Electronic Supplementary Information (ESI) for

Chemically modified Adeno-associated virus capsid as a promising targeting alternative for gene delivery

Mathieu Mével,[a,*] Mohammed Bouzelha,[a] Aurélien Leray,[a,c] Simon Pacouret,[a] Mickäel Guilbaud,[a] Magalie Penaud-Budloo,[a] Dimitri Alvarez-Dorta,[c] Laurence Dubreil,[b] Sébastien G. Gouin,[c] Jean Philippe Combal,[d] Mirja Hommel,[e] Gloria Gonzalez-Aseguinolaza,[d,e] Véronique Blouin,[a] Philippe Moullier,[a] Oumeya Adjali,[a] David Deniaud,[c,*] and Eduard Ayuso,[a,*].

a. INSERM UMR 1089, Université de Nantes, CHU de Nantes, 44200 Nantes, France. Email: mathieu.mevel@univ-nantes.fr and eduard.ayuso@univ-nantes.fr.

b. PanTher-UMR 703, INRA-ONIRIS, Atlanpole-Chanterie, 44307 Nantes, France.

c. LUNAM Université, CEISAM, Chimie Et Interdisciplinarité, Synthèse, Analyse, Modélisation, UMR CNRS 6230, UFR des Sciences et des Techniques, 44322 Nantes, France. Email: david.deniaud@univ-nantes.fr

d. Vivet Therapeutics SAS, Paris, France.

e. Gene Therapy and Regulation of Gene Expression Program, CIMA, FIMA, University of Navarra, Navarra Institute for Health Research (IdisNA), Pamplona, Spain.

*Corresponding authors

Email address: mathieu.mevel@univ-nantes.fr, david.deniaud@univ-nantes.fr and eduard.ayuso@univ-nantes.fr.

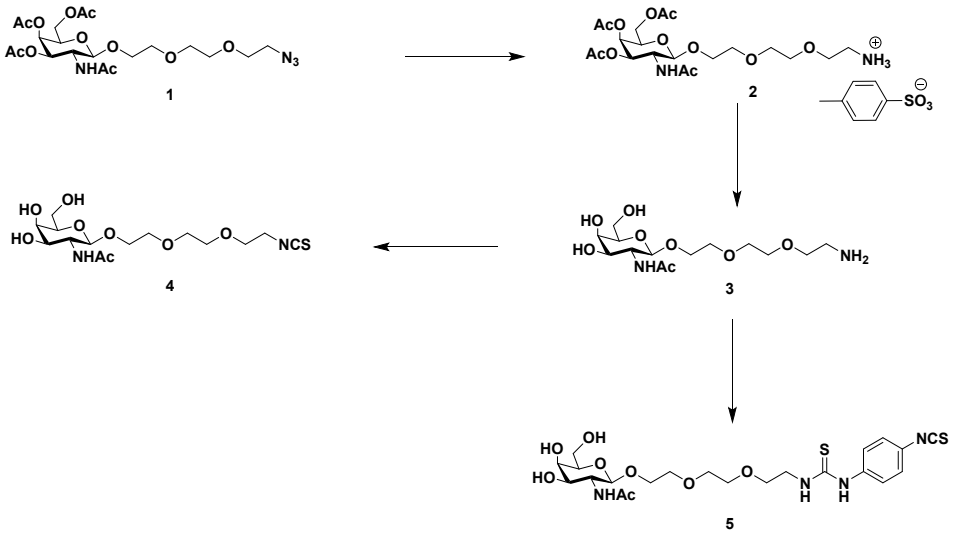
Materials

All chemical reagents were purchased from Acros Organics or Aldrich and were used without further purification. Rabbit polyclonal anti-AAV capsid proteins antibody (cat N° 61084), mouse monoclonal B1 (cat N° 61058) and mouse monoclonal A20 antibody (cat N° 61055) were obtained from PROGEN Biotechnik. Anti-Fluorescein-AP Fab fragment antibody (cat N° 11426338910) for the detection of fluorescein-labeled compounds was obtained from Sigma-Aldrich. FITC-Soybean Agglutinin lectin (SBA) was purchased from Vector laboratories. Reactions requiring anhydrous conditions were performed under nitrogen atmosphere. All chemically synthetized compounds were characterized by ^{1}H (400.133 or 300.135 MHz), ^{13}C (125.773 or 75.480 MHz) NMR spectroscopy (Bruker Avance 300 Ultra Shield or Bruker Avance III 400 spectrometer). Chemical shifts are reported in parts per million (ppm); coupling constants are reported in units of Hertz [Hz]. The following abbreviations were used: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, br = broad singlet. When needed, ^{13}C heteronuclear HMQC and HMBC were used to unambiguously establish structures. High-resolution mass spectra (HRMS) were recorded with a Thermofisher hybrid LTQ-orbitrap spectrometer (ESI^{+}) and a Bruker Autoflex III SmartBeam spectrometer (MALDI). HPLC analysis were performed on an HPLC autopurification system (WATERS) equipped with BGM 2545 binary pump, a 2767 Sample Manager and equipped with an UV-visible diode array detector and evaporative light scattering detector in series (WATERS 996 PDA and WATERS 2424 ESLD). Chromatographic separation was performed on an Atlantis T3 (WATERS; 4.6 x 150 mm, 5µ). The mobile phase consisted of water (solvent A) and acetonitrile (solvent B). Gradient mix: 0.0 min [95% A; 5% B], 20.0 min [77% A; 23% B], flowrate: 1 mL/min.

