCGGC-3′. The short (n=0; 16 nt) version of the 3′-alkyne capture oligonucleotide was used in even-numbered selection rounds, and the long (n=11; 49 nt) version in odd-numbered rounds, to avoid enriching noncatalytic DNA sequences that have abnormal PAGE migration positions. The substrate for forming the product migration standard by CuAAC during each selection round was the uncatalyzed background product between the peptide and the N_{40} or N_{20} pool (pH 8.5, 1 mM peptide, 37 °C, 16 h; ~60% yield under these conditions, which use higher peptide concentration than in any selection). The black arrow marks the site of 3′-truncation for each new DNAzyme. For assays of individual DNAzymes prepared by solid-phase synthesis, the DNAzyme was changed to begin with 5′-TCC rather than 5′-CGAA, to allow full base pairing with the 3′-terminus of the TFPE oligonucleotide. Before activation as the 5′-aryl ester, the 5′-CO_2H oligonucleotide was internally radiolabeled by splint ligation (T4 DNA ligase, splint 5′-AACAACAACAACAACAACCCGCGTGCCTCCGTCGCCATCTCTTC-3′) to the 9-mer 5′-GGCAC-GCGG-3′ that was 5′-^{32}P-radiolabeled using γ-^{32}P-ATP and T4 polynucleotide kinase.

Procedures for in vitro selection and cloning

Procedure for phosphorylation and ligation in round 1. A 10 µL sample containing 1 nmol of N_{40} or N_{20} DNA pool, 1 mM ATP, 1× T4 PNK buffer A, and 10 units of T4 polynucleotide kinase (Thermo Fisher) was incubated at 37 °C for 2.5 h. After phenol/chloroform extraction and ethanol precipitation, the sample was redissolved in 20 µL of 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA containing 1.1 nmol of DNA splint and 1.2 nmol of 5′-CO_2H DNA substrate and annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. The ligation reaction was initiated by bringing the sample to a final volume of 40 µL containing 1× T4 DNA ligase buffer and 5 units of T4 DNA ligase (Thermo Fisher). The sample was incubated at 37 °C for 16 h and purified by 8% PAGE.

Procedure for phosphorylation and ligation subsequent rounds. A 17 µL sample containing the PCR-amplified N_{40} or N_{20} DNA pool (~5–10 pmol), 30 pmol of DNA splint, and 60 pmol of 5′-CO_2H DNA 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. The ligation reaction was initiated by bringing the sample to a final volume of 20 µL containing 1× T4 DNA ligase buffer and 1 unit of T4 DNA ligase. The sample was incubated at 37 °C for 16 h and purified by 8% PAGE.

Procedure for aryl ester activation of 5′-CO_2H in round 1. Immediately prior to performing the aryl ester activation and selection steps, a 10 µL sample containing 200 pmol of the 5′-CO_2H ligation product in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA was annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. The 5′-CO_2H ligation product was activated in a total volume of 16 µL containing 75 mM 1-ethyl 3-(3-dimethylaminopropyl)carbodiimide (EDC), added from a freshly prepared 600 mM stock solution in water, and 75 mM 2,3,5,6-tetrafluorophenol (TFP), added from a 600 mM stock solution in DMF, at room temperature for 2 h.

Procedure for aryl ester activation of 5′-CO_2H in subsequent rounds. An 8 µL sample containing ~5–10 pmol of 5′-CO_2H ligation product in 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA was annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. The 5′-CO_2H ligation product was activated in a total volume of 12 µL containing 50 mM EDC and 50 mM TFP at room temperature for 2 h.

Procedure for enrichment step in round 1. Each selection experiment was initiated by bringing a 16 µL sample containing 200 pmol of annealed 5′-aryl ester activated ligation product (containing 75 mM each EDC and TFP) to 40 µL total volume with one of three sets of contents: (a) 100 µM peptide N_3-ASQSWK (added from a 1 mM stock in water), 105 mM HEPES, pH 7.5, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$ and 150 mM NaCl, with 30 mM each EDC/TFP and final pH of 6.5; (b) 100 µM peptide N_3-ASQSWK (added from a 1 mM stock in water), 220 mM HEPES, pH 8.0, 40 mM $MgCl_2$, 20 mM $MnCl_2$, 1 mM $ZnCl_2$ and 150 mM NaCl, with 30 mM each EDC/TFP and final pH of 7.4; or (c) 10 µM peptide N_3-ASQSWK (added from a 100 µM stock in water), 240 mM CHES, pH 9.0, 40 mM $MgCl_2$, and 150 mM NaCl, with 30 mM each EDC/TFP and final pH of 8.5. The Mg^{2+} and Mn^{2+} were added from 10× stock solutions containing 400 mM $MgCl_2$ and 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. For samples at pH 6.5 and 7.4, 20 mM of the HEPES was from the $ZnCl_2$ stock solution. The metal ion stocks and then the peptide were added last to the final sample. For samples at all three pH values, the required amount of buffer to make 85 mM HEPES, 200 mM HEPES, or 240 mM CHES, respectively, was first evaporated to dryness on a SpeedVac and then all remaining steps were followed, because otherwise the final total volume would exceed 40 µL (all analogous to what we previously reported[1]). The sample was incubated at 37 °C for 16 h, followed by phenol/chloroform extraction and ethanol precipitation.

Procedure for enrichment step in subsequent rounds. A 12 µL sample containing ~5–10 pmol of annealed 5′-aryl ester activated ligation product (containing 50 mM each EDC and TFP) was brought to 20 µL total volume with one of three sets of contents (a), (b), or (c) as described for round 1, using the same stocks, order of addition, and buffer evaporation. The sample was incubated at 37 °C for 16 h, followed by ethanol precipitation (no phenol/chloroform extraction).

After the selection process was completed, we simplified the method by which samples were assembled for the single-turnover DNAzyme assay procedure using individual DNAzymes (see main text Experimental Section), relative to the analogous enrichment step of selection as detailed above. Instead of evaporating the required amount of buffer to dryness, which was necessary for the pH 8.5 samples due to the relatively low aqueous solubility of CHES and was also performed for the pH 6.5 and 7.4 samples to maintain parallel processing, we replaced CHES with a different buffer compound. By using 1.2 M AMPSO, pH 9.0, as the 5× stock solution in place of 0.5 mM CHES, pH 9.0, the pH 8.5 assay samples could be assembled without any buffer evaporation.

Procedure for CuAAC capture step in round 1. A 17.5 µL sample containing the precipitated round 1 enrichment product (~200 pmol), 250 pmol of DNA splint for capture, and 300 pmol of 49 nt 3′-alkyne capture oligonucleotide in 50 mM HEPES, pH 7.5, and 150 mM NaCl was annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. The sample was brought to 25 µL total volume containing 2.8 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 0.8 mM sodium ascorbate, and 0.4 mM $CuSO_4$, respectively added from 28 mM (10×), freshly prepared 8 mM (10×), and 4 mM (10×) stock solutions in water; the three reagents were mixed together and then added at once to the other components. The sample was incubated at 37 °C for 1 h and separated by 8% PAGE.

Procedure for CuAAC capture step in subsequent rounds. A 17.5 µL sample containing the precipitated enrichment step product (~5–10 pmol), 50 pmol of DNA splint for capture, and 100 pmol of 16 nt (even round numbers) or 49 nt (odd round numbers) 3′-alkyne capture oligonucleotide was treated via the same procedure as for round 1 capture.

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 $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, and 0.1% Triton X-100], and 0.5 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 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), 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 DNAzymes. The PCR primers used for cloning were 5′-CGAAGTCG-CCATCTCTTC-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 10^3-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/µL. A 4 µL 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 main text section "Single-turnover DNAzyme assay procedure".

Yield data for DNAzymes

DNAzyme	0.5 h yield, %	2 h yield, %	16 h yield, %	48 h yield, %
12LK206	5.2	12	51	68
12LK208	21	36	65	72
12LK210	24	47	70	73
12LK213	6.0	19	52	63
12LK228	8.3	19	60	73
11LM201	11	17	24	23
11LM224	7.0	12	19	18
13LN204	19	35	73	80
13LN205	5.7	15	55	73
13LN220	6.4	16	58	71
13LN238	9.2	16	52	59

Table S1. Initial yield data for peptide acylation by DNAzymes. The kinetic data of Figure 4A are from a different experiment.

