Electronic Supplementary Material (ESI) for RSC Chemical Biology.

Supplementary Information

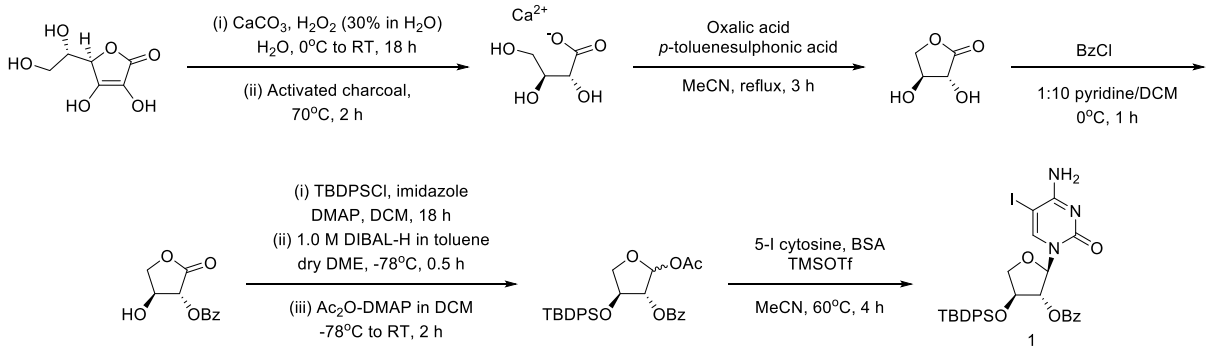
Increasing the Functional Density of Threose Nucleic Acid

Biju Majumdar[1□], Daisy Sarma[1⊥], Yutong Yu[1], Adriana Lozoya Colinas[1], and John C. Chaput[1-4*]

[1]Department of Pharmaceutical Sciences, University of California, Irvine, CA 92697-3958, United States
[2]Department of Chemistry, University of California, Irvine, CA 92697-3958, United States
[3]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697-3958, United States
[4]Department of Chemical and Biomolecular Engineering, University of California, Irvine, CA 92697-3958, United States

Contents

This file contains Supplementary Scheme 1, Supplementary Figures 1-3, experimental methods, and compound characterization.



Scheme S1. Chemical synthesis of 5-iodo-1-(2'-*O*-benzoyl-α-L-threofuranosyl)-cytidine.

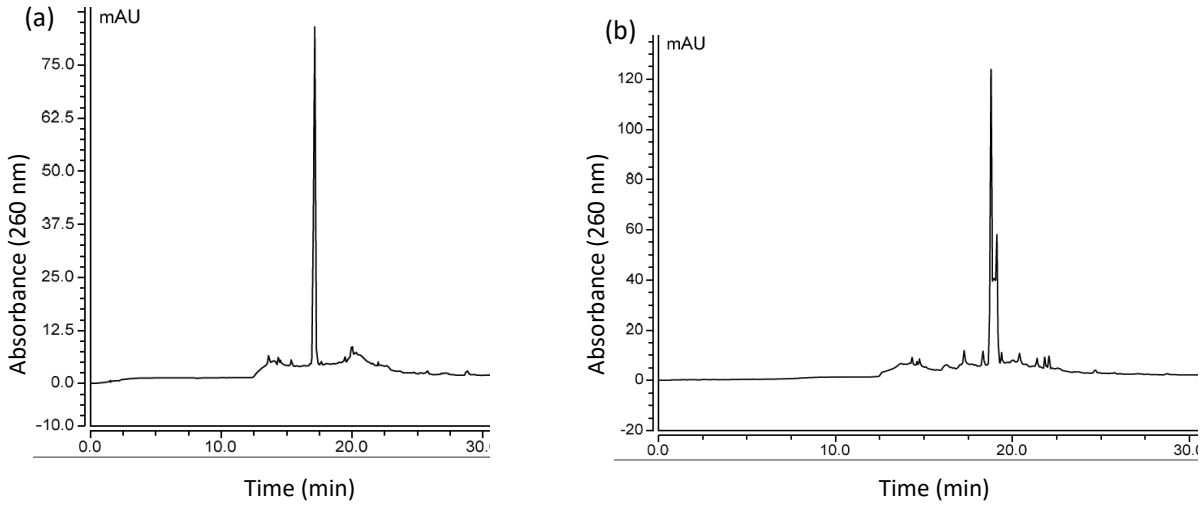
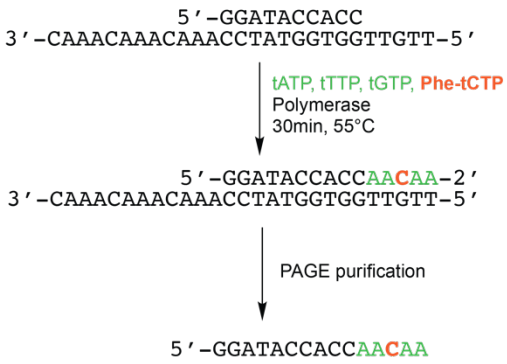


Figure S1. HPLC analysis of C-5 modified (a) benzylamine and (b) propylphenylamine α-L-threofuranosyl cytidine triphosphates.

A.



B.

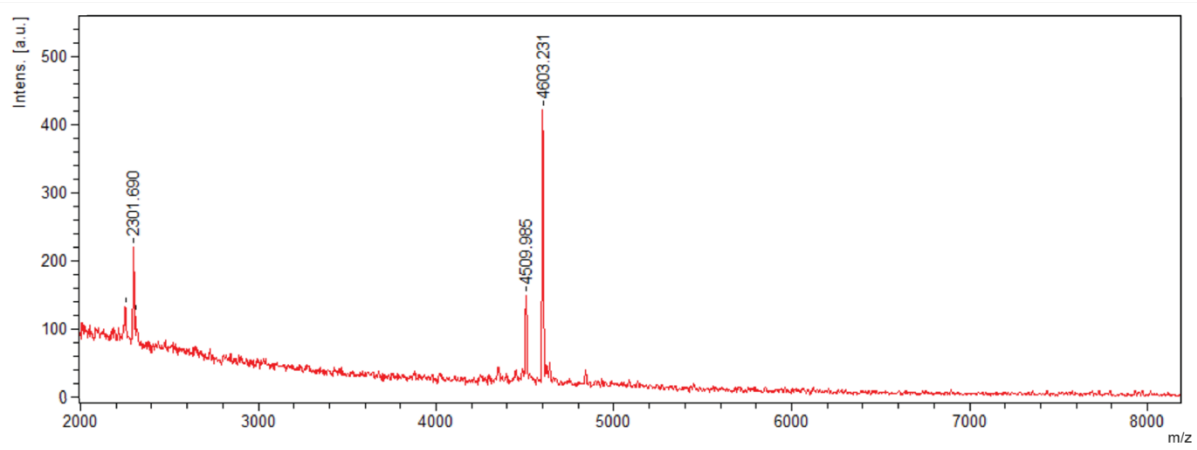


Figure S2. MALDI-ToF mass spectrometry analysis confirming the enzymatic incorporation of a base-modified residue into a TNA oligonucleotide. Product mass calculated 4602 and observed 4603.231.

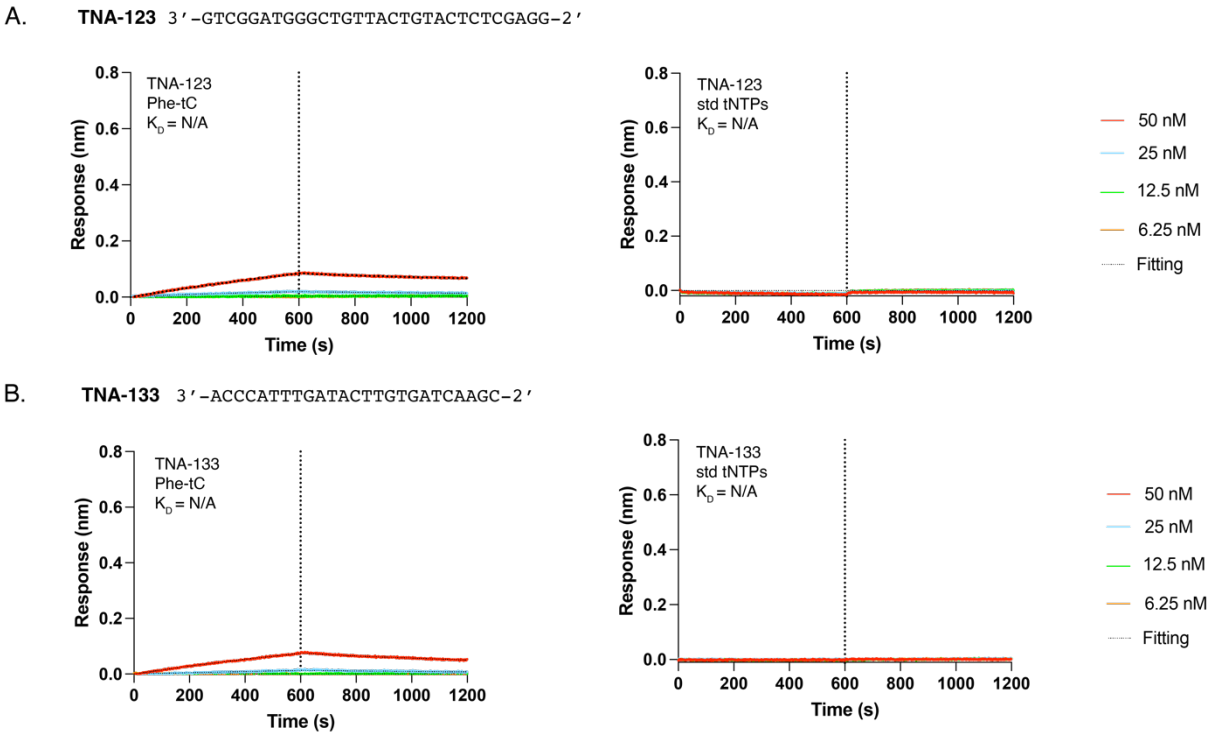


Figure S3. BLI control experiments showing that aptamer protein binding is abrogated when the S1 aptamer sequences are prepared with standard bases only or as singly modified aptamers with tC^{Phe}.

Methods

General Information

All moisture sensitive reactions were performed in anhydrous solvents under an argon or nitrogen atmosphere. All commercial reagents, solvents and anhydrous solvents were used without further purification. Reaction progresses were monitored by thin layer chromatography using glass-backed analytical Silica Plate with UV active F254 indicator. Flash column chromatography was performed with Silica Flash® P60 silica gel (40-63 μm particle size) for most of the crude reaction mixture. Yields are reported as isolated yields of pure compounds. ^{1}H, ^{13}C, and ^{31}P NMR spectra were analysed on 400 MHz NMR spectrometer (Bruker DRX) at the University of California, Irvine NMR Facility. ^{1}H values are reported in parts per million relative to Me_4Si or corresponding deuterium solvents used as an internal standard. ^{13}C values are reported in parts per million relative to corresponding deuterium solvents used as an internal standards. ^{31}P NMR values are reported in parts per million relative to an external standard of 85% H_3PO_4. Splitting patterns are designated as follows: s, singlet; d, doublet; dd, doublet of doublets; t, triplet; m, multiplet. High resolution mass spectrometry (HRMS) data were acquired using the electrospray ionization time-of-flight (ESI-TOF) method at the University of California, Irvine Mass Spectrometry Core Facility.

2-(4-amino-5-iodo-2-oxopyrimidin-1(2H)-yl)-4-((tert-butyldiphenylsilyl)oxy)tetrahydro furan-3-yl benzoate (1): To a stirred suspension of 5-iodo-cytosine (5.9 g, 25.0 mmol, 1.2 equiv) in anhydrous MeCN (70 mL) was added *N*,*O*-bis(trimethylsilyl)acetamide (12.5 mL, 50.0 mmol, 2.4 equiv) and the mixture heated at 60°C for 1.5 h. After cooling to rt, a solution of 1-*O*-acetyl-2-*O*-benzoyl-3-*O*-tertbutyldiphenylsilyl- L-threofuranose (10.5 g, 20.81 mmol, 1.0 equiv) in anhydrous MeCN (40 mL) was added dropwise to the reaction mixture, followed by dropwise addition of TMSOTf (12.0 mL, 66.14 mmol, 3.18 equiv). The resulting mixture was heated at 60°C for 4 h, cooled to rt and concentrated under reduced pressure. The resulting syrup was dissolved in EtOAc (200 mL), washed with saturated $NaHCO_3$ and brine, dried over Na_2SO_4 and concentrated. The crude mixture was purified by column chromatography using deactivated (5% Et_3N/CH_2Cl_2) silica gel eluting with 0-2% $MeOH/CH_2Cl_2$ to afford **1** (10.8 g, 15.85 mmol, 76.6%) as a white solid; ^{1}H NMR (400 MHz, CD_3OD) δ 8.17 (s, 1H), 7.91 (d, J = 7.9 Hz, 1H), 7.64 (m, 2H), 7.57 (m, 3H), 7.44-7.38 (m, 8H), 5.90 (s, 1H), 5.59 (s, 1H), 4.34 – 4.25 (m, 2H), 4.12 (dd, J = 10.1, 3.5 Hz, 1H), 1.03 (s, 9H); ^{13}C NMR (101 MHz, CD_3OD) δ 164.85, 164.53, 155.83, 147.10, 135.58, 135.53, 134.58, 133.32, 132.24, 131.98, 129.98, 129.91, 129.34, 128.98, 128.26, 127.75, 127.64, 127.19, 91.36, 81.71, 76.65, 75.61, 56.04, 26.23, 18.58; HRMS (ESI-TOF) calcd for $C_{31}H_{32}IN_3O_5SiNa$ $[M+ Na]^+$ 704.1053; found 704.1047.

1-(2′-*O*-Benzoyl-3′-*O*-tertbutyldiphenylsilyl α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine (2a): To a 100 mL round-bottomed flask equipped with a stir-bar was added **1** (1.4 g, 2.05 mmol, 1.0 equiv) and $Pd(PPh_3)_4$ (56 mg, 0.0484 mmol, 0.02 equiv). After the vessel was sealed and evacuated for 10-15 min, DMF (20.0 mL) was added, followed by Et_3N (1.4 mL, 10.04 mmol, 5.0 equiv) and benzylamine (2.69 mL, 24.63 mmol, 12 equiv). The vessel was then equipped with 2 balloons filled with CO and the reaction stirred at rt for 48 h. The crude reaction mixture was diluted with CH_2Cl_2 (300 mL), washed with water (3 x 200 mL) and then brine (3 x 200 mL). The combined organics were dried over $MgSO_4$ and concentrated under diminished pressure. The crude residue was dry loaded with neutral alumina and purified by silica gel chromatography eluting with (0-1%) $MeOH/CH_2Cl_2$ containing 1% Et_3N to afford product **2a (**0.950 g, 1.38 mmol, 68%) as light brown solid; ^{1}H NMR (400 MHz, CD_3OD) δ 8.54 (s, 1H), 7.89 (d, J = 8.2 Hz, 2H), 7.58 – 7.20 (m, 18H), 5.93 (s, 1H), 5.69 (s, 1H), 4.43 (s,

2H), 4.34 (d, J = 16.4 Hz, 2H), 4.16 – 4.12 (m, 1H), 0.91 (s, 9H); ^{13}C NMR (101 MHz, CD_3OD) δ 165.60, 164.65, 164.45, 155.25, 147.11, 143.28, 138.64, 135.55, 135.45, 133.30, 132.14, 132.00, 129.97, 129.86, 129.33, 128.98, 128.25, 128.17, 127.74, 127.70, 127.63, 127.60, 127.50, 127.30, 126.94, 126.86, 99.26, 91.59, 91.35, 81.53, 81.48, 76.90, 76.67, 75.61, 75.33, 42.88, 26.22, 25.92, 18.38; HRMS (ESI-TOF) calcd for $C_{39}H_{41}N_4O_6Si$ [M+ H]$^+$ 689.2795; found 689.2795.

***N*^{4}-Benzoyl 1-(2′-*O*-Benzoyl-3′-*O*-tertbutyldiphenylsilyl α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine (3a):** A round bottomed flask equipped with a stir bar was charged with **2a** (1.35 g, 1.96 mmol, 1.0 equiv), then sealed and evacuated for 10-15 min. After attaching a N_2 balloon, anhydrous pyridine (8.0 mL) was added and the solution cooled to 0˚C whereupon benzoyl chloride (0.45 mL, 3.92 mmol, 2.0 equiv) was added and the resulting mixture stirred at 0˚C for 1.5 h. After warming to rt, the solvent was removed under reduced pressure and the resulting residue dissolved in CH_2Cl_2. The solution was washed with water (3 x 200 mL), brine (3 x 200 mL), dried over Na_2SO_4 and then concentrated under reduced pressure. The crude product was purified by silica gel (deactivated with 3% Et_3N/CH_2Cl_2) chromatography, dry loading with Celite and eluting with 0-0.75% MeOH/CH_2Cl_2 to afford **3a** (1.027 g, 1.30 mmol, 66.2%) as light brown solid. 1H NMR (400 MHz, CD_3OD) δ 9.01 (s, 1H), 7.93 (d, J = 7.7 Hz, 2H), 7.80 – 7.06 (m, 24H), 5.98 (s, 1H), 5.70 (s, 1H), 4.47 (m, 5H), 0.94 (s, 9H); 13C NMR (101 MHz, CD_3OD) δ 178.76, 164.50, 161.84, 158.39, 147.56, 146.89, 137.88, 135.97, 135.51, 133.49, 133.49, 132.50, 132.23, 131.95, 129.95, 129.42, 128.86, 128.62, 128.33, 128.16, 128.07, 104.80, 91.80, 80.90, 77.39, 75.33, 43.34, 26.06, 18.52.