All chemically synthetized products were purified by flash chromatography (GRACE REVELERIS Flash Chromatography System) equipped with UV and DLS detectors.

Compounds **1**, **6** and **9** were synthetized according to the literature.[1]

Synthesis



Compound 3

To a solution of the azide **1** (284 mg, 0.56 mmol) in MeOH (6 mL) was added PTSA (107 mg, 0.56 mmol), followed by the addition of 10% of Pd-C (10% w). The resulting suspension was stirred under H_2 atmosphere (1 atm) for 12h. The Pd/C was removed by filtration through Celite® and the filtrate was evaporated under reduced pressure to give the ammonium salt, as confirmed by ^{1}H NMR.

The crude of the reaction was dissolved in a mixture 1:1 H_2O/MeOH (20 mL), followed by addition of Amberlite IRN78 basic resin. After 3h of stirring at 20°C, the reaction mixture was filtrated and evaporated under reduce pressure to give the deprotected amine **3** as a white solid (yield: **3** = 77%) used in the next step without further purification. ^{1}H NMR (MeOD): 1.99 (s, 3H, NHAc), 2.83 (t, 2H, **CH_2**NH_2, $J_{H,H}$ = 5.4 Hz), 3.4-4.0 (m, 18H, 5 x **CH_2**O, H-2, H-3, H-4, H-5, H-6a, H-6b, NH_2), 4.42 (d, 1H, H-1, $J_{1,2}$ = 8.4 Hz) ; ^{13}C NMR (MeOD): 23.1, 42.1, 54.3,

62.6, 69.7, 69.8, 71.3, 71.5, 71.6, 73.5, 73.6, 76.8, 103.1, 174.2 ; HRMS (MALDI) for $C_{14}H_{29}N_2O_8$ $[M+H]^+$, calcd 353.1924 found 353.1918. (**Fig. S3 and S4**)

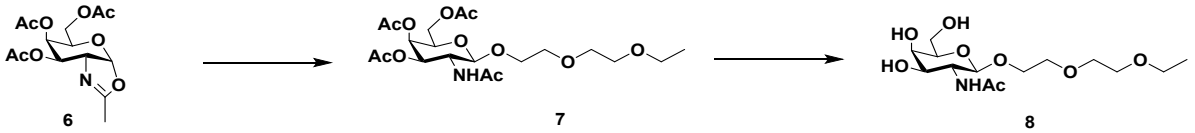
Compound 4

To a solution of **3** (162 mg, 0.46 mmol, crude) in dry DMF (5 mL) was added 1,1'-thiocarbonyldi-2(1H)-pyridone (117 mg, 0.51 mmol) and the reaction was stirred overnight at 20°C under N_2 atmosphere. The reaction mixture was then evaporated under reduced pressure and the residue purified by flash chromatography (SiO_2, DCM/MeOH: 100/0 → 80/20) to yield the thioisocyanate **4** (yield: **4** = 88%). ^{1}H NMR (MeOD): 2.0 (s, 3H, NHAc), 3.4-4.0 (m, 18H, 5 x **CH$_2$**O, **CH$_2$**NCS, H-2, H-3, H-4, H-5, H-6a, H-6b), 4.44 (d, 1H, H-1, $J_{1,2}$ = 8.4 Hz) ; ^{13}C NMR (MeOD): 23.1, 46.3, 54.3, 62.6, 69.7, 69.8, 70.5, 71.5, 71.6, 71.7, 73.6, 76.8, 103.1, 133.2, 174.2 ; HPLC: t_R = 11.71 min, the mobile phase consisted of water (solvent A) and acetonitrile (solvent B). Gradient mix: 0.0 min [95% A; 5% B], 20.0 min [77% A; 23% B], flowrate: 1 mL/min.; HRMS (MALDI) for $C_{15}H_{26}N_2O_8NaS$ $[M+Na]^+$, calcd 417.1308 found 417.1318. (**Fig. S5, S6 and S7**)

Compound 5

To a solution of **3** (64 mg, 0.18 mmol, crude) in dry DMF (1.8 mL) was added, dropwise and under N_2 atmosphere, *p*-phenylene diisothiocyanate (175 mg, 0.9 mmol, dissolved in 1.8 mL of dry DMF) using a siring pump (1 h of addition). After 2h of stirring at 20°C, the solvent was evaporated under reduced pressure and the residue was purified by flash chromatography (SiO_2, DCM/MeOH: 100/0 → 80/20) to yield the thioisocyanate **5** as a lightly yellow solid (Yield: **5** = 85%). ^{1}H NMR (MeOD): 1.99 (s, 3H, NHAc), 3.4-4.0 (m, 18H, 5 x **CH$_2$**O, **CH$_2$**NH, H-2, H-3, H-4, H-5, H-6a, H-6b), 4.43 (d, 1H, H-1, $J_{1,2}$ = 8.4 Hz), 7.24 (d, 2H, $J_{H,H}$ = 8.8 Hz), 7.53 (d, 2H, $J_{H,H}$ = 8.8 Hz) ; ^{13}C NMR (MeOD): 23.2, 45.4, 54.3, 62.6, 69.7, 69.9, 70.3, 71.4, 71.5,

71.6, 73.4, 76.8, 103.2, 2 x 125.8, 2 x127.1, 128.5, 136.6, 139.8, 174.2, 182.5 ; HPLC: t_R = 16.02 min, the mobile phase consisted of water (solvent A) and acetonitrile (solvent B). Gradient mix: 0.0 min [95% A; 5% B], 20.0 min [77% A; 23% B], flowrate: 1 mL/min.; HRMS (MALDI) for $C_{22}H_{33}N_4O_8S_2$ $[M+H]^+$, calcd 545.1740 found 545.1742. (**Fig. S8, S9 and S10**)

