**Supporting information**

# Synthesis and properties of the anticodon stem-loop of human mitochondrial tRNA$^{Met}$ containing the disease-related G or m$^1$G nucleosides at position 37

Karolina Podskoczyj,[a] Katarzyna Kulik,[b] Joanna Wasko,[a] Barbara Nawrot,[b] Tsutomu Suzuki[c] and Grazyna Leszczynska[*a]

[a] *Institute of Organic Chemistry, Lodz University of Technology, Zeromskiego 116, 90-924 Lodz, Poland*
[b] *Department of Bioorganic Chemistry, Centre of Molecular and Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 112, 90-363 Lodz, Poland*
[c] *Depatrment of Chemistry and Biotechnology, Graduate School of Engineering, University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*
*E-mail:* grazyna.leszczynska@p.lodz.pl

**Table of Contents**

## I. General remarks

All solid compounds were dried under high vacuum prior to use. Thin layer chromatography (TLC) was carried out on silica gel coated plates ($60F_{254}$). Column chromatography was performed on silica gel 60 (230-400 mesh, Fluka).

Analytical HPLC was performed with a Shimadzu Prominence HPLC system equipped with a SPD-M20A spectral photodiode array detector using a Reprospher column (RP, 100 Å, C18, 5 µm, 250 x 4.6 mm, Dr. Maisch). Analysis of the nucleoside composition was performed on column eluted with a linear gradient of buffer A (10 mM $KH_2PO_4$, pH 5.3) and buffer B (20% MeOH in 10 mM $KH_2PO_4$, pH 5.1) at a constant flow rate of 0.75 mL/min. Control of ASL's purity was performed on column eluted with a linear gradient of buffer A (0.1 M $CH_3COONH_4$, pH 6.0) and buffer B (ACN) at a constant flow rate 1 mL/min.

Anion-exchange HPLC was performed with a Waters 515 HPLC system equipped with a 996 spectral diode array detector using a Source column (15Q 4.6/100 PE, GE Healthcare). The column was eluted with a linear gradient 50-650 mM NaBr in a 20 mM $Na_2HPO_4$-$NaH_2PO_4$ buffer solution (pH 7.5) containing 50 µM EDTA and 10% ACN at a constant flow rate of 1 mL/min.

NMR spectra were recorded on a Bruker Avance DPX 250 spectrometer at 250 (for $^{1}H$), and 101 (for $^{31}P$) MHz or a Bruker Avance II Plus 700 spectrometer at 700 (for $^{1}H$) and 176 (for $^{13}C$) MHz. Chemical shifts (δ) are reported in ppm relative to TMS (internal standard) for $^{1}H$ and $^{13}C$ or 85% $H_3PO_4$ (external standard) for $^{31}P$. Multiplicities are described as s (singlet), d (doublet), dd (doublet of doublets), dt (doublet of triplets), t (triplet), q (quartet), m (multiplet), and bs (broad singlet). Coupling constants (*J*) are reported in hertz.

High-resolution mass spectrometry (HRMS) measurements were performed using Synapt G2Si mass spectrometer (Waters) equipped with an ESI source and quadrupole-Time-of-flight mass analyzer. The measurement was performed in negative ion mode with capillary voltage set to 2.7 kV and sampling cone to 20 V. The source temperature was 110 °C. To ensure accurate mass measurements, data were collected in centroid mode and mass was corrected during acquisition using leucine enkephalin solution as an external reference (Lock-SprayTM), which generated reference ion at m/z 554.2614Da ([M+H]-) in negative ESI mode. The results of the measurements were processed using the MassLynx 4.1 software (Waters).

Electrospray mass spectrometry measurements were performed using Synapt G2-Si mass spectrometer (Waters) equipped with quadrupole-Time-of-flight mass analyser. The mass spectrometer was operated in the negative ion detection mode. The results of the measurements were processed using the MassLynx 4.1 software (Waters) incorporated with the instrument.

## II. Chemical synthesis of 5-(1,2-diacethoxyethyl)-*N*4-acetylcytidine phosphoramidite 1

### 2'-*O*-(*tert*-butyldimethylsilyl)-3',5'-*O*-(di-*tert*-butylsilylene)cytidine (3)

Title compound **3** was synthesized following a procedure described in the literature.[1] Cytidine **2** (5.00 g, 20.56 mmol) was dissolved in anhydrous DMF (50 mL) and stirred at 0 ºC. Next, triflic acid (1.8 mL, 20.56 mmol) and di-*tert*-butylsilyl bis(trifluoromethanesulfonate) (8.0 mL, 24.67 mmol) were added. After 45 minutes, imidazole (7.00 g, 102.75 mmol) was added, and the reaction was warmed to room temperature over a period of 30 min. Subsequently, *tert*-butyldimethylsilyl chloride (4.00 g, 26.72 mmol) was added in one portion, and the reaction was stirred at 60 ºC for 1.5 h. Then, the reaction mixture was diluted with EtOAc (150 mL) and extracted with sat. $NaHCO_3$ (2 x 100 mL) and $H_2O$ (2 x 100 mL). The organic layer was dried over $MgSO_4$, filtered and evaporated under reduced pressure. The product **3** was isolated by column chromatography (silica gel, 0-3% MeOH in $CH_2Cl_2$) as a white solid (8.58 g, 84%). TLC: $R_f$ = 0.36 ($CHCl_3$/MeOH, 95:5, v/v); **$^1$H NMR** (700 MHz, $CDCl_3$) δ: 7.35 (d, 1H, *J* = 7.7 Hz, H-6), 5.71 (d, 1H, *J* = 7.7 Hz, H-5), 5.68 (s, 1H, H-1'), 4.51 (dd, 1H, *J* = 9.1 Hz, *J* = 5.6 Hz, H-5'), 4.32 (d, 1H, *J* = 4.9 Hz, H-2'), 4.22-4.19 (m, 1H, H-4'), 3.98 (dd, 1H, *J* = 9.1 Hz, *J* = 1.4 Hz, H-5"), 3.85 (dd, 1H, *J* = 9.8 Hz, *J* = 4.9 Hz, H-3'), 1.04 (s, 9H, Si-C$(CH_3)_3$), 1.02 (s, 9H, Si-C$(CH_3)_3$), 0.94 (s, 9H, Si-C$(CH_3)_3$), 0.22 (s, 3H, Si-$CH_3$), 0.15 (s, 3H, Si-$CH_3$). Analytical data was consistent with the literature.[1]

### 2'-*O*-(*tert*-butyldimethylsilyl)-3',5'-*O*-(di-*tert*-butylsilylene)-5-iodocytidine (4)

Title compound **4** was synthesized following a procedure described in the literature.[2] To a solution of compound **3** (3.50 g, 7.03 mmol) in anhydrous $CH_3CN$ (50 mL), iodine (1.34 g, 5.27 mmol) and ammonium cerium (IV) nitrate (5.70 g, 10.4 mmol) were added. The mixture was stirred at 80 ºC for 3 h under reflux. Subsequently, the reaction mixture was diluted with EtOAc (250 mL) and extracted with sat. $NaHCO_3$ (200 mL) and sat. $Na_2S_2O_5$ (3 x 100 mL) until the yellow color disappeared. The organic phase was dried over $MgSO_4$, filtered and evaporated under reduced pressure. The product **3** was isolated by column chromatography (silica gel, 0-1% MeOH in $CHCl_3$) as a yellow solid (3.28 g, 75%). TLC: $R_f$ = 0.43 ($CHCl_3$/MeOH, 95:5, v/v); **$^1$H NMR** (700 MHz, $CDCl_3$) δ: 7.66 (s, 1H, H-6), 5.65 (s, 1H, H-1'), 4.54 (dd, 1H, *J* = 9.1 Hz, *J* = 5.6 Hz, H-5'), 4.31 (d, 1H, *J* = 4.2 Hz, H-2'), 4.24-4.20 (m, 1H, H-4'), 3.98 (dd, 1H, *J* = 9.1 Hz, *J* = 1.4 Hz, H-5"), 3.82 (dd, 1H, *J* = 9.1 Hz, *J* = 4.2 Hz, H-3'), 1.05 (s, 9H, Si-C$(CH_3)_3$), 1.03 (s, 9H, Si-C$(CH_3)_3$), 0.94 (s, 9H, Si-C$(CH_3)_3$), 0.23 (s, 3H, Si-$CH_3$), 0.15 (s, 3H, Si-$CH_3$). Analytical data was consistent with the literature.[2]

### 2'-*O*-(*tert*-butyldimethylsilyl)-3',5'-*O*-(di-*tert*-butylsilylene)-5-vinylcytidine (5)

Compound **4** (3.00 g, 4.81 mmol) was co-evaporated with anhydrous DMF (2 x 15 mL), dissolved in anhydrous DMF (30 mL) and treated with $(Ph_3P)_2PdCl_2$ (0.34 g, 0.48 mmol). Then, tributyl(vinyl)tin (1.82 mL, 6.25 mmol) was added dropwise and the reaction mixture was stirred at 80 ºC for 1.5 h. The mixture was diluted with EtOAc (120 mL) and filtered through a Celite pad. The organic layer was washed with 5% aq. $NaHCO_3$ (2 x 50 mL), $H_2O$ (50 mL) and sat. NaCl (50 mL), dried over anhydrous $MgSO_4$ and filtered. The filtrated solution

was evaporated under reduced pressure. The oily residue was co-evaporated with anhydrous toluene (2 x 15 mL). The product **5** was isolated by column chromatography (silica gel, 0-2% MeOH in $CH_2Cl_2$) as a yellow solid (1.92 g, 76%). TLC: $R_f$ = 0.32 ($CHCl_3$/MeOH, 95:5, v/v); **$^1$H NMR** (700 MHz, DMSO-$d_6$) δ: 7.62 (s, 1H, H-6), 7.49 (bs, 1H, NH), 7.12 (bs, 1H, NH), 6.56 (dd, 1H, *J* = 17.5 Hz, *J* = 11.2 Hz, H-1"), 5.64 (s, 1H, H-1'), 5.44 (dd, 1H, *J* = 17.5 Hz, *J* = 1.4 Hz, H-2"(E)), 5.13 (dd, 1H, *J* = 11.2 Hz, *J* = 1.4 Hz, H-2"(Z)), 4.43 (dd, 1H, *J* = 4.9 Hz, *J* = 0.7 Hz, H-2'), 4.40 (dd, 1H, *J* = 8.4 Hz, *J* = 4.2 Hz, H-5'), 4.16 (dd, 1H, *J* = 9.1 Hz, *J* = 5.6 Hz, H-4'), 3.95-3.87 (m, 2H, H-3', H-5"), 1.03 (s, 9H, Si-C($CH_3$)$_3$), 0.99 (s, 9H, Si-C($CH_3$)$_3$), 0.88 (s, 9H, Si-C($CH_3$)$_3$), 0.13 (s, 3H, Si-$CH_3$), 0.08 (s, 3H, Si-$CH_3$); **$^{13}$C NMR** (176 MHz, DMSO-$d_6$) δ: 163.51 (C-4), 153.81 (C-2), 138.45 (C-6), 128.07 (C-1"), 113.98 (C-2"), 105.02 (C-5), 94.83 (C-1'), 75.22 (C-4'), 74.10 (C-2'), 73.70 (C-3'), 67.07 (C-5'), 27.22 (Si-C(<u>C</u>$H_3$)$_3$), 26.75 (Si-C(<u>C</u>$H_3$)$_3$), 25.68 (Si-C(<u>C</u>$H_3$)$_3$), 22.22 (Si-<u>C</u>($CH_3$)$_3$), 19.86 (Si-<u>C</u>($CH_3$)$_3$), 17.99 (Si-<u>C</u>($CH_3$)$_3$), -4.63 (Si-$CH_3$), -5.11 (Si-$CH_3$); HRMS (ESI) calcd for $C_{25}H_{46}N_3O_5Si_2$ $[M + H]^+$ 524.2976, found 524.2979.

**2'-*O*-(*tert*-butyldimethylsilyl)-3',5'-*O*-(di-*tert*-butylsilylene)-5-(1,2-dihydroxyethyl)cytidine (6):**

To a solution of **5** (1.50 g, 2.86 mmol) in acetone/$H_2O$/*t*-BuOH (40 mL, 4:1:1 v/v/v), *N*-methylmorpholine *N*-oxide (0.67 g, 5.72 mmol) and a solution of $OsO_4$/*t*-BuOH (5.0 mg/mL, 7.27 mL, 0.143 mmol) were added. The resulting mixture was stirred at room temperature for 4 h. Subsequently, the reaction was quenched with 10% aq. $Na_2S_2O_3$ (70 mL) and extracted with EtOAc (3 x 100 mL). The organic layer was washed with 10% aq. $Na_2S_2O_3$ (100 mL), sat. NaCl (100 mL), dried over anhydrous $MgSO_4$ and filtered. The filtrate was evaporated under reduced pressure. The product **6** was isolated by column chromatography (silica gel, 0-8% MeOH in $CHCl_3$) as a light yellow solid (1.26 g, 79%). TLC: $R_f$ = 0.46 ($CHCl_3$/MeOH, 85:15, v/v); **$^1$H NMR** (700 MHz, DMSO-$d_6$) δ: 7.41 (s, 0.45H, H-6 isomer A), 7.36 (s, 0.56H, H-6 isomer B), 5.65 (bs, 0.55H, H-1' isomer B), 5.64 (bs, 0.45H, H-1' isomer A), 5.44 (d, 0.55H, *J* = 4.9 Hz, CH-O<u>H</u> isomer B), 5.33 (d, 0.44H, *J* = 4.9 Hz, CH-O<u>H</u> isomer A), 4.71 (t, 0.55H, *J* = 5.6 Hz, $CH_2$-O<u>H</u> isomer B), 4.68 (t, 0.45H, *J* = 5.6 Hz, $CH_2$-O<u>H</u> isomer A), 4.45-4.39 (m, 2H, C<u>H</u>-OH, H-5'), 4.26 (dd, 1H, *J* = 6.3 Hz, *J* = 4.2 Hz, H-2'), 4.03-3.99 (m, 1H, H-4'), 3.97-3.95 (m, 2H, H-3', H-5"), 3.42 (m, 2H, C<u>H$_2$</u>-OH), 1.02 (s, 9H, Si-C($CH_3$)$_3$), 0.99 (s, 9H, Si-C($CH_3$)$_3$), 0.90 (s, 4.1H, Si-C($CH_3$)$_3$ isomer A), 0.89 (s, 4.9H, Si-C($CH_3$)$_3$ isomer B), 0.15 (s, 1.37H, Si-$CH_3$ isomer A), 0.14 (s, 1.63H, Si-$CH_3$ isomer B), 0.01 (s, 1.36H, Si-$CH_3$ isomer A), 0.09 (s, 1.64H, Si-$CH_3$ isomer B); **$^{13}$C NMR** (176 MHz, DMSO-$d_6$) δ: 164.23 (C-4 isomer B), 164.19 (C-4 isomer A), 154.14 (C-2 isomer A), 154.11 (C-2 isomer B), 138.78 (C-6 isomer B), 138.70 (C-6 isomer A), 106.63 (C-5 isomer A), 106.63 (C-5 isomer B), 93.73 (C-1' isomer B), 93.67 (C-1' isomer A), 75.14 (C-4' isomer B), 74.10 (C-4' isomer A), 74.62 (C-2' isomer B), 74.59 (C-2' isomer A), 73.79 (C-3' isomer A), 73.76 (C-3' isomer B), 68.64 (CH-OH isomer B), 67.70 (CH-OH isomer A), 67.02 (C-5'), 64.50 ($CH_2$-OH isomer A), 64.28 ($CH_2$-OH isomer B), 27.25 (Si-C(<u>C</u>$H_3$)$_3$), 26.73 (Si-C(<u>C</u>$H_3$)$_3$), 25.69 (Si-C(<u>C</u>$H_3$)$_3$), 22.17 (Si-<u>C</u>($CH_3$)$_3$), 19.87 (Si-<u>C</u>($CH_3$)$_3$), 17.95 (Si-<u>C</u>($CH_3$)$_3$ isomer A), 17.93 (Si-<u>C</u>($CH_3$)$_3$ isomer B), -4.67 (Si-$CH_3$), -5.03 (Si-$CH_3$ isomer A), -5.05 (Si-$CH_3$ isomer B); HRMS (ESI) calcd for $C_{25}H_{48}N_3O_7Si_2$ $[M + H]^+$ 558.3031, found 558.3026.

