

Synthesis of Nucleotidylated Polio Virus VPg Proteins

Gerbrand J. van der Heden van Noort, Herman S. Overkleeft, Gijsbert A. van der Marel, Dmitri V. Filippov*

Leiden Institute of Chemistry, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands.

filippov@chem.leidenuniv.nl

Table of contents

General information	S1
Preparation of compound 5	S1
Preparation of compound 6	S2
Preparation of compound 8	S2
Preparation of compound 9	S3
Preparation of compound 11	S3/S4
Preparation of compound 12	S4/S5
Preparation of compound 14	S5
Preparation of compound 15	S5/S6
Preparation of compound 21	S6
Preparation of compound 22	S7
NMR spectra of compound 4	S8-S9
NMR spectra of compound 5	S10-S11
NMR spectra of compound 6	S12-S13
NMR spectra of compound 7	S14-S16
NMR spectra of compound 8	S17-S19
NMR spectra of compound 9	S20-S22
NMR spectra of compound 10	S23-S25
NMR spectra of compound 11	S26-S28
NMR spectra of compound 12	S29-S34
NMR spectra of compound 13	S35-S37
NMR spectra of compound 14	S38-S40
NMR spectra of compound 15	S41-S43
Spectra of VPgpC 20	S44-S47
Spectra of VPgpA 21	S48-S51
Spectra of VPgpG 22	S52-S55

General Information

All chemicals were used as received and reactions were carried out dry, under an argon atmosphere, at ambient temperature, unless stated otherwise. Column chromatography was performed on silica gel 60 (0.040-0.063 mm). TLC analysis was conducted on HPTLC aluminium sheets (silica gel 60, F245). Compounds were visualized by UV absorption (245 nm), by spraying with a solution of $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}\cdot 4\text{H}_2\text{O}$ 25 g/l and $(\text{NH}_4)_4\text{Ce}(\text{SO}_4)_4\cdot 2\text{H}_2\text{O}$ 10 g/l, in 10% aqueous H_2SO_4 followed by charring at +/- 140 °C. Optical rotation measurements ($[\alpha]_D^{20}$) were performed on an automated polarimeter (Sodium D-line, $\lambda = 589$ nm) with a concentration of 10 mg/ml ($c = 1$). Infrared spectra were recorded on a FT-IR spectrometer. ^{31}P , ^1H , and ^{13}C APT NMR spectra were recorded on a 400 MHz NMR device (161.7, 400 and 100 MHz respectively) or a 600 MHz NMR device. NMR spectra were recorded in CDCl_3 with chemical shift (δ) relative to tetramethylsilane, unless stated otherwise. COSY-GR and HSQC measurements were used to assign signals. High resolution mass spectra were recorded by direct injection (2 μl of a 2 μM solution in water/acetonitrile; 50/50; v/v and 0.1 % formic acid) on a mass spectrometer (LTQ Orbitrap) equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10, capillary temperature 250 °C) with resolution $R = 60000$ at m/z 400 (mass range $m/z = 150$ -2000) and dioctylphthalate ($m/z = 391.28428$) as a lock mass. The high resolution mass spectrometer was calibrated prior to measurements with a calibration. TLC-MS (ESI) spectra of phosphoramidites were obtained by analysis of TLC-plates with a TLC-MS interface coupled to a mass instrument.

6-*N*-Benzoyl-2', 3'-di-*O*-isobutyryl- β -D-adenosine (5)

6-*N*-Benzoyl- β -D-adenosine **1** (1.8 g, 4.8 mmol) was dissolved in pyridine (25 mL) after coevaporation with pyridine. TBDMS-Cl (10 mmol, 1.5 g) was added and the reaction mixture was stirred for 1 hour at room temperature. After this isobutyric anhydride (20 mmol, 3.7 mL) was added and the reaction mixture was stirred for 48 hours. The reaction mixture was concentrated *in vacuo* and taken up in DCM. After washing with 5% citric acid and H_2O the organic phase was dried (MgSO_4) and concentrated *in vacuo*. The residue was taken up in 4:1 MeCN/ H_2O (50 mL, v/v) and *p*-TsOH (5 mmol, 0.8 g) was added after which the reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was concentrated and purified with silica gel column chromatography EA/PE (0/100 – 45/55) to afford the title compound as a white foam (1.9 g, 3.9 mmol, 78%). ^{13}C -NMR (100 MHz, CDCl_3); δ 175.6, 175.1, 164.6, 152.3, 150.9, 150.2, 142.2, 133.3, 132.8 (Arom. Bz), 128.8 (Arom. Bz), 127.9 (Arom. Bz), 124.4, 88.4 (C1'), 86.3 (C4'), 72.8 (C2'), 72.2 (C3'), 62.4 (C5'), 33.8, 33.5 (CH *i*Bu), 18.89, 18.75, 18.67, 18.59 (CH₃ *i*Bu). ^1H -NMR (400 MHz, CDCl_3); δ 9.26 (bs, 1H), 8.76 (s, 1H), 8.14 (s, 1H), 8.03, 8.01 (d, 2H, Bz), 7.62 - 7.49 (2x t, 3H, Bz), 6.14, 6.12 (d, 1H, H1'), 6.03 – 5.99 (dd, 1H, H2'), 5.90 (bs, 1H), 5.71 - 5.70 (dd, 1H, H3'), 4.35, 4.34 (d, 1H, H4'), 4.01 – 3.87 (ddd, 2H, H5'), 2.69 – 2.47 (2x m, 2H, CH *i*Bu), 1.29 – 1.23, 1.18 – 1.07 (2x q, 12H, CH₃ *i*Bu). IR: 3274, 2977, 1743, 1701, 1609, 1582, 1454, 1246, 1153, 1095. LCMS (10-90% B in 15 min), $R_t = 6.57$. ESI-MS: m/z 512.2 $[\text{M} + \text{H}]^+$. HRMS [$\text{C}_{25}\text{H}_{29}\text{N}_5\text{O}_7 + \text{H}]^+$; calcd. 512.2135, found 512.2139.

2-*N*-phenoxyacetyl-2', 3'-di-*O*-isobutyryl-β-D-guanosine (6)

2-*N*-phenoxyacetyl-β-D-guanosine **3** (4.17 g, 10 mmol) was dissolved in pyridine (50 mL) and TBDMSCl (3.0 g, 20.0 mmol) was added. When TLC analysis indicated complete silylation, isobutyric anhydride (6.6 mL, 40.0 mmol) was added and the reaction mixture was stirred overnight. After concentration the residue was redissolved in EtOAc (100 mL) and washed with sat. aq. NaHCO₃, 5% citric acid and water. The organic layer was dried (MgSO₄) and directly desilylated without further purification. For this the residue was dissolved in MeCN/water (100 mL, 4:1, v/v) and excess *p*-TsOH (2.4 g, 14.8 mmol) was added. After concentration the residue was taken up in EtOAc and washed with sat. aq. NaHCO₃ and water. Silica gel column chromatography with a gradient of EtOAc in toluene (0/100 – 70/30) yielded the title compound as an off-white powder (4.5 g, 8.5 mmol, 85 %). ¹³C-NMR (100 MHz, CDCl₃); δ 175.7, 175.2, 169.9, 156.4, 155.0, 147.0, 146.5, 138.9, 129.7, 122.6, 114.7 (Arom. Pac), 87.5 (C1'), 85.0 (C4'), 72.6 (C2'), 71.4 (C3'), 66.5 (CH₂ Pac), 61.9 (C5'), 33.7, 33.5 (CH *i*Bu), 18.8, 18.7, 18.6, 18.5 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); δ 9.7 (s, 1H, NH), 7.92 (H8), 7.35, 7.33, 7.31 (t, 2H, Arom. Pac), 7.07, 7.05, 7.03, 7.01, (m, 3H, Arom. Pac), 7.06, 7.04 (d, 2H, Arom. Pac), 5.98, 5.97 (d, 1H, H1'), 5.94, 5.92, 5.91 (t, 1H, H2'), 5.73, 5.73, 5.72, 5.71 (dd, 1H, H3'), 4.77 (s, 2H, CH₂ Pac), 4.31, 4.30 (d, 1H, H4'), 4.00, 3.98, 3.89, 3.86 (m, 2H, H5'), 2.65 - 2.60, 2.53 - 2.50 (2 x m, 2H, CH *i*Bu), 1.27 - 1.18 (d, 6H, CH₃ *i*Bu), 1.13 - 1.05 (m, 6H, CH₃ *i*Bu). IR: 2974, 1684, 1608, 1558, 1471, 1407, 1188, 1150, 1096. LCMS (10-90% B in 15 min), Rt = 7.94. ESI-MS: *m/z* 558.0 [M+H]⁺. HRMS: [C₂₆H₃₁N₅O₉ + H]⁺ calcd. 558.2194, found 558.2194.

2-cyanoethoxy-*N,N'*-diisopropylamino-(6-*N*-benzoyl-2',3'-di-*O*-isobutyryl-β-D-adenyl-5'-yl) phosphine (8)

To a stirred solution of 6-*N*-benzoyl-2', 3'-di-*O*-isobutyryl-β-D-adenosine **5** (1.3 g, 2.6 mmol), coevaporated with MeCN, in DCM (15 mL), containing TEA (1.2 mL, 8.5 mmol), 2-cyanoethoxy-*N,N'*-diisopropylaminochlorophosphine (0.7 mL, 3.2 mmol) was added under an argon atmosphere. The reaction mixture was stirred for 45 min at room temperature. After ³¹P NMR showed complete consumption of the 2-cyanoethoxy-*N,N'*-diisopropylaminochlorophosphine (400 MHz, CDCl₃; δ 180.4) and formation of the product (400 MHz, CDCl₃; δ 150.0, 149.5), 30 mL DCM was added and the reaction mixture was washed with water and 5% NaHCO₃. The organic phase was dried (MgSO₄) and concentrated. The residue was applied on a silica gel column (neutralized with TEA) and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as an off-white foam (1.1 g, 1.6 mmol, 60%). ¹³C-NMR (100 MHz, CDCl₃); δ 175.91, 175.86, 175.4, 175.3, 164.6, 154.9, 152.0, 151.9, 149.6, 141.4, 141.3, 133.7, 132.8, 128.9, 128.8, 127.8 (Arom. Bz), 122.9, 117.7, 117.6, 85.4, 85.0 (C1'), 83.9, 83.8, 83.6, 83.5 (C4'), 74.0, 73.9 (C2'), 72.0, 71.7 (C3'), 63.2, 63.0, 63.0, 62.8 (C5'), 58.8, 58.6, 58.4 (CH₂ CNEO), 43.2, 43.1 (CH *i*Pr), 33.9, 33.8, 33.59, 33.57 (CH *i*Bu), 24.7, 24.6 (CH₃ *i*Pr), 20.40, 20.36, 20.32, 20.29 (CH₂ CNEO), 18.9, 18.8, 18.7, 18.6 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); 9.26 (bs, 1H, NH), 8.83, 8.82 (2x s, 1H, H2), 8.55, 8.52 (2x s, 1H, H8), 8.05, 8.03 (d, 2H, Bz), 7.61 - 7.51 (2x m, 3H, Bz), 6.46, 6.1 (2x d, 1H, H1'), 5.86 - 5.78 (2x dd, 1H, H2'), 5.65 - 5.762 (2x dd, 1H, H3'), 4.41 - 4.40 (m, 1H, H4'), 3.97 - 3.84 (m, 4H, H5', CH₂ CNEO), 3.68 - 3.62 (m, 2H, CH *i*Pr), 2.76 - 2.64, 2.52 - 2.49 (2x m, 2H, CH *i*Bu), 1.25 - 1.18, 1.11 - 1.05 (2x m, 24H, CH₃ *i*Bu, CH₃ *i*Pr). ³¹P-NMR (161 MHz, CDCl₃); δ 150.0, 149.5. IR: 2968, 1746, 1702, 1610, 1582, 1508, 1458, 1245, 1184, 1153. LCMS (10-90% B in 15 min), Rt = 7.37. ESI-MS: *m/z* 629.1 [M + H]⁺ of the corresponding H-Phosphonate. TLCMS 712.6 [M + H]⁺. HRMS: [C₃₅H₄₈N₇O₁₀P + H]⁺; 712.3218 calcd., 712.3229 found.

2-cyanoethoxy-*N,N'*-diisopropylamino-(2-*N*-phenoxyacetyl-2',3'-di-*O*-isobutyryl- β -D-guanyl -5'-yl) phosphine (9)

To a stirred solution of 2-*N*-phenoxyacetyl-2', 3'-di-*O*-isobutyryl- β -D-guanosine **6** (0.94 g, 1.77 mmol), coevaporated with MeCN in DCM (10 mL), containing freshly distilled TEA (0.8 mL, 5.8 mmol), 2-cyanoethoxy-*N,N'*-diisopropylaminochlorophosphine (0.5 mL, 2.2 mmol) was added under an argon atmosphere. The reaction mixture was subsequently stirred for 45 min at room temperature. After ^{31}P NMR showed complete consumption of the 2-cyanoethoxy-*N,N'*-diisopropylaminochlorophosphine (200 MHz, CDCl_3 ; δ 180.4) and formation of the phosphoramidite (^{31}P NMR, 200 MHz: δ 150.0, 149.1), 20 mL DCM was added and the reaction mixture was washed with water and 5% NaHCO_3 . The organic phase was dried (MgSO_4) and concentrated. The residue was applied on a silica gel column (neutralized with TEA) and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as an off-white foam (1.1 g, 1.4 mmol, 78%). ^{13}C -NMR (100 MHz, CDCl_3); δ 175.82, 175.76, 175.2, 175.1, 169.6, 169.45, 156.5, 156.4, 155.4, 148.0, 147.9, 146.4, 138.1, 137.7, 130.02 (Arom. Pac), 129.97 (Arom. Pac), 123.05 (Arom. Pac), 122.98 (Arom. Pac), 117.7, 114.9 (Arom. Pac), 85.7, 85.3 ($\text{C}1'$), 83.3, 83.2, 82.9, 82.8 ($\text{C}4'$), 73.44, 73.37 ($\text{C}2'$), 71.4, 71.3 ($\text{C}3'$), 67.0 (CH_2 Pac), 62.92, 62.85, 62.75, 62.67 ($\text{C}5'$), 58.7, 58.5, 58.3 (CH_2 CNEO), 43.26, 43.21, 43.13, 43.08 (CH *iPr*), 33.8, 33.6, (CH *iBu*), 24.71, 24.67, 24.64, 24.60 (CH_3 *iPr*), 20.39, 20.35, 20.31 20.28 (CH_2 CNEO), 18.91, 18.87, 18.8, 18.6 (CH_3 *iBu*). ^1H -NMR (400 MHz, CDCl_3); δ 8.14, 8.13 (2x s, 1H), 7.42 – 7.38 (t, 2H, Arom. Pac), 7.15 - 7.11 (t, 1H, Arom. Pac), 7.09 – 7.06 (d, 2H, Arom. Pac), 6.13 - 6.09 (m, 1H, $\text{H}1'$), 5.92 - 5.88 (m, 1H, $\text{H}2'$), 5.76 - 5.74 (m, 1H, $\text{H}3'$), 4.73 (s, 2H, CH_2 Pac), 4.39 – 3.38, (m, 3H, $\text{H}4'$), 3.99 – 3.80 (m, 4H, $5'$, CH_2 CNEO), 3.67 – 2.60 (m, 2H, CH *iPr*), 2.75 – 2.54 (3x m, 3H, CH *iBu*, CH_2 CNEO), 1.24 - 1.05 (2x m, 24H, CH_3 *iBu*, CH_3 *iPr*). ^{31}P -NMR (161 MHz, CDCl_3); δ 150.1, 149.7. IR: 2972, 1745, 1685, 1599, 1560, 1497, 1388, 1364, 1183, 1154. LCMS (10-90% B in 15 min), R_t = 7.93. ESI-MS: m/z 674.9 $[\text{M}+\text{H}]^+$ of the corresponding H-Phosphonate. TLCMS: 758.5 $[\text{M}+\text{H}]^+$. HRMS: $[\text{C}_{35}\text{H}_{48}\text{N}_7\text{O}_{10}\text{P} + \text{H}]^+$; 758.3273 calcd., 758.3280 found.

2-cyanoethoxy-(*N* ^{α} - Fmoc-tyrosin-4-yl allyl ester)-(6-*N*-benzoyl-2',3'-di-*O*-isobutyryl- β -D-adenyl-5'-yl) phosphate (11)

N ^{α} - Fmoc-tyrosine-allyl ester [Fmoc-Tyr-OAll] (0.26 g, 0.58 mmol) was added to amidite **8** (1.0 eq., 0.58 mmol) and the two compounds were coevaporated with MeCN and DCE and then dissolved in 1:1 MeCN/DCM (10 mL, v/v). Tetrazole (0.12 g, 1.74 mmol) was added and the reaction mixture was stirred for 45 min. until ^{31}P NMR (200 MHz: δ 135.36, 134.91) showed complete consumption of the phosphoramidite (200 MHz, CDCl_3 ; δ 150.31, 148.83). The reaction mixture was cooled to 0 °C and *m*-CPBA was added in portions until ^{31}P NMR (200 MHz: δ -6.17, -6.31) showed complete consumption of the phosphite. The reaction mixture was diluted in DCM and washed with 5% $\text{Na}_2\text{S}_2\text{O}_3$ and water. The organic phase was dried (MgSO_4) and concentrated *in vacuo*. The residue was applied to a silica gel column and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as a colourless gel (0.39 g, 0.37 mmol, 64%)

^{13}C -NMR (100 MHz, CDCl_3); δ 175.7, 175.6, 175.2, 175.1, 171.0, 170.8, 165.0, 164.9, 155.6, 152.7, 151.4, 151.1, 149.6, 149.5, 148.92, 148.85, 148.78, 143.5, 143.4, 143.2, 141.23, 141.16, 141.0, 140.9, 140.5, 139.6, 134.0, 133.6, 133.4, 133.2, 132.3, 131.1, 131.0, 130.8, 128.63, 128.55, 128.5, 128.1, 127.9, 127.8, 127.7, 127.6, 127.5 127.4, 127.3, 127.1, 127.0, 126.8, 126.7,

126.6, 124.7, 124.5, 124.0, 122.7, 122.0, 119.9, 119.6, 119.50, 119.45, 119.4, 119.3, 119.1, 116.14, 116.09, 85.6, 84.7 (C1'), 81.7, 81.6, 81.5, 81.4 (C4'), 73.5, 73.1 (C2'), 70.3, 70.2 (C3'), 67.2, 66.1, 66.0, 65.9, 65.4 (CH₂ Fmoc, CH₂ Allyl, CH₂ CNEO), 63.0, 62.90, 62.86, 62.8, 62.1 (C5'), 54.4, 54.1 (C_α), 47.1, 47.0 (CH Fmoc), 36.7, 36.3 (C_β), 33.6, 33.4 (CH *i*Bu), 19.5, 19.4 (CH₂ CNEO), 18.7, 18.64, 18.56, 18.51, 18.47 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); δ 9.30, 9.28 (d s, 1H), 8.79, 8.76 (d, 1H), 7.95 - 6.89 (m, 17H), 6.54, 6.52 (d, 1H), 6.38, 6.36, 6.33, 6.32 (2x d, 1H, H1'), 5.92 - 5.68 (m, 1H, CH Allyl, H3'), 5.54 - 5.51 (m, 1H, H2'), 5.28 - 5.15 (m, 2H, CH₂ Allyl), 4.65 - 4.60 (m, 1H, CH α), 4.56 - 4.48, 4.46 - 4.32 (m, 9H, C4', C5', CH₂ Fmoc, CH₂ Allyl, CH₂ CNEO), 4.27 - 3.98 (m, 1H, CH Fmoc), 3.12 - 3.03 (m, 2H, CH₂ β), 2.92 - 2.88 (m, 2H, CH₂ CNEO), 2.82 - 2.65, 2.57 - 2.47 (m, 2H, CH *i*Bu), 1.26 - 1.22, 1.13 - 1.04 (m, 12H, CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -6.97, 6.98. IR: 2976, 1738, 1699, 1608, 1583, 1505, 1456, 1247, 1189, 1155, 1029, 961. LCMS (50 - 90 % B in 15 min), Rt = 10.25. ESI-MS: *m/z* 1070.2 [M+H]⁺. HRMS: [C₅₅H₅₆N₇O₁₄P + H]⁺; 1070.3695 calcd., 1070.3706 found.

2-cyanoethoxy-(*N*^α- Fmoc-tyrosin-4-yl-allyl ester)-(2-*N*-phenoxyacetyl-2',3'-di-*O*-isobutyryl-β-D- guanyl -5'-yl) phosphate (12)

N^α- Fmoc-tyrosine-allyl ester (0.78 g, 1.77 mmol) was added to amidite **9** (1.0 eq., 1.77 mmol) and the two compounds were coevaporated with MeCN and DCE and then dissolved in 1:1 MeCN/DCM (10 mL, v/v). Tetrazole (295 mg, 4.2 mmol) was added and the reaction mixture was stirred for 45 min. until ³¹P NMR showed complete consumption of the phosphoramidite (³¹P NMR, 200 MHz: δ 134.18, 134.06). The reaction mixture was cooled to 0 °C and *m*-CPBA was added in portions until ³¹P NMR showed complete oxidation of the phosphite (³¹P NMR, 200 MHz: δ -6.56, -6.98). The reaction mixture was diluted in 20 mL DCM and washed with 5% Na₂S₂O₃ and water. The organic phase was dried (MgSO₄) and concentrated. The residue was applied to a silica gel using a gradient of EtOAc in PE (50/50 - 70/30) to yield the separate stereoisomers as white foams, 233 mg (mmol, 12 %) (Rt = 0.5, EtOAc) and 476 mg (mmol, 27 %) (Rt = 0.6, EtOAc) of which analytical samples were repurified using silica gel column chromatography using a gradient of 2 % MeOH in DCM, and a mix of both isomers (0.43 g, 0.38 mmol, 24 %) respectively. A total of 1.14 g, 1.0 mmol, 56 %. ¹³C-NMR (100 MHz, CDCl₃), isomer A (Rt = 0.6), 95 % pure, 93 % ee.; δ 175.6, 175.1, 170.8, 170.1, 156.9, 155.5, 155.0, 148.4, 148.3, 147.4, 146.5, 143.6, 140.9, 134.0, 131.2, 130.3, 129.4, 127.4, 126.8, 124.9, 122.6, 121.9, 119.8, 119.7, 119.6, 118.8, 116.1, 114.6, 87.0 (C1'), 81.0, 80.9 (C4'), 71.2 (C2'), 70.0 (C3'), 67.0, 66.6, 66.64 (CH₂ Fmoc, CH₂ Pac, C5'), 65.8 (CH₂ Allyl), 62.8, 62.7 (CH₂ CNEO), 54.6 (C_α), 46.8 (CH Fmoc), 36.8 (C_β), 33.5, 33.3 (CH *i*Bu), 19.3, 19.2 (CH₂ CNEO), 18.9, 18.6, 18.5, 18.4, 18.3 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); δ 11.74, 10.33 (2 x s, 1H), 7.77 (s, 1H), 7.75, 7.71 (d, 2H), 7.56, 7.55, 7.53 (t, 2H), 7.37, 7.35, 7.33 (t, 2H), 7.31 - 7.24 (m, 4H), 7.01 - 6.95 (m, 6H), 6.12 - 6.09 (m, 1H, H2'), 5.97, 5.95 (d, 2H, H1'), 5.89 - 5.78 (m, 2H, CH Allyl, H3'), 5.31, 5.27, 5.24, 5.21 (m, 2H, CH₂ Allyl), 4.60 - 4.52 (m, 7H, CH₂ Allyl, CH₂ Fmoc, H5', Hα), 4.43 (m, 1 H, H4'), 4.36 - 4.23 (m, 2H, CH₂ CNEO), 4.19, 4.17, 4.15 (t, 1H, CH Fmoc), 3.12 - 2.97 (m, 2H, CH₂ β), 2.73 - 2.60 (m, 2H, CH₂ CNEO), 2.56 - 2.49 (m, 2H, CH *i*Bu), 1.23, 1.21 (d, 6H, CH₃ *i*Bu), 1.14, 1.12, 1.10, 1.06 (dd, 6H, CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ - 6.94. IR: 2970, 1695, 1610, 1558, 1539, 1497, 1472, 1448, 1188, 1153, 1041. LCMS (50-90% B in 15 min), Rt = 7.56 ESI-MS: *m/z* 1116.6 [M+H]⁺. α_D (CHCl₃): -17.4. HRMS: C₅₆H₅₈N₇O₁₆P + H]⁺ calcd. 1116.3750, found 1116.3765.

¹³C-NMR (100 MHz, CDCl₃), isomer B (Rt = 0.5), 90 % pure, 96 % ee.; δ 175.6, 175.1, 170.7, 170.1, 156.9, 155.5, 155.0, 148.4, 148.3, 147.4, 146.5, 143.6, 140.9, 134.0, 131.2, 130.3 (CH Allyl), 129.4, 127.4, 126.8, 124.9, 122.6, 122.0, 119.7, 119.6, 118.8 (CH₂ Allyl), 116.1, 114.6, 87.0 (C1'), 81.0, 80.9 (C4'), 71.2 (C2'), 69.9 (C3'), 67.0, 66.6, 66.4, 65.8, 65.87 (CH₂ Allyl, C5', CH₂ Fmoc, CH₂ Pac), 62.8, 62.7 (CH₂ CNEO), 54.6 (C_α), 46.8 (CH Fmoc), 36.8 (C_β), 33.5, 33.3 (CH *i*Bu), 19.3, 19.2 (CH₂ CNEO), 18.9, 18.6, 18.5, 18.4 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); δ 11.84, 10.42 (2 x s, 1H), 7.75, 7.75, 7.73 (m, 3H), 7.57, 7.56 (m, 2H), 7.39, 7.37, 7.35 (m, 2H), 7.29 – 7.25 (m, 4H), 7.09 – 6.97 (m, 6H), 6.01 – 5.95 (m, 2H, H1', H2'), 5.93 – 5.82 (m, 1H, CH Allyl), 5.79 – 5.76 (m, 1H, H3'), 5.33 – 5.23 (m, 2H, CH₂ Allyl), 4.69, 4.68 (m, 2H, CH₂ Pac), 4.64 – 4.60 (m, 3H, CH_α, CH₂ Allyl), 4.48 (m, 2H, CH₂ Fmoc), 4.38 – 4.23 (m, 2H, H4', H5'), 4.21 – 4.15 (m, 3H, CH Fmoc, CH₂ CNEO), 3.16 – 3.06 (m, 2H, CH₂ β), 2.65 – 2.49 (m, 4H, CH *i*Bu, CH₂ CNEO), 1.21, 1.19 (d, 6H, CH₃ *i*Bu), 1.15, 1.13, 1.11, 1.09 (dd, 6H, CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -7.3. IR: 2970, 1695, 1610, 1558, 1539, 1497, 1471, 1450, 1387, 1190, 1153, 1033. LCMS (50-90% B in 15 min), Rt = 7.61 ESI-MS: *m/z* 1116.5 [M+H]⁺. α_D (CHCl₃): -24.8. HRMS: [C₅₆H₅₈N₇O₁₆P + H]⁺ calcd. 1116.3750, found 1116.3764.

