Angewandte Chemie

Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

© Wiley-VCH 2008

69451 Weinheim, Germany

DNA-Catalyzed Formation of Nucleopeptide Linkages

P. I. Pradeepkumar, Claudia Höbartner, Dana A. Baum, and Scott K. Silverman*

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

Table of Contents

Oligonucleotide synthesis

DNA oligonucleotides and DNA-RNA chimeric oligonucleotides were prepared at IDT (Coralville, IA). RNA oligonucleotides were prepared by in vitro transcription with T7 RNA polymerase and a synthetic DNA template.[1] All oligonucleotides, including those incorporating amino acid moieties, were purified by denaturing PAGE with running buffer 1× TBE (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3) as described previously.[2,3] For synthesis of DNA-peptide-DNA strands, the following reagents were purchased from Glen Research: 5′-amino-dT-CE phosphoramidite, 1*H*-tetrazole; deblocking solution [3% (w/v) trichloroacetic acid (TCA) in CH_2Cl_2] and oxidizing reagent (0.02 M I_2 in THF/pyridine/water 7:2:1). Bpoc-Alanine and *O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) were purchased from Chem-Impex International. 1-Hydroxybenzotriazole (HOBt) and *N*-methylmorpholine (NMM) were purchased from Acros.

DNA-peptide-DNA strands were synthesized on the 0.2 µmol scale via a multi-step assembly process (Scheme S1). The sequences are shown in Table S1. The preparative route involved synthesis of the 3′-portion of DNA on the solid phase; attachment of the amino acid to the DNA; and synthesis of the 5′-portion of DNA via a phosphoramidate linkage to the α-amino group. An oligonucleotide representing the 3′-portion of the DNA with 5′-terminal 5′-amino-5′-deoxythymidine was synthesized on the solid support using an ABI 394 DNA synthesizer (W. M. Keck Center for Comparative and Functional Genomics, UIUC). The final coupling of the series was performed manually using the 5′-amino-modified thymidine phosphoramidite from Glen Research in which the 5′-NH_2 group is protected using the acid-labile MMT group.[4,5] With the oligonucleotide still attached to the solid support, the MMT group was deprotected using 3% TCA in CH_2Cl_2, and the free 5′-NH_2 was manually coupled with Bpoc-Ala or with a MMT-protected amino acid (Ser, Tyr, Phe, or Lys; see below), using HATU/HOBt as coupling agent and NMM as the base in DMF.[5] Deprotection of Bpoc or MMT was achieved with 3% TCA in CH_2Cl_2, unmasking the free α-amino group. The α-amino group was manually coupled to the dG phosphoramidite using standard coupling conditions, leading to a phosphoramidate linkage.[6] After each of the three manual coupling steps, capping was omitted (for the first two manual couplings, capping leads to substantial truncation of synthesis via acetylation of free NH_2 groups[5]). After manual oxidation of P(III) to P(V) and TCA deprotection of the DMT group, synthesis of the 5′-portion of the DNA was completed on the DNA synthesizer. The final DNA-peptide-DNA product was removed from the solid support and fully deprotected with 29% aqueous ammonia at 55 ºC for 12 h and purified by 20% denaturing PAGE. The detailed procedures for all manually performed steps are described below. The products were quantified by UV absorbance (A_{260}); yields were 18–70 nmol of each DNA-peptide-DNA strand. Product identities were confirmed by MALDI-TOF mass spectrometry (Table S1).

	m/*z* [M+H]⁻	
DNA-peptide-DNA	calcd.	found
5'-GGATAATACG-Ala-TTCACTGCG-3'	5897	5898 ± 6
5'-GGATAATACG-Ser-TTCACTGCG-3'	5913	5914 ± 6
5'-GGATAATACG-Tyr-TTCACTGCG-3'	5990	5992 ± 6
5'-GGATAATACG-Lys-TTCACTGCG-3'	5954	5952 ± 6
5'-GGATAATACG-Phe-TTCACTGCG-3'	5974	5973 ± 6

Table S1. Sequences and MALDI-TOF mass spectrometry data for DNA-peptide-DNA strands.
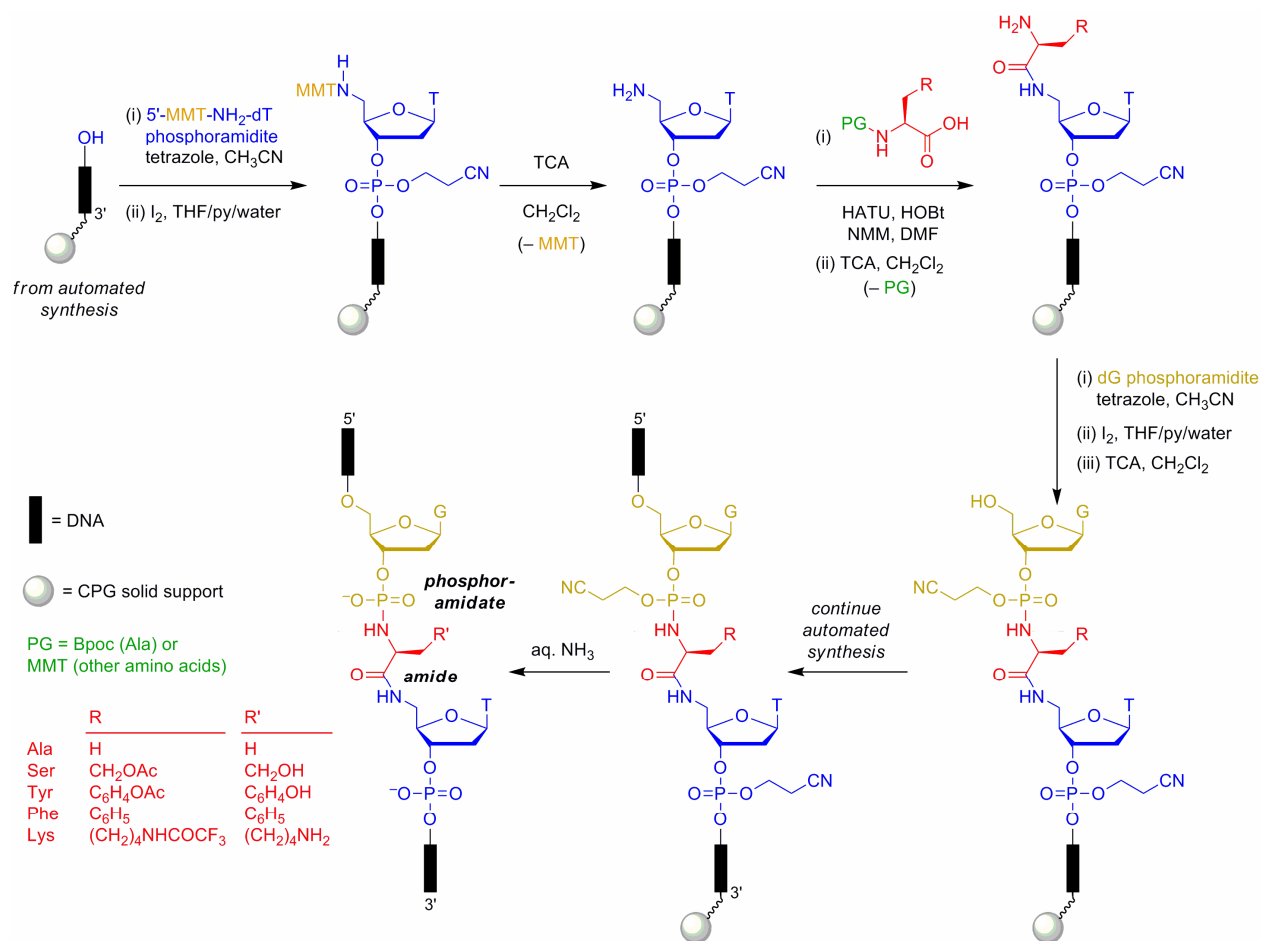
Scheme S1. Multi-step assembly route for synthesis of DNA-peptide-DNA strands.

Manual coupling of 5'-NH$_2$-dT phosphoramidite to 5'-OH of DNA, or of dG phosphoramidite to α-amino group. Solutions of phosphoramidite (200 μL, 100 mg/mL in CH$_3$CN) and tetrazole (500 μL, 0.45 M in CH$_3$CN) were stored separately over 4Å molecular sieves for 16 h, then mixed together and introduced into the DNA synthesis column containing the CPG. The coupling reaction was performed for 12 min, with manual mixing using two syringes every 3 min. The CPG was washed with CH$_3$CN (3 × 5 mL) and dried under a stream of nitrogen. For the coupling of dG to the α-amino group, a second coupling–oxidation cycle was performed to increase the efficiency.

Manual oxidation of P(III) to P(V). Oxidizing reagent (1 mL, 0.02 M I$_2$ in THF/pyridine/water) was introduced into the DNA synthesis column containing the CPG. The oxidation reaction was performed for 15 min, with manual mixing using two syringes every 5 min. The CPG was washed with DMF (2 × 5 mL) and CH$_3$CN (2 × 5 mL) and dried under a stream of nitrogen.
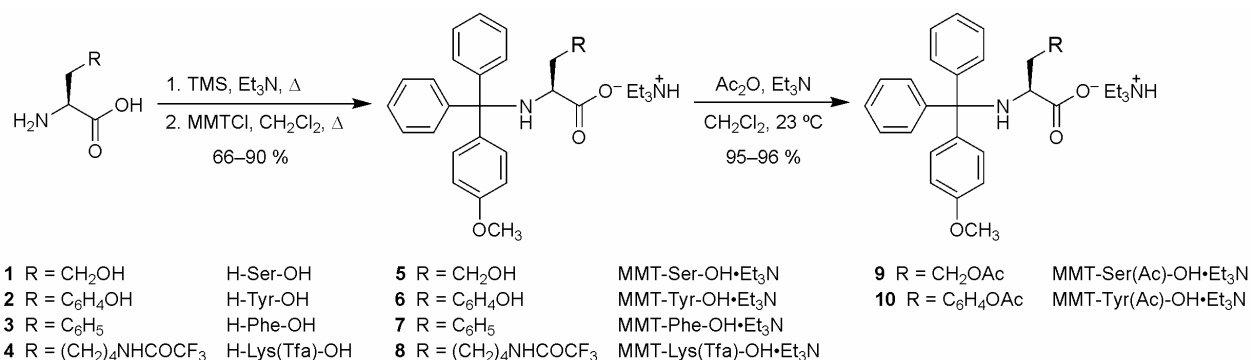
Manual removal of MMT/DMT protecting group. Deblocking solution (1 mL, 3% TCA in CH$_2$Cl$_2$) was slowly passed during the course of 1 min though the column containing the CPG. The CPG was rinsed with CH$_2$Cl$_2$ (2 × 5 mL) and CH$_3$CN (2 × 5 mL). After MMT release, the efficiency of the next coupling reaction was increased by slowly passing during the course of 1 min a solution of 2% (v/v) NMM in CH$_3$CN (1.5 mL) through the CPG, followed by washing with CH$_3$CN (2 × 5 mL).

Manual coupling of protected amino acid with 5'-NH$_2$-T to create amide linkage. A mixture of Bpoc- or MMT-protected amino acid (100 μmol), HATU (38 mg, 100 μmol), and HOBt (13 mg, 100 μmol) was dissolved in DMF (500 μL), and NMM (26 μL, 230 μmol) was added. The well-mixed solution was immediately introduced into the DNA synthesis column containing the CPG. The coupling reaction was

performed for 50 min, with manual mixing using two syringes every 5 min. The CPG was washed with DMF (2×5 mL), CH_3CN (2×5 mL), and CH_2Cl_2 (1×5 mL) and dried under a stream of nitrogen.