**2'-*O*-(*tert*-butyldimethylsilyl)-3',5'-*O*-(di-*tert*-butylsilylene)-5-(1,2-diacetoxyethyl)-*N*4-acetylcytidine (7):**

To a solution of compound **6** (1.20 g, 2.15 mmol) in dry pyridine (28 mL), DMAP (25.7 mg, 0.21 mmol) and $Ac_2O$ (0.91 mL, 9.67 mmol) were added. The mixture was stirred at room temperature for 20 h. Then, EtOH (28 mL) was added and the resulting mixture was evaporated under reduced pressure. The residue was co-evaporated with toluene (30 mL), dissolved in EtOAc (300 mL) and washed with $H_2O$ (3 x 130 mL), and sat. NaCl (130 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered and evaporated under reduced pressure. The crude product was purified by column chromatography (silica gel, $CHCl_3$) to give **7** as a white foam (1.15 g, 78%). TLC: $R_f$ = 0.39 ($CHCl_3$/MeOH, 98:2, v/v); **$^1$H NMR** (700 MHz, $CDCl_3$) δ: 13.03 (bs, 1H, NH), 7.48 (s, 1H, H-6), 6.17-6.15 (m, 1H, C<u>H</u>-OAc), 5.68 (s, 1H, H-1'), 4.59 (dd, 1H, *J* = 9.1 Hz, *J* = 5.6 Hz, H-5'), 4.44 (dd, 1H, *J* = 11.9 Hz, *J* = 2.8 Hz, C<u>H$_2$</u>-OAc), 4.29-4.24 (m, 3H, C<u>H$_2$</u>-OAc, H-2', H-4'), 3.99-3.96 (m, 1H, H-5"), 3.79 (dd, 1H, *J* = 9.8 Hz, *J* = 4.2 Hz, H-3'), 2.24 (s, 3H, NH-CO-C<u>H$_3$</u>), 2.12 (s, 3H, CO-$CH_3$), 2.04 (s, 3H, CO-$CH_3$), 1.04 (s, 9H, Si-C($CH_3$)$_3$), 1.03 (s, 9H, Si-C($CH_3$)$_3$), 0.95 (s, 9H, Si-C($CH_3$)$_3$), 0.21 (s, 3H, Si-$CH_3$), 0.15 (s, 3H, Si-$CH_3$); **$^{13}$C NMR** (176 MHz, $CDCl_3$) δ: 188.00 (NH-<u>C</u>O-$CH_3$), 170.53 (<u>C</u>O-$CH_3$), 169.45 (<u>C</u>O-$CH_3$), 156.23 (C-4), 146.99 (C-2), 138.62 (C-6), 109.80 (C-5), 94.43 (C-1'), 76.00 (C-4'), 75.50 (C-2'), 74.91 (C-3'), 68.01 (<u>C</u>H-OAc), 67.67 (C-5'), 64.21 (<u>C</u>H$_2$-OAc), 29.00 (CO-<u>C</u>H$_3$), 27.59 (Si-C(<u>C</u>H$_3$)$_3$), 27.10 (Si-C(<u>C</u>H$_3$)$_3$), 25.97 (Si-C(<u>C</u>H$_3$)$_3$), 22.95 (Si-<u>C</u>(CH$_3$)$_3$), 20.96 (CO-<u>C</u>H$_3$), 20.94 (CO-<u>C</u>H$_3$), 20.49 (Si-<u>C</u>(CH$_3$)$_3$), 18.38 (Si-<u>C</u>(CH$_3$)$_3$), -4.18 (Si-$CH_3$), -4.91 (Si-$CH_3$); HRMS (ESI) calcd for $C_{31}H_{54}N_3O_{10}Si_2$ $[M + H]^+$ 684.3348, found, 684.3340.

**2'-*O*-(*tert*-butyldimethylsilyl)-5-(1,2-diacetoxyethyl)-*N*4-acetylcytidine (8):**

Compound **7** (1.00 g, 1.46 mmol) was dissolved in anhydrous $CH_2Cl_2$ (12 mL) and cooled to 0 ºC. A mixture of anhydrous pyridine (900 μL) and HF-pyridine (146 μL, 5.62 mmol) was added to the solution. The reaction mixture was stirred at 0 ºC for 1.5 h. Then, the reaction was diluted with $CH_2Cl_2$ (100 mL) and extracted with sat. $NaHCO_3$ (2 x 150 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered and evaporated under reduced pressure. The residue was co-evaporated with toluene (2 x 15 mL), and the product **8** was isolated by column chromatography (silica gel, 0-2% MeOH in $CH_2Cl_2$) as a white foam (0.53 g, 67%). TLC: $R_f$ = 0.36 ($CHCl_3$/MeOH, 95:5, v/v); **$^1$H NMR** (700 MHz, DMSO-$d_6$) δ: 10. 21 (s, 0.29H, NH isomer A), 10.14 (s, 0.71H, NH isomer B), 8.71 (s, 0.28H, H-6 isomer A), 8.70 (s, 0.73H, H-6 isomer B), 5.87 (m, 1H, C<u>H</u>-OAc), 5.71 (d, 0.71H, *J* = 2.8 Hz, H-1' isomer B), 5.70 (d, 0.29H, *J* = 2.1 Hz, H-1' isomer A), 5.32 (t, 0.48H, *J* = 4.2 Hz, 5'-OH isomer A), 5.30 (t, 0.78H, *J* = 4.2 Hz, 5'-OH isomer B), 5.00 (d, 0.29H, *J* = 5.6 Hz, 3'-OH isomer A), 4.99 (d, 0.71H, *J* = 5.6 Hz, 3'-OH isomer B), 4.29-4.25 (m, 2H, C<u>H$_2$</u>-OAc), 4.13-4.10 (m, 1H, H-2'), 3.98-3.95 (m, 2H, H-3', H-4'), 3.82-3.79 (m, 1H, H-5'), 3.66-3.63 (m, 1H, H-5"), 2.22 (s, 2.1H, NH-CO-C<u>H$_3$</u> isomer B), 2.20 (s, 0.9H, NH-CO-C<u>H$_3$</u> isomer A), 2.02 (s, 6H, 2xCO-$CH_3$), 0.86 (s, 9H, Si-C($CH_3$)$_3$), 0.07 (s, 3H, Si-$CH_3$), 0.05 (s, 3H, Si-$CH_3$); **$^{13}$C NMR** (176 MHz, DMSO-$d_6$) δ: 170.96 (NH-<u>C</u>O-$CH_3$), 170.01 (<u>C</u>O-$CH_3$), 169.50 (<u>C</u>O-$CH_3$), 160.17 (C-4), 153.52 (C-2), 144.47 (C-6), 106.46 (C-5), 90.68 (C-1' isomer A), 90.48 (C-1' isomer B), 83.90 (C-4' isomer B), 83.57 (C-4' isomer A), 76.63 (C-2' isomer B), 76.20 (C-2' isomer A), 67.95

(C-3' isomer B), 67.57 (C-3' isomer A), 66.98 (CH-OAc), 63.97 ($CH_2$-OAc), 59.15 (C-5' isomer B), 58.94 (C-5' isomer A), 25.72 (Si-C($CH_3$)$_3$), 24.51 (NH-CO-$CH_3$ isomer B), 24.34 (NH-CO-$CH_3$ isomer A), 20.67 (CO-$CH_3$), 20.48 (CO-$CH_3$), 17.92 (Si-C($CH_3$)$_3$), -4.84 (Si-$CH_3$), -4.97 (Si-$CH_3$); HRMS (ESI) calcd for $C_{23}H_{38}N_3O_{10}Si$ $[M + H]^+$ 544.2326, found 544.2330.

**2'-*O*-(*tert*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-(1,2-diacetoxyethyl)-*N*4-acetylcytidine (9):**

To a solution of **8** (0.45 g, 0.83 mmol) in anhydrous pyridine (10 mL) DMTr-Cl (0.43 g, 1.25 mmol) was added. The reaction mixture was stirred for 24 h at room temperature. Then, the reaction was quenched with $H_2O$ (10 mL), and extracted with $CH_2Cl_2$ (3 x 25 mL). Combined organic layers were dried over anhydrous $MgSO_4$, filtered and evaporated under reduced pressure. The oily residue was co-evaporated with toluene (3 x 15 mL). The product **9** was isolated by column chromatography (silica gel, 0-8% acetone in $CH_2Cl_2$) as a yellow oil (0.56 g, 80%). TLC: $R_f$ = 0.39 ($CHCl_3$/MeOH, 98:2, v/v); **$^1$H NMR** (700 MHz, DMSO-$d_6$) δ: 10.21 (bs, 0.43H, NH isomer A), 10.09 (bs, 0.37H, NH isomer B), 7.95 (s, 0.44H, H-6 isomer A), 7.91 (s, 0.36H, H-6 isomer B), 7.42-7.40 (m, 2H, $H_{Ar}$ DMTr), 7.32-7.27 (m, 6H, $H_{Ar}$ DMTr), 7.23-7.21 (m, 1H, $H_{Ar}$ DMTr), 6.90-6.88 (m, 4H, $H_{Ar}$ DMTr), 5.89-5.88 (m, 0.48H, CH-OAc isomer A), 5.81-5.79 (m, 0.43H, CH-OAc isomer B), 5.78 (d, 0.46H, $J$ = 3.5 Hz, H-1' isomer B), 5.77 (d, 0.54H, $J$ = 2.8 Hz, H-1' isomer A), 5.05 (d, 0.5H, $J$ = 6.3 Hz, 3'-OH isomer A), 5.02 (d, 0.43H, $J$ = 6.3 Hz, 3'-OH isomer B), 4.25-4.05 (m, 4H, H-2', H-4', $CH_2$-OAc), 3.85-3.81 (m, 0.54H, H-3' isomer A), 3.80-3.76 (m, 0.46H, H-3' isomer B), 3.73 (m, 6H, 2xO-$CH_3$ DMTr), 3.29-3.26 (m, 2H, H-5', H-5"), 2.24 (s, 1.16H, CO-$CH_3$ isomer B), 2.22 (s, 1.32H, CO-$CH_3$ isomer A), 1.95-1.94 (m, 3H, CO-$CH_3$), 1.78 (s, 1.28H, CO-$CH_3$ isomer A), 1.73 (s, 1.18H, CO-$CH_3$ isomer B), 0.86 (s, 9H, Si-C($CH_3$)$_3$), 0.07-0.03 (m, 6H, 2xSi-$CH_3$); **$^{13}$C NMR** (176 MHz, DMSO-$d_6$) δ: 170.88 (NH-CO-$CH_3$), 169.76 (CO-$CH_3$), 169.26 (CO-$CH_3$ isomer A), 169.08 (CO-$CH_3$ isomer B), 160.64 (C-4 isomer B), 160.24 (C-4 isomer A), 158.11 ($C_{Ar}$ DMTr), 153.27 (C-2), 144.75 ($C_{Ar}$ DMTr isomer A), 144.72 ($C_{Ar}$ DMTr isomer B), 144.02 (C-6 isomer A), 143.43 (C-6 isomer B), 135.45 ($C_{Ar}$ DMTr isomer B), 135.39 ($C_{Ar}$ DMTr isomer A), 135.28 ($C_{Ar}$ DMTr), 129.65 ($C_{Ar}$ DMTr), 127.86 ($C_{Ar}$ DMTr isomer B), 127.82 ($C_{Ar}$ DMTr isomer A), 127.63 ($C_{Ar}$ DMTr isomer A), 127.59 ($C_{Ar}$ DMTr isomer B), 126.68 ($C_{Ar}$ DMTr), 113.23 ($C_{Ar}$ DMTr), 113.19 ($C_{Ar}$ DMTr), 106.64 (C-5 isomer A), 106.15 (C-5 isomer B), 91.89 (C-1' isomer A), 91.24 (C-1' isomer B), 85.62 ($C^{IV}$ DMTr isomer A), 85.54 ($C^{IV}$ DMTr isomer B), 82.74 (C-4' isomer A), 82.51 (C-4' isomer B), 75.46 (C-2' isomer B), 75.27 (C-2' isomer A), 69.56 (C-3' isomer A), 69.49 (C-3' isomer B), 66.97 (CH-OAc isomer A), 66.87 (CH-OAc isomer B), 64.02 (C-5'), 63.84 ($CH_2$-OAc), 54.97 (O-$CH_3$ DMTr), 25.66 (Si-C($CH_3$)$_3$), 24.59 (CO-$CH_3$ isomer A), 24.40 (CO-$CH_3$ isomer B), 20.41 (CO-$CH_3$ isomer A), 20.38 (CO-$CH_3$ isomer B), 20.21 (CO-$CH_3$), 17.90 (Si-C($CH_3$)$_3$), -4.81 (Si-$CH_3$ isomer B), -4.84 (Si-$CH_3$ isomer A), -5.09 (Si-$CH_3$); HRMS (ESI) calcd for $C_{44}H_{56}N_3O_{12}Si$ $[M + H]^+$ 846.3633, found 846.3640.