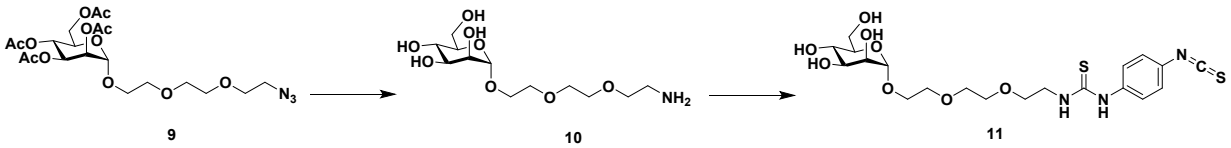


Compound 7

To a solution of the oxazoline **6** (981 mg, 2.98 mmol) and 2-(2-Ethoxyethoxy)ethanol (404 µL, 2.98 mmol) in dry DCM (3 mL) was added 4 Å MS. The resulting solution was stirred at 20°C under N_2 atmosphere. After 30 min, TMSOTf (270 µL, 1.49 mmol) was added dropwise at 0°C under N_2 atmosphere. The mixture of the reaction was stirred for 30 min at 0°C and warmed up to 20°C. After 12 h of stirring, the MS was filtered and the solvent evaporated. The residue was redissolved in DCM, washed respectively with saturated aqueous $NaHCO_3$, water and brine, dried over $MgSO_4$, filtered and evaporated. The residue was purified by flash chromatography (SIO_2, DCM/MeOH: 100/0 → 95/5) to yield the derivative **7** as a colorless oil. (Yield: **7** = 73%). ^{1}H NMR (MeOD): 1.19 (t, 3H, **CH_3**CH_2O, $J_{H,H}$ = 7.2 Hz), 1.93 (s, 3H, AcO), 1.94 (s, 3H, AcO), 2.03 (s, 3H, AcO), 2.14 (s, 3H, NHAc), 3,52 (q, 2H, CH_3**CH_2**O, $J_{H,H}$ = 7.2 Hz), 3.54-4.20 (m, 12H, 4 x **CH_2**O, H-2, H-5, H-6a, H-6b), 4.67 (d, 1H, H-1, $J_{1,2}$ = 8.4 Hz), 5.04 (dd, 1H, H-3, $J_{3,4}$ = 3.2 Hz, $J_{3,2}$ = 11.2 Hz), 5.33 (d, 1H, H-4, $J_{3,4}$ = 3.2 Hz) ; ^{13}C NMR (MeOD): 15.4, 20.5, 2*20.6, 22.9, 51.6, 54.8, 62.7, 67.6, 68.2, 70.0, 70.9, 71.6, 71.8, 72.3, 102.8, 171.7, 2*172.1, 173.5 ; HRMS (MALDI) for $C_{20}H_{34}NO_{11}$ $[M+H]^+$, calcd 464.2132 found 464.2141. (**Fig. S11 and S12**)

Compound 8

Acetylated sugar **7** (152 mg, 0.328 mmol) was dissolved in a mixture 1:1 H_2O/MeOH (20 mL), followed by addition of Amberlite IRN78 basic resin. After 3h of stirring at 20°C, the reaction mixture was filtrated and evaporated under reduced pressure to give the final compound **8** as a white solid (Yield: **8** = 77%). ^{1}H NMR (MeOD): 1.2 (t, 3H, **CH$_3$**CH$_2$O, $J_{H,H}$ = 6.8 Hz), 1.99 (s, 3H, NHAc), 3.4-4.0 (m, 16H, 4 x **CH$_2$**O, CH$_3$**CH$_2$**O, H-2, H-3, H-4, H-5, H-6a, H-6b), 4.45 (d, 1H, H-1, $J_{1,2}$ = 8.4 Hz) ; ^{13}C NMR (MeOD): 15.4, 23.1, 54.3, 62.6, 67.6, 69.7, 69.8, 71.0, 71.6, 71.7, 73.6, 76.8, 103.1, 174.2 ; HPLC: t_R = 7.57 min, the mobile phase consisted of water (solvent A) and acetonitrile (solvent B). Gradient mix: 0.0 min [95% A; 5% B], 20.0 min [77% A; 23% B], flowrate: 1 mL/min.; HRMS (MALDI) for $C_{14}H_{28}NO_8$ $[M+H]^+$, calcd 338.1815 found 338.1819. (**Fig. S13, S14 and S15**)



9 10 11

Compound 10

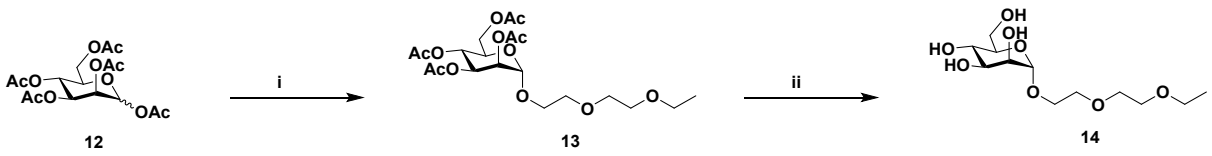
To a solution of **9** (225 mg, 0.445 mmol) in MeOH (4.5 mL) was added PTSA (77 mg, 0.449 mmol), followed by the addition of 10% of Pd-C (10% w). The resulting suspension was stirred under H_2 atmosphere (1 atm) for 12h. The Pd/C was removed by filtration through Celite® and the filtrate was evaporated under reduced pressure to give the ammonium salt, as confirmed by ^{1}H NMR.