***N*^{4}-Benzoyl 1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5 benzylcarbamoylcytosine (4a):** To a stirred 0˚C solution of **3a** (1.02 g, 1.29 mmol) in THF (6.0 mL) was added dropwise a 1M THF solution of TBAF (5.0 mL, 5.0 mmol). After stirring at 0°C for 2 h, the reaction was concentrated under reduced pressure and the residue was dissolved in CH_2Cl_2. The solution was washed with water and then brine, dried over Na_2SO_4 and then concentrated under reduced pressure. The crude product was purified by silica gel (deactivated with 3% Et_3N/CH_2Cl_2) chromatography, dry loading with Celite and eluting with 0-0.5% MeOH/CH_2Cl_2 to afford **4a** (0.438 g, 0.79 mmol, 61%) as white solid.; HRMS (ESI-TOF) calcd for $C_{30}H_{26}N_4O_7Na$ [M+ Na]$^+$ 577.1699; found 577.1700.

***N*^{4}-Benzoyl 1-(2′-*O*-Benzoyl-3′-*O*-[(2-cyanoethoxy)(diisopropylamino)phosphino]α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine (5a):** To a stirred solution of **4a** (0.178 g, 0.32 mmol) and DMAP (15 mg, 0.12 mmol, 0.36 equiv) in CH_2Cl_2 (30 mL) under argon at rt was added (*i*-Pr)$_2$NEt (1.2 mL, 6.91 mmol, 21.6 equiv) followed by 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphine (0.534 mL, 2.38 mmol, 7.44 equiv). After stirring for 2 h at rt, the solution was diluted with CH_2Cl_2 (200 mL), washed with water and then brine, dried over Na_2SO_4 and concentrated under reduced pressure. The resulting crude mixture was purified by silica gel automated column chromatography, dry loading with Celite and eluting using 5-10% acetone/CH_2Cl_2 to afford product **5a (**0.165 g, 0.219 mmol, 68%) as white solid; ^{31}P NMR (162 MHz, CD_3CN) δ 150.9, 150.3.

***N*^{4}-Benzoyl 1-(2′-*O*-Benzoyl-3′-*O*-[bis(2-cyanoethyl)- phosphoryl]-α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine (6a):** To a stirred solution of **5a** (0.132 g, 0.174 mmol) in MeCN (10 mL) was added 3-hydroxypropionitrile (0.058 mL, 0.848 mmol, 4.9 equiv) followed by a solution of 0.4 M tetrazole in MeCN (2.7 mL, 0.742 mmol, 4.26 equiv). After 1 h of stirring at rt, complete consumption of the starting material was observed by TLC whereupon H_2O_2 (30% in H_2O, 0.03 mL, 0.394 mmol, 2.26 equiv) was added and reaction stirred until complete conversion of the intermediate product was observed. The solvent was removed under diminished pressure and the resulting residue dissolved in CH_2Cl_2 (200 mL) washed with water

(3 x 200 mL) and brine (3 x 200 mL), dried over Na_2SO_4 and concentrated to afford **6a** (0.116 g, 0.157 mmol, 90%) as white foam. The obtained compound was pure enough to be carried into the next step without further purification; ^{31}P NMR (162 MHz, $CDCl_3$) δ -4.1.

1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine 3′ monophosphate (7a): Compound **6a** (0.153 g, 0.206 mmol) was suspended in NH_4OH (7 mL, 30-33% NH_3 in H_2O) and stirred for 20 h at 37°C. The mixture was cooled to rt, diluted with water and washed with CH_2Cl_2 (3 x 40 mL). The aqueous layer was then lyophilized (at 10 mT and -105°C) to afford the desired monophosphate **7a** (0.082 g, 0.192 mmol, 93%) as a white solid; ^{31}P NMR (162 MHz, D_2O) δ 0.8; HRMS (ESI-TOF) calcd for $C_{16}H_{20}N_4O_8P$ $[M+H]^+$ 427.1019; found 427.1004.

1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine 3′ monophosphor-2-methylimidazolide (8a): To a solution of the nucleoside monophosphate **7a** (0.050 g, 0.12 mmol) in anhydrous DMF (2 mL) under a nitrogen atmosphere was slowly added anhydrous Et_3N (0.044 mL, 0.312 mmol, 2.6 equiv) at 0 °C. After stirring for 5 min, 2-methylimidazole (0.089 g, 1.08 mmol, 9 equiv) was added followed by PPh_3 (0.066 g, 0.24 mmol, 2.0 equiv). After stirring at room temperature for 10 min, 2,2′-dipyridyl disulfide (0.072 g, 0.324 mmol, 2.7 equiv) was added and the mixture further stirred for 5 h at rt with monitoring by analytical HPLC (mobile phase: MeCN/0.05 M TEAB buffer, from 0% to 70% over 42 minutes). Upon consumption of the starting material, the product was precipitated by adding the reaction mixture dropwise to vigorously stirred Et_2O (100 mL). The precipitate was pelleted by centrifugation at 4400 rpm for 5 minutes at rt. After discarding the supernatant, the pellet was resuspended in a minimal amount of DMF (2 mL). The resulting solution was then precipitated a second time by adding it dropwise to a solution of $NaClO_4$ (150 mg, 1.23 mmol) in EtOAc/Et_3N (15 mL/0.6 mL) The resulting suspension was centrifuged at 4400 rpm for 5 min at rt, the supernatant discarded, the pellet washed with 1:2 Et_2O/EtOAc (2 x 40 mL) and then dried under high vacuum to afford the product **8a** (0.038 g, 0.074 mmol, 66.6%) as white solid.

1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5-benzylcarbamoyl-cytosine 3′ triphosphate (9a): To a mixture containing activated nucleoside monophosphate **8a** (0.02 g, 0.039 mmol), and 1-(2-(pyrenesulfonyl)ethyl)pyrophosphate (0.022 g, 0.0467 mmol, 1.2 equiv) was added a 1.0 M solution of $ZnCl_2$ in anhydrous DMF (0.350 mL, 0.35 mmol, 10 equiv) under a nitrogen atmosphere. The mixture was stirred at rt for 6 h and the reaction progress monitored by HPLC (MeCN/0.05 M TEAB buffer, from 0% to 70% over 42 minutes). Upon complete consumption of the starting material, the product was precipitated by adding the reaction mixture dropwise to stirred EtOAc (100 mL). The precipitate was pelleted by centrifuging at 4400 rpm for 5 min at rt. After discarding the supernatant, the pellet was resuspended in 20:80 water/MeCN containing 2% (*i*-Pr)$_2$NEt and centrifuged at 4400 rpm for 5 min and the supernatant collected. The solid pellet was again resuspended by 20:80 water/MeCN containing 2% (*i*-Pr)$_2$NEt and centrifuged at 4400 rpm for 5 min. The combined supernatants were evaporated under diminished pressure and the resulting crude dry-loaded in celite onto a silica gel column (deactivated with 100 mL 4% (*i*-Pr)$_2$NEt/CH_2Cl_2) and eluted with (5-10% water/acetone with 2% (*i*-Pr)$_2$NEt). The fractions containing the product were collected and evaporated under diminished pressure at 30-40°C to afford the fully protected nucleoside triphosphate. The fully protected nucleoside triphosphates were subsequently dissolved in 30-33% aqueous NH_4OH (6.0 mL) and stirred for 12 h at rt in a sealed tube. After concentrating the solution under reduced pressure in rotavapor, the resulting solid was resuspended with minimum volume of MilliQ water (4-5 mL) and the solution extracted with CH_2Cl_2 (10 mL) followed by EtOAc (10 mL). The aqueous layer was then filtered through a 0.22 μm syringe filter and concentrated under diminished pressure. To the concentrated aqueous extract was added $NaClO_4$ (0.072 g,

0.585 mmol, 15 equiv) in acetone (100 mL, 40 volumes) dropwise at rt. The resulting suspension was centrifuged at 4400 rpm for 5 min at rt, the supernatant discarded, and the pellet washed with organic solution (10:1 acetone/CH_2Cl_2, 2 x 50 mL), and then dried under vacuum to afford the desired product **9a** (0.01 g, 0.015 mmol, 38.5% yield over two steps) as white solid; ^{1}H NMR (400 MHz, D_2O) δ 8.3 (s, 1H), 7.4-7.3 (m, 5H), 5.8 (s, 1H), 4.9 (m, 1H), 4.6-4.4 (m, 5H); ^{31}P NMR (162 MHz, D_2O) δ -5.3, -12.4, -19.3; HRMS (ESI-TOF) calcd for $C_{16}H_{21}N_5O_{14}P_3Na_3$ $[M-3H+3Na+NH_3]^+$ 670.1051; found 670.1049.

1-(2′-*O*-Benzoyl-3′-*O*-tertbutyldiphenylsilyl-α-L-threofuranosyl)-5-propylphenyl carbamoyl-cytosine (2b): To a 100 mL round-bottomed flask equipped with a stir-bar was added **1** (1.13 g, 1.66 mmol, 1.0 equiv) and $Pd(PPh_3)_4$ (38.4 mg, 0.0332 mmol, 0.02 equiv). After the vessel was sealed and evacuated for 10-15 min, DMF (6.0 mL) was added, followed by Et_3N (8.3 mL, 8.3 mmol, 5.0 equiv) and phenylpropylamine (1.5 mL, 10.5 mmol, 6.3 equiv). The vessel was then equipped with 2 balloons filled with CO and the reaction stirred at rt for 48 h. The crude reaction mixture was diluted with CH_2Cl_2 (300 mL), washed with water (3 x 200 mL) and then brine (3 x 200 mL). The combined organics were dried over $MgSO_4$ and concentrated under diminished pressure. The crude residue was dry loaded with neutral alumina and purified by silica gel chromatography eluting with (0-1%) MeOH/CH_2Cl_2 containing 1% Et_3N to afford product **2b (**0.430 g, 0.6 mmol, 36.1%) as light brown solid; ^{1}H NMR (400 MHz, THF-d_8) δ 8.32 (s, 1H), 7.94 (d, *J* = 7.5 Hz, 2H), 7.77 (t, *J* = 5.5 Hz, 1H), 7.66-7.51 (ddd, *J* = 28.9, 17.3, 7.5 Hz, 5H), 7.47 – 7.00 (m, 15H), 6.00 (s, 1H), 5.77 (s, 1H), 4.42 (s, 1H), 4.23 (d, *J* = 9.7 Hz, 1H), 4.04 (d, *J* = 9.8 Hz, 1H), 3.39 – 3.17 (m, 2H), 2.68 – 2.58 (m, 2H), 1.91 – 1.78 (m, 2H), 0.99 (s, 9H); ^{13}C NMR (101 MHz, THF-d_8) δ 167.54, 166.55, 165.93, 155.52, 144.17, 143.80, 137.55, 134.84, 133.97, 131.68, 131.42, 129.90, 129.51, 127.36, 100.86, 93.28, 83.80, 78.16, 77.43, 40.88, 35.04, 33.26, 28.12, 20.46; HRMS (ESI-TOF) calcd for $C_{41}H_{44}N_4O_6SiNa$ $[M+ Na]^+$ 739.2928; found 739.2902.

***N*^{4}-Benzoyl 1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5 propylphenylcarbamoylcytosine (3b):** A round bottomed flask equipped with a stir bar was charged with 2**b** (0.577 g, 0.805 mmol, 1.0 equiv), then sealed and evacuated for 10-15 min. After attaching a N_2 balloon, anhydrous pyridine (5.0 mL) was added and the solution cooled to 0˚C whereupon benzoyl chloride (0.21 mL, 1.81 mmol, 2.24 equiv) was added and the resulting mixture stirred at 0˚C for 2 h. After warming to rt, the solvent was removed under reduced pressure and the resulting residue dissolved in CH_2Cl_2. The solution was washed with water (3 x 200 mL), brine (3 x 200 mL), dried over Na_2SO_4 and then concentrated under reduced pressure. The crude product was purified by silica gel (deactivated with 3% Et_3N/CH_2Cl_2) chromatography, dry loading with Celite and eluting with 0-0.25% MeOH/CH_2Cl_2 to afford **3b** (0.652 g, 0.794 mmol, 98.7%) as light brown solid; HRMS (ESI-TOF) calcd for $C_{48}H_{48}N_4O_7SiNa$ $[M+ Na]^+$ 843.3190; found 843.3214.

***N*^{4}-Benzoyl 1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5 propylphenylcarbamoylcytosine (4b):** To a stirred 0˚C solution of **3b** (0.623 g, 0.759 mmol) in THF (6.0 mL) was added dropwise a 1M THF solution of TBAF (3.1 mL, 3.1 mmol). After stirring at 0°C for 3 h, the reaction was concentrated under reduced pressure and the residue was dissolved in CH_2Cl_2. The solution was washed with water and then brine, dried over Na_2SO_4 and then concentrated under reduced pressure. The crude product was purified by silica gel (deactivated with 3% Et_3N/CH_2Cl_2) chromatography, dry loading with Celite and eluting with 0-0.5% MeOH/CH_2Cl_2 to afford **4b** (0.308 g, 0.529 mmol, 69.6%) as white solid; ^{1}H NMR (400 MHz, THF-d_8) δ 13.77 (s, 1H), 10.03 (s, 1H), 9.08 (s, 1H), 8.07 (dd, *J* = 25.9, 7.8 Hz, 4H), 7.60-7.37 (m, 7H), 7.21 – 7.10 (m, 4H), 6.12 (s, 1H), 5.44 (s, 1H), 4.43 – 4.35 (m, 2H), 4.28 (d, *J* = 9.8 Hz, 1H), 3.46 (dd, *J* = 12.9, 6.6 Hz, 2H), 2.73 – 2.64 (m, 2H), 2.02 – 1.91 (m, 2H); $C_{32}H_{30}N_4O_7Na$ $[M+ Na]^+$ 605.2012; found 605.2017.

***N*⁴-Benzoyl 1-(2′-*O*-Benzoyl-3′-*O*-[(2-cyanoethoxy)(diisopropylamino)phosphino]α-L-threofuranosyl)-5-propylphenylcarbamoyl-cytosine (5b):** To a stirred solution of **4b** (0.221 g, 0.379 mmol) and DMAP (23 mg, 0.184 mmol, 0.48 equiv) in CH_2Cl_2 (30 mL) under argon at rt was added (*i*-Pr)$_2$NEt (1. mL, 5.76 mmol, 15.2 equiv) followed by 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphine (0.350 mL, 1.56 mmol, 4.2 equiv). After stirring for 2 h at rt, the solution was diluted with CH_2Cl_2 (200 mL), washed with water and then brine, dried over Na_2SO_4 and concentrated under reduced pressure. The resulting crude mixture was purified by silica gel automated column chromatography, dry loading with Celite and eluting using 5-10% acetone/CH_2Cl_2 to afford product **5b (**0.205 g, 0.262 mmol, 69.2%) as white solid; ^{1}H NMR (400 MHz, CD_3CN) δ 13.65 (s, 1H), 10.03 (s, 1H), 8.83 (s, 1H), 8.05 (m, 4H), 7.66-7.44 (m, 7H), 7.22 – 7.14 (m, 4H), 6.16 (m, 1H), 5.73-5.55 (m, 1H), 4.64 – 2.49 (m, 15H), 10.06 (m, 12H); ^{31}P NMR (162 MHz, CD_3CN) δ 150.9, 150.25; ^{13}C NMR (101 MHz, CD_3CN) δ 179.08, 164.70, 161.15, 159.18, 147.67, 146.87, 141.88, 136.63, 133.88, 132.99, 129.76, 129.15, 128.76, 128.64, 128.39, 128.35, 125.86, 118.59, 118.22, 105.46, 91.40, 80.48, 80.04, 76.83, 74.90, 74.71, 58.96, 58.86, 58.78, 58.66, 43.23, 43.10, 38.83, 32.93, 31.31, 23.97, 23.90, 23.80, 23.73, 23.66, 19.91, 19.84; $C_{41}H_{47}N_6O_8PNa$ $[M+ Na]^+$ 805.3091; found 805.3085.

***N*⁴-Benzoyl 1-(2′-*O*-Benzoyl-3′-*O*-[bis(2-cyanoethyl)- phosphoryl]-α-L-threofuranosyl)-5-propylphenylcarbamoyl-cytosine (6b):** To a stirred solution of **5b** (0.204 g, 0.260 mmol) in MeCN (10 mL) was added 3-hydroxypropionitrile (0.150 mL, 2.19 mmol, 8.42 equiv) followed by a solution of 0.4 M tetrazole in MeCN (2.5 mL, 1.04 mmol, 4.0 equiv). After 1 h of stirring at rt, complete consumption of the starting material was observed by TLC whereupon H_2O_2 (30% in H_2O, 0.150 mL, 1.47 mmol, 5.65 equiv) was added and reaction stirred until complete conversion of the intermediate product was observed. The solvent was removed under diminished pressure and the resulting residue dissolved in CH_2Cl_2 (200 mL) washed with water (3 x 200 mL) and brine (3 x 200 mL), dried over Na_2SO_4 and concentrated to afford **6b** (0.194 g, 0.252 mmol, 97%) as white foam. The obtained compound was pure enough to be carried into the next step without further purification; ^{31}P NMR (162 MHz, $CDCl_3$) δ -3.83.