**2'-*O*-(*tert*-butyldimethylsilyl)-3'-(2-cyanoethyl-*N,N*-diisopropylphosphoramidite)-5'-*O*-(4,4'-dimethoxytrityl)-5-(1,2-diacetoxyethyl)-*N*4-acetylcytidine (1):**

Compound **9** (0.35 g, 0.41 mmol) was dissolved in anhydrous $CH_2Cl_2$ (2.1 mL) and DIPEA (286 μL, 1.64 mmol), and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (183 μL, 0.82 mmol) were added. The reaction mixture was stirred at room temperature for 4 h under argon atmosphere. Then, the reaction was diluted with $CH_2Cl_2$ (8 mL) and washed with 5% aq. $NaHCO_3$ (10 mL) and $H_2O$ (10 mL). The organic layer was dried over anhydrous $MgSO_4$, filtered and evaporated under reduced pressure. The product **1** was isolated by flash chromatography (silica gel, 33% acetone in petroleum ether) as a white solid (0.34 g, 79%). TLC: $R_f$ = 0.61 (petroleum ether/acetone, 2:1, v/v); **$^1$H NMR** (700 MHz, $CDCl_3$) δ: 7.97-7.61 (m, 1H, H-6), 7.47-7.44 (m, 2H, $H_{Ar}$ DMTr), 7.36-7.28 (m, 6H, $H_{Ar}$ DMTr), 7.24-7.22 (m, 1H, $H_{Ar}$ DMTr), 6.86-6.81 (m, 4H, $H_{Ar}$ DMTr), 6.12-5.83 (m, 2H, C<u>H</u>-OAc, H-1'), 4.46-3.87 (m, 5H, C<u>H$_2$</u>-OAc, H-2', H-3', H-4'), 3.78 (s, 6H, 2xO-$CH_3$ DMTr), 3.75-3.21 (m, 6H, H-5', H-5", $CH_2$-OP, 2xN-C<u>H</u>$(CH_3)_2$), 2.66-1.97 (m, 11H, $CH_2$-CN, 3xCO-$CH_3$), 1.26-0.87 (m, 21H, 2xN-CH(C<u>H$_3$</u>)$_2$, Si-C$(CH_3)_3$), 0.25-0.03 (m, 6H, 2xSi-$CH_3$); **$^{31}$P NMR** (101 MHz, $C_6H_6$) δ: 150.88, 150.21, 149.47; HRMS (ESI) calcd for $C_{53}H_{73}N_5O_{13}SiP$ $[M + H]^+$ 1046.4712, found 1046.4734.