Organic synthesis of MMT-protected amino acids

The synthetic route is shown in Scheme S2. Reagents were commercial grade and used without purification unless otherwise indicated. Dry solvents were obtained from Aldrich Sure/Seal or Acros Acroseal bottles or by drying over freshly activated 4Å molecular sieves overnight. All reactions were performed under argon or nitrogen unless otherwise noted. Thin-layer chromatography (TLC) was performed on silica gel plates pre-coated with fluorescent indicator with visualization by UV light (254 nm) or by dipping into a solution of 5% (v/v) conc. H_2SO_4 in ethanol and heating. Flash column chromatography was performed with silica gel (230–400 mesh). Silica gel columns were packed with the initial solvent additionally containing 2% triethylamine. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Unity 500 instrument. The chemical shifts in parts per million (δ) are reported downfield from TMS (0 ppm) and referenced to the residual proton signal of the deuterated solvent, as follows: $CDCl_3$ (7.26 ppm) for ^{1}H NMR spectra; $CDCl_3$ (77.0 ppm) for ^{13}C NMR spectra. Multiplicities of ^{1}H NMR spin couplings are reported as s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), or m (multiplet and overlapping spin systems). Values for apparent coupling constants (*J*) are reported in Hz. Mass spectrometry data were obtained at the UIUC School of Chemical Sciences mass spectrometry laboratory using a Micromass Q-Tof Ultima instrument (HR-ESI).

1 R = CH_2OH	H-Ser-OH	**5** R = CH_2OH	MMT-Ser-OH•Et_3N	**9** R = CH_2OAc	MMT-Ser(Ac)-OH•Et_3N
2 R = C_6H_4OH	H-Tyr-OH	**6** R = C_6H_4OH	MMT-Tyr-OH•Et_3N	**10** R = C_6H_4OAc	MMT-Tyr(Ac)-OH•Et_3N
3 R = C_6H_5	H-Phe-OH	**7** R = C_6H_5	MMT-Phe-OH•Et_3N		
4 R = $(CH_2)_4NHCOCF_3$	H-Lys(Tfa)-OH	**8** R = $(CH_2)_4NHCOCF_3$	MMT-Lys(Tfa)-OH•Et_3N		

Scheme S2. Synthesis of MMT-protected amino acids.

MMT-Ser-OH•Et₃N (5)

The general approach was adapted from a previous report.[7] Portions of TMSCl (0.9 mL, 7.9 mmol) and Et₃N (1.1 mL, 7.0 mmol) were added to a suspension of H-Ser-OH **1** (210 mg, 2 mmol) in CH_2Cl_2 (15 mL) at room temperature. The mixture was heated at reflux (bath temperature 60 ºC) for 2 h. After cooling to room temperature, MMTCl (620 mg, 2.0 mmol) was added and the solution was heated at reflux for 2 h. After cooling to room temperature, the solution was cooled further in an ice bath, and Et₃N (4 mL, 29 mmol) and MeOH (2 mL, 50 mmol) were added. After 30 min, the solution was diluted with CH_2Cl_2 (200 mL) and washed with 5% (w/v) aqueous citric acid (50 mL) and saturated NaCl (2 × 50 mL). The organic layer was dried over sodium sulfate and concentrated. The oily residue was purified by chromatography using 0–4% methanol in CH_2Cl_2 with 2% (v/v) Et₃N.

Yield: 633 mg of **5** (66%) as a white solid.

TLC: R_f = 0.6 [4% methanol in CH_2Cl_2 with 2% (v/v) Et₃N].

^{1}H NMR (500 MHz, CDCl₃): δ 7.51-7.49 (m, 4H), 7.39 (d, J = 9.0 Hz, 2H), 7.24 (t, J = 7.5 Hz, 4H), 7.17-7.15 (m, 2H), 6.78 (d, J = 9.0 Hz, 2H), 3.76 (s, 3H), 3.32 (dd, J = 10.0, 3.5 Hz, 1 H), 3.26 (dd, J = 7.0, 3.5 Hz, 1 H), 3.01-2.93 (m, 7H), 1.24 (t, J =7.5 Hz, 9H) ppm.

^{13}C NMR (500 MHz, CDCl₃): δ 178.9, 157.8, 146.8, 138.4, 130.3, 128.9, 128.8, 127.6, 126.2, 112.9, 70.9, 64.3, 58.2, 55.1, 44.9, 8.4 ppm.

HRMS: m/z calcd. for $C_{23}H_{22}NO_4$ [M]⁻ 376.1549, found: 376.1550 (Δm 0.0001, error 0.3 ppm).
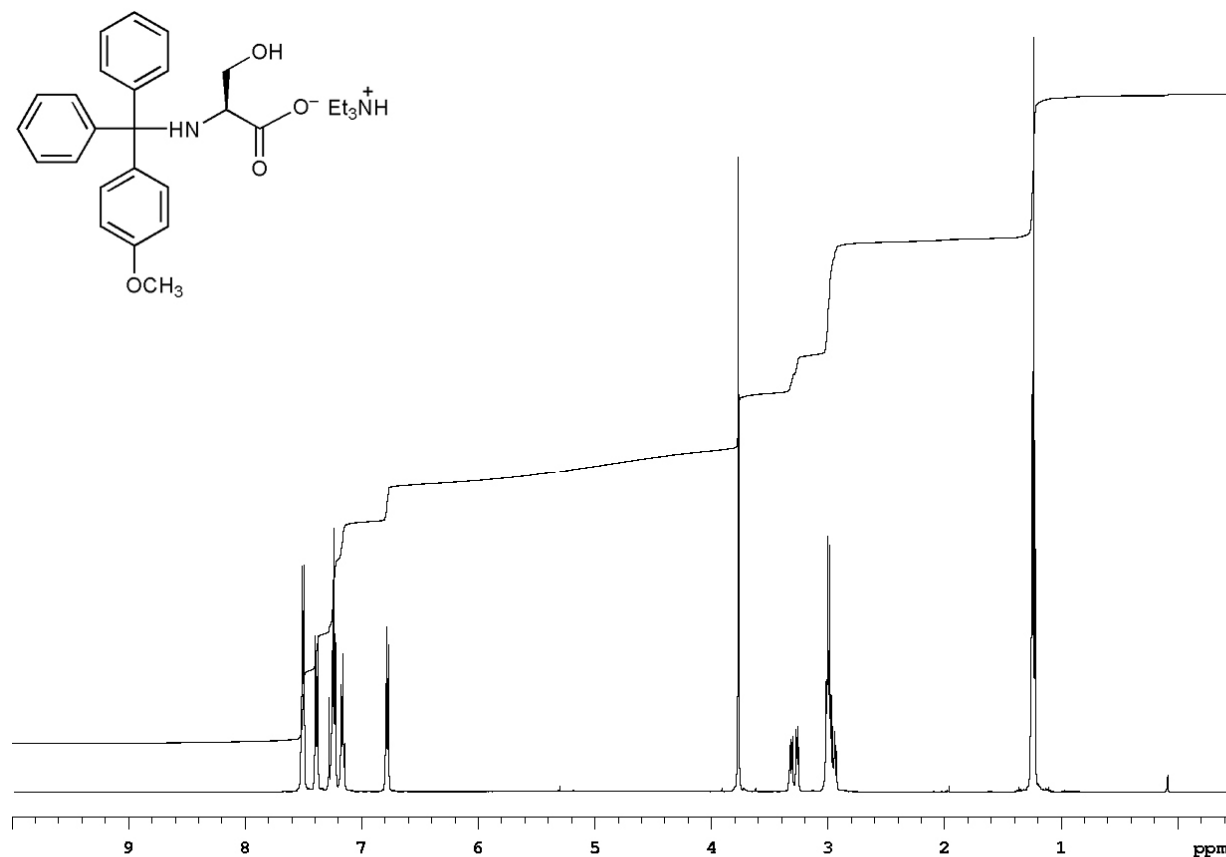
MMT-Tyr-OH•Et₃N (6)

Portions of TMSCl (0.45 mL, 3.5 mmol) and Et$_3$N (0.55 mL, 3.6 mmol) were added to a suspension of H-Tyr-OH **2** (181 mg, 1.0 mmol) in CH$_2$Cl$_2$ (7 mL) at room temperature. The mixture was heated at reflux (bath temperature 60 ºC) for 2 h. After cooling to room temperature, MMTCl (309 mg, 1.0 mmol) was added and the solution was heated at reflux for 2 h. After cooling to room temperature, the solution was cooled further in an ice bath, and Et$_3$N (2 mL, 14 mmol) and MeOH (1 mL, 25 mmol) were added. After 30 min, the solution was diluted with CH$_2$Cl$_2$ (125 mL) and washed with 5% (w/v) aqueous citric acid (25 mL) and saturated NaCl (2 × 25 mL). The organic layer was dried over sodium sulfate and concentrated. The oily residue was purified by chromatography using 0–3% methanol in CH$_2$Cl$_2$ with 2% (v/v) Et$_3$N.
Yield: 500 mg of **6** (90%) as a white solid.
TLC: R$_f$ = 0.4 [4% methanol in CH$_2$Cl$_2$ with 2% (v/v) Et$_3$N].
^{1}H NMR (500 MHz, CDCl$_3$): δ 7.53-7.51 (m, 4H), 7.40 (d, J = 9.0 Hz, 2H), 7.18 (t, J = 8.0 Hz, 4H), 7.11-7.08 (m, 2H), 6.99 (d, J = 8.0 Hz, 2H), 6.73 (d, J = 9.0 Hz, 2H), 6.47 (d, J = 8.5 Hz, 2H), 3.73 (s, 3H), 3.47 (X of ABX, J_{AX} = 5.0 Hz, J_{BX} = 6.5 Hz, 1 H), 2.79 (A of ABX, J_{AB} = 13.5 Hz, J_{AX} = 5.0 Hz, 1H), 2.73 (q, J = 7.5 Hz, 6H), 2.53 (B of ABX, J_{AB} = 13.5 Hz, J_{BX} = 6.5 Hz, 1H), 1.07 (t, J =7.5 Hz, 9H) ppm.
^{13}C NMR (500 MHz, CDCl$_3$): 179.6, 157.5, 155.6, 147.5, 147.2, 139.1, 130.6, 130.4, 129.4, 128.9, 128.8, 127.5, 125.8, 114.9, 112.7, 70.8, 59.3, 54.9, 55.0, 44.8, 40.8, 8.4 ppm.
HRMS: m/z calcd. for C$_{29}$H$_{26}$NO$_4$ [M]$^-$ 452.1862, found 452.1862 (Δm 0.0000, error 0.0 ppm).
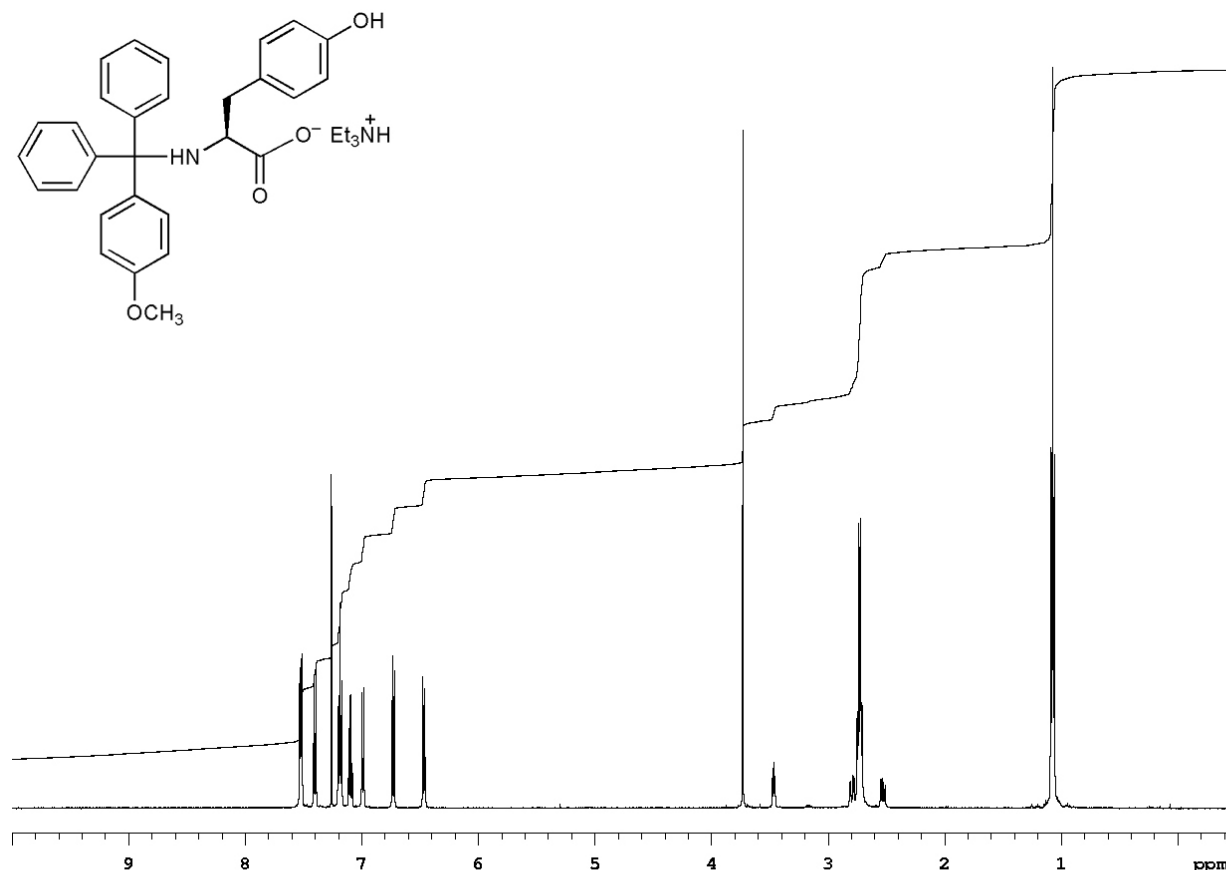
MMT-Phe-OH•Et₃N (7)