The crude of the reaction was dissolved in a mixture 1:1 H_2O/MeOH (20 mL), followed by addition of Amberlite IRN78 basic resin. After 3 h of stirring at 20°C, the reaction mixture was filtrated and evaporated under reduced pressure to give the deprotected amine **10** as a colorless oil (yield: **10** = 99 %) used in the next step without further purification. ^{1}H NMR (MeOD): 2.81 (m, 2H, **CH$_2$**NH$_2$), 2.15-3.9 (m, 16H, 5 x **CH$_2$**O, H-2, H-3, H-4, H-5, H-6a, H-6b), 4.81 (d, 1H, H-1, $J_{1,2}$ = 1.7Hz); ^{13}C NMR (MeOD): 41.9, 62.9, 67.7, 68.6, 71.3, 71.4, 71.6, 72.1, 72.5, 72.8,

74.6, 101.7; HRMS (ESI) for $C_{12}H_{26}NO_8$ $[M+H]^+$, calcd 312.1658 found 312.1651. (**Fig. S16 and S17**)

Compound 11

To a solution of **10** (138 mg, 0.444 mmol) in dry DMF (4.0 mL) was added, dropwise and under N_2 atmosphere, *p*-phenylene diisothiocyanate (426 mg, 0.2.219 mmol, dissolved in 4.0 mL of dry DMF) using a siring pump (1h of addition). After 2h of stirring at 20°C, the solvent was evaporated under reduced pressure and the residue was purified by flash chromatography (SiO_2, DCM/MeOH: 100/0 → 80/20) to yield the thioisocyanate **11** (Yield: **11** = 44 %) as a lightly yellow solid. ^{1}H NMR (DMSO-d6): 3.29-3.72 (m, 18H, 5 x **CH_2**O, **CH_2**NH, H-2, H-3, H-4, H-5, H-6a, H-6b), 4.45 (t, 1H, OH, $J_{OH,H}$ = 5.9 Hz), 4.58 (d, 1H, OH, $J_{OH,H}$ = 5.9 Hz), 4.64 (d, 1H, H-1, $J_{1,2}$ = 1.4 Hz), 4.73 (d, 1H, OH, $J_{OH,H}$ = 4.5 Hz), 4.76 (d, 1H, OH, $J_{OH,H}$ = 4.7 Hz), 7.37 (d, 2H, $J_{H,H}$ = 8.9 Hz), 7.59 (d, 2H, $J_{H,H}$ = 8.9 Hz), 7.93 (bs, 1H, NH), 9.79 (bs, 1H, NH); ^{13}C NMR (DMSO-d6): 43.6, 61.3, 65.7, 67.0, 68.5, 69.5, 69.6, 69.8, 70.3, 70.9, 73.9, 100.0, 123.1, 124.7, 126.2, 132.6, 139.3, 180.3; HPLC: t_R = 8.17 min, the mobile phase consisted of water (solvent A) and acetonitrile (solvent B). Gradient mix: 0.0 min [95% A; 5% B], 20.0 min [77% A; 23% B], flowrate: 1 mL/min.; HRMS (ESI) for $C_{20}H_{29}N_3O_8S_2Na$ $[M+Na]^+$, calcd 526.1294 found 526.1292. (**Fig. S18, S19 and S20**).



Compound 13

To a solution of the acetyl glycoside **12** (500 mg, 1.282 mmol) and 2-(2-Ethoxyethoxy)ethanol (258 mg, 1.923 mmol) in dry DCM (13 mL) was added 4 Å MS. The resulting solution was stirred at 20°C under N_2 atmosphere. After 30 min, $F_3B{\cdot}OEt_2$ (791 µL, 6.410 mmol) was added dropwise at 0°C under N_2 atmosphere. The mixture of the reaction was stirred for 30 min at

0°C and warmed up to 20°C. After 12h of stirring, the MS was filtered and the solvent evaporated. The residue was redissolved in DCM, washed respectively with saturated aqueous $NaHCO_3$, water and brine, dried over $MgSO_4$, filtered and evaporated. The residue was purified by flash chromatography (SiO_2, Cy/AcOEt 70/30) to yield the derivative **7** (Yield: **13** = 72 %) as a colorless oil. ^{1}H NMR ($CDCl_3$): 1.20 (t, 3H, **CH_3**CH_2O, J = 7.0 Hz), 1.97 (s, 3H, AcO), 2.02 (s, 3H, AcO), 2.09 (s, 3H, AcO), 2.14 (s, 3H, AcO), 3.52 (q, 2H, CH_3**CH_2**O, J = 7.0 Hz), 3.56-3.86 (m, 8H, 4 x **CH_2**O), 4.08 (m, 2H, H-5, H-6a), 4.28 (dd, 1H, H-6b, $J_{6b,6a}$ = 12.6 Hz, $J_{6b,5}$ = 5.3 Hz), 4.89 (d, 1H, H-1, $J_{1,2}$ = 1.5 Hz), 5.25 (dd, 1H, H-2, $J_{2,3}$ = 3.3 Hz, $J_{2,1}$ = 1.8 Hz), 5.27 (t, 1H, H-4, $J_{4,3}$ = $J_{4,5}$ = 9.7 Hz), 5.35 (dd, 1H, H-3, $J_{3,4}$ = 10.0 Hz, $J_{3,2}$ = 3.4 Hz); ^{13}C NMR ($CDCl_3$): 15.1, 20.6, 20.7, 20.8, 62.4, 66.1, 66.6, 67.4, 68.4, 69.1, 69.6, 69.8, 70.0, 70.8, 97.7, 169.7, 169.8, 170.0, 170.1; HRMS (ESI) for $C_{20}H_{32}O_{12}Na$ $[M+Na]^+$, calcd 487.1791 found 487.1784. (**Fig. S21 and S22**).