III. **NMR spectra of synthesized compounds 3-9 and 1**

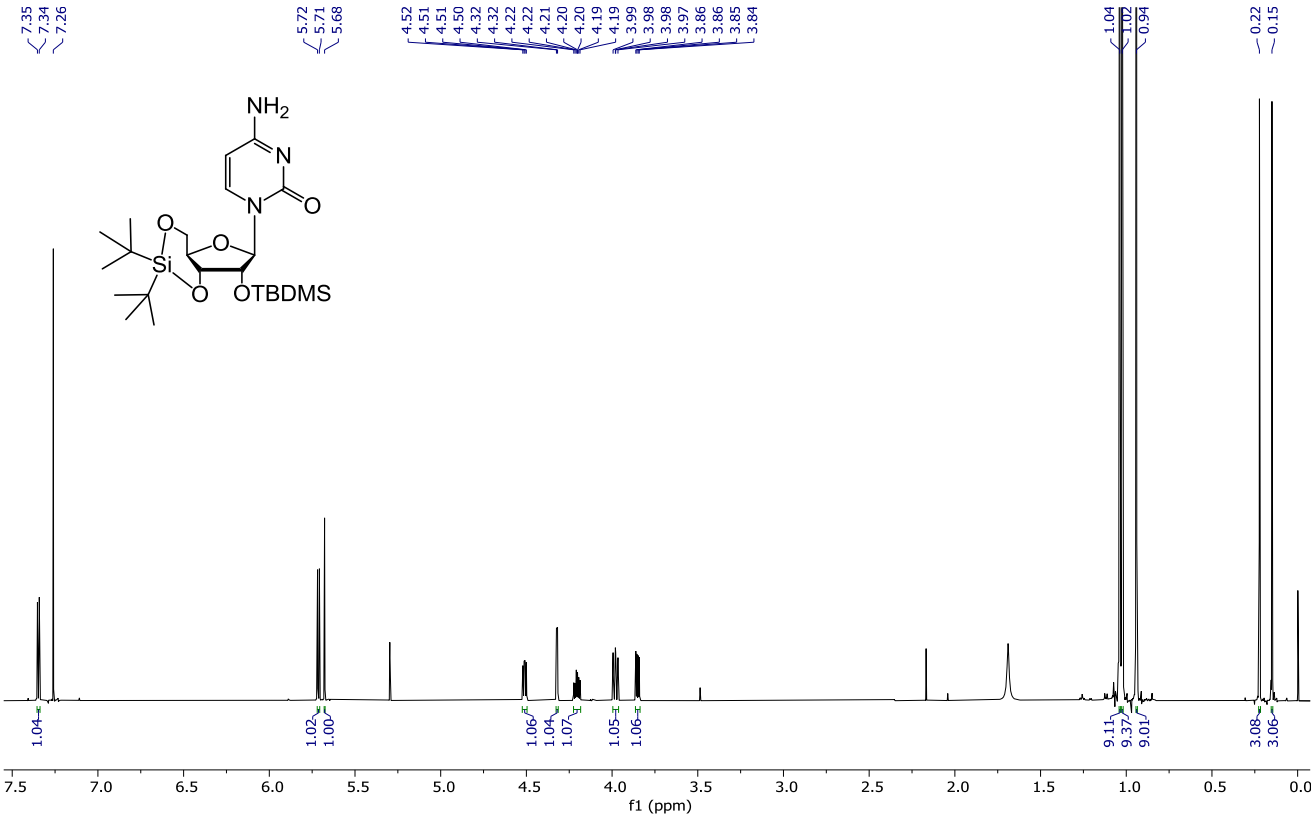


**Figure S1.** $^{1}$H NMR spectrum of **3** ($CDCl_3$).

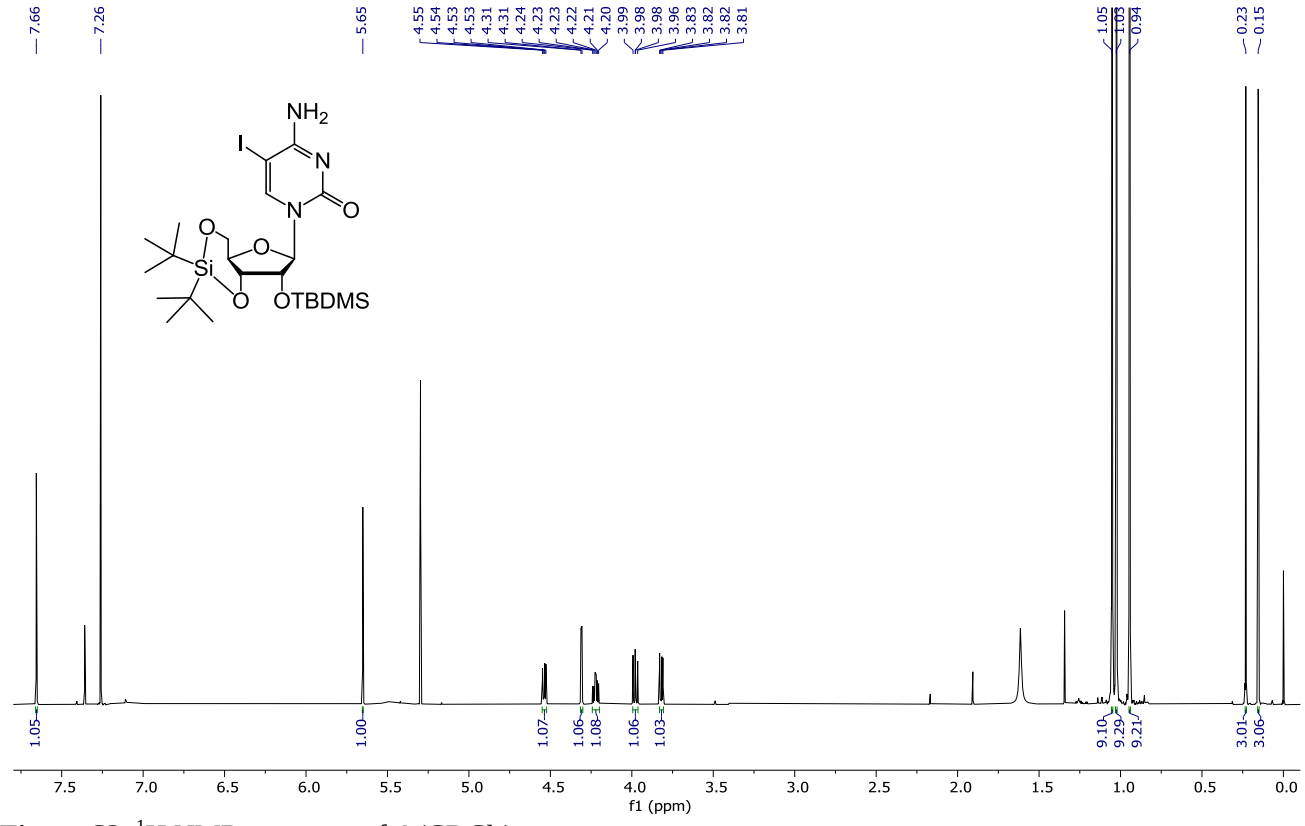


**Figure S2.** $^{1}$H NMR spectrum of **4** ($CDCl_3$).

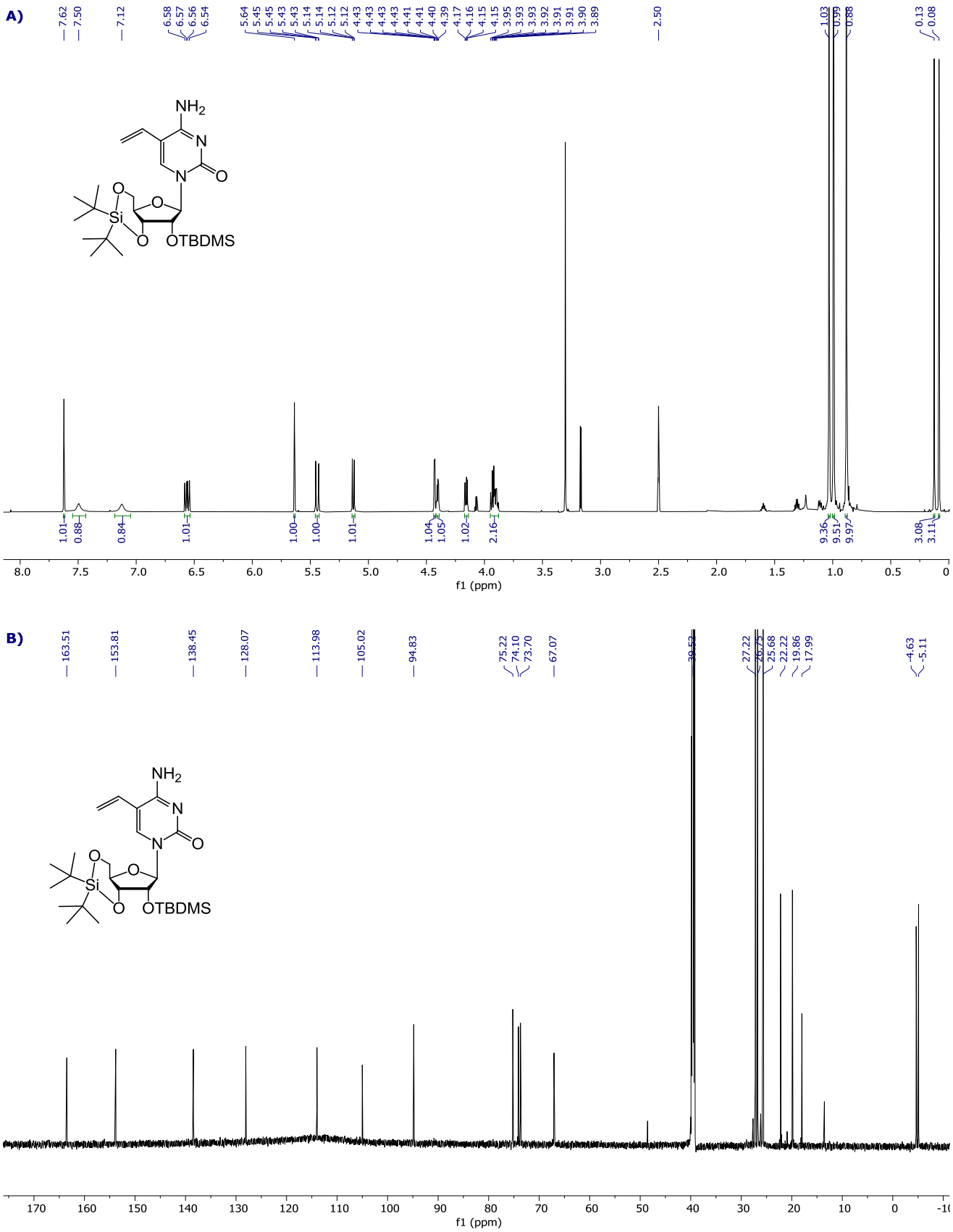


**Figure S3. A)** $^{1}$H NMR spectrum of **5** (DMSO-$d_6$); **B)** $^{13}$C NMR spectrum of **5** (DMSO-$d_6$).

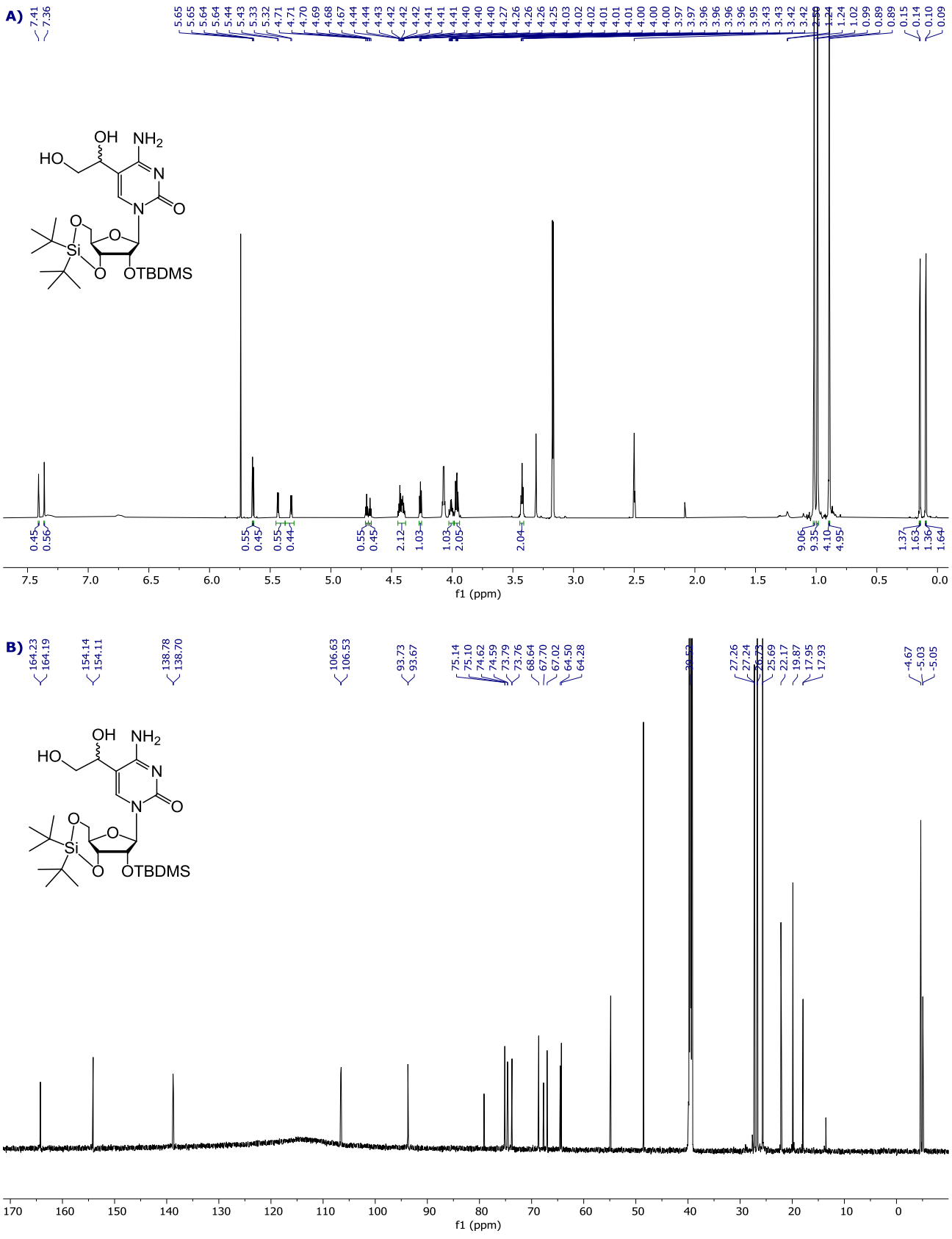


**Figure S4. A)** $^{1}$H NMR spectrum of **6** (DMSO-$d_6$); **B)** $^{13}$C NMR spectrum of **6** (DMSO-$d_6$).

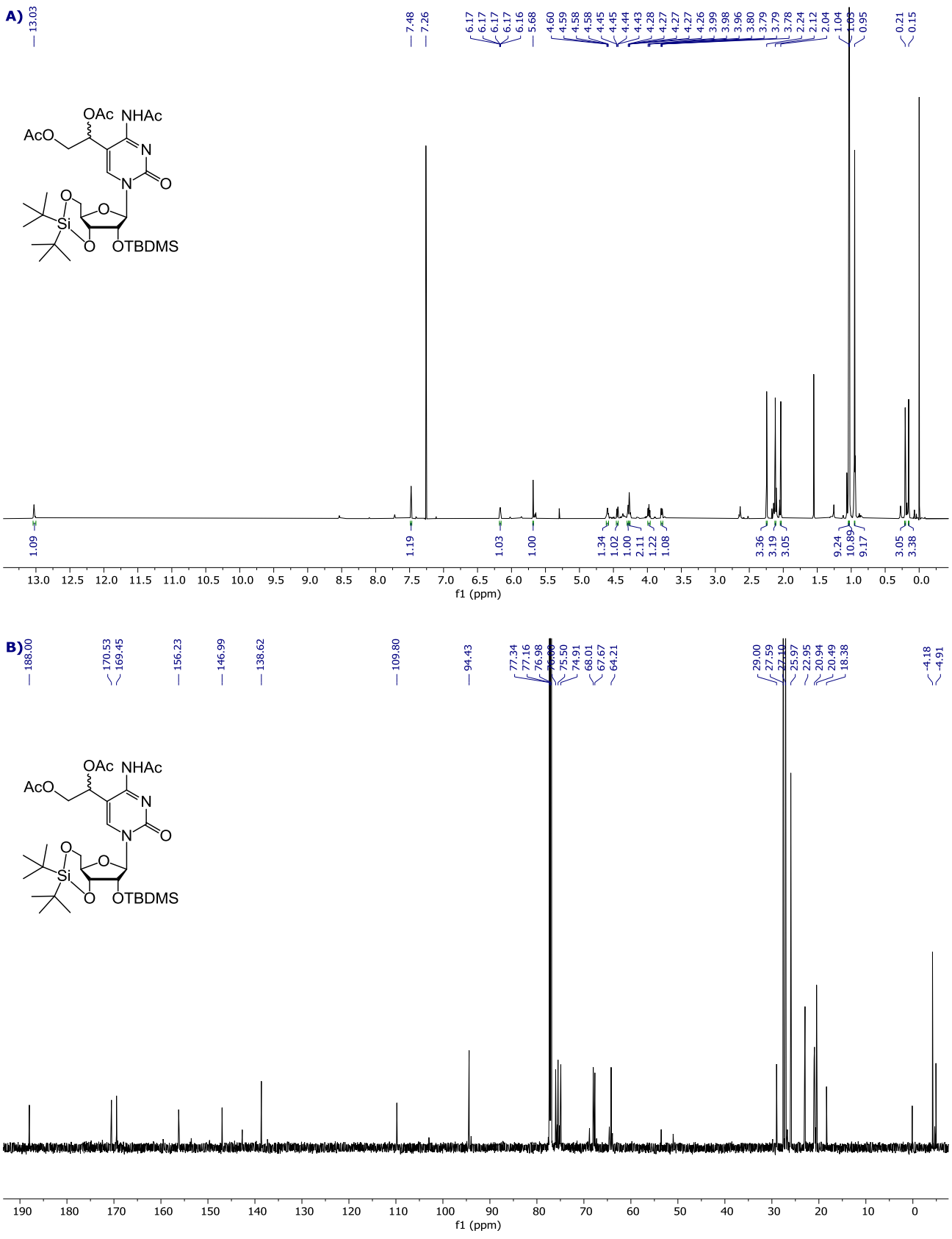


**Figure S5. A)** $^{1}$H NMR spectrum of **7** ($CDCl_3$); **B)** $^{13}$C NMR spectrum of **7** ($CDCl_3$).

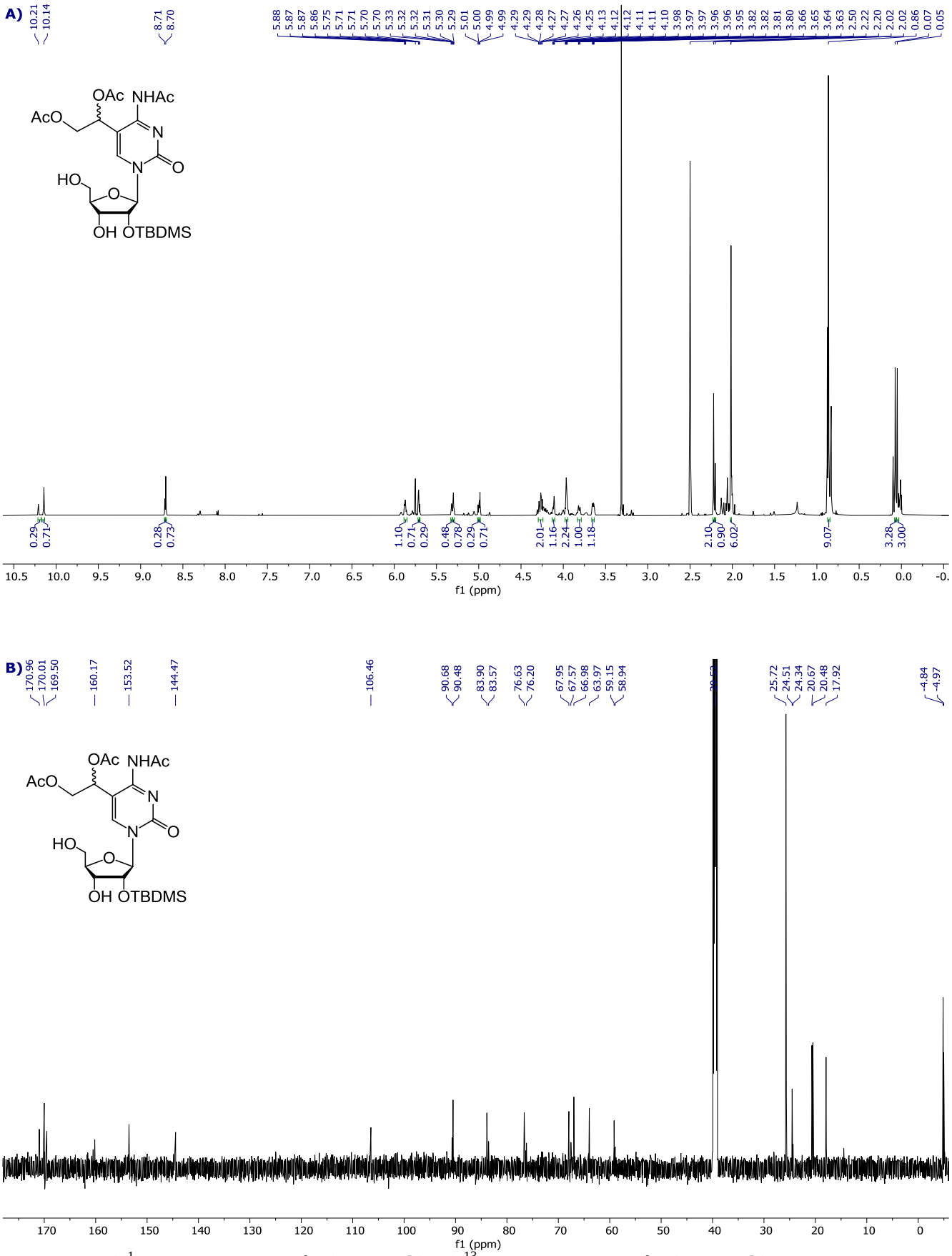


**Figure S6. A)** $^{1}$H NMR spectrum of **8** (DMSO-$d_6$); **B)** $^{13}$C NMR spectrum of **8** (DMSO-$d_6$).

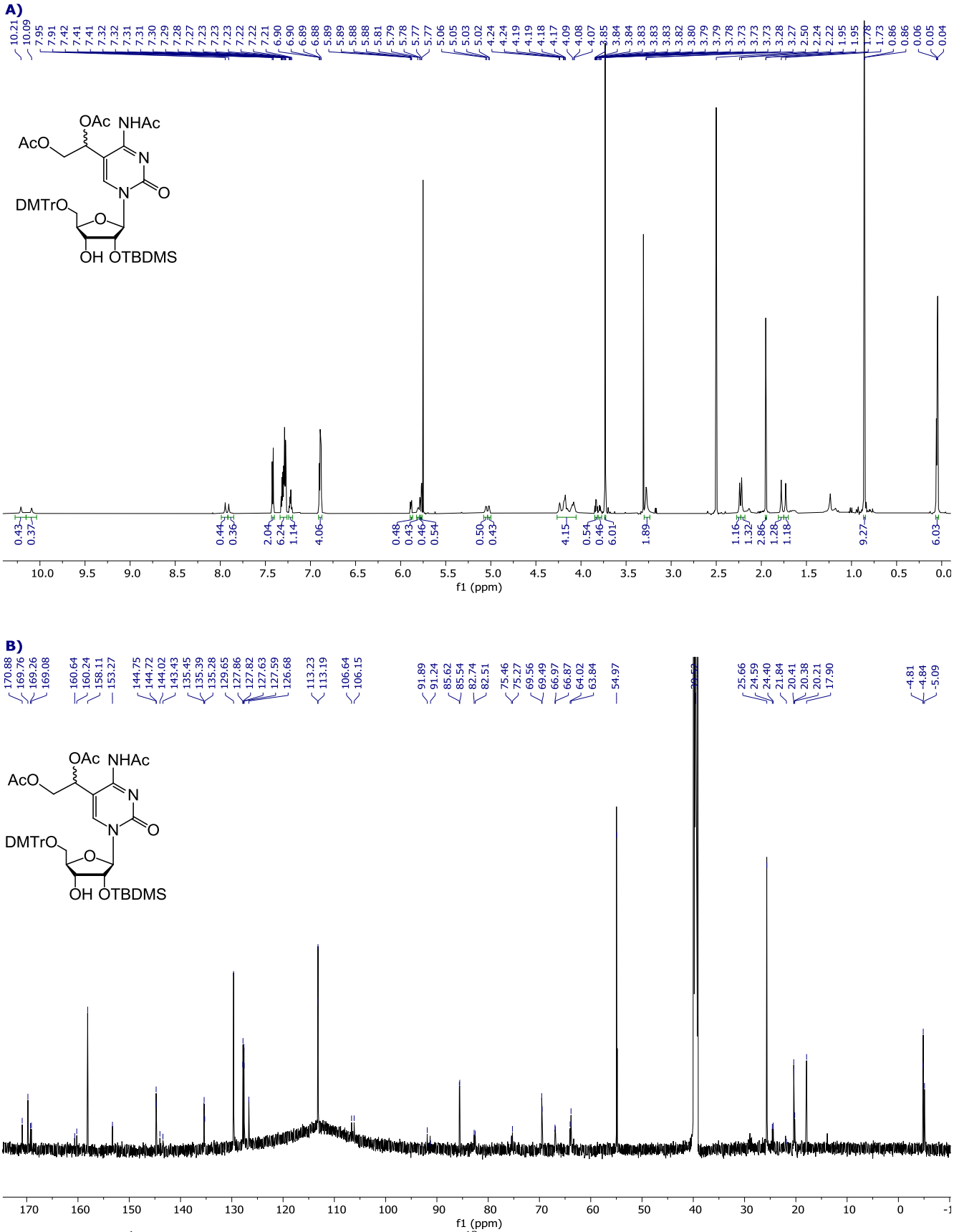


**Figure S7. A)** $^{1}$H NMR spectrum of **9** (DMSO-$d_6$); **B)** $^{13}$C NMR spectrum of **9** (DMSO-$d_6$).

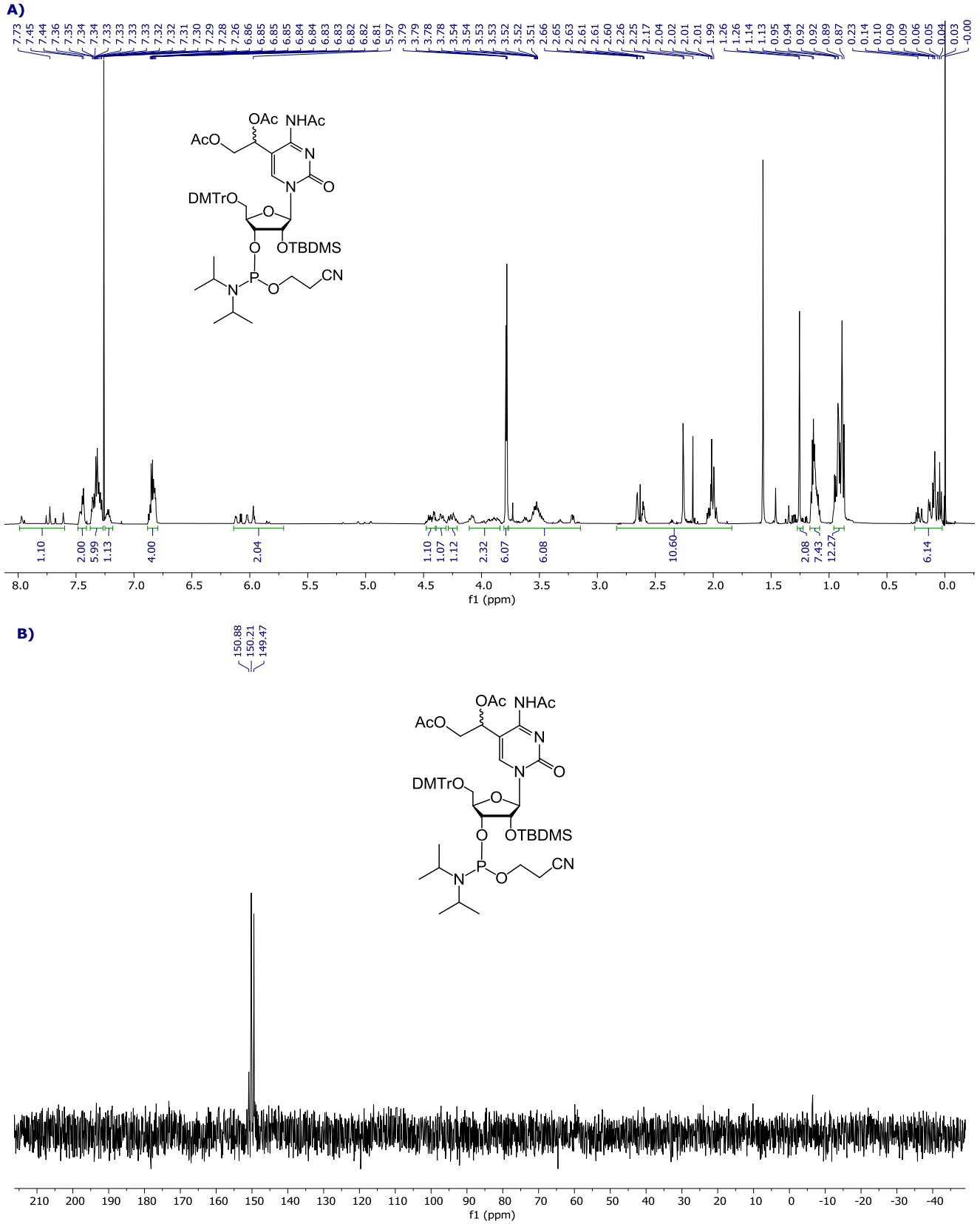


**Figure S8. A)** $^{1}$H NMR spectrum of **1** ($CDCl_3$); **B)** $^{31}$P NMR spectrum of **1** ($C_6H_6$).

## IV. **Chemical synthesis of precursor oligonucleotides**

**Table S1.** Sequences of synthesized oligonucleotides.

| Symbol | Sequence |
|---|---|
| pON1 | 5'-UCGGGCC**dhe$^5$C**AU**A**CCCCGA$_m$-3' |
| pON2 | 5'-UCGGGCC**dhe$^5$C**AU**G**CCCCGA$_m$-3' |
| pON3 | 5'-UCGGGCC**dhe$^5$C**AU**m$^1$G**CCCCGA$_m$-3' |
| ON1 | 5'-UCGGGCC**f$^5$C**AU**A**CCCCGA$_m$-3' |
| ON2 | 5'-UCGGGCC**f$^5$C**AU**G**CCCCGA$_m$-3' |
| ON3 | 5'-UCGGGCC**f$^5$C**AU**m$^1$G**CCCCGA$_m$-3' |

Precursor oligonucleotides **pON1-pON3** were synthesized manually on a 5 µmol scale using 2'-*O*-Me rA(pac)-succinyl-CPG (ChemGenes) support. The commercially available 5'-*O*-DMTr-2'-*O*-TBDMS phosphoramidites of A, C, U, G with additional protection of exocyclic amine functions with 4-*tert*-butylphenoxyacetyl (tac) (Proligo) were used as 0.1 M solutions in anhydrous ACN. Incorporation of A, U, C, and G amidites was performed in 8 molar excess with a coupling time of 8 min, while modified monomers were used in 14 molar excess and coupled twice, each time using 7 molar excess of amidite with a coupling time of 10 min. Condensation steps were carried out in the presence of 0.25 M solution of 5-(3,5-bis(trifluoromethyl)phenyl)-1H-tetrazole in ACN (Activator 42). The capping was performed with the mixture of Cap A (THF/tac$_2$O, 100:5 v/w), and Cap B (THF/N-methylimidazole, 84:16, v/v) for 2 min. A 0.02 M iodine solution in THF/$H_2O$/Py (90.54:9.05:0.41, v/v/v) was used as an oxidizing agent for 2 min. After the last coupling DMTr group was removed, and the solid support was washed with ACN, dried with argon, and transferred from the column to a screw cap glass vial.