Portions of TMSCl (0.7 mL, 5.5 mmol) and Et₃N (0.8 mL, 5.7 mmol) were added to a suspension of H-Phe-OH **3** (248 mg, 1.5 mmol) in CH_2Cl_2 (12 mL) at room temperature. The mixture was heated at reflux (bath temperature 60 ºC) for 2 h. After cooling to room temperature, MMTCl (462 mg, 1.5 mmol) was added and the solution was heated at reflux for 2 h. After cooling to room temperature, the solution was cooled further in an ice bath, and Et₃N (2 mL, 14 mmol) and MeOH (2 mL, 50 mmol) were added. After 30 min, the solution was diluted with CH_2Cl_2 (150 mL) and washed with 5% (w/v) aqueous citric acid (30 mL) and saturated NaCl (2×30 mL). The organic layer was dried over sodium sulfate and concentrated. The oily residue was purified by chromatography using 0–2% methanol in CH_2Cl_2 with 2% (v/v) Et₃N.

Yield: 650 mg of **7** (80%) as a white solid.

TLC: R_f = 0.3 [2% methanol in CH_2Cl_2 with 2% (v/v) Et₃N].

^{1}H NMR (500 MHz, CDCl₃) δ 7.47-7.43 (m, 4H), 7.33 (d, J = 9.0 Hz, 2H), 7.27-7.08 (m, 11H), 6.71 (d, J = 9.0 Hz, 2H), 3.74 (s, 3H), 3.48 (X of ABX, J_{AX} = 6.0 Hz, J_{BX} = 5.5 Hz, 1 H), 2.79 (A of ABX, J_{AB} = 13.0 Hz, J_{AX} = 6.0 Hz, 1H), 2.75 (q, J = 7.5 Hz, 6H), 2.70 (B of ABX, J_{AB} = 13.5 Hz, J_{BX} = 5.5 Hz, 1H), 1.06 (t, J = 7.5 Hz, 9H) ppm.

^{13}C NMR (500 MHz, CDCl₃) δ 179.1, 157.9, 147.7, 147.4, 139.9, 139.6, 130.7, 130.6, 129.3, 129.2, 128.1, 127.9, 126.2, 126.1, 113.1, 71.0, 59.4, 55.4, 44.7, 42.3, 8.9 ppm.

HRMS: *m/z* calcd. for $C_{29}H_{26}NO_3$ [M]⁻ 436.1913; found: 436.1900 (Δm 0.0013, error 3.0 ppm).
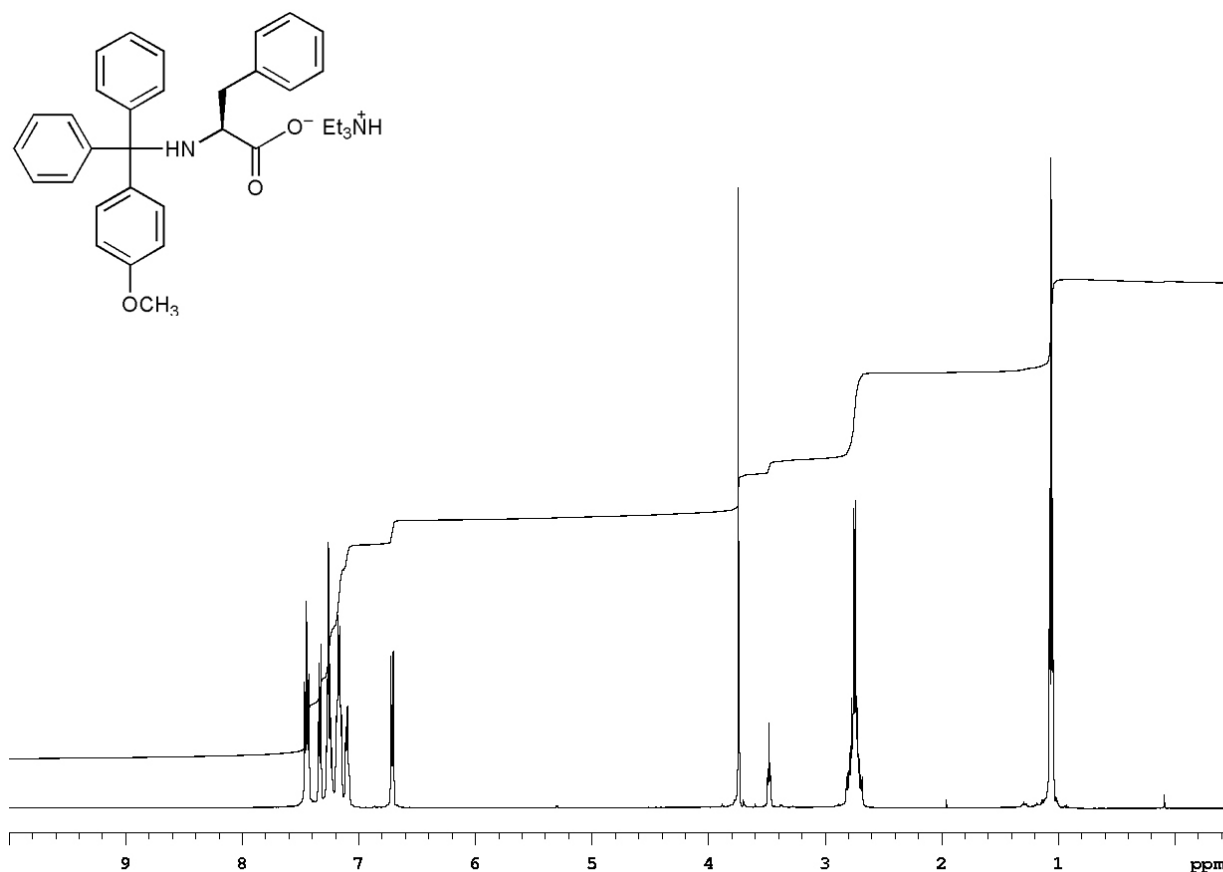
MMT-Lys(Tfa)-OH•Et₃N (8)

Portions of TMSCl (0.45 mL, 3.5 mmol) and Et₃N (0.48 mL, 3.5 mmol) were added to a suspension of H-Lys(TFA)-OH **4** (242 mg, 1.0 mmol, EMD Biosciences) in CH_2Cl_2 (7 mL) at room temperature. The mixture was heated at reflux (bath temperature 60 ºC) for 2 h. After cooling to room temperature, MMTCl (309 mg, 1.0 mmol) was added and the solution was heated at reflux for 2 h. After cooling to room temperature, the solution was cooled further in an ice bath, and Et₃N (2 mL, 14 mmol) and MeOH (1 mL, 25 mmol) were added. After 30 min, the solution was diluted with CH_2Cl_2 (125 mL) and washed with 5% (w/v) aqueous citric acid (25 mL) and saturated NaCl (2 × 25 mL). The organic layer was dried over sodium sulfate and concentrated. The oily residue was purified by chromatography using 0–3% methanol in CH_2Cl_2 with 2% (v/v) Et₃N.