1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5-propylphenylcarbamoyl-cytosine-3′-monophosphate (7b): Compound **6b** (0.188 g, 0.245 mmol) was suspended in NH_4OH (7 mL, 30-33% NH_3 in H_2O) and stirred for 17 h at 37°C. The mixture was cooled to rt, diluted with water and washed with CH_2Cl_2 (3 x 40 mL). The aqueous layer was then lyophilized (at 10 mT and -105°C) to afford the desired monophosphate **7b** (0.1 g, 0.22 mmol, 90%) as a white solid; ^{31}P NMR (162 MHz, D_2O) δ 0.4996; $C_{18}H_{22}N_4O_8P$ $[M-H]^-$ 453.1175; found 453.1155.

1-(2'-*O*-Benzoyl-α-L-threofuranosyl)-5-propylphenylcarbamoyl-cytosine-3′-monophosphor-2-methylimidazolide (8b): To a solution of the nucleoside monophosphate **7b** (0.065 g, 0.143 mmol) in anhydrous DMF (2 mL) under a nitrogen atmosphere was slowly added anhydrous Et_3N (0.1 mL, 0.709 mmol, 4.96 equiv) at 0 °C. After stirring for 5 min, 2-methylimidazole (0.106 g, 1.29 mmol, 9 equiv) was added followed by PPh_3 (0.098 g, 0.356 mmol, 2.5 equiv). After stirring at room temperature for 10 min, 2,2′-dipyridyl disulfide (0.086 g, 0.387 mmol, 2.7 equiv) was added and the mixture further stirred for 5 h at rt with monitoring by analytical HPLC (mobile phase: MeCN/0.05 M TEAB buffer, from 0% to 70% over 42 minutes). Upon consumption of the starting material, the product was precipitated by adding the reaction mixture dropwise to vigorously stirred Et_2O (100 mL). The precipitate was pelleted by centrifugation at 4400 rpm for 5 minutes at rt. After discarding the supernatant, the pellet was resuspended in a minimal amount of DMF (2 mL). The resulting solution was then precipitated a second time by adding it dropwise to a solution of $NaClO_4$ (150 mg, 1.23 mmol) in EtOAc/Et_3N (15 mL/0.6 mL) The resulting suspension was centrifuged at 4400 rpm for 5 min at rt, the supernatant discarded, the pellet washed with 1:2 Et_2O/EtOAc (2 x 40 mL) and

then dried under high vacuum to afford the product **8b** (0.049 g, 0.091 mmol, 66.2%) as white solid.

1-(2′-*O*-Benzoyl-α-L-threofuranosyl)-5-propylphenylcarbamoyl-cytosine 3′-triphosphate (9b): To a mixture containing activated nucleoside monophosphate **8b** (0.049 g, 0.0945 mmol), and 1-(2-(pyrenesulfonyl)ethyl)pyrophosphate (0.054 g, 0.115 mmol, 1.22 equiv) was added a 1.0 M solution of $ZnCl_2$ in anhydrous DMF (1 mL, 1.0 mmol, 10.6 equiv) under a nitrogen atmosphere. The mixture was stirred at rt for 6 h and the reaction progress monitored by HPLC (MeCN/0.05 M TEAB buffer, from 0% to 70% over 42 minutes). Upon complete consumption of the starting material, the product was precipitated by adding the reaction mixture dropwise to stirred EtOAc (100 mL). The precipitate was pelleted by centrifuging at 4400 rpm for 5 min at rt. After discarding the supernatant, the pellet was resuspended in 20:80 water/MeCN containing 2% $(i\text{-Pr})_2NEt$ and centrifuged at 4400 rpm for 5 min and the supernatant collected. The solid pellet was again resuspended by 20:80 water/MeCN containing 2% $(i\text{-Pr})_2NEt$ and centrifuged at 4400 rpm for 5 min. The combined supernatants were evaporated under diminished pressure and the resulting crude dry-loaded in celite onto a silica gel column (deactivated with 100 mL 4% $(i\text{-Pr})_2NEt/CH_2Cl_2$) and eluted with (5-10% water/acetone with 2% $(i\text{-Pr})_2NEt$). The fractions containing the product were collected and evaporated under diminished pressure at 30-40°C to afford the fully protected nucleoside triphosphate. The fully protected nucleoside triphosphates were subsequently dissolved in 30-33% aqueous NH_4OH (6.0 mL) and stirred for 6 h at rt in a sealed tube. After concentrating the solution under reduced pressure in rotavapor, the resulting solid was resuspended with minimum volume of MilliQ water (4-5 mL) and the solution extracted with CH_2Cl_2 (10 mL) followed by EtOAc (10 mL). The aqueous layer was then filtered through a 0.22 μm syringe filter and concentrated under diminished pressure. To the concentrated aqueous extract was added $NaClO_4$ (0.072 g, 0.585 mmol, 15 equiv) in acetone (100 mL, 40 volumes) dropwise at rt. The resulting suspension was centrifuged at 4400 rpm for 5 min at rt, the supernatant discarded, and the pellet washed with organic solution (10:1 acetone/CH_2Cl_2, 2 x 50 mL), and then dried under vacuum to afford the desired product **9b** (0.012 g, 0.017 mmol, 20.7% yield over two steps) as white solid; ^1H NMR (400 MHz, D_2O) δ 7.99 (s, 1H), 7.33-7.18 (m, 5H), 5.78 (s, 1H), 4.58-4.38 (m, 4H), 3.41-3.17 (m, 2H), 2.68(m, 2H), 1.91 (m, 2H); ^{31}P NMR (162 MHz, D_2O) δ -5.31, -12.47, -19.66; $C_{18}H_{25}N_5O_{14}P_3Na_3$ $[M\text{-}3H+3Na+NH_3]^+$ 698.0402; found 698.0400.

Molecular Biology

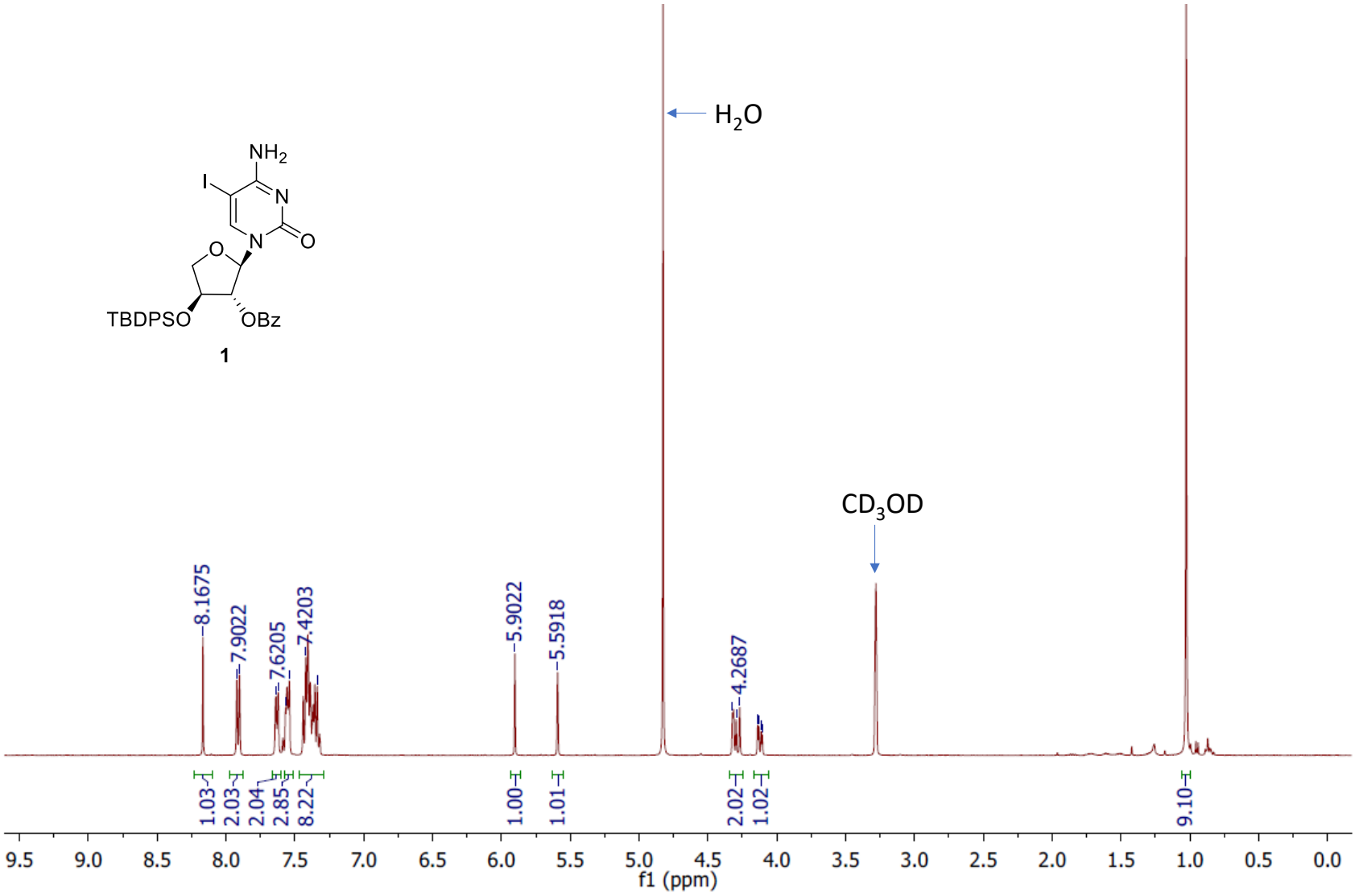
Primer extension assay. The primer-extension reactions were done using 1 μM of IR680-labeled DNA primer and equimolar amount of DNA template with 2 μM of Kod-RSGA in 1x Thermopol buffer (20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton® X-100, pH 8.8 at 25°C) and 100 μM of each tNTP in 20 μL volume. The template and primer were annealed by heating in the buffer to 95°C for 5 min and cooled on ice for 5 min. Then the reactions were initiated by adding the tNTPs and polymerase and incubating at 55°C for 2 hours. The reactions were quenched by adding stop buffer (95% form amide, 25 mM EDTA) and analyzed by denaturing PAGE, the gels were imaged in a LI-COR Odyssey CLx imager.

BLI assay. Aptamers 123 and 133 were prepared as biotin-labeled TNA molecules using the 5' biotinylated PBS8 20 nt DNA primer and the corresponding template, which contained an additional 20 nucleotides, $(AAAC)_5$, at the 3′ end to facilitate band separation by PAGE. A 750 μL reaction volume containing 1 μM of both primer and template as well as 1× ThermoPol

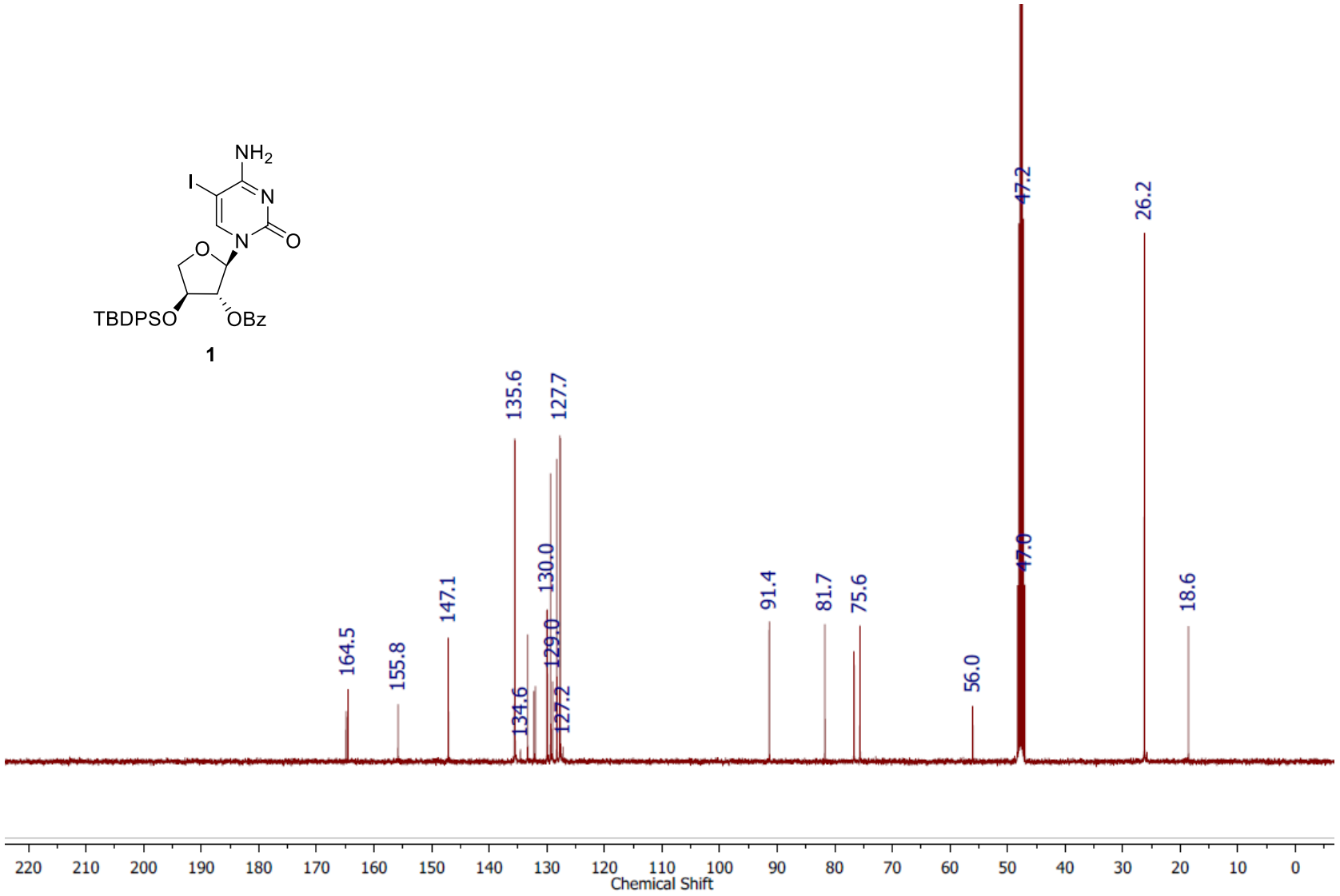
buffer was heated for 5 min at 90°C and then cooled on ice for 5 min to anneal. TNA polymerization was initiated by adding 100 μM of each tNTP and 2 μM of Kod-RSGA TNA polymerase and then incubated for 2 h at 55°C. Full-length biotinylated aptamers were purified by 10% denaturing polyacrylamide gel electrophoresis for 1.5 h at a constant 16 W. TNA was recovered from the gel by electroelution at 60 V overnight and then buffer-exchanged into water and concentrated using a YM-10 Centricon centrifugal filter device. TNA concentration was quantified by measuring its absorbance at 260 nm with a NanoDrop spectrophotometer.

Biotinylated aptamers at a concentration of 50 nM were folded into their active conformation in BLI binding buffer (10 mM HEPES pH 7, 150 mM NaCl, 3 mM EDTA, 0.05% Tween 20) by heating to 90°C for 5 minutes and then cooling on ice for 10 minutes. Streptavidin-coated optical biosensors were previously equilibrated in BLI binding buffer for 30 min. For measuring full kinetics, the BLI run was performed with the following steps: a buffer only baseline for 60 s to equilibrate sensors, loading the aptamer for 200 s, a second buffer only baseline for 200 s, an association phase with the target protein for 600 s, and a dissociation phase for 600 s. Serial diluted target concentration ranged from 50 nM to 3.125 nM. The experiments were conducted at 30ºC. Data were analyzed using the Octet Data Analysis HT software. The buffer only baseline measurement was used to subtract the background from all samples before applying Savitzky-Golay filtering and fitting the association and dissociation curves and applying a global fit to determine K_D.