## V. **Deprotection and purification of precursor oligonucleotides (pON1-pON3)**

Each of CPG-linked precursor oligonucleotides (0.2 µmol) was treated with TEA-ACN (264 µL, 1:1, v/v) for 20 min at room temperature. The TEA solution was removed and the resin was washed with anhydrous ACN (3 x 200 µL), and dried *in vacuo* for 30 min. Subsequently, the oligonucleotides were cleaved from the solid support with the simultaneous removal of the standard nucleobase protecting groups by treatment with 30% aq. $NH_3$-EtOH (300 µL, 3:1, v/v) for 16 h at 37 ºC. The supernatant was collected, and the solid support was washed with 50% aq. EtOH (3 x 200 µL). The combined fractions containing oligonucleotide were evaporated to dryness on a Speed-Vac. The resulting solid residue was co-evaporated with anhydrous EtOH and dried for 3 h on a Speed-Vac. The 2'-*tert*-butyldimethylsilyl protecting groups were removed by treatment of oligonucleotides with TEAx3HF-NMP (120 µL, 1:1, v/v) for 2.5 h at 65 ºC. Then, the reaction was precipitated using ethoxytrimethylsilane (240 µL), and *tert*-butyl methyl ether (600 µL). The resulting suspension was centrifuged for 2 min at 4 ºC, and the collected RNAs were washed with *tert*-butyl methyl ether (2 x 600 µL). Each of the fully deprotected oligonucleotides were purified by ion exchange HPLC (Source 15Q 4.6/100 PE

column) at a constant flow rate of 1 mL/min. The column was eluted with a linear gradient 50-650 mM NaBr in a 20 mM $Na_2HPO_4$-$NaH_2PO_4$ buffer solution (pH 7.5) containing 50 μM EDTA and 10% ACN. Fractions containing RNA were collected, concentrated on Speed-Vac, then loaded on a Sep-Pak C18 cartridge (Waters), washed with $H_2O$ (miliQ quality, 10 mL), and eluted with $H_2O$-ACN (1:1, v/v). The collected fractions were evaporated to dryness and lyophilized to give 7.6 $OD_{260}$, 6.5 $OD_{260}$, and 3.4 $OD_{260}$ units of **pON1**, **pON2** and **pON3**, respectively. The obtained products were analyzed by ESI mass spectrometry (**pON1**: *m/z* calcd 5431.7, found 5431.8; **pON2**: *m/z* calcd 5447.7, found 5447.8; **pON3**: *m/z* calcd 5461.7, found 5461.8).

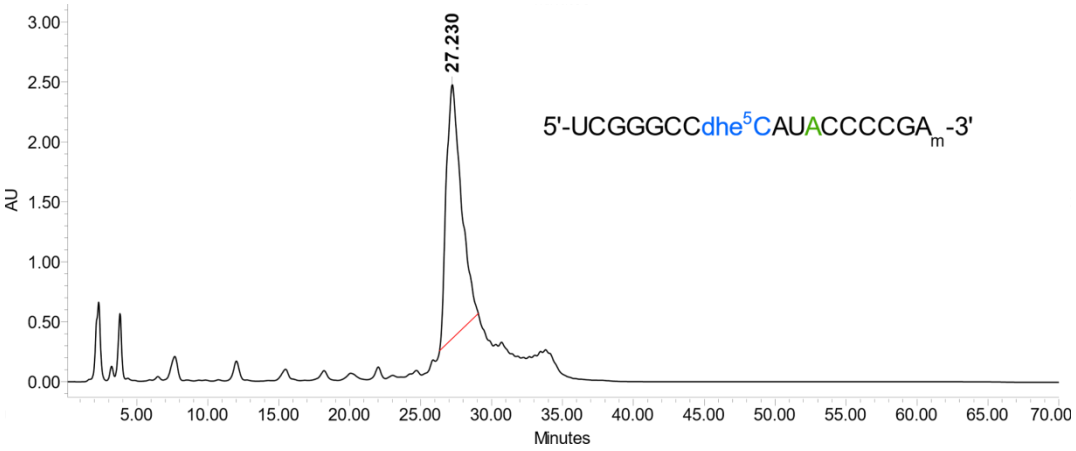


**Figure S9.** IE-HPLC of oligonucleotide **pON1**.

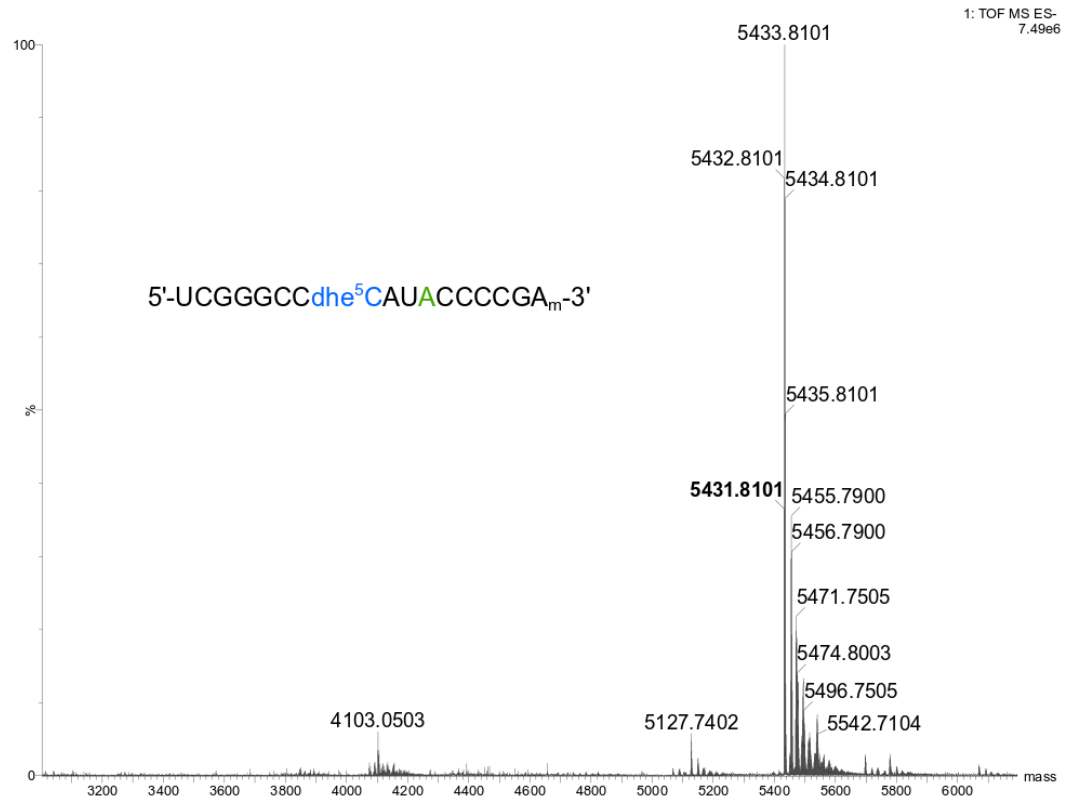


**Figure S10.** ESI MS spectrum of oligonucleotide **pON1**; calculated monoisotopic mass is 5431.76; measured *m/z* is 5431.81.