Metal ion dependence of DNAzymes

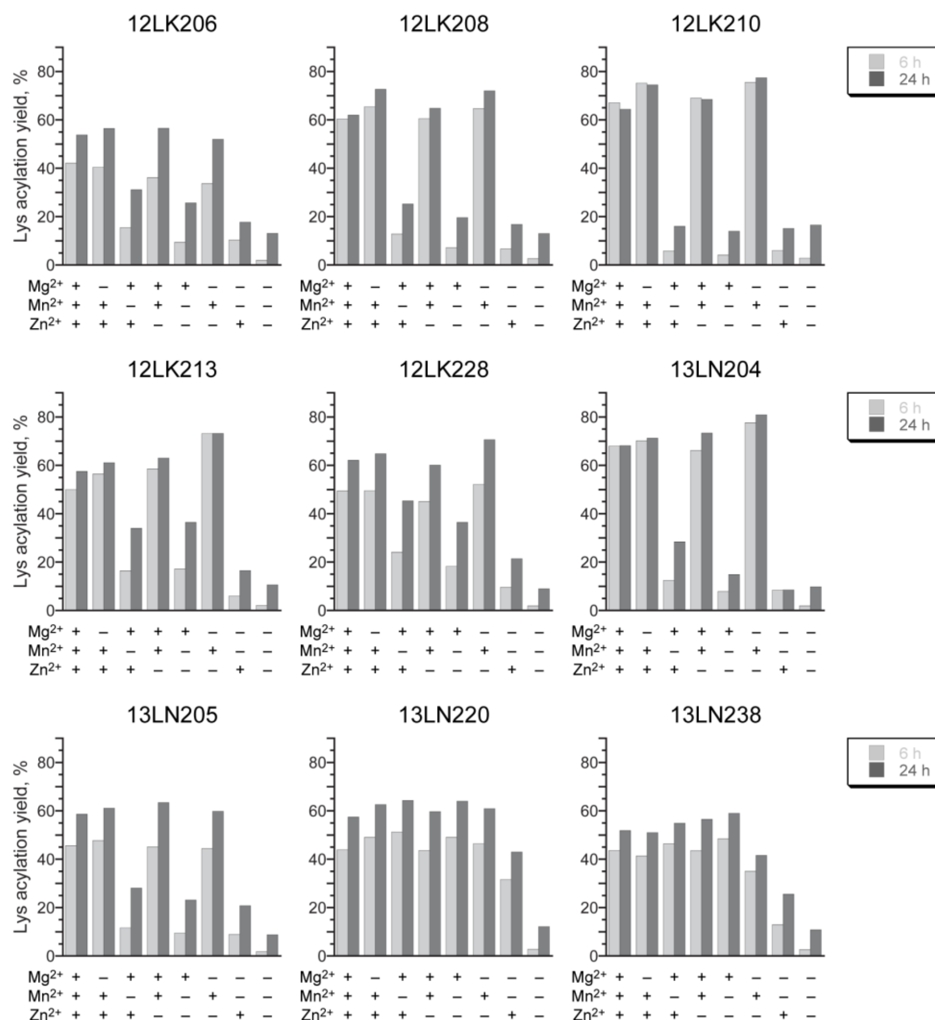


Figure S3. Metal ion requirements of the nine pH 6.5 DNAzymes. Each bar plot shows the DNAzyme yield at 6 and 24 h with the indicated divalent metal ion or ions. Mn^{2+} alone is sufficient for full catalytic activity for all nine DNAzymes, but some of the DNAzymes (e.g., 13LN220 and 13LN238) clearly do not require Mn^{2+}.

Mass spectrometry of EDC adducts and DNAzyme products

EDC adduct or DNAzyme product	*m/z* calcd.	*m/z* found	Δ
EDC adduct of substrate	5740.0	5742.3	+0.02%
EDC adduct of product	6565.9	6562.3	–0.04%
12LK210	6410.7	6411.8	+0.01%
12LK213	6410.7	6411.8	+0.01%
13LN204	6410.7	6413.4	+0.03%
11LM201	6410.7	6413.7	+0.03%

Table S2. MALDI-TOF mass spectrometry data of EDC adducts and DNAzyme products. All *m/z* values are for $[M+H]^+$. EDC adducts of substrate and product are shown in Figure 4A.

Peptide concentration dependence of DNAzymes

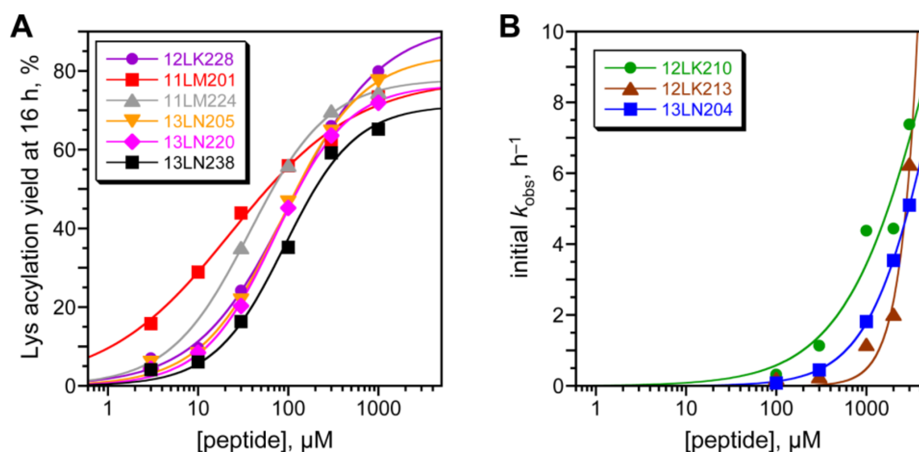


Figure S4. Peptide concentration dependence of DNAzymes. (A) Determining the peptide concentration at which 50% of the maximal yield was observed (C_{50} value). See Figure 4B for the data for the other four of 11 DNAzymes, and for all C_{50} values. (B) Determining the peptide K_m by initial-rate kinetics for three of the DNAzymes (K_m >1 mM).

Evaluating DNAzymes with the tethered peptide substrate

The DNA-tethered peptide was prepared as follows. A 2 nmol sample of 3′-alkyne oligonucleotide was brought to 40 µL total volume containing 100 mM HEPES, pH 7.5, 1 mM N_3-ASQSWK peptide, 40 mM tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), 10 mM sodium ascorbate, and 5 mM $CuSO_4$. The four latter reagents were added from 10 mM (10×), 1 M (25×), freshly prepared 100 mM (10×), and 100 mM (20×) stock solutions in water, respectively; the four reagents were mixed together and then added at once to the other components. The sample was incubated at 37 °C for 1 h and separated by 20% PAGE. The assays were performed using 0.5 pmol of 5′-^{32}P-radiolabeled DNA-tethered peptide, 5 pmol of DNAzyme, and 10 pmol of 5′-TFPE acyl donor oligonucleotide.

