Electronic Supplementary Material (ESI) for ChemComm.

Supporting information to:

Assembly of split aptamers by a dynamic pH-responsive covalent ligation

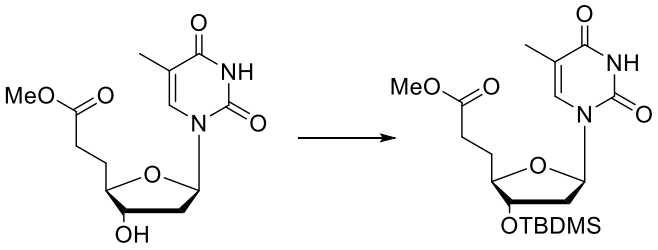
Aapo Aho and Pasi Virta* (pamavi@utu.fi)

Department of Chemistry, University of Turku, 20500 Turku, Finland

Contents

Synthesis of small molecules

Synthesis of 3



2 **3**

Nucleoside **2** (173 mg, 0.58 mmol, see synthesis on reference S1) was dissolved in dry DMF (5.0 ml). To the solution was added TBDMSCl (105 mg, 0.70 mmol) and TEA (242 µl, 1.74 mmol) and a small crystal of DMAP. The reaction was mixed in room temperature overnight after which TLC analysis (4% MeOH in DCMor) indicated the reaction was partially complete. More TBDMSCl (75 mg, 0.5 mmol) and TEA (150 µl, 1.08 mmol) was added to the mixture. After stirring the mixture for overnight, the TLC indicated complete transformation of the starting material into a non-polar product. The reaction mixture was evaporated in a vacuum into a syrup, which was extracted using ethyl acetate (100 ml). The organic phase was washed with $NaHCO_3$ (2 × 25 ml) and brine (25 ml). The organic phase was then dried using Na_2SO_4, filtered, evaporated to dryness, and purified by silica column chromatography (2.5% MeOH in DCM). The resulting product (157 mg, 65%) was amber and highly viscous oil. **^{1}H NMR** ($CDCl_3$, 600 MHz): δ 9.43 (1H, br, NH), 7.14 (1H, d, J = 1.2 Hz, H6), 6.18 (1H, t, J = 6.0 Hz, H1´), 4.08 (1H, dt, J_1 = 6.6 Hz, J_2 = 4.8 Hz, H3´), 3.78 (1H, dt, J_1 = 9.6 Hz, J_2 = 4.8 Hz, H4´), 3.68 (3H, s, OMe), 2.49 (2H, m, H7), 2.59 (1H, ddd, J_1 = 13.6 Hz, J_2 = 6.6 Hz, J_3 = 4.2 Hz, H2´), 2.07 (1H, ddd, J_1 = 13.6 Hz, J_2 = 6.6 Hz = 6.6 Hz, H2´´), 2.04 (1H, m, H5´), 1.94 (3H, d, J = 1.2 Hz, MeT), 1.90 (1H, m, H5´´), 0.89 (9H, s, TBDMS), 0.08 (3H, s, TBDMS), 0.07 (3H, s, TBDMS). **^{13}C NMR** ($CDCl_3$): δ 173.5 (C8), 163.9 (C4), 150.3 (C2), 135.3 (C6), 111.1 (C5), 85.5 (C4´), 84.7 (C1´), 74.8 (C3´), 51.7 (OMe), 40.5 (C2´), 30.5 (C7), 28.5 (C5´), 25.7 (TBDMS), 17.9 (TBDMS), 12.6 (MeT), –4.5 (TBDMS), –4.8 (TBDMS). **HRMS-ESI** (*m*/*z*) calc. for $C_{19}H_{32}N_2NaO_6Si$ [M + Na^+]: 435.1927; Found: 435.1947.

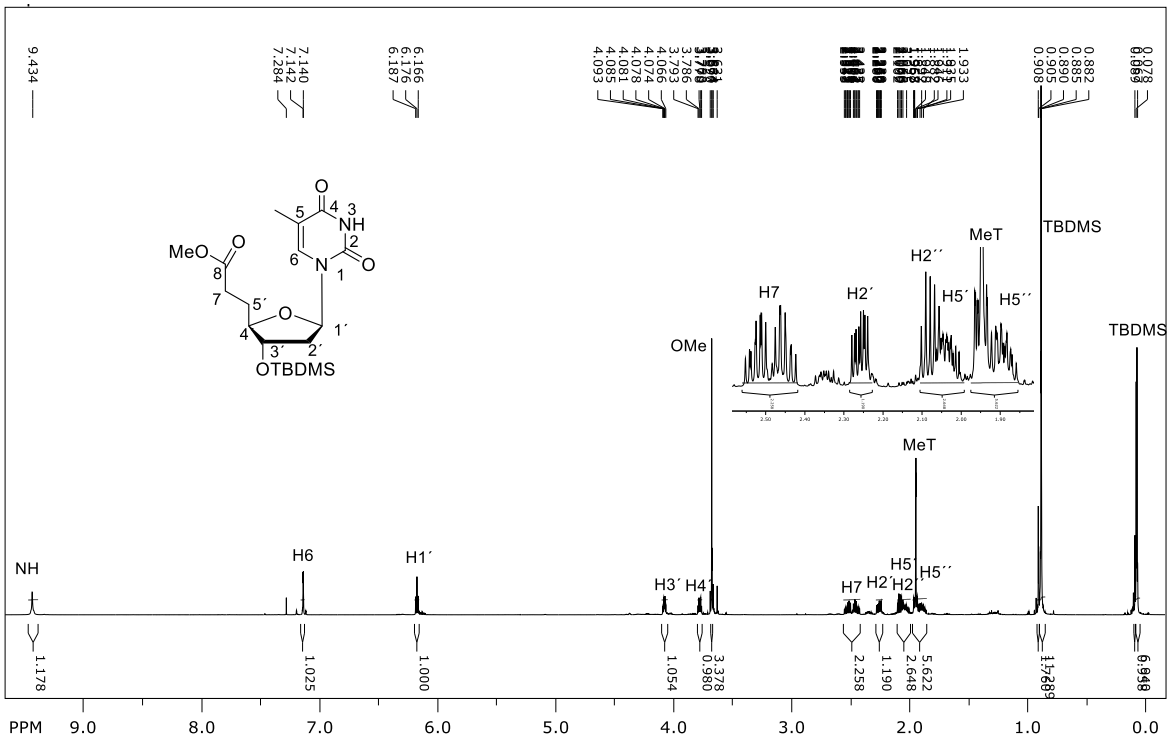


Figure S1 ^{1}H NMR (600 MHz, $CDCl_3$) spectrum of **3**.

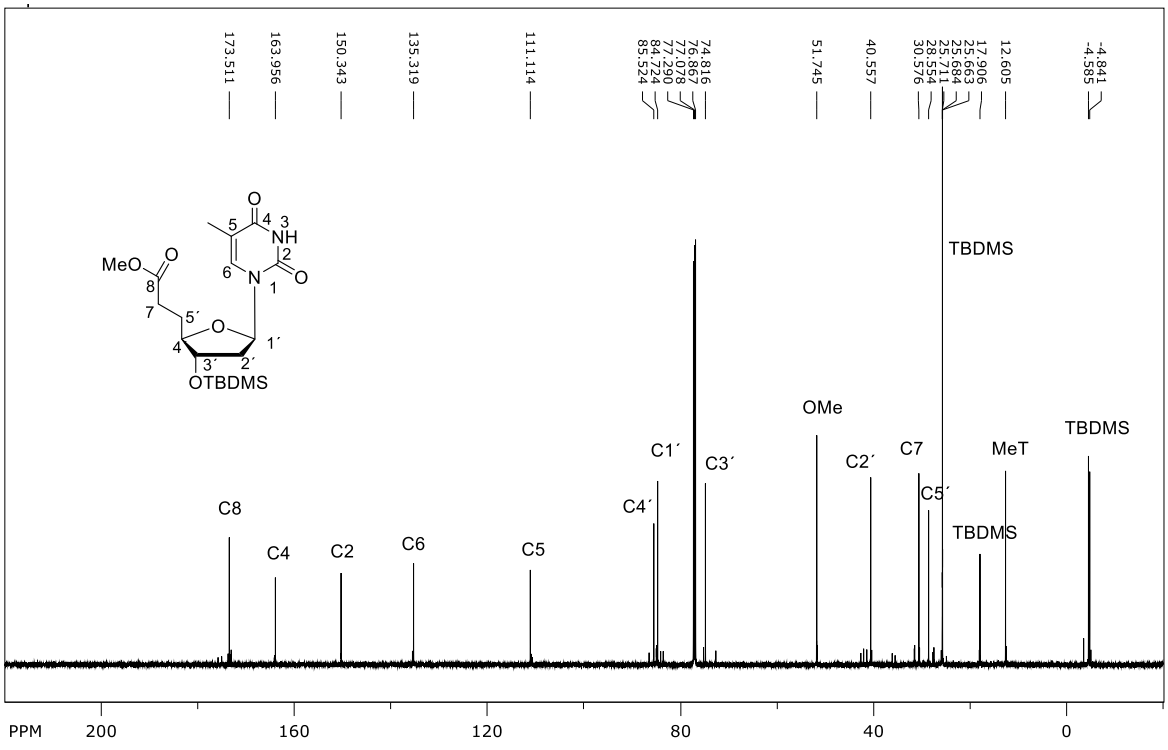
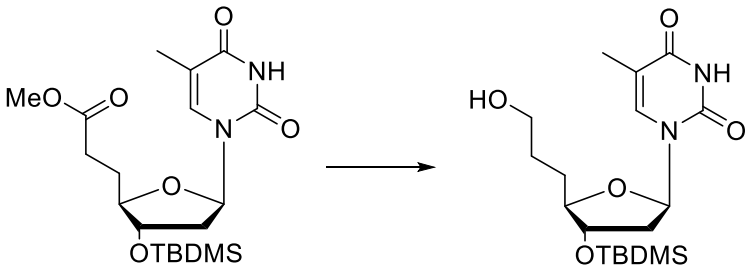


Figure S2 ^{13}C NMR (150 MHz, $CDCl_3$) spectrum of **3**.

Synthesis of 4



3 **4**

Nucleoside **3** (157 mg, 0.38 mmol) was dissolved in dry THF (4.0 ml) under nitrogen and cooled down to 0 °C using ice–water-bath. The mixture was stirred vigorously using a magnetic stirrer during which was added dropwise a 1.0 M solution of $LiAlH_4$ (390 µl, 0.39 mmol) over a period of 30 min. 30 minutes after final addition, the reaction was monitored using TLC (5% MeOH in DCM), which indicated the reaction was partially complete. More of the 1.0M solution of $LiAlH_4$ (100 µl, 0.10 mmol) was added to the mixture during a period of 15 min. After addition, TLC indicated a complete consumption of the starting material into a more polar product. The reaction was quenched carefully by addition of cold H_2O (11 ml). The quenched reaction mixture was extracted using ethyl acetate (3 × 30 ml). The organic phase was washed using aq. 10% NH_4Cl (20 ml) and brine (20 ml) and then dried using Na_2SO_4, filtered, evaporated to dryness, and purified by silica column chromatography (4.5% MeOH in DCM). The resulting product (91 mg, 62%) was amorphous colorless solid. **^{1}H NMR** ($CDCl_3$, 500 MHz): δ 8.95 (1H, br, NH), 7.16 (1H, d, *J* = 1.0 Hz, H6), 6.19 (1H, t, *J* = 6.6 Hz, H1′), 4.10 (1H, dt, J_1 = 6.7 Hz, J_2 = 4.2 Hz, H3′), 3.83 (1H, ddd, J_1 = 8.0 Hz, J_2 = J_3 = 4.2 Hz, H4′), 3.73 (2H, m, H8), 2.28 (1H, ddd, J_1 = 13.6 Hz, J_2 = 6.8 Hz, J_3 = 4.2 Hz, H2′), 2.10 (1H, ddd, J_1 = 13.6 Hz, J_2 = J_3 = 6.8 Hz, H2′′), 1.95 (3H, d, *J* = 1.0 Hz, MeT), 1.86–1.65 (3H, m, H5′ & H5′′), 1.86–1.65 (3H, m, H7 & H7′), 0.91 (9H, s, TBDMS), 0.10 (3H, s, TBDMS), 0.09 (3H, s, TBDMS). **^{13}C NMR** ($CDCl_3$): δ 163.7 (C4), 150.2 (C2), 135.4 (C6), 111.1 (C5), 86.5 (C4′), 84.9 (C1′), 75.0 (C3′), 62.4 (C8), 40.5 (C2′), 30.0 (C7), 29.2 (C5′), 25.7 (TBDMS), 17.9 (TBDMS), 12.7 (MeT), –4.5 (TBDMS), –4.8 (TBDMS). **HRMS-ESI** (*m*/*z*) calc. for $C_{18}H_{33}N_2O_5Si$ [M + H^+]: 385.2153; Found: 385.2100.

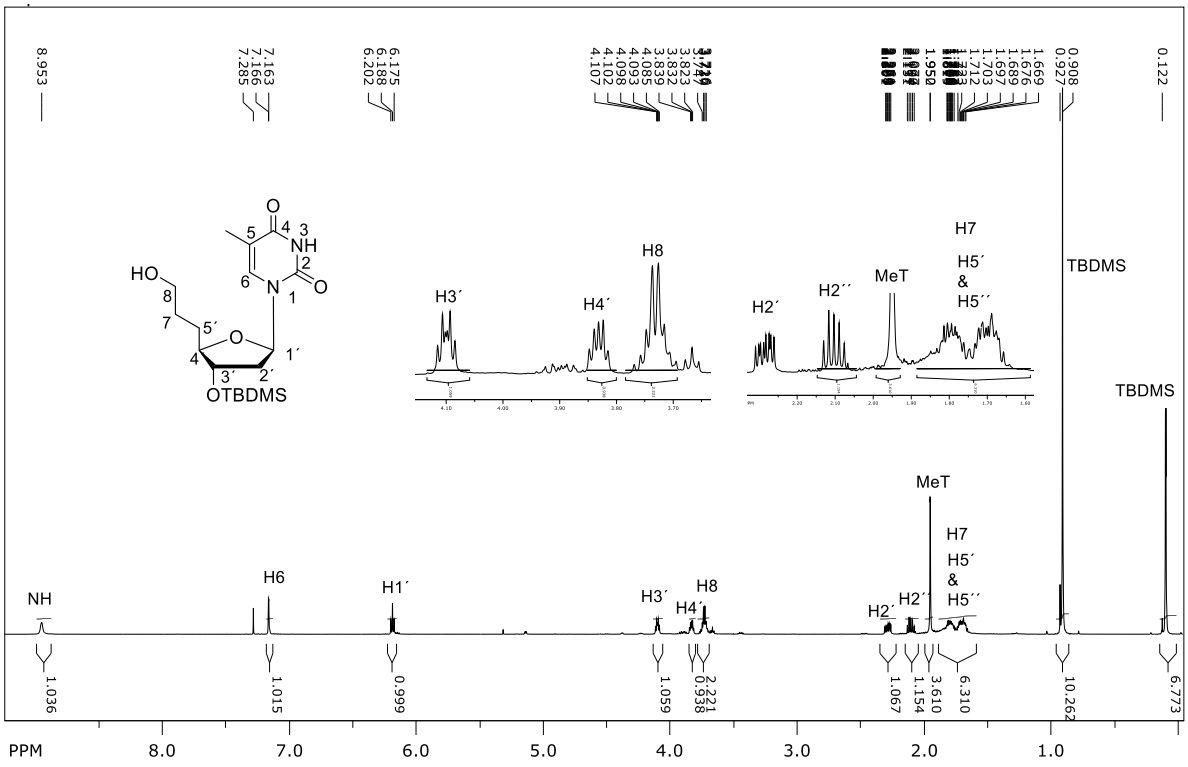


Figure S3 ^{1}H NMR (500 MHz, $CDCl_3$) spectrum of **4**.

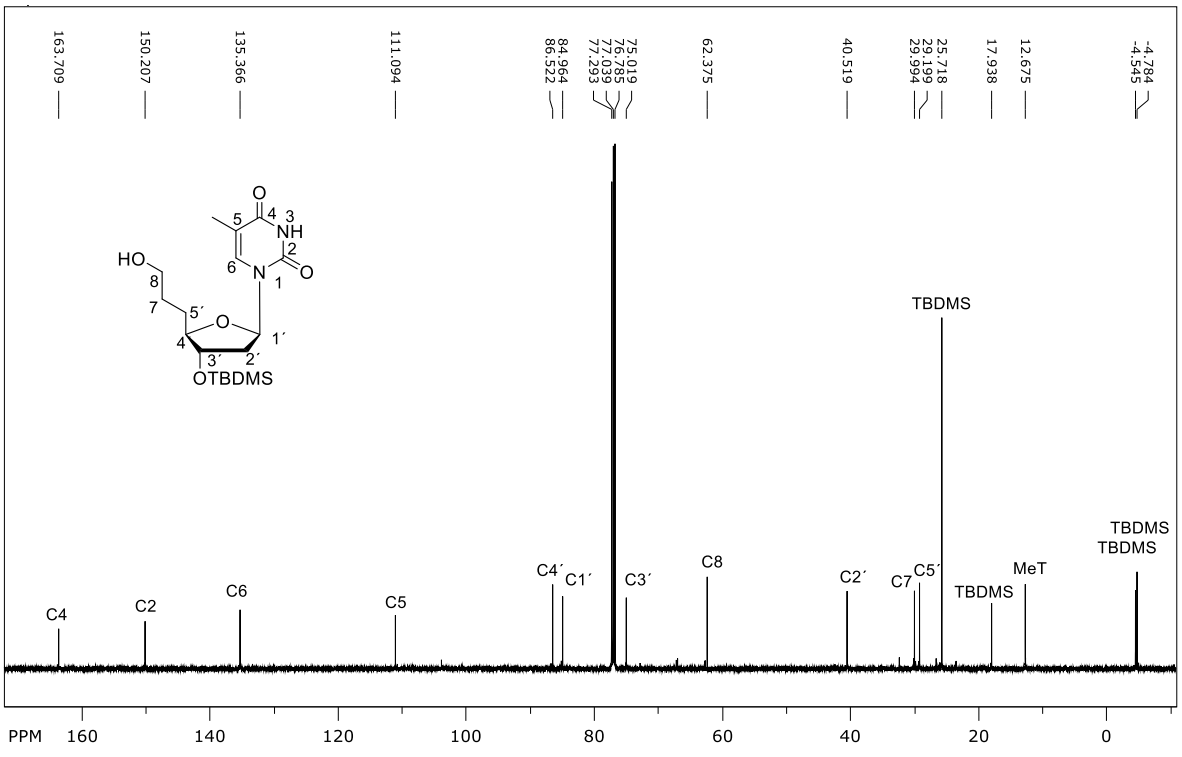
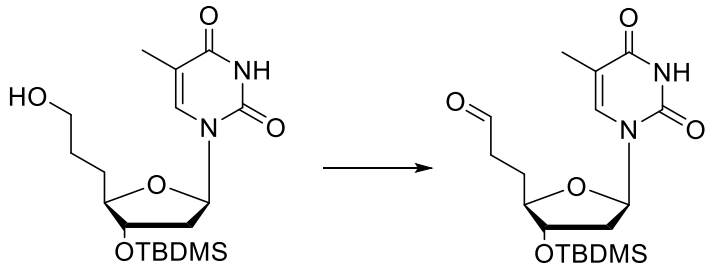


Figure S4 ^{13}C NMR (125 MHz, $CDCl_3$) spectrum of **4**.