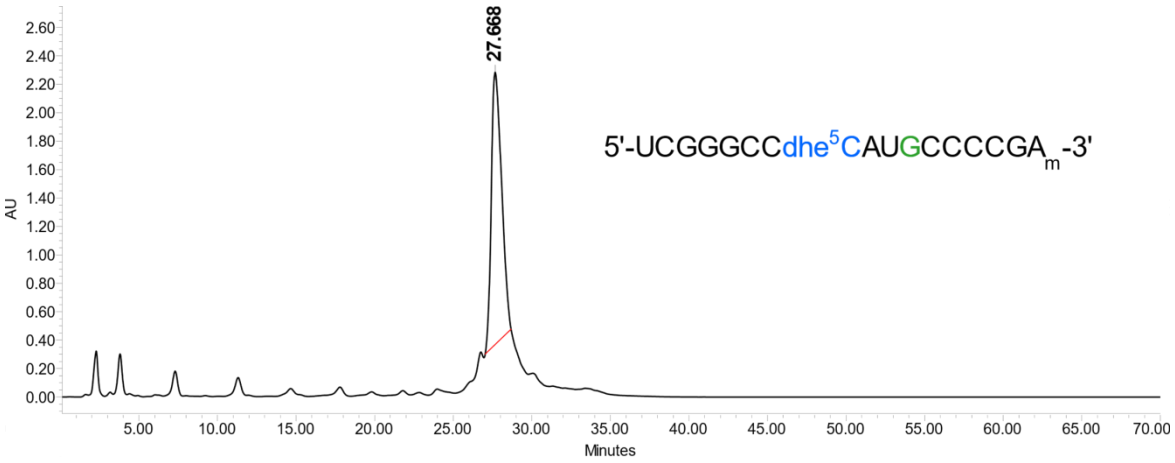


**Figure S11.** IE-HPLC of oligonucleotide **pON2**.

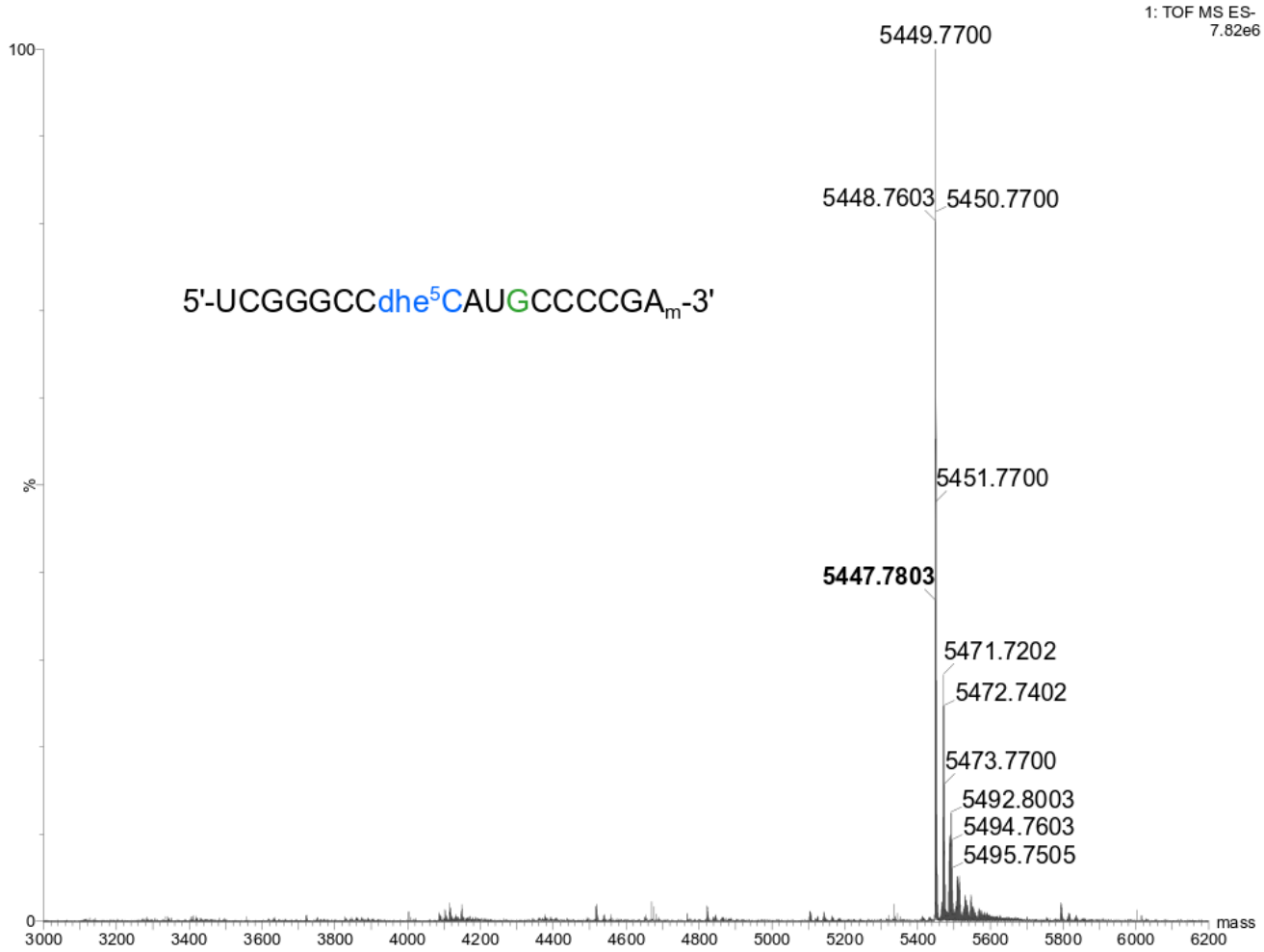


**Figure S12.** ESI MS spectrum of oligonucleotide **pON2**; calculated monoisotopic mass is 5447.76; measured *m/z* is 5447.78.

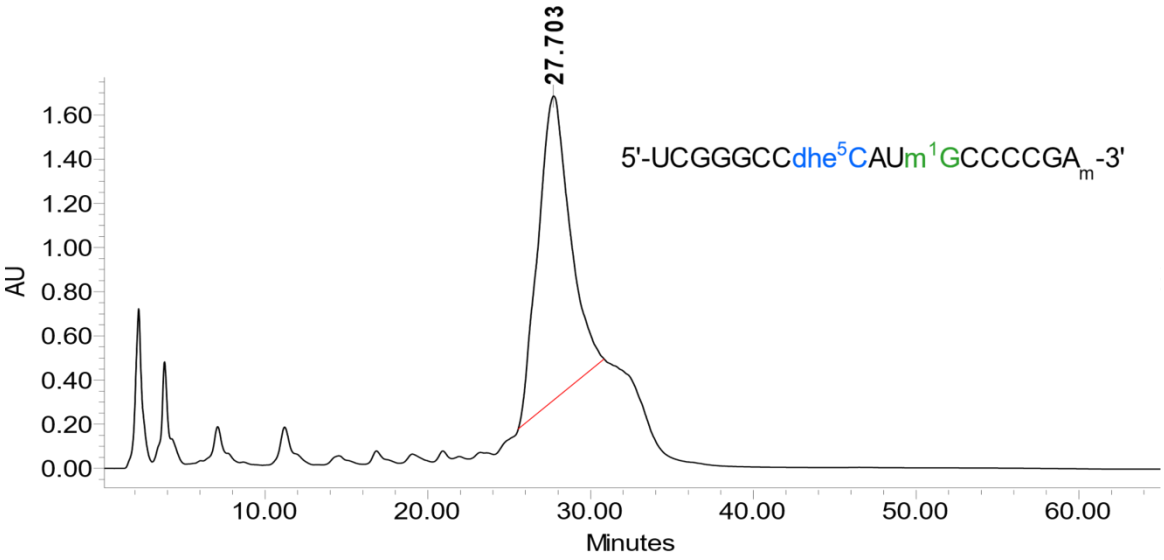


**Figure S13.** IE-HPLC of oligonucleotide **pON3**.

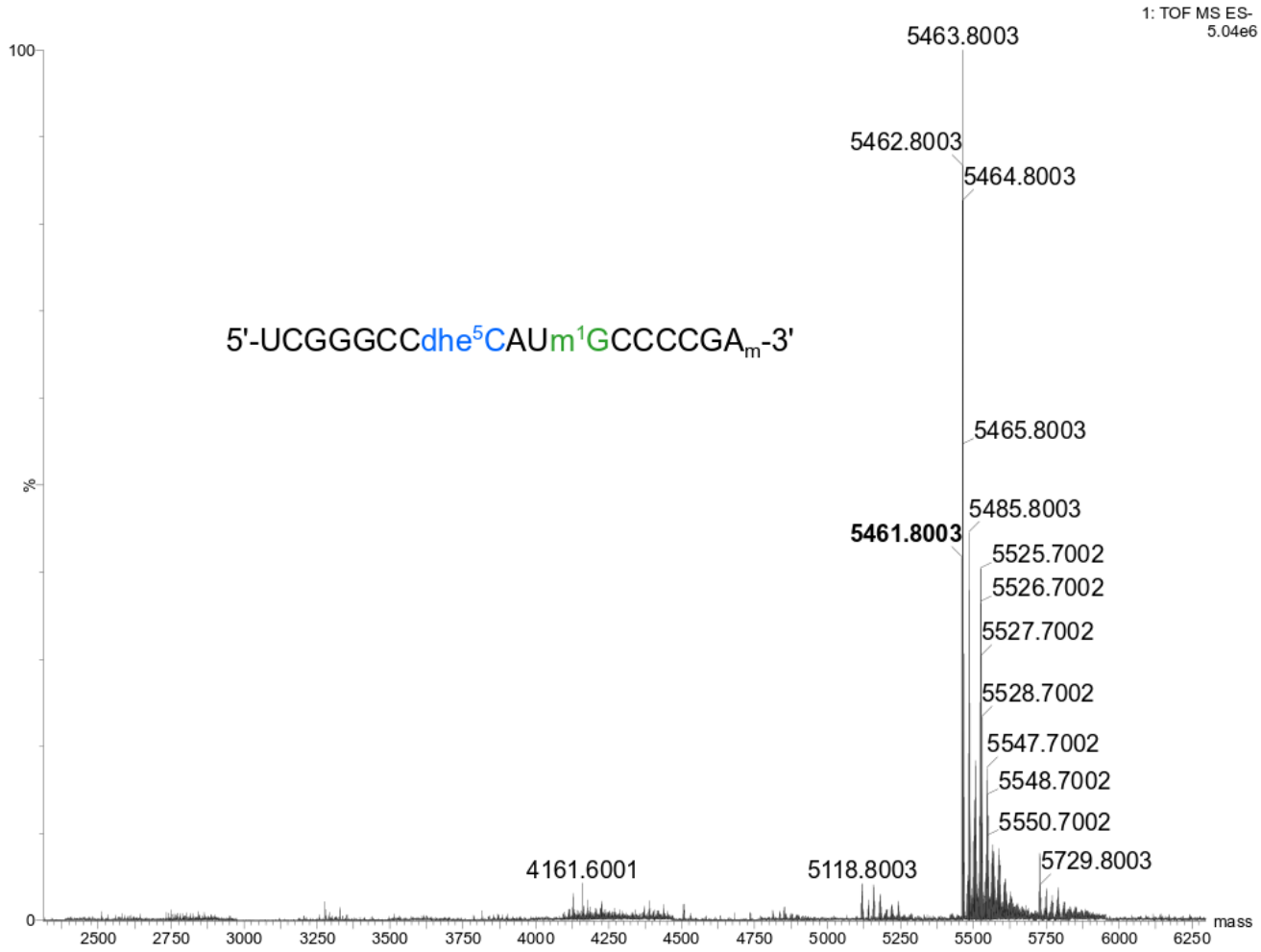


**Figure S14.** ESI MS spectrum of oligonucleotide **pON3**; calculated monoisotopic mass is 5461.7; measured *m/z* is 5431.8.

## VI. Post-synthetic transformation of pON1-pON3 and characteristic of ON1-ON3 using IE-HPLC, RP-HPLC and ESI MS

Each fully deprotected oligonucleotide **pON1-pON3** containing 5-(1,2-dihydroxyethyl)cytidine (3.0 $OD_{260}$) was treated with 0.05 M aq. $NaIO_4$ (225 µL) at 0 ºC for 30 min. The resulting solution was diluted with 0.1 M AcONa and desalted on the Sep-Pak C18 cartridge (Waters). The collected fraction was evaporated on Speed-Vac. The oligonucleotides were purified by ion exchange chromatography, using a Source 15Q 4.6/100 PE column with a linear gradient 50-650 mM NaBr in a 20 mM $Na_2HPO_4$-$NaH_2PO_4$ buffer solution (pH 7.5) containing 50 µM EDTA and 10% ACN. The ONs were desalted using the Sep-Pak C18 cartridge (Waters) to give 2.4 $OD_{260}$ (81% yield), 2.2 $OD_{260}$ (74% yield), and 2.6 $OD_{260}$ (88% yield) units of $f^5C$-containing **ON1**, **ON2** and **ON3** respectively. The obtained oligonucleotides were analyzed by ESI mass spectrometry (**ON1**: *m/z* calcd 5399.7, found 5399.8; **ON2**: *m/z* calcd 5415.7, found 5415.8; **ON3**: *m/z* calcd 5429.7, found 5429.8). In addition, **ON1-ON3** oligomers were analyzed by RP-HPLC, using a Reprospher RP 100 Å C18, 5 µm, 250 x 4.6 mm, column with a linear gradient of buffer A (0.1 M $CH_3COONH_4$, pH 6.0) and buffer B (ACN) at a constant flow rate 1 mL/min.

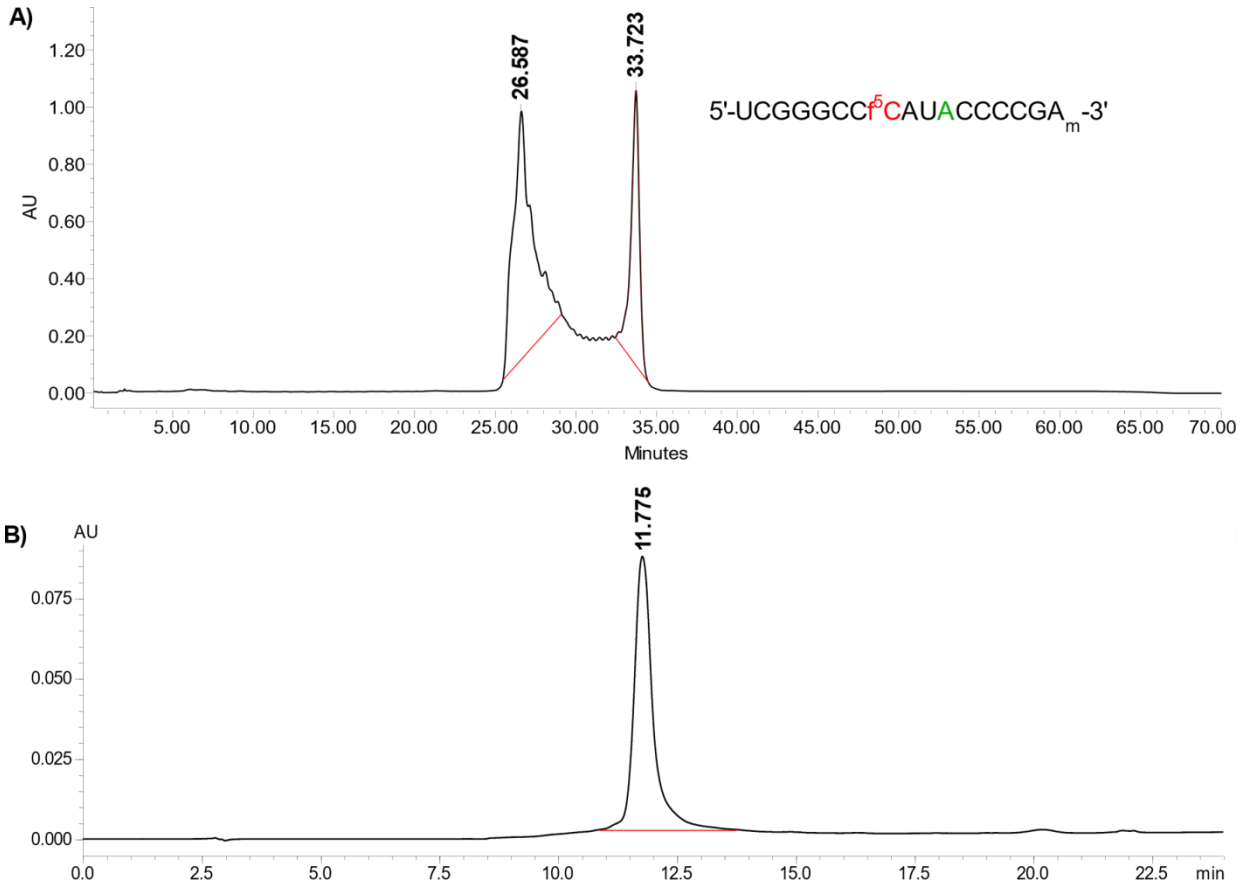


**Figure S15. A)** IE-HPLC analysis of oligonucleotide **ON1**; **B)** RP-HPLC analysis of **ON1**.

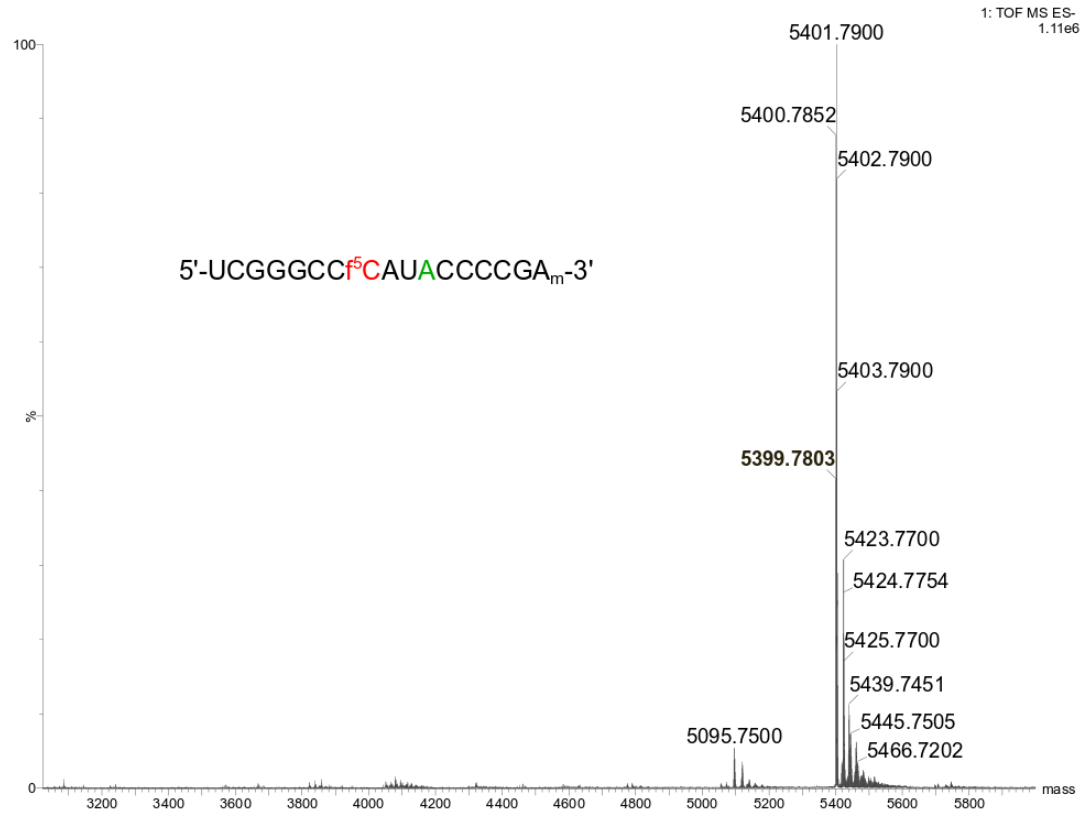


**Figure S16.** ESI MS spectrum of oligonucleotide **ON1**; calculated monoisotopic mass is 5399.76; measured *m/z* is 5399.78.