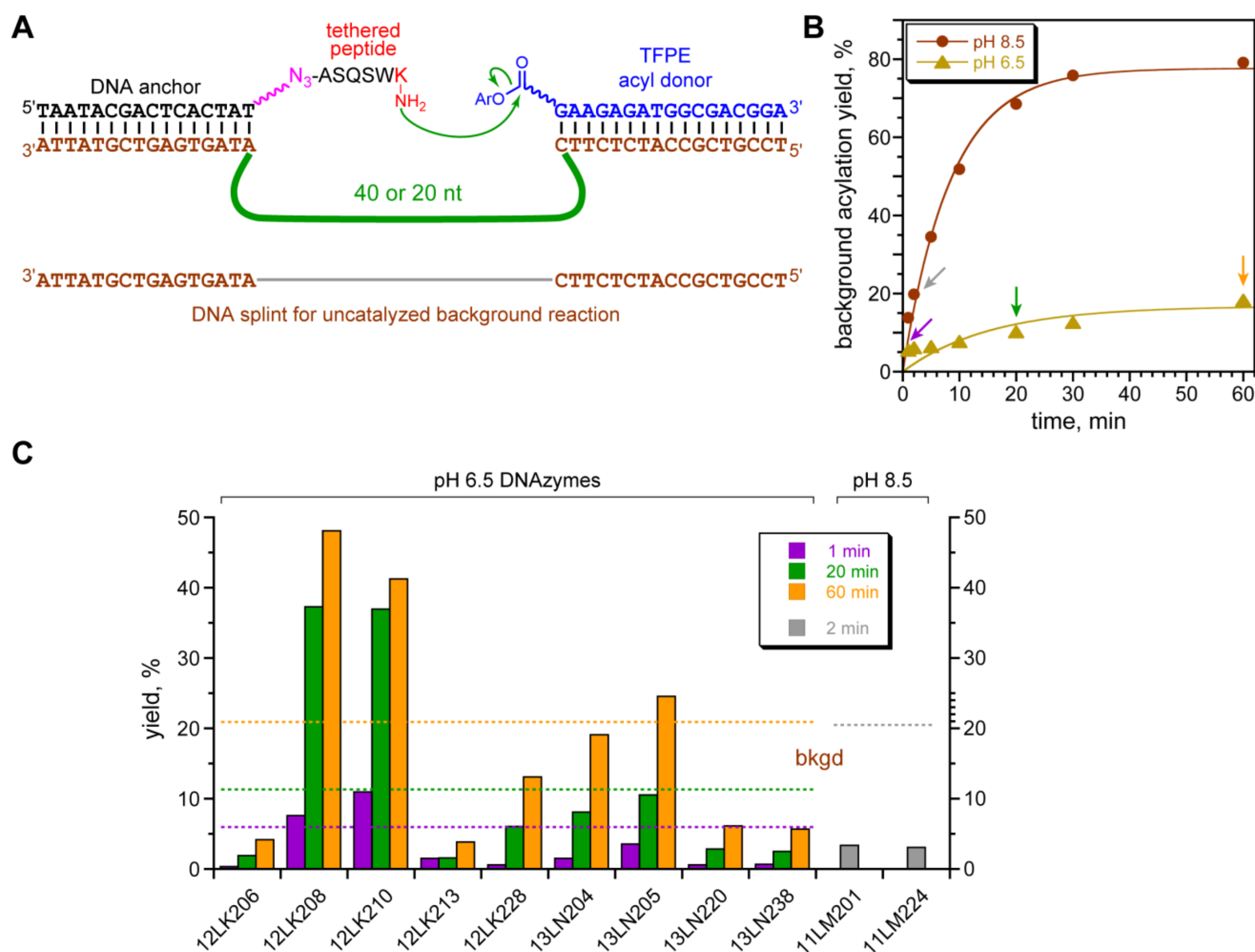


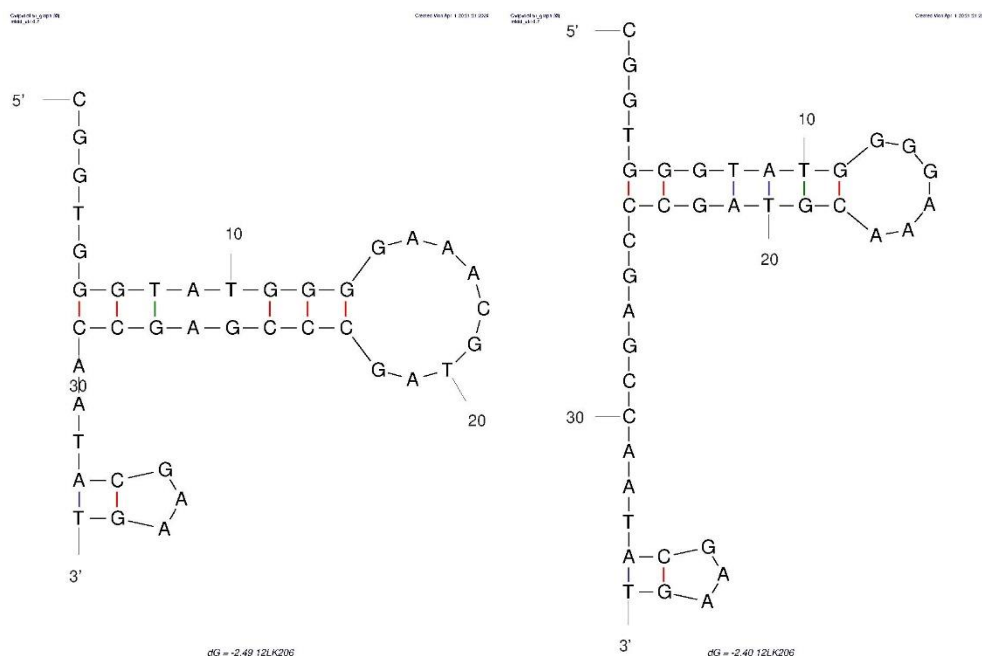
Figure S5. Evaluating DNAzymes with the tethered peptide substrate. (A) Assay arrangement. The DNA-tethered peptide was 5′-^{32}P-radiolabeled. See Figure S2 for full structures of the 3′-alkyne oligonucleotide (same 3′-alkyne modifier as was used for the capture oligonucleotide), azido-modified hexapeptide, and 5′-TFPE acyl donor oligonucleotide. At the bottom is the splint for the uncatalyzed background reaction. This splint has the two DNA binding arms directly connected, without any intervening N_{40}/N_{20} nucleotides. (B) We first assayed the DNA-splinted uncatalyzed background reaction (t = 1, 2, 5, 10, 20, 30, 60 min), to identify reaction times with sufficiently low background yield (colored arrows) that allow for the possibility of rate enhancement by a DNAzyme. (C) We then assayed all 11 DNAzymes at those reaction times, with the data plotted using the same colors as the corresponding arrows in panel B, and the background yields from panel B shown as horizontal dashed lines. Two DNAzymes, 12LK208 and 12LK10, have clear rate enhancement above the background, most evident at 20 min. The other nine DNAzymes each have yield either similar to or much lower than the background.

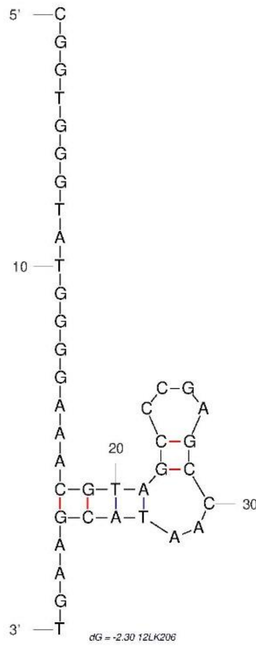
DNAzyme secondary structure predictions using mfold

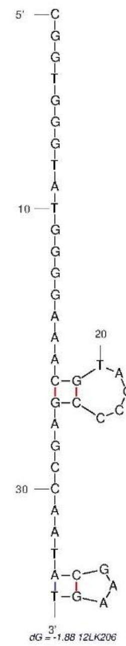
DNAzyme	number of structures	lowest ΔG, kcal/mol	DNAzyme	number of structures	lowest ΔG, kcal/mol
12LK206	8	–2.5	13LN204	7	+0.6
12LK208	2	–4.0	13LN205	1	–1.6
12LK210	4	–6.4	13LN220	1	–1.8
12LK213	1	–4.3	13LN238	5	+0.4
12LK228	6	–1.3			
11LM201	7	–4.2			
11LM224	4	–1.4			

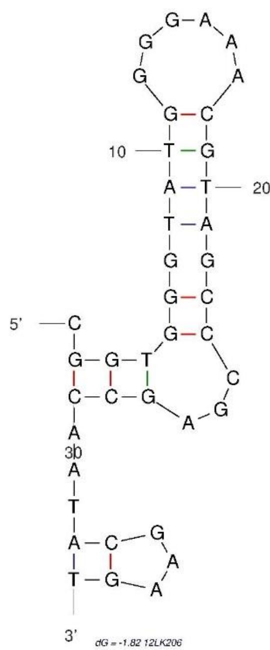
Table S3. Summary of mfold-predicted[3] secondary structures of all 11 new DNAzymes reported in this study. The default settings were used for the sequences of the initially random N_{40} regions with the DNA Folding Form at http://unafold.rna.albany.edu/?q=mfold, adjusted to 150 mM Na^{+} and 40 mM Mg^{2+}. All predicted secondary structures are shown in Figure S6.