2-cyanoethoxy-(N^α- Fmoc-tyrosin-4-yl)-(6-N-benzoyl-2',3'-di-O-isobutyryl-β-D-adenyl-5'-yl) phosphate (14)

Compound **11** (0.39 g, 0.37 mmol) was coevaporated with MeCN and dissolved in 1:1 THF/DCM (10 mL, v/v). After addition of AcOH (1.7 mmol, 0.10 mL), Bu₃SnH (0.21 mL) and Pd(PPh₃)₄ (18 mg) the reaction was stirred for 2 hours. The reaction mixture was concentrated *in vacuo* and the residue was applied to a silica gel column and eluted with a gradient of MeOH in DCM (0/100 – 5/95) to yield the title compound as a white foam (0.32 g, 0.31 mmol, 85 %). ¹³C-NMR (100 MHz, CDCl₃); δ 175.5, 175.3, 174.4, 174.0, 165.7, 155.4, 153.1, 151.6, 149.8, 149.0, 143.75, 143.68, 143.61, 141.3, 141.2, 140.4, 134.0, 133.5, 132.7, 131.2, 131.0, 128.5, 128.5, 128.2, 128.2, 127.7, 127.6, 127.0, 125.0, 125.0, 124.9, 124.7, 122.0, 121.8, 119.9, 119.9, 119.7, 116.2, 116.1, 85.2, 84.8 (C1'), 82.0, 81.9 (C4'), 73.6, 73.3 (C2'), 70.4, 70.0 (C3'), 67.6, 67.1, 66.6, 66.4 (CH₂ Fmoc, CH₂ CNEO), 63.0, 63.0, 62.8, 62.8 (C5'), 55.0, 54.8 (CH α), 47.2 (CH Fmoc), 38.3, 37.9 (CH₂ β), 33.79, 33.77, 33.56, 33.53 (CH *i*Bu), 19.7, 19.6 (CH₂ CNEO), 18.9, 18.8, 18.7, 18.6, 18.60, 18.57 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); δ 8.78, 8.77 (2x s, 1H), 8.12, 7.86 (2x s, 1H), 7.98 – 6.89 (m, 17H), 6.43, 6.42, 6.39, 6.37 (2x d, 1H, H1'), 5.97, 5.95 (2x s, 1H), 5.75 – 5.59 (m, 2H, C3', C2'), 4.52 – 4.36 (m, 7H, CH₂ Fmoc, CH α, C4', CH₂ CNEO, CH Fmoc), 4.17 – 4.15 (m, 2H, H5'), 3.27 – 3.21 (m, 2H, CH₂ β), 2.98 – 2.92 (m, 2H, CH₂ CNEO), 2.72 – 2.48 (2x m, 2H, CH *i*Bu), 1.29 – 1.20, 1.11 – 1.06 (m, 12H, CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -6.8, -7.1. IR: 2976, 2360, 1738, 1714, 1700, 1584, 1505, 1455, 1247, 1155, 1032. LCMS (10 – 90 % B in 15 min) Rt = 9.39. ESI-MS: *m/z* 1030.1 [M+H]⁺. HRMS: [C₅₂H₅₂N₇O₁₄P + H]⁺; 1030.3383 calcd., 1030.3394 found

2-cyanoethoxy-(N^α- Fmoc-tyrosin-4-yl)-(2-N-phenoxyacetyl-2',3'-di-O-isobutyryl-β-D-guanyl-5'-yl) phosphate (15)

A mixture of stereoisomers of Fmoc-Tyr(*i*Bu₂)pG(Pac))-OAll **12** (0.43 g, 0.38 mmol) was coevaporated with MeCN and dissolved in 1:1 THF/DCM (10 mL, v/v). After addition of AcOH (2.0 mmol, 0.11 mL), Bu₃SnH (0.36 mL) and Pd(PPh₃)₄ (20 mg) the reaction was stirred for 2 hours. The reaction mixture was concentrated and the residue was applied to a silica gel column and eluted with a gradient of MeOH in DCM (0/100 – 4/96) containing 1% of AcOH to yield the title compound as a white foam (0.23 g, 0.21 mmol, 56 %). ¹³C-NMR (100 MHz, CDCl₃); δ

176.2, 175.9, 175.5, 174.0, 170.3, 156.8, 155.4, 155.4, 148.9, 148.8, 148.5, 147.8, 146.9, 143.9, 143.8, 141.3, 134.2, 131.2, 129.8, 127.7, 127.1, 125.2, 122.6, 121.6, 119.9, 119.7, 116.6, 114.9, 87.0 (C1'), 80.9, 80.8 (C4'), 72.5 (C2'), 69.8 (C3'), 67.1, 66.9, 66.8 (C5', CH₂ PAC, CH₂ Fmoc), 63.1, 63.03 (CH₂ CNEO), 54.9 (C_α), 47.2 (CH Fmoc), 37.1 (C_β), 33.8, 33.7 (CH *i*Bu), 19.5, 19.4 (CH₂ CNEO), 18.9, 18.84, 18.81, 18.7 (CH₃ *i*Bu). ¹H-NMR (400 MHz, CDCl₃); δ 7.88 (s, 1H), 7.75 - 7.73 (m, 3H), 7.59 - 7.57 (m, 2H), 7.39 - 7.35 (m, 2H), 7.31 - 7.27 (m, 4H), 7.13 - 7.12 (d, 2H), 7.04 - 6.99 (m, 6H), 6.00, 5.98 (d, 1H, H1'), 5.96 - 5.92 (m, 1H, H3'), 5.88 - 5.85 (m, 1H, H2'), 4.68 (m, 2H, CH₂ Pac), 4.48 - 4.17 (m, 9H, CH α, CH₂ Fmoc, H4', H5', CH Fmoc, CH₂ CNEO), 3.15 (m, 1H, CH₂ β), 2.67 - 2.55 (m, 4H, CH *i*Bu, CH₂ CNEO), 1.27 - 1.10 (m, 12H, CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -7.0, -7.1. IR: 2951, 1686, 1574, 1560, 1412, 1248, 1209, 1153, 1070. LCMS (10-90% B in 15 min), Rt = 8.96, ESI-MS: *m/z* 1076.6 [M+H]⁺. HRMS: C₅₃H₅₄N₇O₁₆P + H⁺ calcd. 1076.3437, found 1076.3449.

VPgpA, (H₂N-Gly-Ala-Tyr(pA)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Glu-OH) (21).

The protected peptide **16** was prepared via solid phase peptide synthesis starting from Tentagel S RAM resin (200 mg, 50 μmol) to which Fmoc-Glu-OtBu was attached via the side-chain carboxylic acid functionality using an automated peptide synthesizer. As repetitive cycle was used: 1) Fmoc cleavage using 20% piperidine in NMP, 2.) Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. HCTU in NMP (0.25M) and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour, 3.) Capping using 0.5 M acetic anhydride, 0.125 M DipEA and 0.015 M HOBt in NMP for 1 min.

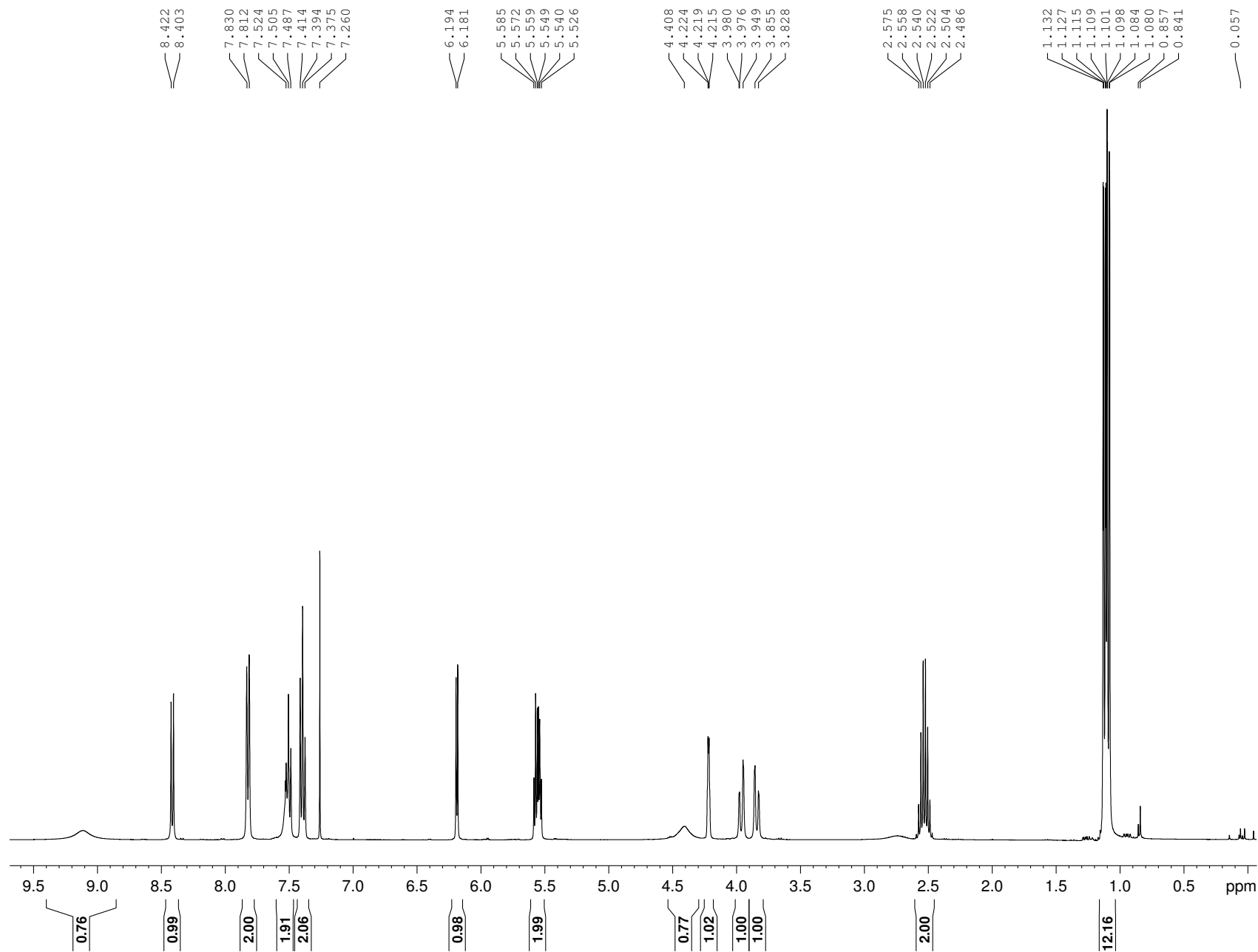
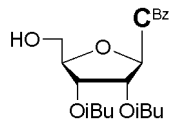
The last three amino acids were introduced manually on 25 μmol **16** by a double coupling protocol: 1) Fmoc cleavage using 2% DBU in DMF, 2.) Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. HATU in NMP (0.22M) and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour, 3.) Capping using 0.5 M acetic anhydride, 0.125 M DipEA and 0.015 M HOBt in NMP for 1 min. The crude partially protected nucleopeptide was cleaved of the resin using TFA/TIS/H₂O (95/5/5, v/v/v, 5 mL) by shaking for 4 hours. After filtration into cold Et₂O the precipitate was concentrated *in vacuo* and RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C₁₈ column (250 x 10.00 mm, 5 μ, flow: 4 mL/min). The applied eluent system was; A: 0,5 % TFA in H₂O, B: MeCN, Gradient: 20- 50 % B in 3 CV and detection at 280 nm. to give partially protected **18**. This nucleopeptide was then treated with 25% aq. ammonia (1.25 mL) in 1,4-dioxane (1.25 mL). After stirring for 5 hours careful concentration of the reaction mixture afforded the crude title compound. RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C₁₈ column (250 x 10.00 mm, 5 μ, flow: 4 mL/min). The applied eluents system was; A: 0,5 % TFA in H₂O, B: MeCN and detection at 280 nm. Gradient: 10- 25 % B in 3 CV. The collected fractions were lyophilized, yielding the target nucleopeptide **21**; 6.1 mg, 2.3 μmol (9% based on initial loading of the resin). ³¹P-NMR (161 MHz, D₂O); δ - 4.21. LCMS (00 - 90 % B in 15 min.), Rt = 4.18 min. ESI-MS: *m/z* 1345.0 [M]²⁺, 896.6 [M]³⁺, 672.7[M]⁴⁺. HRMS: [C₁₁₅H₁₈₉N₃₆O₃₆P + H]²⁺ calcd. 1341.6974, found 1341.6949, [M + H]³⁺ calcd. 894.8007, found 894.8008, [M + H]⁴⁺ calcd. 671.3524, found 671.3532, [M + H]⁵⁺ calcd. 537.2833, found 537.2839.

VPgpG, (H₂N-Gly-Ala-Tyr(pG)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Glu-OH (22)

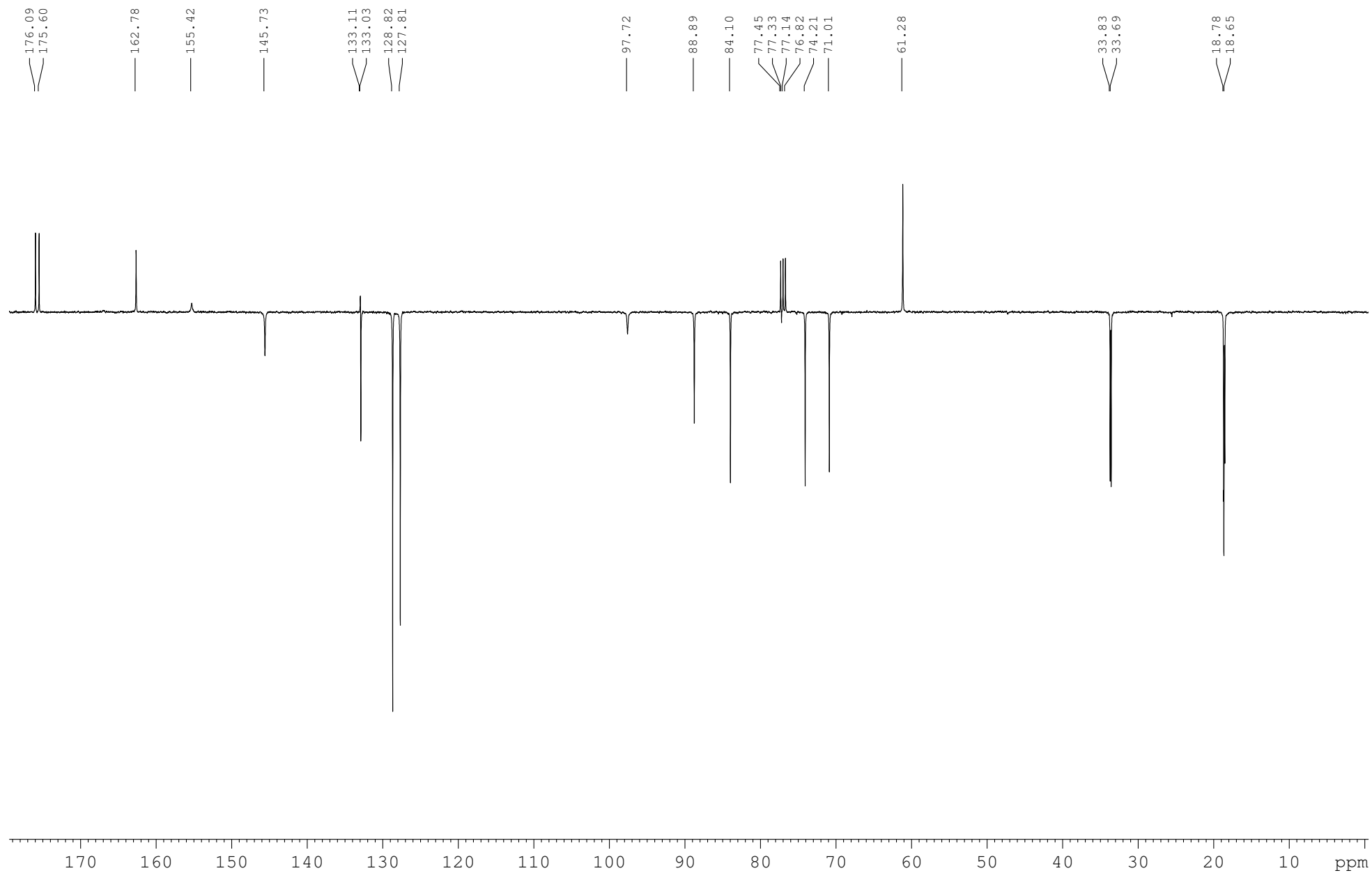
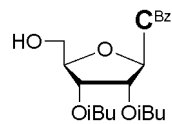
The protected peptide **16** was prepared via solid phase peptide synthesis starting from Tentagel S RAM resin (200 mg, 50 μmol) to which Fmoc-Glu-OtBu was attached via the side-chain carboxylic acid functionality using an automated peptide synthesizer. As repetitive cycle was used: 1) Fmoc cleavage using 20% piperidine in NMP, 2.) Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. HCTU in NMP (0.25M) and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour, 3.) Capping using 0.5 M acetic anhydride, 0.125 M DipEA and 0.015 M HOBt in NMP for 1 min.

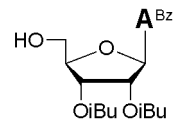
The last three amino acids were introduced manually on 12.5 μmol **16** by a double coupling protocol: 1) Fmoc cleavage using 2% DBU in DMF, 2.) Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. HATU in NMP (0.22M) and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour, 3.) Capping using 0.5 M acetic anhydride, 0.125 M DipEA and 0.015 M HOBt in NMP for 1 min. The crude partially protected nucleopeptide was cleaved of the resin using TFA/TIS/H₂O (95/5/5, v/v/v, 5 mL) by shaking for 2.5 hours. After filtration into cold Et₂O the precipitate was concentrated *in vacuo*. The residue was then treated with 25% aq. ammonia (1.25 mL) in 1,4-dioxane (1.25 mL). After stirring for 5 hours careful concentration of the reaction mixture afforded the crude title compound. RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C₁₈ column (250 x 10.00 mm, 5 μ, flow: 4 mL/min). The applied eluent system was; A: H₂O, B: MeCN, C: 0,5 % TFA in H₂O, in a gradient of 10- 30 % B (10% C as stationary phase) in 3 CV and detection at 280 nm The collected fractions were lyophilized, yielding the target nucleopeptide **22**; 9.8 mg, 3.6 μmol (29% based on initial loading of the resin). ³¹P-NMR (161 MHz, D₂O); δ -4.21. LCMS (00 – 90% B in 15 min.), Rt = 4.21 min., ESI-MS: *m/z* 2700.2 [M + H]⁺, 1350.2[M+H]²⁺, 900.6 [M+H]³⁺. HRMS: [C₁₁₅H₁₈₉N₃₆O₃₇P + H]²⁺ calcd. 1349.6948, found 1349.6952, [M + H]³⁺ calcd. 900.1323, found 900.1327, [M + H]⁴⁺ calcd. 675.3511, found 675.3508, [M + H]⁵⁺ calcd. 540.4823, found 540.4816.

¹H NMR, 400MHz, CDCl₃, compound 4

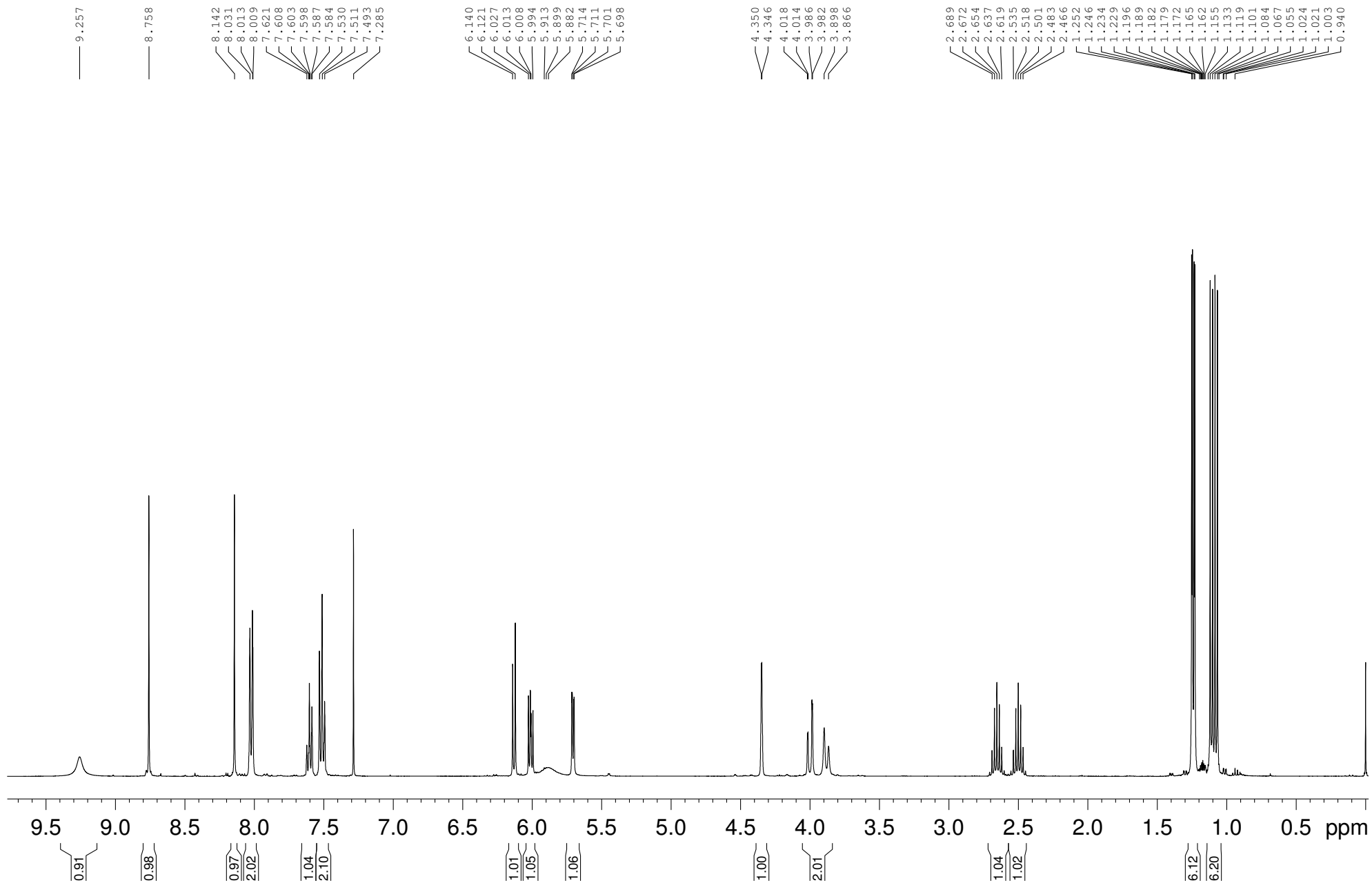


¹³C APT NMR, 100MHz, CDCl₃, compound 4

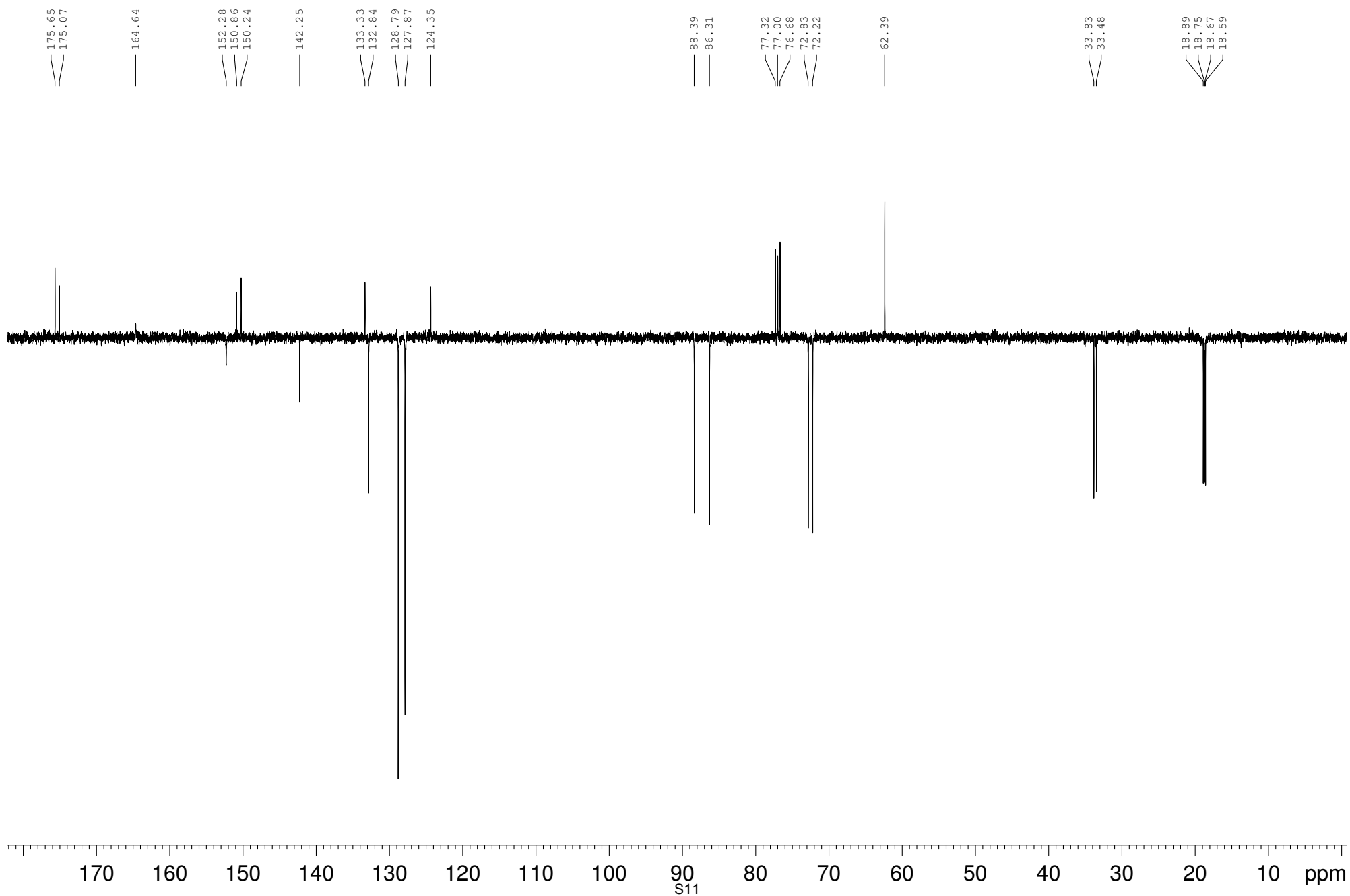
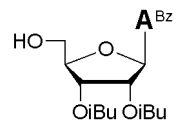