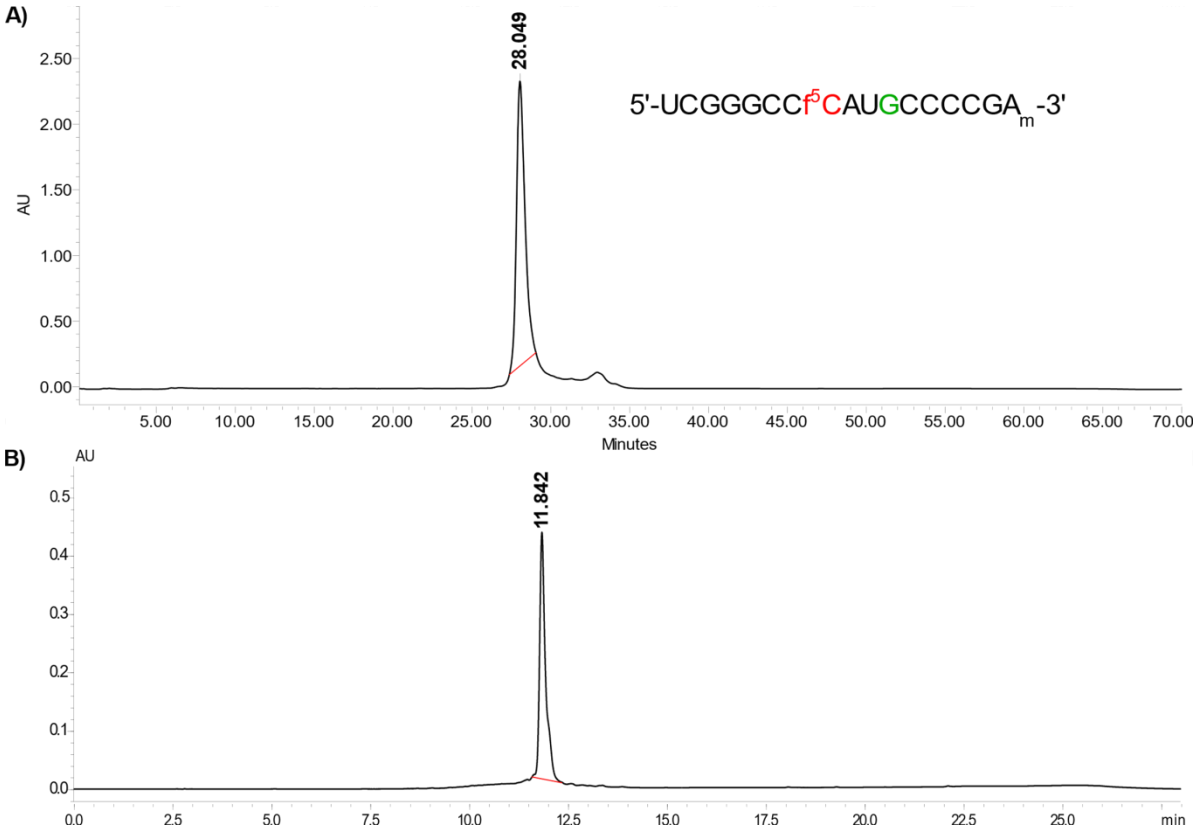


**Figure S17. A)** IE-HPLC analysis of oligonucleotide **ON2**; **B)** RP-HPLC analysis of oligonucleotide **ON2**.

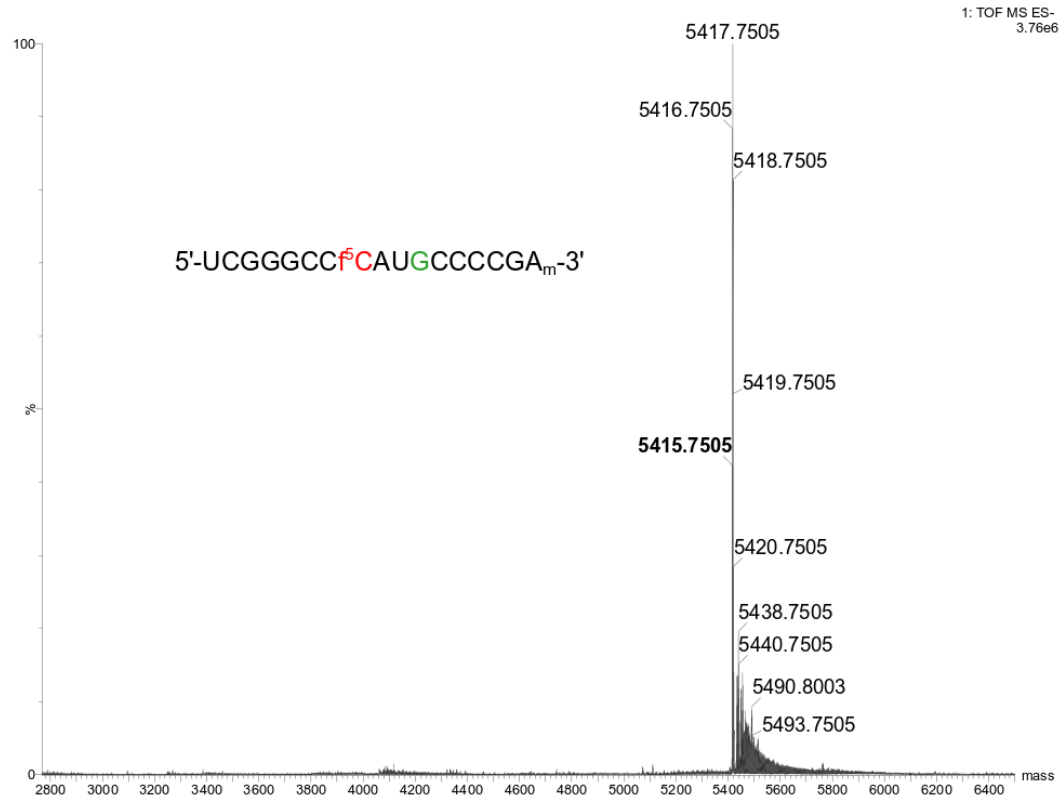


**Figure S18.** ESI MS spectrum of oligonucleotide **ON2**; calculated monoisotopic mass is 5415.76; measured *m/z* is 5415.75.

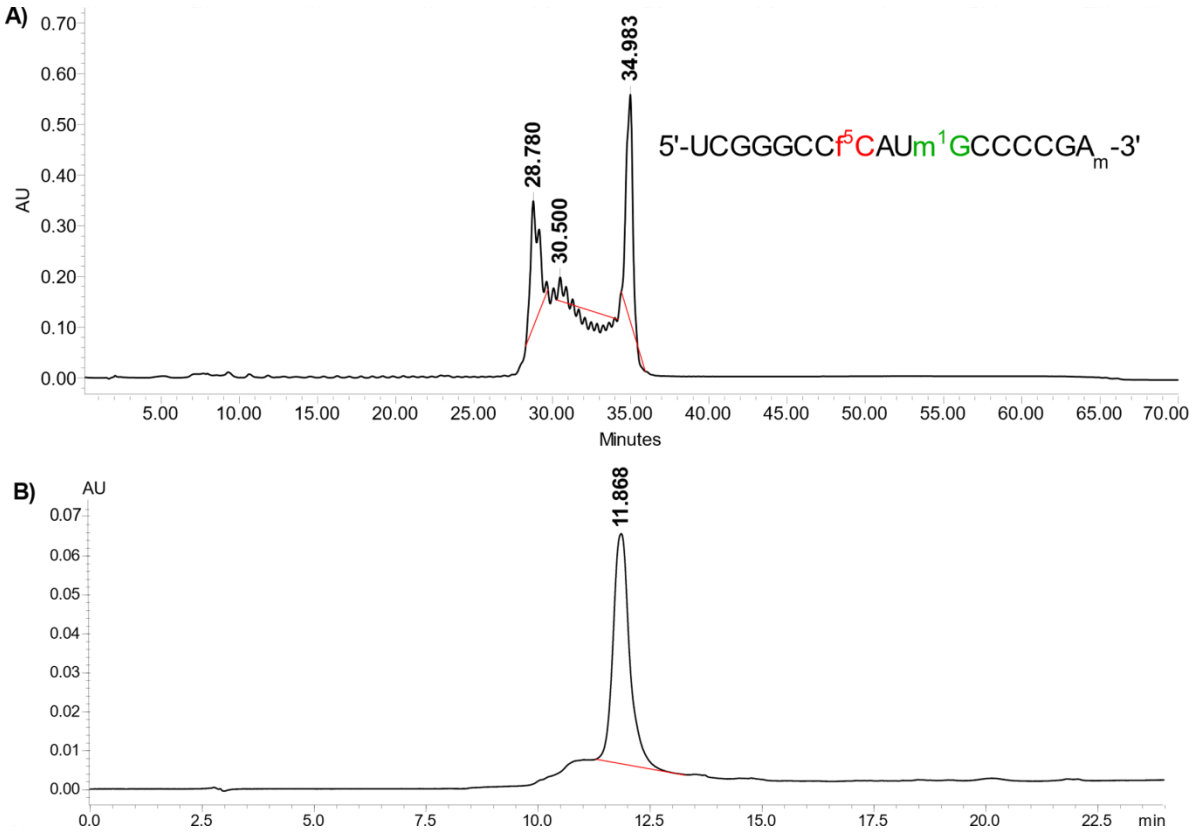


**Figure S19. A)** IE-HPLC analysis of oligonucleotide **ON3**; **B)** RP-HPLC analysis of oligonucleotide **ON3**.

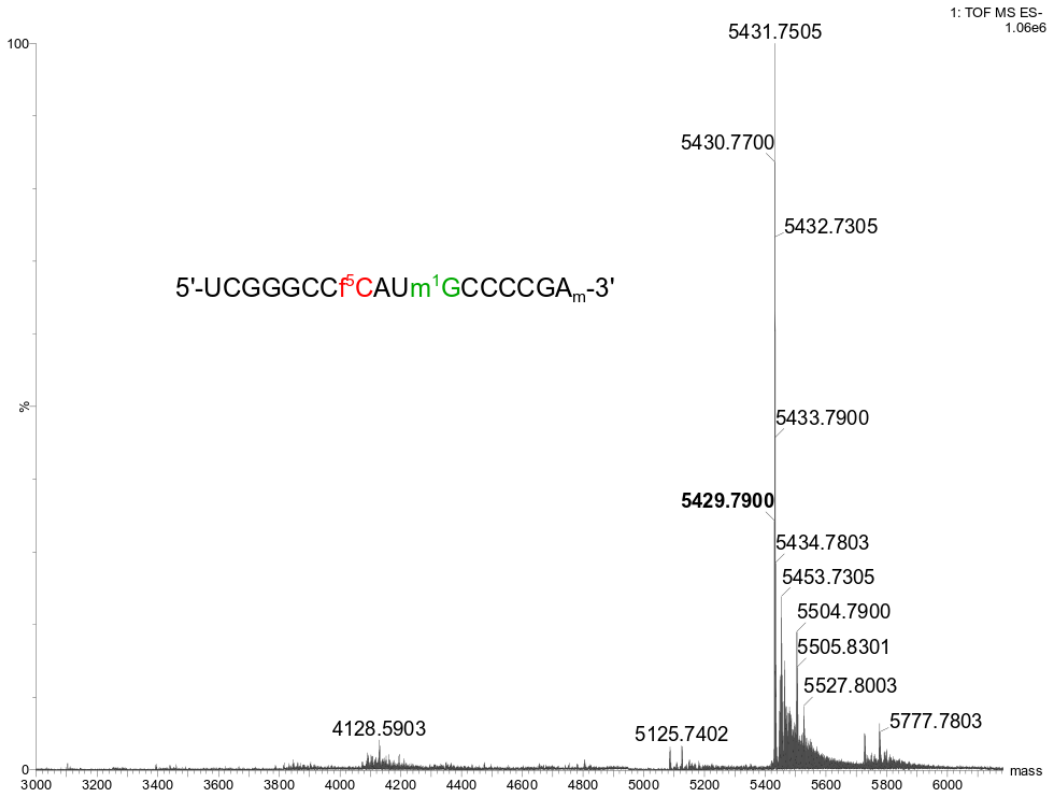


**Figure S20.** ESI MS spectrum of oligonucleotide **ON3**; calculated monoisotopic mass is 5429.76; measured *m/z* is 5429.79.

## VII. **Native polyacrylamide gel electrophoresis**

A 24% polyacrylamide gel containing no urea was pre-incubated in 89 mM Tris-borate and 2 mM EDTA (pH 8.0) running buffer for 1 hour. Each oligonucleotide **ON1-ON3** (0.3 $OD_{260}$) was dissolved in 6 µL $ddH_2O$ and then 4 µL of gel loading solution was added. Each sample was applied to a separate lane of the gel. Electrophoresis was conducted for 8 h. Following electrophoresis, the gel was stained for 20 to 30 minutes in 0.5 µg/L aqueous solution of ethidium bromide, and then the gel was photographed.