Compound 14

Acetylated sugar **13** (465 mg, 0.667 mmol) was dissolved in a mixture 1:1 H_2O/MeOH (12 mL), followed by addition of Amberlite IRN78 basic resin. After 3h of stirring at 20°C, the reaction mixture was filtrated and evaporated under reduced pressure to give the final compound **14** (Yield: **14** = 99 %) as a white solid. ^{1}H NMR (MeOD): 1.19 (t, 3H, **CH_3**CH_2O, J = 7.0 Hz), 3.54 (q, 2H, CH_3**CH_2**O, J = 7.0 Hz), 3.56-3.86 (m, 14H, 4 x **CH_2**O, H-2, H-3, H-4, H-5, H-6a, H-6b), 4.79 (d, 1H, H-1, $J_{1,2}$ = 1.7 Hz); ^{13}C NMR (MeOD): 15.4, 62.9, 67.6, 67.8, 68.6, 70.9, 71.4, 71.6, 72.1, 72.5, 74.6, 101.7; HPLC: t_R = 2.19 min, the mobile phase consisted of water (solvent A) and acetonitrile (solvent B). Gradient mix: 0.0 min [95% A; 5% B], 20.0 min [77% A; 23% B], flowrate: 1 mL/min.; HRMS (ESI) for $C_{12}H_{24}O_8Na$ $[M+Na]^+$, calcd 319.1369 found 319.1371. (**Fig. S23, S24 and S25**).

AAV2 production and purification

AAV vectors were produced from two plasmids: (i) pHelper, PDP2-KANA encoding AAV Rep2-Cap2 and adenovirus helper genes (E2A, VA RNA, and E4) for AAV2 vectors or PDP3-KANA encoding AAV Rep2-Cap3b and adenovirus helper genes (E2A, VA RNA, and E4) for AAV3b vectors and (ii) the pVector ss-CAG-eGFP containing the ITRs. All vectors were produced by transient transfection of HEK293 cells with calcium phosphate-HeBS method. AAV2 transfected cells were harvested 48h after transfection and treated with Triton-1% and benzonase (25U/mL) for 1h at 37°C. AAV3b transfected cells were also harvested at 48h but without Triton/benzonase treatment. The bulk was subjected to freeze-thaw cycles to release vector particles. The cellular debris were removed by centrifugation at 2500 rpm for 15 min. Cell lysates were precipitated with PEG overnight and clarified by centrifugation at 4000 rpm for 1h. The precipitates were then incubated with benzonase for 30 min at 37 °C and collected after centrifugation at 10000 g for 10 min at 4°C. Vectors were purified by double cesium chloride (CsCl) gradient ultracentrifugation. The viral suspension was then subjected to four successive rounds of dialysis under slight stirring in a Slide-a-Lyzer cassette (Pierce) against dPBS (containing Ca^{++} and Mg^{++}).

Coupling and purification

AAV2-GFP or AAV3b (10^{12} vg, 2.49 nmol) were added to a solution of TBS buffer pH = 9.3 containing FITC, **4**, **5**, **8**, **11** or **14** at different molar ratios (3E5, 3E6 or 1.5E7), as stated in the results section for each case, and incubated during 4h at RT. The solutions containing the vectors were then dialyzed against dPBS + 0.001% Pluronic to remove free molecules that were not bond to the AAV capsid.

Quantification of AAV vector genomes

3 µL of AAV were treated with 20 units of DNase I (Roche #04716728001) at 37°C for 45 min to remove residual DNA in vector samples. After the treatment with DNase I, 20µL of proteinase K 20 mg/mL (MACHEREY-NAGEL # 740506) was then added and incubated at 70°C for 20 min. Extraction columns (NucleoSpin®RNA Virus) were then used to extract DNA from purified AAV vectors.

Quantitative real time PCR (qPCR) was performed with a StepOnePlus™ Real-Time PCR System Upgrade (Life technologies). All PCRs were performed in a 20µL final volume PCR including primers and probe targeting the ITR2 sequence,[2] PCR Master Mix (TaKaRa) and 5µL of template DNA (plasmid standard, or sample DNA). qPCR was carried out with an initial denaturation step at 95°C for 20 seconds, followed by 45 cycles of denaturation at 95°C for 1 second and annealing/extention at 56°C for 20 seconds. Plasmid standards were generated with seven serial dilutions (containing 10^8 to 10^2 copies of plasmid) according to D'Costa *et al.*[2]

Western blotting and silver staining

All vectors were denatured at 100°C for 5 min using Laemmli sample buffer and separated by SDS-PAGE 10% Tris-glycine polyacrylamide gels (Life Technologies). Precision plus Protein All Blue Standards (BioRad) was used as a molecular-weight size marker. Following electrophoresis, gels were either silver stained (PlusOne Silver staining Kit, protein, GE Healthcare) or transferred onto nitrocellulose membranes for Western blot analysis. After transferring the proteins to nitrocellulose membrane using a transfer buffer (25mM tris/192mM Glycine/0.1 (w/v) SDS/20%MeOH) for 1h at 150 mA in a Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad), the membrane was saturated with 5% semi-skimmed milk in PBS-Tween (0.1%) or with 1% gelatin, 0.1% Igepal in PBS-Tween (0.01%) during 2h at RT. After saturation, the membrane was probed with the corresponding antibody (anti-capsid polyclonal, B1 monoclonal or anti-fluorescein-AP), with FITC-Soybean Agglutinin lectin (GalNAc

detection) or with FITC-Concanavalin A lectin (mannose detection) overnight at 4°C. Three washes were carried out between each stage to remove unbound reagents with PBS-Tween (0.1%) for 15 min at RT. Bands were visualized by chemiluminescence using alkaline phosphatase (AP) or horseradish peroxidase (HRP)-conjugated secondary antibodies and captured on X-ray film.