Synthesis of 6



4 **5**

Nucleoside **4** (0.90 g, 2.37 mmol) was dissolved in dry dioxane (30 ml) under argon. The solution was bubbled with argon gas for 20 min. To the solution was added 2-iodoxybenzoic acid (2.0 g of 45%wt in stabilizer (benzoic acid, isophtalic acid), 3.1 mmol) and DMSO (0.40 ml, 5.70 mmol), and a stirring magnet. The reaction vessel was tightly sealed and placed in an oil bath (70 °C) and mixed with a magnetic stirrer. After two hours the reaction was monitored by LC/MS (C18, 150 × 4.6 mm, 5 µm) a linear gradient of 0–100% MeCN over 30 min) which indicated the starting material was fully consumed. The reaction was cooled using ice bath and diluted with a mixture (1:1, *v*/*v*) of MeCN/H_2O (20.0 ml). The mixture was vacuum-filtered using a size 4 glass frit (10–16 µm pore size). The filtered solids were washed with a mixture (10:2, *v*/*v*) of MeCN/H_2O (8 ml). The filtrate was evaporated to dryness using a vacuum. The evaporation residue was suspended in ethyl acetate (100 ml) and filtered again. The filtrate was washed with H_2O (4 × 25 ml), sat. aq. Na_2HCO_3 (2 × 25 ml), and brine (2 × 25 ml). The organic phase was dried using Na_2SO_4 and evaporated to dryness. The product (0.89 g, 99%) was slightly yellow amorphous solid and produced a single spot (R_f = 0.5) on TLC analysis (6% MeOH in DCM). **^{1}H NMR** ($CDCl_3$, 500 MHz): δ 10.30 (1H, br, NH), 9.81 (1H, s, H8), 7.14 (1H, d, J = 1.1 Hz, H6), 6.15 (1H, t, J = 6.5 Hz, H1′), 4.10 (1H, dt, J_1 = 7.0 Hz, J_2 = 4.6 Hz, H3′), 3.77 (1H, ddd, J_1 = 9.0 Hz, J_2 = J_3 = 4.6 Hz, H4′), 2.65 (2H, m, H7), 2.27 (1H, ddd, J_1 = 13.6 Hz, J_2 = 6.7 Hz, J_3 = 4.7 Hz, H2′), 2.15 (1H, ddd, J_1 = 13.6 Hz, J_2 = J_3 = 6.8 Hz, H2′′), 2.03 (1H, m, H5′), 1.95 (3H, dd, J = 1.1 Hz, MeT), 1.91 (1H, m, H5′′), 0.89 (9H, s, TBDMS), 0.08 (3H, s, TBDMS), 0.08 (3H, s, TBDMS). **^{13}C NMR** ($CDCl_3$): δ 201.5 (C8), 164.8 (C4), 150.5 (C2), 135.9 (C6), 111.2 (C5), 85.4 (C4′), 85.0 (C1′), 74.7 (C3′), 40.5 (C2), 40.4 (C7), 25.4 (C5′), 17.9 (TBDMS), 12.6 (MeT), –4.6 (TBDMS), –4.8 (TBDMS). **HRMS-ESI** (*m*/*z*) calc. for $C_{18}H_{30}N_2O_5Si$ [M + H^+]: 383.1997; Found: 383.1958.

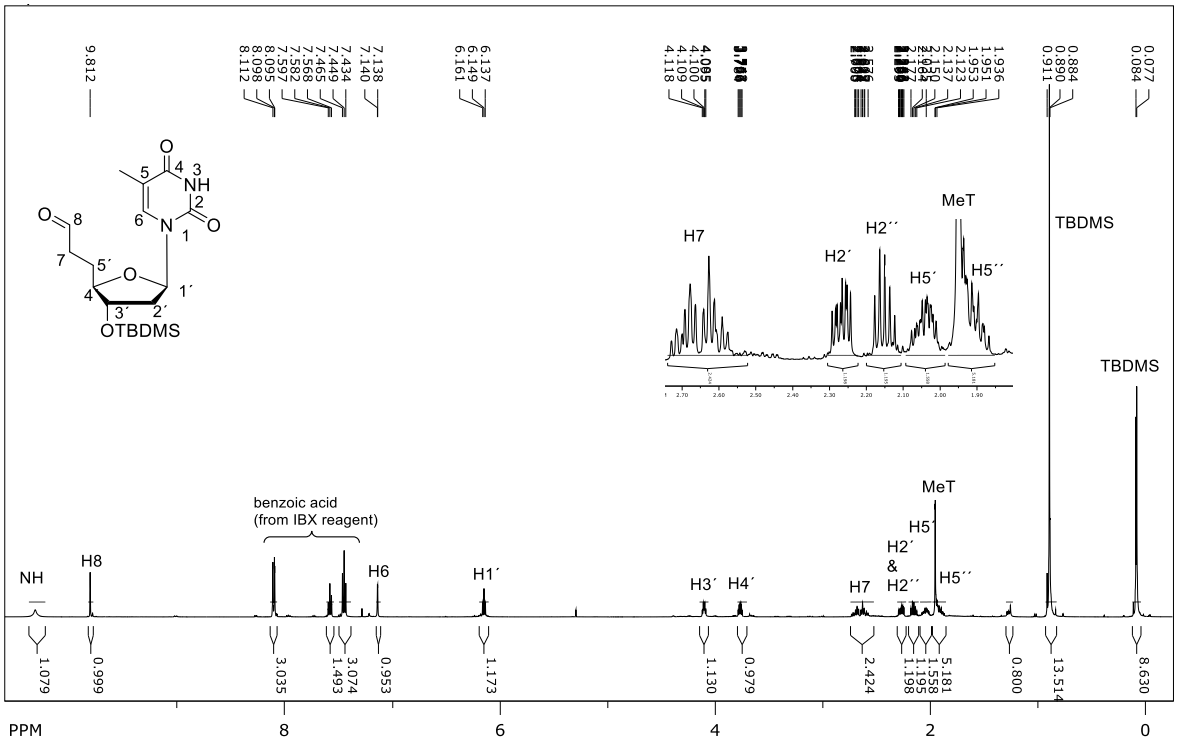


Figure S5 ^{1}H NMR ($CDCl_3$, 500 MHz) spectrum of **5**.

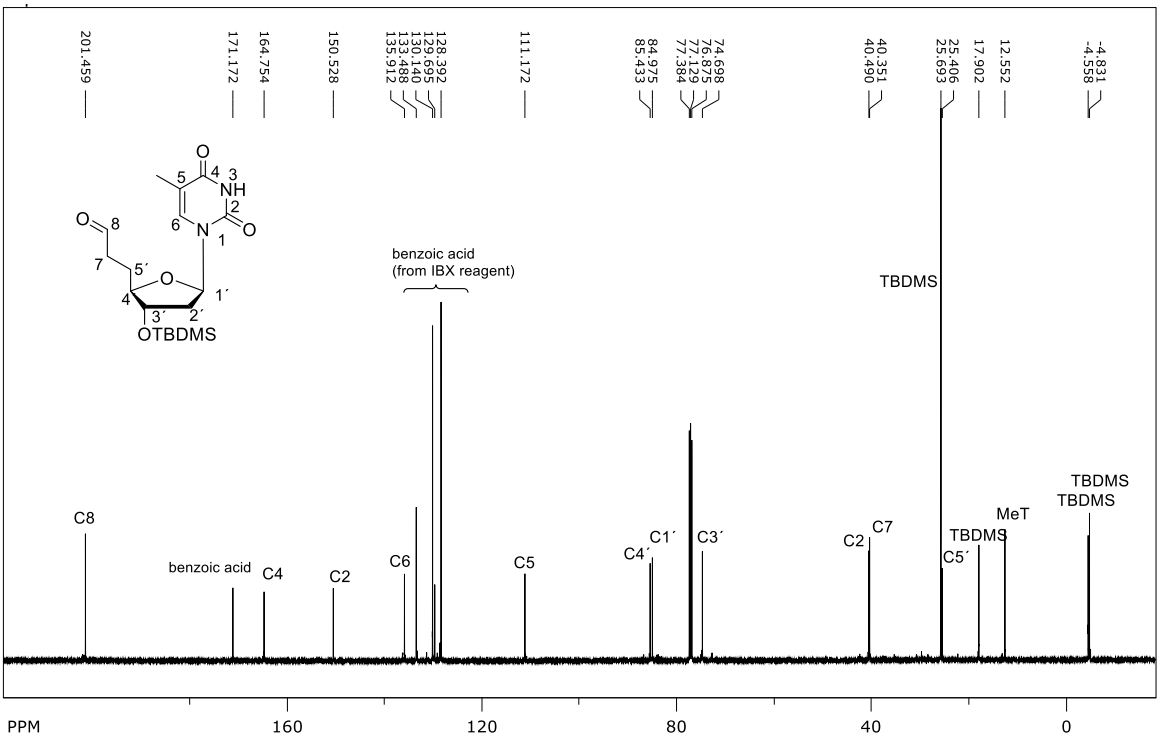
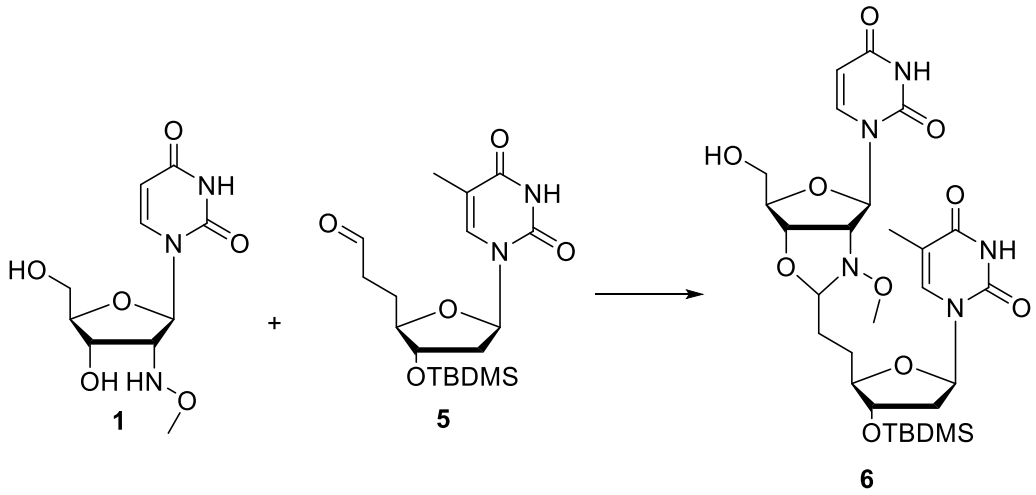


Figure S6 ^{13}C NMR ($CDCl_3$, 125 MHz) spectrum of **5**.

Synthesis of 6



Nucleosides **1** (0.31 g, 1.15 mmol) and **5** (0.44 g, 1.15 mmol) were dissolved in a mixture of DMSO/AcOH (3:1, *v*/*v*, 2 ml). The mixture was placed in an oil bath (40 °C) and stirred with a magnetic stirrer. The reaction was monitored by TLC, which indicated starting materials were consumed after 30 min. The reaction mixture was quenched by cooling on an ice bath and adding aq. sat. ammonia (2.0 ml) and H_2O (20 ml). The reaction mixture was extracted using ethyl acetate (3 × 50 ml), and the organic phase was washed with aq. sat. $NaHCO_3$ (25 ml) and brine (25 ml). The organic phase was dried using Na_2SO_4, filtered, and evaporated to dryness. The crude product was purified by silica column chromatography (0.1% TEA and 5.0% MeOH in DCM). The resulting product (0.56 g, 76%) was a white solid foam. A mixture of isomers in 2:1 ratio was obtained determined by the ratio of uracil H6 NMR signals. **^{1}H NMR** (CD_3CN, 600 MHz): δ 11.15 (1H, br, NH), 7.82 (1H, d, *J* = 8.2 Hz, H6U minor) 7.73 (1H, d, *J* = 8.1 Hz, H6U major), 7.30 (1H, s, H6T), 6.18 (1H, t, *J* = 6.8 Hz, H1´T), 6.02 (1H, br, H1´U minor) 5.72 (1H, d, *J* = 8.1 Hz, H5U), 5.70 (1H, d, *J* = 8.6 Hz, H1´U major), 4.70 (1H, dd, J_1 = 7.3 Hz, J_2 = 3.6 Hz, H3´U), 4.62 (1H, br, H8), 4.29 (1H, m, H2´U), 4.25 (1H, m, H3´T), 4.00 (1H, m, H4´U), 3.82–3.64 (3H, m, H4´T & H5´U & H5´U´), 3.54 (3H, s, OMe minor), 3.50 (3H, s, OMe major), 2.19 (2H, m, H2´T), 2.00 (2H, br, H7), 1.87 (3H, m, MeT), 1.87 (1H, m, H5´T), 1.78 (1H, m, H5´T´), 0.92 (9H, s, TBDMS), 0.13 (3H, s, TBDMS), 0.11 (3H, s, TBDMS). **^{13}C NMR** (CD_3CN, 150 MHz): 164.1 (C4T), 163.4 (C4U), 150.7 (C2T), 150.6 (C2U), 141.6 (C6U), 141.5 (C6U), 136.1 (C6T), 136.0 (C6T), 110.5 (C5T), 110.5 (C5T), 102.4 (C5U), 102.2 (C5U), 101.5 (C8), 95.9 (C8), 88.2 (C1´U), 86.1 (C4´T), 86.0 (C4´U), 84.2 (C1´T), 84.1 (C1´T), 78.7 (C3´U), 77.4 (C2´U), 74.9 (C3´T), 74.8 (C3´T), 61.9 (C5´U), 61.7 (C5´U), 60.8 (OMe), 39.5 (C2´T), 29.5 (C5´T), 25.2 (TBDMS), 24.9 (C7), 17.6 (TBDMS), 11.6 (MeT), 11.6 (MeT), -5.3 (TBDMS), -5.3 (TBDMS). **HRMS-ESI** (*m*/*z*) calc. for $C_{28}H_{44}N_5O_{10}Si$ [M + H^+]: 638.2852; Found: 638.2924.

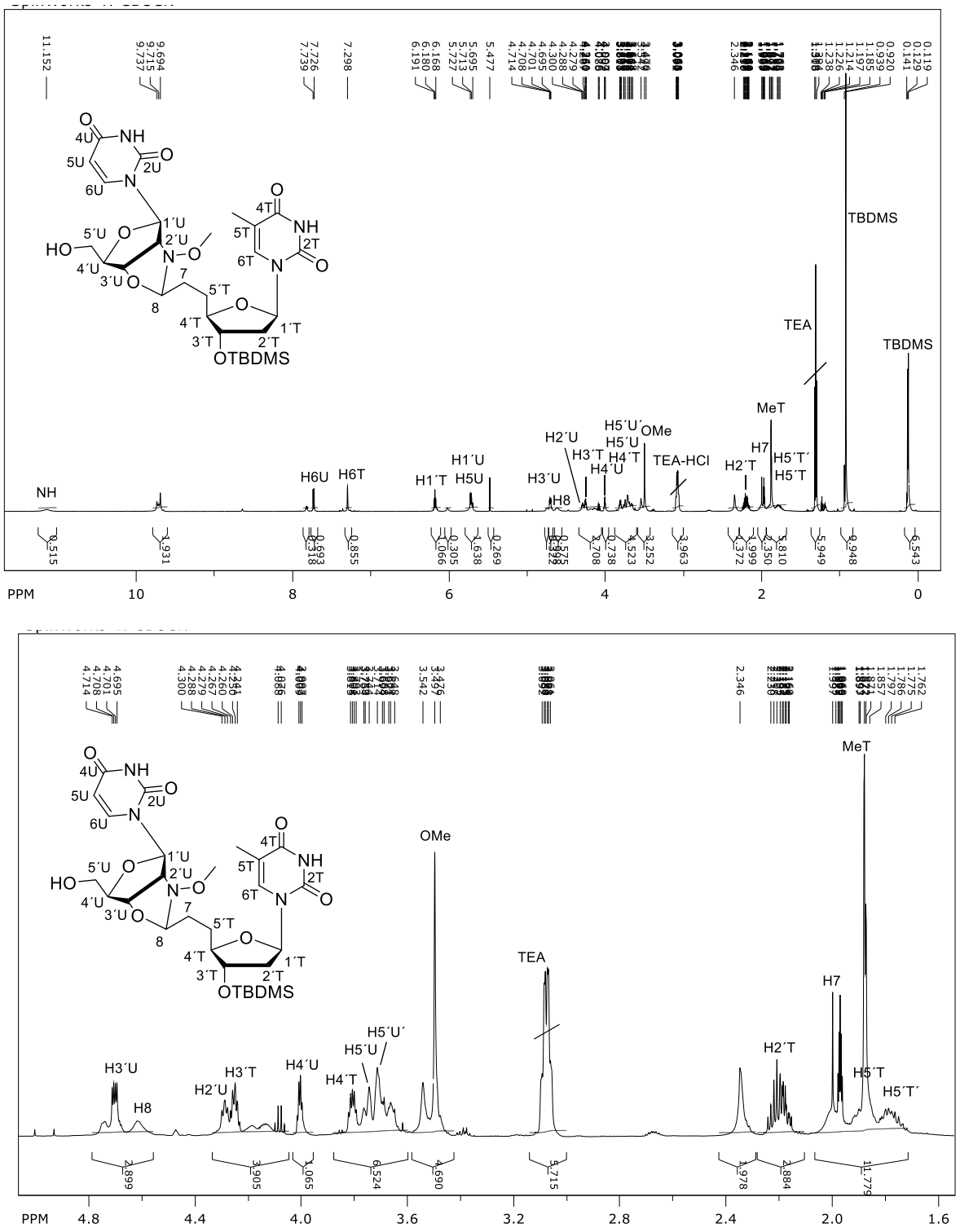


Figure S7 ^{1}H NMR (CD_3CN, 600 MHz) spectrum of **6**.