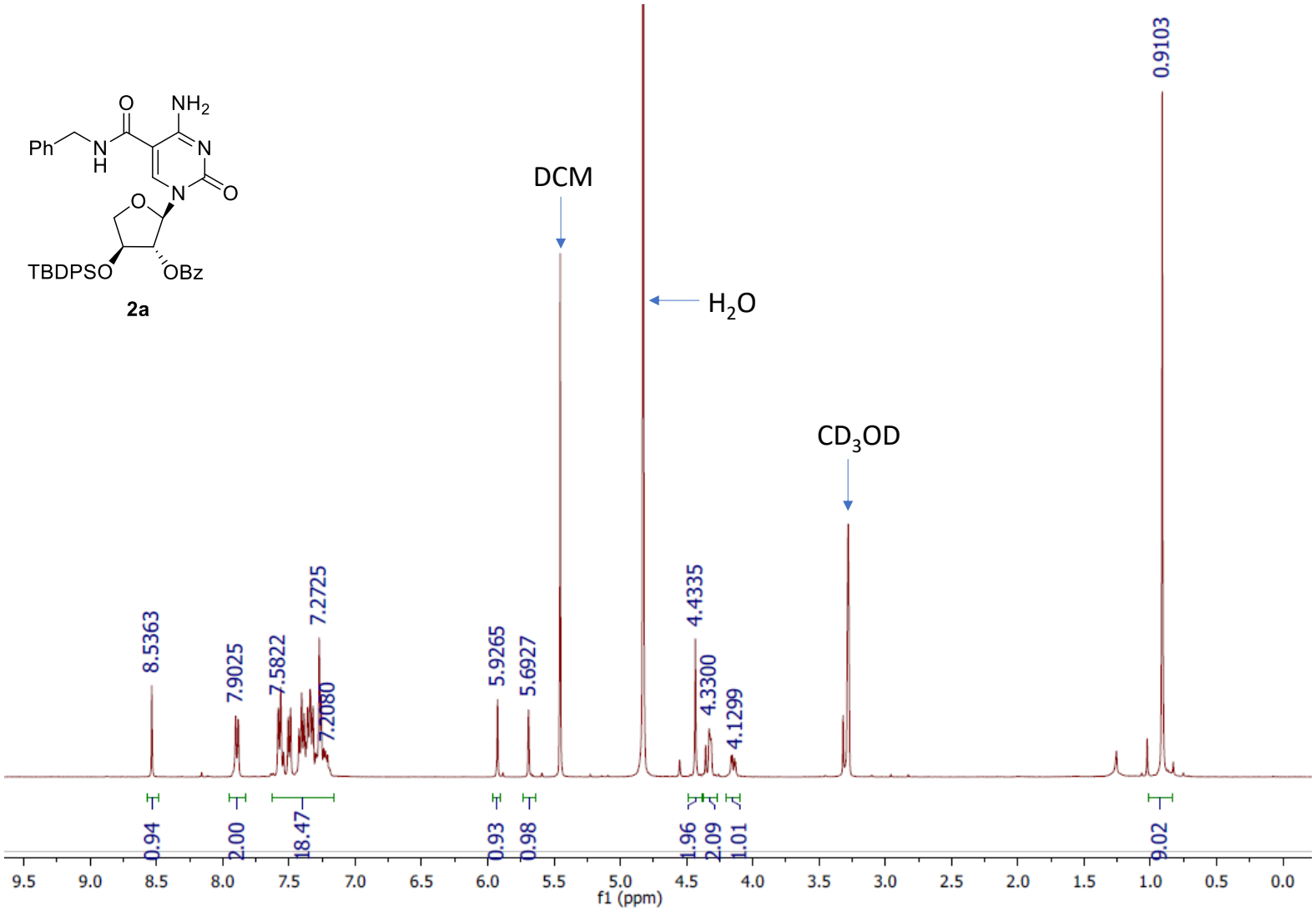
MALDI-TOF analysis. Oligonucleotides were prepared using the 10nt PBS8 DNA primer and the corresponding template, which contained an additional 12 nucleotides, (AAAC)3, at the 3′ end to facilitate the band separation by PAGE. A 750 μL reaction volume containing 1 μM of both primer and template as well as 1× ThermoPol Buffer was heated for 5 min at 90°C and then cooled on ice for 5 min to anneal. TNA polymerization was initiated by adding 100 μM of each tNTP and 2 μM of Kod-RSGA TNA polymerase and then incubated for 30 min at 55°C. Full-length TNA oligonucleotides were purified by 20% denaturing polyacrylamide gel electrophoresis for 2 h at a constant 16 W. TNA was recovered from the gel by electroelution at 60 V overnight and then buffer-exchanged into water and concentrated using a YM-30 Centricon centrifugal filter device. TNA concentration was quantified by measuring its absorbance at 260 nm with a NanoDrop spectrophotometer. Oligonucleotides (20 μM in de-ionized water) were spotted onto a MALDI-TOF steel target plate (1 μL), heated at 55°C for <3 minutes to dry, and covered with 1 μL of matrix (18 mg/mL 2,4,6-trihydroxy-acetophenone (THAP) and 6 mg/mL dibasic ammonium citrate in HPLC grade 50:50 acetonitrile/water) for characterization on a Bruker Daltronics Microflex MALDI-TOF mass spectrometer.



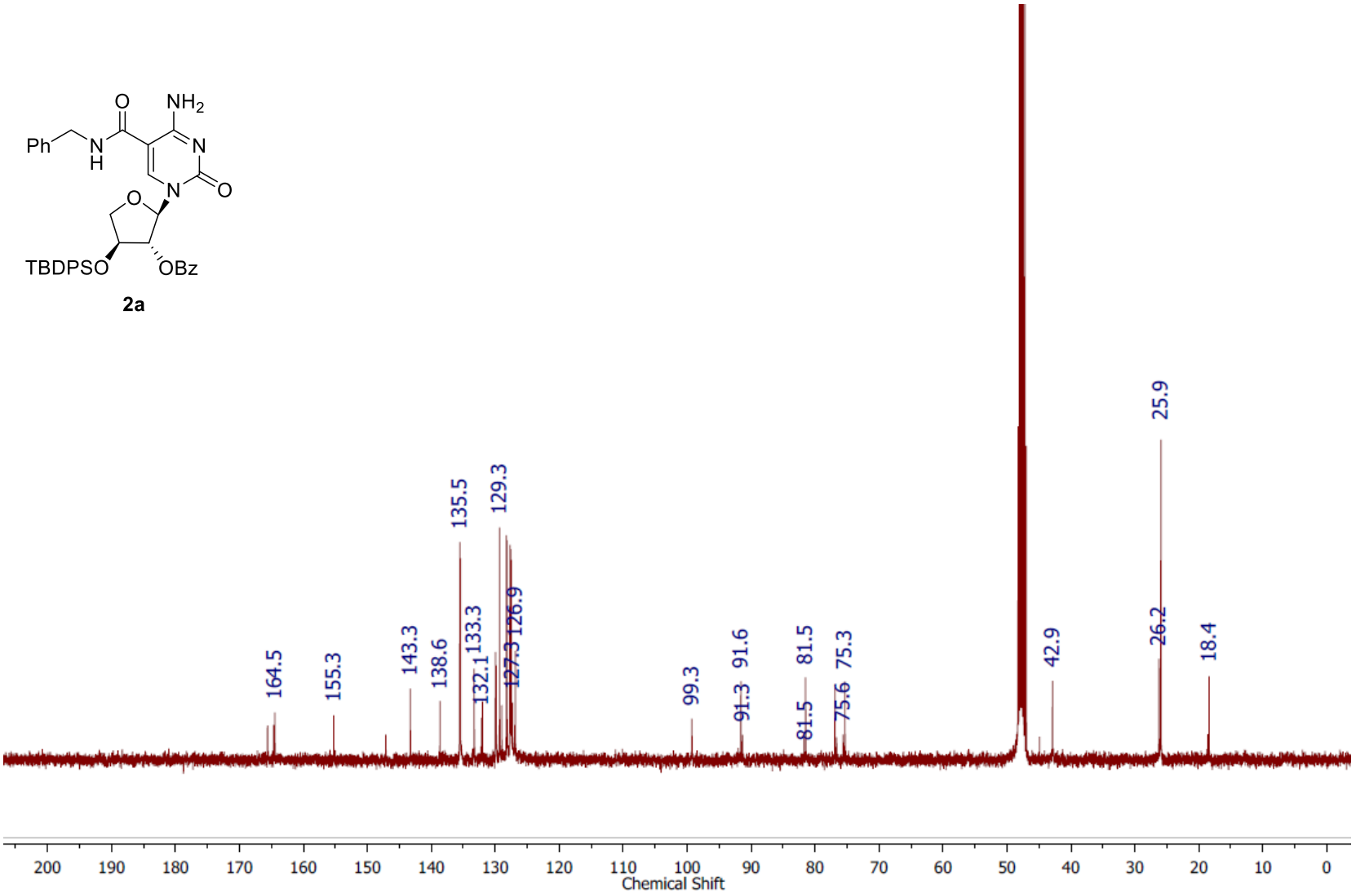
^{1}H NMR spectrum of compound **1** (400 MHz, CD_3OD)



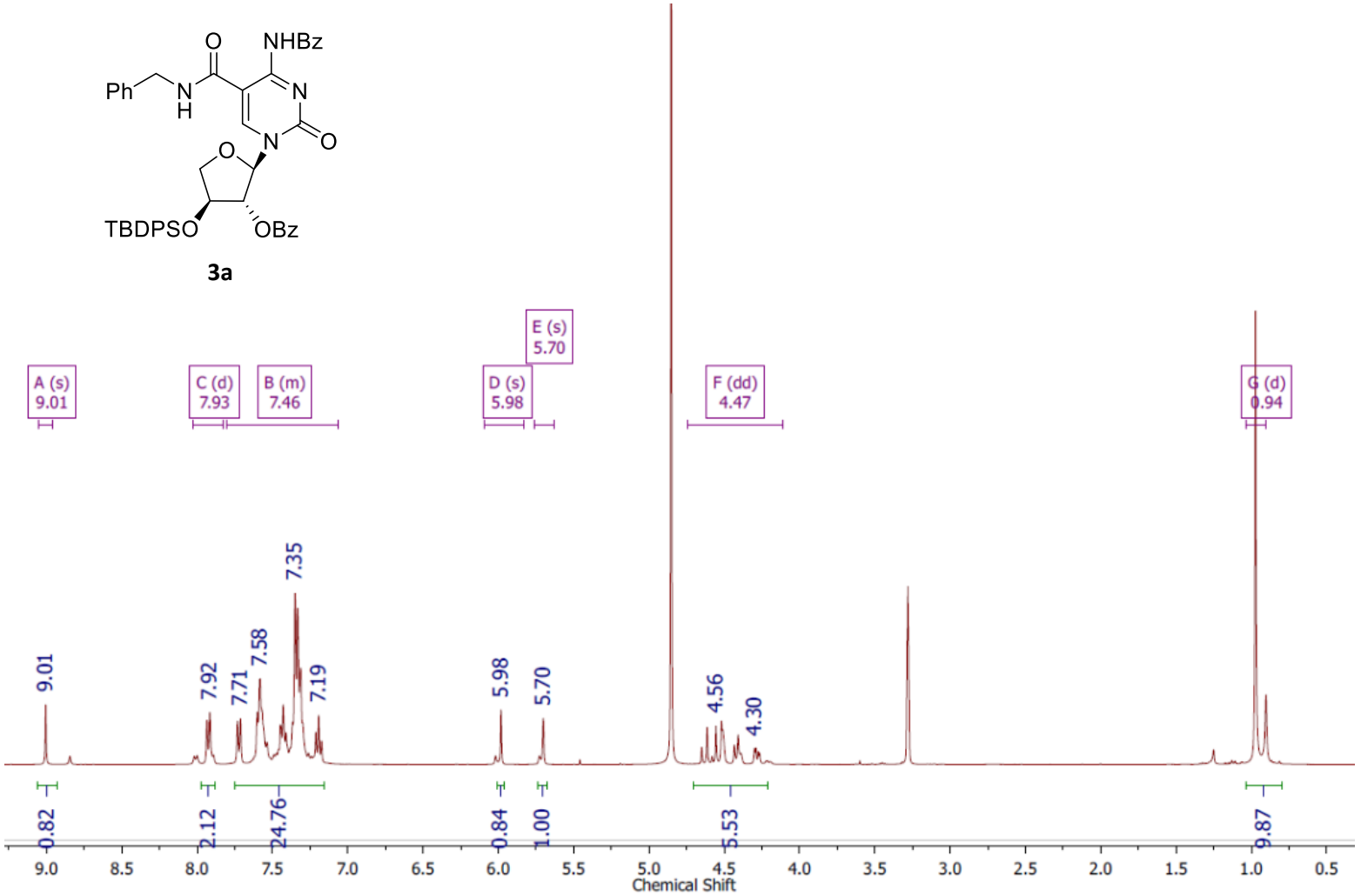
^{13}C NMR spectrum of compound **1** (101 MHz, CD_3OD)



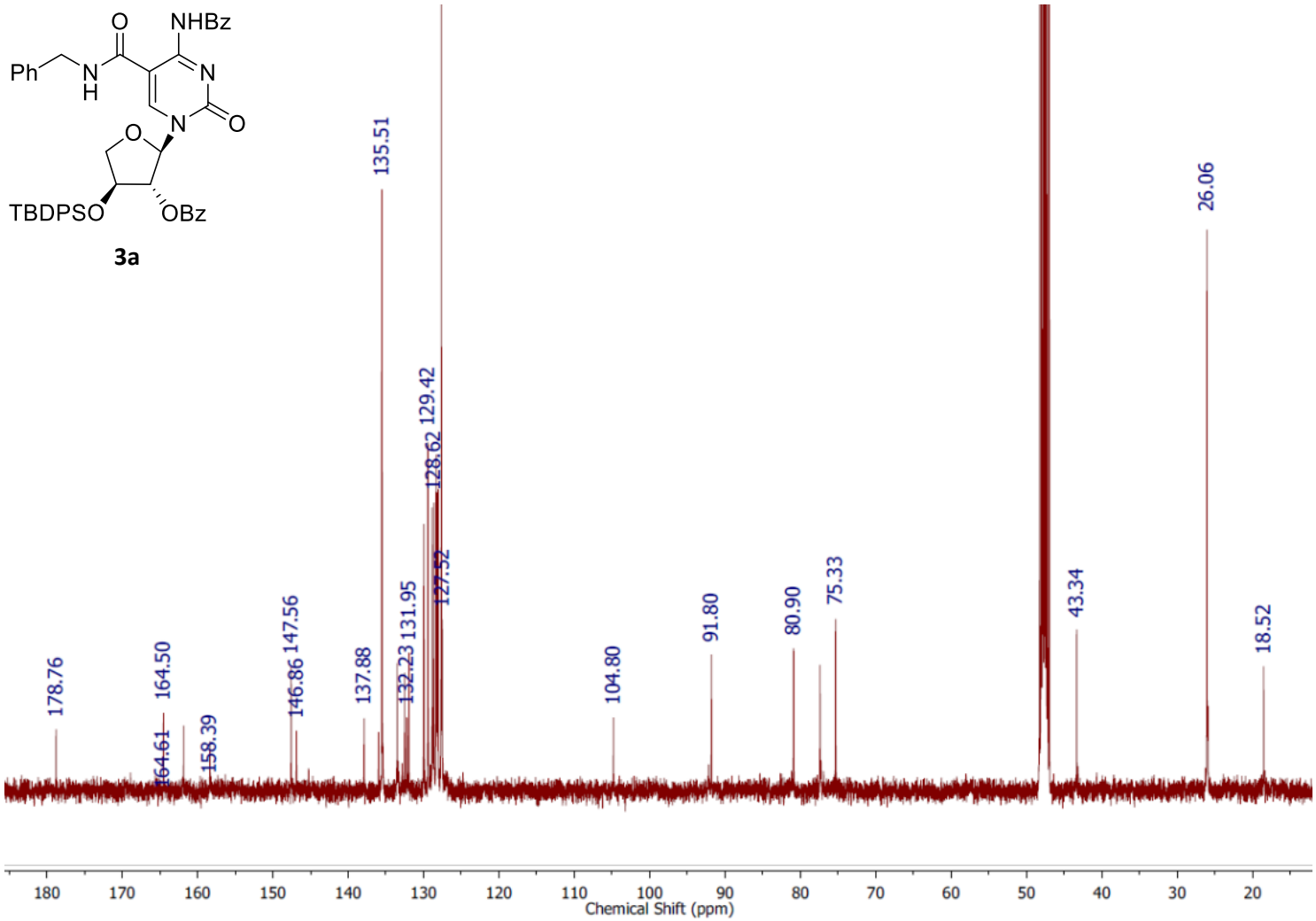
^{1}H NMR spectrum of compound **2a** (400 MHz, CD_3OD)