Yield: 452 mg of **8** (74%) as a white solid.
TLC: R_f = 0.3 [1% methanol in CH_2Cl_2 with 2% (v/v) Et₃N].
¹H NMR (500 MHz, CDCl₃) δ 7.51-7.49 (m, 4H), 7.38 (d, J = 9.0 Hz, 2H), 7.22-7.19 (m, 4H), 7.14-7.11 (m, 2H), 6.74 (d, J = 9.0 Hz, 2H), 3.75 (s, 3H), 3.29-3.18 (m, 3H), 2.86 (q, J = 7.0 Hz, 6H), 1.51-1.44 (m, 4H), 1.29-1.19 (m, 2H), 1.14 (t, J =7.0 Hz, 9H) ppm.
¹³C NMR (500 MHz, CDCl₃) δ 179.8, 157.6, 147.4, 147.2, 139.1, 130.3, 128.9, 127.5, 125.9, 112.7, 70.8, 57, 55.1, 44.5, 39.6, 34.4, 28.2, 21.8, 8.5 ppm.
HRMS: *m/z* calcd. for $C_{28}H_{28}N_2O_4F_3$ [M]⁻ 513.1996, found: 513.2001 (Δm 0.0005, error 1.0 ppm).
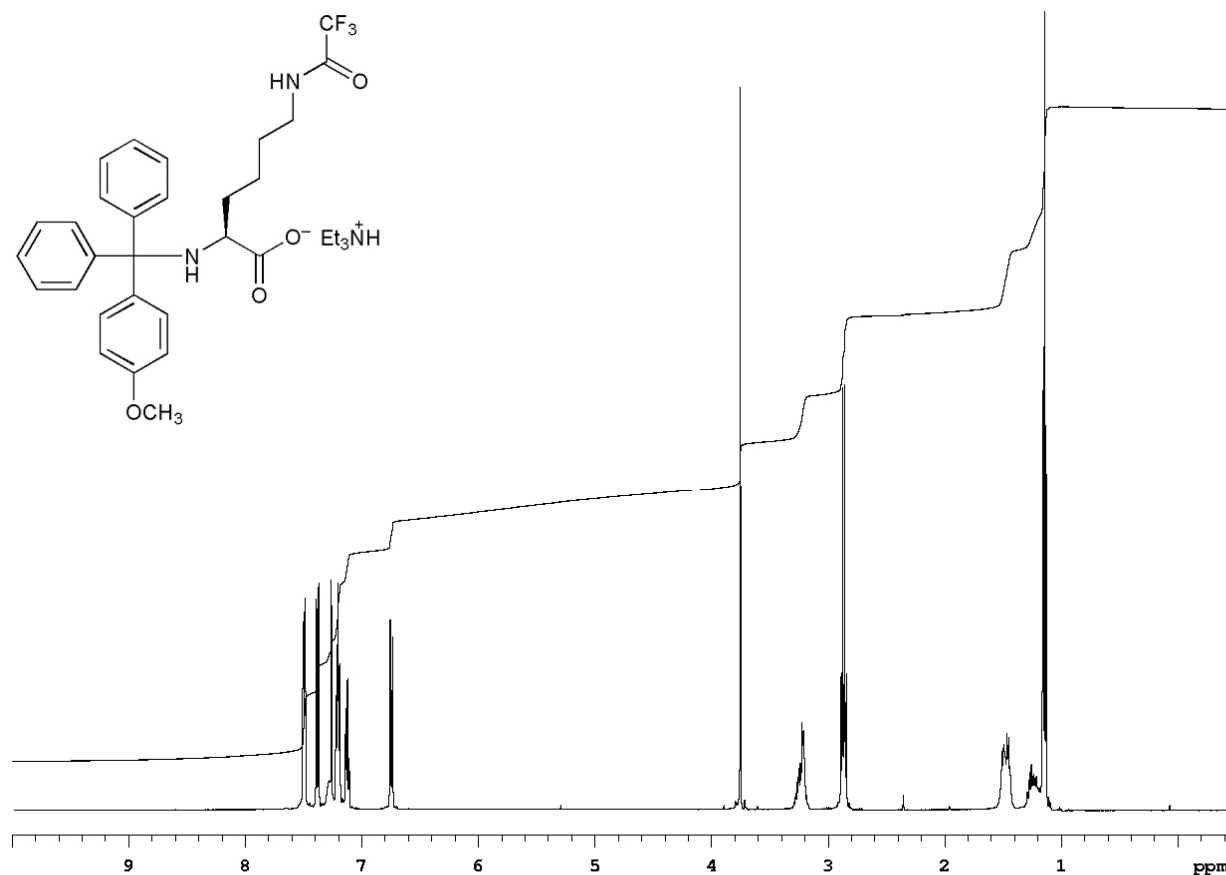
MMT-Ser(Ac)-OH•Et₃N (9)

To a stirred solution of **5** (230 mg, 0.48 mmol) in CH_2Cl_2 (6 mL) was added Et_3N (0.42 mL, 3.0 mmol) and Ac_2O (0.28 mL, 3.0 mmol). The solution was stirred at room temperature for 4 h, diluted with CH_2Cl_2 (50 mL), and poured into saturated $NaHCO_3$ (25 mL). The aqueous layer was back-extracted with CH_2Cl_2 (2 × 40 mL), and the combined organic layers were washed with saturated NaCl (25 mL). The organic layer was dried over sodium sulfate and concentrated. The oily residue was purified by chromatography using 0–3% methanol in CH_2Cl_2 with 2% (v/v) Et_3N.

Yield: 237 mg of **9** (95%) as a white solid.

TLC: R_f = 0.7 [4% methanol in CH_2Cl_2 with 2% (v/v) Et_3N].

^{1}H NMR (500 MHz, CDCl₃) δ 7.52-7.50 (m, 4H), 7.4 (d, J = 9.0 Hz, 2H), 7.21-7.17 (m, 4H), 7.17-7.15 (m, 2H), 7.11-7.01 (m, 2H), 6.74 (d, J = 10.0 Hz, 2H), 4.16 (A of ABX, J_{AB} = 11.0 Hz, J_{AX} = 4.0 Hz, 1H), 3.98 (B of ABX, J_{AB} = 10.5 Hz, J_{BX} = 4.5 Hz, 1H), 2.92 (s, 3H), 3.35 (t, X of ABX, J = 4.5 Hz, 1 H), 2.83 (q, J = 7.5 Hz, 6H), 2.05 (s, 3H), 1.15 (t, J =7.5 Hz, 9H) ppm.

^{13}C NMR (500 MHz, CDCl₃) δ 176.6, 170.9, 157.6, 147, 146.7, 138.6, 130.2, 128.8, 128.7, 127.5, 125.9, 112.7, 70.5, 67.7, 56.6, 54.9, 44.4, 20.5, 8.2 ppm.

HRMS: *m/z* calcd. for $C_{25}H_{24}NO_5$ [M]⁻ 418.1654, found: 418.1669 (Δm 0.0015, error 3.6 ppm).
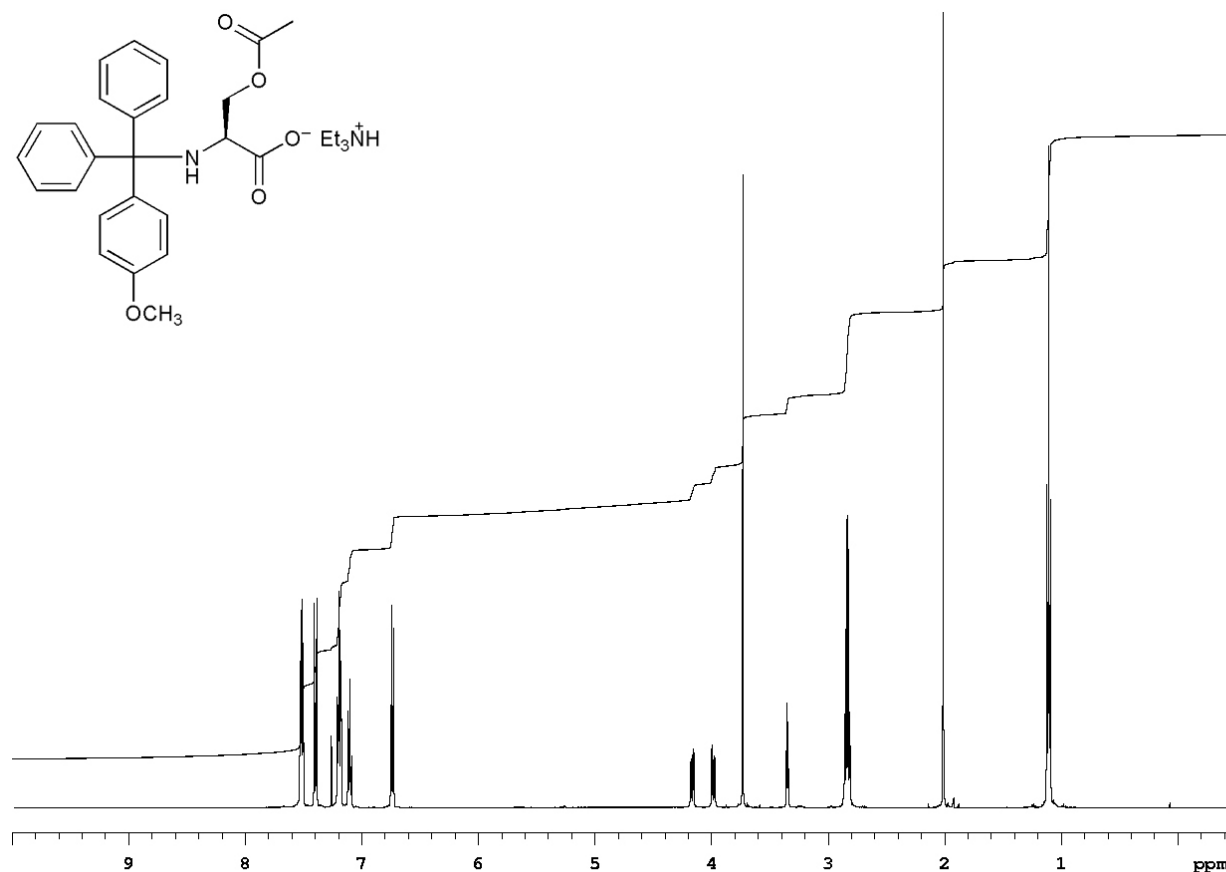
MMT-Tyr(Ac)-OH•Et₃N (10)

To a stirred solution of **6** (250 mg, 0.45 mmol) in CH_2Cl_2 (6 mL) was added Et_3N (0.42 mL, 3.0 mmol) and Ac_2O (0.28 mL, 3.0 mmol). The solution was stirred at room temperature for 4 h, diluted with CH_2Cl_2 (50 mL), and poured into saturated $NaHCO_3$ (25 mL). The aqueous phase was back-extracted with CH_2Cl_2 (2 × 40 mL), and the combined organic layers were washed with saturated NaCl (25 mL). The organic layer was dried over sodium sulfate and concentrated. The oily residue was purified by chromatography using 0–3% methanol in CH_2Cl_2 with 2% (v/v) Et_3N.

Yield: 260 mg of **10** (96%) as a white solid.

TLC: R_f = 0.6 [4% methanol in CH_2Cl_2 with 2% (v/v) Et_3N].

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.42-7.40 (m, 4H), 7.29 (d, J = 9.0 Hz, 2H), 7.20 (d, J = 8.5 Hz, 2H) 7.19-7.16 (m, 4H), 7.12-7.09 (m, 2H), 6.95 (d, J = 8.5 Hz, 2H), 6.72 (d, J = 8.5 Hz, 2H), 3.74 (s, 3H), 3.49 (X of ABX, J_{AX} = 6.0 Hz, J_{BX} = 5.5 Hz, 1 H), 2.78 (A of ABX, J_{AB} = 13.0 Hz, J_{AX} = 6.0 Hz, 1H), 2.75 (q, J = 7.5 Hz, 6H), 2.59 (B of ABX, J_{AB} = 13.5 Hz, J_{BX} = 5.5 Hz, 1H), 2.27 (s, 3H), 1.05 (t, J = 7.5 Hz, 9H) ppm.

^{13}C NMR (500 MHz, $CDCl_3$) δ 178.7, 169.9, 157.9, 149.2, 147.5, 147.3, 139.2, 137.4, 131.3, 130.6, 129.2, 127.8, 126.2, 120.9, 113.1, 70.9, 59.2, 55.3, 44.8, 41.4, 21.4, 8.7 ppm.