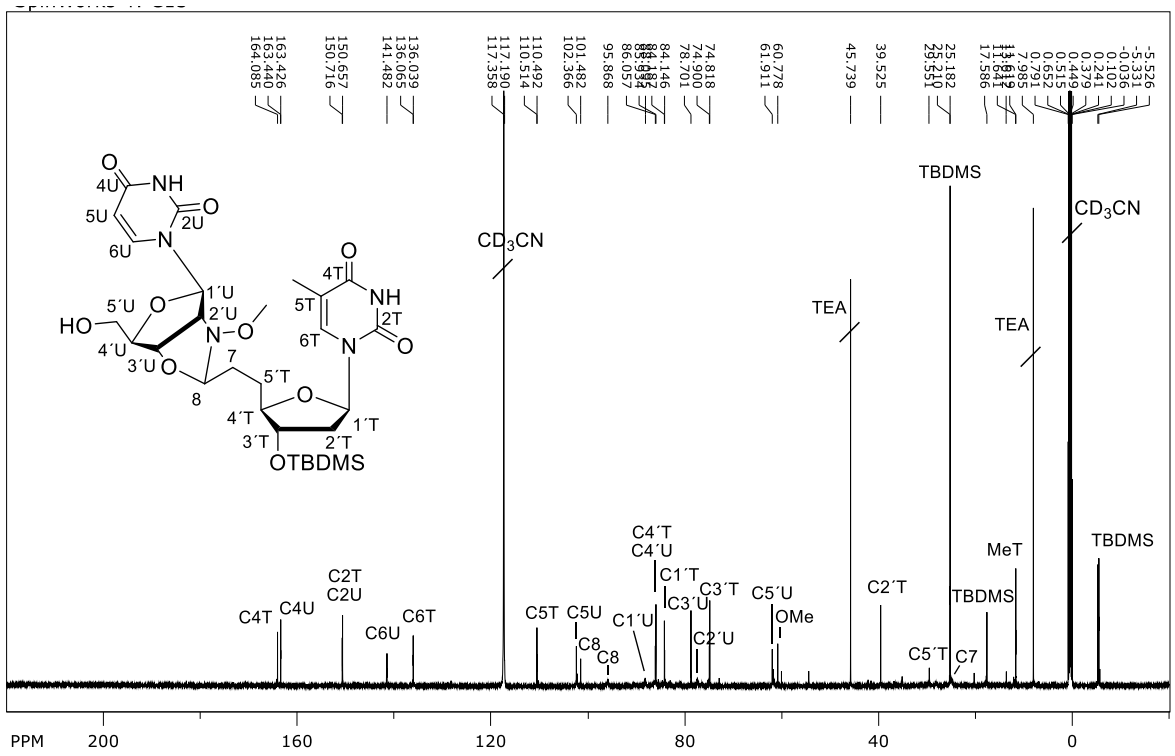
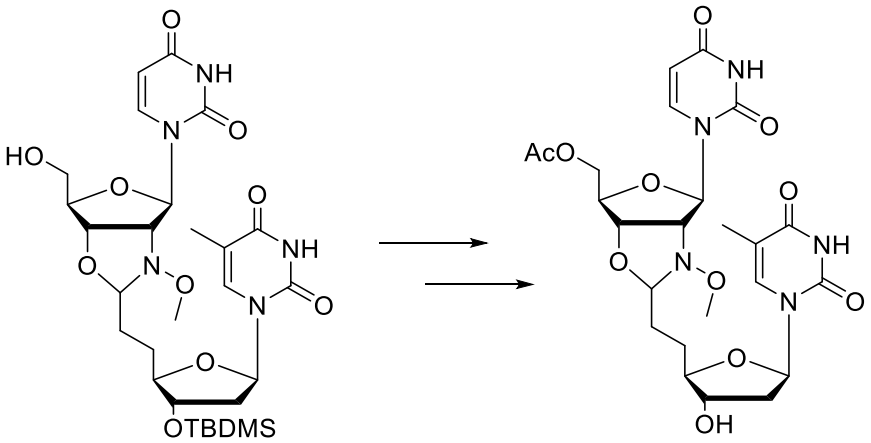


Figure S8 ^{13}C NMR (CD_3CN, 150 MHz) spectrum of **6**.

Synthesis of 7



6 **7**

Nucleoside dimer **6** (0.55 g, 0.86 mmol) was dissolved in pyridine (15 ml). To the mixture was added acetic anhydride (0.20 ml, 1.72 mmol) and a small crystal of DMAP. After one hour the starting material was completely consumed according to TLC (6% MeOH in DCM). The reaction mixture was concentrated using a vacuum and then diluted with ethyl acetate (150 ml) and washed with aq. sat. $NaHCO_3$ (2 × 30 ml) and brine (30 ml). The organic phase was dried using Na_2SO_4, filtered, and evaporated to dryness to yield a solid white foam, which was dried in a vacuum desiccator overnight. The evaporation residue was dissolved in dry THF (10 ml) and tetrabutylammonium fluoride (1.0 M in THF, 2.5 ml, 2.5 mmol) was added. The reaction vessel was tightly sealed and placed in an oven (55 °C). After one hour the solvents were removed in vacuum. The evaporation residue was dissolved in water (50 ml) to which was added dropwise aq. $NaClO_4$ (3.0 M) until forming of a white solid was ceased.

The mixture was filtered through a glass frit and then a syringe filter (0.2 µM pore size). The filtrate was salted with NaCl (18 g, 0.31 mol) and then extracted using ethyl acetate (6 × 100 ml). The organic phase was dried with Na_2SO_4, filtered, and evaporated to dryness. The evaporation residue was adsorbed on silica gel and purified by silica column chromatography (0.1% TEA and 7.0% MeOH in DCM). The product (0.34 g, 70%) was a white solid foam. A mixture of isomers in 1:1 ratio was obtained determined by the ratio of uracil H6 NMR signals. **^{1}H NMR** (CD_3CN, 500 MHz): δ 9.95 (1H, br, NH), 7.55 (0.5H, d, *J* = 8.1 Hz, H6U), 7.54 (0.5H, d, *J* = 8.1 Hz, H6U), 7.31 (1H, m, H6T), 6.19 (1H, t, *J* = 6.5 Hz, H1´T), 6.00 (0.5H, d, *J* = 3.2 Hz, H1´U) 5.76 (0.5H, d, *J* = 8.1 Hz, H5U), 5.75 (0.5H, d, *J* = 8.1 Hz, H5U), 5.72 (0.5H, br, H1´U), 4.70 (1H, br, H8), 4.65 (1H, m, H3´U), 4.37 (1H, m, H2´U), 4.30 (1H, dd, J_1 = 11.9 Hz, J_2 = 3.5 Hz, H5´U), 4.23 (1H, m, H5´U´), 4.18 (1H, m, H3´T), 4.15 (1H, m, H4´U), 3.83 (1H, m, H4´T), 3.56 (1.5H, s, OMe), 3.50 (1.5H, s, OMe), 2.20 (2H, m, H2´T), 2.05 (1.5H, s, Ac), 2.04 (1.5H, s, Ac), 2.00 (2H, br, H7), 1.86 (3H, m, MeT), 1.86 (1H, m, H5´T), 1.79 (1H, m, H5´T´). **^{13}C NMR** (CD_3CN, 125 MHz): δ 170.6 (Ac), 170.6 (Ac), 164.4 (C4T), 163.7 (C4U), 163.6 (C4U), 150.8 (C2T), 150.7 (C2U), 141.4 (C6U), 136.2 (C6T), 136.0 (C6T), 110.6 (C5T), 102.5 (C5U), 102.5 (C5U), 101.6 (C8), 88.5 (C1´U), 86.0 (C4´T), 85.7 (C4´T), 84.2 (C1´T), 84.2 (C1´T), 82.9 (C4´U), 78.8 (C3´U), 77.0 (C2´U), 73.8 (C3´T), 73.7 (C3´T), 64.2 (C5´U), 64.0 (C5´U), 61.8 (OMe, minor isomer), 60.9 (OMe, major isomer), 39.0 (C2´T), 39.0 (C2´T), 29.6 (C5´T), 20.1 (Ac), 20.1 (Ac), 11.7 (MeT), 11.7 (MeT). **HRMS-ESI** (*m/z*) calc. for $C_{24}H_{31}N_5O_{11}$ [M + H^+]: 566.2093; Found: 566.2.

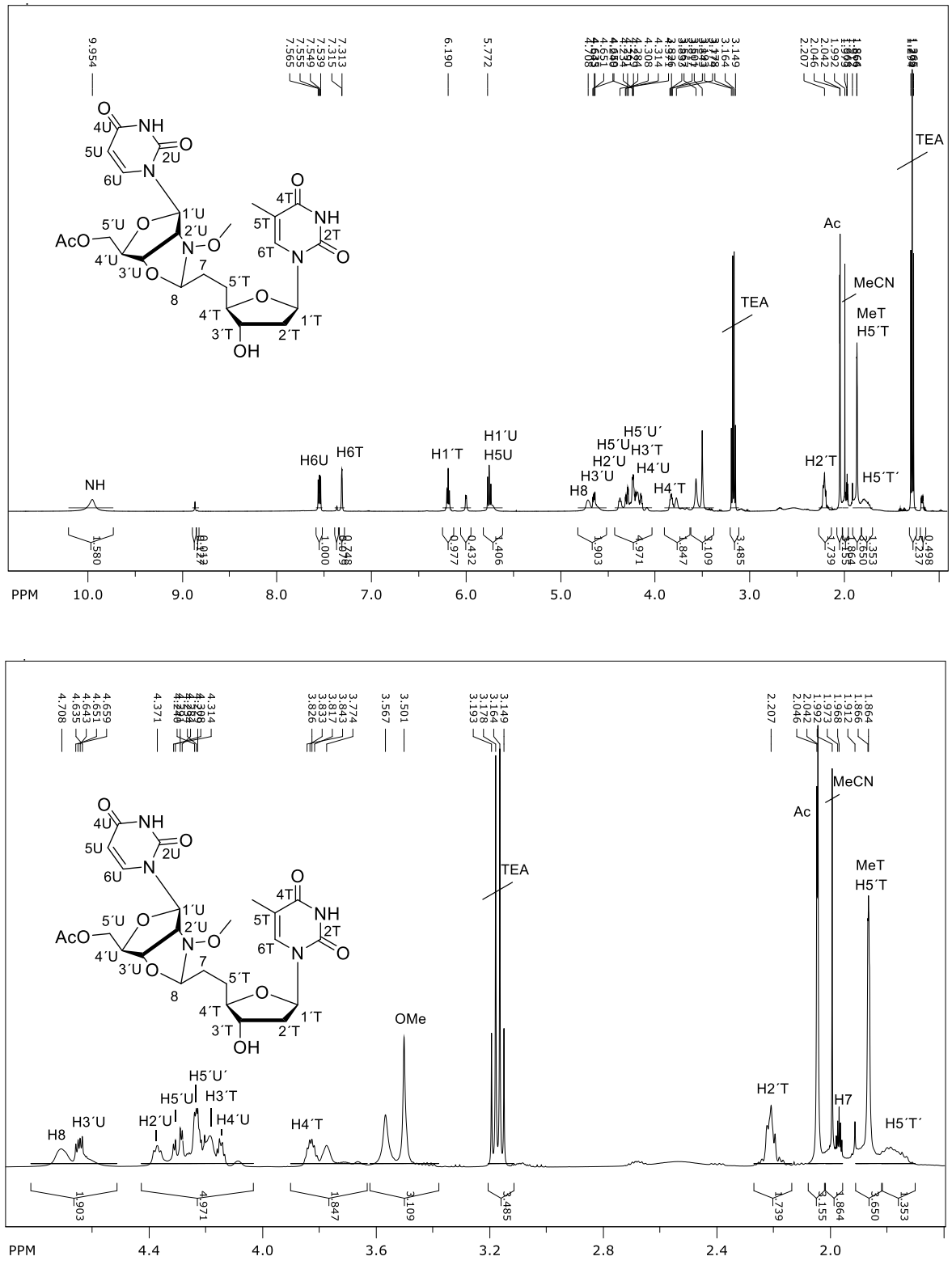


Figure S9 ^{1}H NMR (CD_3CN, 500 MHz) spectrum of **7**.

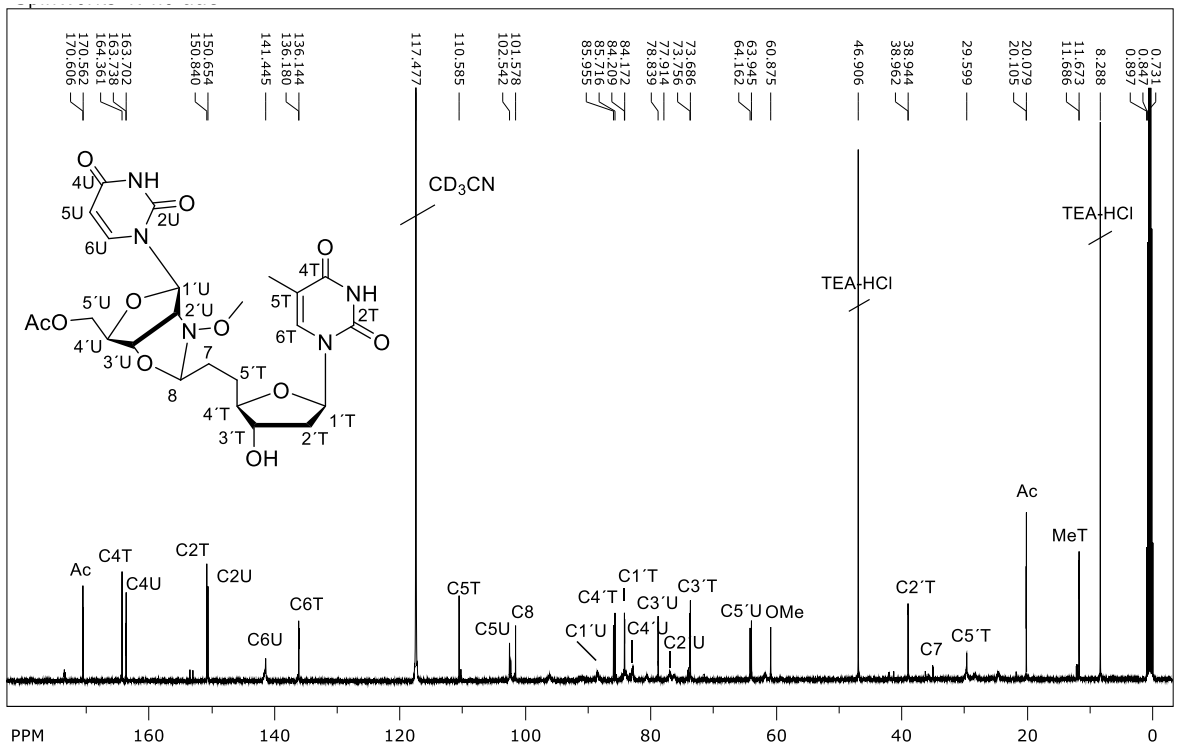
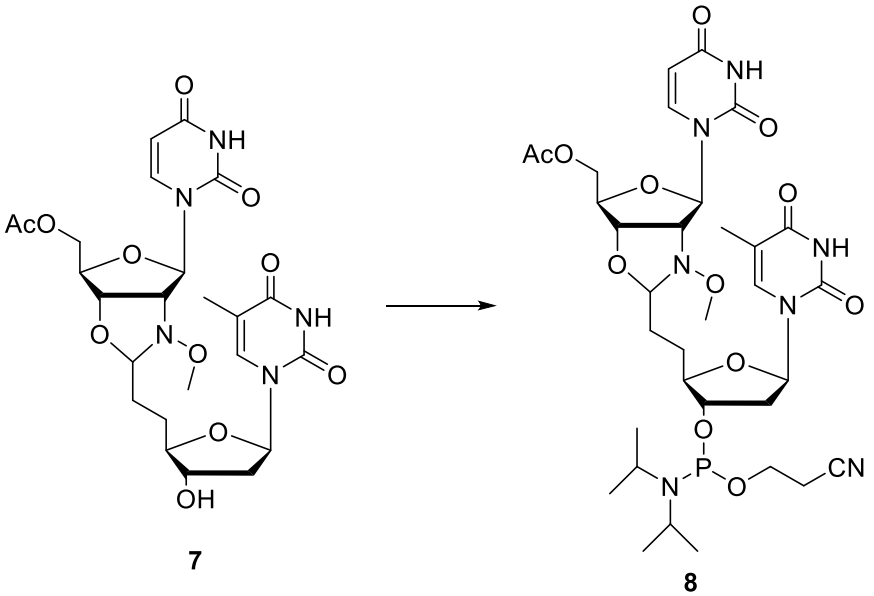


Figure S10 ^{13}C NMR (CD_3CN, 125 MHz) spectrum **7**.

Synthesis of 8



Nucleoside **7** (170 mg, 0.30 mmol) was dissolved in dry acetonitrile (2.5 ml) under nitrogen. To the solution was added DIPEA (125 µl, 0.72 mmol) and *N*,*N*-diisopropylcyanoethylchlorophosphoramidite (80 µl, 0.36 mmol). After one hour, the reaction mixture was concentrated in a vacuum. The evaporation residue was purified using silica column chromatography (1.0% TEA and 4% MeOH in DCM). The product (143 mg, 62%) was colorless amorphous solid. A mixture of multiple isomers was obtained, for which the number of isomers and their ratios could not be determined due to

overlapping of ^{1}H NMR signals. **^{1}H NMR** (CD_3CN, 500 MHz): δ 9.20 (1H, br, NH), 7.50 (1H, m, H6U), 7.28 (1H, s, H6T), 6.19 (1H, m, H1′T), 6.00 (0.5H, m, H1′U), 5.71 (1H, m, H5U), 5.70 (0.5H, m, H1′U), 4.71 (1H, br, H8), 4.63 (1H, m, H3′U), 4.38 (1H, m, H3′T), 4.36 (1H, m, H2′U), 4.29 (1H, m, H5′U), 4.23 (0.5H, m, H5′U′), 4.23 (0.5H, m, H4′U), 4.14 (0.5H, m, H4′U), 4.00 (0.5H, m, H4′T), 3.95 (0.5H, m, H4′T), 3.85 (1H, m, $POCH_2$), 3.78 (1H, M, $POCH_2$), 3.58 (m, OMe), 3.52 (s, OMe), 3.50 (s, OMe), 2.70 (1H, t, *J* = 6.2 Hz, CH_2CN), 2.69 (1H, t, *J* = 6.2 Hz, CH_2CN), 2.42–2.25 (2H, m, H2′T), 2.18 (1H, br, H7), 2.05 (3H, s, Ac), 1.94 (1H, m, H5′T), 1.87 (3H, m, MeT), 1.82 (1H, m, H5′T′), 1.22 (s, iPr), 1.20 (s, iPr) . **^{13}C NMR** (CD_3CN, 125 MHz): δ 170.4 (Ac), 170.4 (Ac), 163.7 (C4T), 163.0 (C4U), 163.0 (C4U), 150.4 (C2T), 150.4 (C2U), 141.4 (C6U), 141.1 (C6U), 135.9 (C6T), 135.8 (C6T), 135.8 (C6T), 118.7 (CN), 118.6 (CN), 110.6 (C5T), 110.6 (C5T), 110.5 (C5T), 110.5 (C5T), 102.5 (C5U), 102.2 (C5U), 101.4 (C8), 101.3 (C8), 88.4 (C1′U), 84.8 (C4′T), 84.8 (C4′T), 84.7 (C4′T), 84.7 (C4′T), 84.3 (C1′T), 84.2 (C1′T), 84.2 (C1′T), 84.1 (C1′T), 78.8 (C3′U), 76.3–75.6 (8 peaks, C2′U, C3′T, C4′U), 64.1 (C5′U), 63.9 (C5′U), 63.8 (C5′U), 61.7 (OMe), 60.8 (OMe), 58.6 ($POCH_2$), 58.4 ($POCH_2$), 58.3 ($POCH_2$) 58.3 ($POCH_2$), 43.1 (iPr), 43.1 (iPr), 43.1 (iPr), 43.0 (iPr), 38.2 (C2′T), 38.2 (C2′T), 38.0 (C2′T), 38.0 (C2′T), 35.0 (C7), 29.4 (C5′T), 24.0 (iPr), 23.9 (iPr), 23.9 (iPr), 23.8 (iPr), 20.1 (Ac), 20.0 (Ac), 20.0 (Ac), 20.0 (Ac), 11.5 (MeT). **MS-ESI** (*m*/*z*) calc. for $C_{33}H_{48}N_7NaO_{12}P$ [M + H^+]: 766.3; Found: 766.3. **HRMS-ESI** (obtained for the excepted hydrolysis product H-phosphonate) (m/z) calc. for $C_{27}H_{36}N_6O_{13}P$ [M + H^+]: 683.2072; Found: 683.1989.