Immuno dot-blot analysis

AAV vectors were loaded at a dose of 10^{10} vg on a nitrocellulose paper soaked briefly in PBS prior to assembling the dot blot manifold (Bio-Rad). Nitrocellulose membrane containing the vectors was treated as for Western blotting.

Dynamic light scattering

DLS was performed using a Malvern zetasizer nano ZS. The calibration was controlled beforehand by using 30 and 300 nm solution of Nanosphere size standard. 50 µL of each vector were placed in a specific cuvette DTS0118 from Malvern and analysed by volume.

AAV infectious titre measurement

A dose of 10^{12} vg of AAV2-GFP was diluted in different buffer solutions (dPBS, PBS, TBS buffer at pH 7.5, 8.5 and 9.3 or NaHCO3-Na2CO3 Buffer at pH 9.3) and incubated at RT for 4h. The infectivity of each sample was measured as follows. In short, HEK293 cells were seeded in six-well culture plates at a density of 10^6 cells/well in 2mL of DMEM growth medium. Cells were then incubated overnight at 37°C to reach 50% confluence. The viral stock was then diluted 10-fold by serial dilutions. Next, 2µL of each dilution was added to separate wells of the six-well plates. Plates were then incubated at 37°C for 24h. The infectivity of AAV2-GFP control was measured immediately upon thawing of the sample. The same procedure was used

for the GalNAc particles. Detection of AAV-GFP-infected cells was performed using a Fluorescence Microscopy.

The transducing unit (TU) titer was calculated using the following formula:

TU/mL = (4040 x NGFP x dilutions x 1000)/V

Where NGFP is the mean number of GFP-positive cells per well.

V is the volume of vector used to infect cells in µL.

Confocal microscopy

Immunolabeling was performed on fixed cells grown on Lab-Tek chamber slides (8 wells, Nalge Nunc International, Rochester, NY). Cells were fixed in 2% PFA (10 minutes at RT) and permeabilized with 0.5% of Triton X-100 (30 minutes at room temperature). After incubation in blocking buffer (2% of goat serum diluted in PBS, 1 hour at RT), cells were incubated overnight at 4°C with the primary antibody A20 (Dako) diluted in blocking buffer (1:50). Three washes of 10 min were performed in PBS before the incubation with the secondary antibody, a goat anti-mouse labeled with Alexa fluor 647 (1:500 in PBS; Invitrogen, Carlsbad, CA, 1h at RT). Finally nuclei were stained using DAPI (15 minutes at RT).

3D modelling of the capsid

Structural analysis was performed using the PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC. AAV2 crystal structure was obtained from the Protein Data Bank (PDB ID: 1LP3).[3] The viral capsid was reconstituted from the monomer crystal structure using the *BiologicalUnit/Quat* script (https://pymolwiki.org/index.php/BiologicalUnit/Quat). The analysis of surface-exposed tyrosine residues was performed on a single monomer, surrounded by seven adjacent monomers (one, two and four from the 2-fold, 3-fold and 5-fold symmetry axis, respectively), using the *findSurfaceResidues* script (https://pymolwiki.org/index.php/

FindSurfaceResidues), with a solvent-exposure cut-off value of 5 Å. Lysine residues located on the internal capsid surface were identified via manual inspection and discarded, resulting in the list of surface-exposed lysine residues represented in this paper.

Neutralizing assay

AAV2-directed neutralizing factors (NFs) were assessed as previously described in Guilbaud *et al*.[4] Briefly, 2h prior to AAV2 transduction, HeLa cells were infected with wild-type adenovirus serotype 5. During this incubation time, AAV2, GalNAc-AAV2 (3E5 or 3E6) expressing the beta-galactosidase reporter protein were incubated with twofold dilutions of a pool of non-human primate serum known to be neutralizing. Dilutions ranged from 1/20 to 1/5120. After removing the adenovirus-containing medium, the mix (serum + AAV2 modified or non-modified) was added to the cells. Forty-eight hours later, the cells were washed in PBS 1X, lysed and incubated with Galacton-Star substrate in accordance with manufacturer recommendations (Galacto-Star, Thermo). Cheluminescence was measured 1h later with a Victor X3 Plate Reader (Perkin Elmer) and results were expressed in % of inhibition of the detected signal with AAV2, GalNAc-AAV2 (3E5 or 3E6) in absence of serum.

Transduction of primary mouse hepatocytes

Primary mouse hepatocytes and culture medium were purchased from BIOPREDIC international (Rennes, FRANCE). Mouse hepatocytes were seeded in a 24-well plate at a density of approximately $2.5x10^5$ cells/well. After reception, cell culture medium was removed and replaced with 1mL of basal medium (MIL600) with additives (ADD222) and incubated 2h at 37°C and 5% CO_2. Primary mouse hepatocytes were transduced at MOI of 10^5 with AAV2 or chemically modified AAV2 vectors in 0.5 mL of culture medium. 6h after the transduction, 0.5 mL of fresh culture medium was added to each well. All AAV vectors encoded for GFP.