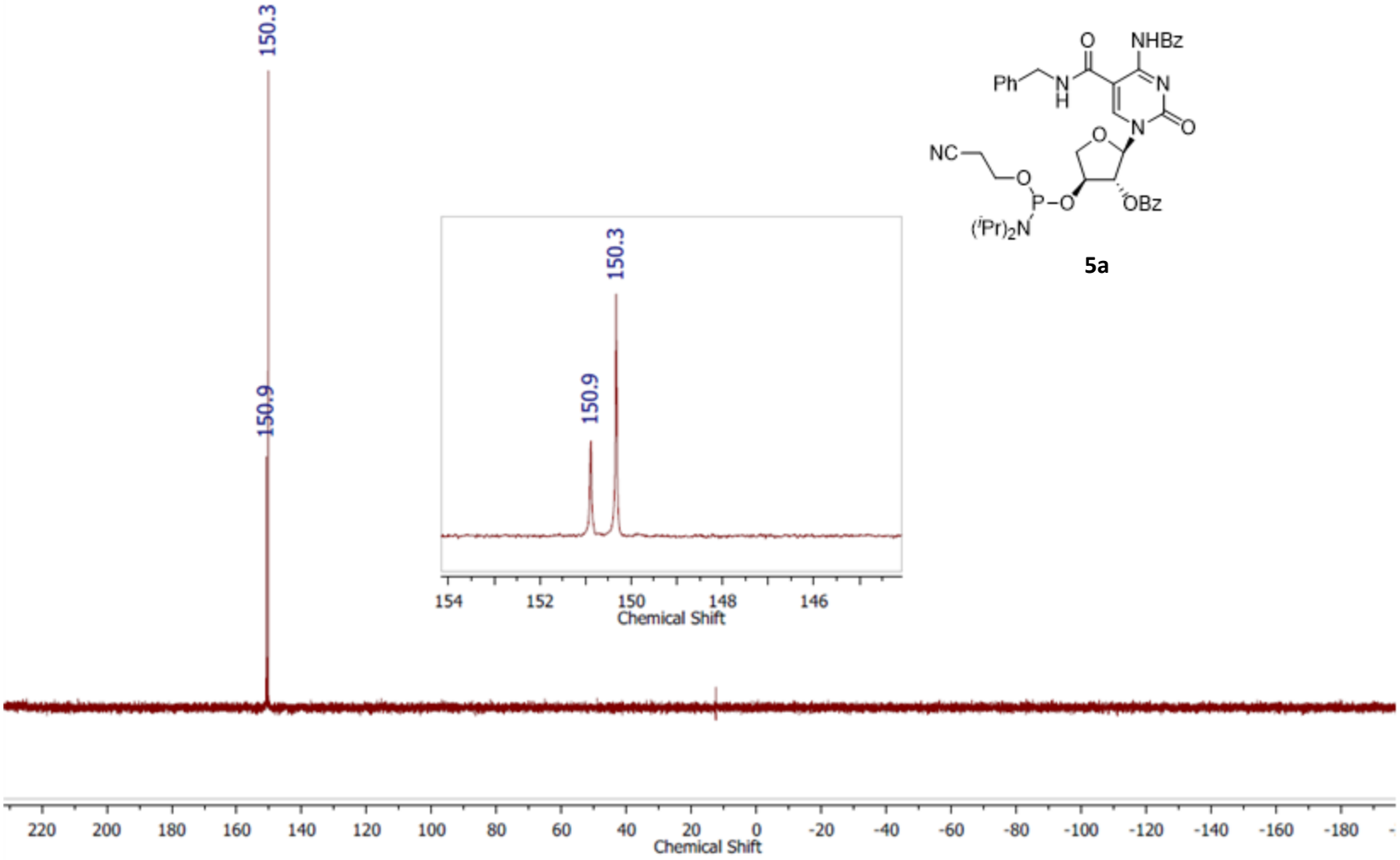
^{13}C NMR spectrum of compound **2a** (101 MHz, CD_3OD)



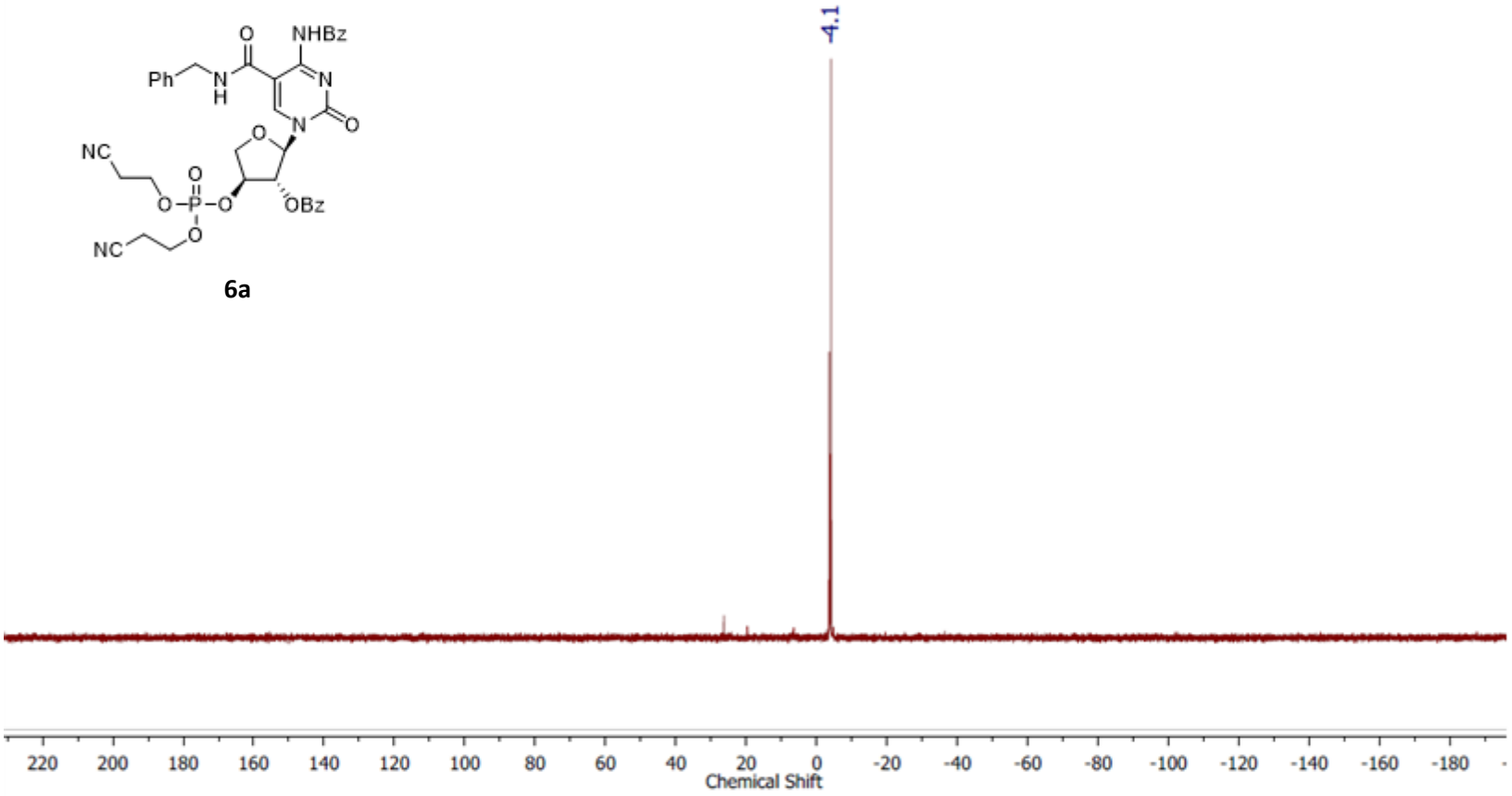
^{1}H NMR spectrum of compound **3a** (400 MHz, CD_3OD)



^{13}C NMR spectrum of compound **3a** (101 MHz, CD_3OD)

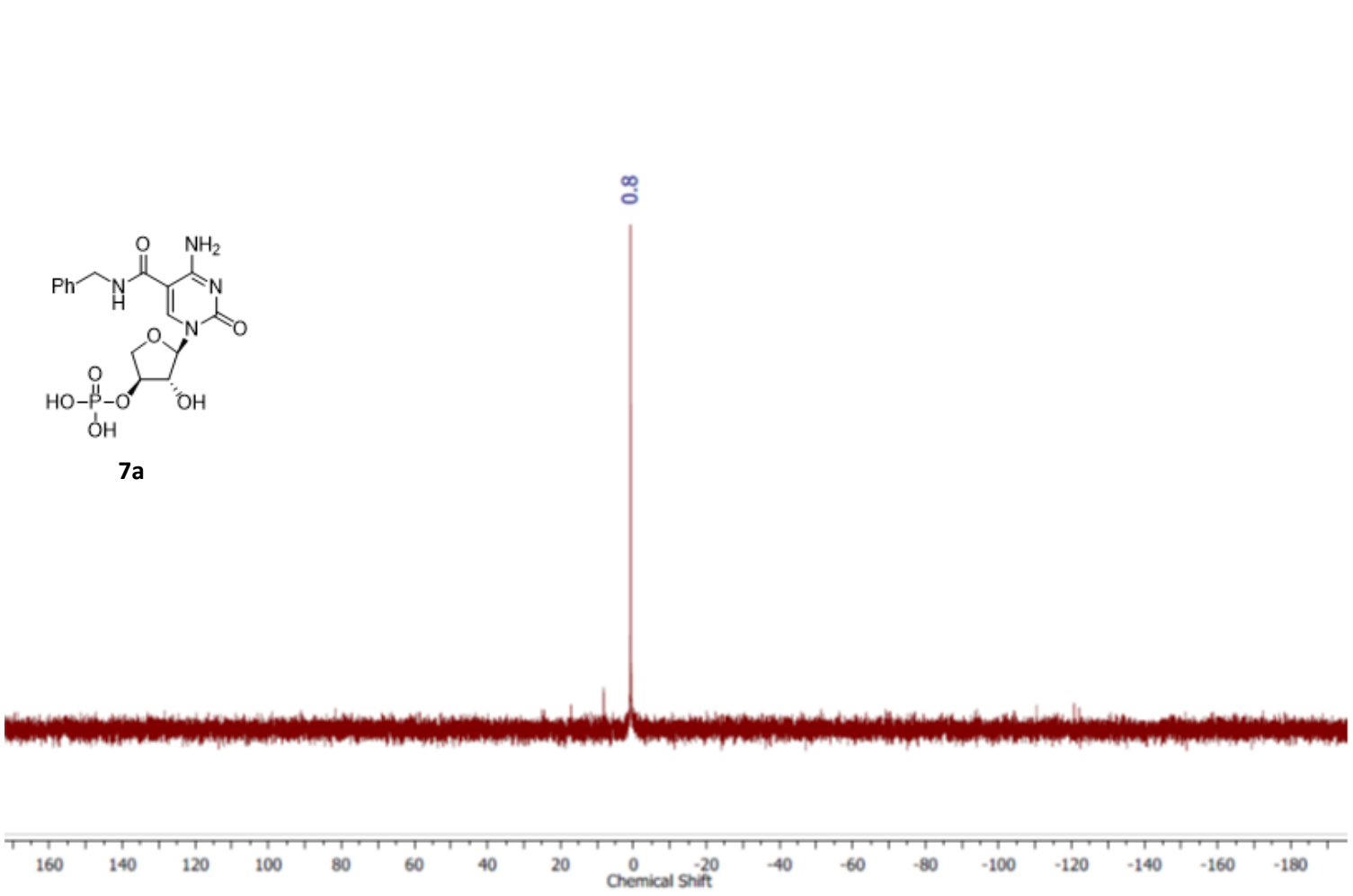


^{31}P NMR spectrum of compound **5a** (161 MHz, CD_3CN)

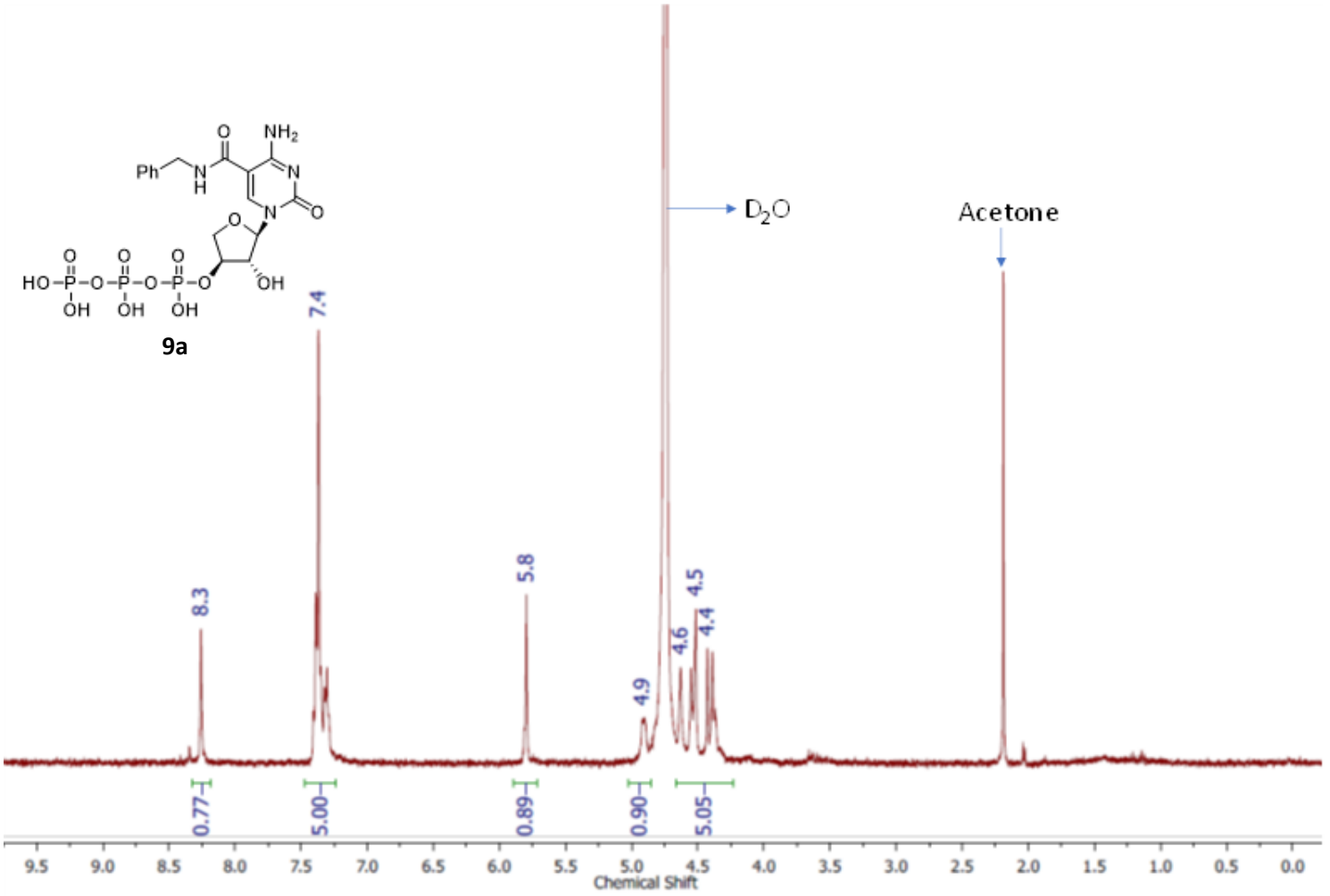


6a

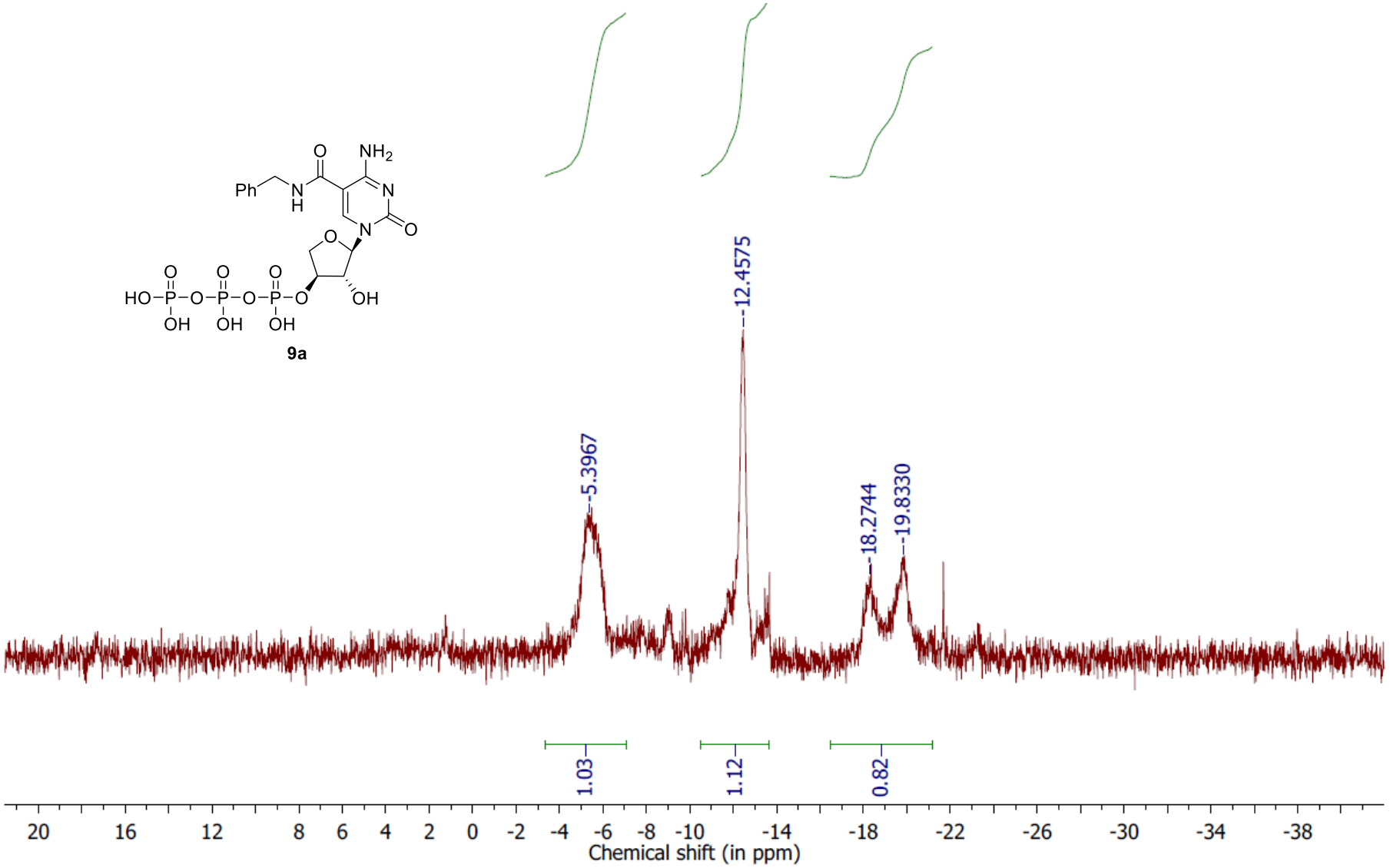
^{31}P NMR spectrum of compound **6a** (161 MHz, $CDCl_3$)