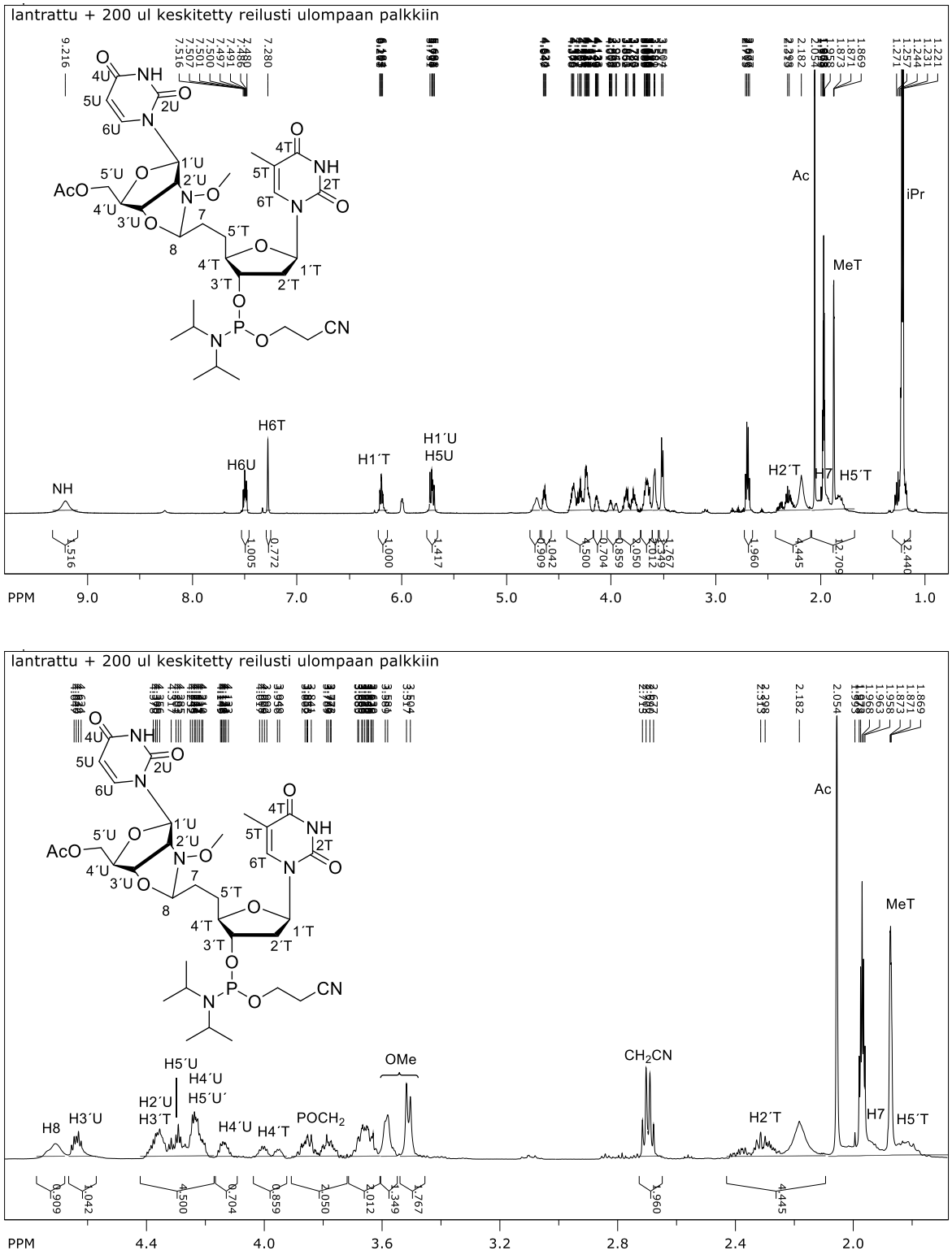


Figure S11 ^{1}H NMR (CD_3CN, 500 MHz) spectrum of **8**.

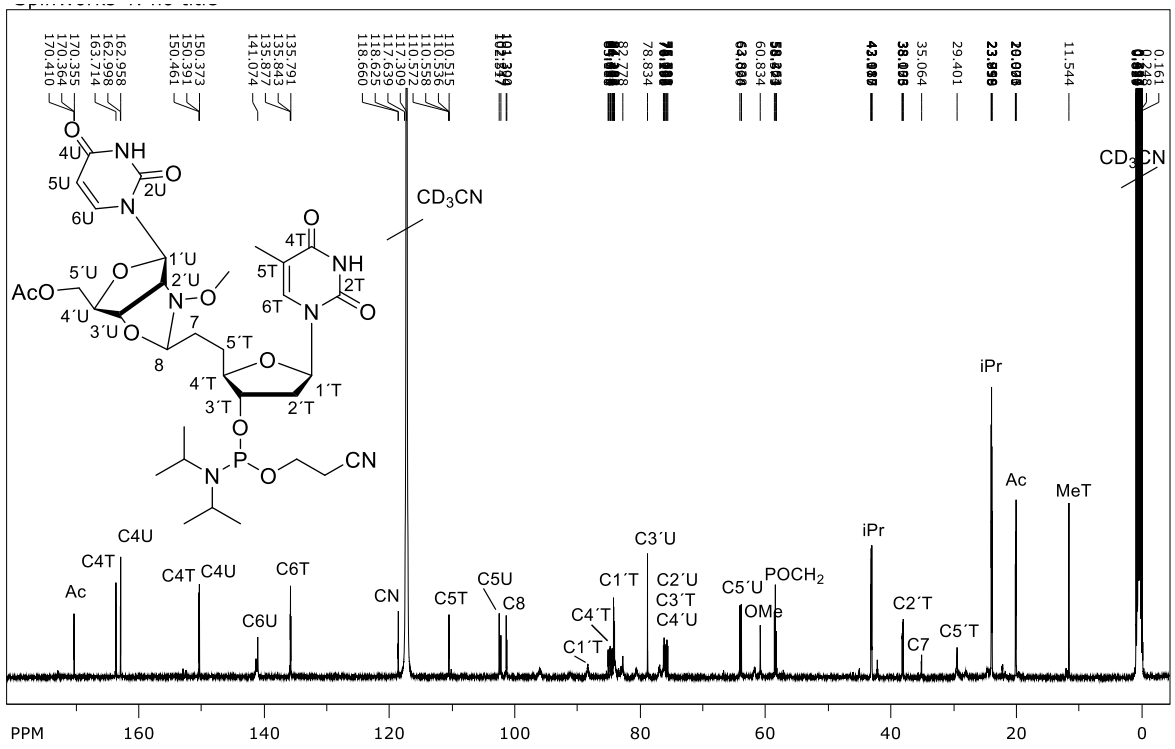


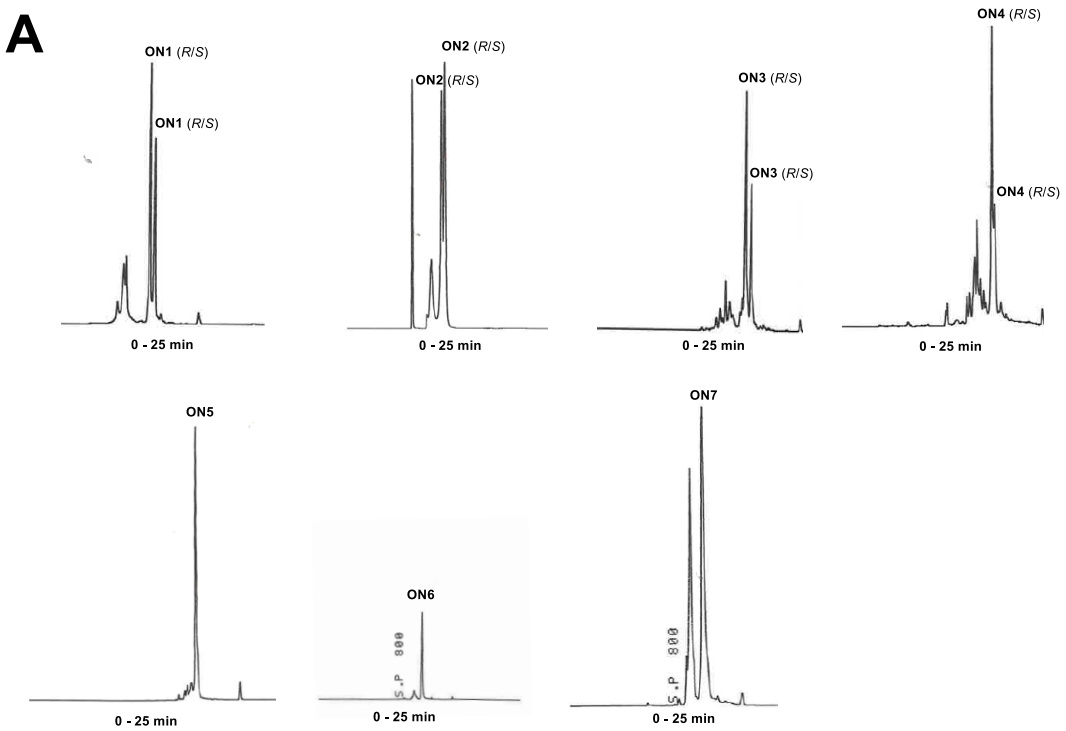
Figure S12 ^{13}C NMR (CD_3CN, 125 MHz) spectrum of **8**.

Synthesis of oligonucleotides

All oligonucleotides were synthesized on 1.0 µmol scale using an automatic DNA/RNA-synthesizer, commercially available phosphoramidite building blocks and phosphoramidite **8**, commercial solid support and solid support **9**. Benzylthiotetrazol was used as an activator. Coupling times of 2 min for commercial building blocks and 30 min for phosphoramidite **8** were used. After the chain assembly, the solid-supported oligonucleotides were exposed to aq. 25% NH_4OH for 16 h in oven at 55 °C. The resulting suspensions were filtered and the filtrates were evaporated to dryness. The residues were dissolved in water and purified by RP HPLC After lyophilization, the oligonucleotides were characterized by MS (ESI-TOF) and quantified UV spectrophotometrically.

Table S1 Molecular masses for the synthesized ONs.

Oligonucleotide	Isolated yield (%)	Observed molecular mass	Calculated molecular mass
ON1	15	3571.7	3572.5
ON2	13	8001.3	8002.3
ON3	12	4303.9	4305.0
ON4	5	4554.9	4556.1
ON5	22	7119.4	7119.7
ON6	13	9210.8	9211.1
ON7	11	4905.0	4905.4



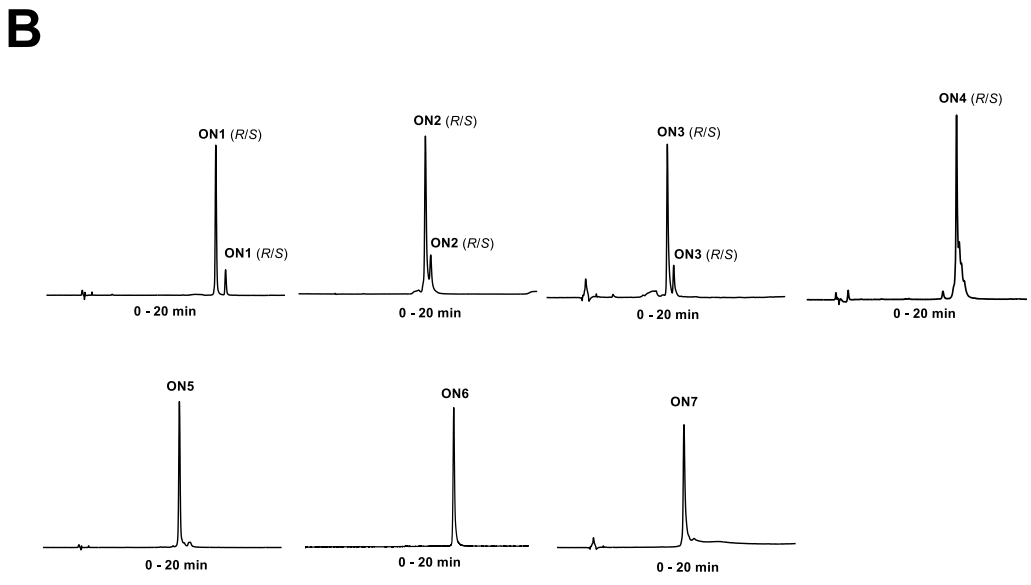


Figure S13 RP HPLC of the synthesized ONs **(ON1, ON2, ON3, ON4, ON5, ON6,** and **ON7**). A) Crude products. B) After purification. Note, that isolation of single isomer is difficult due to rapid isomerization in aqueous conditions. RP HPLC analyses were run using C18-column (250 × 10 mm, 5 µM) with a linear gradient of MeCN (5–40% in 25 min) over 50 mM aq. triethylammonium acetate (pH 7). A flow rate of 2.5 ml min-1 and a detection wavelength of 260 nm were employed.

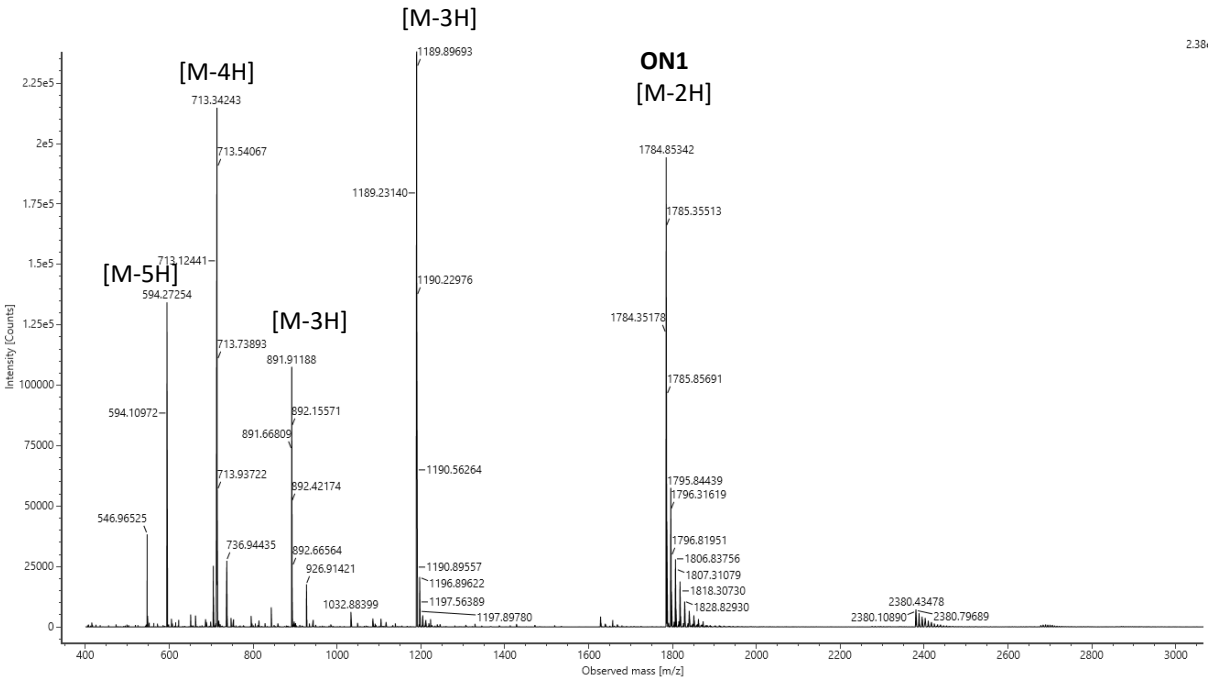


Figure S14 Mass spectrum for **ON1**.

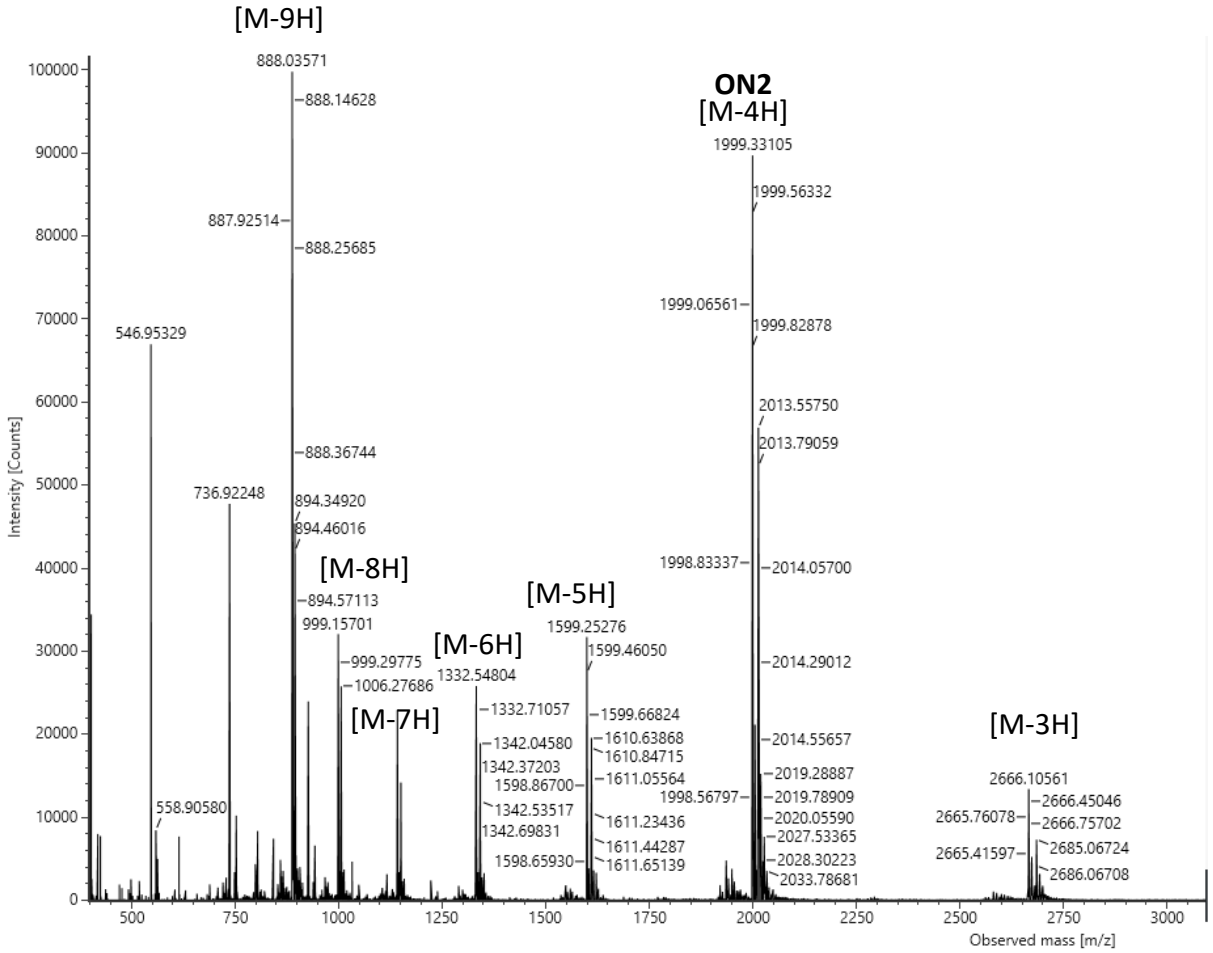


Figure S15 Mass spectrum for **ON2**.