The percentage of GFP positive cells was measured by FACS analysis 72h after the transduction. Cells were dissociated with Trypsin-EDTA (Sigma-Aldrich), fixed with 4% paraformaldehyde and analysed on a BD-LSRII Flow Cytometer (BD Bioscience). All data were processed by FlowJo (V10; Flowjo LLC, Ashland, OR).

Animals

The experimental design was approved by the Ethics Committee for Animal Testing of the University of Navarra and all experiments were performed in accordance with relevant institutional and national Spanish guidelines.

C57BL/6 (B6) mice were purchased from Harlan Laboratories (Barcelona, Spain). They were housed at the animal facility at CIMA under SPF2 conditions with 12-hour light cycles and ad libitum access to food and water.

Mice were injected intravenously (i.v.) with a dose of 5×10^{12} vg/kg of either GalNAc-coupled eGFP-expressing AAV2 or unmodified eGFP-expressing AAV2. All experiments were performed in agreement with the protocol approved by the Animal Ethics Committee of the University of Navarra (Spain).

Quantification of RNA expression in liver tissue

The extraction of total RNA was performed by an initial step of mechanical homogenization in TRIzol (Invitrogen, #15596018). Subsequently 1 µg of total RNA was treated with Turbo DNAse (TURBO DNA-free™ Kit, #AM1907, Ambion) and reverse transcribed (RT) using random priming (Invitrogen, #48190-011) and Moloney murine leukemia virus (M-MLV) reverse transcriptase (Invitrogen, #28025013) according to the manufacturer's instructions. Total cDNA was analysed by quantitative PCR (qPCR) using *GoTaq*® qPCR *Master Mix (Promega – #A600A)* in a CFX96 Real-Time Detection System (BioRad). Mouse histone was used as reference for normalization.

Total and neutralizing antibodies in mice

Total anti-AAV antibodies were measured by a standard ELISA against AAV2, whereas neutralizing antibodies were determined by measuring infection of HEK293T cells by AAV2-luciferase. In short, HEK-293T cells were seeded into 96-well plates (ViewPlate TM - 96 BLACK, PerkinElmer - Ref: 6005182) and incubated overnight at 37˚C and 5% CO_2. The next day serial serum dilutions were prepared in a separate plate and AAV2-luciferase was added at a pre-determined MOI. Plates were incubated for 2h at 37˚C and 5% CO_2. The supernatant of the HEK293T cells was removed and the serum/AAV2 mix was added. Plates were incubated for a further 48h. Luciferin was added and plates were read in luminometer.

Statistical analysis

All the experiments are shown as mean ± standard error (SEM). GraphPad Prism 5 software was used for statistical analysis. Data were subjected to one-way analysis of variance (ANOVA). Samples were considered significantly different if $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