HRMS: m/z calcd. for $C_{31}H_{28}NO_5$ [M]⁻ 494.1967, found: 494.1964 (Δm 0.0003, error 0.6 ppm).
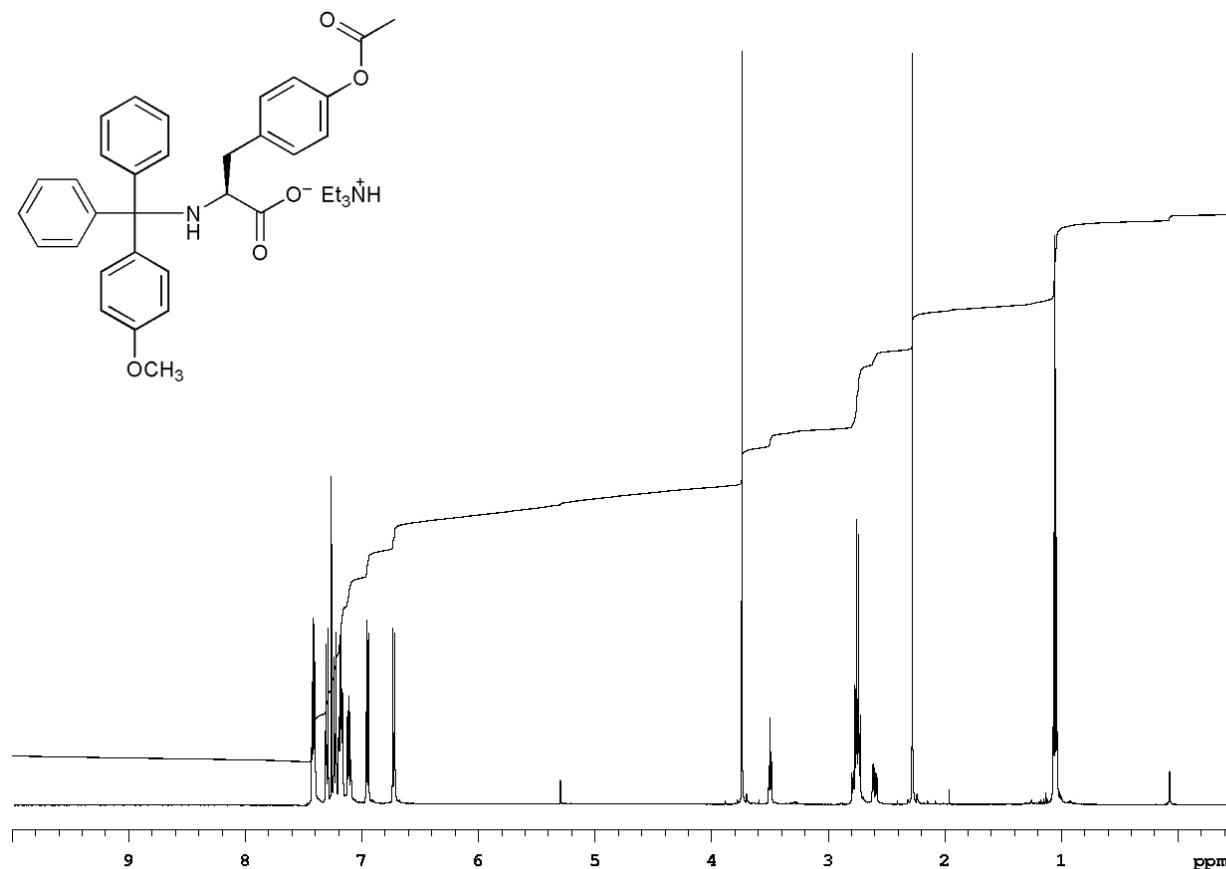
In vitro selection procedure, activity assays, and preparative procedures

The in vitro selection experiments and deoxyribozyme cloning procedures were performed essentially as described previously.[2] The key selection steps used incubation conditions of 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM MgCl$_2$ at 37 °C for 2 h. Sequences are shown in Figure 1 (the loop sequence at the far right was 5′-GGCGAA-3′, where the first two nucleotides are RNA and the next four are DNA).

For activity assays using deoxyribozymes prepared by solid-phase synthesis, the general approach was described previously.[2] For brevity, the DNA-X-DNA substrate is designated as the L (left-hand) substrate, and the 5′-triphosphate-RNA substrate is designated as the R (right-hand) substrate. The 5′-^{32}P-radiolabeled L substrate was the limiting reagent relative to the deoxyribozyme (E) and the R substrate. The L:E:R ratio was <1:3:10, with the concentration of E equal to ~0.3 µM in 10 µL assays. Values of k_{obs} were obtained by fitting the yield versus time data directly to first-order kinetics; i.e., yield = $Y(1 - e^{-kt})$, where $k = k_{obs}$ and Y is the final yield. However, when k_{obs} was low (<0.001 min^{-1}), its value was estimated from a linear fit to the time points.

The 2′,5′-branched product from the 9HR17 deoxyribozyme [from the DNA-rA-DNA positive control selection; see section below] was synthesized from 500 pmol L, 600 pmol E, and 750 pmol R. These oligonucleotides were annealed in 45 µL of 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 2 min and then cooling at room temperature for 45 min. To the sample was added 12 µL of 5× CHES buffer and 3 µL of 800 mM MgCl$_2$, making the final concentrations 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM MgCl$_2$. The sample was incubated at 37 °C for 3 h and purified by 20% denaturing PAGE. After 5′-^{32}P-radiolabeling, the product was used for analytical assays.

The product from the Tyr1 deoxyribozyme [identified in the DNA-Tyr-DNA selection; see section below] was synthesized from 1.0 nmol L, 1.2 nmol E, and 1.5 nmol R. These strands were annealed in 90 µL of 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 2 min and then cooling at room temperature for 45 min. To the sample was added 24 µL of 5× CHES buffer and 6 µL of 800 mM MgCl$_2$, making the final concentrations 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM MgCl$_2$. The sample was incubated at 37 °C for 20 h, precipitated with ethanol, and purified by 20% denaturing PAGE. The product was used directly for MALDI-TOF MS analysis or (after 5′-^{32}P-radiolabeling) for analytical assays.

The product from the Tyr13 deoxyribozyme [identified in the DNA-Tyr-DNA selection; see section below] was synthesized from 6 pmol 5′-^{32}P-radiolabeled L, 25 pmol E, and 75 pmol R. These strands were annealed in 15 µL of 5 mM HEPES, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 2 min and then cooling at room temperature for 30 min. To the sample was added 4 µL of 5× CHES buffer and 1 µL of 800 mM MgCl$_2$, making the final concentrations 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 40 mM MgCl$_2$. The sample was incubated at 37 °C for 20 h and purified by 20% denaturing PAGE.

The linear standards (DNA-X-DNA-RNA connectivity, where X = rA, dA, Tyr, or Phe) were synthesized from 3 pmol 5′-^{32}P-radiolabeled L, 20 pmol splint oligonucleotide, and 30 pmol 5′-phosphate-R (prepared from 5′-triphosphate-R by dephosphorylation with CIP and phosphorylation with ATP and T4 PNK). These oligonucleotides were annealed in 8 µL of 5 mM Tris, pH 7.5, 15 mM NaCl, and 0.1 mM EDTA by heating at 95 °C for 2 min and then cooling at room temperature for 30 min. To the sample was added 1 µL of 10× T4 DNA ligase buffer and 1 µL of T4 DNA ligase (1 U/µL, Fermentas). The sample was incubated at 37 °C for 4 h and purified by 20% denaturing PAGE. The splint sequence (5′-CCGTCGCCATCTCTTCCCGCAGTGAA-3′) was exactly complementary to R (underlined) and the 3′-terminal DNA portion of L.

Deoxyribozymes from selection with the DNA-rA-DNA substrate

Many clones from the DNA-rA-DNA selection were highly active for creating DNA-RNA branched linkages. The most active clone, designated 9HR17 (as the 17th clone from round 9 of the selection arbitrarily designated HR according to our laboratory nomenclature), has one potential secondary structure as shown in Figure S1a. Representative assays for individual active clones are shown in Figures S1b and S1c (branch-site rA and rC, respectively). A partial alkaline hydrolysis experiment that demonstrates the 2′,5′-branched connectivity of the 9HR17 product is shown in Figure S2.
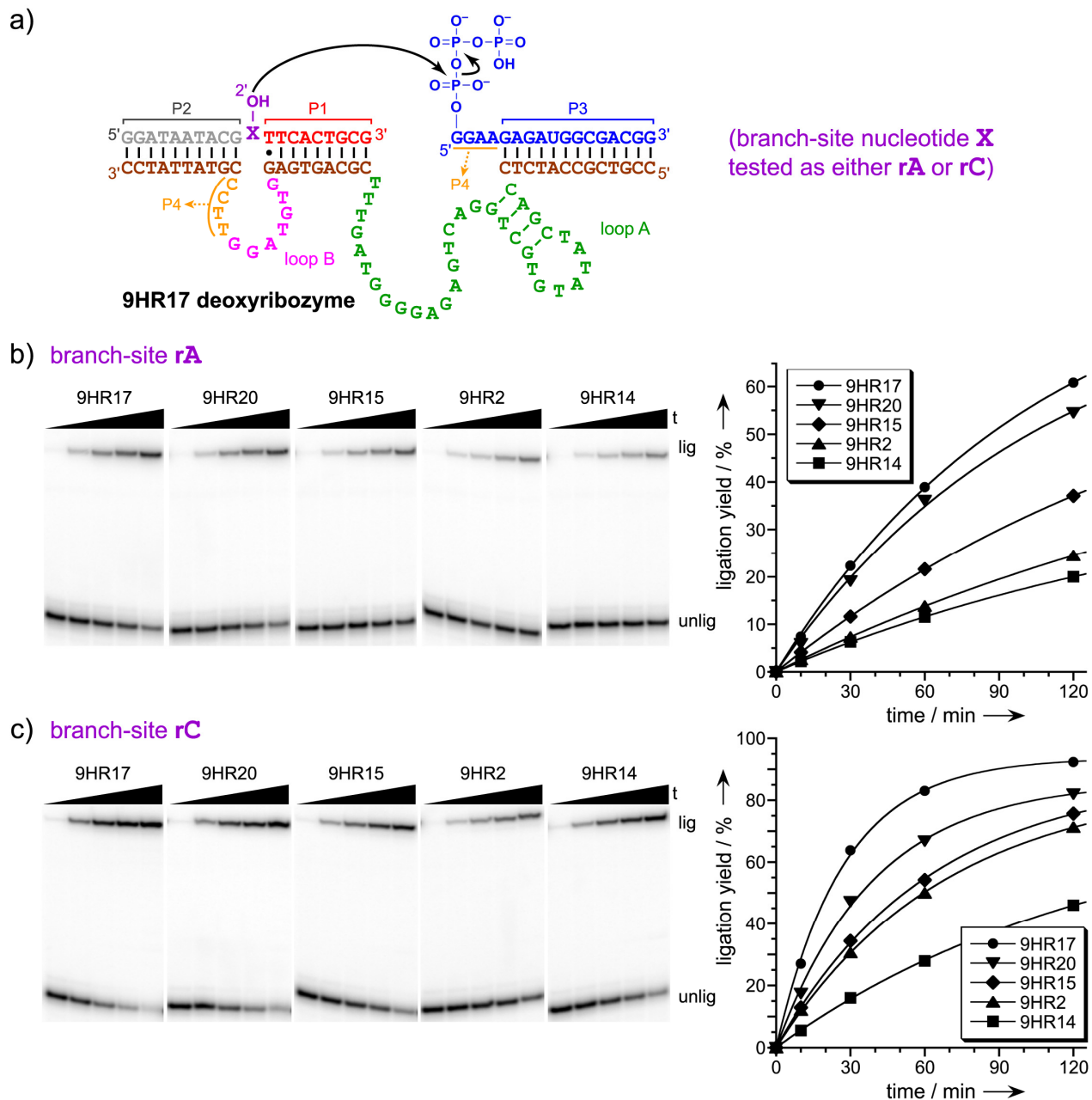
Figure S1. New deoxyribozymes that form DNA-RNA branches. a) One potential secondary structure of the 9HR17 deoxyribozyme, which has the highest ligation activity of all tested clones from this selection experiment. b) Assays for five individual 9HR clones with L substrate that has branch-site rA (t = 0.5, 10, 60, and 120 min; 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, 40 mM MgCl$_2$, 37 °C). k_{obs} values (top to bottom) 0.0094, 0.0093, 0.0054, 0.0042, 0.0053 min^{-1}. c) Assays for the same five 9HR clones with L substrate that has branch-site rC. k_{obs} values (top to bottom) 0.037, 0.025, 0.016, 0.015, 0.0076 min^{-1}.
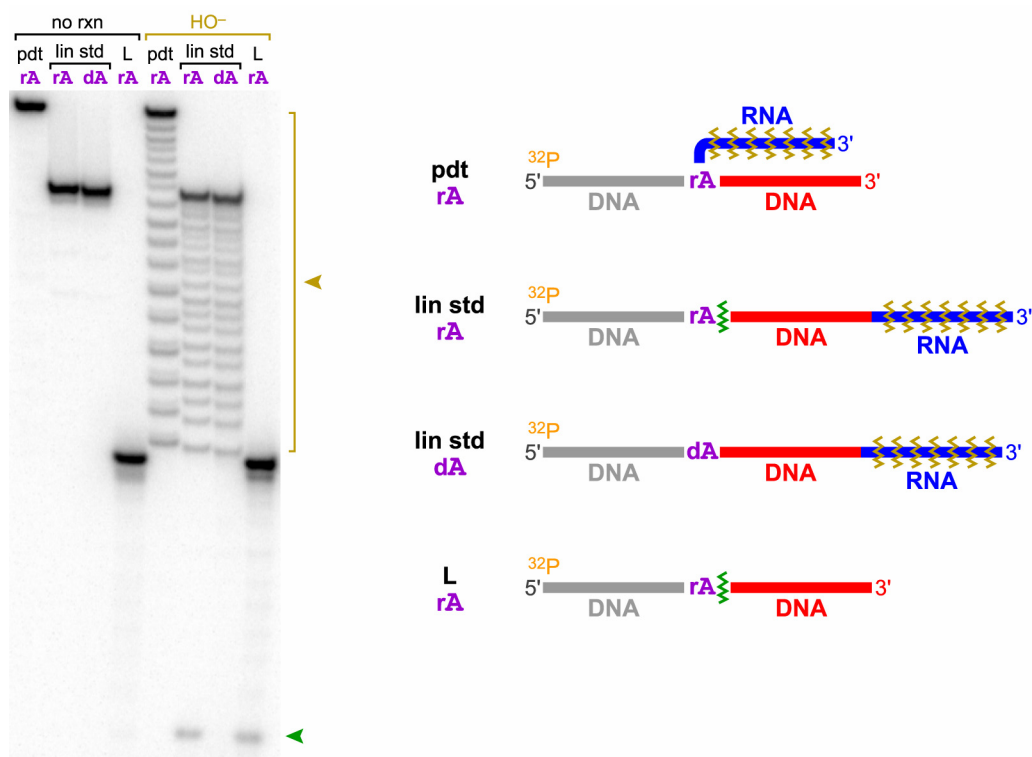
Figure S2. Partial alkaline hydrolysis demonstrates the 2′,5′-branched connectivity of the product formed by the 9HR17 deoxyribozyme using the branch-site rA substrate. Two sets of lanes are shown: the first four lanes with no treatment (no rxn), and the second four lanes with partial alkaline hydrolysis (HO⁻; 50 mM NaHCO₃, pH 9.2, 90 °C, 4 min). Within each set of four lanes, the included material is as illustrated on the right. As denoted with the yellow arrowhead and zigzag lines, the alkaline hydrolysis degrades only the RNA portion of each sample. Note that as expected, the degradation ladder in each case extends almost down to (but does not meet) the L standard, and no degradation ladder is observed for L itself ("L" stands for "left-hand substrate"). As denoted with the green arrowhead and zigzag line, cleavage at the rA embedded within the substrate results in a rapidly migrating band, but only when rA is both present and not involved in a 2′,5′-linkage (i.e., for "lin std rA" and "L rA" but not for "pdt rA" and "lin std dA").