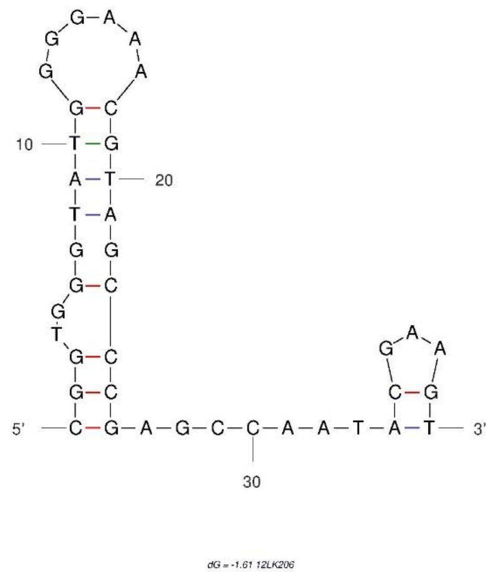
(A) 12LK206

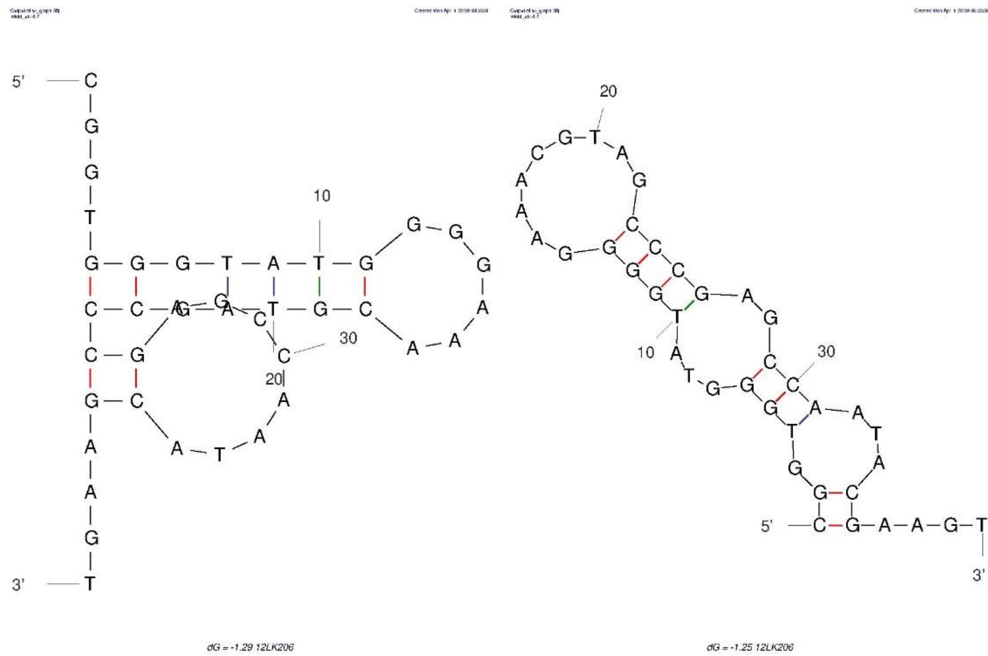



5'
10
20
30
3'
dG = -2.30 12LK206

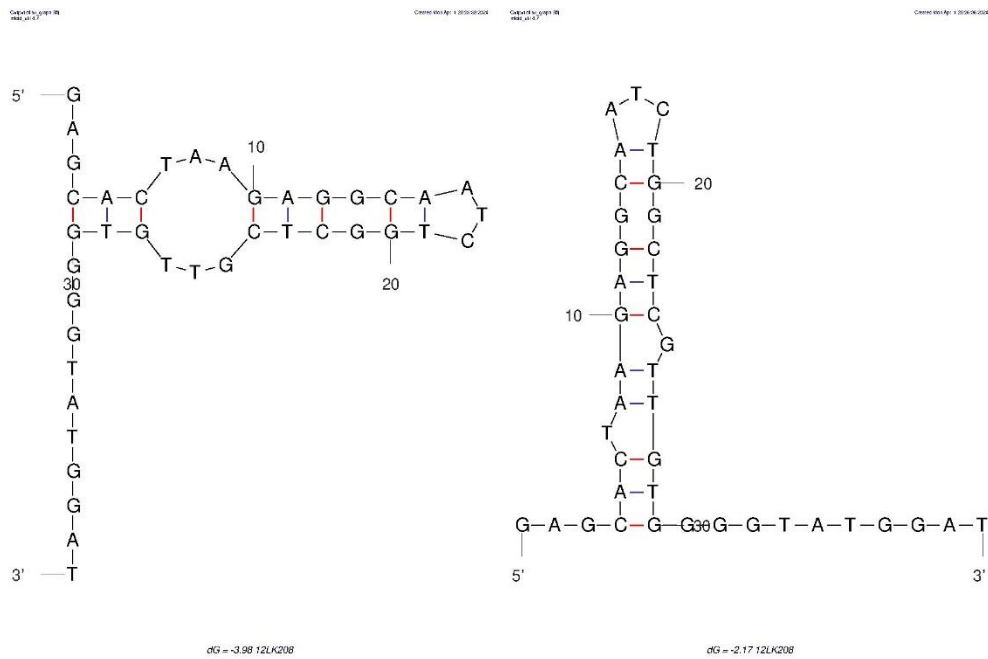

5'
10
20
30
3'
dG = -1.88 12LK206


10
20
5'
30
3'
dG = -1.82 12LK206

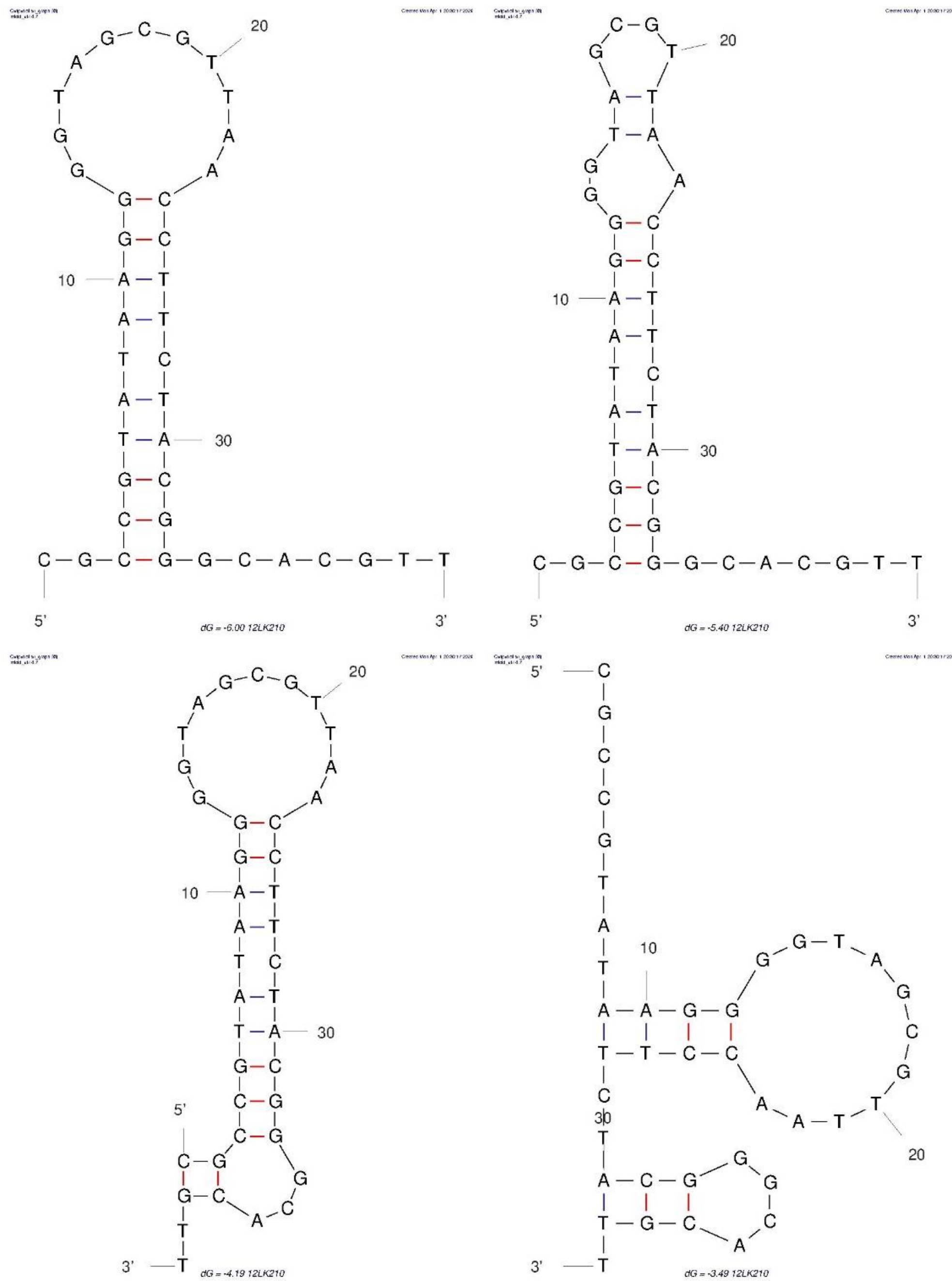

10
20
5'
3'
30
dG = -1.61 12LK206



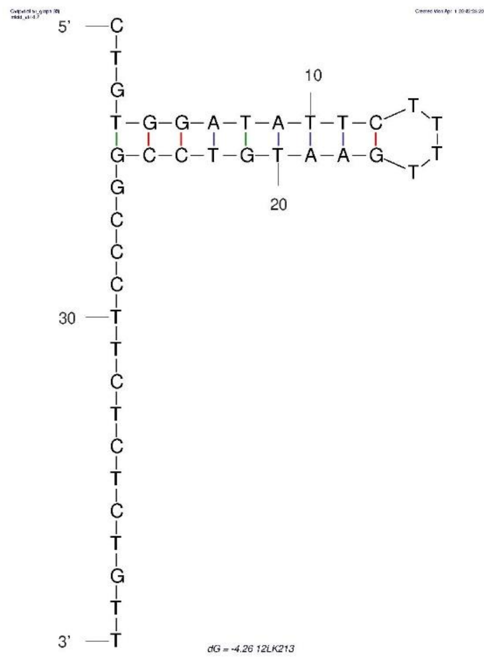
(B) 12LK208



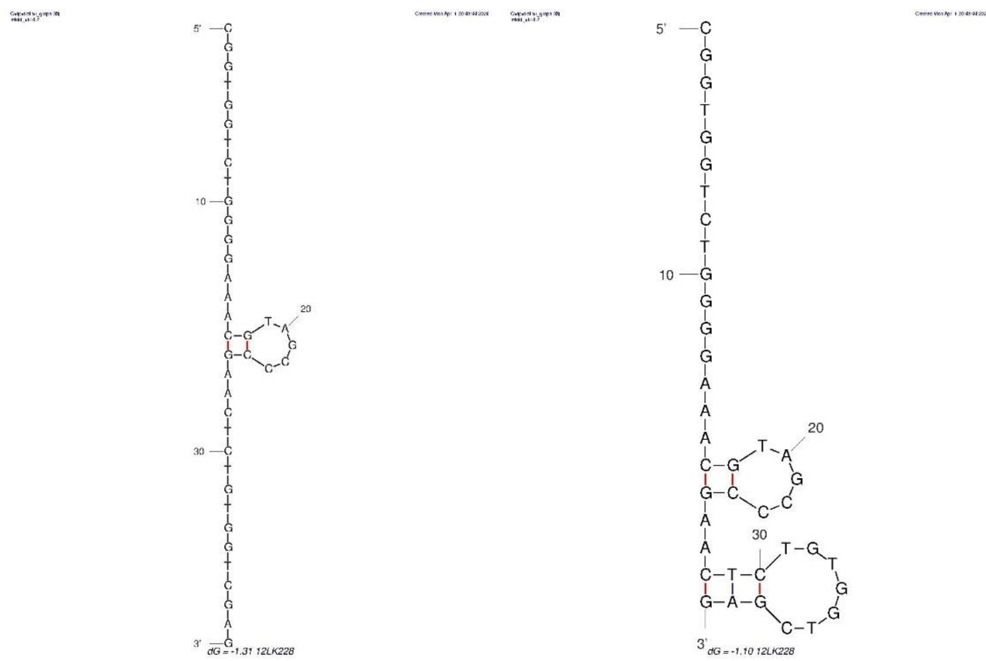
(C) 12LK210

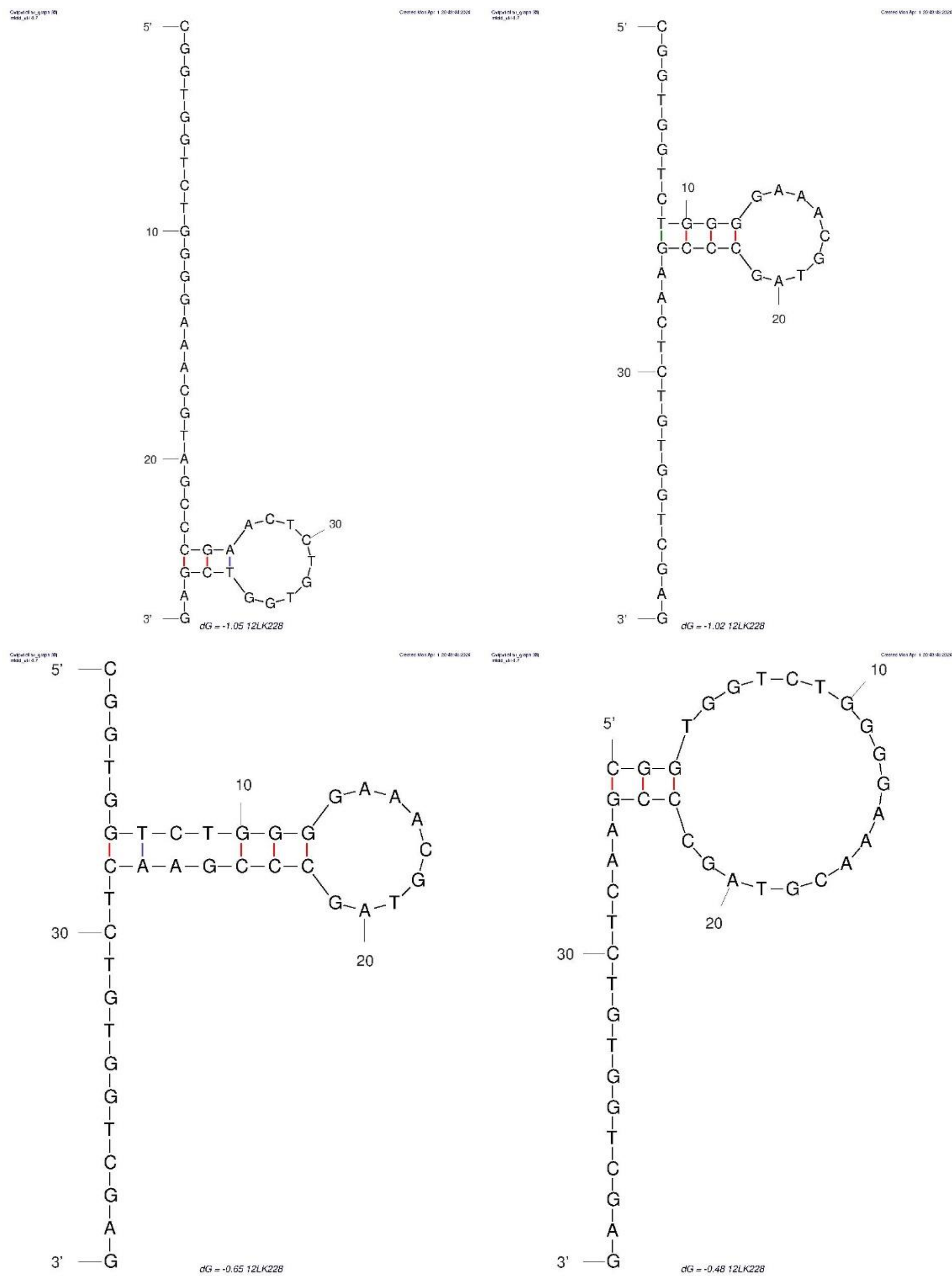


(D) 12LK213

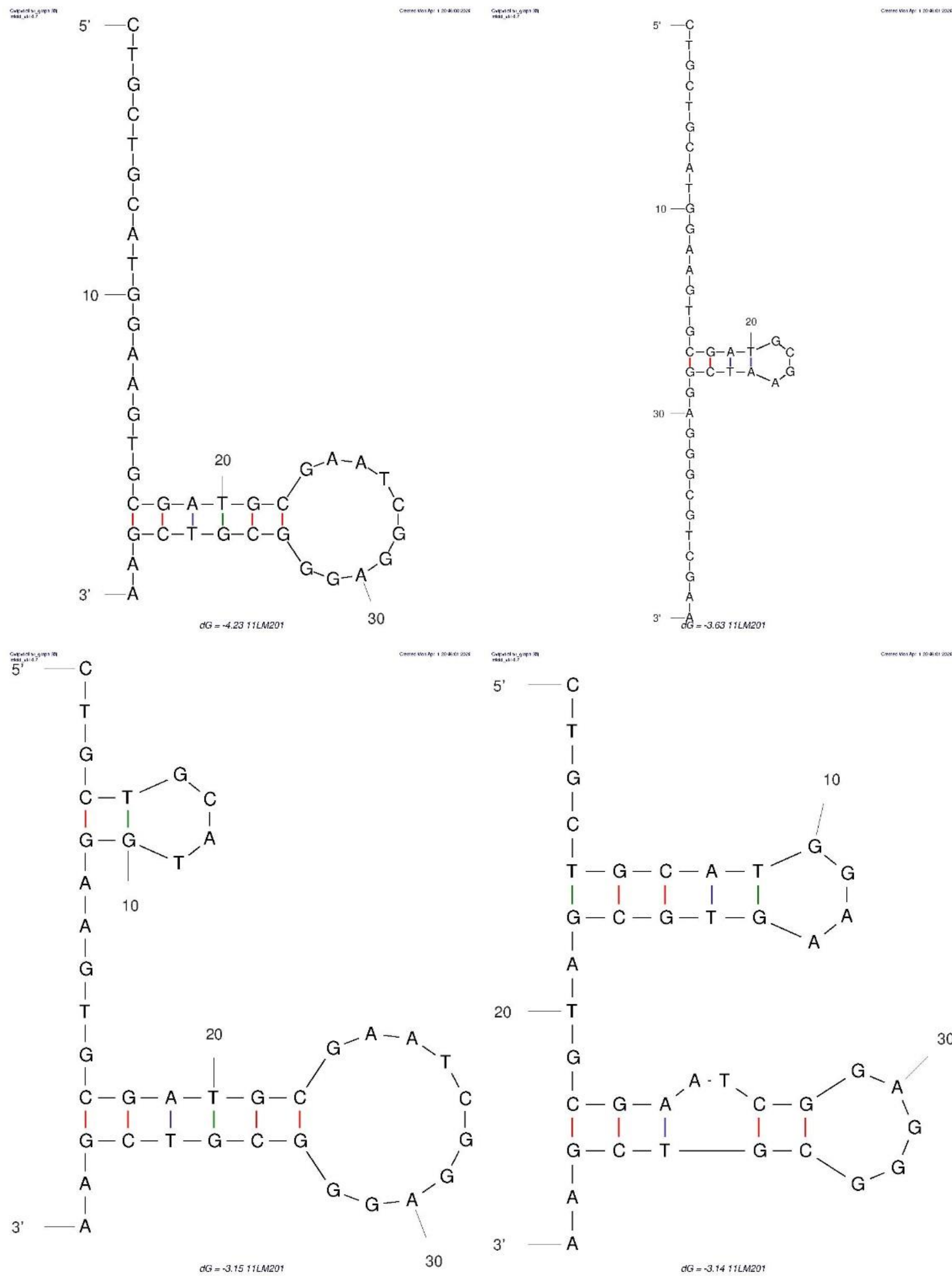