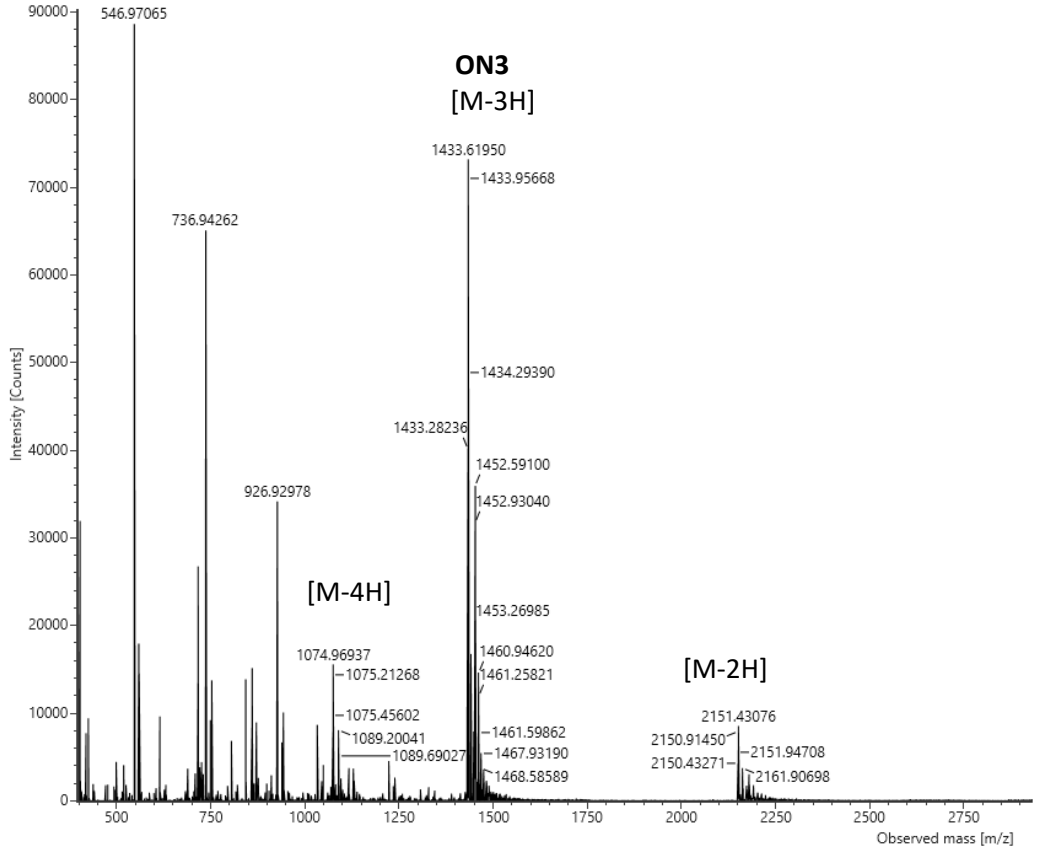


Figure S16 Mass spectrum for **ON3**.

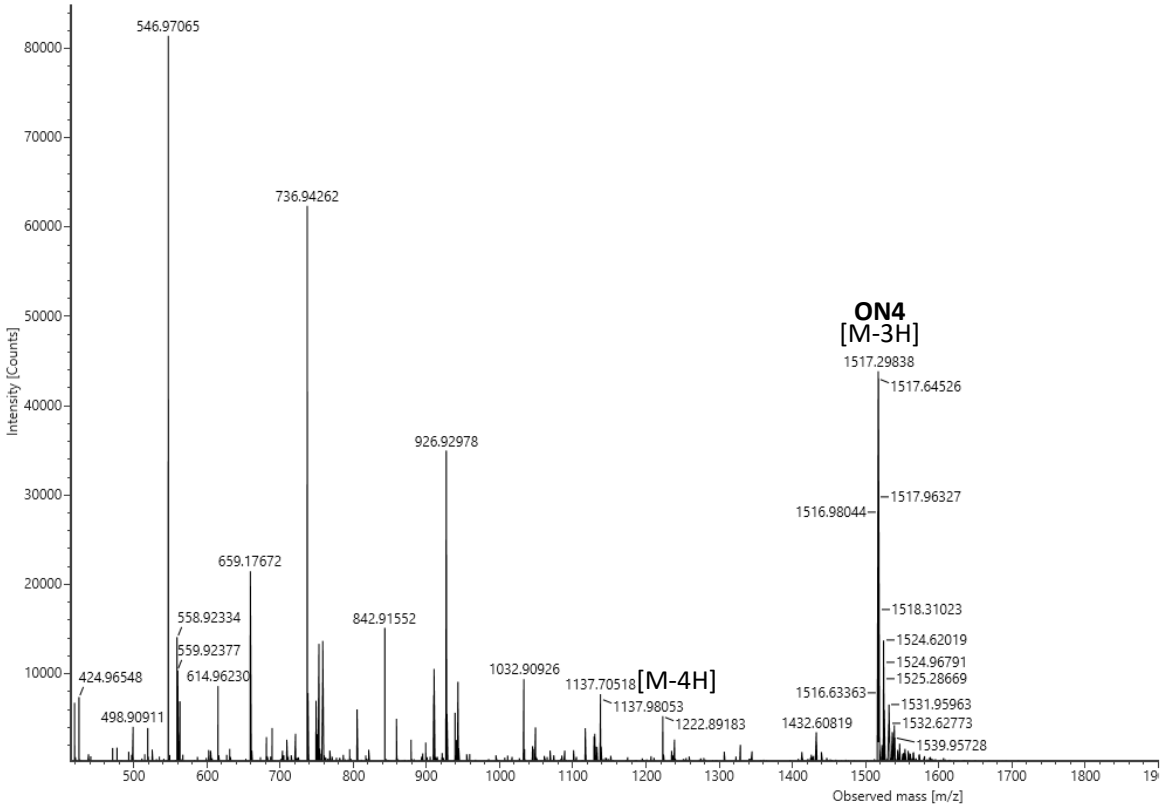


Figure S17 Mass spectrum for **ON4**.

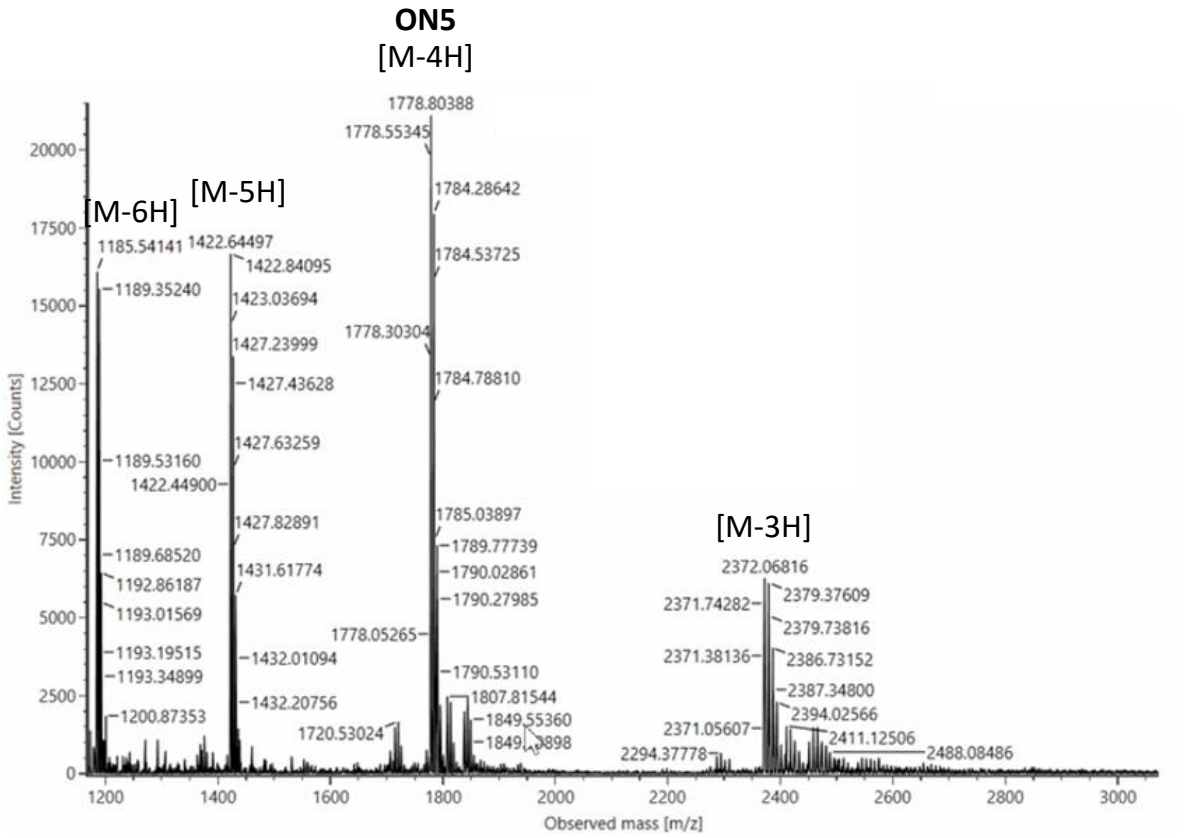


Figure S18 Mass spectrum for **ON5**.

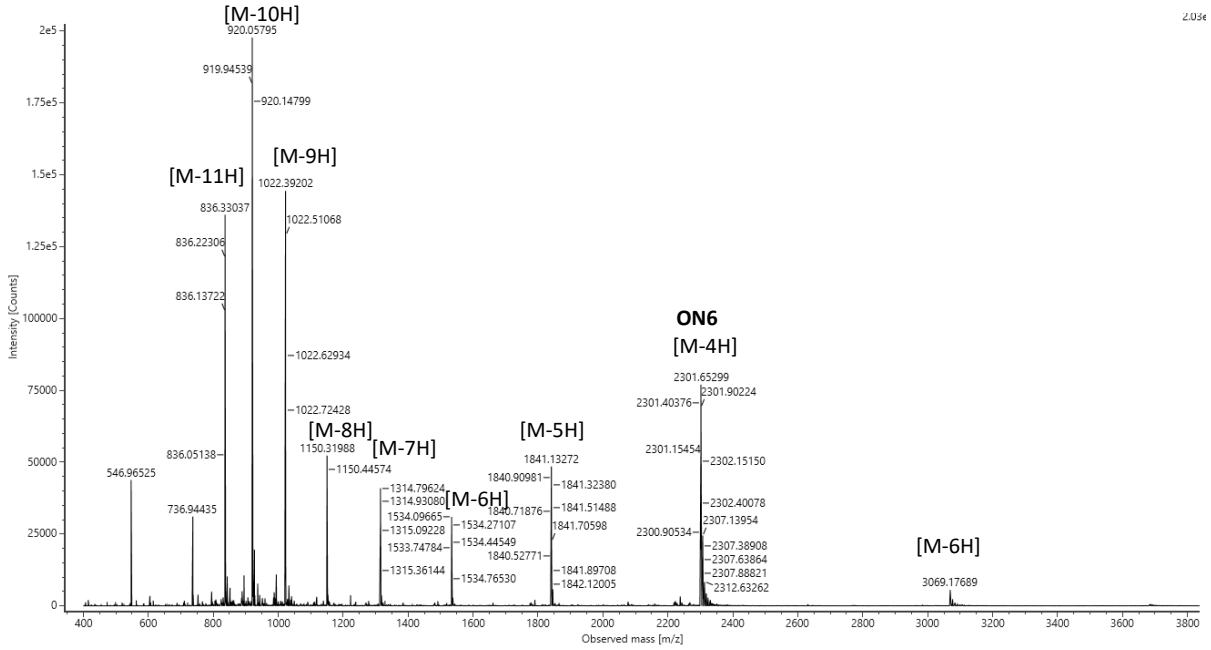


Figure S19 Mass spectrum for **ON6**.

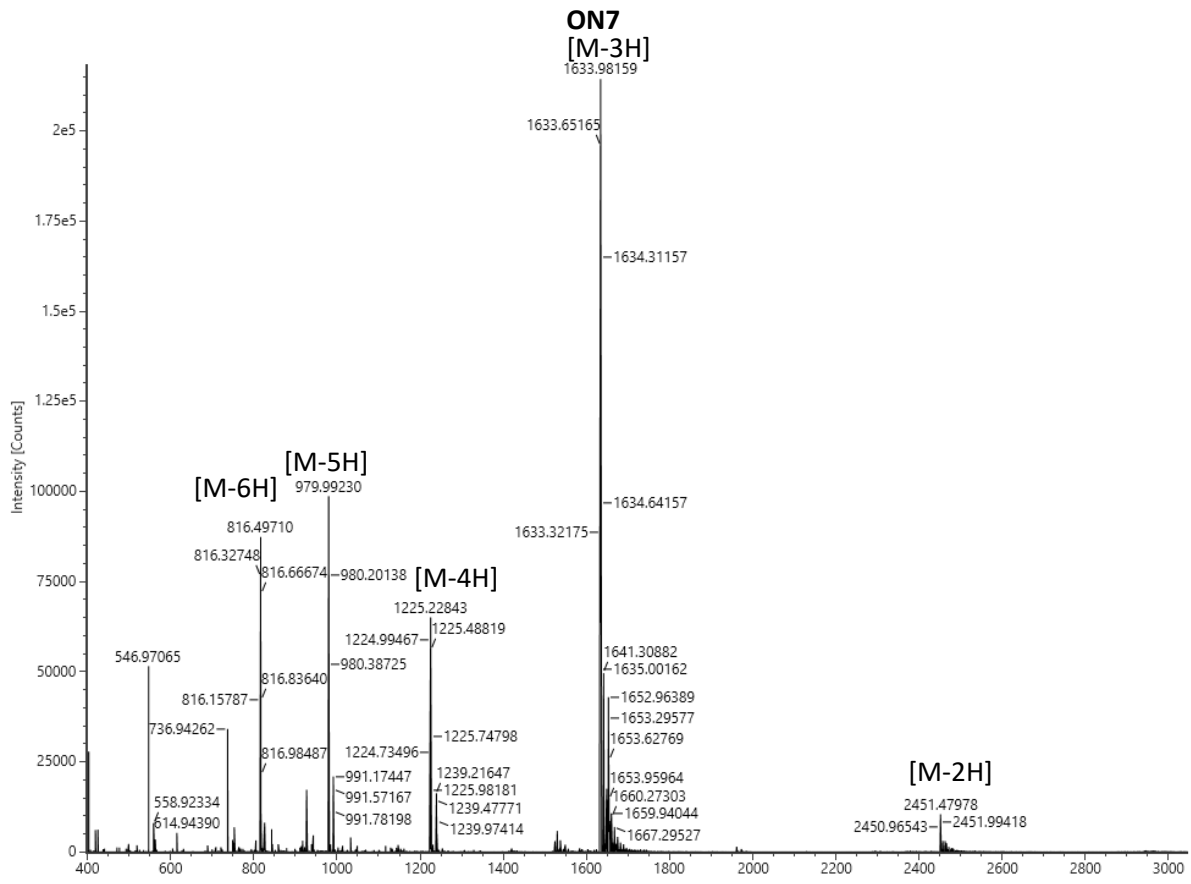


Figure S20 Mass spectrum for **ON7**.

Preliminary DNA-templated ligation studies with ON1, ON4, and ON5 (deprotection, reprotection, and hairpin L1 formation)

General: The buffers used in all DNA-templated hairpin self-assembly studies are listed in **Table S2**.

Table S2 List of the buffers used in DNA-templated ligation studies.

pH	Acid/base	*I* to *c* ratio (adjusted using NaCl)
4	AcOH/AcONa	1
5	AcOH/AcONa	1
6	MES/MES-Na	1
7	HEPES/HEPES-Na	1
8	HEPES/HEPES-Na	1
9	CHES/CHES-Na	1
10	CHES/CHES-Na	1
11	CAPS/CAPS-Na	1

***N,O*-transacetalization: ON5** (750 pmol, 5 µM) and **ON1-Cap** (750 pmol, 5 µM) were dissolved in water (150 µl) which pH was adjusted using a 0.1 M buffer solution. Samples (10 µl) were taken at adequate intervals and basified to pH 8 using 0.4 M buffer solution (90 µl) and analyzed immediately by RP HPLC (**Scheme S1**). The data points were fitted to exponential decay function (Equation S2)

Hydrolysis of ON1: **ON1** (750 pmol, 5 µM) was dissolved in water (150 µl) containing 0.1 M buffer. Samples (10 µl) were taken at adequate intervals and basified to pH 8 using 0.4 M buffer solution (90 µl) and analyzed immediately by RP HPLC (**Scheme S1**). The data data points were fitted to reversible bimolecular integrated rate law (Equation S1).

Reprotection of ON1* and **ON4***: The ON (**ON1** or **ON4**, 825 pmol, 5.5 µM) was dissolved in water (150 µl) adjusted to pH 5 using trace amount of acetic acid and then incubated overnight at 37°C after which RP HPLC analysis indicated the ON was fully deprotected (**ON1*** or **ON4***). Then, the mixture was allowed to cool to room temperature and the pH was adjusted by addition of 1.0 M buffer (15 µl). To the solution was added nucleoside **1** (3.94 µl of stock solution for final concentration of 500 µM). Samples (10 µl) were taken at adequate intervals and quenched to pH 8 using aqueous 0.4M buffer solution (90 µl) and analyzed immediately by RP HPLC (**ON1** in **Scheme S1** and **ON4** in **Scheme S2**). The data data points were fitted to reversible bimolecular integrated rate law (Equation S1).