Deoxyribozymes from selection with the DNA-Tyr-DNA substrate

Seventeen clones from the DNA-Tyr-DNA selection were analyzed for ligation activity. Fifteen of the clones synthesized the intended Tyr-RNA product, whereas two of the clones made the linear DNA-Tyr-DNA-RNA product (both as assessed from the migration rate of the product). All seventeen clones were sequenced. Of the 15 clones that synthesized Tyr-RNA, seven shared the Tyr1 sequence; five differed by only one nucleotide within the random regions; two differed by three nucleotides within the random regions; and one was an unidentifiable mixture of two sequences. Both of the clones that made the linear product shared the Tyr13 sequence.

One potential secondary structure of the Tyr1 deoxyribozyme complexed with its substrates is shown in Figure S3. Determination of $K_{d,app}$ for Mg^{2+} and for Mn^{2+} is shown in Figure S4.
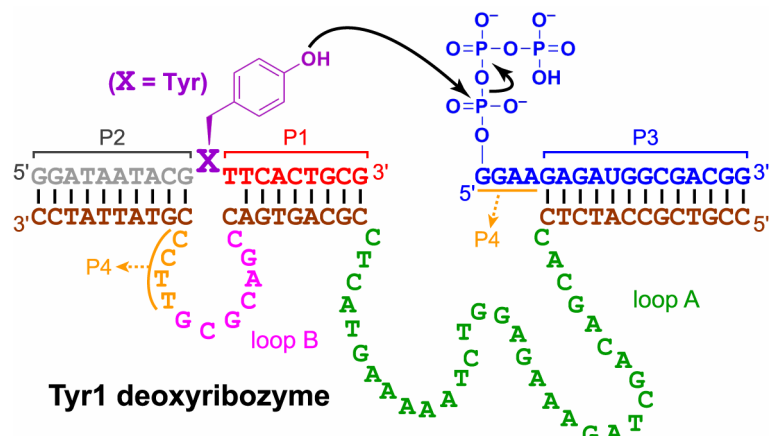

Figure S3. One potential secondary structure of the Tyr1 deoxyribozyme complexed with its substrates. Mfold predicts no strong secondary structure elements within loop A.
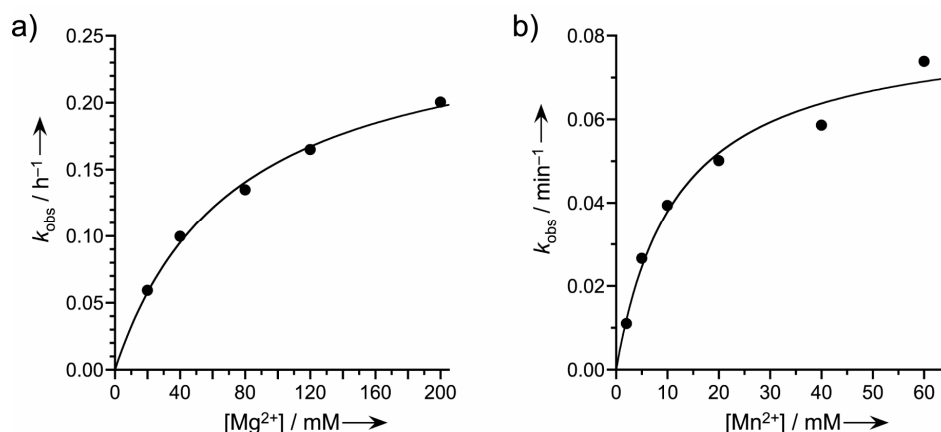

Figure S4. Determination of $K_{d,app}(Mg^{2+})$ and $K_{d,app}(Mn^{2+})$ for the Tyr1 deoxyribozyme using the DNA-Tyr-DNA substrate. a) Values of k_{obs} determined in 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, and 37 °C with various concentrations of MgCl$_2$. K_d,app = 62 ± 14 mM. b) Values of k_{obs} determined in 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl, and 37 °C with various concentrations of MnCl$_2$. K_d,app = 12 ± 3 mM.

One potential secondary structure of the Tyr13 secondary structure is shown in Figure S5a. Data showing that a 3′-phosphate on the DNA-X-DNA substrate entirely blocks the ligation reaction is shown in Figure S5b. Data supporting the conclusion that the 3HJ architecture is not formed by Tyr13 is shown in Figure S6.
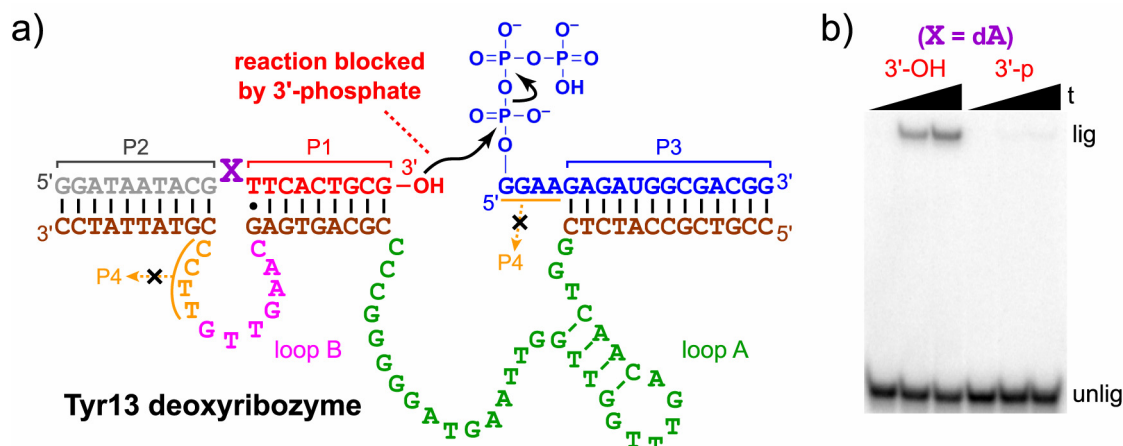
Figure S5. Tyr13 deoxyribozyme and support for the linear connectivity of its product. a) One potential secondary structure of the Tyr13 deoxyribozyme complexed with its substrates. Although the short (4 bp) P4 duplex could in principle form, the data in this figure and the next suggest that P4 does not actually form. See Figure 3b for ligation assays of Tyr13 with X = Tyr, Ala, Ser, rA, and dA. b) Experiment showing that a 3′-phosphate on the DNA-X-DNA (X = dA) substrate entirely suppresses Tyr13 activity, consistent with the linear nature of its DNA-X-DNA-RNA product. t = 0, 5, 10 h at 40 mM Mg^{2+}, pH 9.0, and 37 °C. Because T4 PNK (normally used for 5′-^{32}P-radiolabeling) removes the 3′-phosphate, the 5′-triphosphate-RNA substrate rather than the DNA-X-DNA substrate was 3′-radiolabeled using ^{32}P-pCp and T4 RNA ligase,[8] and the RNA was used as the limiting reagent in the assay (DNA-X-DNA:deoxyribozyme:RNA = 10:3:<1). The very faint band for the reaction with 3′-phosphate substrate was quantified as 0.4% yield in 10 h, versus 20% yield in 10 h for the 3′-OH substrate (ca. 50-fold difference in yield). The small amount of ligation product found for the 3′-phosphate substrate may reasonably be attributed to a small amount of hydrolysis of the 3′-phosphate after several hours at pH 9.0 and 37 °C.
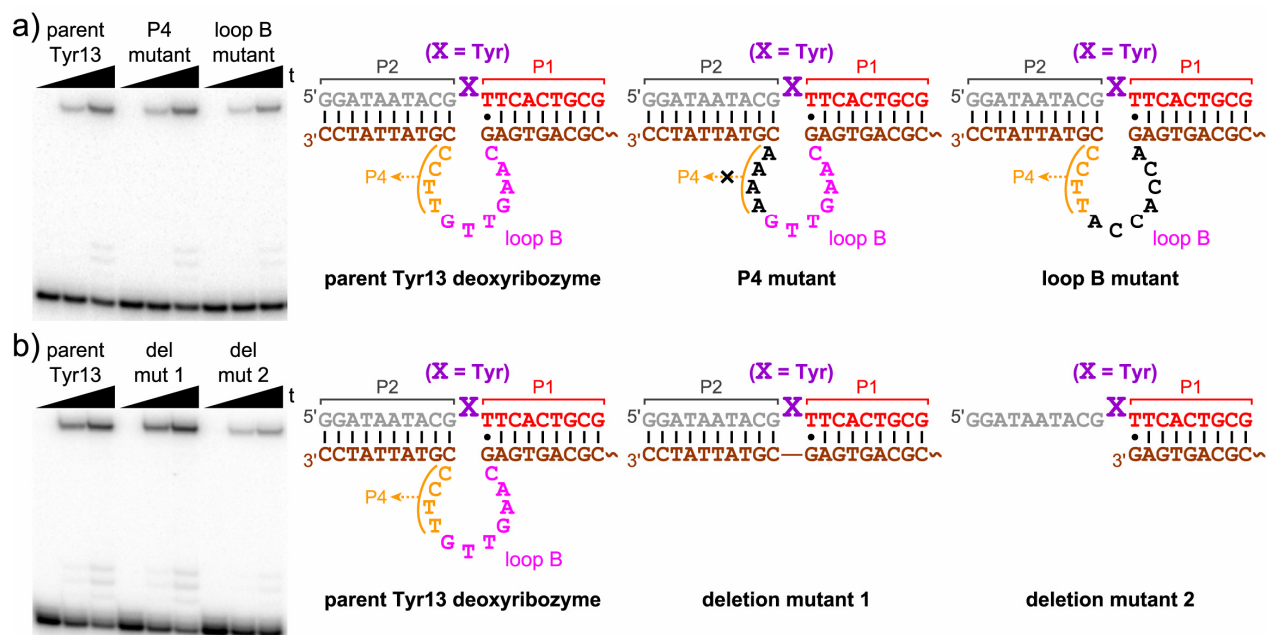