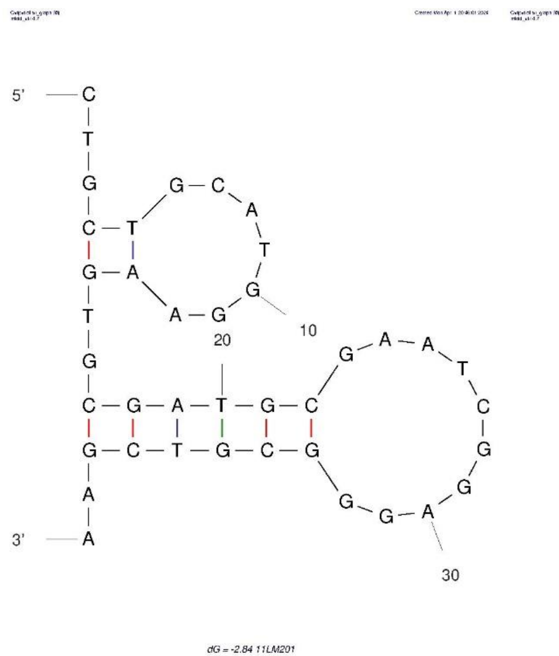
(E) 12LK228

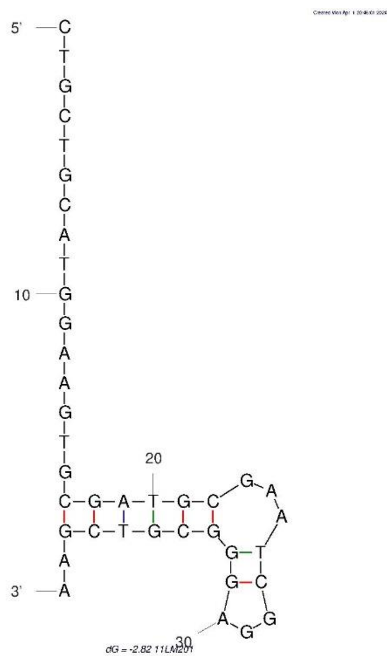


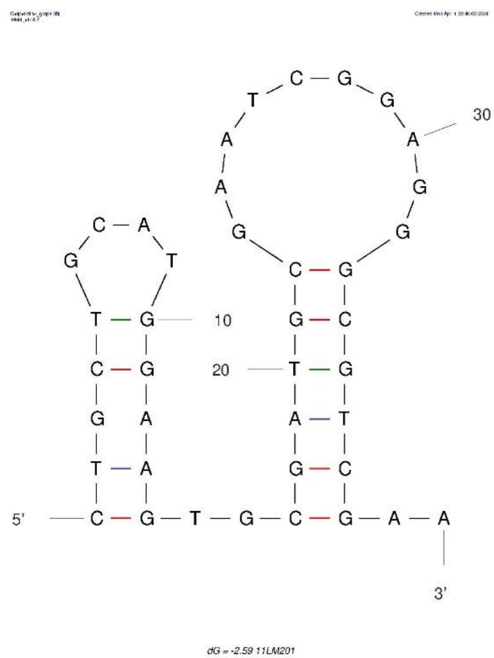

dG = -1.05 12LK228
dG = -1.02 12LK228
dG = -0.65 12LK228
dG = -0.48 12LK228

(F) 11LM201

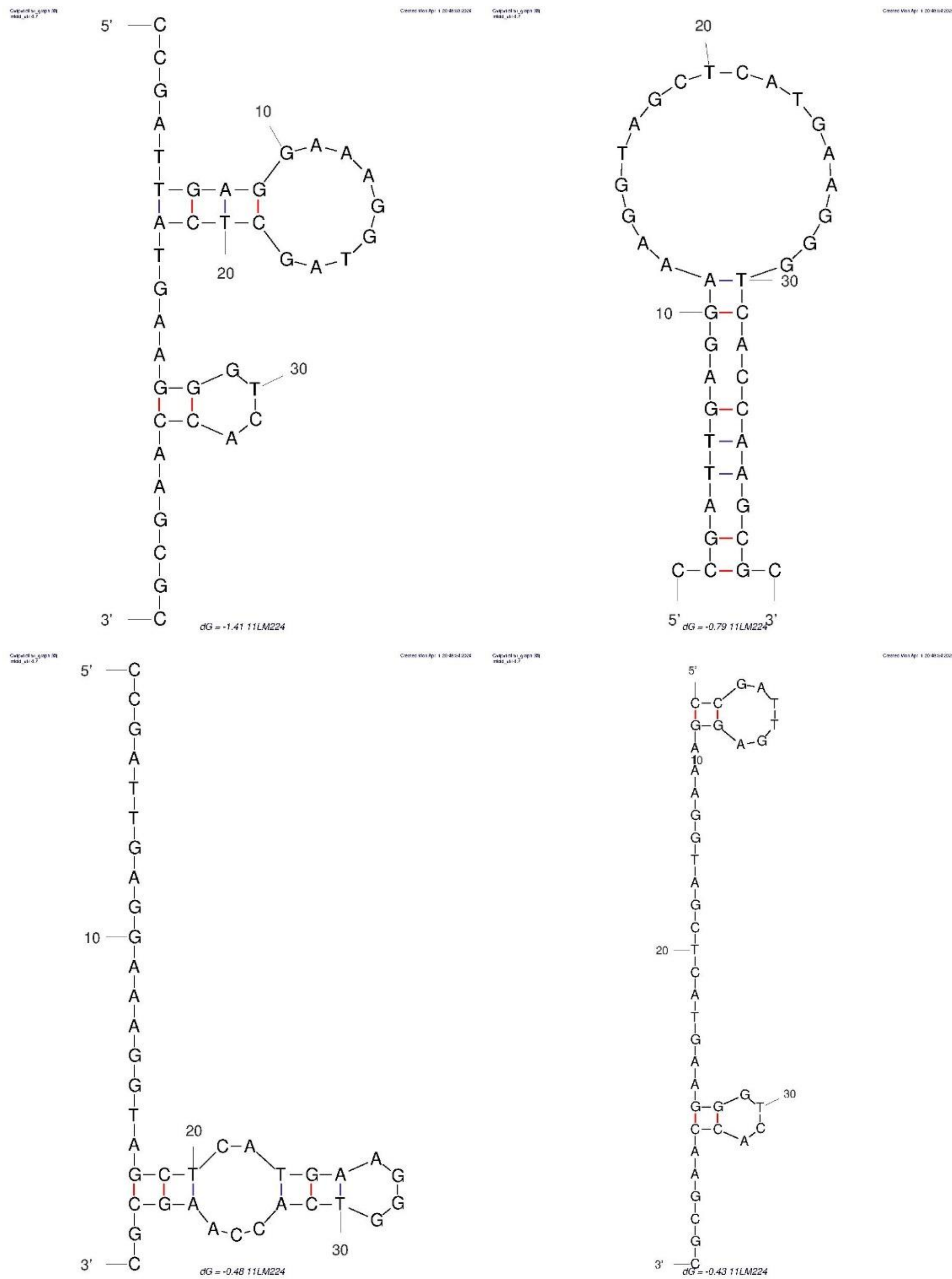



5'
3'
10
20
30
dG = -2.84 11LM201

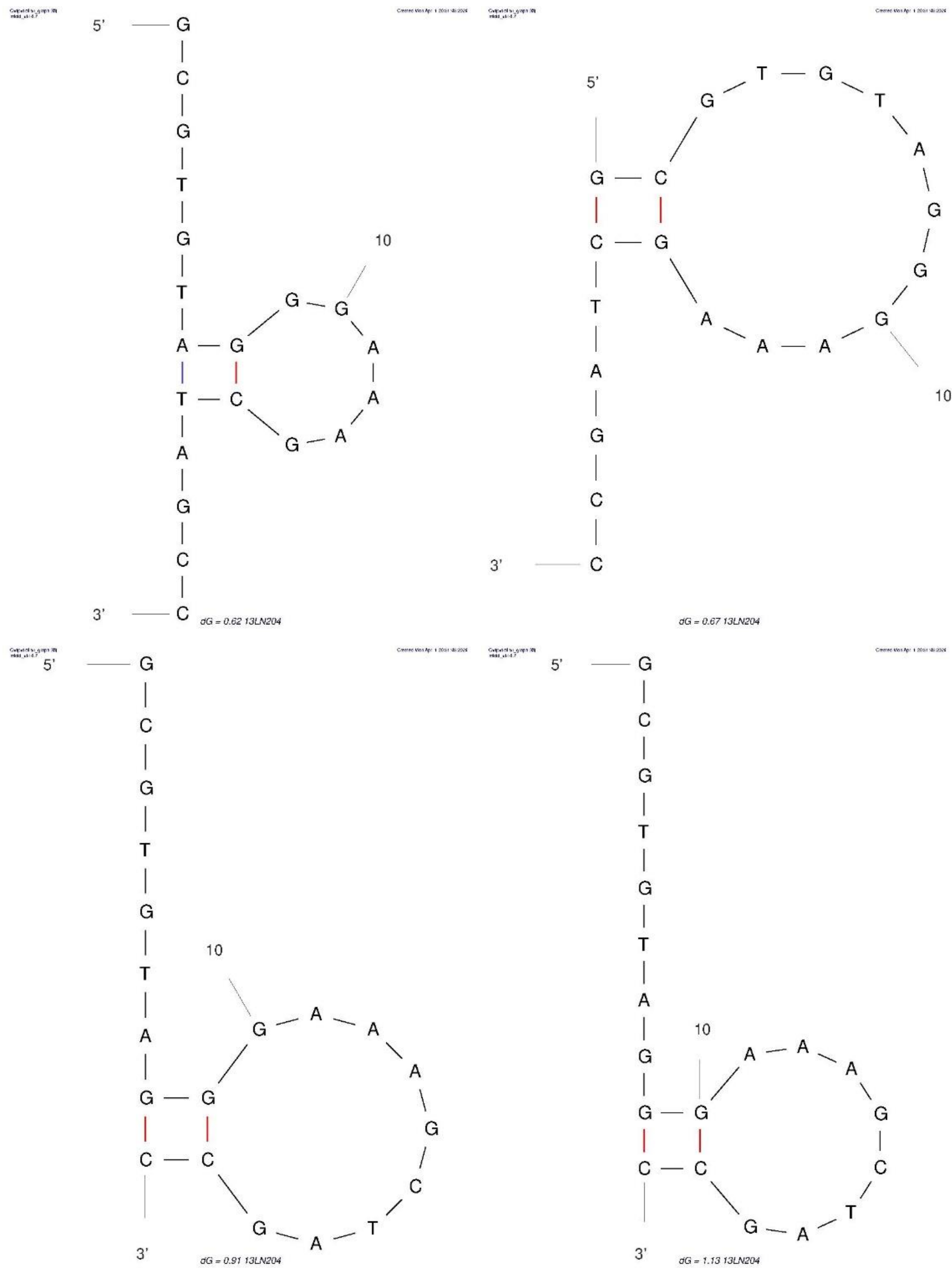

5'
3'
10
20
30
dG = -2.82 11LM201

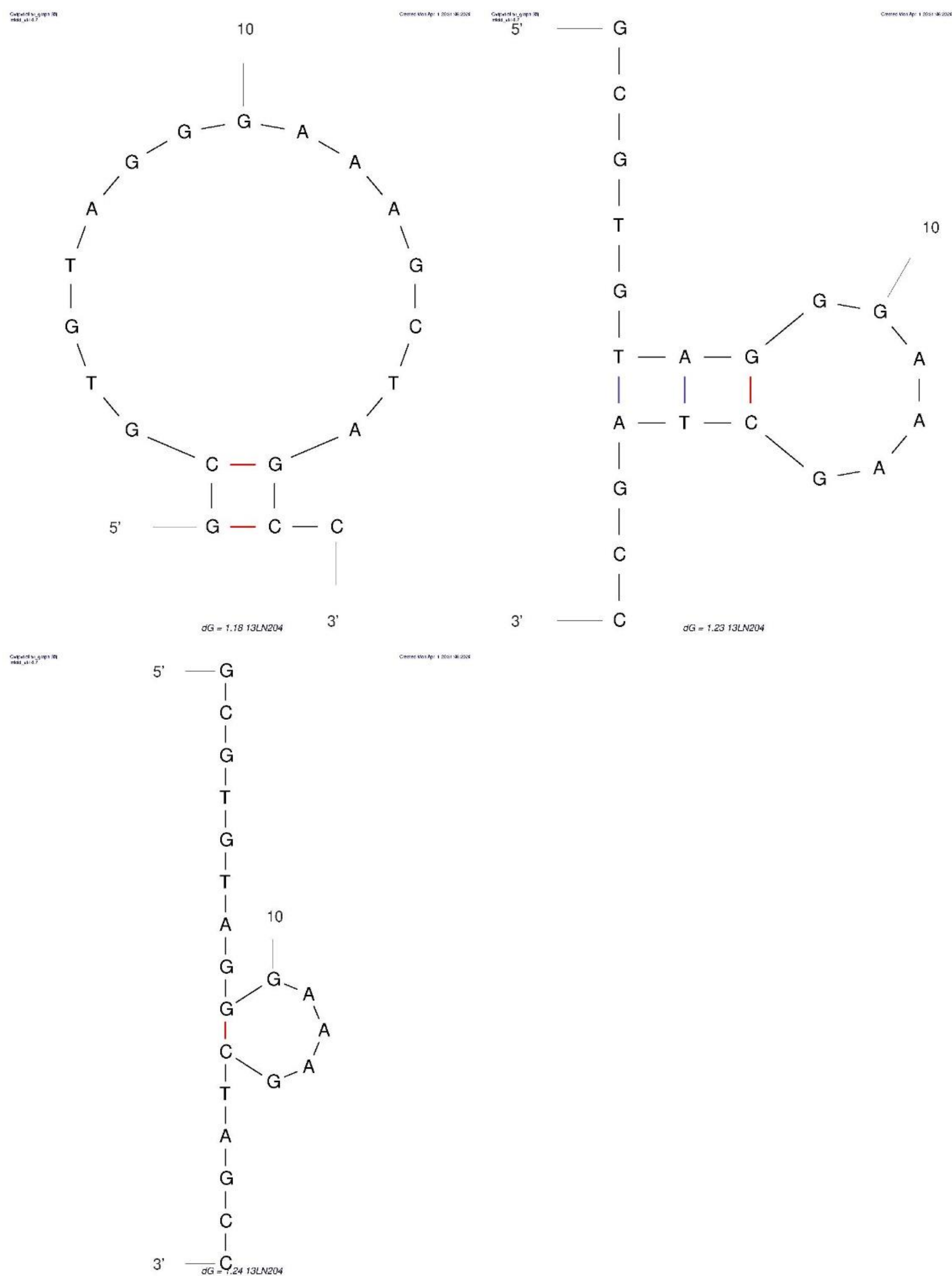

5'
3'
10
20
30
dG = -2.59 11LM201

(G) 11LM224

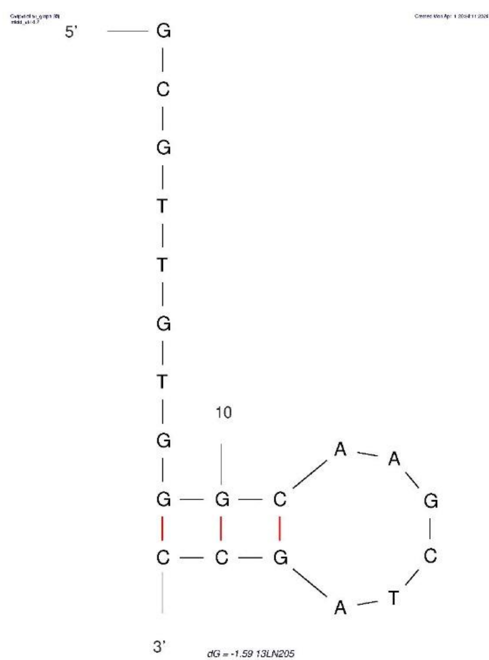