DNA-templated *N*-methoxyoxazolidine formation: **ON1** (825 pmol, 5.5 µM) was dissolved in water (150 µl) adjusted to pH 5 using trace amount of acetic acid. The solution was incubated overnight at 37 °C after which RP HPLC analysis indicated that **ON1** was fully deprotected into **ON1***. The pH of the solution was adjusted by adding aq. 1.0 M buffer (15 µl, see list of buffers below). To the solution was added **ON5** (825 pmol, 5 µM). Samples (10 µl) were taken at adequate intervals and quenched to pH 11 using 0.4 M buffer solution (90 µl) and analyzed immediately by RP HPLC (**Scheme S1**). The data data points were fitted to reversible bimolecular integrated rate law (Equation S1).

Reversible bimolecular rate law for the reaction A + B ⇌ C:

$$C(t) = C_0 + \frac{(1-e^{-\kappa t})}{\beta(1-e^{\kappa t})+\kappa(1+e^{-\kappa t})} \qquad \text{(Equation S1)}$$

where $\alpha = 2(k_1 A_0 B_0 - k_2 C_0)$, $\beta = k_1(A_0+B_0) + k_2$,

and $\kappa = \sqrt{k_1^2(A_0 - B_0)^2 + 2k_1k_2(A_0 + B_0 + 2C_0) + k_2^2}$

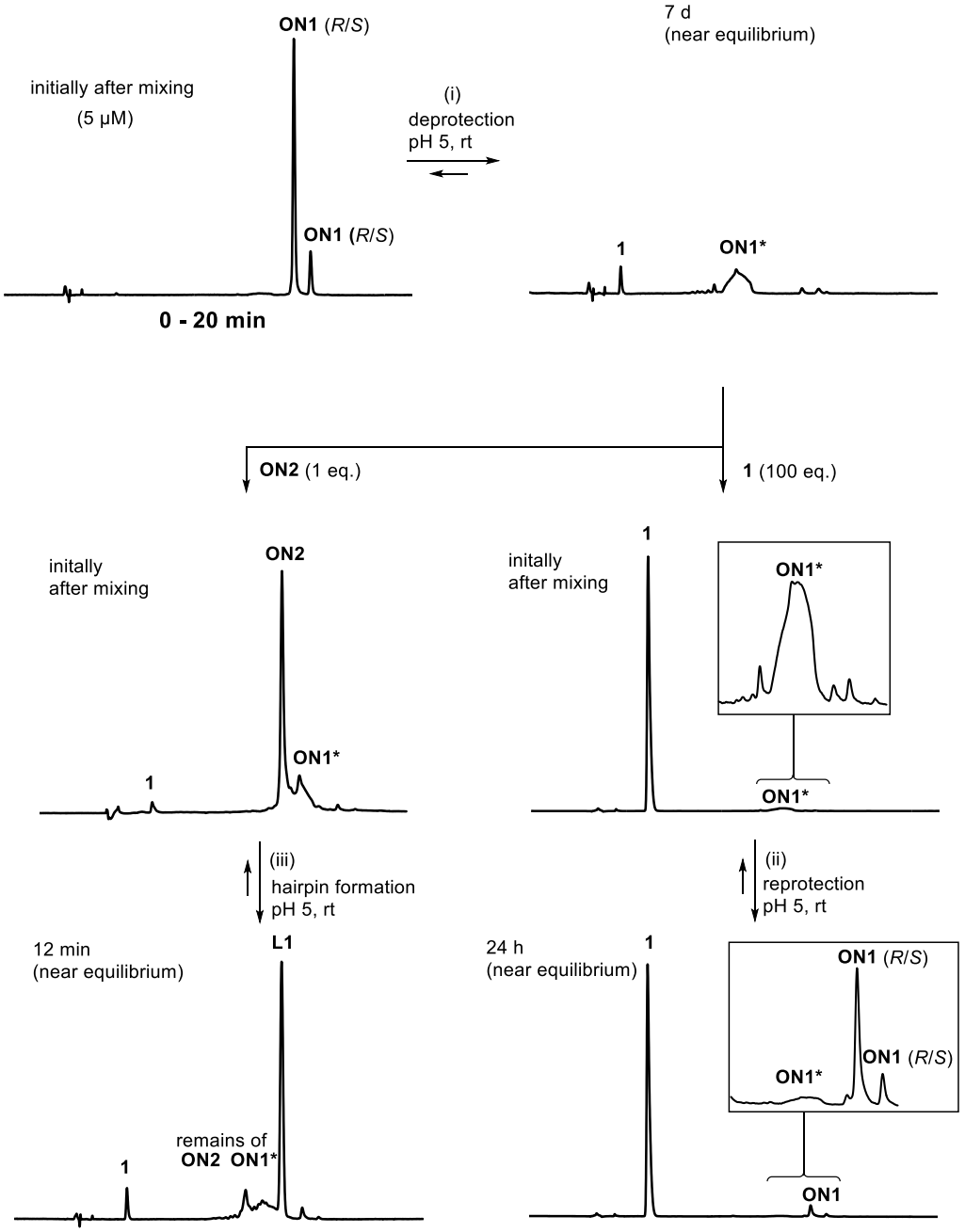
Half-life was determined as the time-point, where the concentration of the product was half from the concentration at the equilibrium (first half-life).

Exponential decay function:

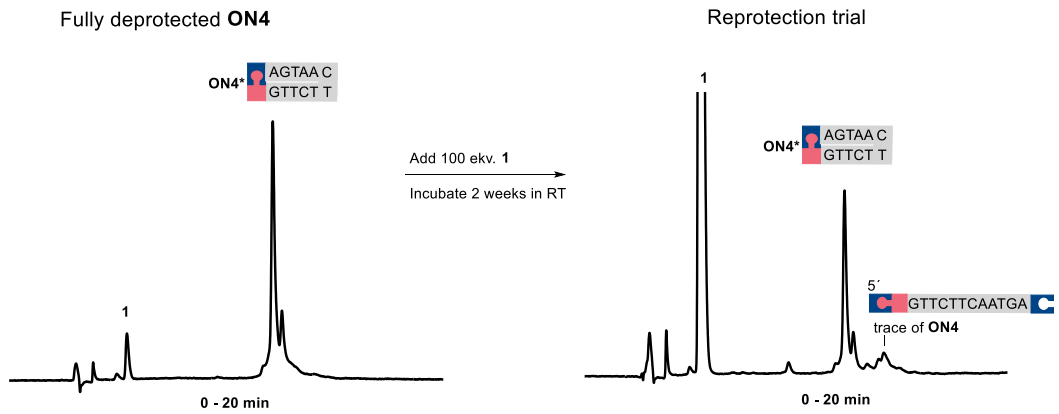
$$y = ae^{-x/t} + b \qquad \text{(Equation S2)}$$

From which half-life ($t_{0.5}$) was calculated:

$$t_{0.5} = \frac{t}{\ln(2)}$$



Scheme S1 Selected examples for the HPLC chromatograms obtained during the preliminary studies with **ON1**, **ON4** and **ON5**. RP HPLC analyses were run using C18-column (250 × 4.6 mm, 5 µm) with a linear gradient of MeCN (5–20% in 25 min) over aq. 50 mM TEAA buffer (pH 7.0). A detection wavelength of 260 nm was employed.



Scheme S2 Reprotection of **ON4.** RP HPLC analyses were run using C18-column (250 × 4.6 mm, 5 µm) with a linear gradient of MeCN (5–20% in 25 min) over aq. 50 mM TEAA buffer (pH 7.0). A detection wavelength of 260 nm was employed.

Table S3 Kinetic and thermodynamic data obtained during the preliminary studies. See Figure S21 for graphical illustration. Equilibrium yields are shown in parentheses.

pH	*Tm* (°C Starting material ON2 + ON4)	*t*½ of *N,O*-transacetalization (ON1 + ON5)	*t*½ of deprotection (ON1)[b]	*t*½ of hairpin formation (ON1* + ON5)[b]	*t*½ of reprotection (ON1* + nucleoside 1)[b]
4	not found	7.69 ± 0.45 h (76 ± 1 %)[a]	8.57 ± 0.14 h (96 ± 2 %)[a]	29.4 ± 0.6 min (76 ± 1 %)[a]	1.85 ± 0.04 h (33 ± 1 %)[a]
5	22.1 ± 0.1	22.3 ± 1.1 h (87 ± 1 %)[a]	30.3 ± 0.5 h (96 ± 1 %)[a]	3.42 ± 1.30 min (94 ± 19 %)[a]	5.00 ± 0.15 h (30 ± 1 %)[a]
6	26.3 ± 0.1	228 ± 38 h (95 ± 6 %)[a]	188 ± 6 h (97 ± 3 %)	5.70 ± 0.58 min (83 ± 5 %)[a]	30.1 ± 0.7 h (34 ± 1 %)[a]
7	31.5 ± 0.1	N/A	70.2 ± 9.9 d (15.9 ± 9.4 %)	29.6 ± 0.7 min (79 ± 1 %)[a]	235 ± 12 h (29 ± 1 %)[a]
8	30.8 ± 0.1	N/A	N/A	3.05 ± 0.07 h (79 ± 1 %)[a]	N/A
9	27.6 ± 0.1	N/A	N/A	13.7 ± 0.3 h (75 ± 2 %)[a]	N/A
10	23.1 ± 0.1	N/A	N/A	27.9 ± 1.3 h (78 ± 2 %)[a]	N/A
11	not found	N/A	N/A	N/A	N/A

N/A: Reactions reactions were not monitored due to their slowness (<4% of product after one month). [a]The equilibrium yield determined for the reaction. [b]Deprotection, reprotection, and **L1** hairpin formation were fitted to reversible bimolecular integrated rate law (Equation S1). [c]DNA-templated transacetalization was fitted to exponential decay function (Equation S2) (Because poorer fit was obtained by using Equation S1).

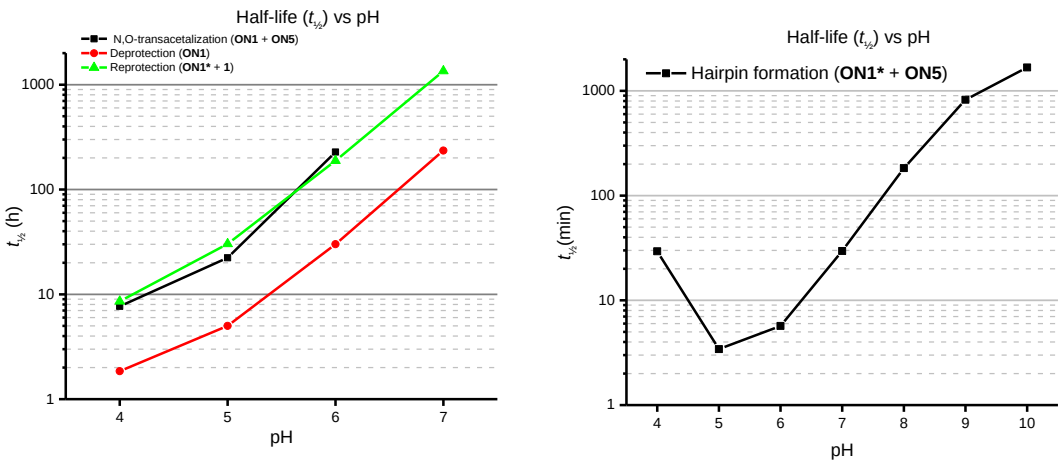


Figure S21 pH-dependent kinetic profiles obtained during the preliminary studies with **ON1**, **ON4**, and **ON5.** See Table S3 above for the results in numerical form and below for individual reaction profiles.

Table S4 Molecular masses obtained from the preliminary studies.

ON	observed molecular weight	calculated molecular weight
ON1*	3317.2	3317.2
ON4*	4298.8	4300.9
L1	10418.9	10418.9

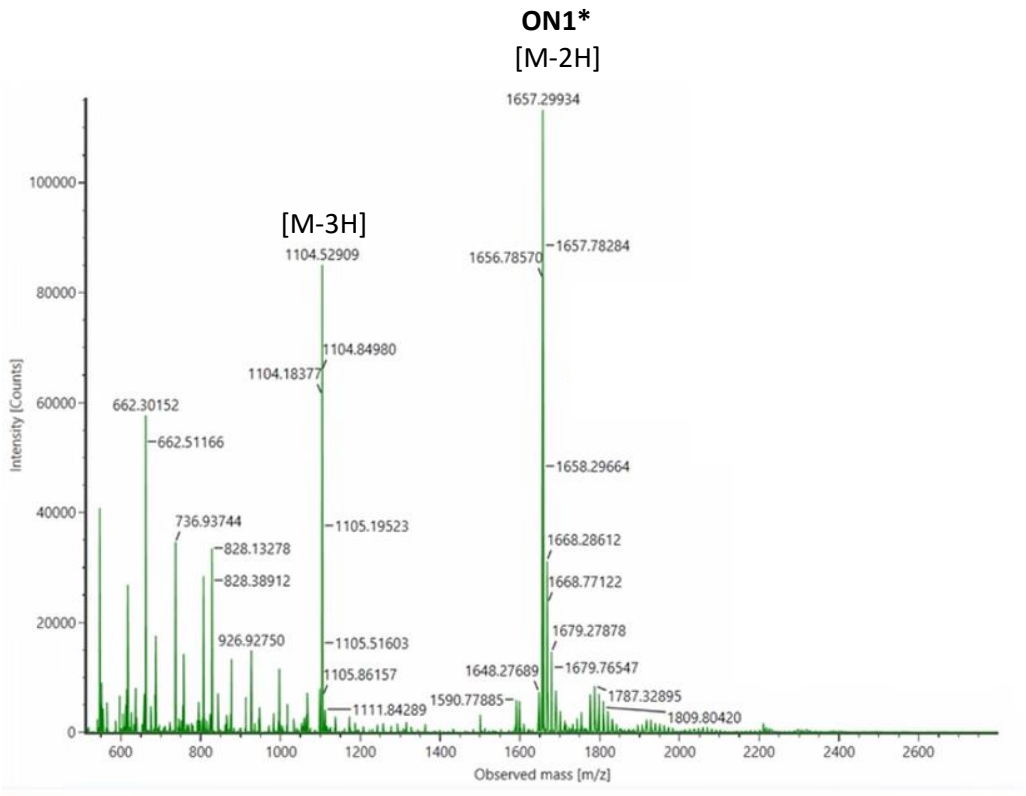


Figure S22 Mass spectrum for **ON1***.

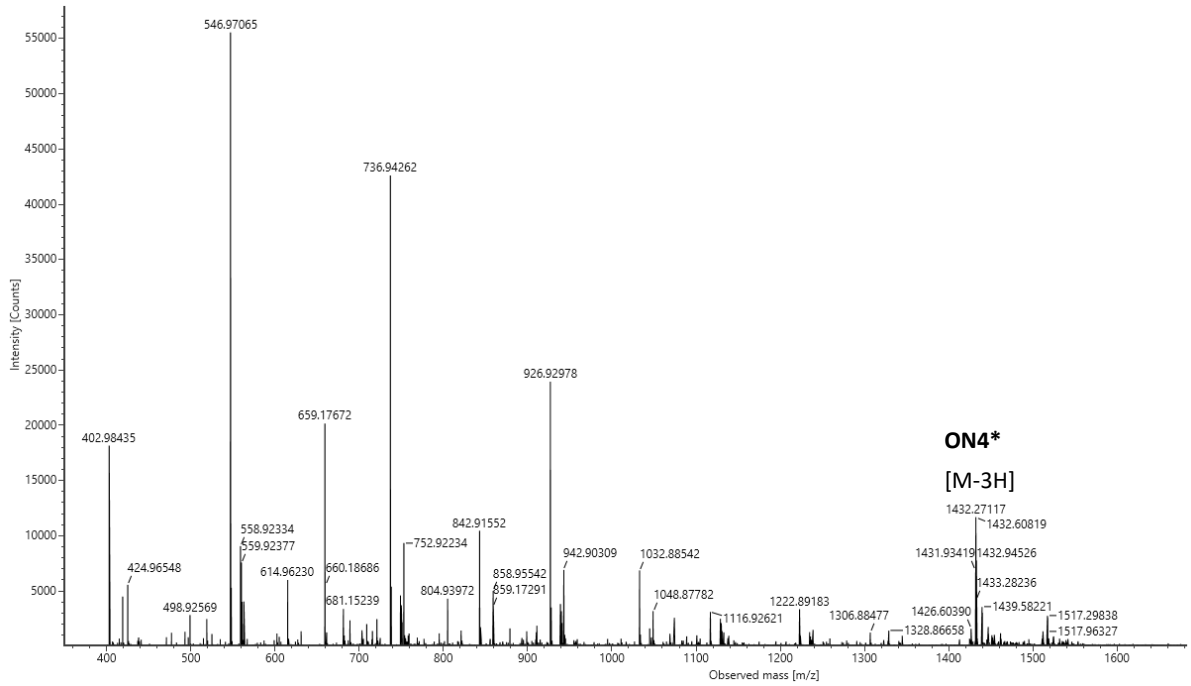


Figure S23 Mass spectrum for **ON4***.

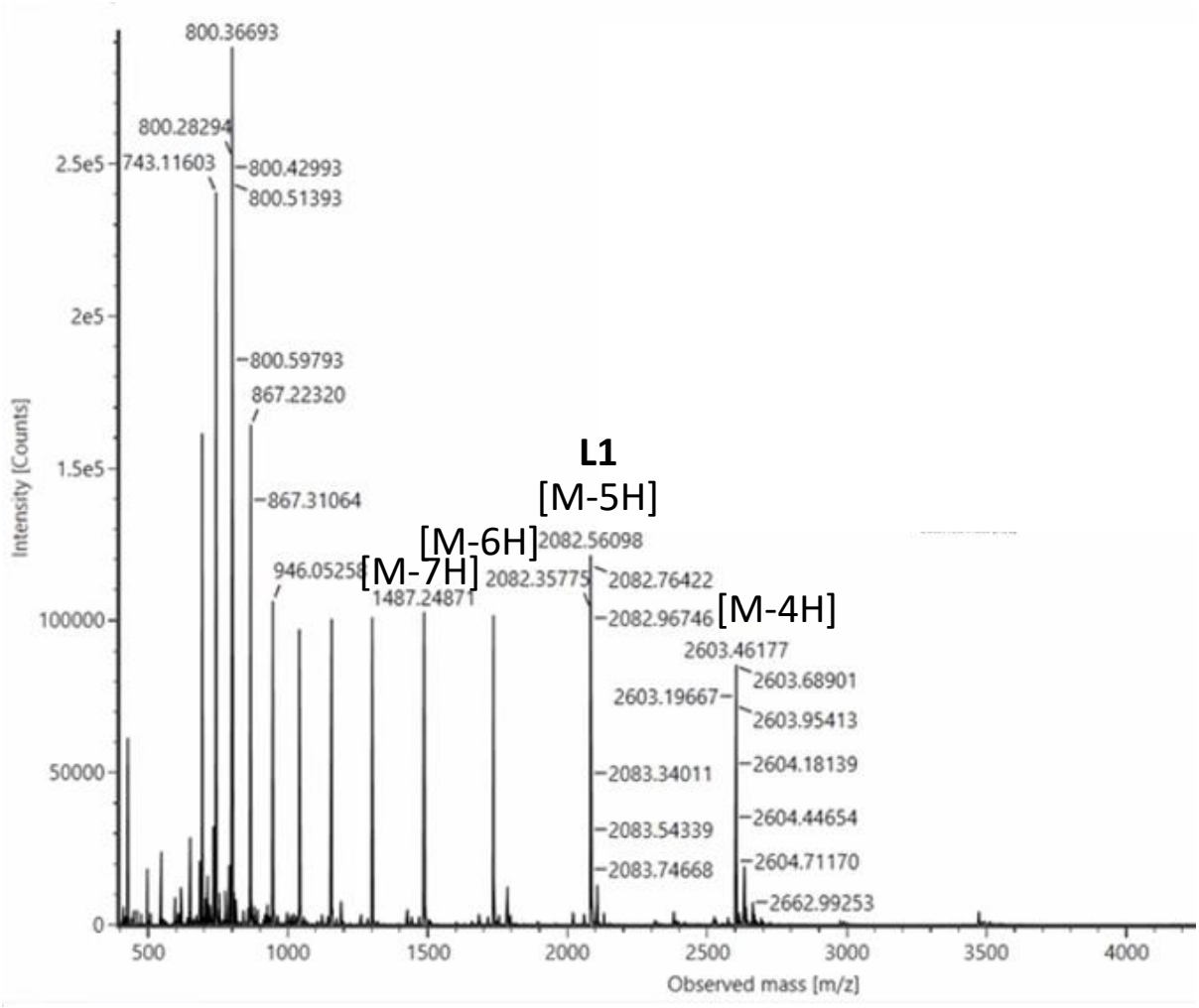


Figure S24 Mass spectrum for **L1**.