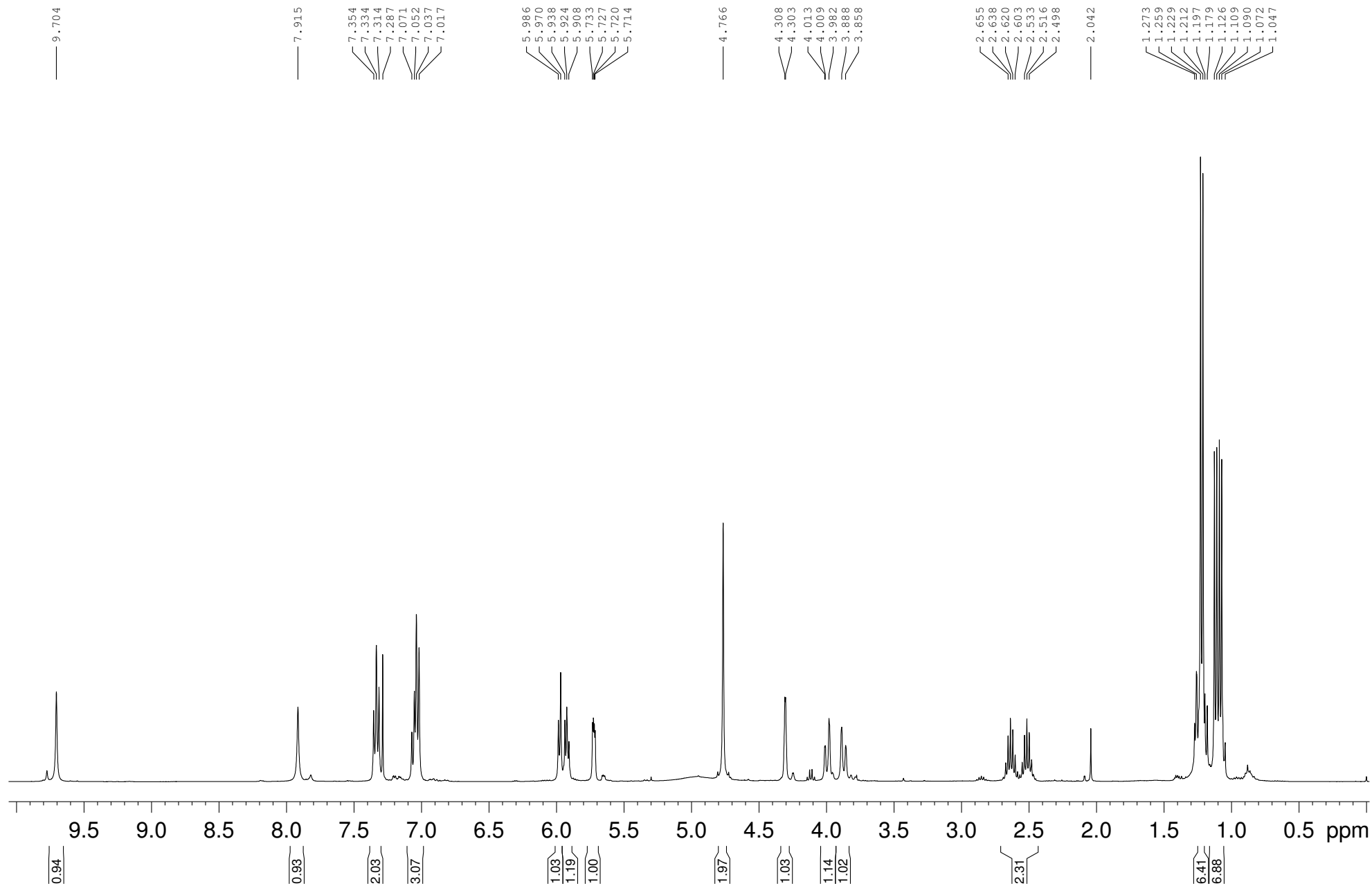
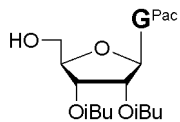
¹H NMR, CDCl₃, 400 MHz, compound 5



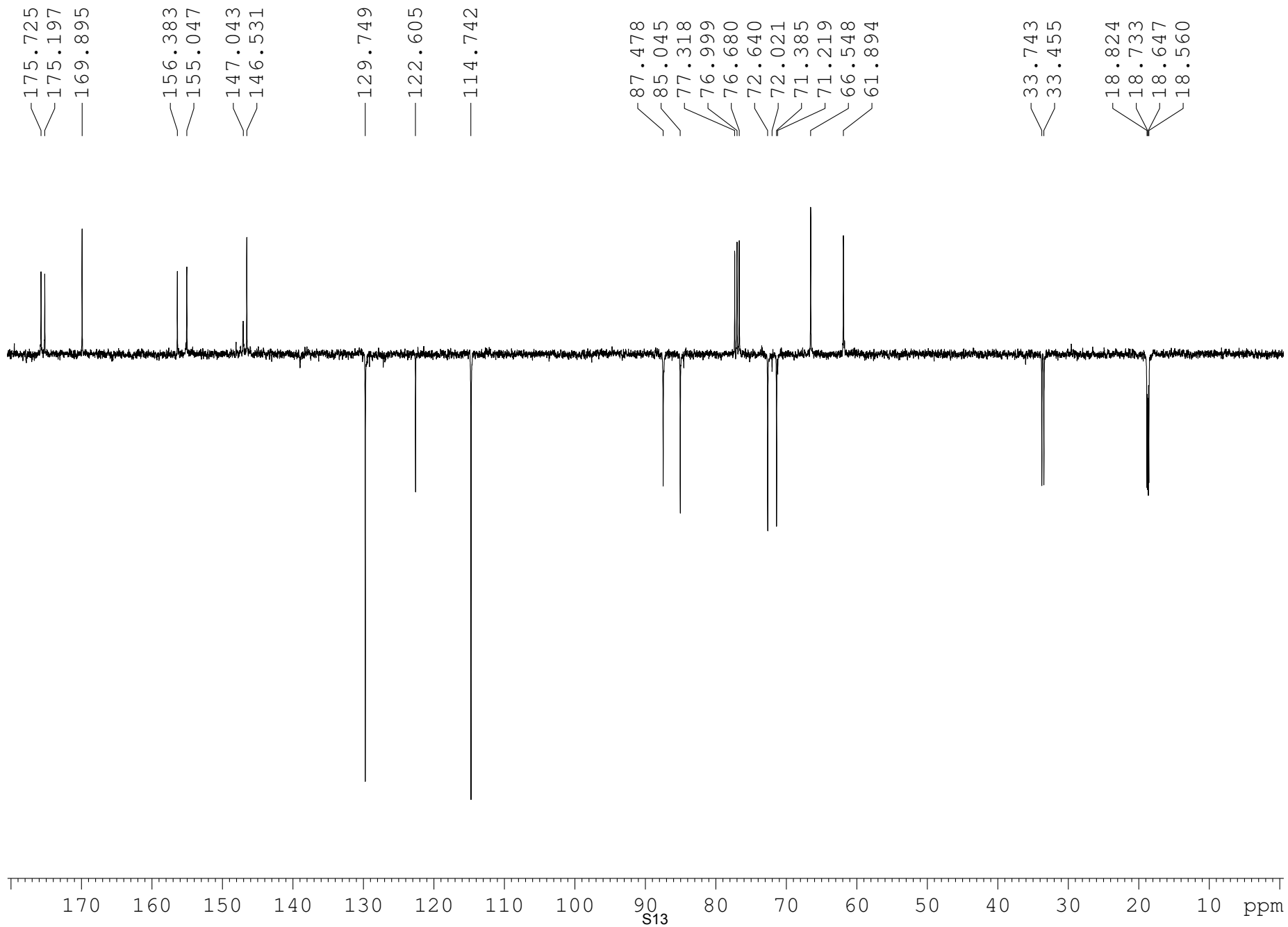
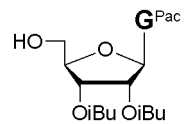
^{13}C APT NMR, CDCl_3 , 100 MHz, compound 5



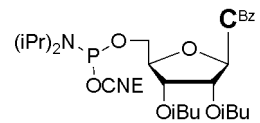
¹H NMR, CDCl₃, 400 MHz, compound 6



¹³C APT NMR, CDCl₃, 100 MHz, compound 6

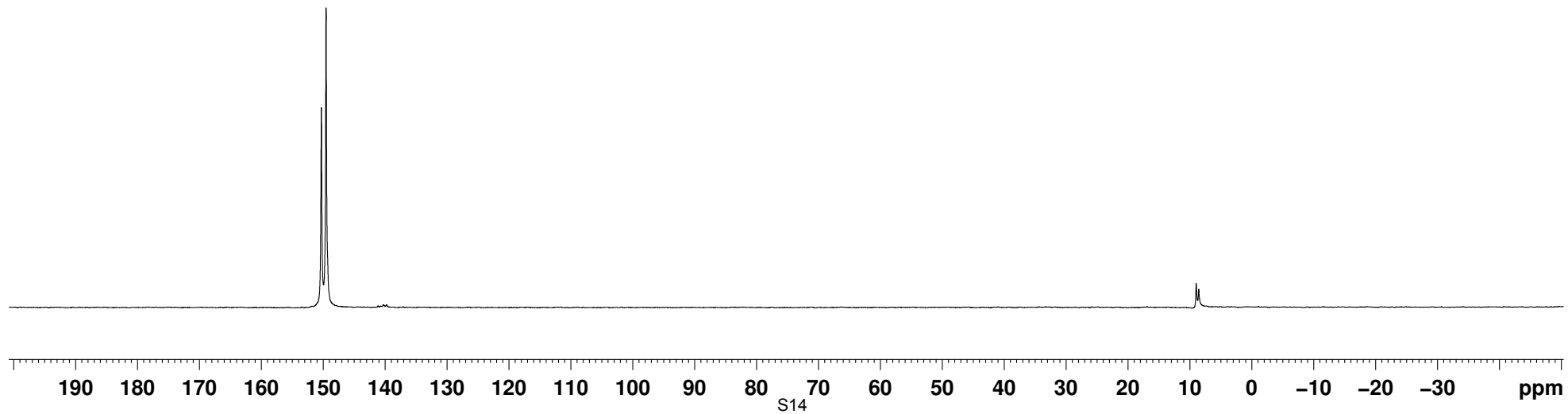


31P NMR, 161MHz, CDCl3, compound 7

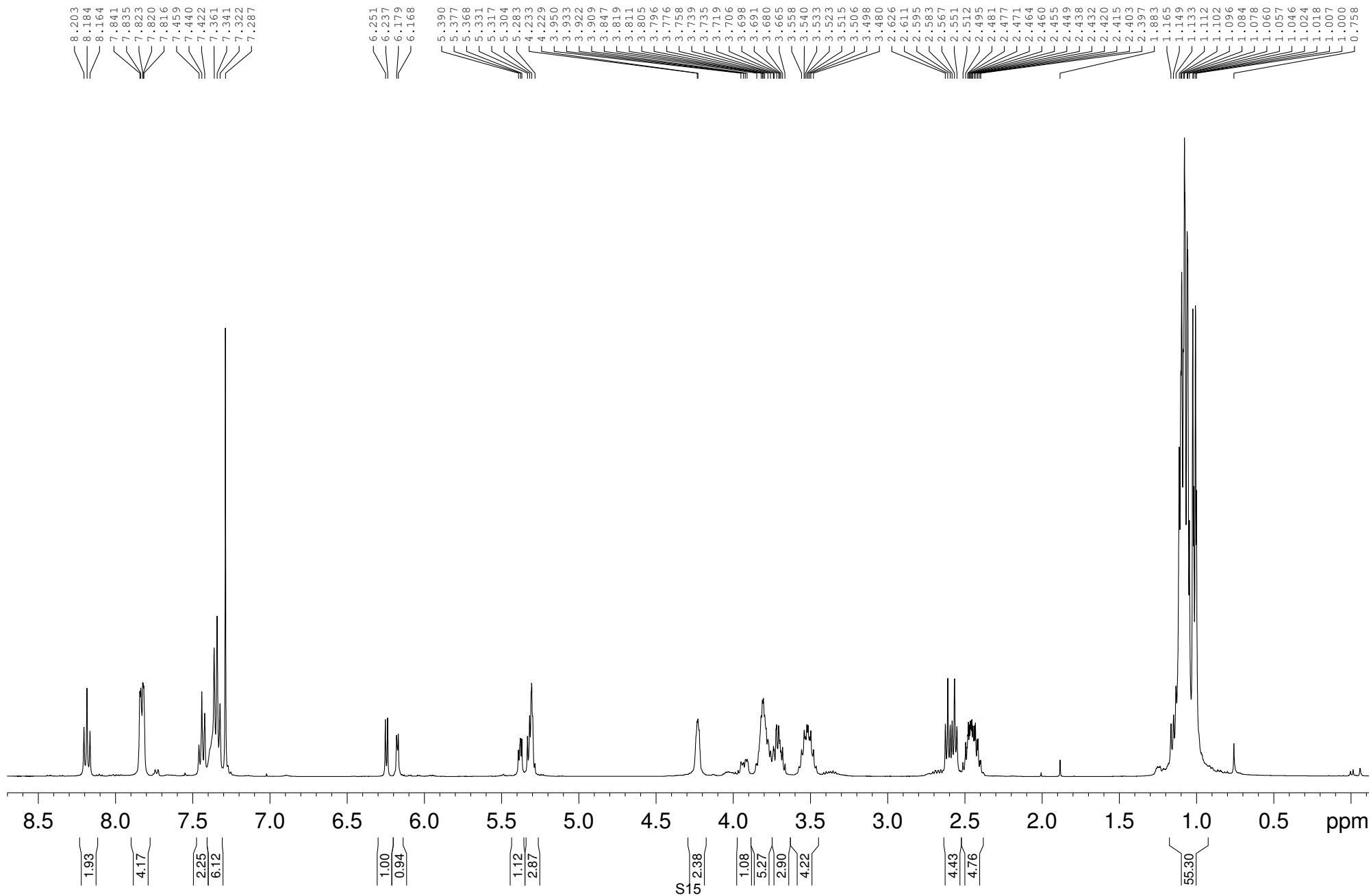
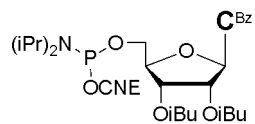


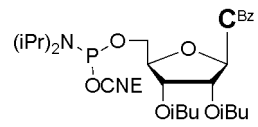
150.32
149.56

8.98

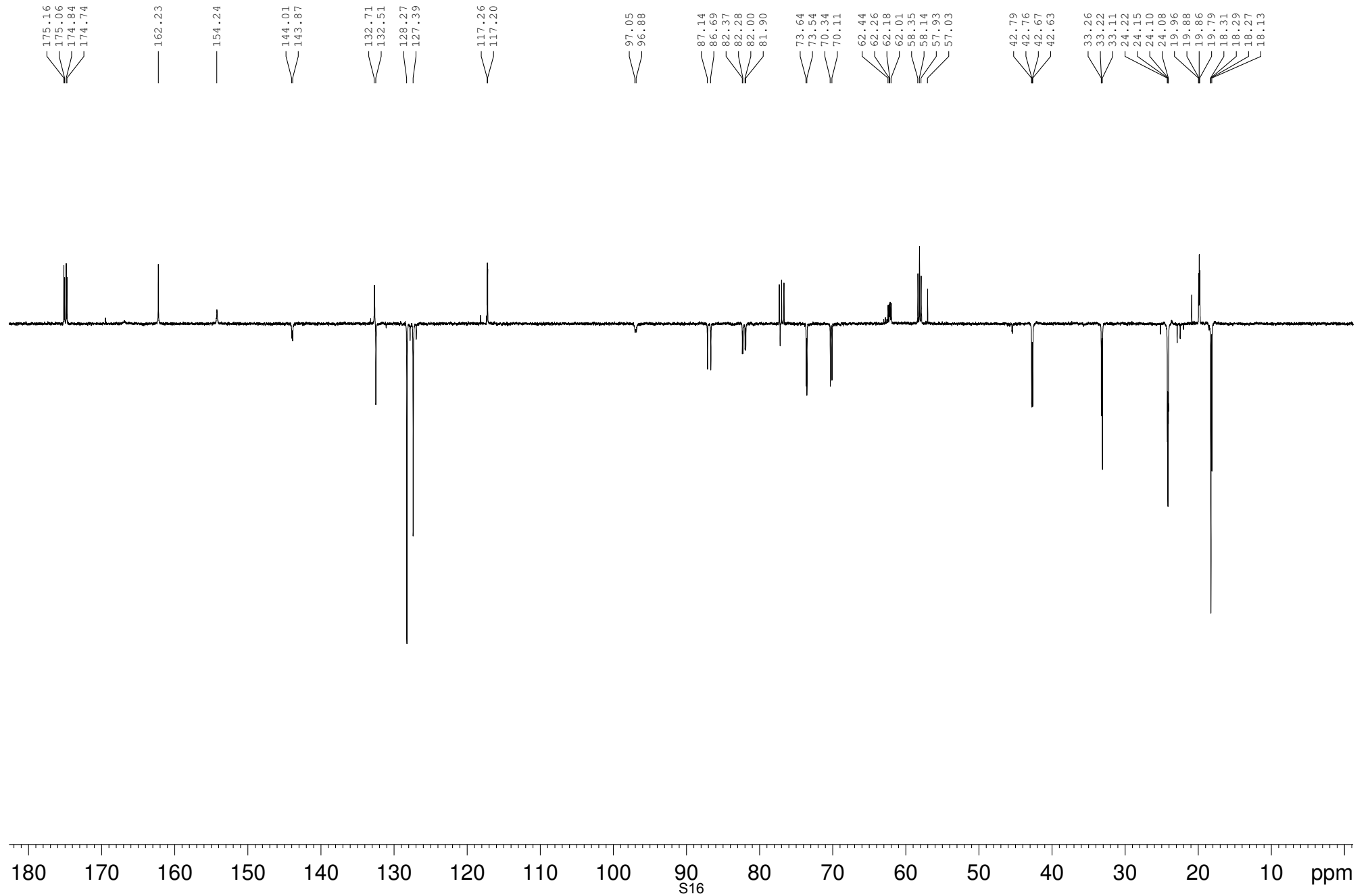


¹H NMR, CDCl₃, 400 MHz, compound 7

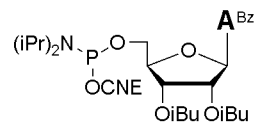




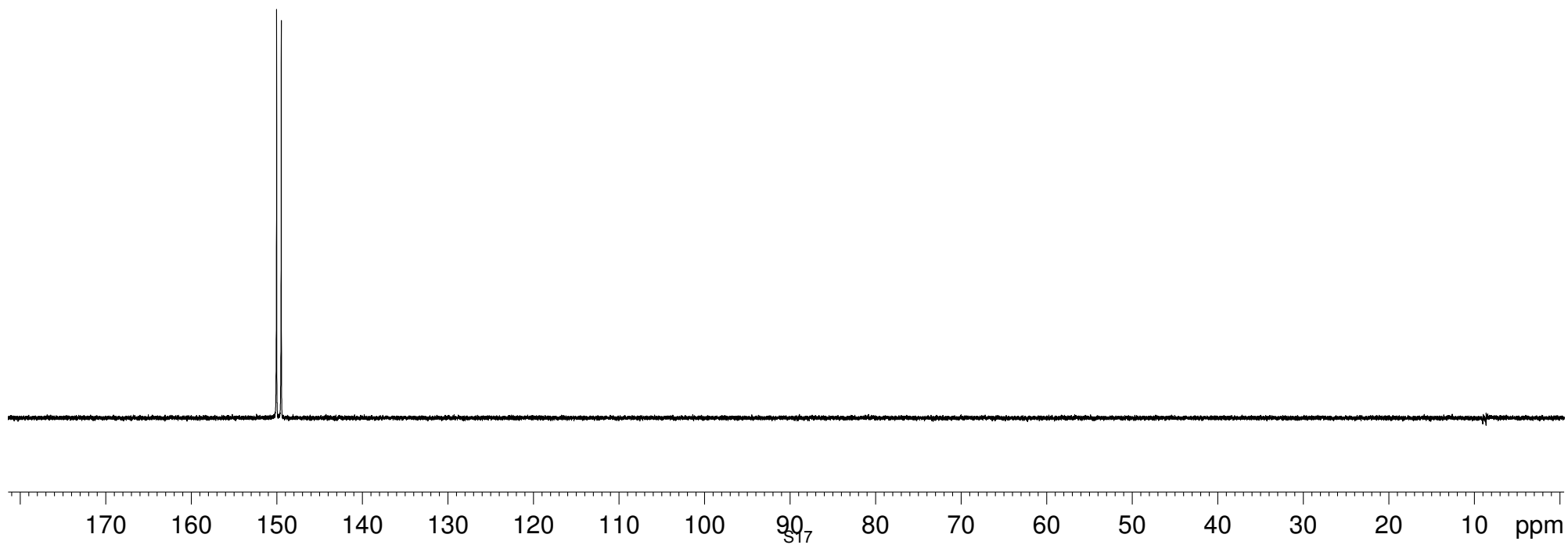
¹³C APT NMR, CDCl₃, 100 MHz, compound 7

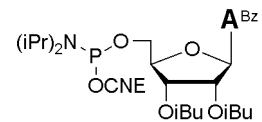


31P NMR, CDCl₃, 161 MHz, compound 8

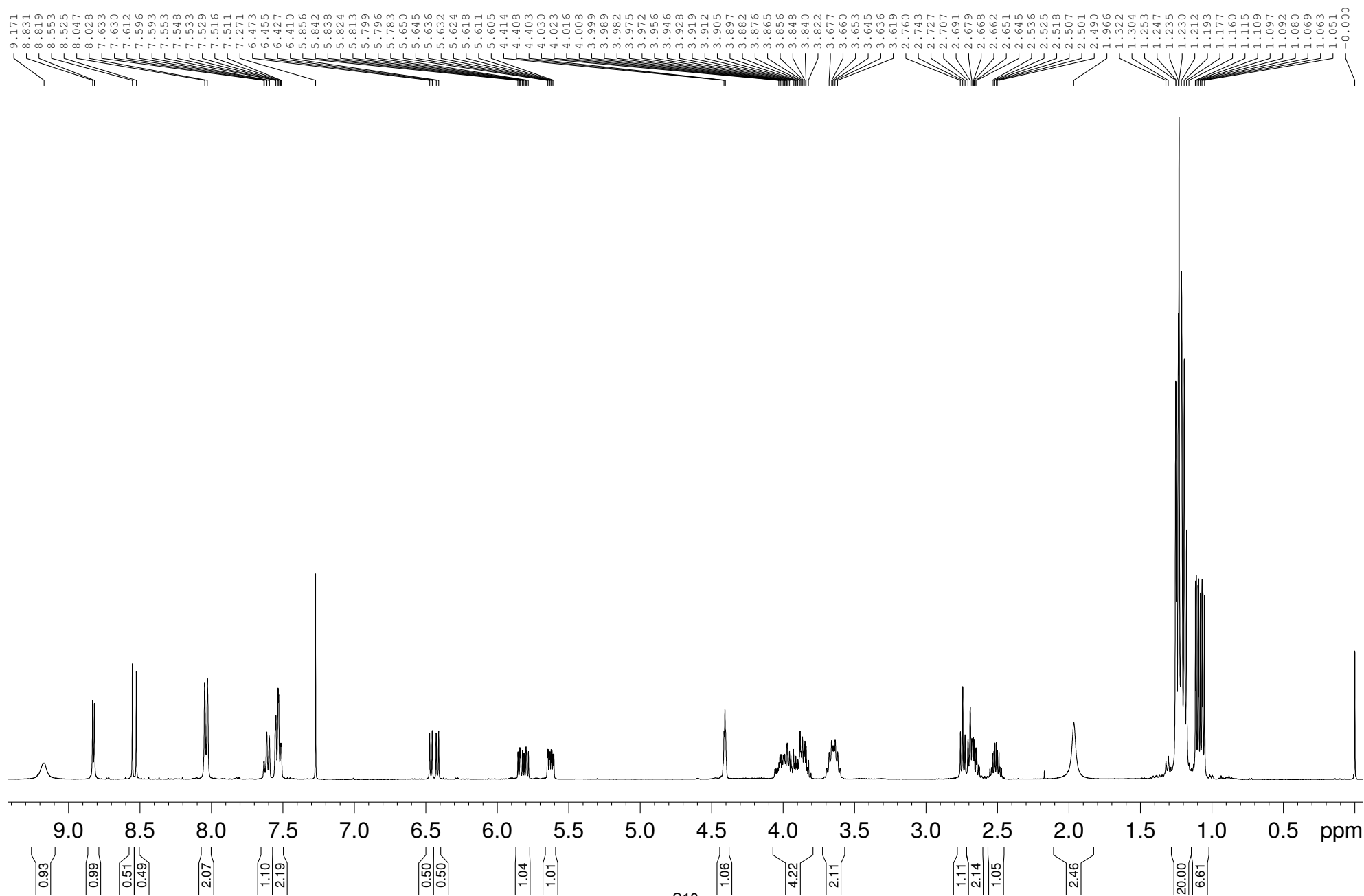


150.03
149.49

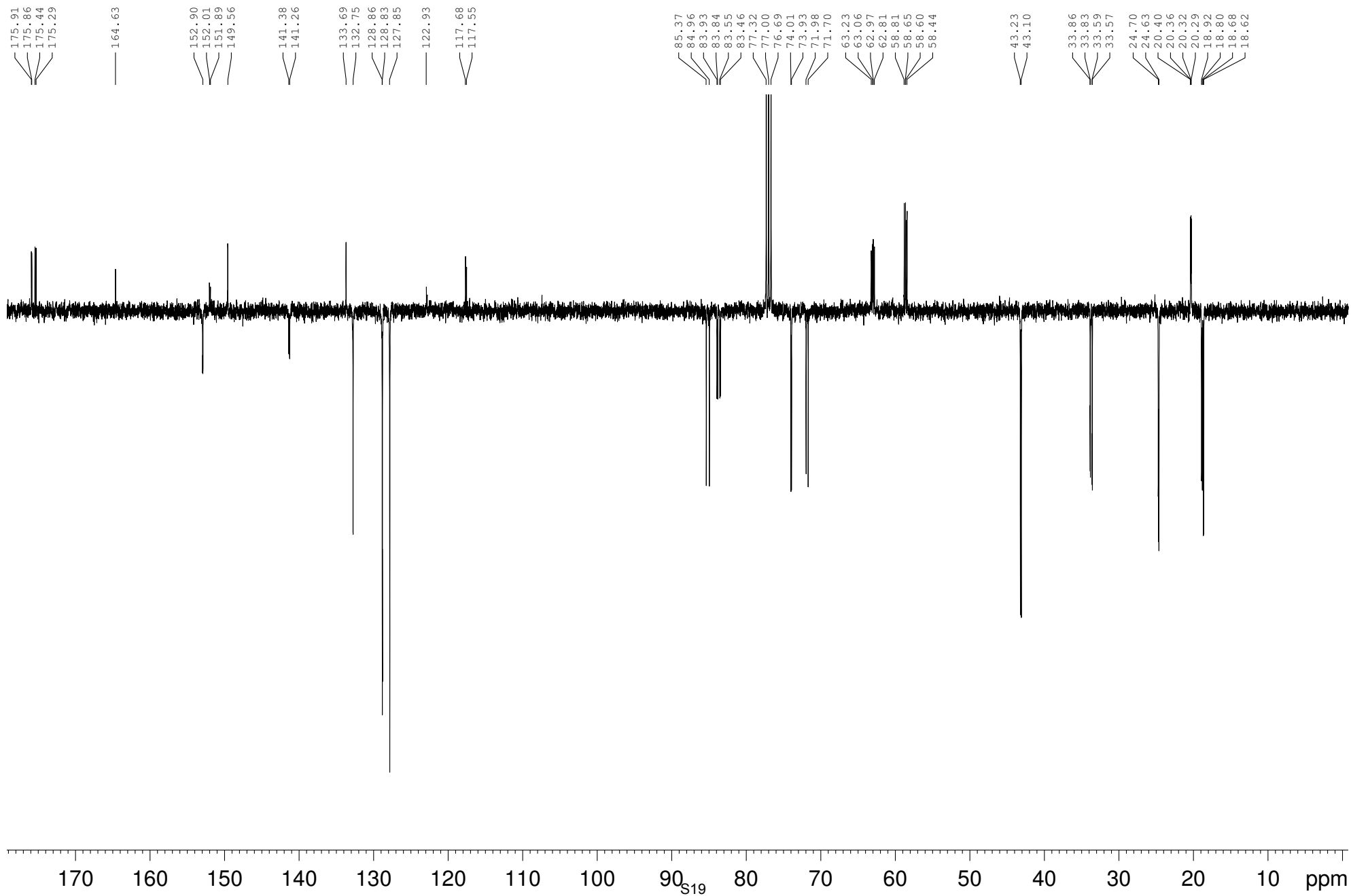
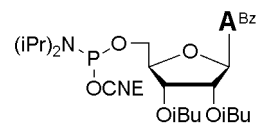