(H) 13LN204

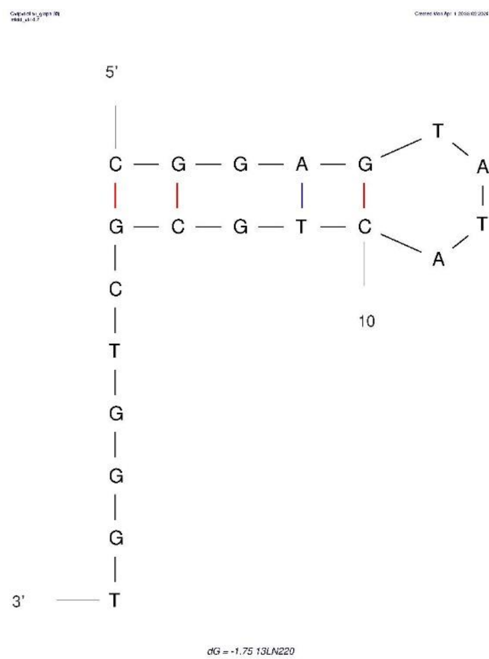



10
5'
3'
dG = 1.18 13LN204
5'
10
3'
dG = 1.23 13LN204
5'
10
3'
dG = 1.24 13LN204

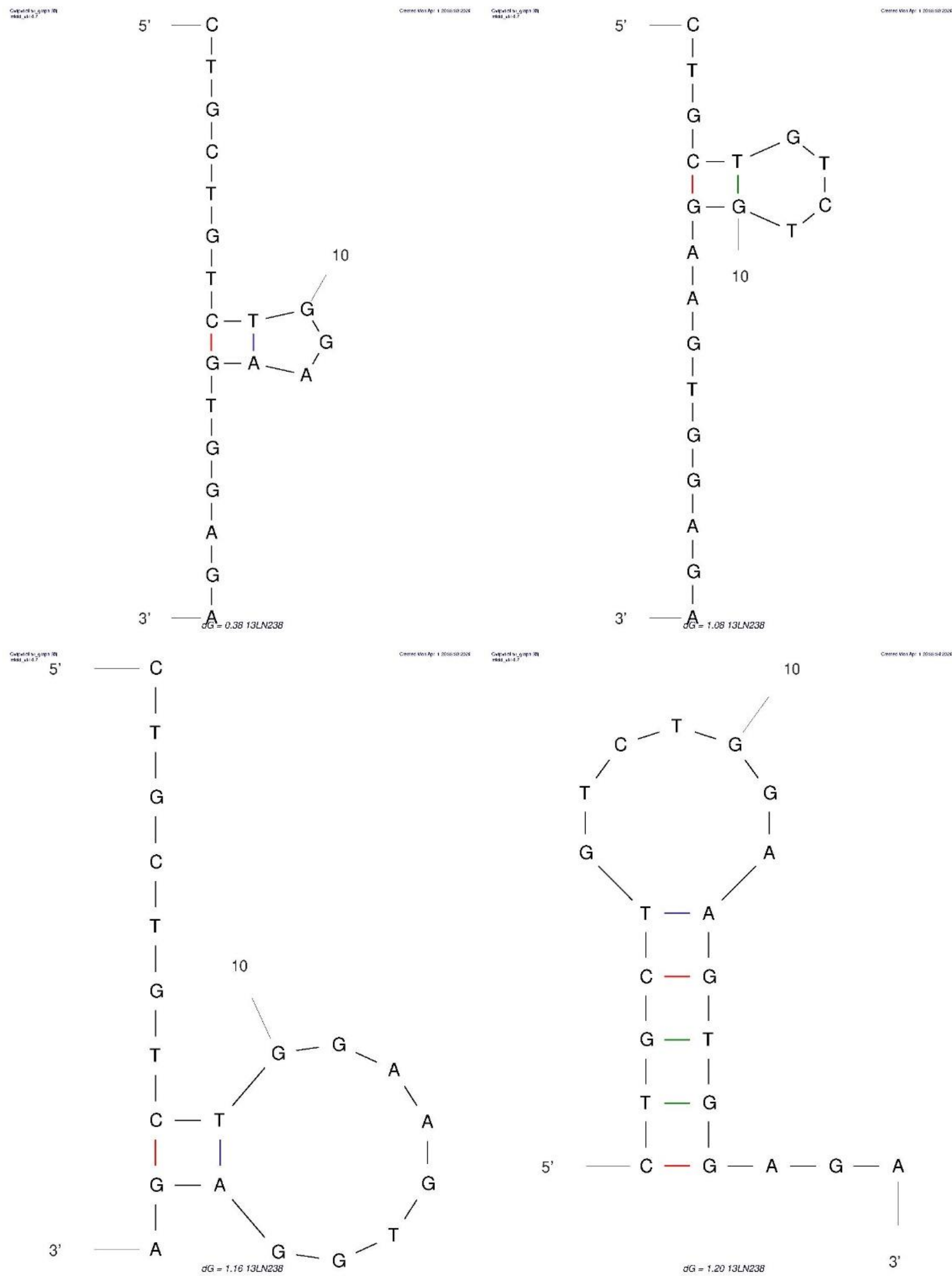
(I) 13LN205



(J) 13LN220



(K) 13LN238



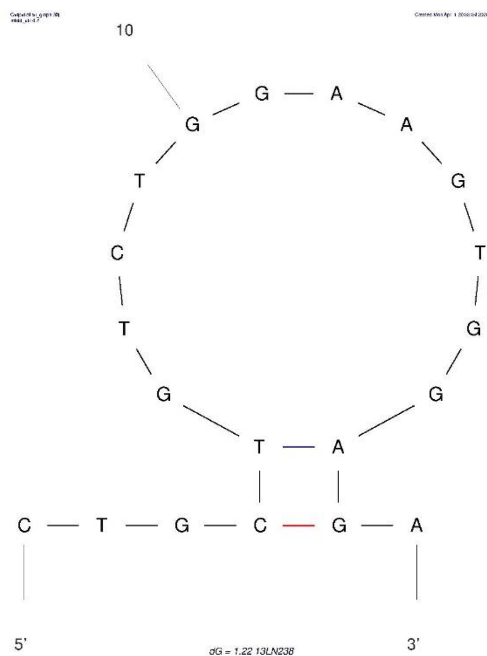


Figure S6. Mfold-predicted[3] secondary structures for all 11 DNAzymes reported in this study. The default settings were used for the initially random N_{40} regions with the DNA Folding Form at http://unafold.rna.albany.edu/?q=mfold, adjusted