Figure S6. Mutants of the Tyr13 deoxyribozyme reveal that the 3HJ architecture is not intact because P4 is not formed. a) The parent deoxyribozyme and the P4 mutant have equivalent activities, while the loop B mutant has only slightly lower activity. The yields at final time points were 30%, 30%, and 17% (t = 0, 5, 17 h at 40 mM Mg^{2+}, pH 9.0, and 37 °C). Only the DNA-Tyr-DNA substrate and a portion of the deoxyribozyme are depicted here (see previous figure for full Tyr13 deoxyribozyme structure). The observation that mutation of either P4 or loop B has little or no effect on activity indicates that these regions do not contribute strongly to catalysis, which is incompatible with the 3HJ architecture. b) Deletion of the loop B and P4 regions has no effect, whereas additional deletion of the P2 binding arm has a modest effect (presumably due to decreased binding strength upon loss of P2). The yields at final time points were 46%, 49%, and 13% (t = 0, 8, 22 h at 40 mM Mg^{2+}, pH 9.0, and 37 °C).

Deoxyribozymes from selection with the DNA-Ser-DNA substrate
One potential secondary structure of the Ser7 deoxyribozyme is shown in Figure S7a. The very modest activity of Ser7 and migration of its product close to the position of the Tyr1 product are depicted in Figure S7b.
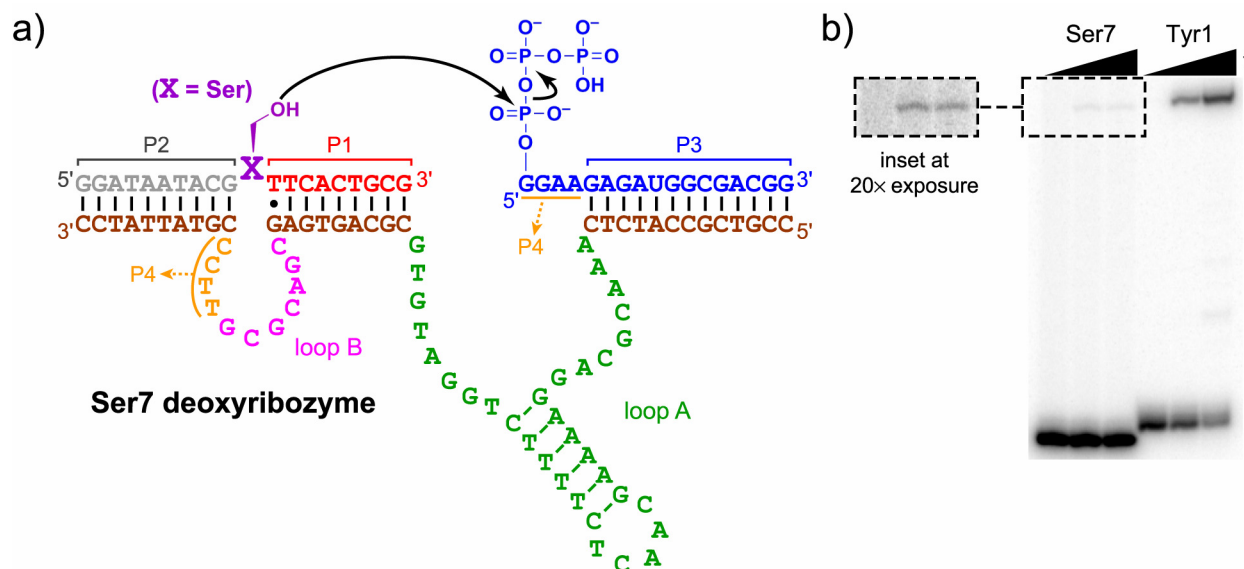

Figure S7. Ser7 deoxyribozyme. a) One potential secondary structure of Ser7. b) Timecourse of Ser7 activity, with Tyr1 for comparison (t = 0, 5, 20 h, 50 mM CHES, pH 9.0, 150 mM NaCl, 2 mM KCl, 40 mM MgCl$_2$, 37 °C). The Ser7 yield at 5 h is only 0.16% (k_{obs} = 5 × 10^{-6} min^{-1}), but this activity is reproducible.

MALDI-TOF mass spectrometry analysis of the product from the Tyr1 deoxyribozyme
The Tyr1 product for MALDI-MS analysis was synthesized as described above. The observed mass value was as expected for the assigned product: *m/z* calcd. 11660; found 11659 ± 12 (Figure S8).
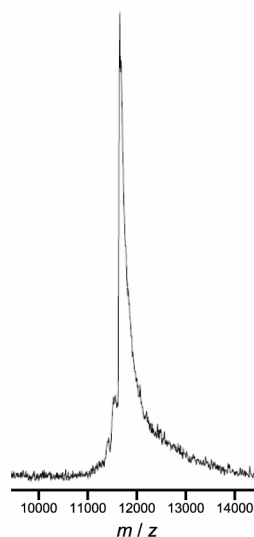

Figure S8. MALDI-TOF MS analysis of the Tyr1 deoxyribozyme product.

SVPD treatment of the Tyr1 product

Because snake venom phosphodiesterase 1 (SVPD) is well-known to cleave Tyr-RNA phosphodiester linkages,[9] we used SVPD to probe for the presence of this linkage in the Tyr1 product. Figure S9 shows an expanded version of the experiment in Figure 4a, along with schematic diagrams to explain the results. The 5′-^{32}P-radiolabeled DNA-X-DNA standards (X = dA or Tyr) and the appropriate Tyr1 or Tyr13 products were synthesized as described above. Each sample (0.1 pmol) was treated with SVPD (from *Crotalus atrox*, Sigma, 10^{-6} units) in 10 μL containing 50 mM Tris, pH 8.0, and 5 mM MgCl$_2$ at 4 °C. Time points were taken at t = 0.5, 2, 10, 30, and 120 min and analyzed by 20% denaturing PAGE.
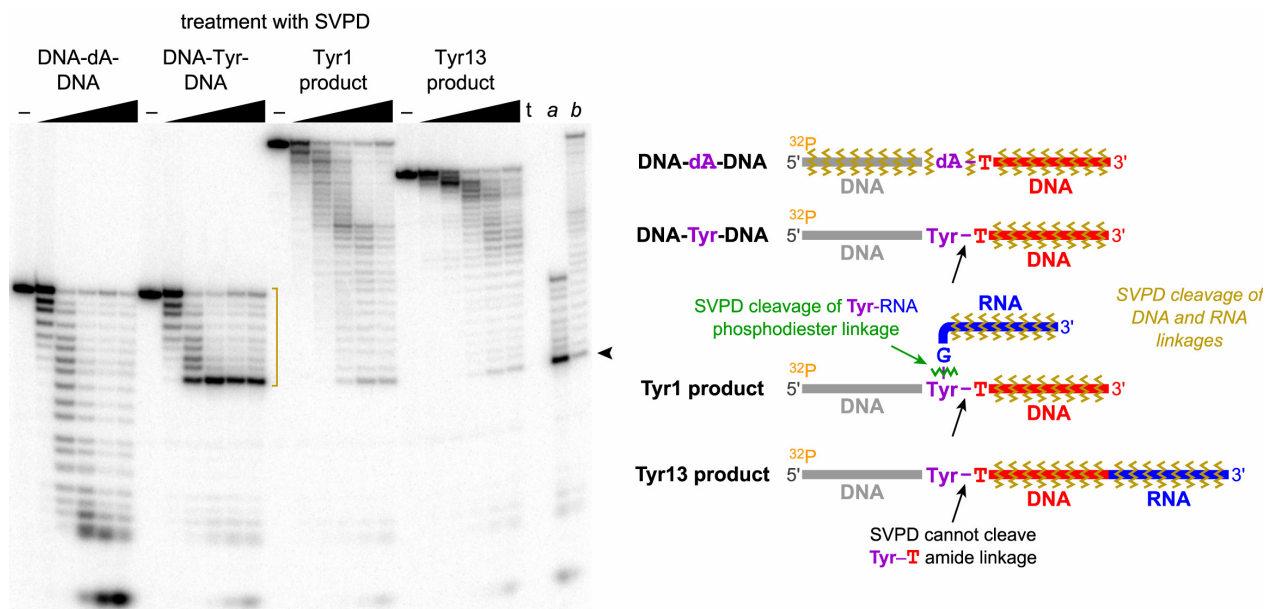

Figure S9. Expanded version of Figure 4a (SVPD treatment). The lanes marked *a* and *b* are the 120 min (final) time points from the DNA-Tyr-DNA and Tyr1 product experiments, respectively, clearly illustrating that the two SVPD degradation products comigrate as marked with the arrowhead. In the DNA-Tyr-DNA lanes, note that the expected number of cleavage bands are observed (as marked with the yellow bracket; 8 cleavage events). This corresponds to degradation of the 3′-terminal portion of the DNA (Tyr-TTCACTGCG-3′; Figure 1) from its 3′-end, while leaving the Tyr–T amide linkage uncleaved by SVPD.