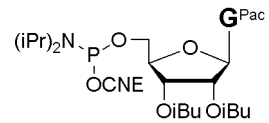
¹H NMR, CDCl₃, 400 MHz, compound 8



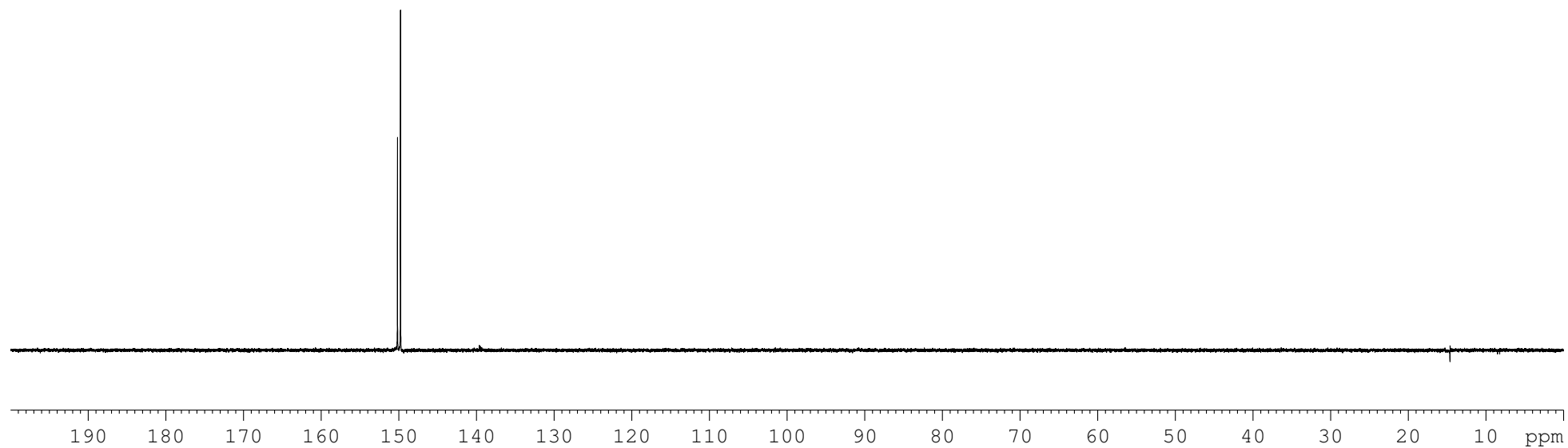
^{13}C APT NMR, CDCl_3 , 100 MHz, compound 8