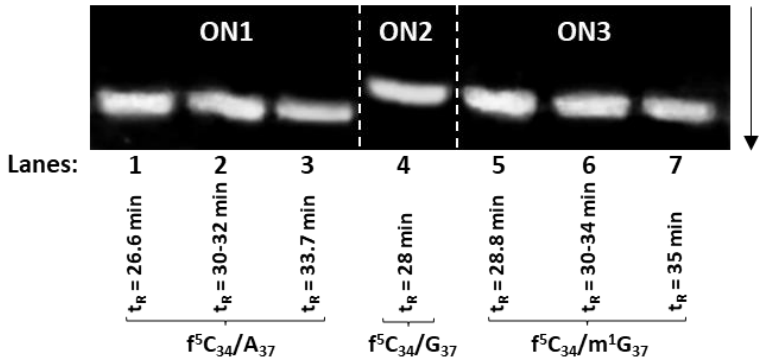


**Figure S21.** Electrophoresis on a 24% polyacrylamide gel. Samples 1-3 and 5-7 refer to the fractions separated during IE HPLC purification of **ON1** and **ON3**, respectively. Lanes: 1) **ON1** collected at $t_R$ = 26.58 min, 2) **ON1** collected at $t_R$ = 30-32 min, 3) **ON1** collected at 33.73 min, 4) **ON2**, 5), **ON3** collected at $t_R$ = 28.78 min, 6) **ON3** collected at $t_R$ = 30-34 min, 7) **ON3** collected at $t_R$ = 34.98 min.

## VIII. **Circular dichroism**

CD spectra were recorded on a Jasco J-1500 spectrophotometer using quartz cell with a 0.1 cm path length. Solutions of **ON1-ON3** oligomers were prepared in 20 mM Na-K phosphate buffer (pH 6.8) at concentrations of 4 μM. Each sample was heated to 85 ºC for 10 minutes and then slowly cooled to room temperature with a temperature gradient of 0.5 ºC/min before data collection. The measurements were recorded at 21 ºC in the 200-300 nm wavelength range with a 5 nm data point interval, 1 nm bandwidth and scan speed 100 nm/min. The buffer spectrum was subtracted from the sample spectra, and the resultant CD spectra were smoothed with a Savitzky-Golay algorithm (5 convolution coefficient).

## IX. **Thermal denaturation experiments**

Solutions of **ON1-ON3** oligomers were prepared in 20 mM Na-K phosphate buffer (pH 6.8) at concentrations of 4 μM. Thermal denaturation was performed on a Jasco V-770 UV−VIS/NIR spectrophotometer equipped with a thermal controller a Peltier Thermocell. The samples were then heated to 90 ºC, and cooled to 15 ºC with a temperature gradient of 1.5 ºC/min. The melting profiles were recorded from 15 to 90 ºC, with the temperature gradient of 0.5 ºC/min. The calculation of thermodynamic parameters (Tm, $\Delta G^o$, $\Delta H^o$ and $\Delta S^o$) was done by numerical fitting of a given melting curve using a two-state model algorithm provided by a MeltWin v.3.5 software. Each result was taken as an averaged one from three independent experiments.

## X. **Enzymatic digestion**

Each ON (0.25 $OD_{260}$) were hydrolyzed with nuclease P1 and alkaline phosphatase in a 20 mM TEAxHCl (pH 7).[3] The resulting nucleoside mixture was analyzed by RP-HPLC (Reprospher 100 C18, 5 μm, 250 x 4.6 mm) at a constant flow rate of 0.75 mL/min. The column was eluted with a linear gradient of buffer A (10 mM $KH_2PO_4$, pH 5.3) and buffer B (20% MeOH in 10 mM $KH_2PO_4$, pH 5.1). The peaks were compared with reference samples of modified units in separate control experiments indicating the presence of modified units ($f^5C$, $m^1G$, $A_m$).

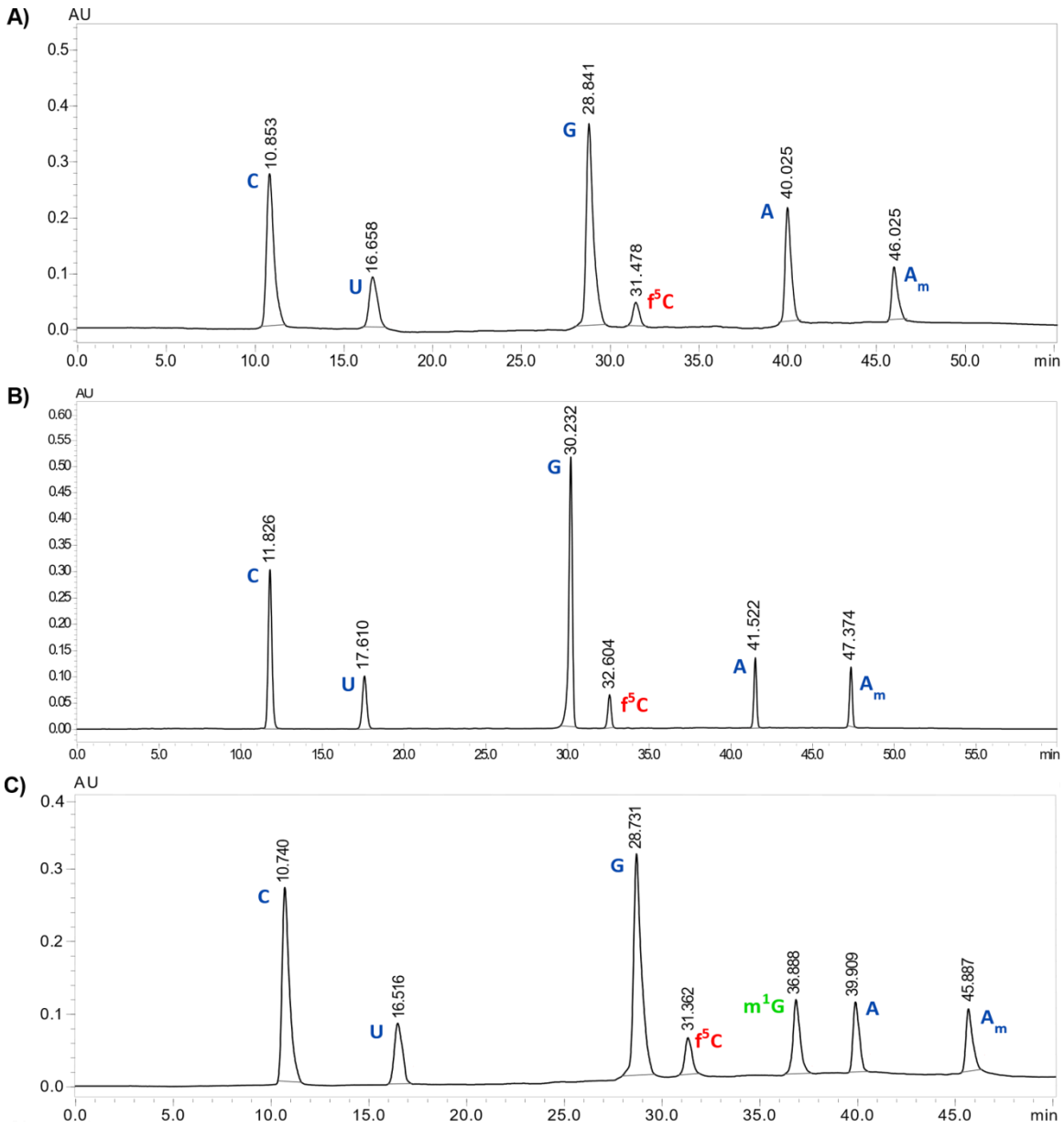


**Figure S22.** RNA enzymatic digestion of $f^5C$-containing oligonucleotides A) **ON1**, B) **ON2**, C) **ON3**.

## XI. **Synthesis and melting profiles of ON6 and ON7**

Both **ON6** and **ON7** oligomers were synthesized using standard protocol of solid-phase phosporamidite chemistry [4] and purified by IE-HPLC. The correct mass of oligomers were confirmed by ESI-MS analysis (Figures S23 and S24).

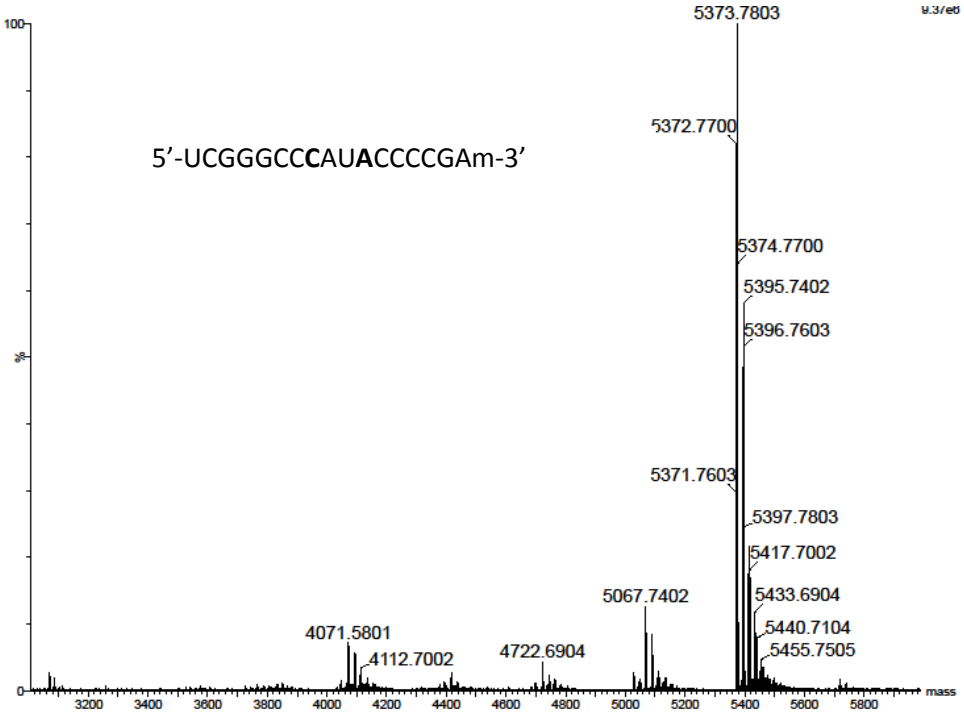


**Figure S23.** ESI MS spectrum of oligonucleotide **ON6**; calculated monoisotopic mass is 5371.7; measured *m/z* is 5371.7.

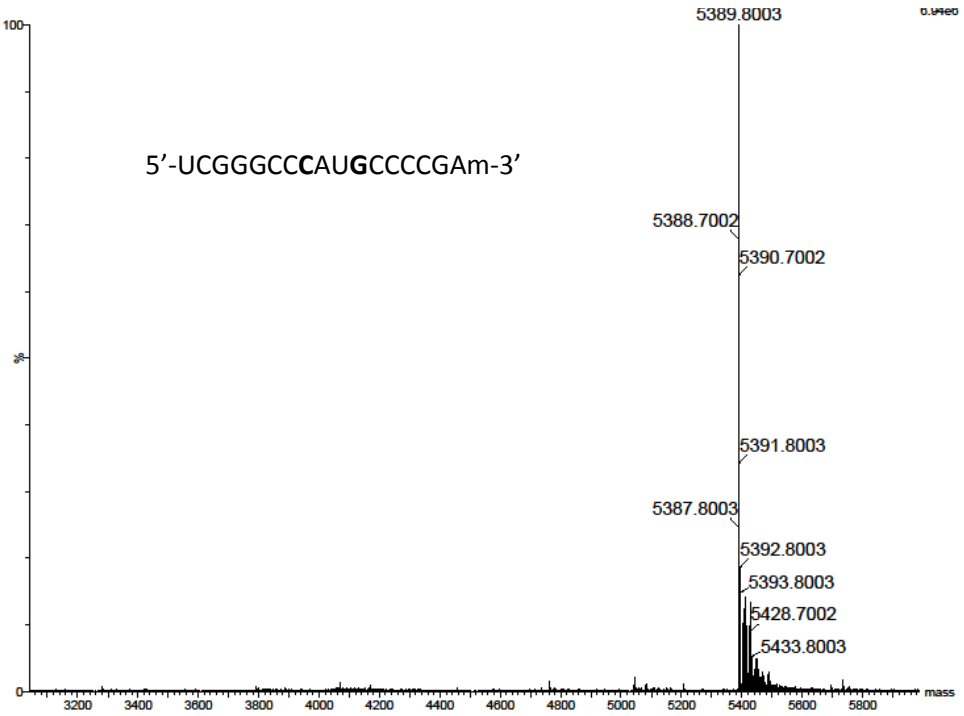


**Figure S24.** ESI MS spectrum of oligonucleotide **ON7**; calculated monoisotopic mass is 5387.7; measured *m/z* is 5387.8.

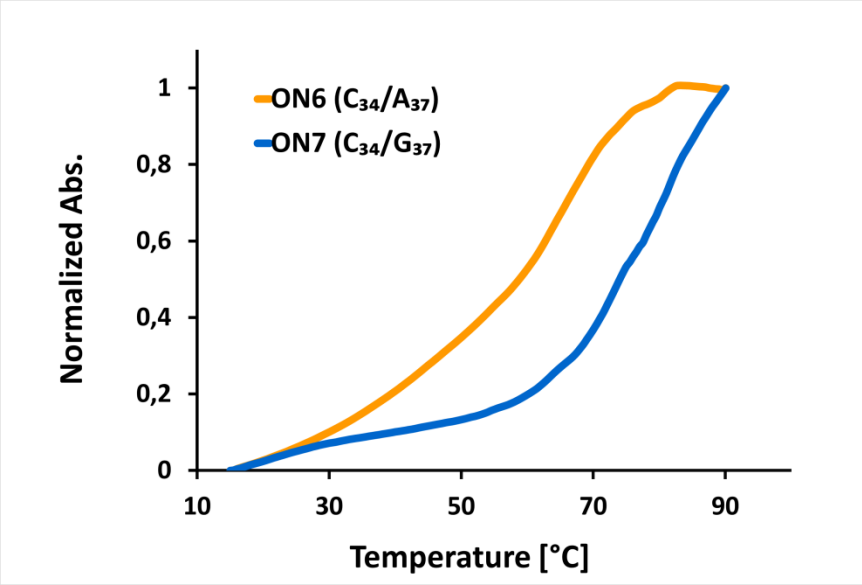


**Figure S25.** Melting profiles of **ON6** and **ON7**. *Conditions: $c_{RNA}$= 4 μM; 20 mM Na-K phosphate buffer, pH 6.8.*

## XII. **References**

[1] V. Serebryany, L. Beigelman, *Tetrahedron Lett.*, 2002, **43**, 1983.

[2] J. Asakura, M. J. Robins, *J. Org. Chem.*, 1990, **55**, 4928.

[3] K. Miyauchi, S. Kimura, T. Suzuki, *Nat. Chem. Biol.*, 2013, **9**, 105.

[4] B. S. Sproat, *Methods Mol. Biol.*, 2005, **288**, 17.