to 150 mM Na^{+} and 40 mM Mg^{2+}. Where multiple structures 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 S3 for full tabulation of number of structures and lowest ΔG value for all 11 DNAzymes identified in this study.

Assays of DNAzymes with mutated peptide substrates

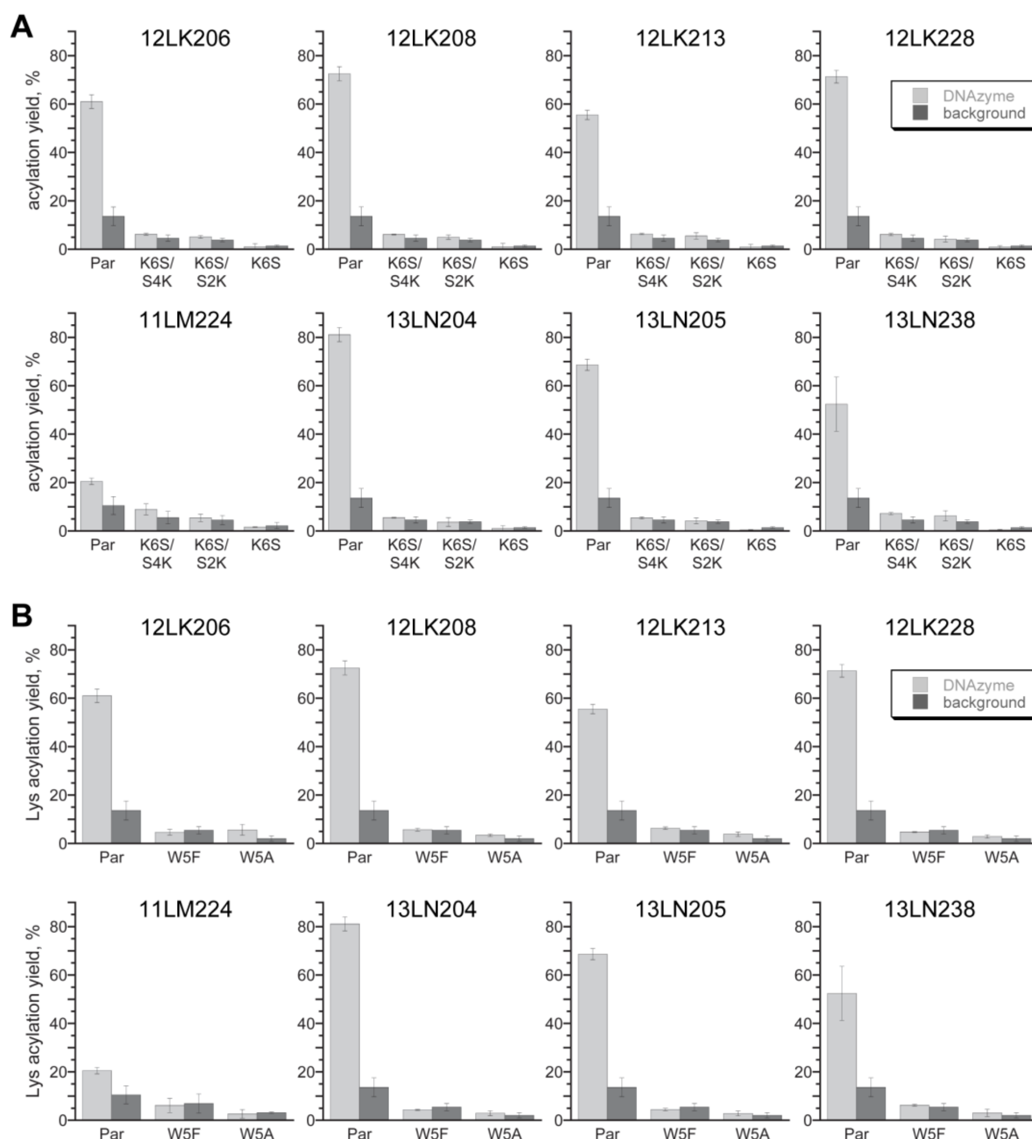


Figure S7. Additional data for peptide sequence requirements of the DNAzymes. All yields at 24 h. (A) Evaluating site selectivity for Lys within the peptide substrate. See Figure 6A for data for 12LK210, 11LM201, and 13LN220. (B) Evaluating sequence context requirements by mutating the Trp adjacent to Lys. See Figure 6B for data for 12LK210, 11LM201, and 13LN220.

References for Supporting Information

[1] M. M. Kennebeck, C. K. Kaminsky, M. A. Massa, P. K. Das, R. D. Boyd, M. Bishka, J. T. Tricarico, S. K. Silverman, *Angew. Chem. Int. Ed.* **2024**, *63*, e202317565.
[2] U. J. Desai, P. K. Pfaffle, *BioTechniques* **1995**, *19*, 780-784.
[3] M. Zuker, *Nucleic Acids Res.* **2003**, *31*, 3406-3415.