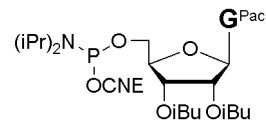


31P - NMR, CDCl3, 161 MHz, compound 9

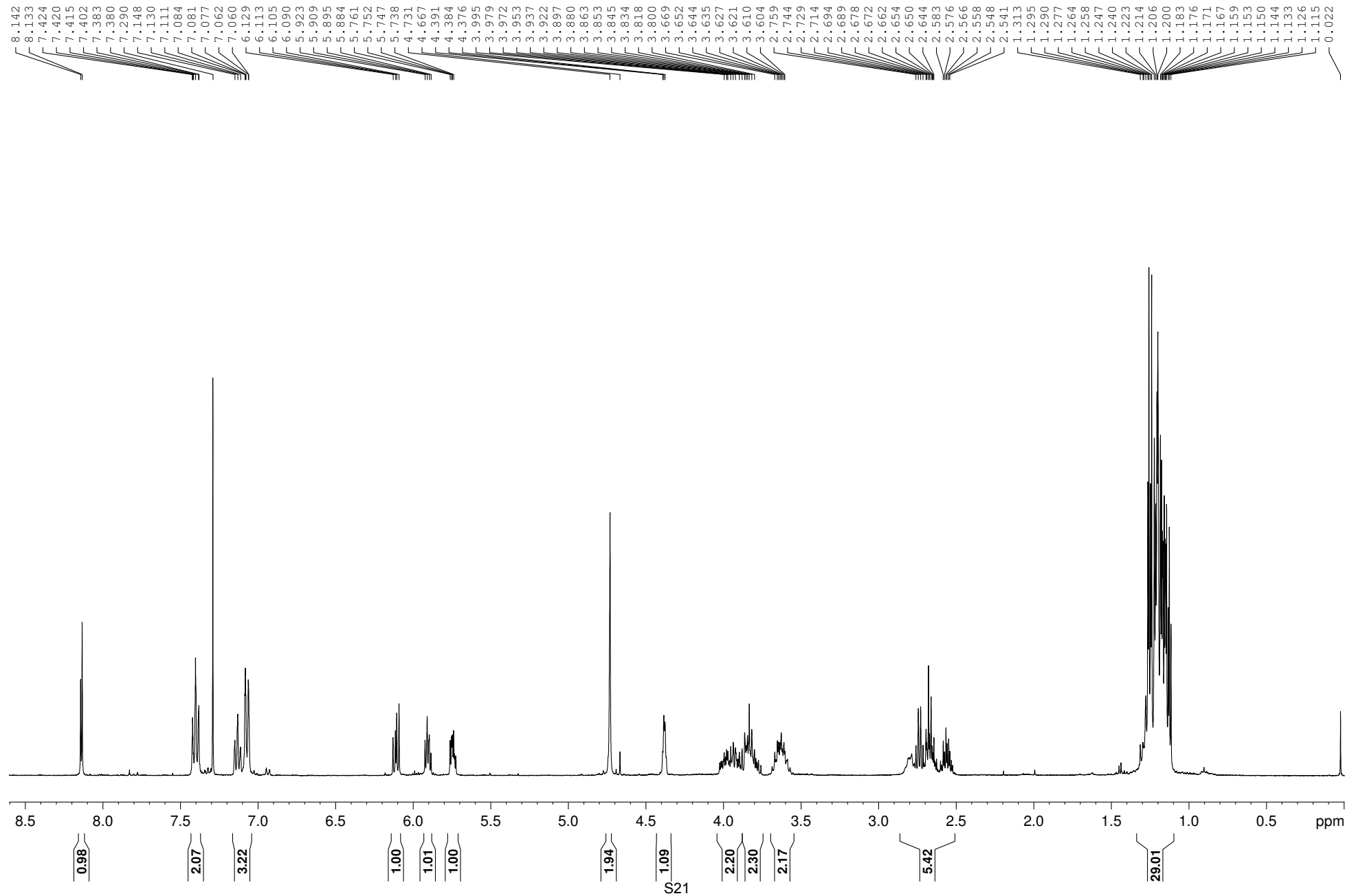


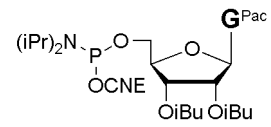
150.14
149.74



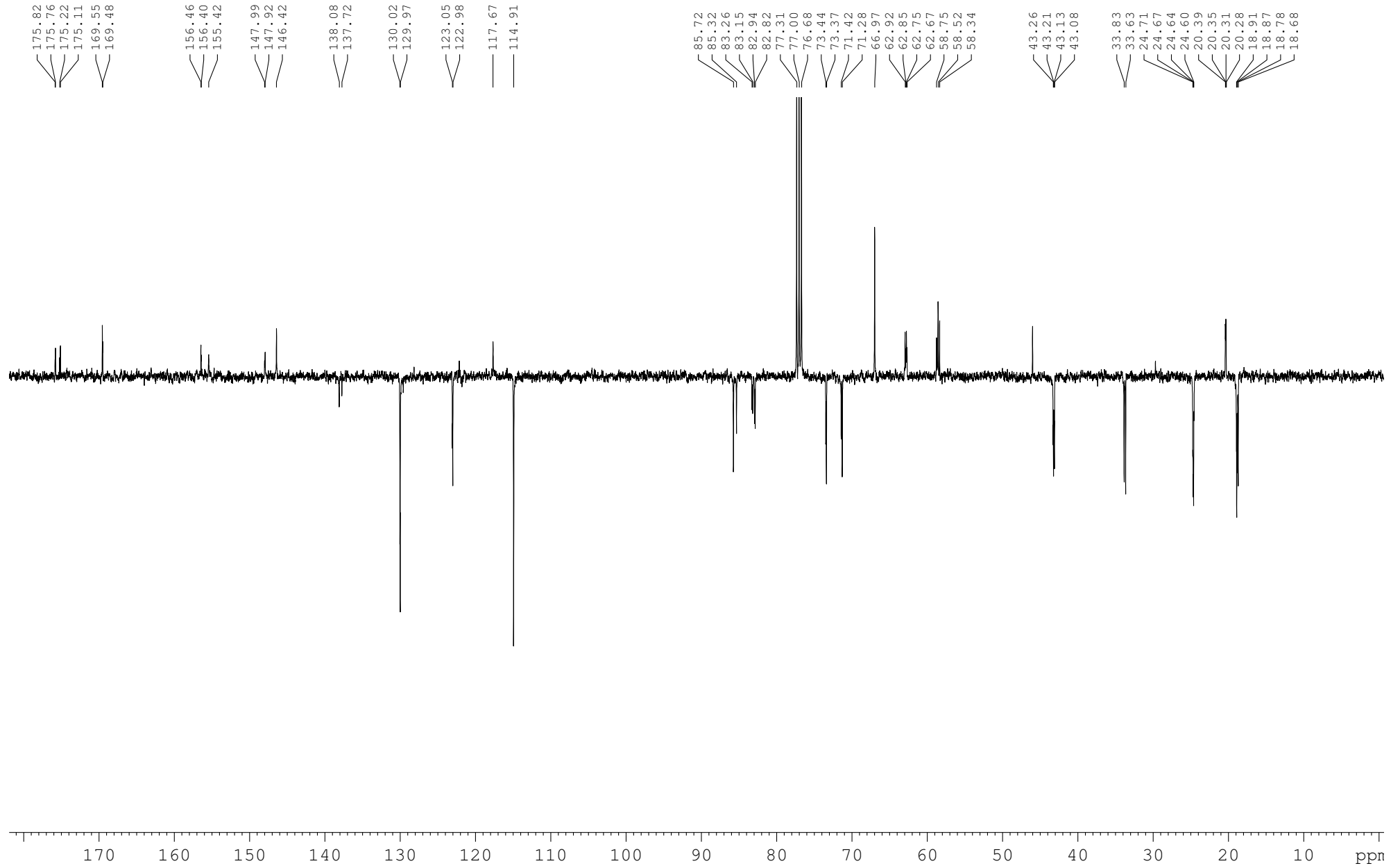


¹H - NMR, CDCl₃, 400 MHz, compound 9

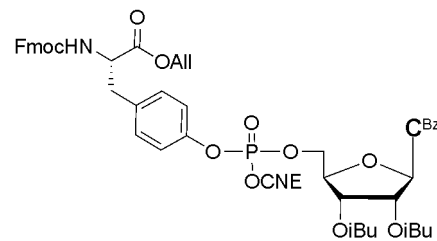




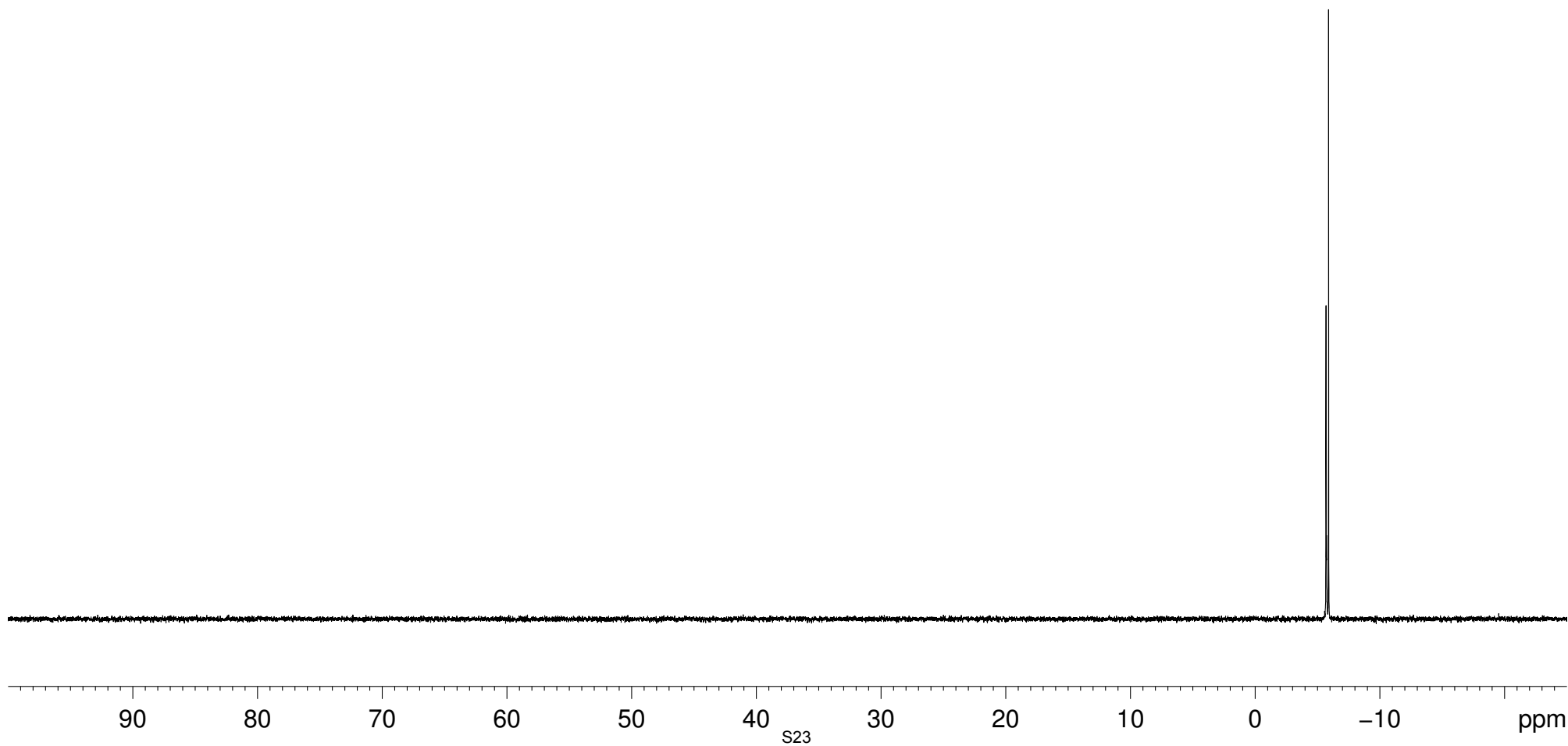
¹³C APT NMR, CDCl₃, 100 MHz, compound 9



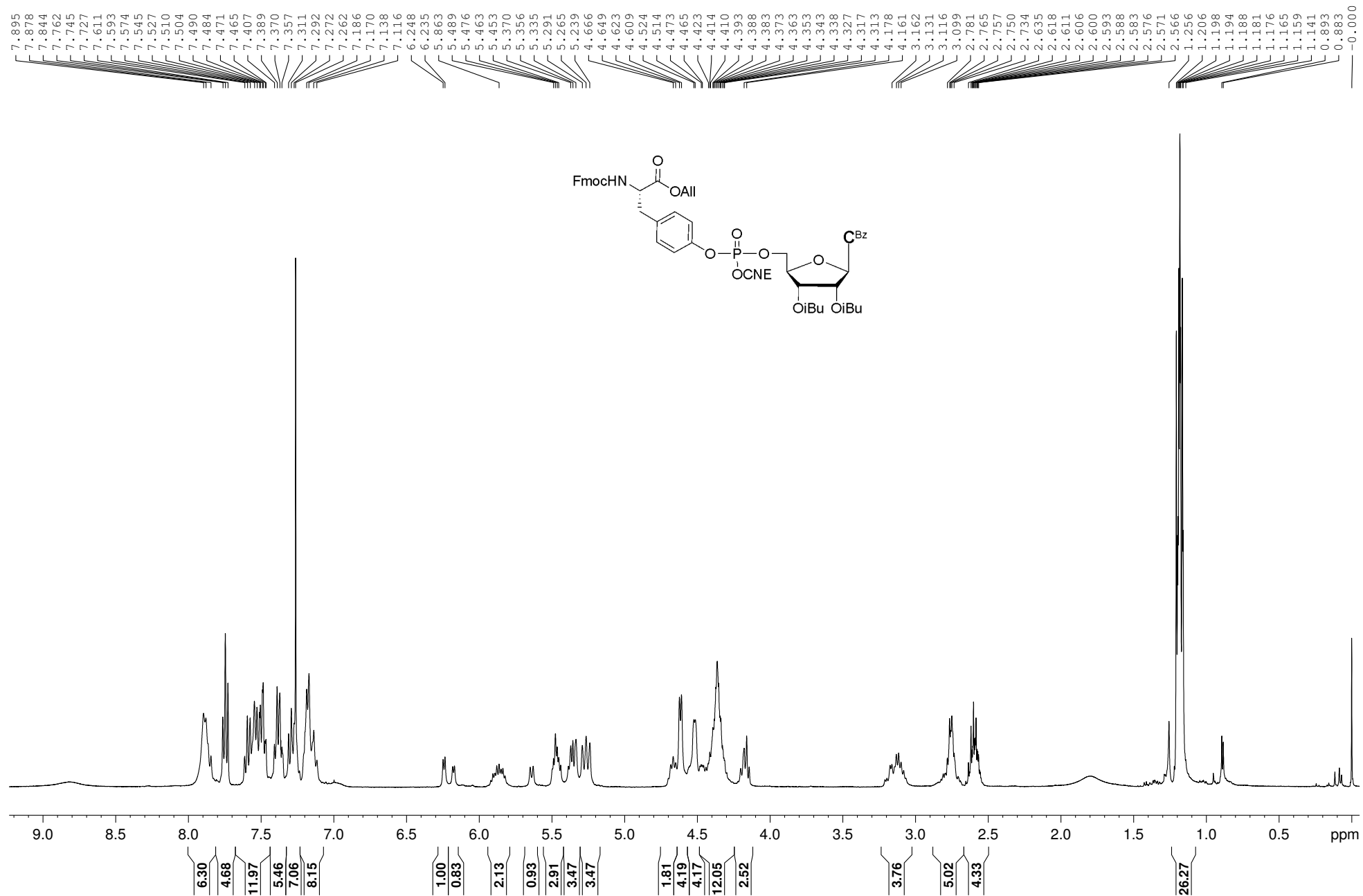
31P NMR , CDCl3, 161 MHz, compound 10



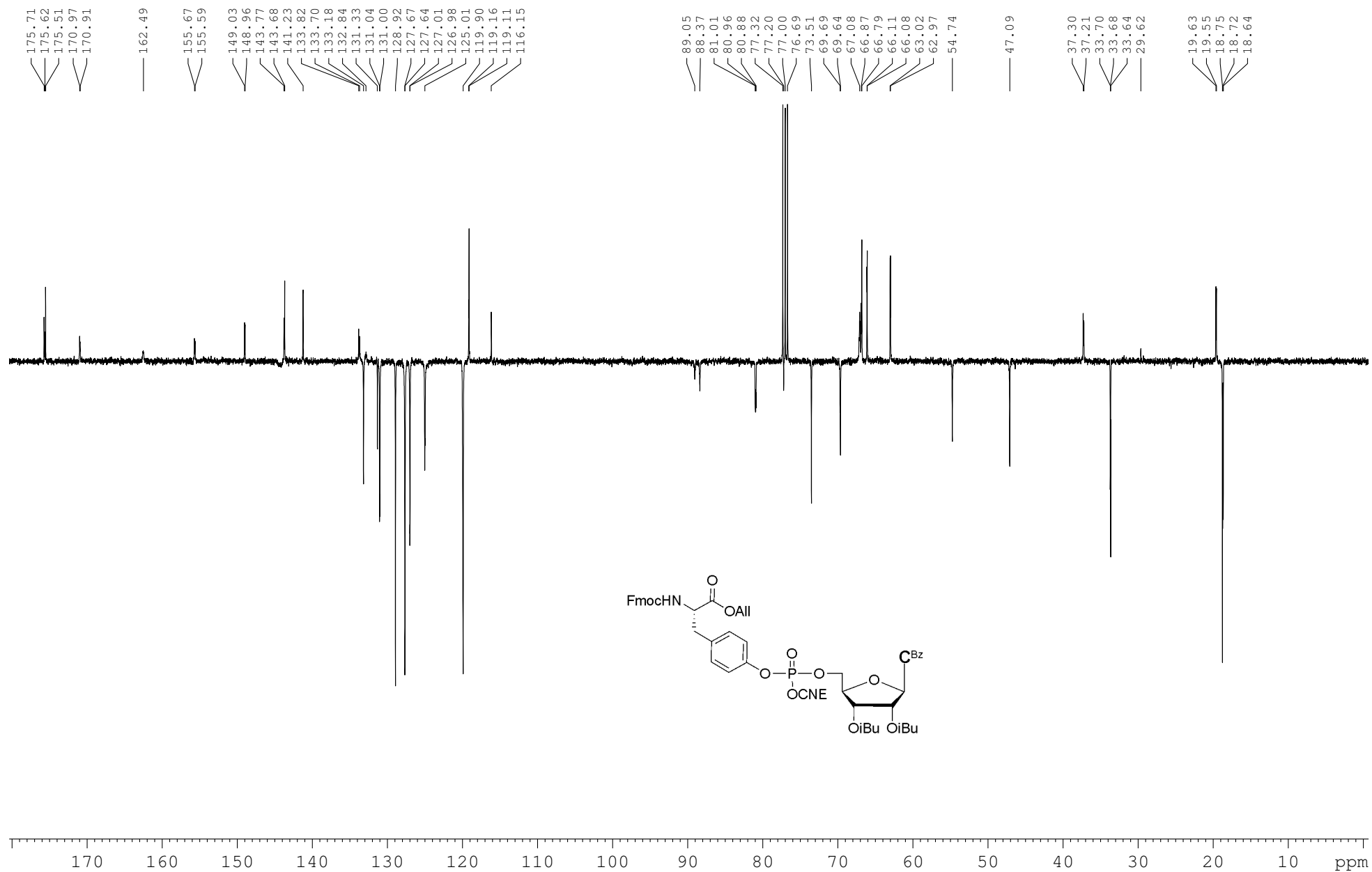
5.68
5.89



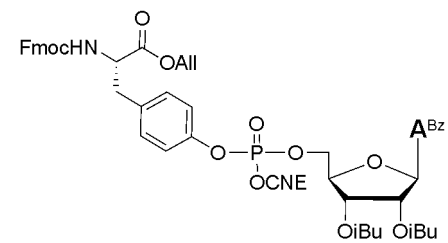
¹H NMR, 400 MHz, CDCl₃, compound 10



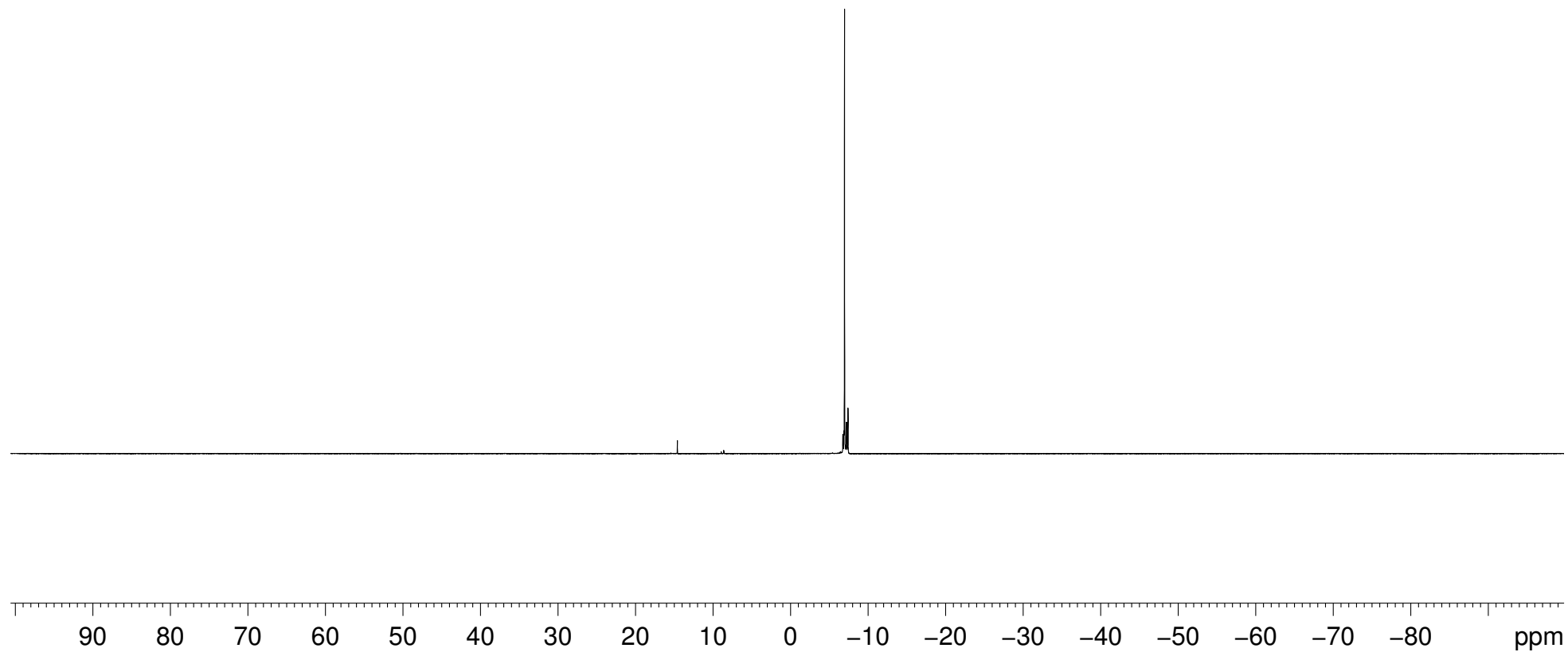
¹³C APT NMR, 100MHz, CDCl₃, compound 10



31P NMR, CDCl3, 161 MHz, compound 11

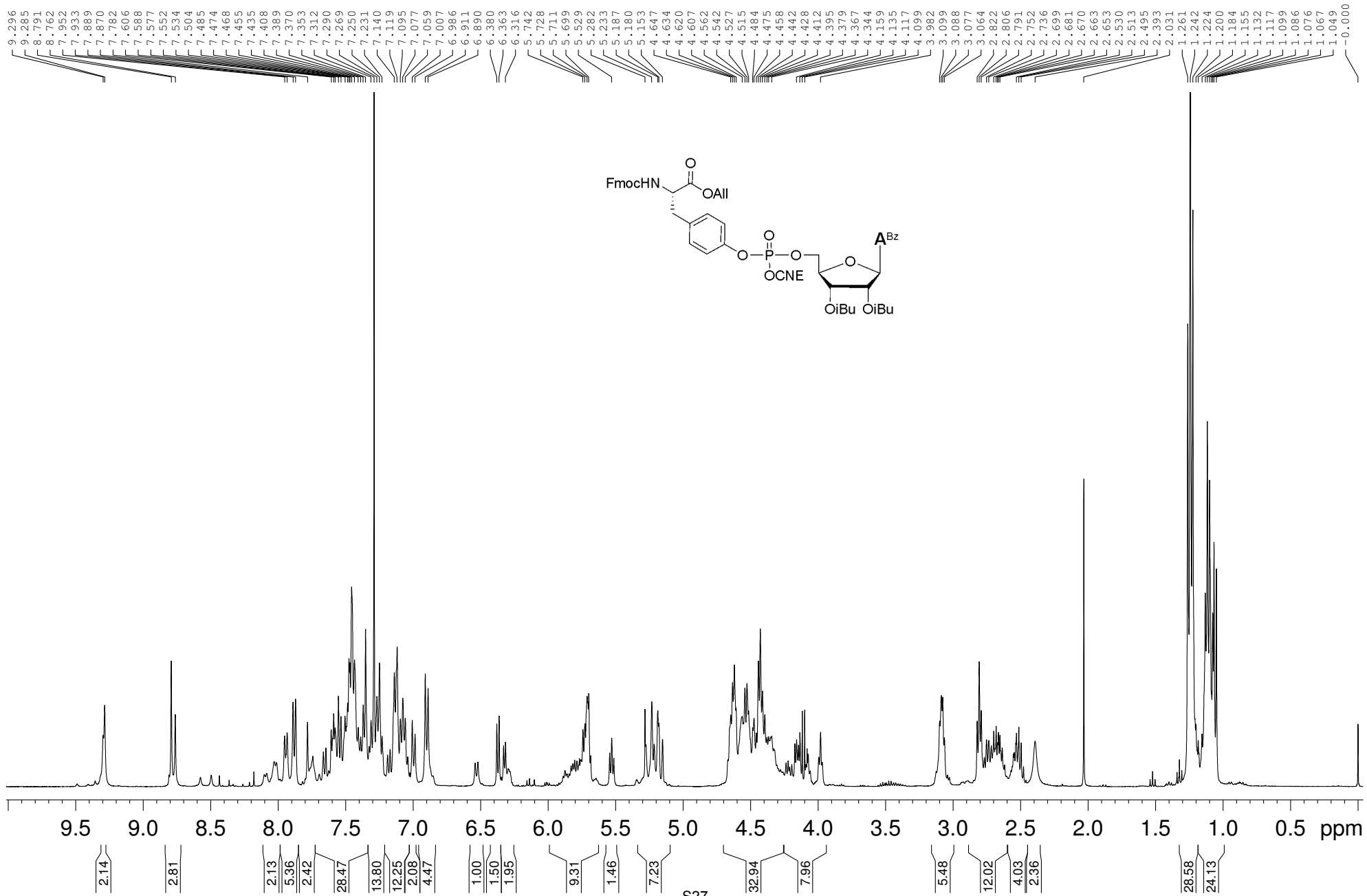


-6.97
-6.98

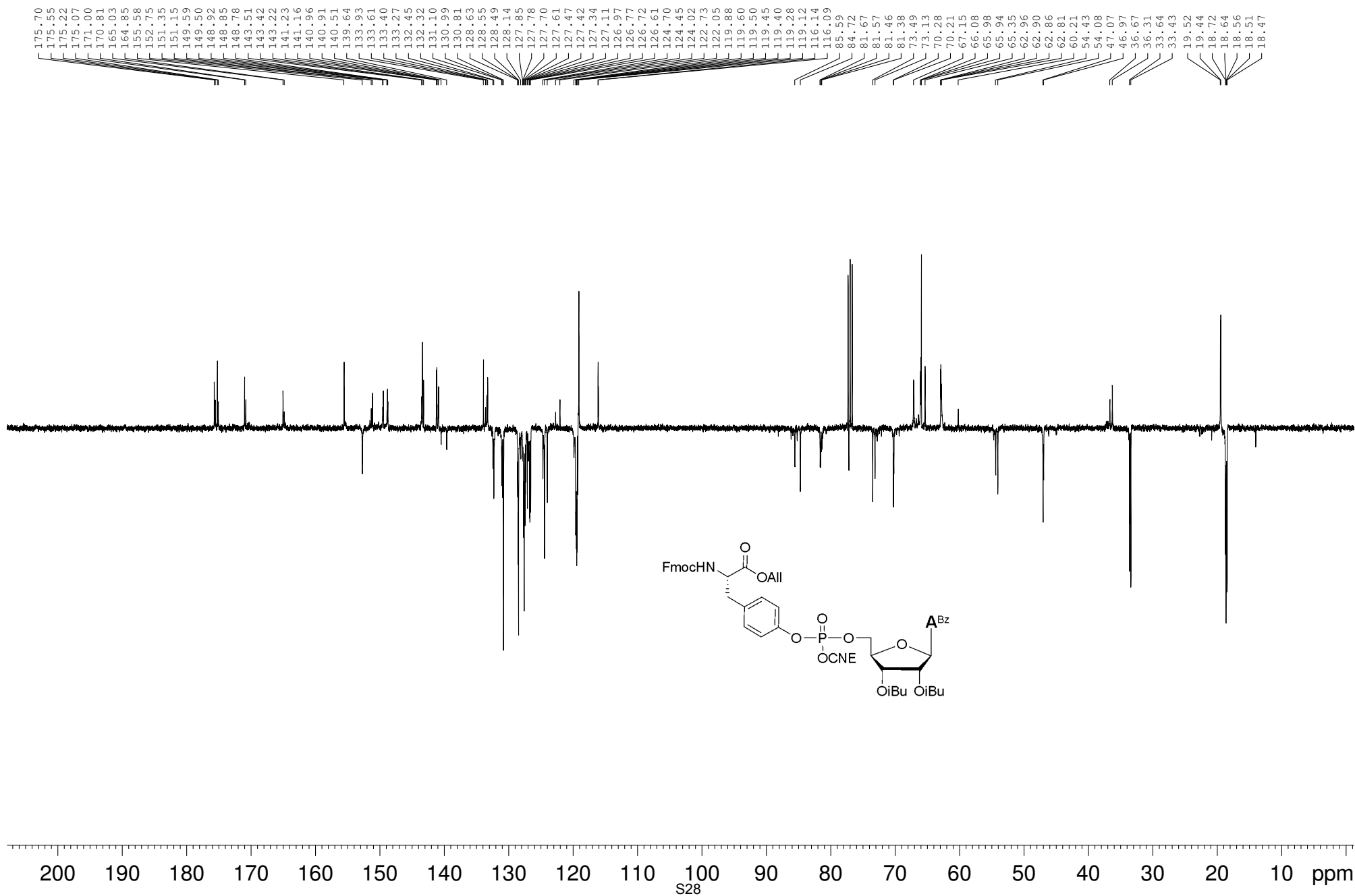


S26

¹H NMR, CDCl₃, 400 MHz, compound 11

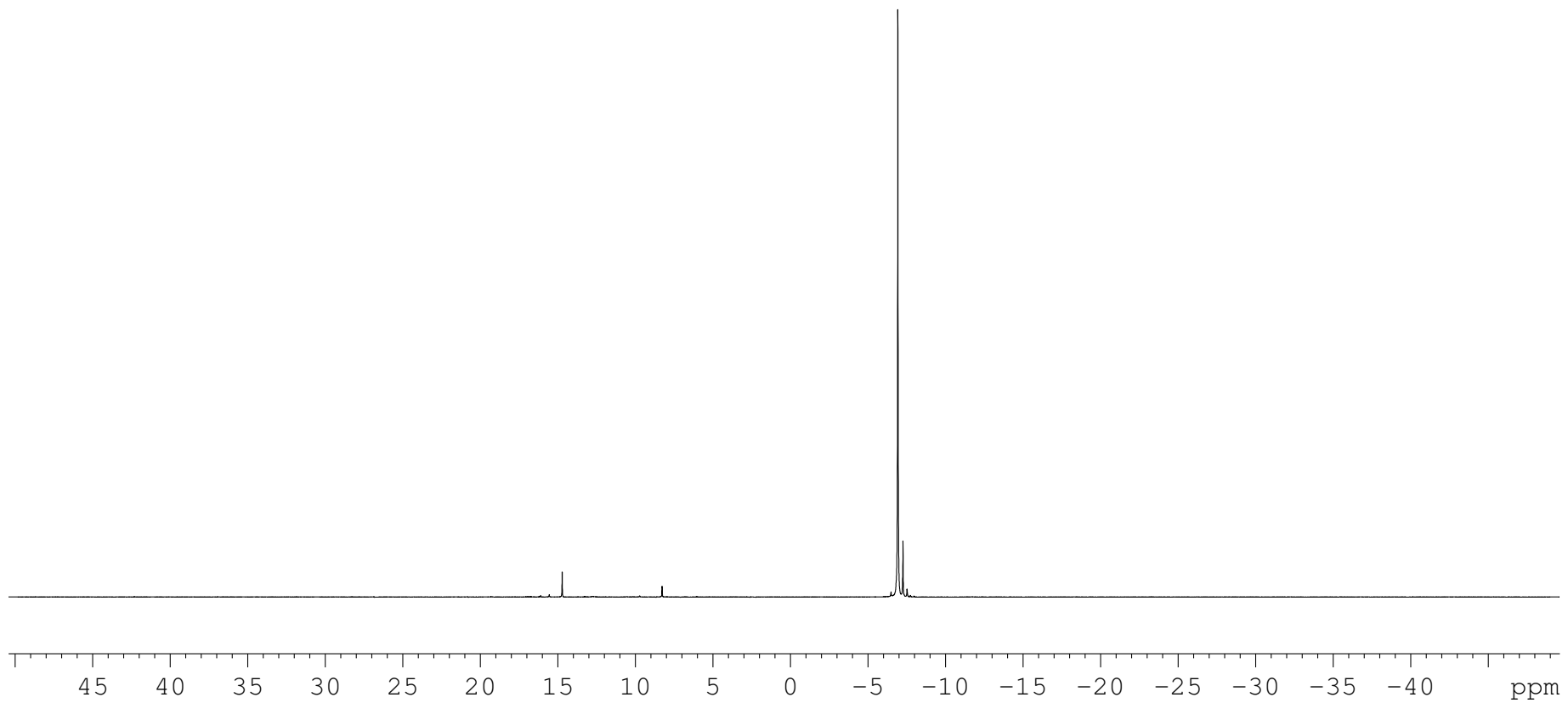
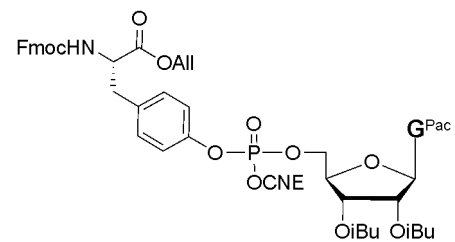


¹³C APT NMR, CDCl₃, 100 MHz, compound 11

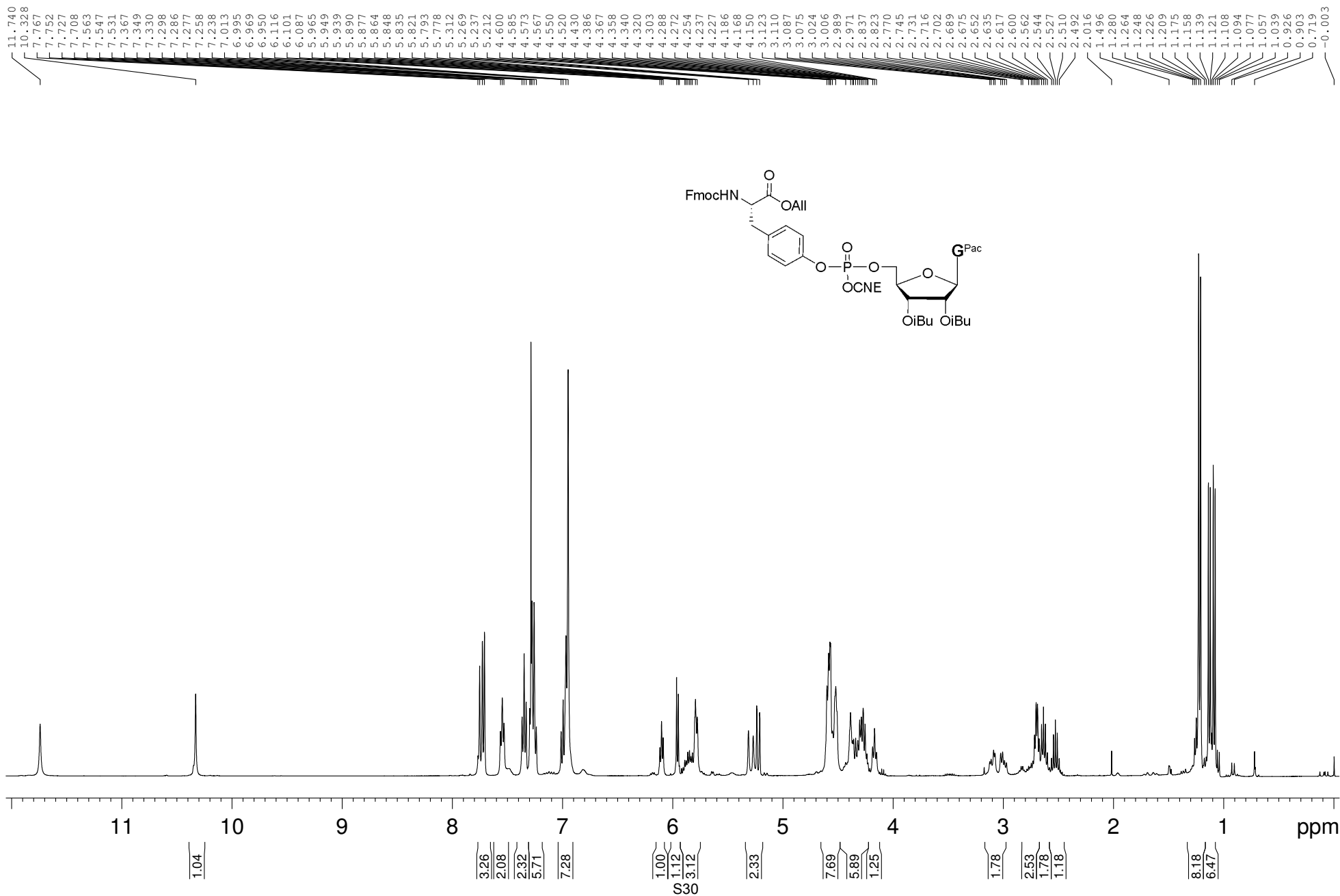


31P NMR, CDCl3, 161 MHz, compound 12A

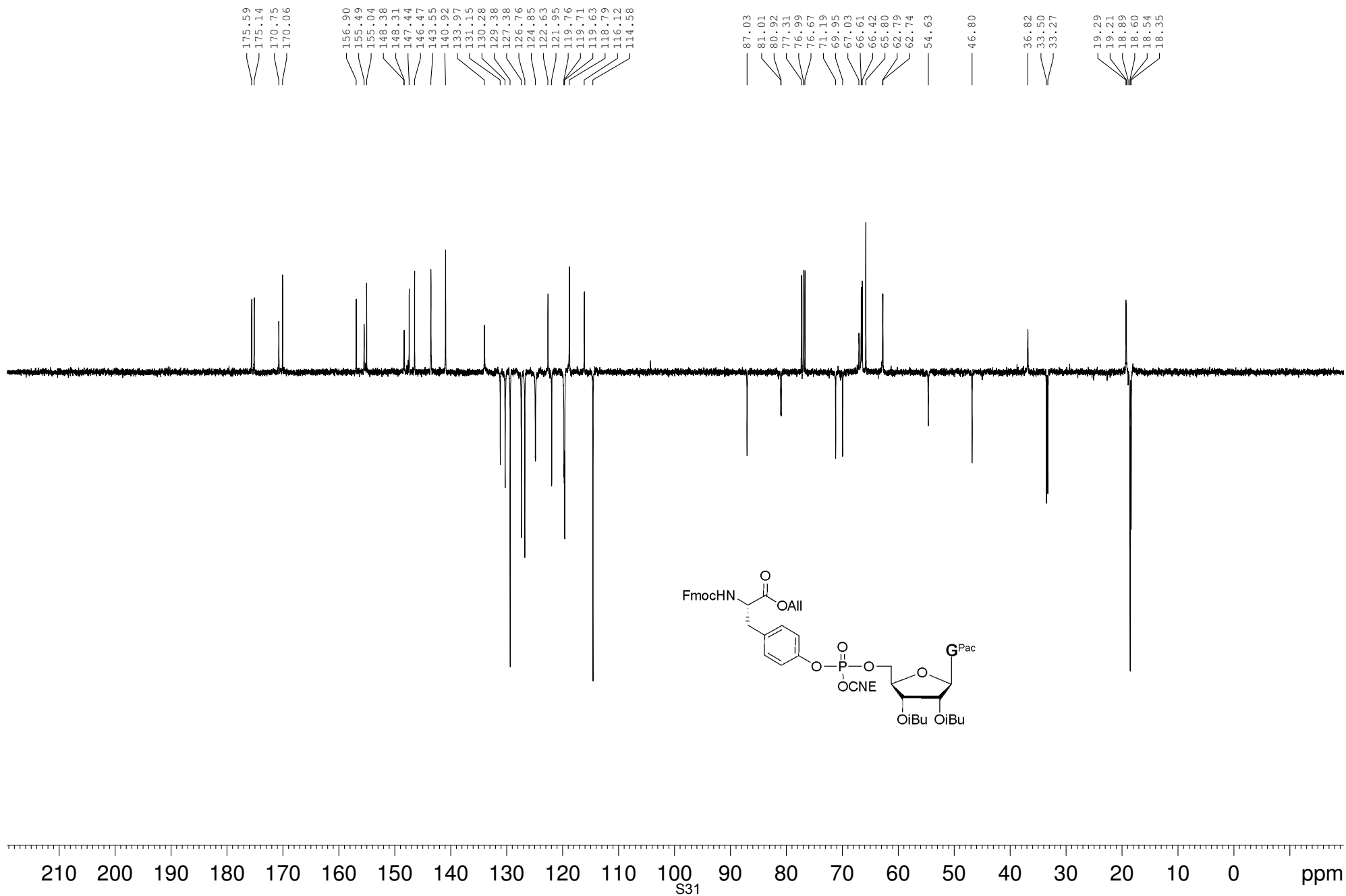
—6.935



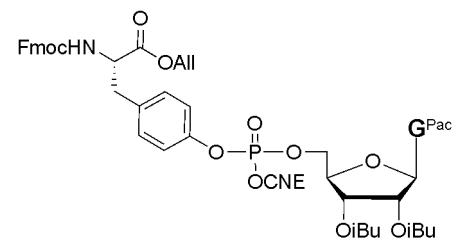
1H NMR, CDCl3, 400 MHz, compound 12A



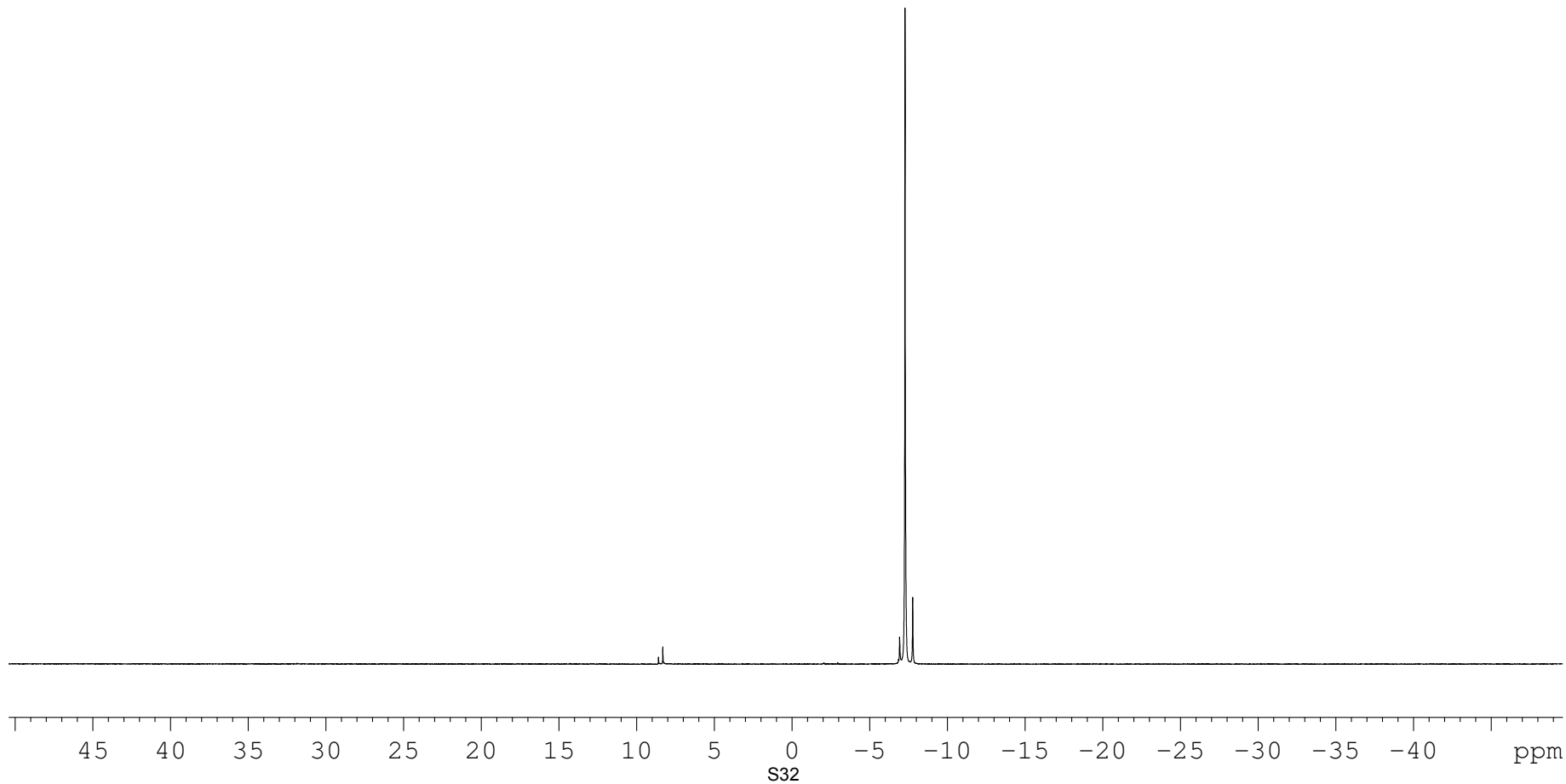
¹³C APT NMR, CDCl₃, 100 MHz, compound 12A