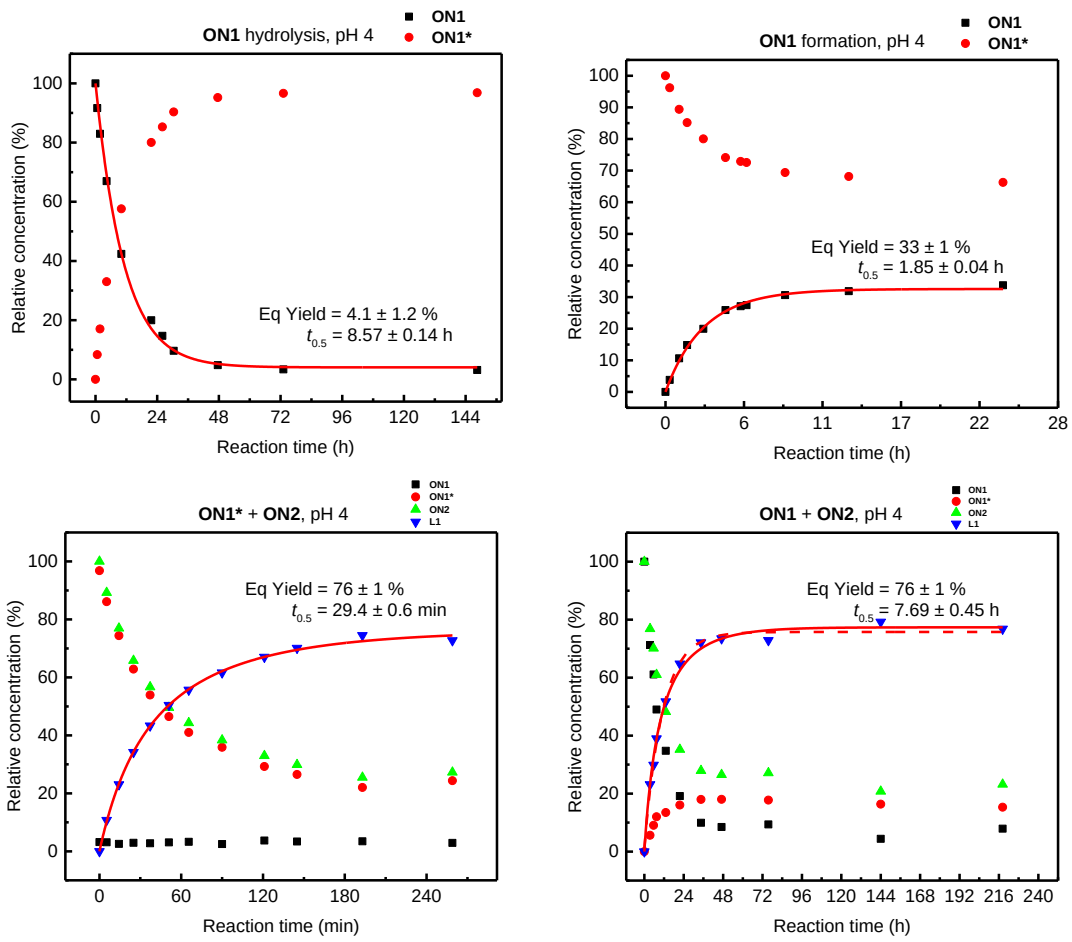


Figure S25 Kinetic profiles for the preliminary studies at pH 4. Solid line was obtained by fitting the data points to Equation S1. Dashed line was obtained by fitting to Equation S2.

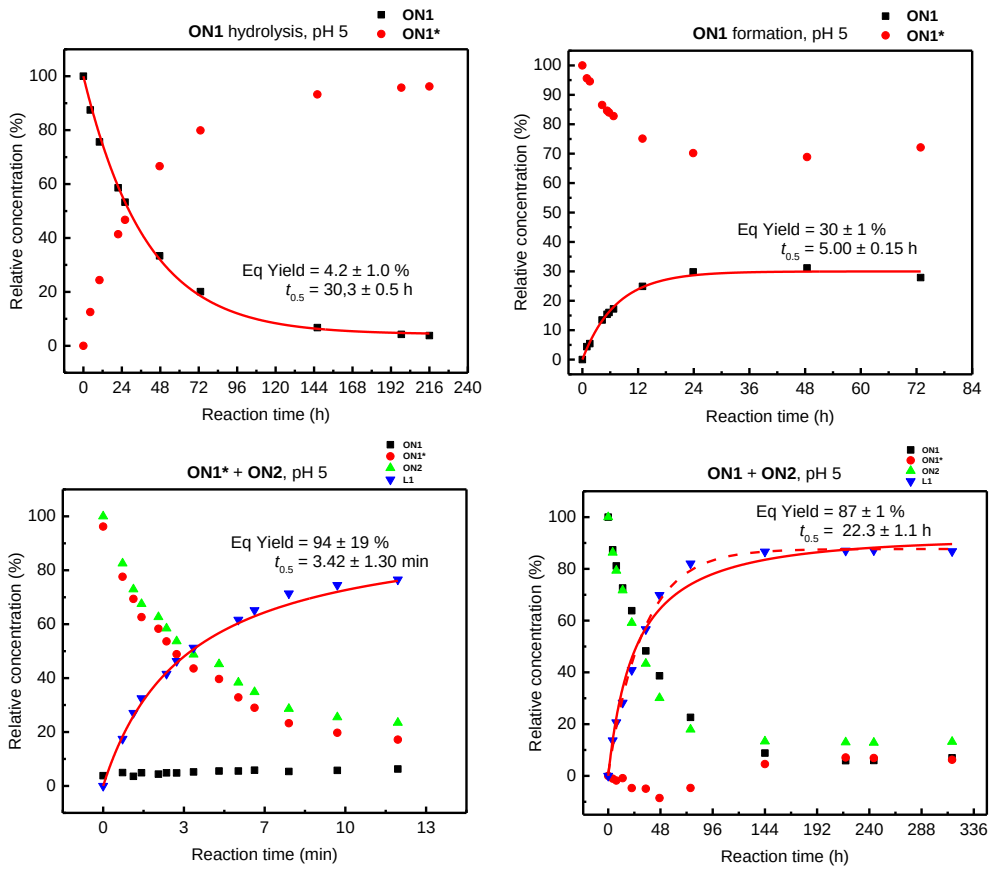


Figure S26 Kinetic profiles for the preliminary studies at pH 5. Solid line was obtained by fitting the data points to equation S1. Dashed line was obtained by fitting to equation S2.

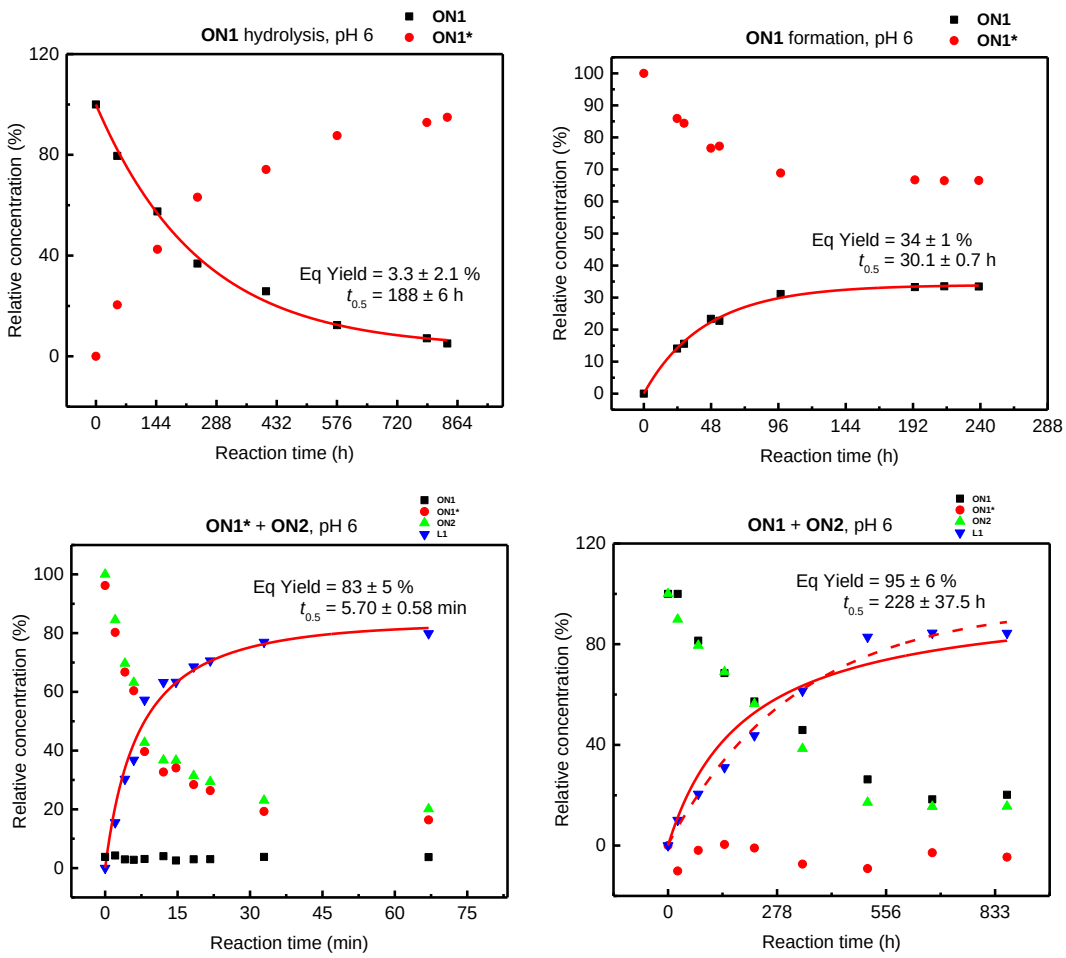


Figure S27 Kinetic profiles for the preliminary studies at pH 6. Solid line was obtained by fitting the data points to equation S1. Dashed line was obtained by fitting to equation S2.

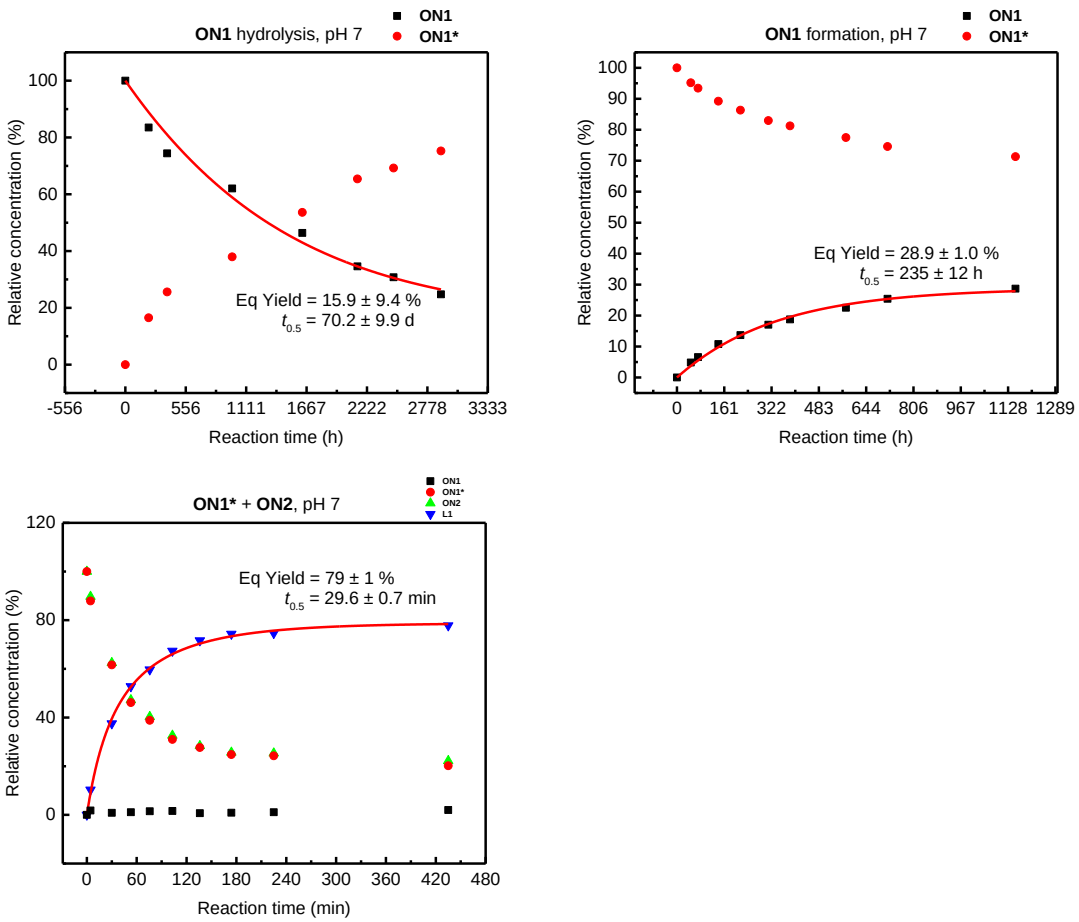


Figure S28 Kinetic profiles for the preliminary studies at pH 7. Solid line was obtained by fitting the data points to equation S1.

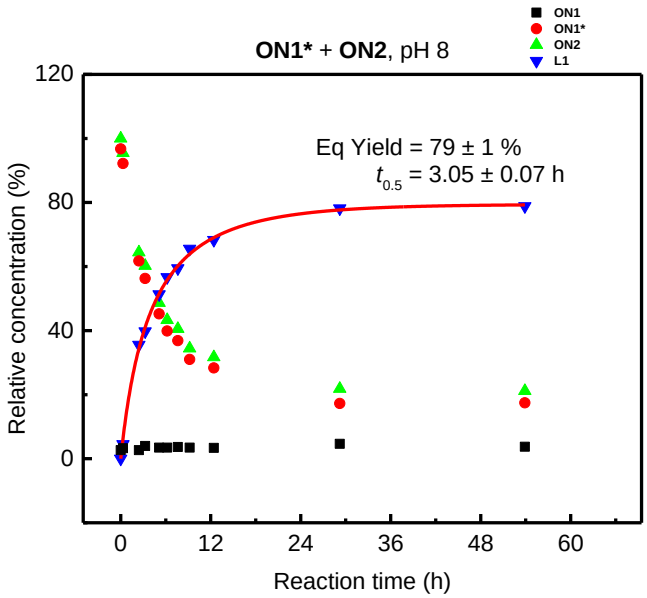


Figure S29 Kinetic profiles for the preliminary studies at pH8. Solid line was obtained by fitting the data points to equation S1.

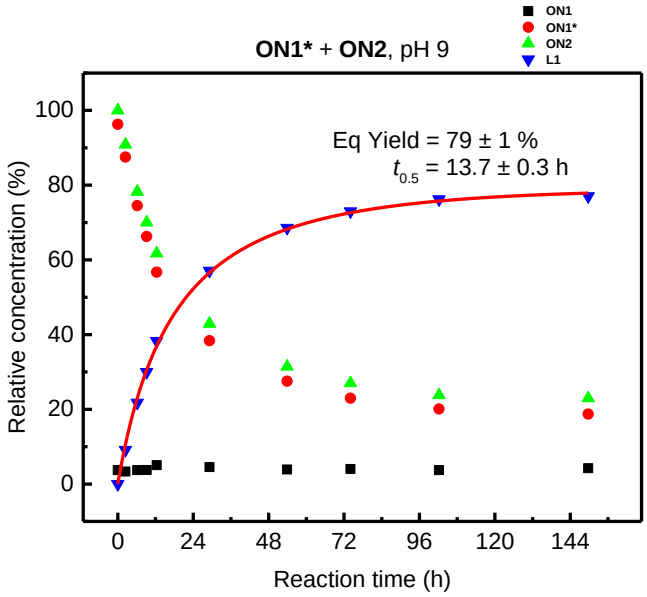


Figure S30 Kinetic profiles for the preliminary studies at pH 9. Solid line was obtained by fitting the data points to equation S1.

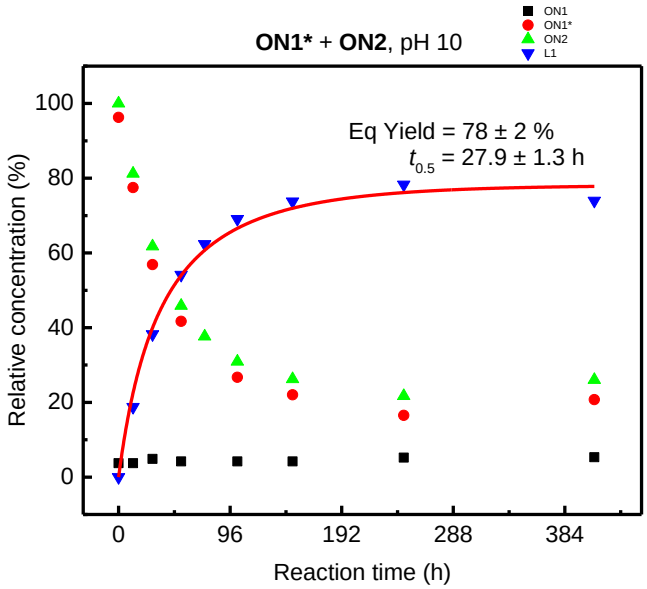


Figure S31 Kinetic profiles for the preliminary studies at pH 10. Solid line was obtained by fitting the data points to equation S1.