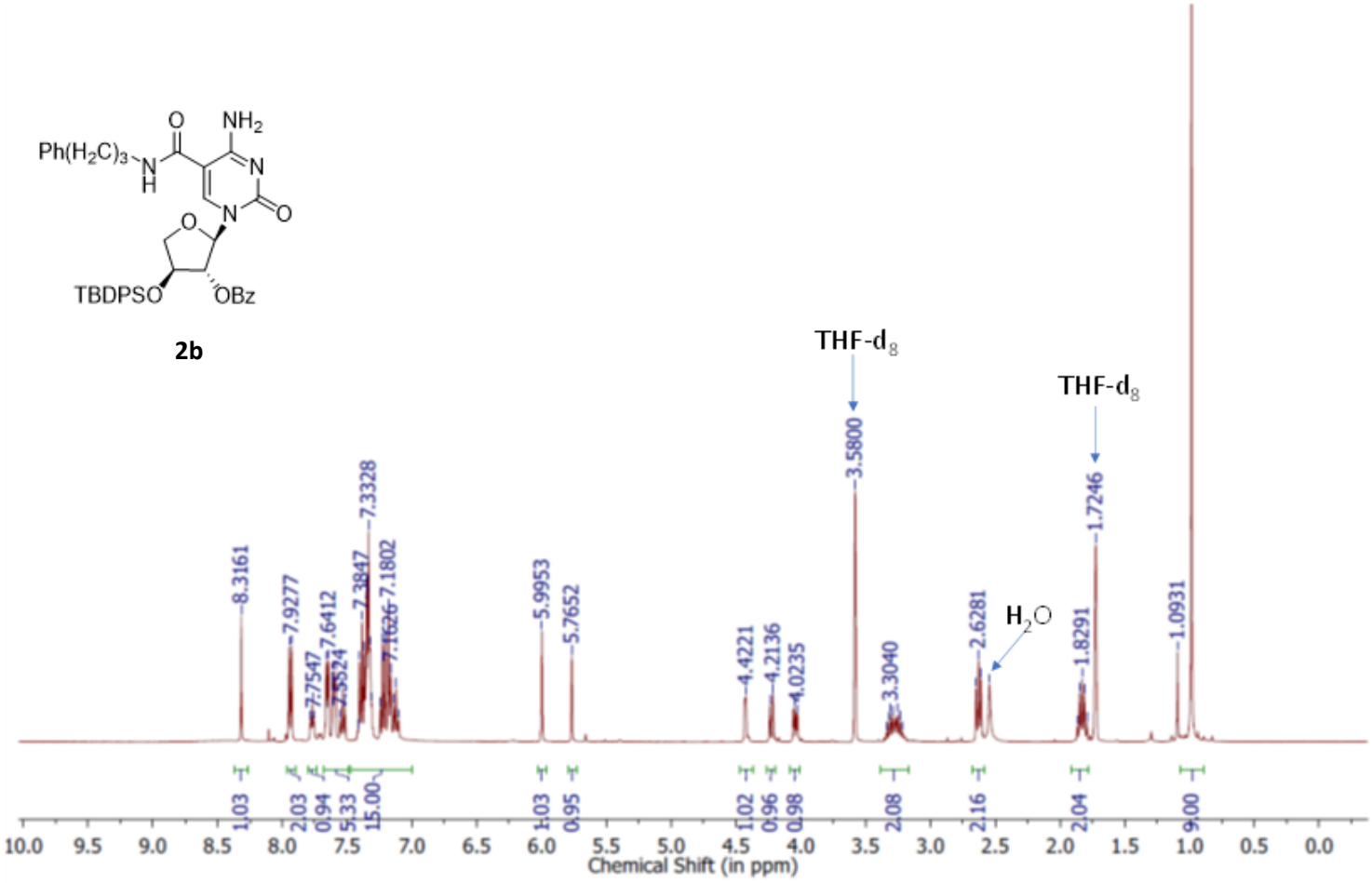
^{31}P NMR spectrum of compound **7a** (161 MHz, D_2O)



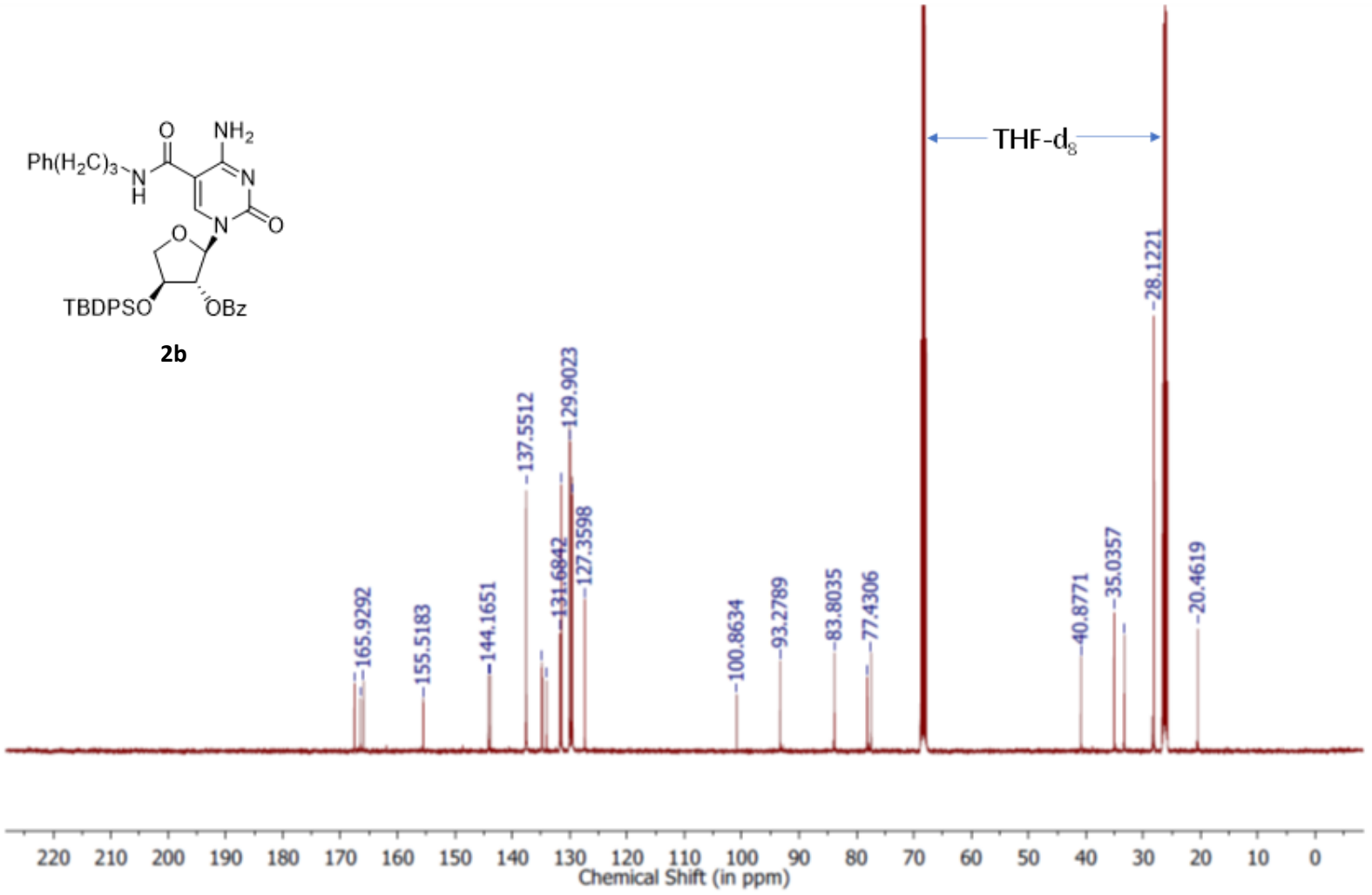
^{1}H NMR spectrum of compound **9a** (400 MHz, D_2O)



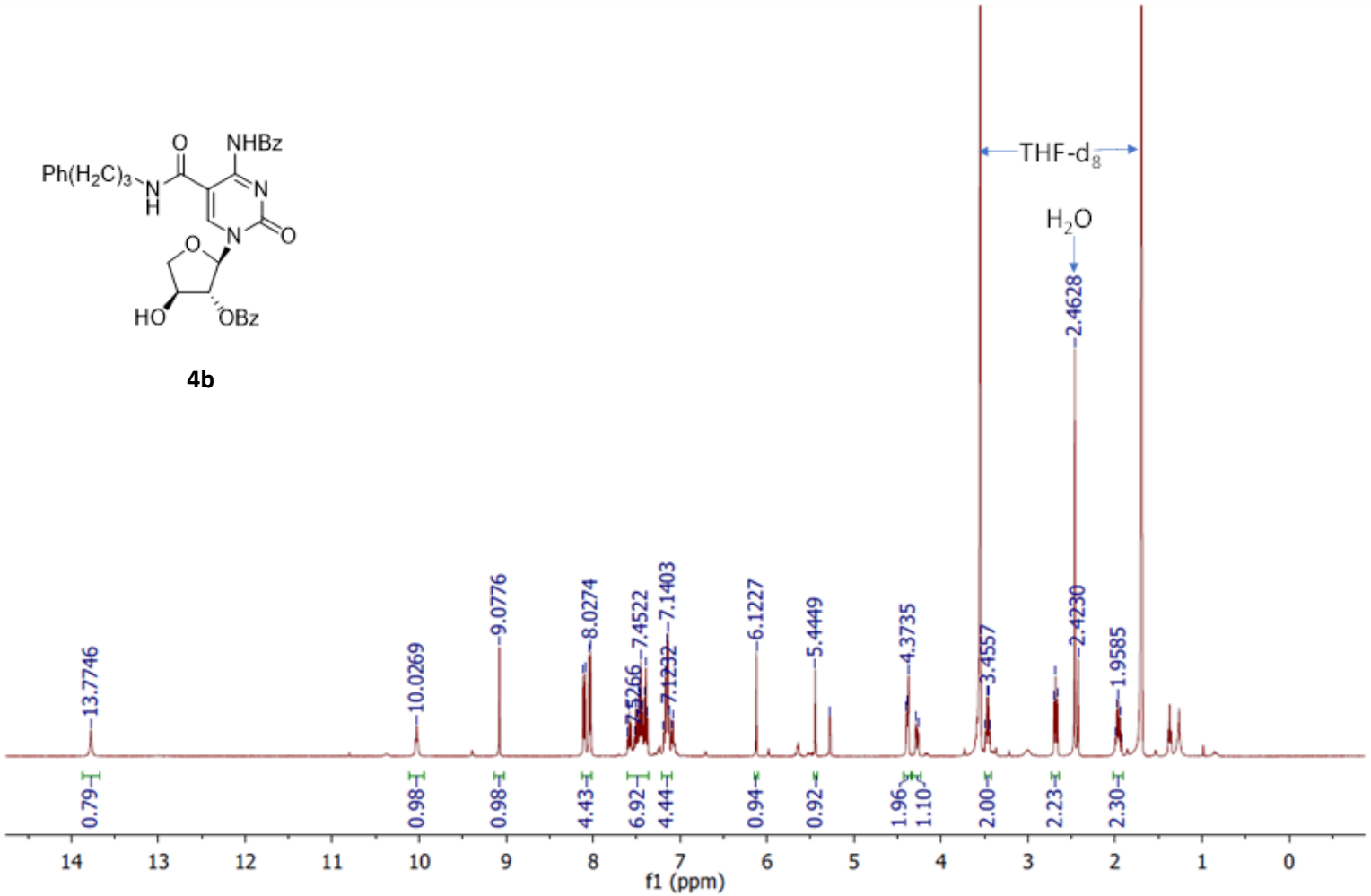
^{31}P NMR spectrum of compound **9a** (161 MHz, D_2O)



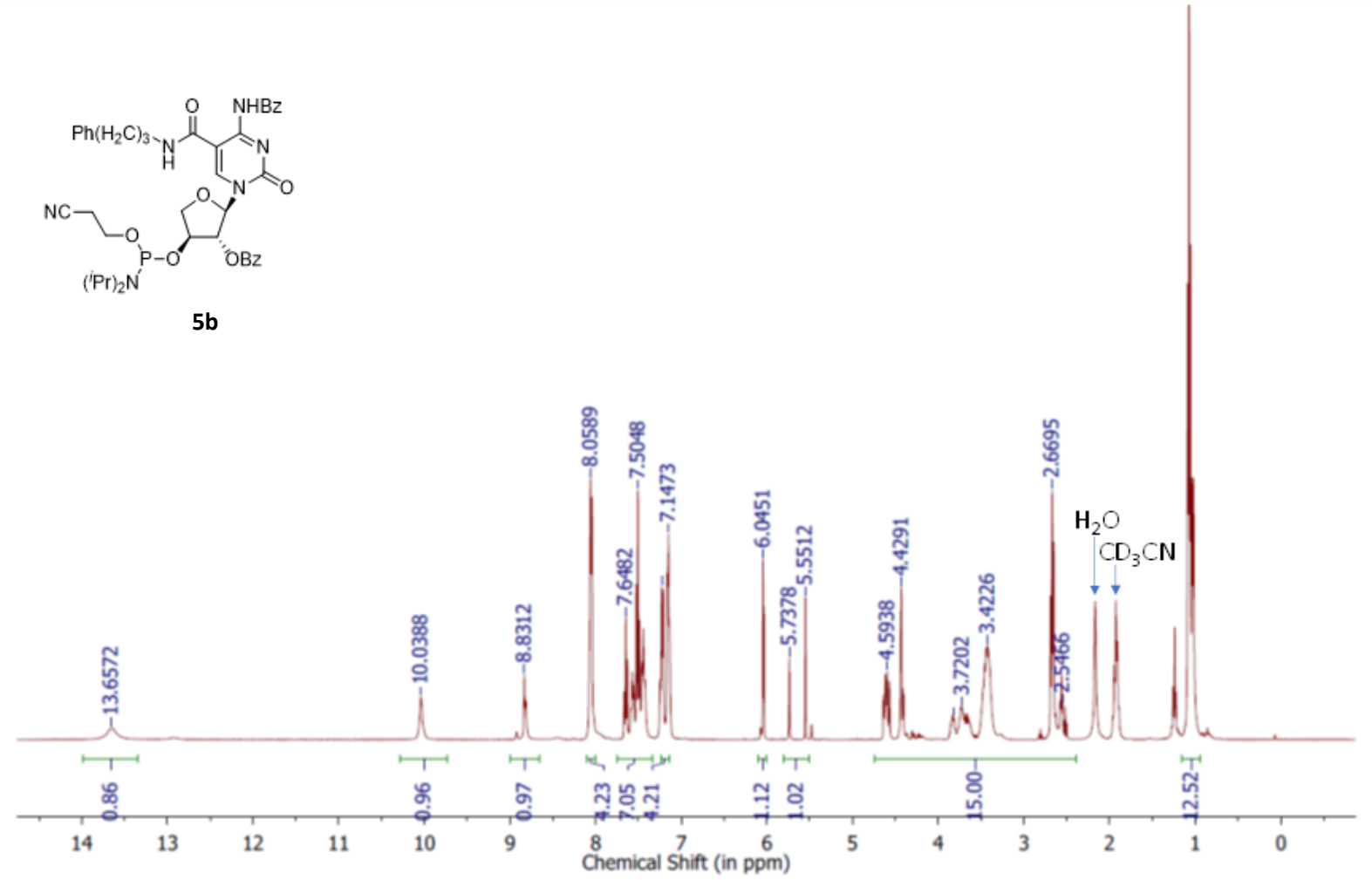
[1]H NMR spectrum of compound **2b** (400 MHz, THF-d8)