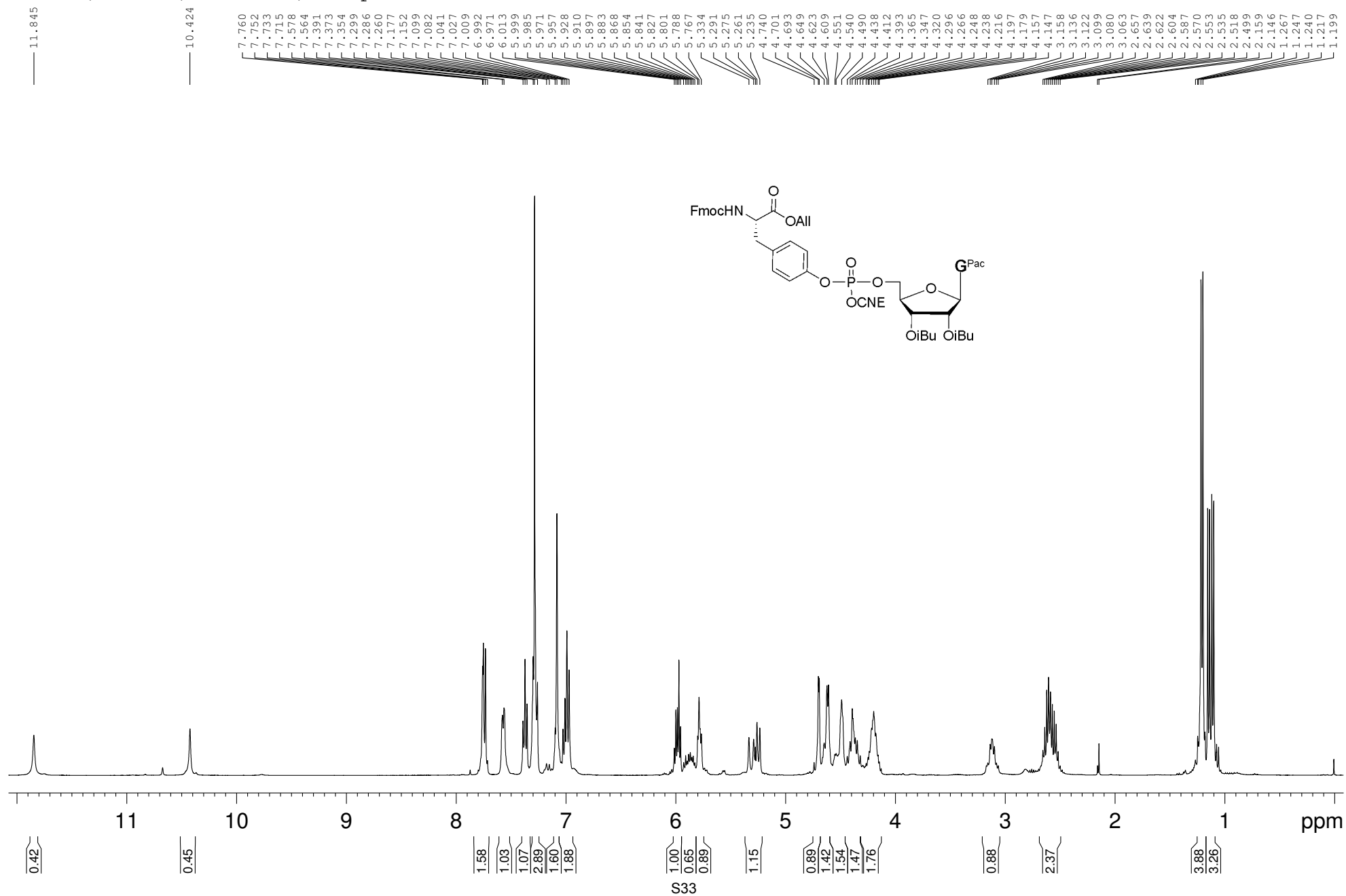
31P NMR, CDCl3, 161 MHz, compound 12B



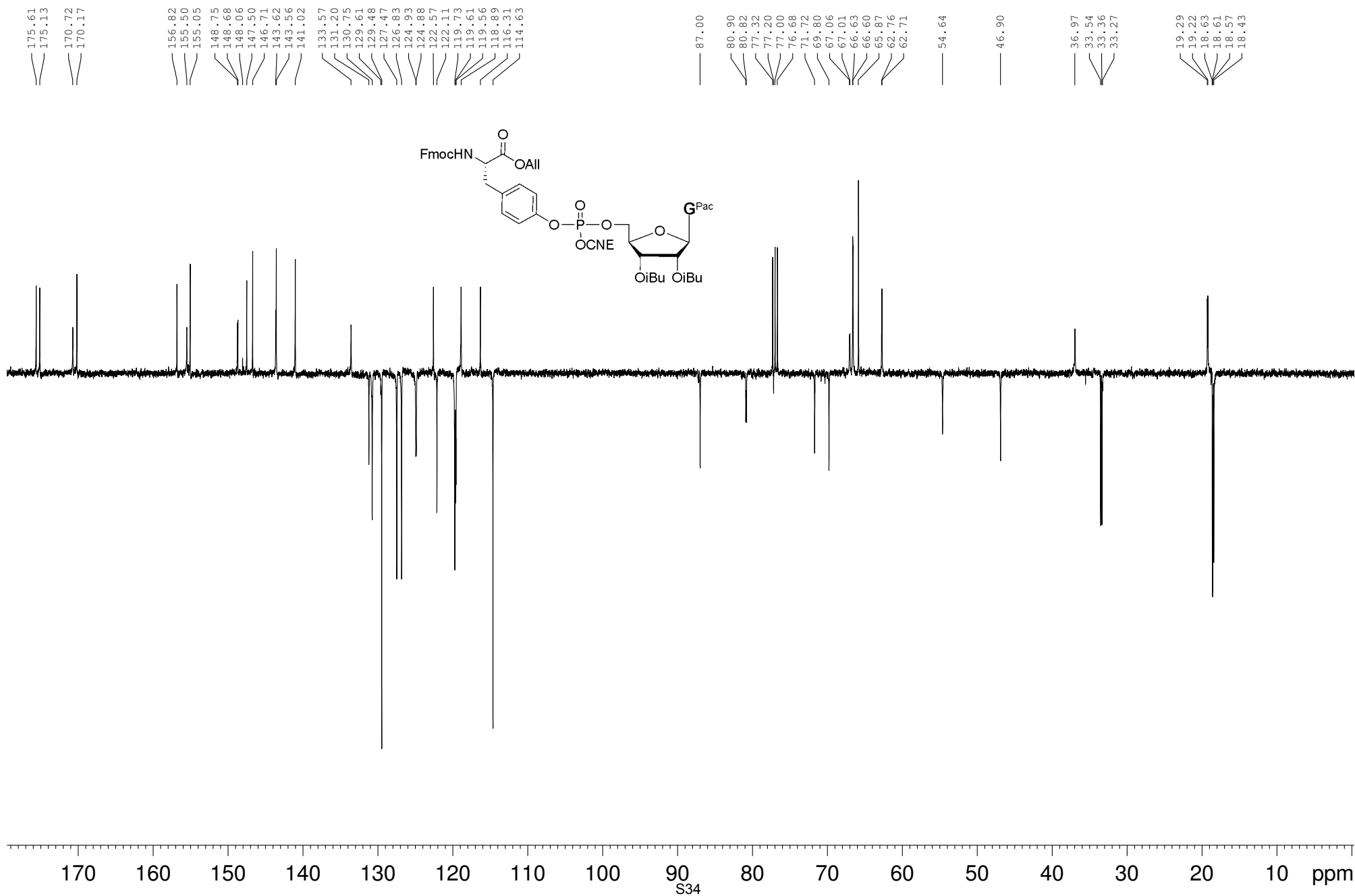
---7.281



¹H NMR, CDCl₃, 400 MHz, compound 12B

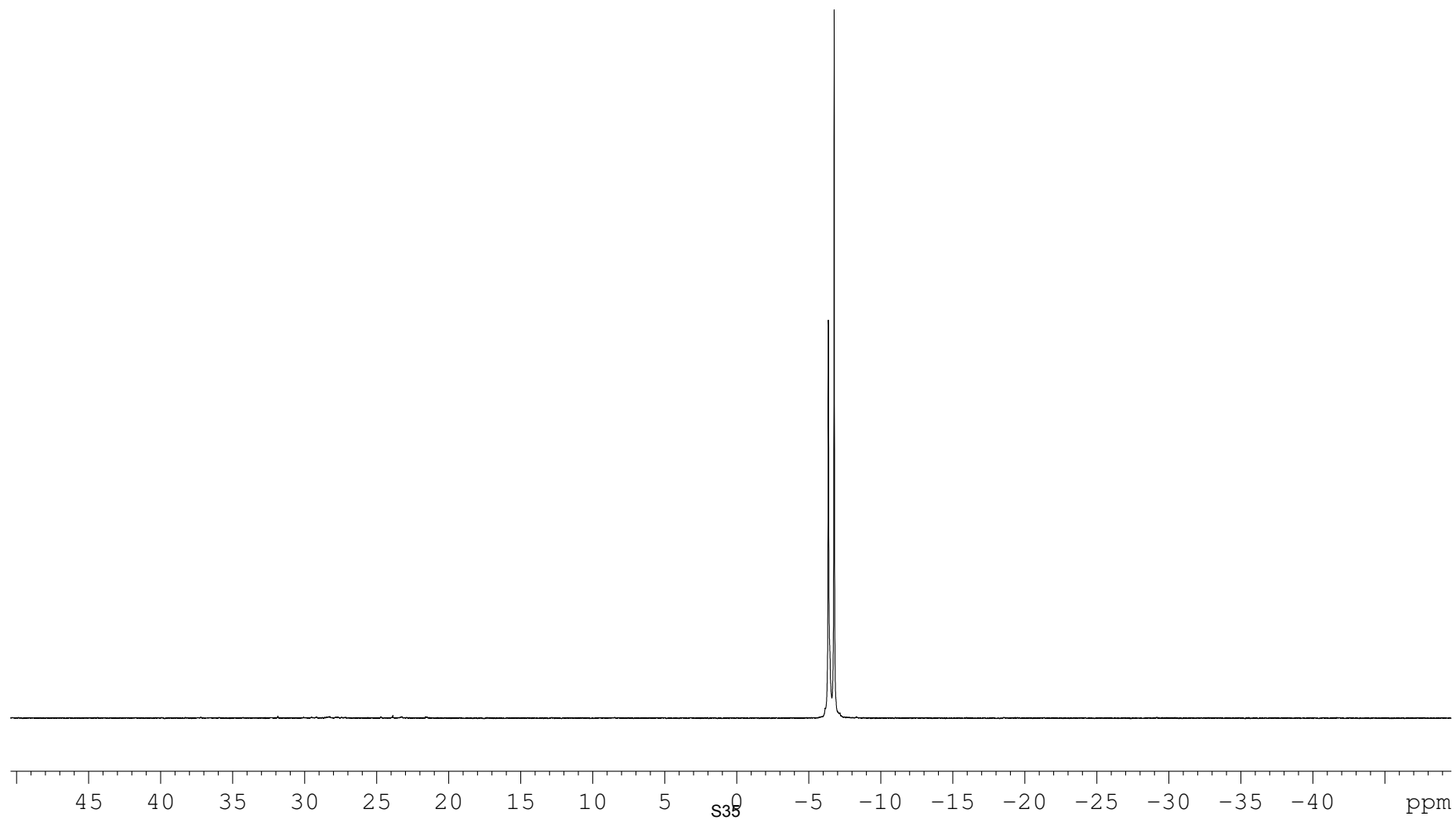
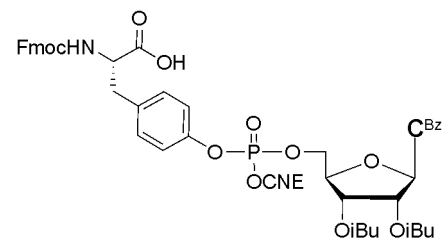


¹³C APT NMR, CDCl₃, 100 MHz, compound 12B

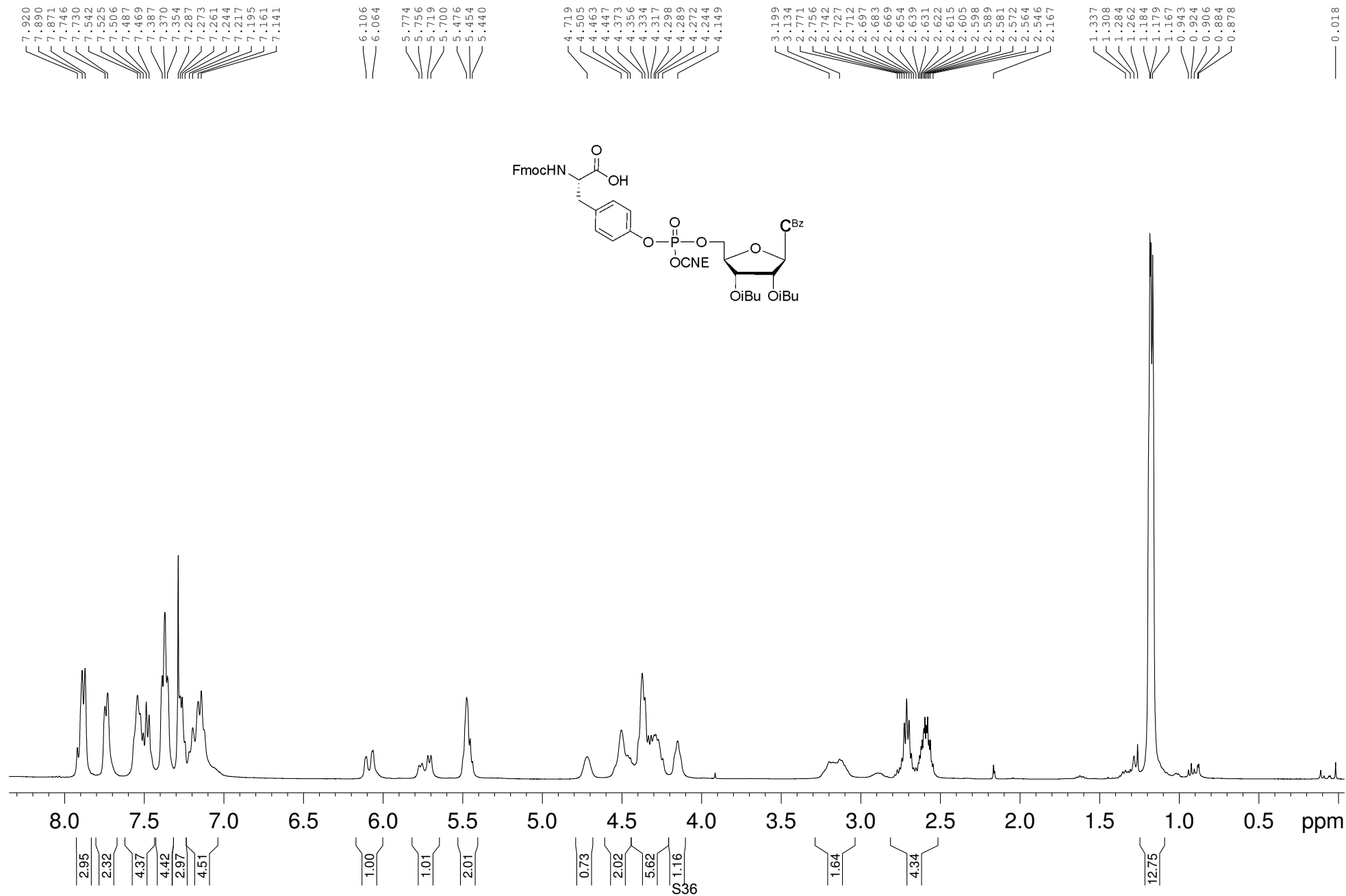


31P NMR, CDCl3, 161 MHz, compound 13

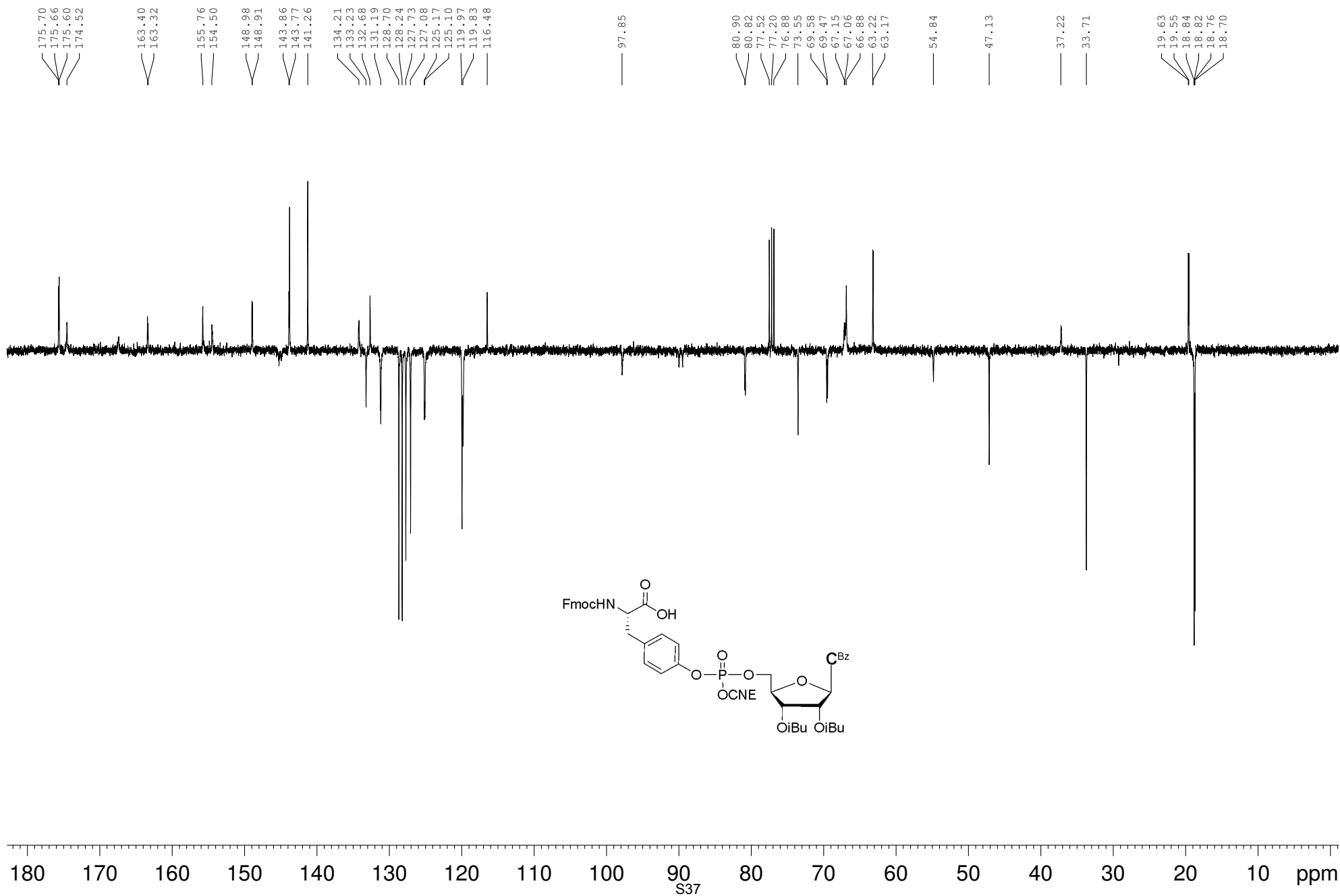
6.373
6.775



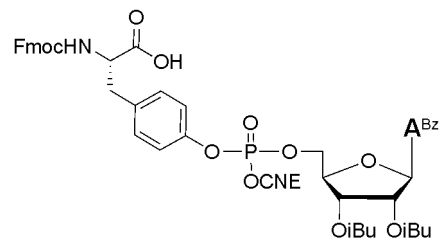
¹H NMR, CDCl₃, 400 MHz, compound 13



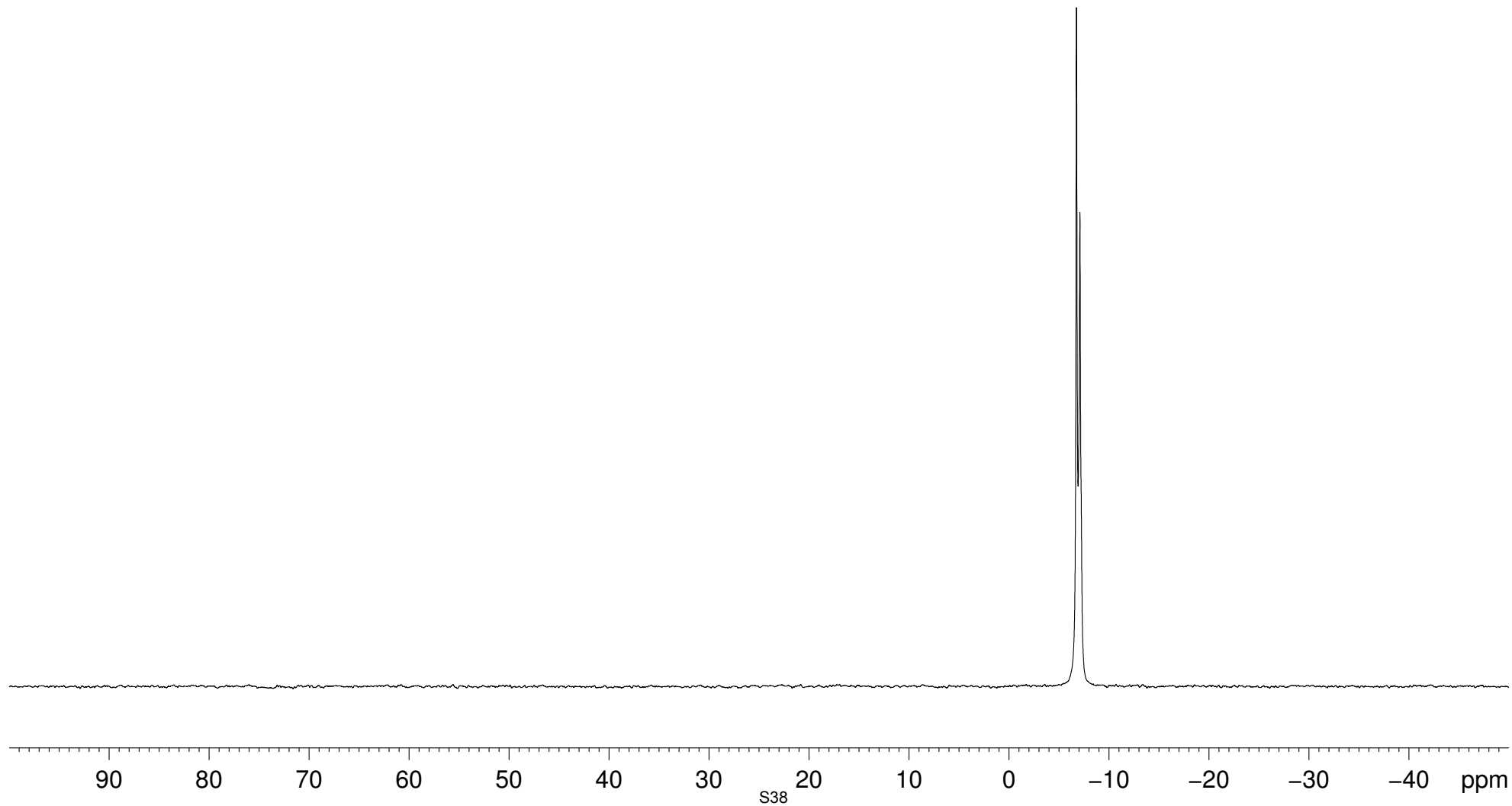
¹³C APT NMR, CDCl₃, 100 MHz, compound 13



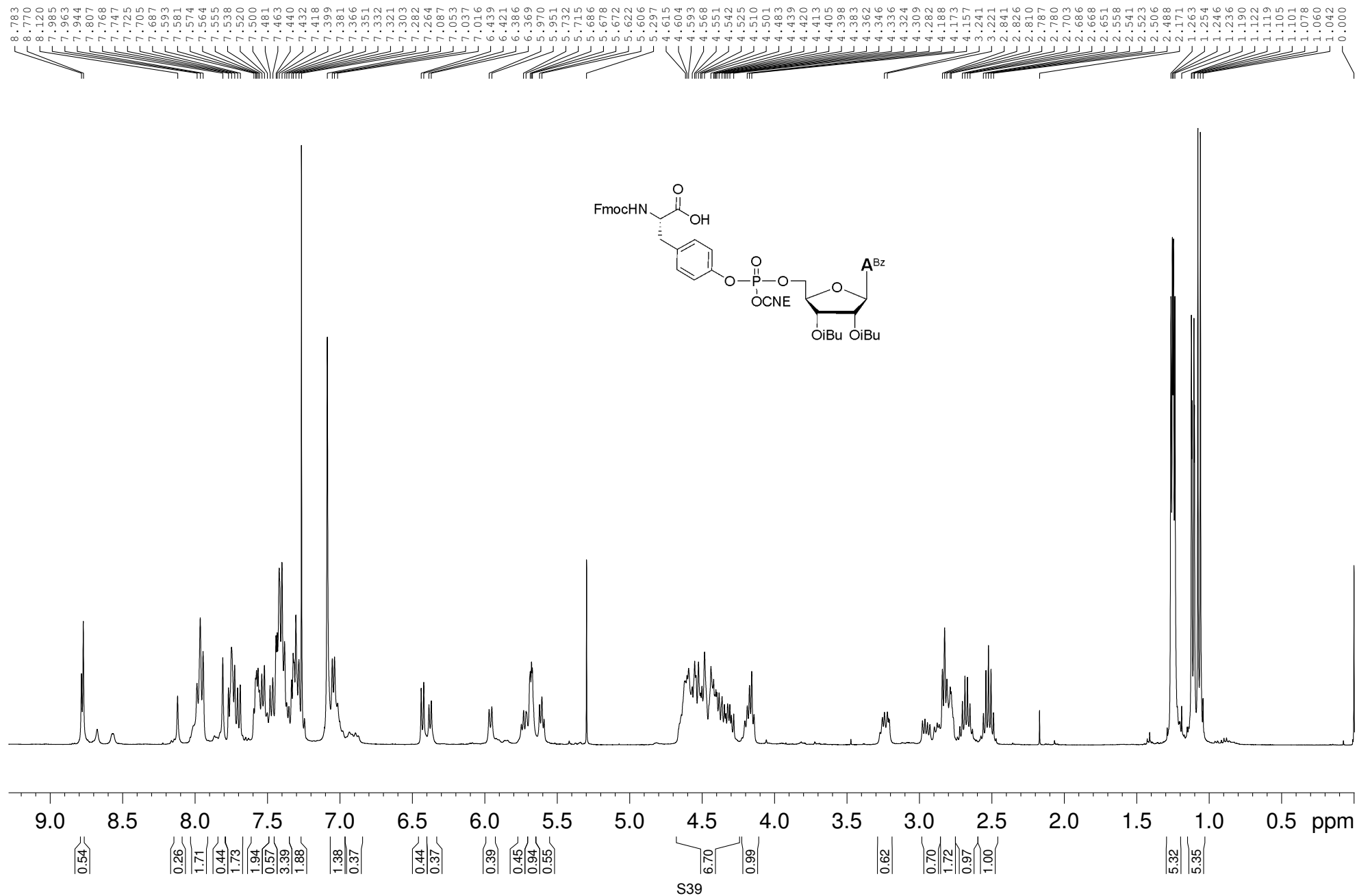
31P-NMR, CDCl3, 161 MHz, compound 14



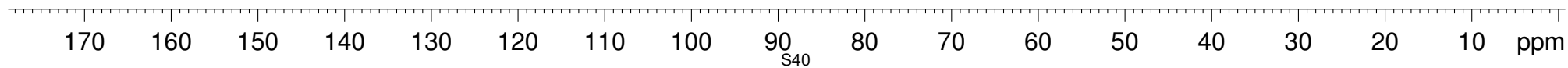
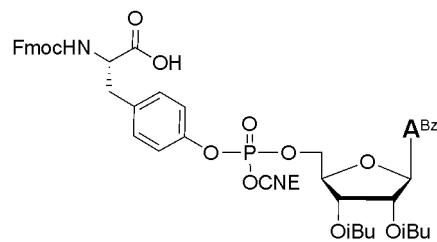
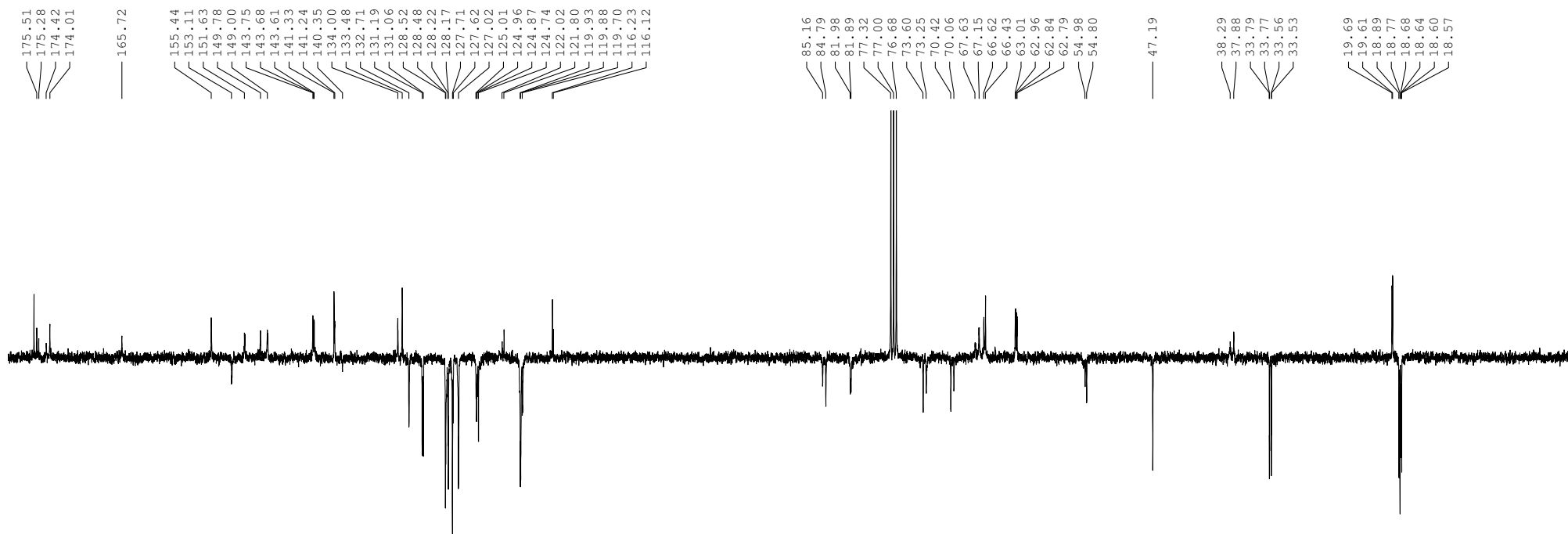
-6.75
-7.09