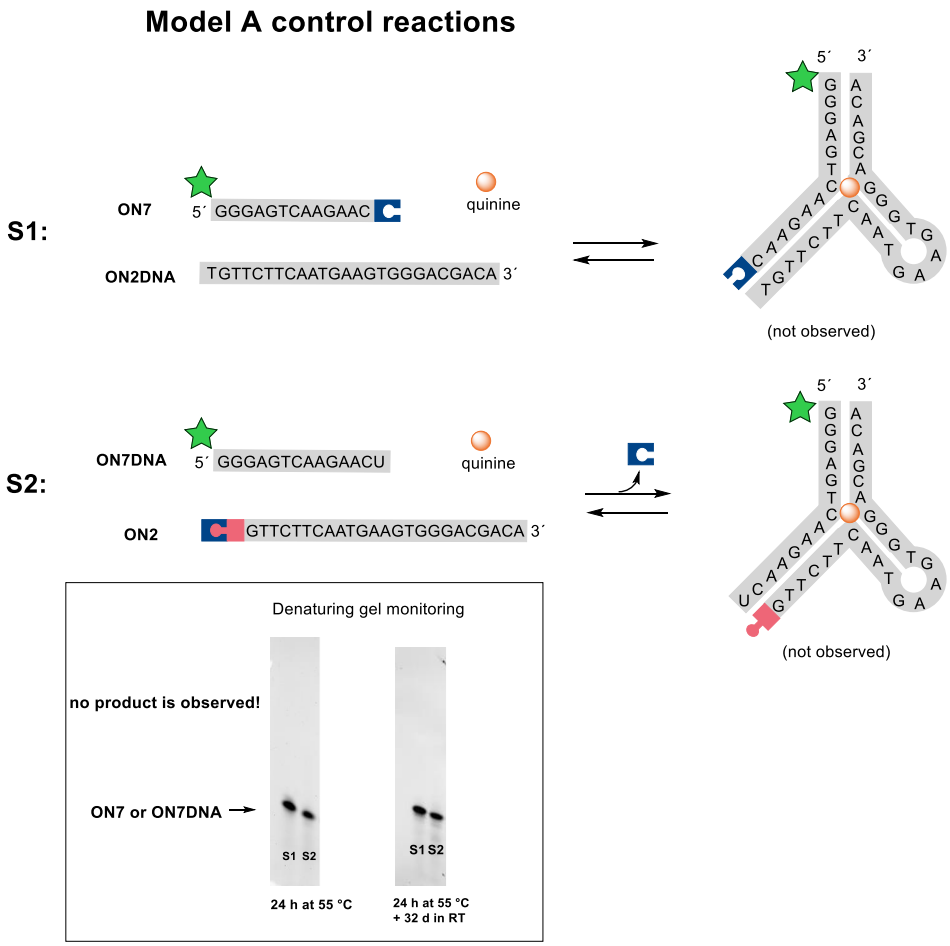
Split aptamer formation

General procedure for the quinine induced split aptamer (CBA) formation: ONs (0.5 – 2.5 nmol, 1–5 µM, see table below) were dissolved in water (500 µl) containing buffer (0.1 M AcOH/AcONa, I = 0.1 M adjusted using NaCl, pH 5) and variable amount of quinine (see table below). The mixture was placed in oven (55 °C) for overnight. A sample (5 µl) was taken from the warm solution and analyzed using denaturing PAGE. The reaction mixture was then allowed to cool to room temperature (21 °C). Samples (5 µl) were taken at intervals (see table below) at room temperature and analyzed using native PAGE. Native PAGE analyses were performed using precast TBE-Urea gel (15% acrylamide, TBE, 7 M urea), which was fixed into a vertical electrophoresis chamber, and the chamber was filled with the running buffer (90 mM Tris, 90 mM borate, and 2 mM EDTA, pH 8.3). The reaction samples (5 µl) were premixed with Novex™ TBE-urea 2x sample buffer (5 µl) and then applied to gel wells. The loaded gels were run at 200 V and 45 mA for 60 min. The completed gels were imaged by the fluorescence label in **ON6** or **ON7 (Figure S32)**. The bands containing **ON6** or **ON7** were quantified by processing the images into 2D density profiles and integrating them using ImageJ software. Mass spectra was obtained for the ligation products (**Table S6** and **Figure S33**—**Figure S36**).

Table S5 List of the split aptamer formation reactions.

SI Entry	Denoted in the main text as	ONs	*c*(quinine) / µM
S1[a]	Control	**ON2**, **ON7DNA**[d]	1000
S2[a]	Control	**ON2DNA**[e], **ON7**	1000
S3[a]	Model A	**ON2**, **ON7**	0
S4[a]	Model A	**ON2**, **ON7**	1
S5[a]	Model A	**ON2**, **ON7**	10
S6[a]	Model A	**ON2**, **ON7**	100
S7[a]	Model A	**ON2**, **ON7**	1000
S8[a]	Model B	**ON3**, **ON6**	0
S9[a]	Model B	**ON3**, **ON6**	1
S10[a]	Model B	**ON3**, **ON6**	10
S11[a]	Model B	**ON3**, **ON6**	100
S12[a]	Model B	**ON3**, **ON6**	1000
S13[b]	Thrice-split (lane 1)	**ON4**, **ON7**	0
S14[b]	Thrice-split (lane 2)	**ON3**, **ON7**	0
S15[b]	Thrice-split (lane 3)	**ON3**, **ON4**, **ON7**	0
S16[b]	Thrice-split (lane 4)	**ON4**, **ON6**	100
S17[b]	Thrice-split (lane 5)	**ON3**, **ON7**	100
S18[b]	Thrice-split (lane 6)	**ON3**, **ON4**, **ON7**	100
S19[b,c]	Thrice-split (lane 7)	**ON3**, **ON4**, **ON7**	100

[a]Concentration of each ON is 1 µM. [b]Concentration of each ON is 5 µM. [c] Additional $MgCl_2$ (50 mM) in the reaction mixture. [d]Regular DNA with 5´-fluoroscein (FAM) and 3´-uridine: FAM-GGGAGTCAAGAACu. [e]Regular DNA: TGTTCTTCAATGAAGTGGGACGACA



Scheme S3 Model A control reactions. For reaction conditions see **Table S5**

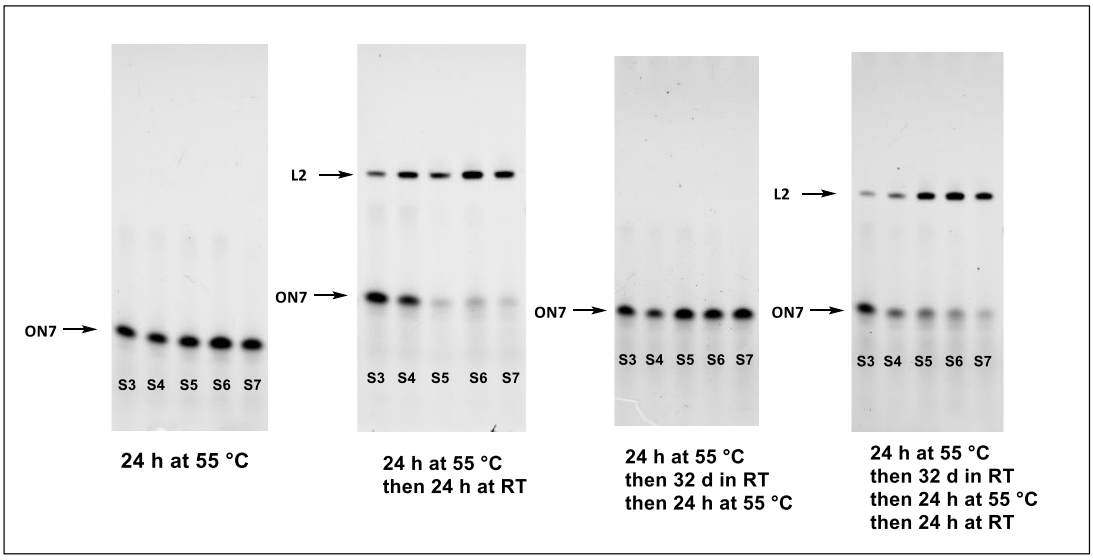


Figure S32 Reversibility of model A.

Table S6 Molecular masses obtained from the split aptamer studies.

ON	observed molecular weight	calculated molecular weight
L2	12634.5	12634.5
L3	13242.0	13244.9
L4	8936.7	8937.1
L5	13218.1	13219.9

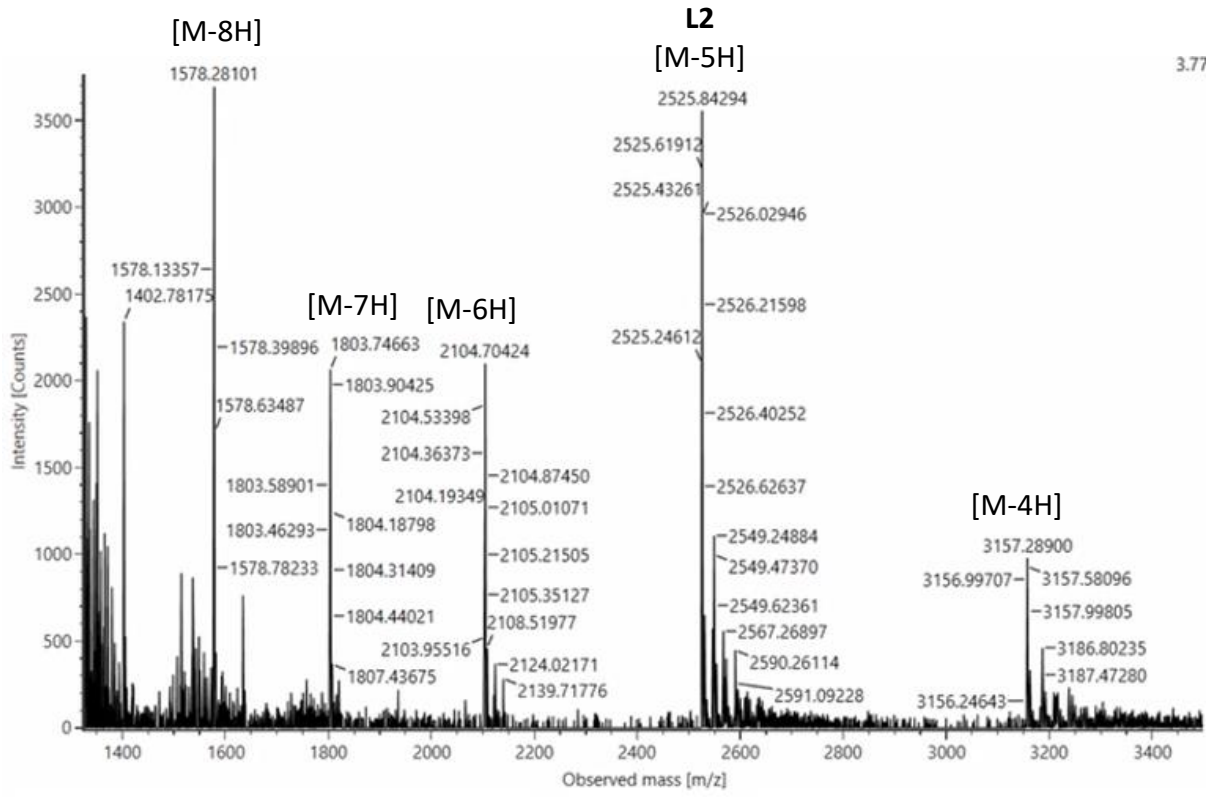


Figure S33 Mass spectrum for **L2**.

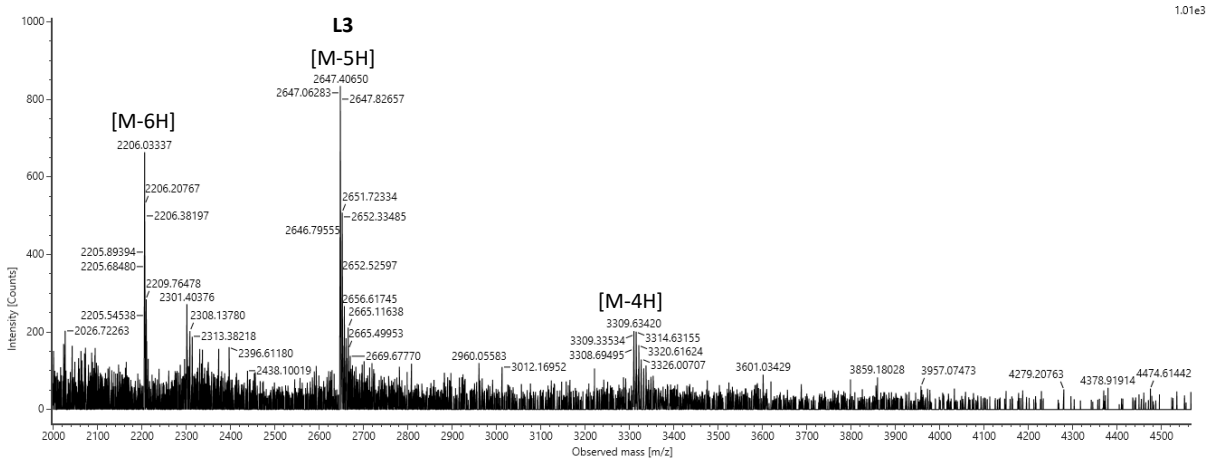


Figure S34 Mass spectrum for **L3**.

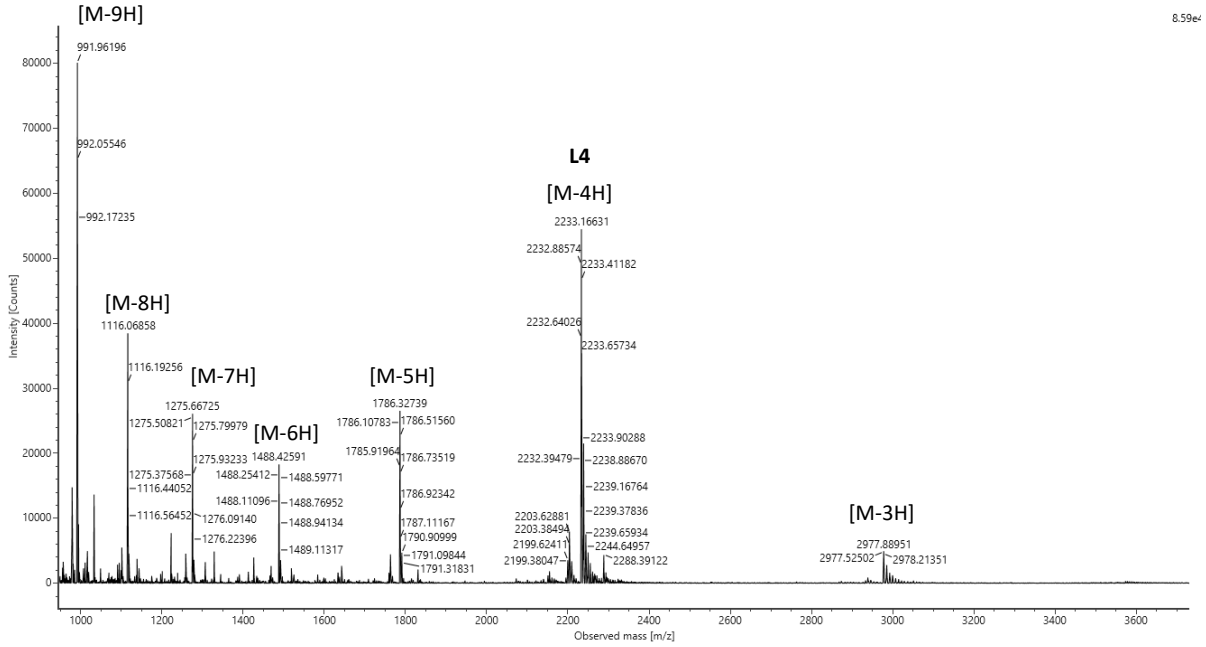


Figure S35 Mass spectrum for **L4**.

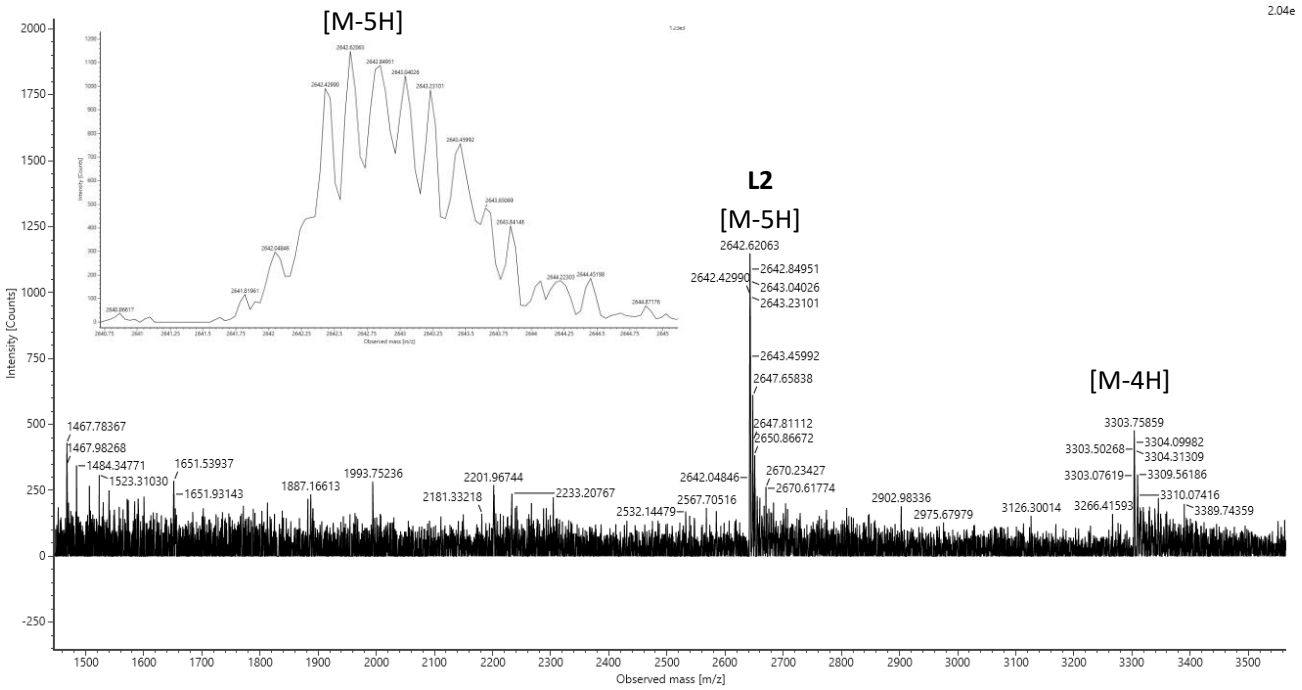


Figure S36 Mass spectrum for **L5**.

References

(S1) Yuta Ito, Airi Kimura, Takashi Osawa, and Yoshiyuki Hari, *The Journal of Organic Chemistry* **2018**, *83*, 10701-10708, DOI: 10.1021/acs.joc.8b00637