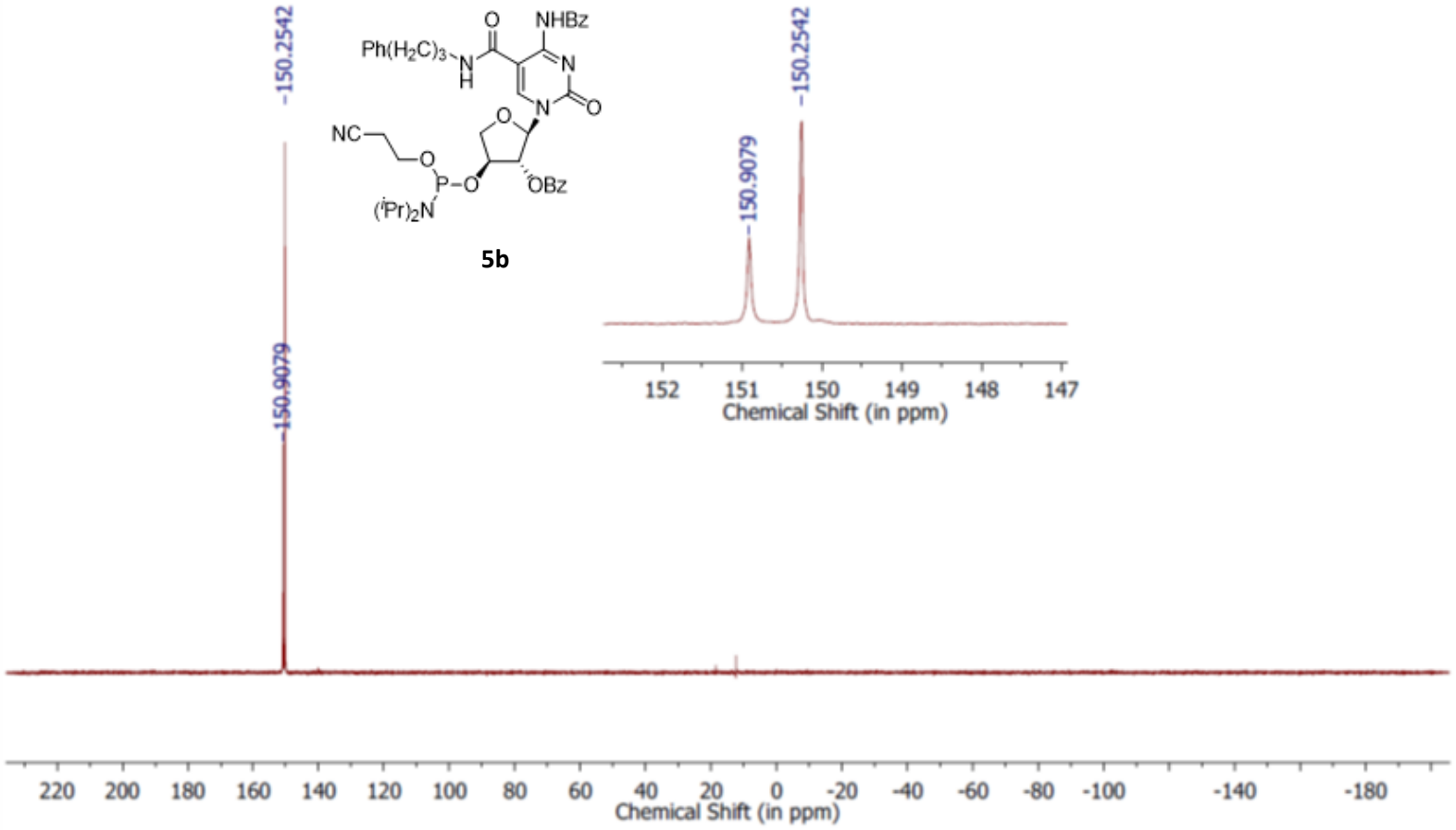
^{13}C NMR spectrum of compound **2b** (101 MHz, THF-d8)



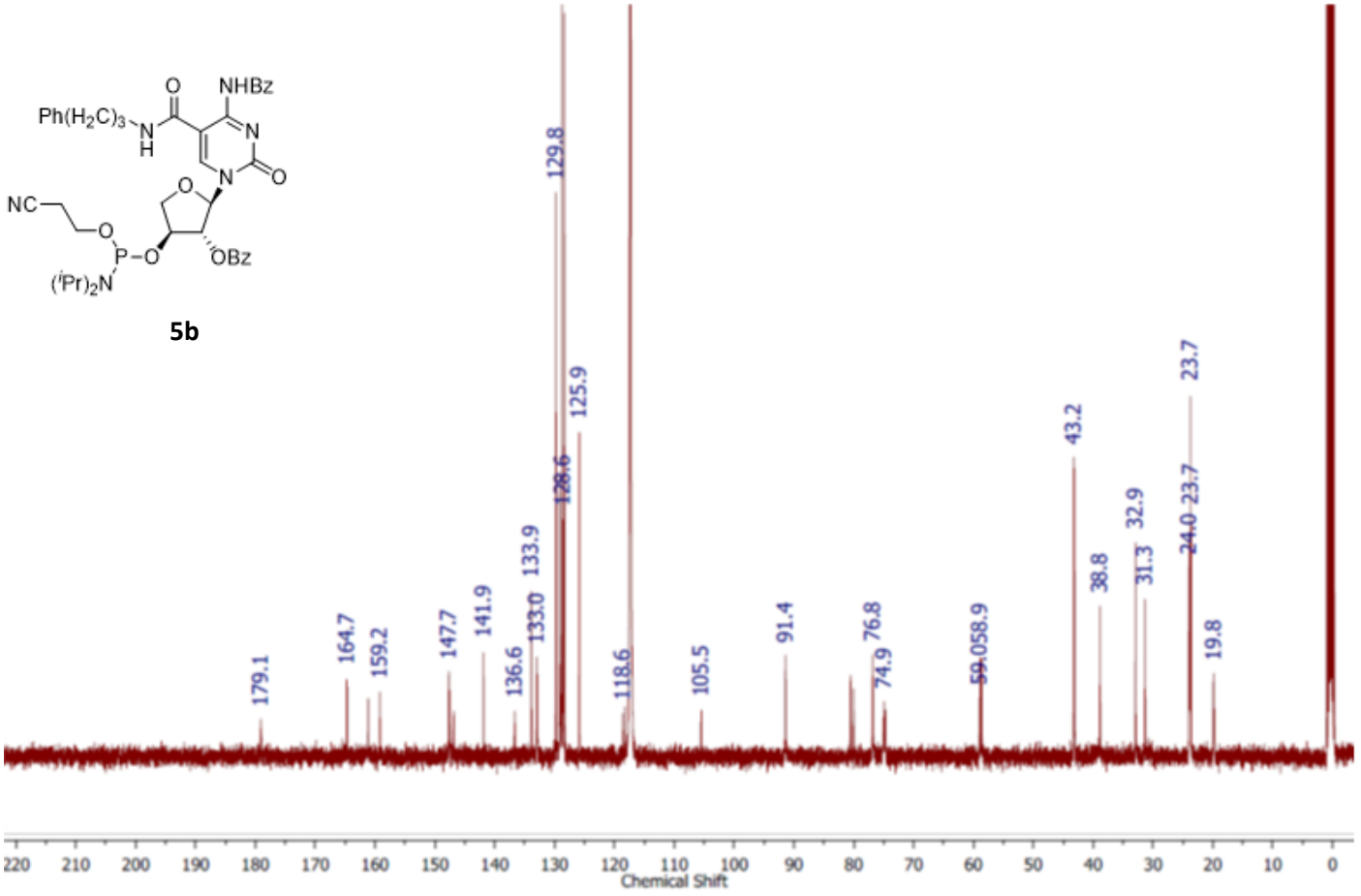
^{1}H NMR spectrum of compound **4b** (400 MHz, THF-d8)



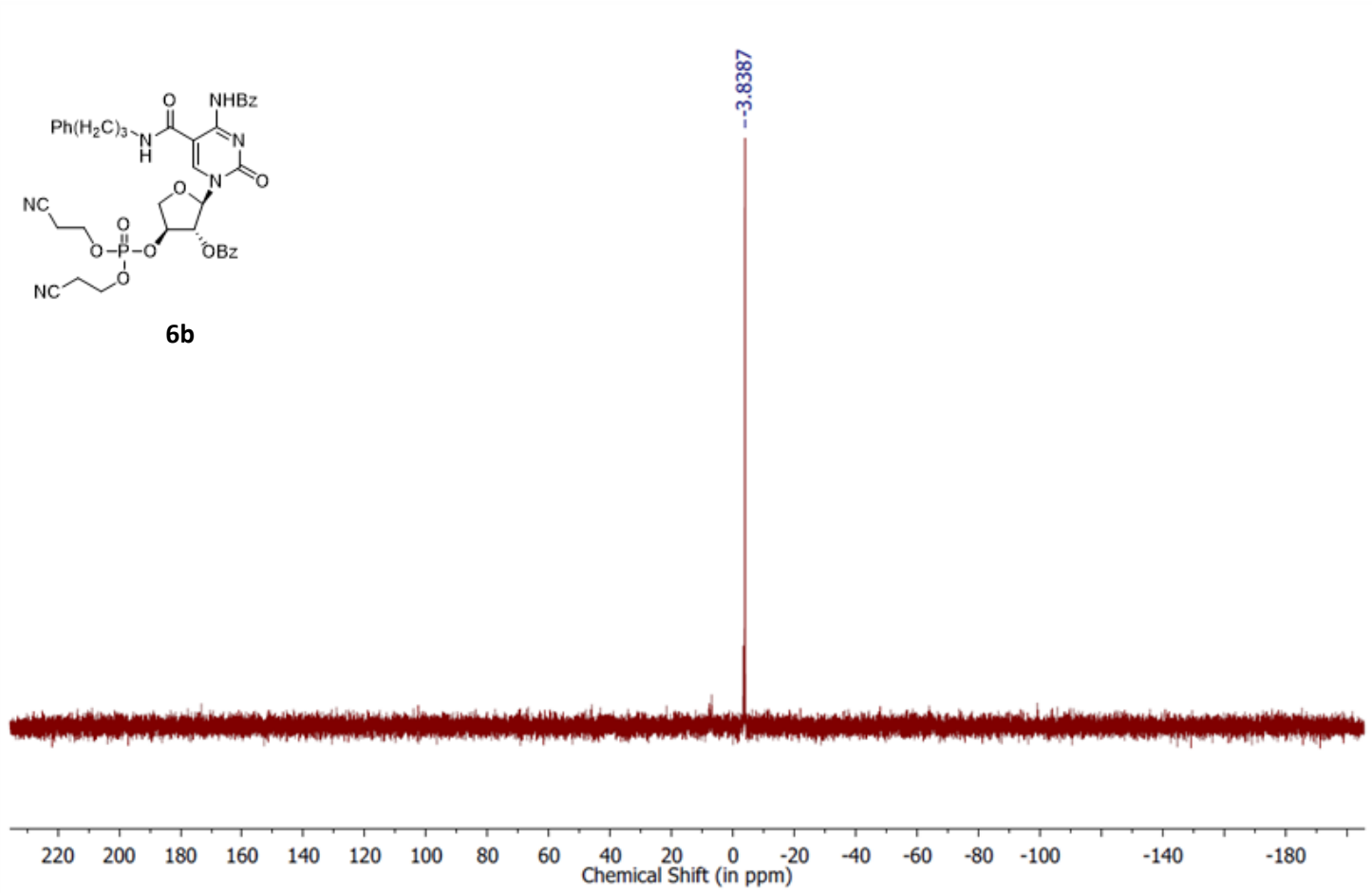
^{1}H NMR spectrum of compound **5b** (400 MHz, CD_3CN)



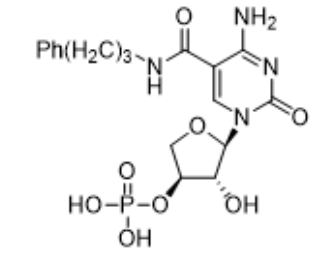
^{31}P NMR spectrum of compound **5b** (161 MHz, CD_3CN)



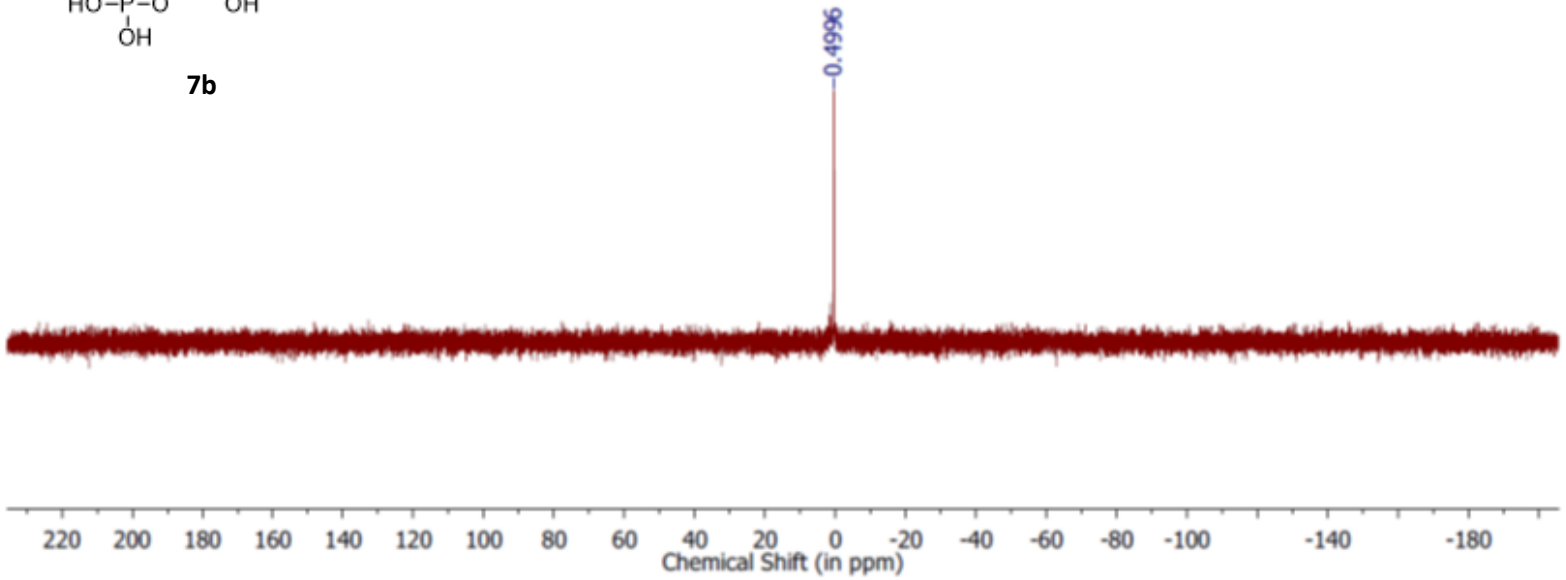
^{13}C NMR spectrum of compound **5b** (101 MHz, THF-d8)



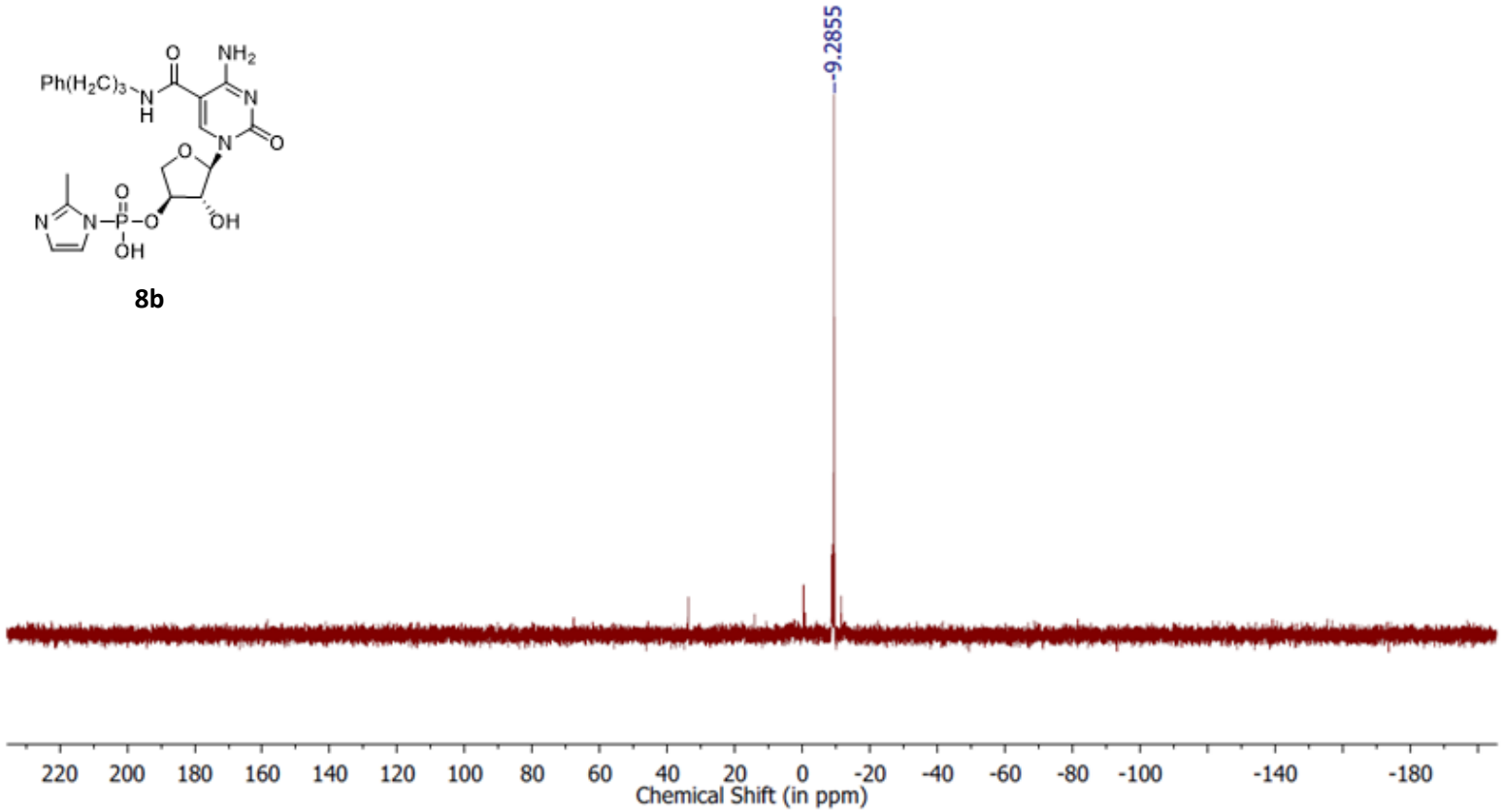
^{31}P NMR spectrum of compound **6b** (161 MHz, $CDCl_3$)