¹H NMR, CDCl₃, 400 MHz, compound 14

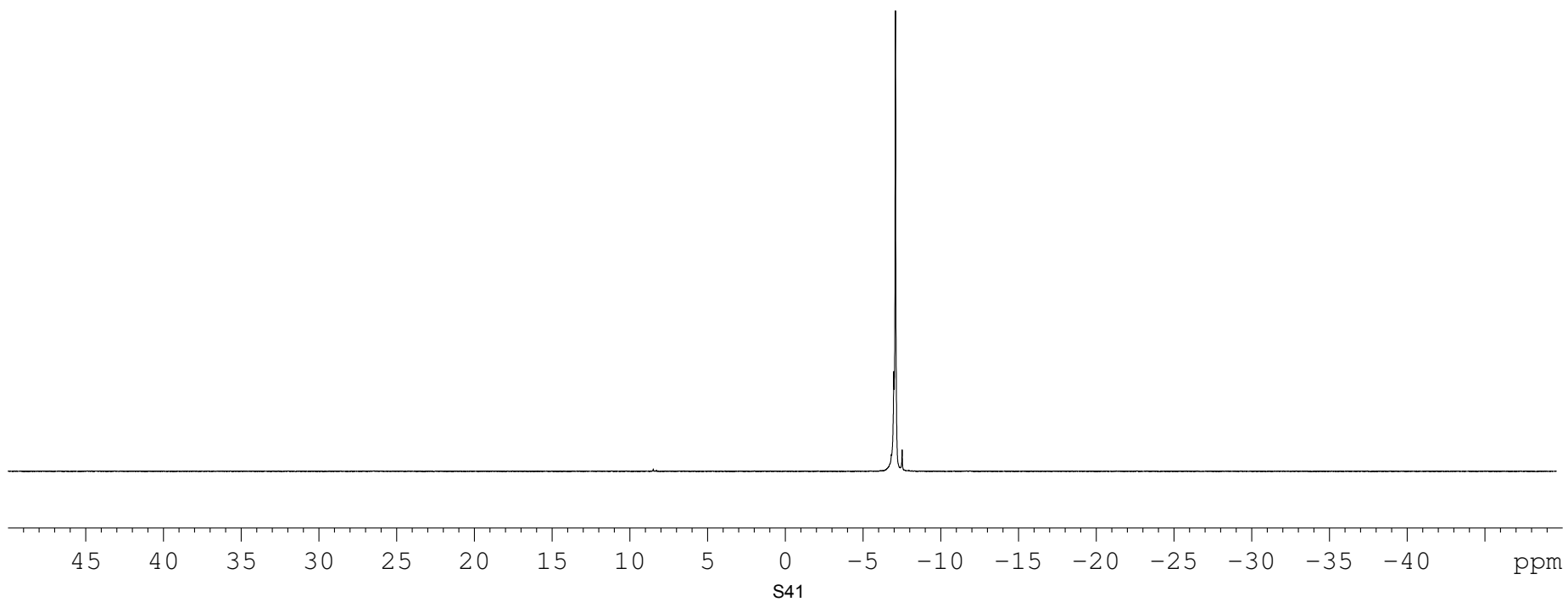
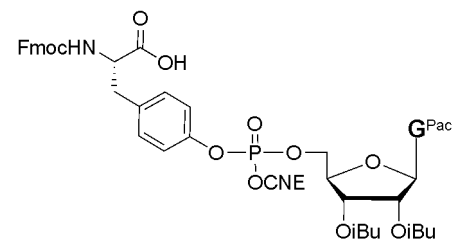


¹³C APT NMR, CDCl₃, 100 MHz, compound 14

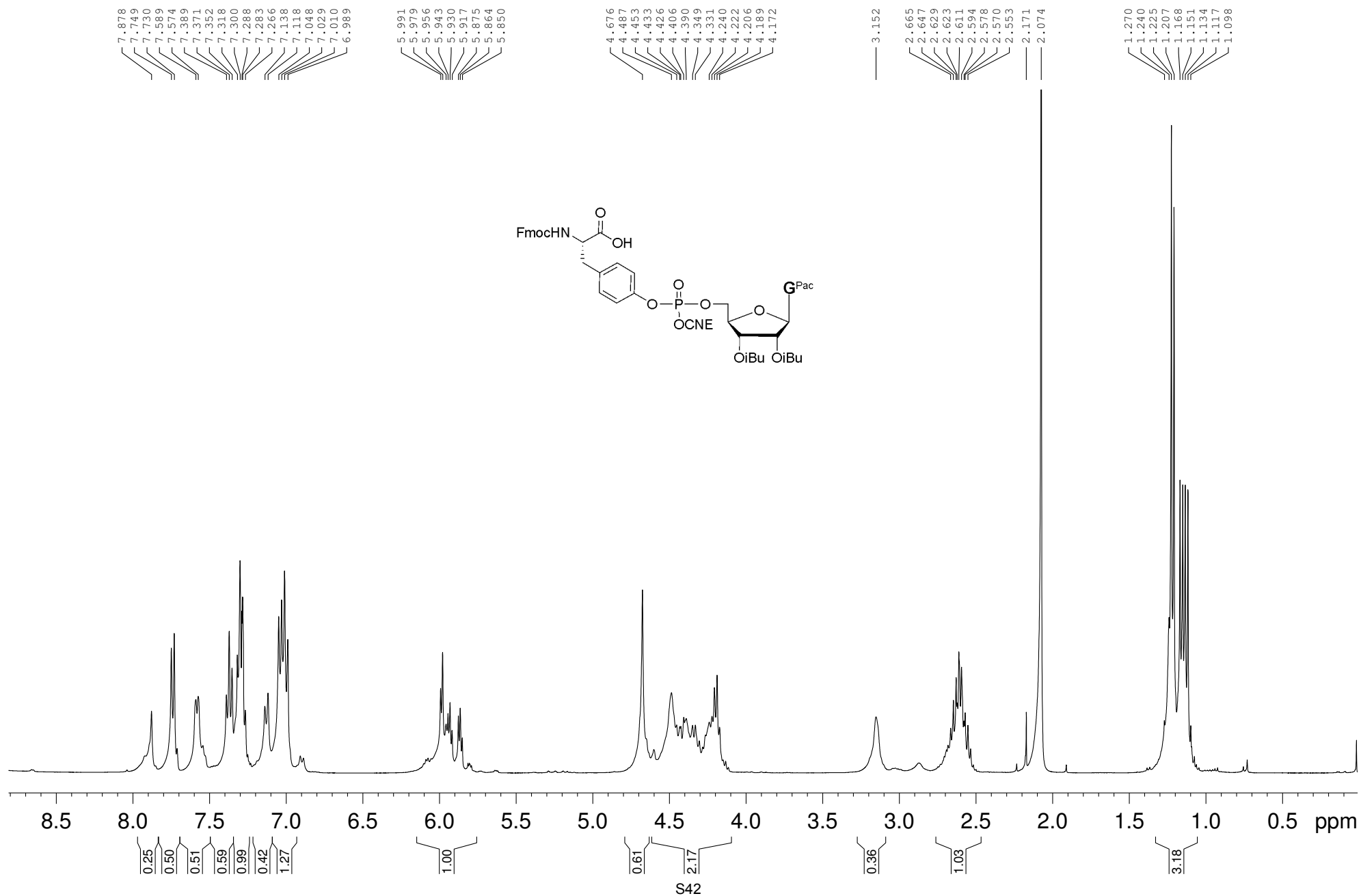


31P NMR, CDCl3, 161 MHz, compound 15

6.988
7.094



¹H NMR, CDCl₃, 400 MHz, compound 15



13C APT NMR, CDCl3, 100 MHz, compound 15

176.17
175.93
173.96
170.34

157.12
156.81
155.92
155.36
148.85
148.77
148.45
147.75
146.94
143.92
143.81
141.26

134.18
131.20
130.71
129.81
129.72
127.71
127.10
125.23
122.59
122.35
121.64
119.94
119.68
119.64
116.57
114.90

87.05

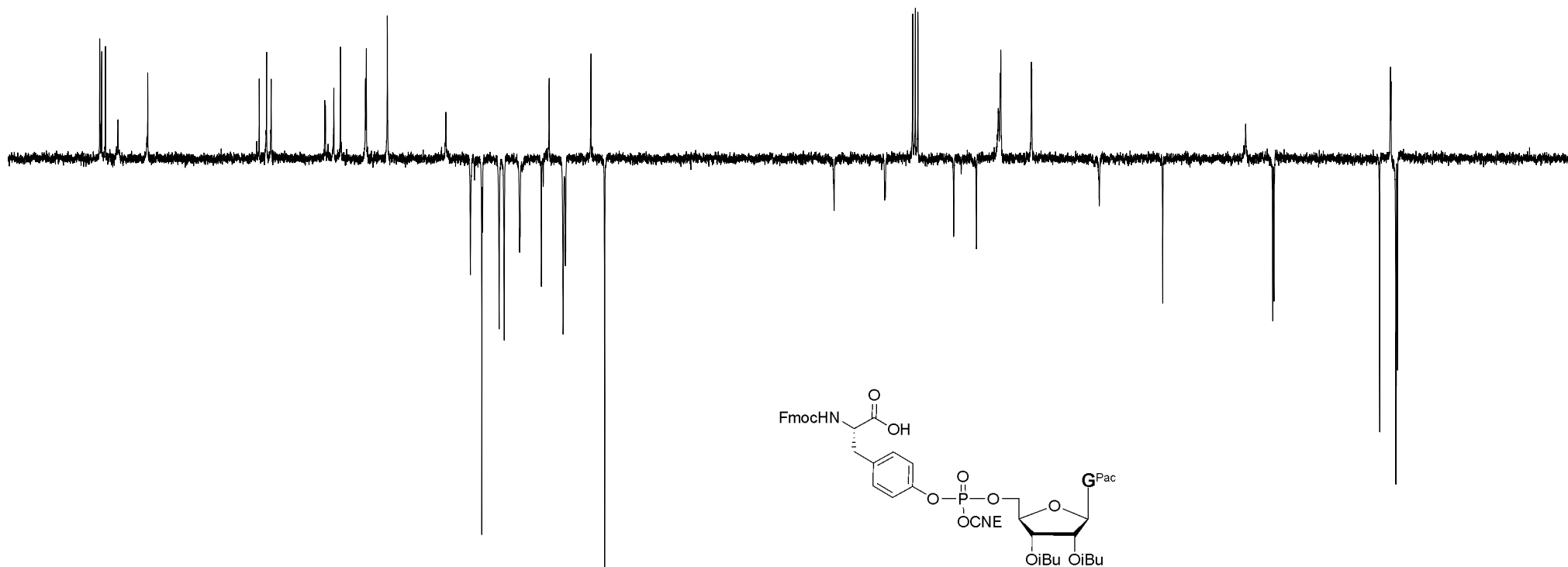
80.91
80.84
77.51
77.39
76.88
72.53
71.62
69.77
67.12
66.89
66.81
63.11
63.06

54.86

47.16

37.10
33.80
33.66

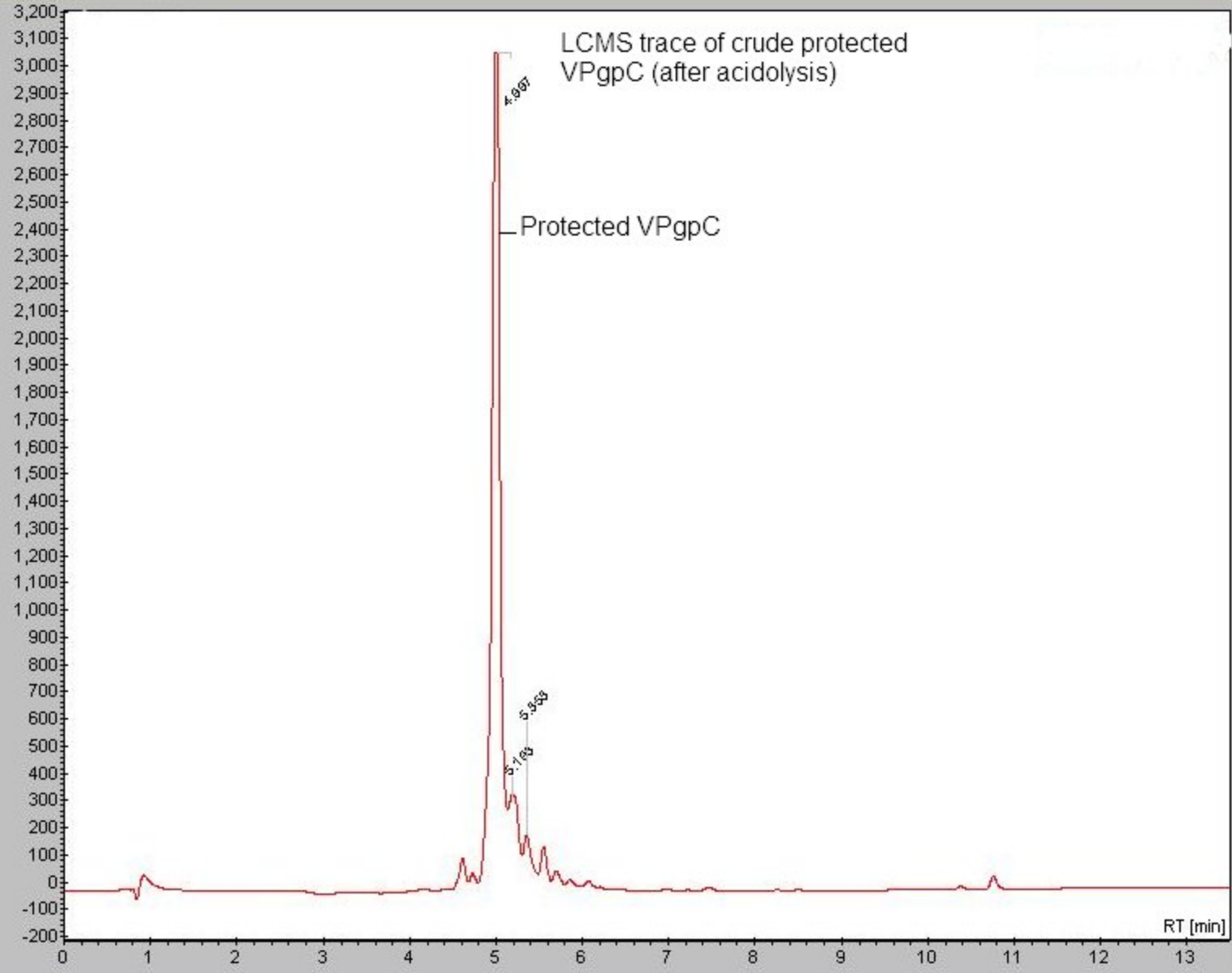
20.83
19.51
19.44
18.86
18.84
18.81
18.66



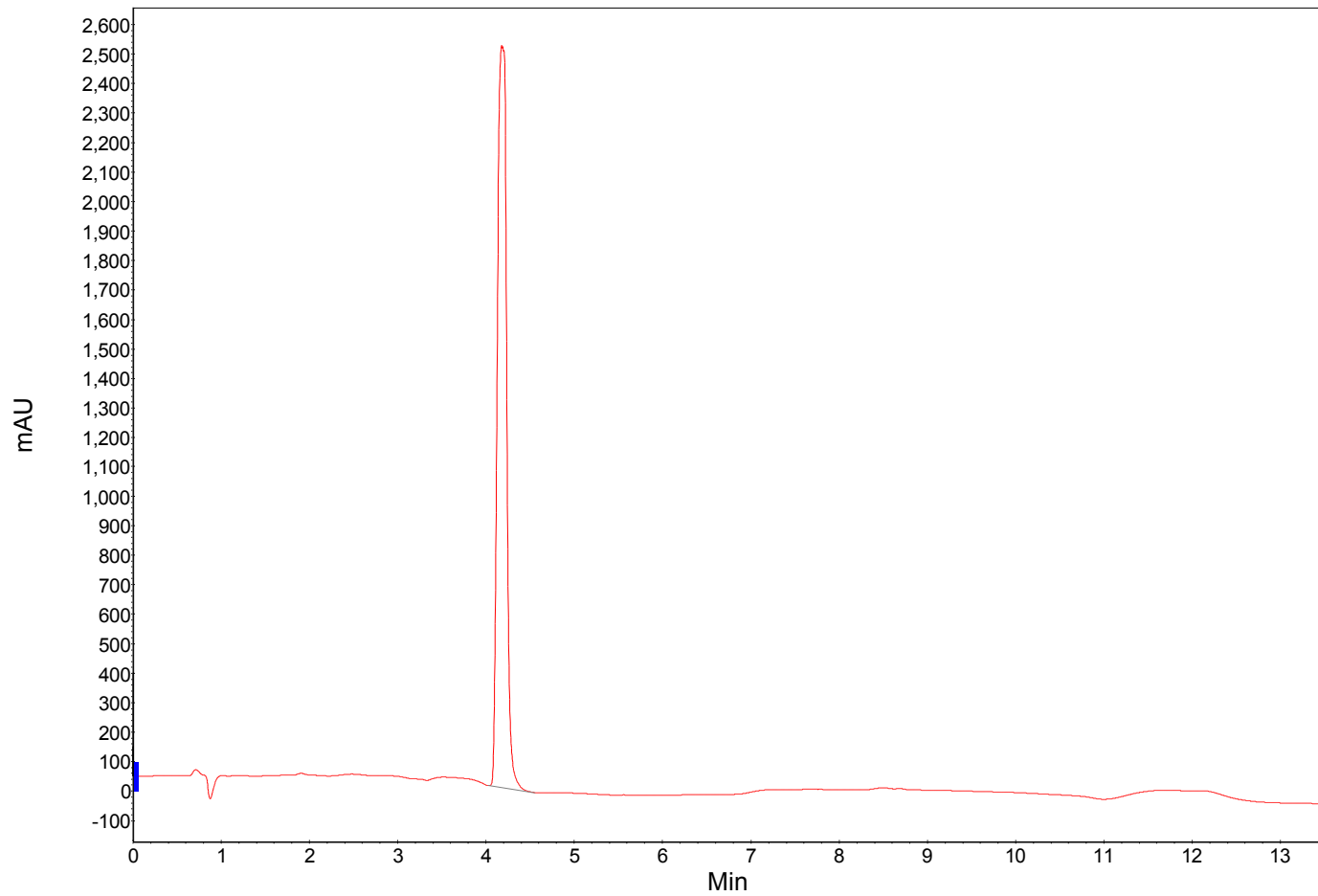
180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 ppm

S43

LCMS trace of crude protected
VPgpC (after acidolysis)

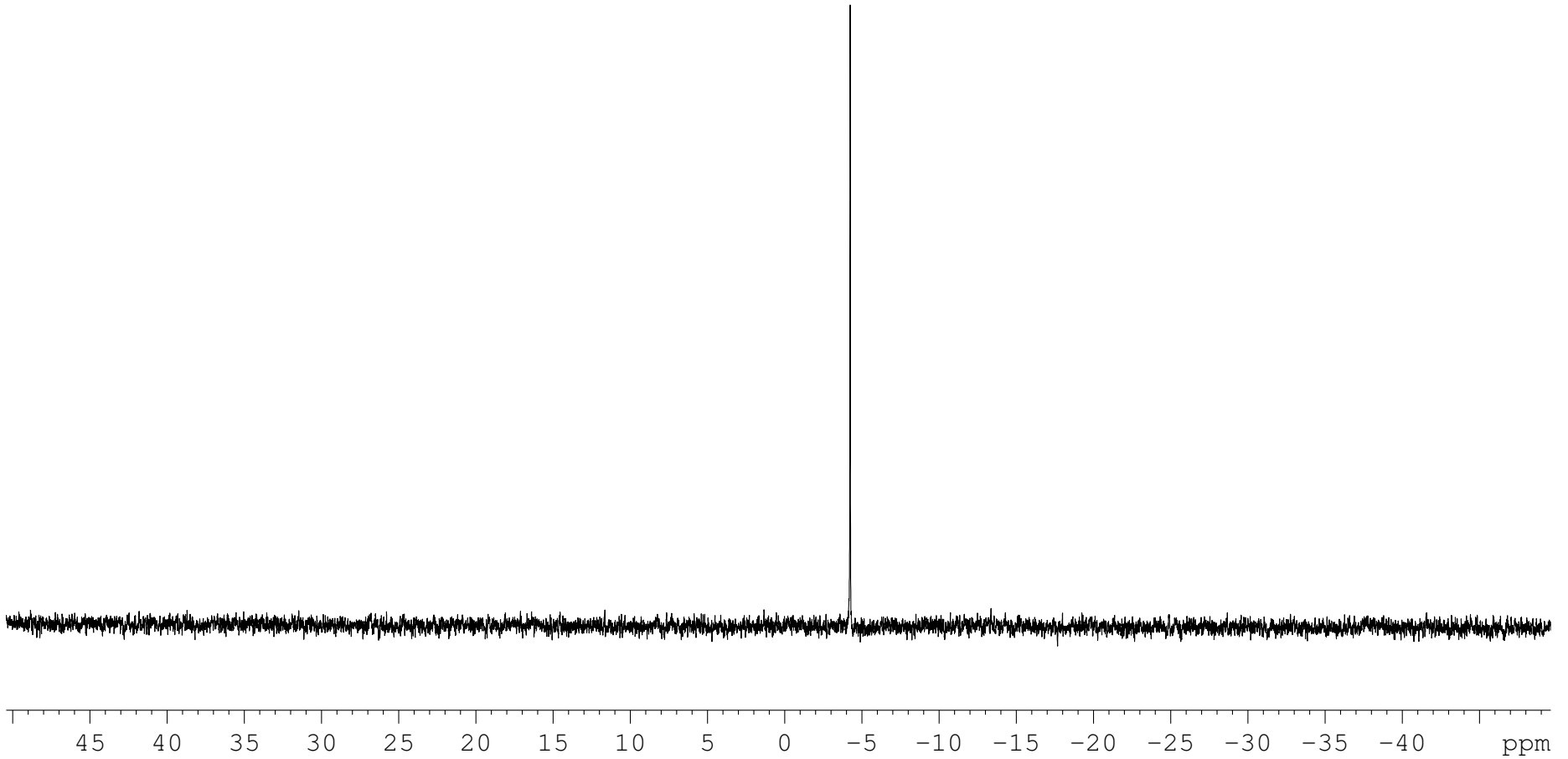


LCMS trace of purified deprotected VPgpC 20



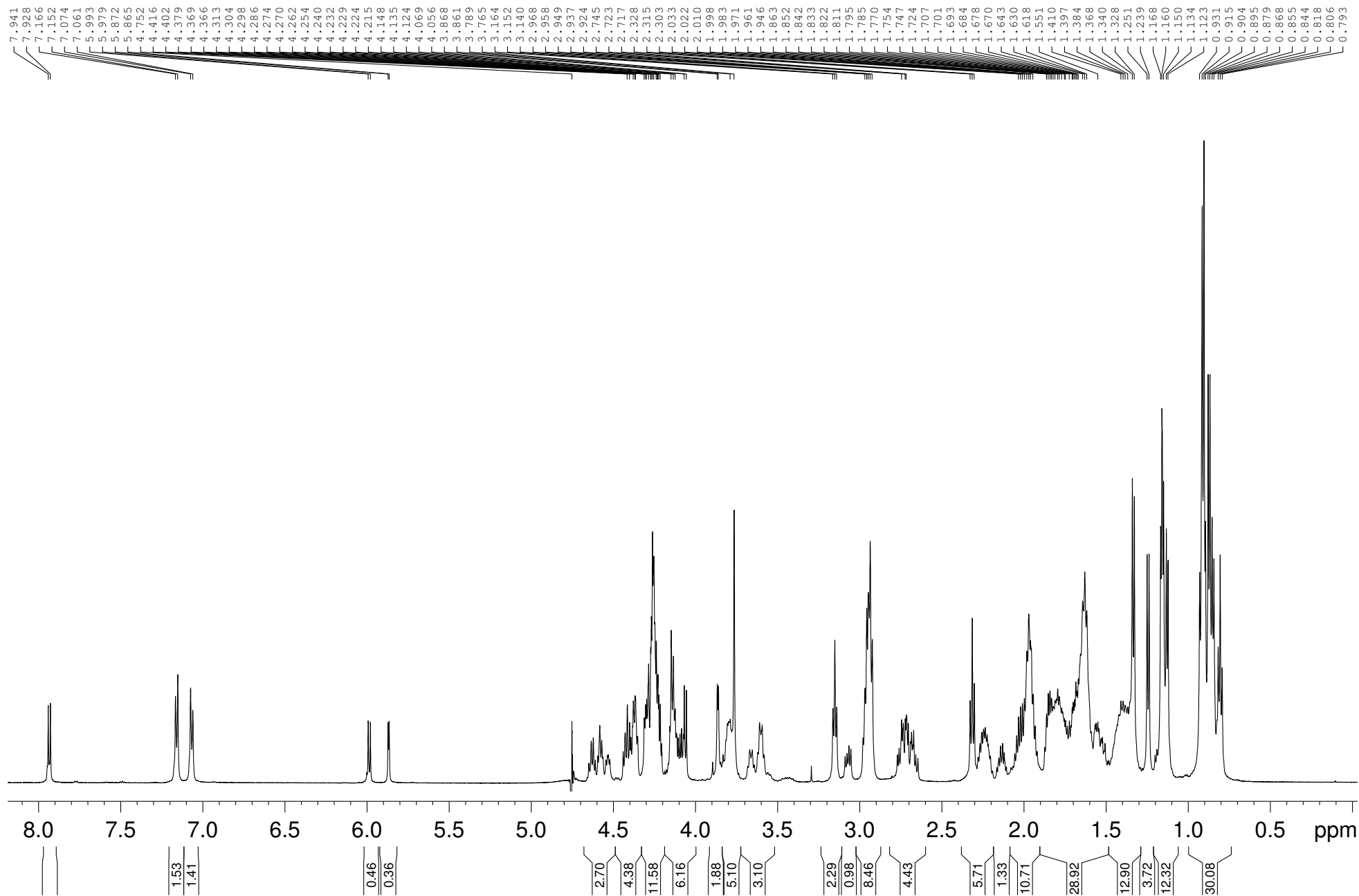
31P NMR, D2O, 161 MHz, compound 20 - VPgpC