Selective chemical reaction of the Tyr –OH sidechain with TAMRA-NHS reagent

For the experiment of Figure 4b, the starting materials for TAMRA-NHS reactions were prepared by RNase T1 pretreatment of the 5′-^{32}P-radiolabeled Tyr1 and Tyr13 products, which were synthesized as described above. Each sample (10 pmol) was treated with RNase T1 (Ambion, 25 U) in 20 µL containing 50 mM Tris, pH 7.4, and 1 mM EDTA at room temperature for 80 min. The samples were purified by 20% denaturing PAGE. Each RNase T1-pretreated product (or DNA-X-DNA where X = Tyr or Phe) was then incubated with TAMRA-NHS [Invitrogen C1171, 5(6)-TAMRA SE], prepared as a 25 mM stock solution in DMSO. Each sample (0.1 pmol) was incubated in 5 µL containing 100 mM HEPES, pH 7.0, 3 mM imidazole, 10% DMSO, and 2.5 mM TAMRA-NHS at room temperature.[10] Time points were taken at t = 0.5, 30, and 60 min and analyzed by 20% denaturing PAGE. Figure S10 shows an expanded version of the experiment in Figure 4b, along with schematic diagrams to explain the results.
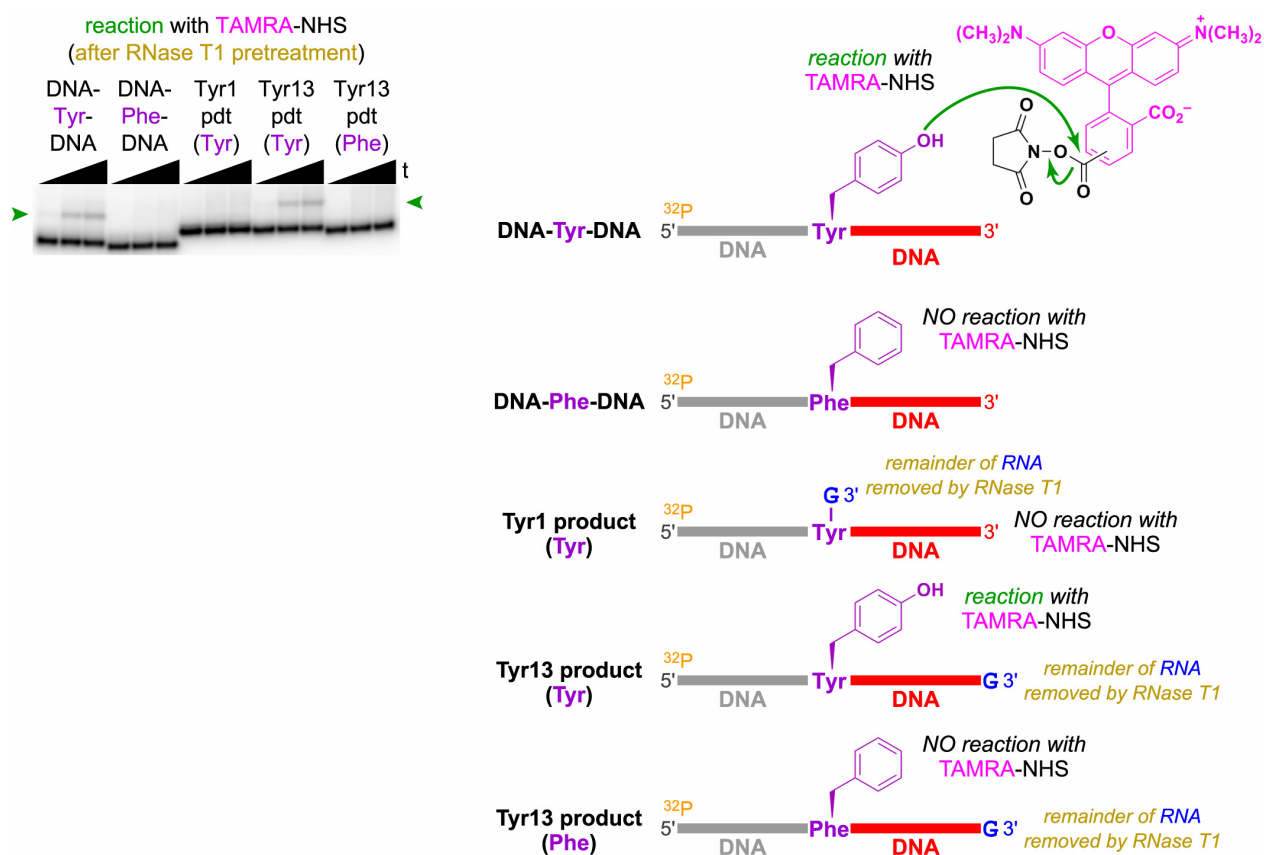

Figure S10. Expanded version of Figure 4b (reaction with TAMRA-NHS after RNase T1 pretreatment), including schematic diagrams to explain the results.

The T–C mismatch adjacent to the reactive Tyr contributes only modestly to Tyr1 activity

The initial deoxyribozyme strand placed a G across from the T that is adjacent to X in the DNA-X-DNA substrate (Figure S11a; circled T•G wobble pair; compare Figure 1). The Tyr1 deoxyribozyme identified after selection had this T•G wobble pair mutated to a T–C mismatch. Using either Mg^{2+} (Figure S11b) or Mn^{2+} (Figure S11c), we showed that the interaction between these two nucleotides contributes only modestly to the Tyr1 activity, because the largest difference in k_{obs} was only a factor of 3 for Mg^{2+} and a factor of 1.5 for Mn^{2+}.
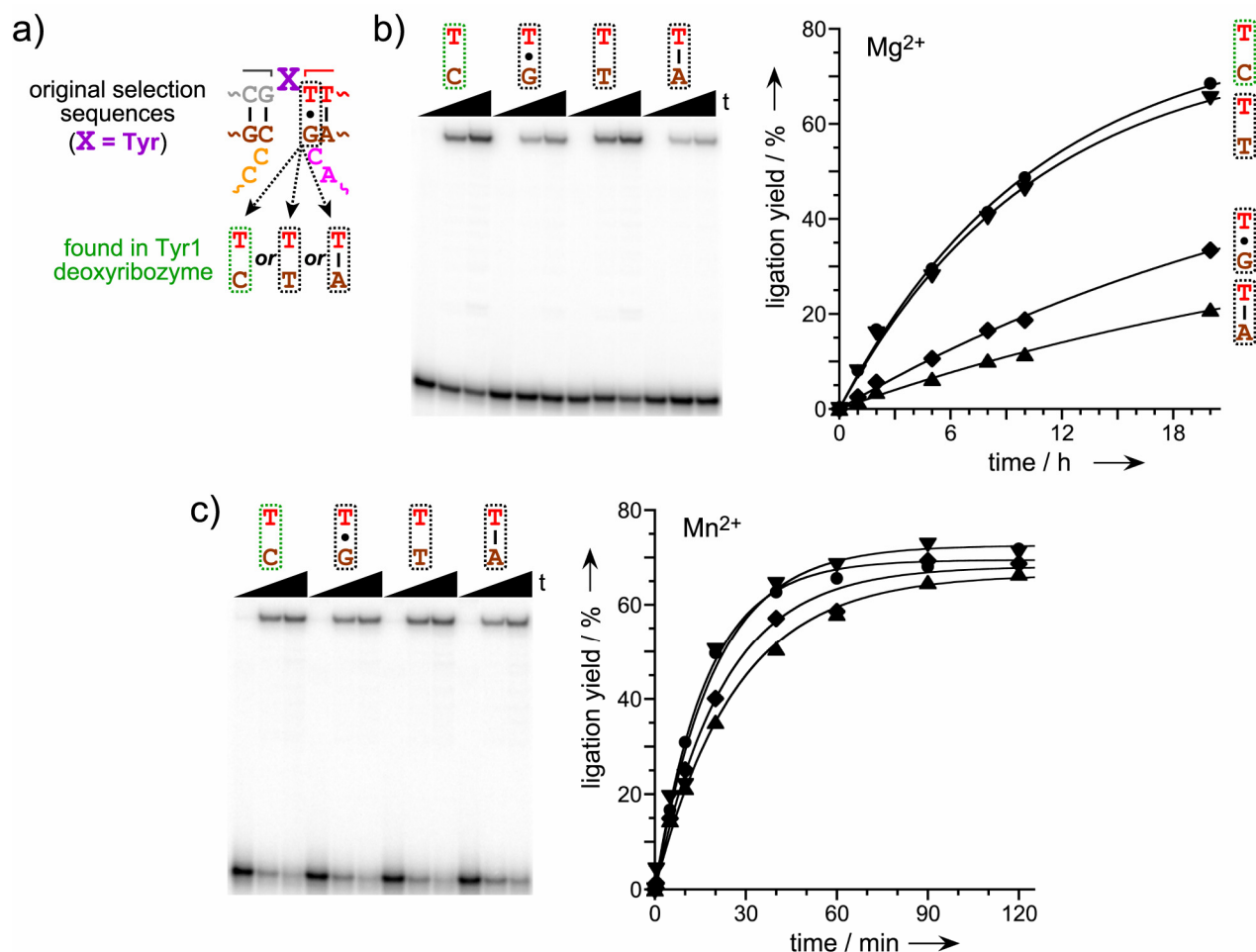

Figure S11. The T–C mismatch adjacent to the reactive Tyr contributes only modestly to Tyr1 activity. a) Structural context; compare to upper part of Figure 1. The circled T•G wobble pair was present at the outset of selection, whereas the T–C mismatch was found in Tyr1. b) Results for all four nucleotide combinations in 40 mM Mg^{2+}, pH 9.0, 37 °C (t = 0, 5, 10 h). Top to bottom in the plot, k_{obs} values were 0.094, 0.095, 0.036, and 0.032 h^{-1}. c) Results for all four nucleotide combinations in 20 mM Mn^{2+}, pH 7.5, 37 °C (t = 0.5, 20, 60 min). Yields are plotted using the same symbols as in panel a; k_{obs} values were 0.059, 0.051, 0.044, and 0.038 min^{-1}.

References for Supporting Information

[1] J. F. Milligan, D. R. Groebe, G. W. Witherell, O. C. Uhlenbeck, *Nucleic Acids Res.* **1987**, *15*, 8783-8798.

[2] A. Flynn-Charlebois, Y. Wang, T. K. Prior, I. Rashid, K. A. Hoadley, R. L. Coppins, A. C. Wolf, S. K. Silverman, *J. Am. Chem. Soc.* **2003**, *125*, 2444-2454.

[3] Y. Wang, S. K. Silverman, *Biochemistry* **2003**, *42*, 15252-15263.

[4] S. Narayanan, A. H. Dalpke, K. Siegmund, K. Heeg, C. Richert, *J. Med. Chem.* **2003**, *46*, 5031-5044.

[5] S. Zaramella, E. Yeheskiely, R. Stromberg, *J. Am. Chem. Soc.* **2004**, *126*, 14029-14035.

[6] F. Bergmann, W. Bannwarth, *Tetrahedron Lett.* **1995**, *36*, 1839-1842.

[7] P. Mamos, C. Sanida, K. Barlos, *Liebigs Ann. Chem.* **1988**, 1083-1084.

[8] D. A. Baum, J. Sinha, S. M. Testa, *Biochemistry* **2005**, *44*, 1067-1077.

[9] a) V. Ambros, D. Baltimore, *J. Biol. Chem.* **1978**, *253*, 5263-5266; b) P. G. Rothberg, T. J. Harris, A. Nomoto, E. Wimmer, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4868-4872; c) S. P. Adler, D. Purich, E. R. Stadtman, *J. Biol. Chem.* **1975**, *250*, 6264-6272; d) Y.-C. Tse, K. Kirkegaard, J. C. Wang, *J. Biol. Chem.* **1980**, *255*, 5560-5565; e) B. Juodka, M. Pfötz, D. Werner, *Nucleic Acids Res.* **1991**, *19*, 6391-6398.

[10] Y. Yu, A. J. Hoffhines, K. L. Moore, J. A. Leary, *Nat. Methods* **2007**, *4*, 583-588.