7b

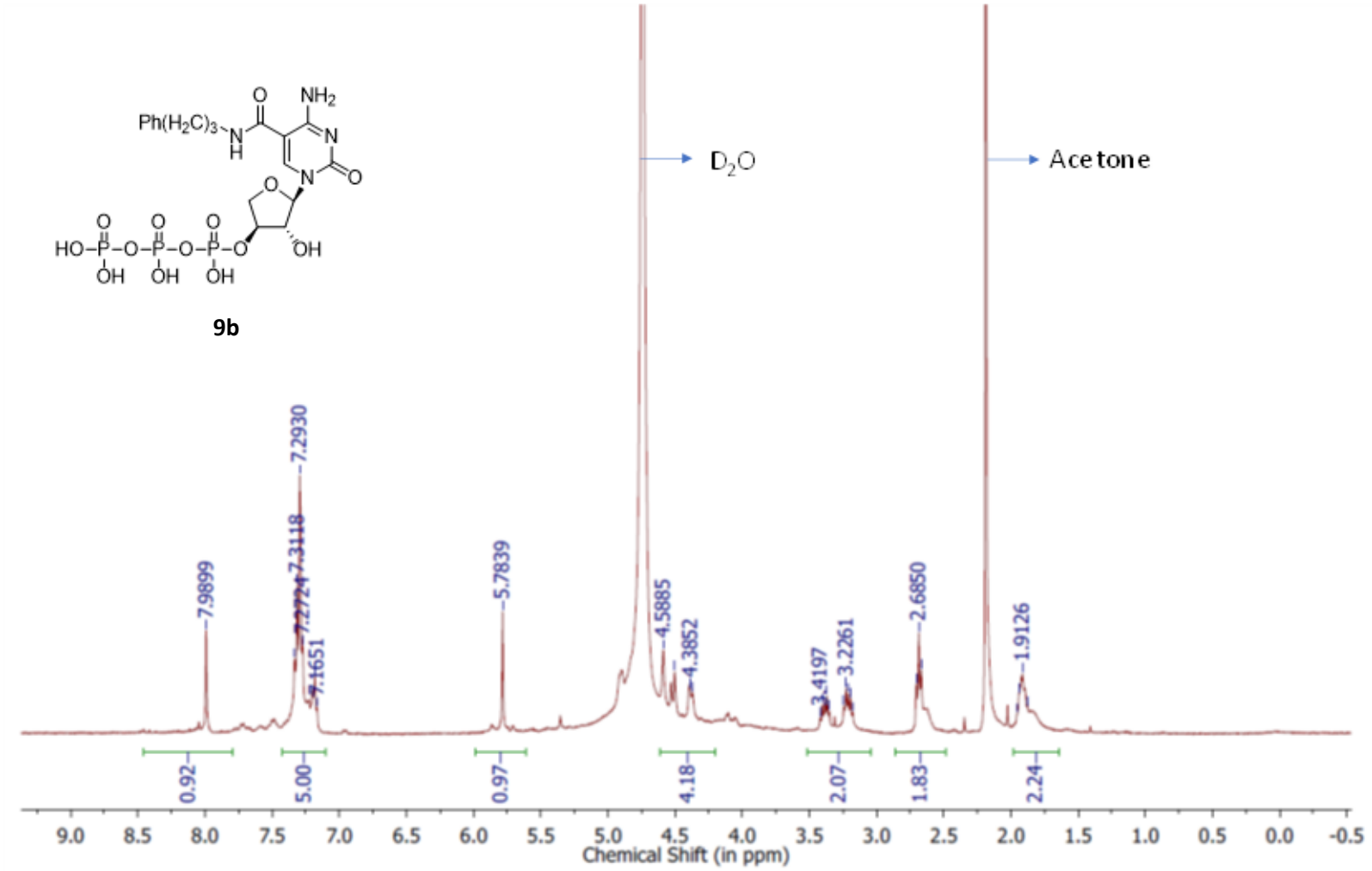


^{31}P NMR spectrum of compound **7b** (161 MHz, D_2O)

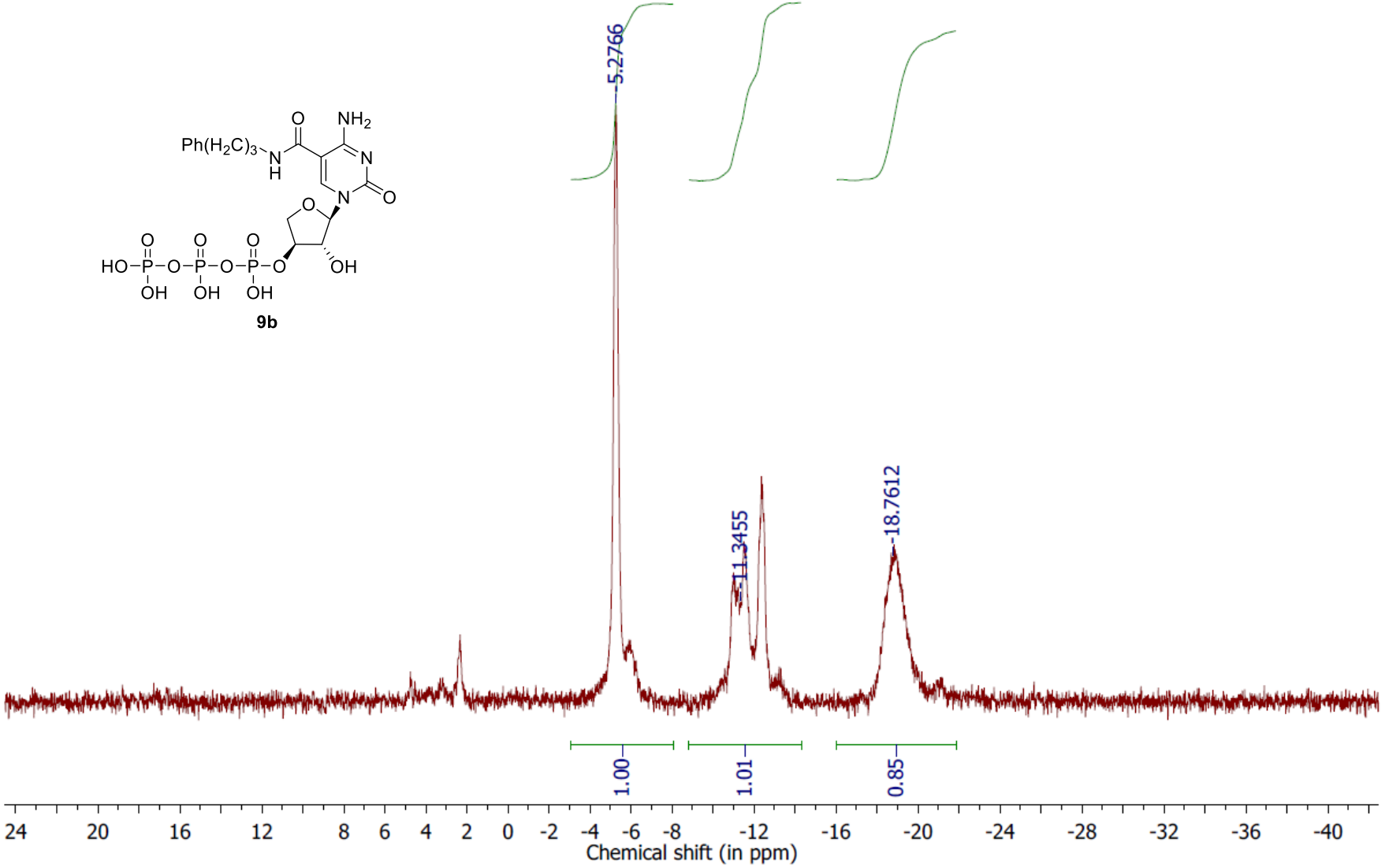


8b

^{31}P NMR spectrum of compound **8b** (161 MHz, D_2O)



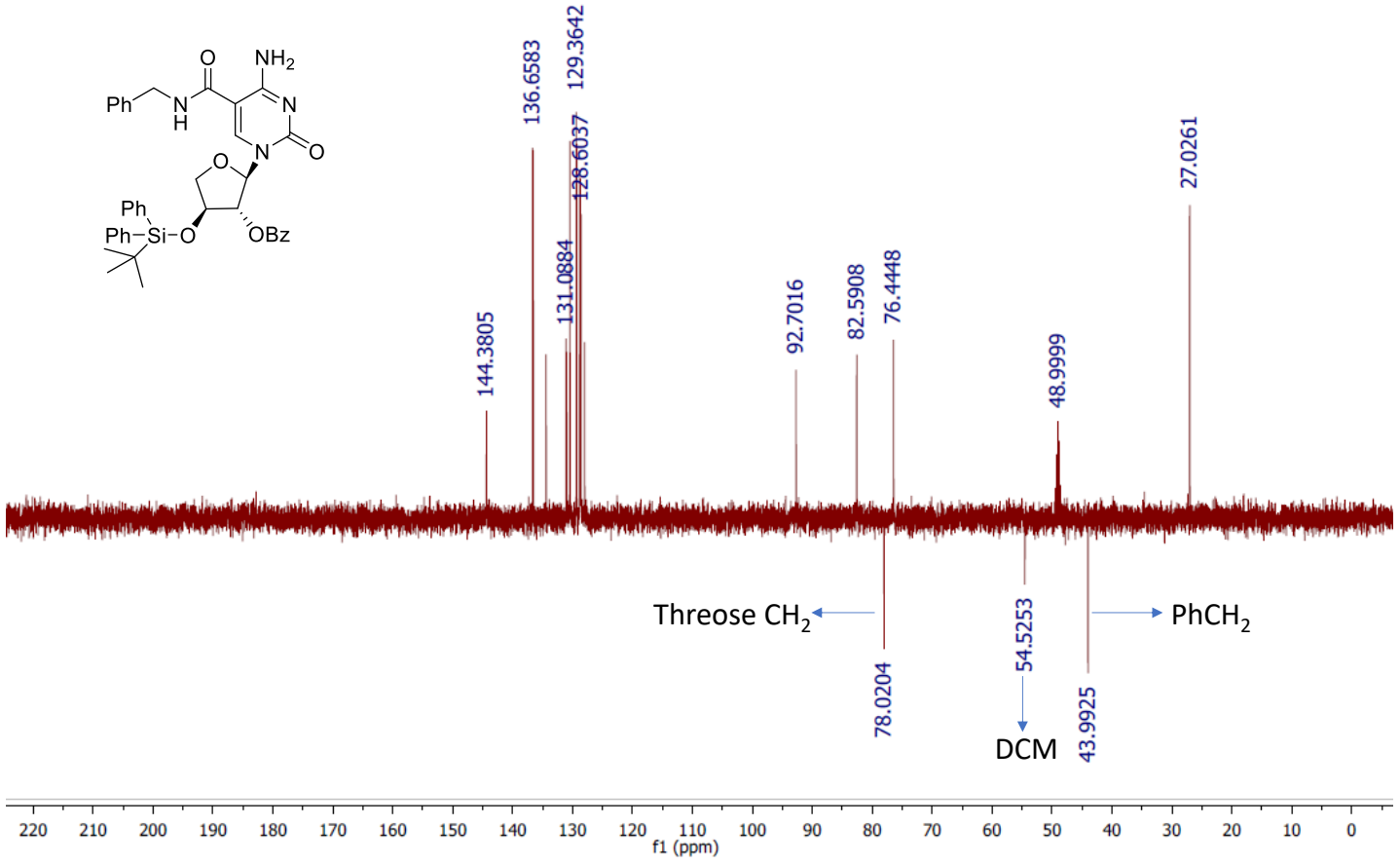
^{1}H NMR spectrum of compound **9b** (400 MHz, D_2O)



9b

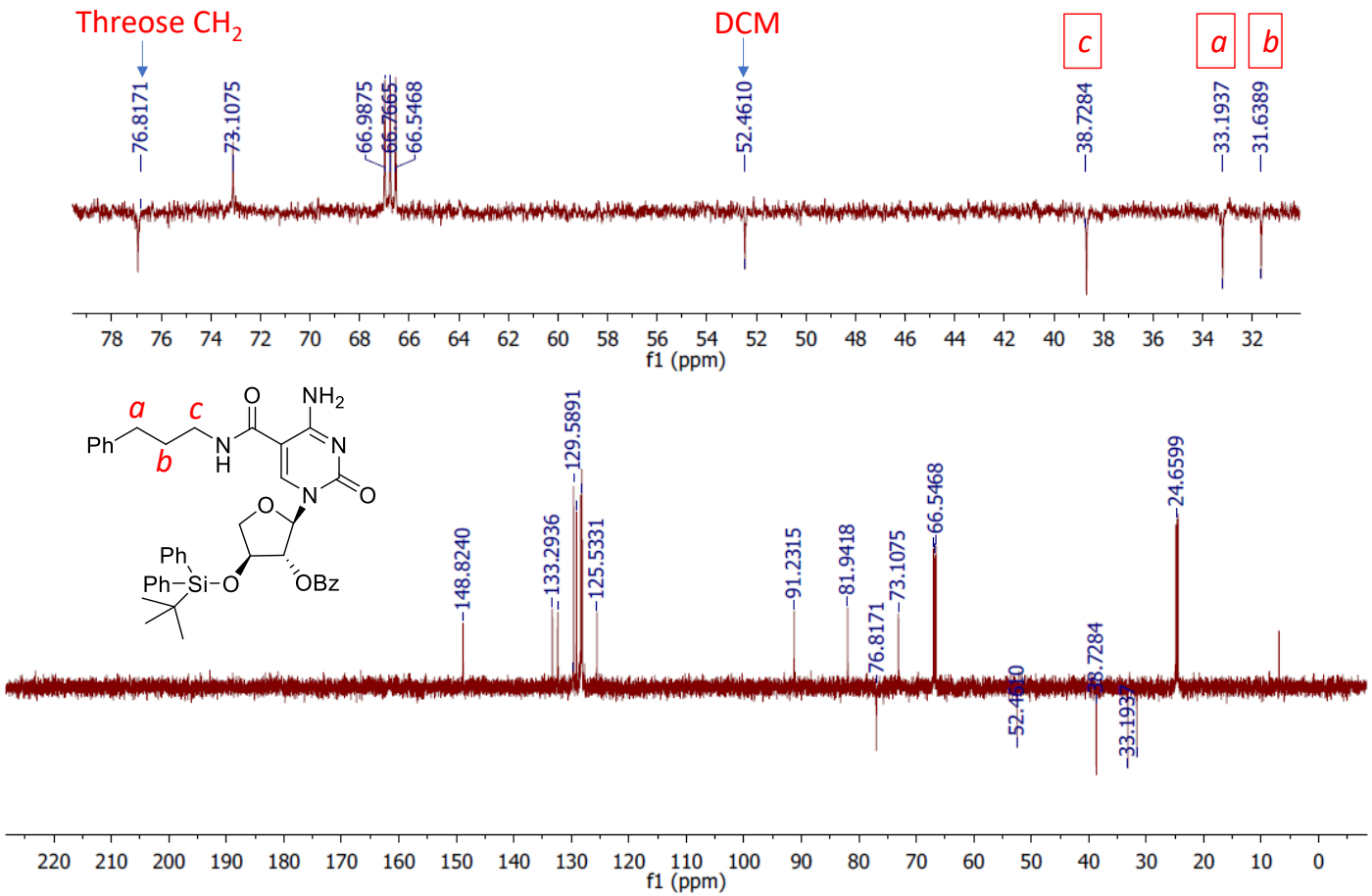
^{31}P NMR spectrum of compound **9b** (161 MHz, D_2O)

Successful C5 carboxyamidation of the nucleoside with the desired arylamine was further confirmed by running DEPT experiment where inverted signals for respective CH_2 groups could be observed. For benzylamine modified nucleoside 2 inverted signals for CH_2 (one from threose sugar and the other from benzylamine) could be observed. (third CH_2 signal is from trace solvent DCM)



DEPT-135 ^{13}C NMR spectrum of compound **2a** (101 MHz, CD_3OD)

For 3-Phenylpropylamine modified nucleoside 4 inverted signals for CH_2 (one from threose sugar and the other three from 3-Phenylpropylamine) could be observed. (fourth CH_2 signal is from trace solvent DCM)



DEPT-135 ^{13}C NMR spectrum of compound **2b** (101 MHz, DMSO-d6)