— -4.243



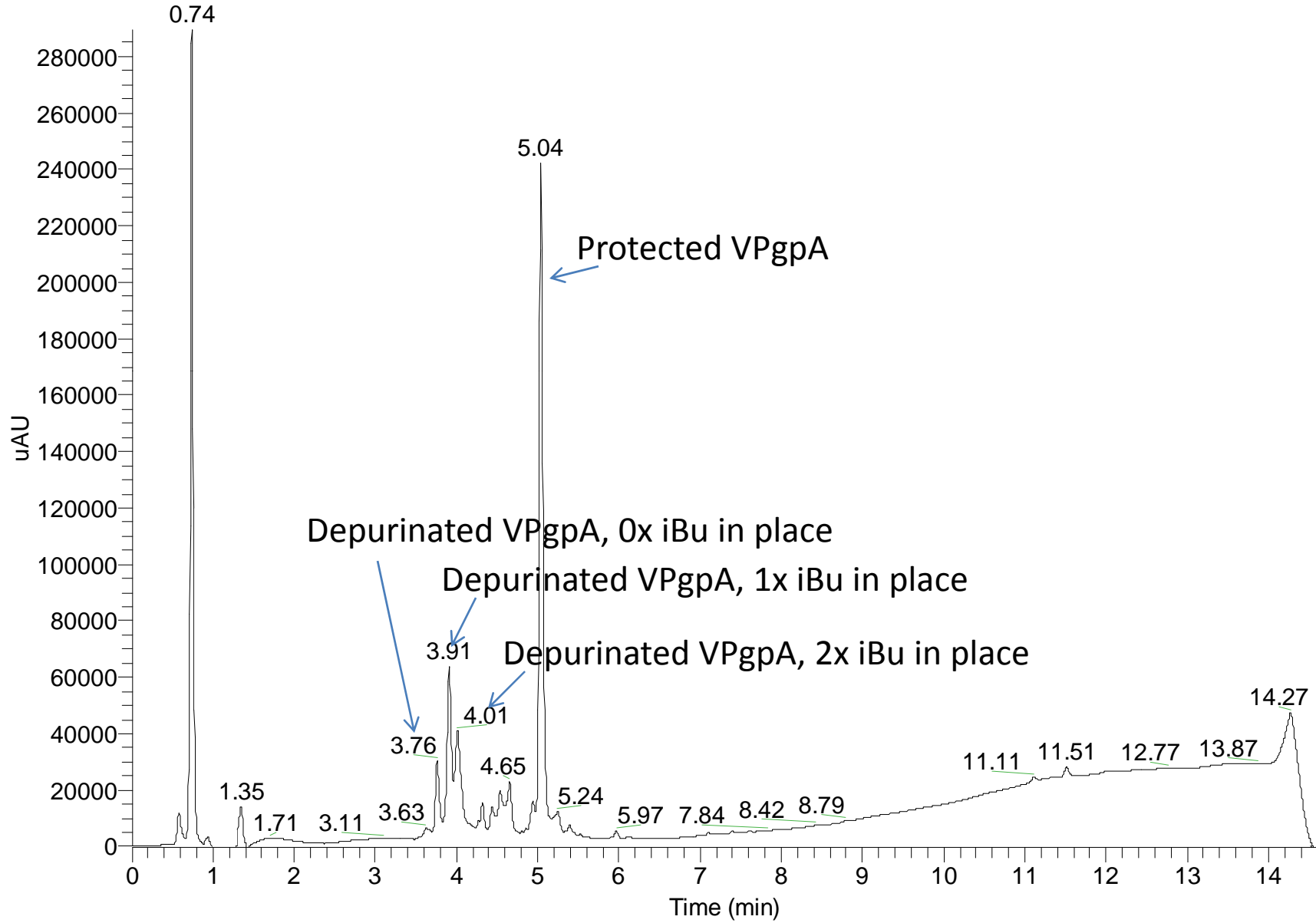
S46

1H NMR, 400 MHz, D2O, compound 20 - VPgpC

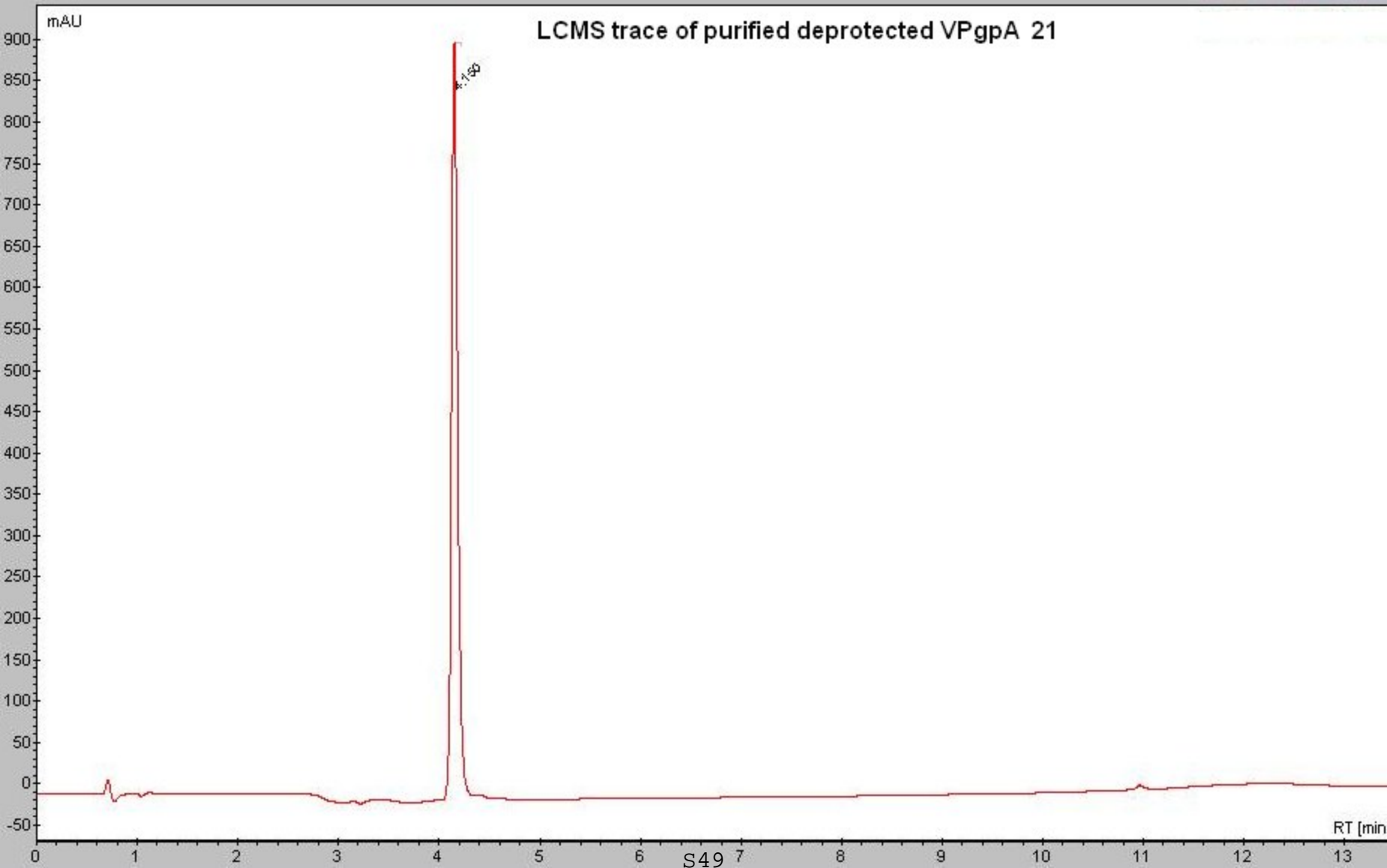


LCMS trace of crude protected VPgpA (after acidolysis)

RT: 0.00 - 14.60



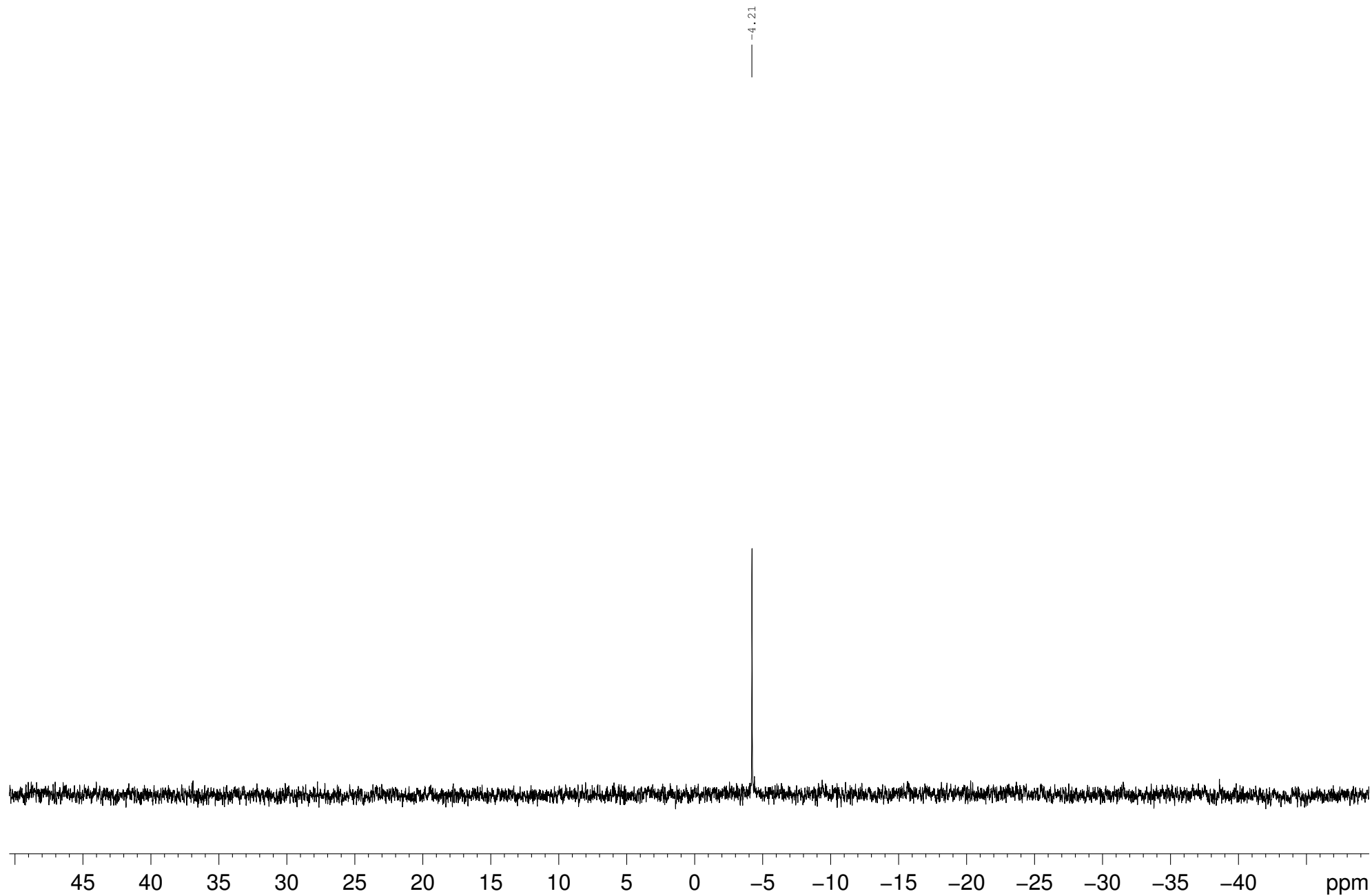
LCMS trace of purified deprotected VPgpA 21



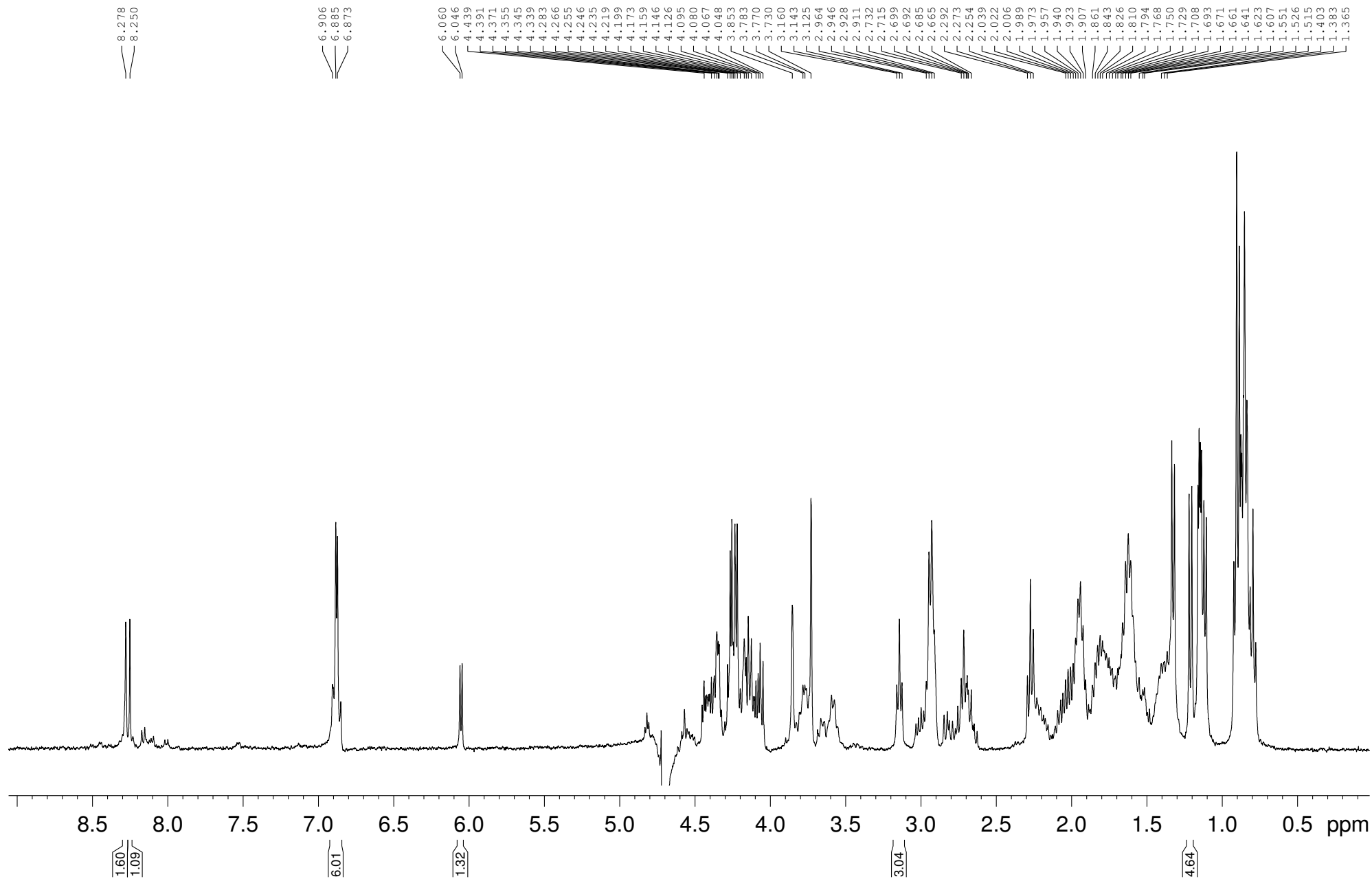
S49

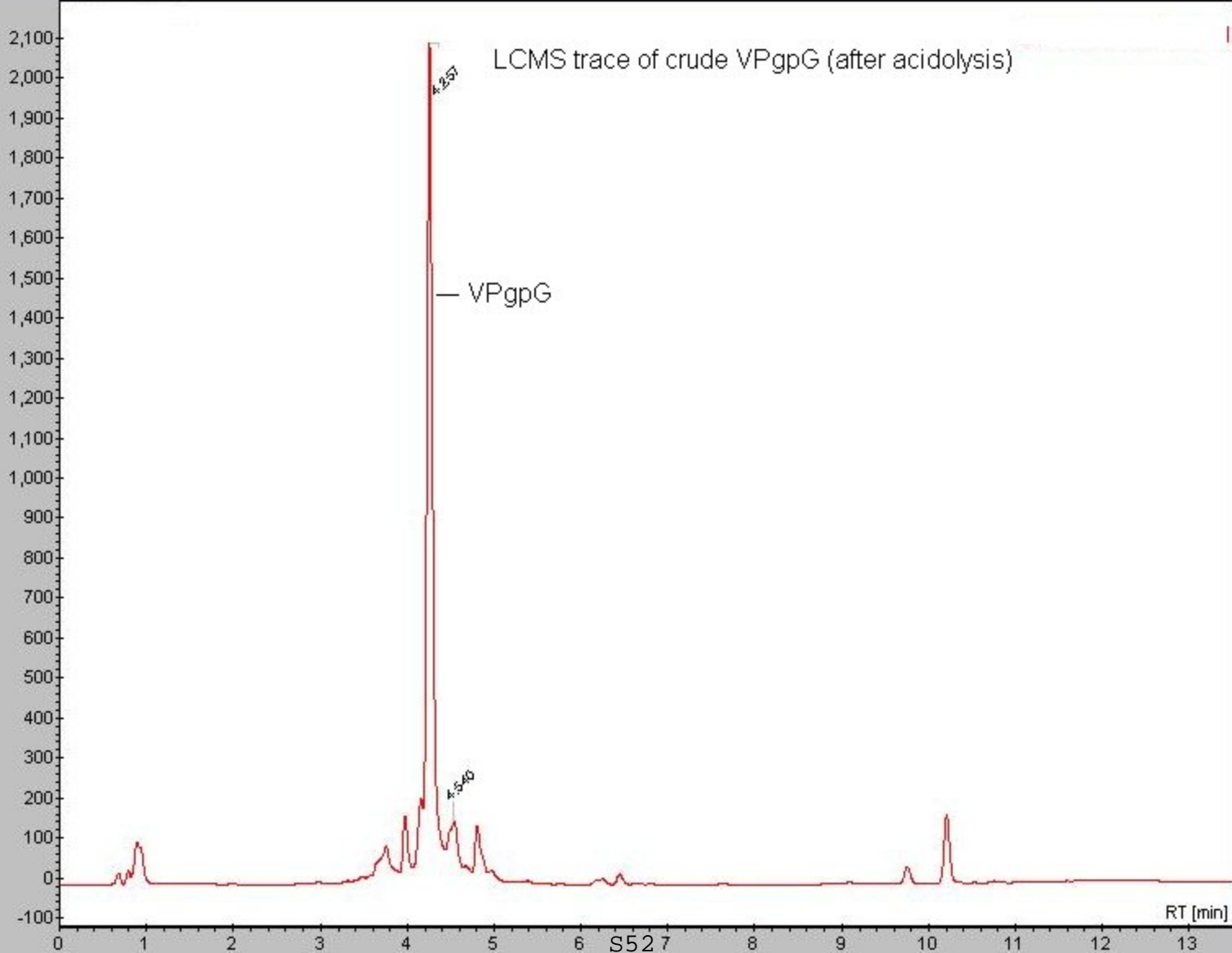
RT [min]

31P-NMR, D2O, 161 MHz, compound 21 - VPgpA

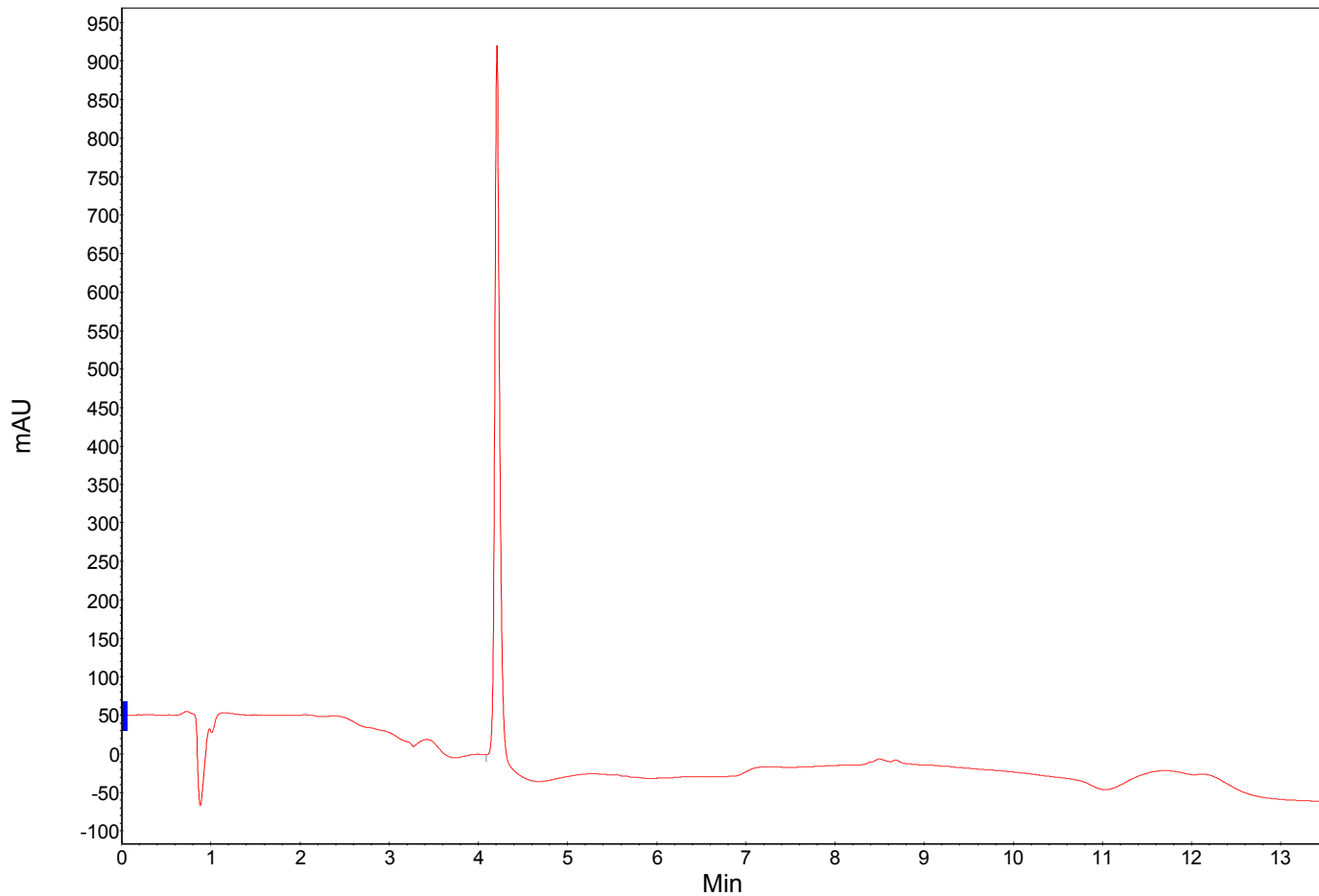


1H NMR, D2O, 400 MHz, compound 21 - VPgpA



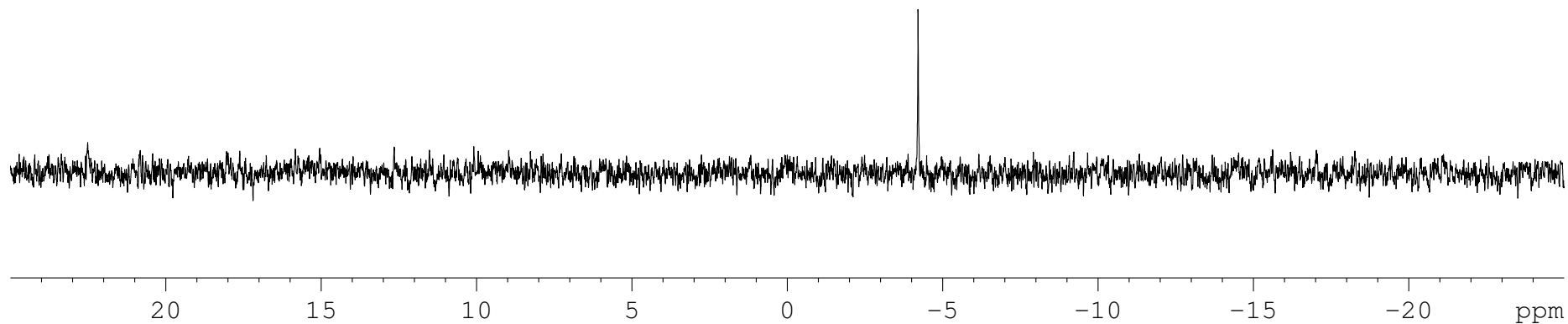


LCMS trace of purified deprotected VPgpG **22**



^{31}P NMR, D_2O , 161 MHz, compound 22 - VPgpG

— -4.209



S54

1H NMR, D2O, 600 Mhz, compound 22 - VPgpG