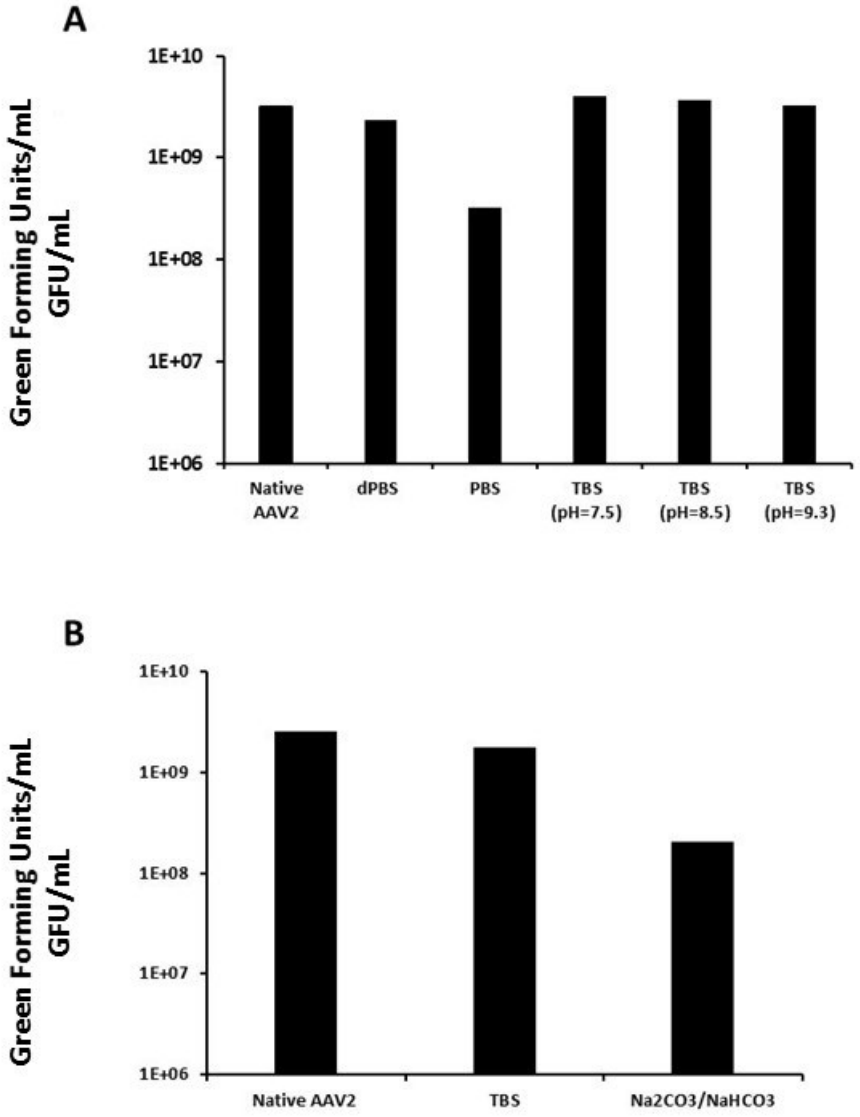


Fig. S1. Stability and infectivity of AAV2 in different buffers. (**A**) native AAV2 (10^{12} vg) stock formulated in dPBS pH=7 was placed in different buffers and vortexed at RT for 4 h. dPBS and PBS were adjusted at pH 7. Infectivity of the native AAV2 samples was tested immediately after thawing to be used a control, without incubation of 4h at RT (**B**) AAV2 (10^{12} vg) was placed in TBS or $NaHCO_3$-Na_2CO_3 buffer at pH 9.3 and vortexed at RT for 4 h. The stability of each sample was evaluated by measuring the GFU/mL titer after 4 h of incubation on HEK293 cells. Results are shown as mean ± SD. n=4 for each condition.

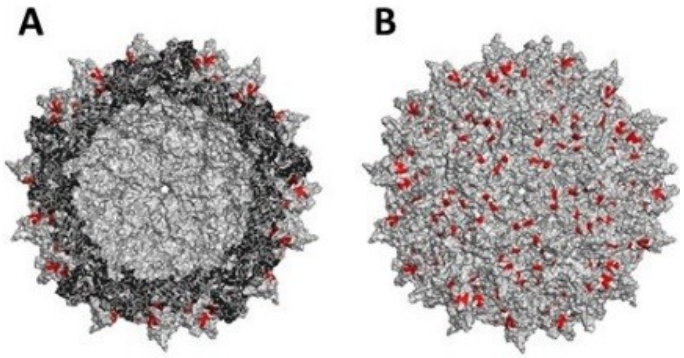


Fig. S2. Lysine residues exposed on the surface of AAV2. Lysine residues 258, 321, 490, 527, 532, 544, 665, 706 (AAV2 VP1 numbering), represented in red, are surface-exposed in AAV2 and potentially amenable to chemical coupling. Transversal (**A**) and external view (**B**) of the AAV2 capsid reconstruction (PDB ID: 1LP3).

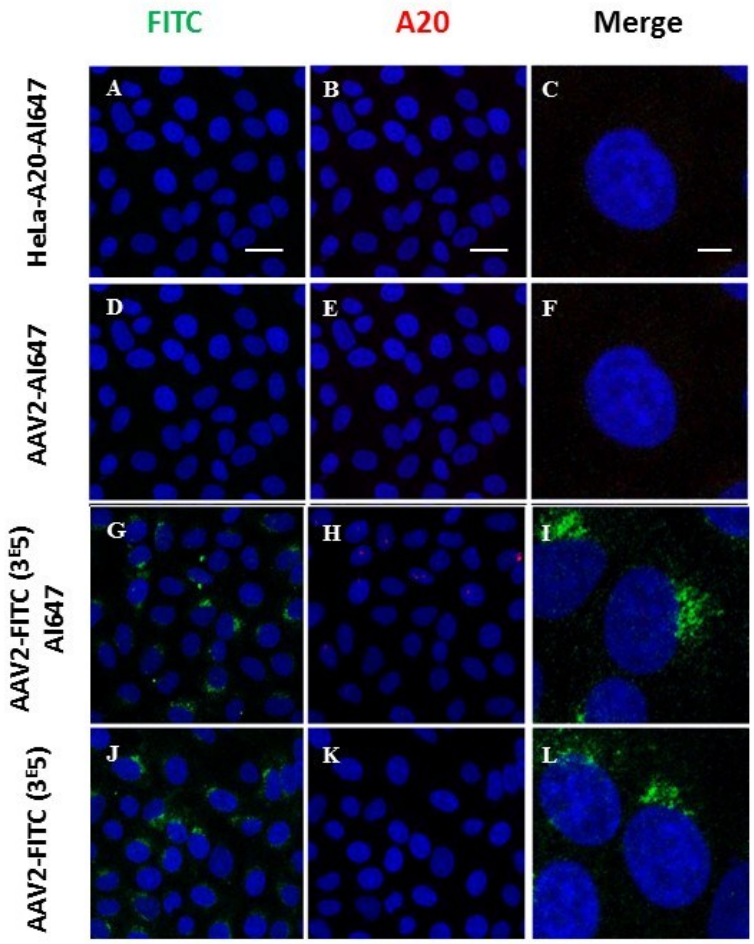


Fig. S3. Confocal imaging of modulated AAV2-FITC

(A-C) Uninfected HeLa cells were incubated with A20 primary antibody and the corresponding Al647 secondary antibody serving as as a negative control for these antibodies.

(D-F) HeLa cells were transduced with native AAV2 and incubated with the secondary Al647 antibody only, being a negative control for Fluorescence FITC detection (**D**) and for A20 immunodetection (**E**).

(G-I) HeLa cells were transduced with AAV2-3E5 FITC and incubated with Al647 only Green FITC fluorescence (**G**) was detected, but no red fluorescence due to the the omission of the primary antibody.

(J-L) HeLa cells were transduced with AAV2-3E5 FITC but not incubated with primary and secondary antibodies for ther capsid. Green FITC fluorescence was detected whereas red fluorescence was not detected.

Cell nuclei were counterstained with DAPI (blue).

Scale bars 25 µm for FITC and A20 images; 5µm for merged images.

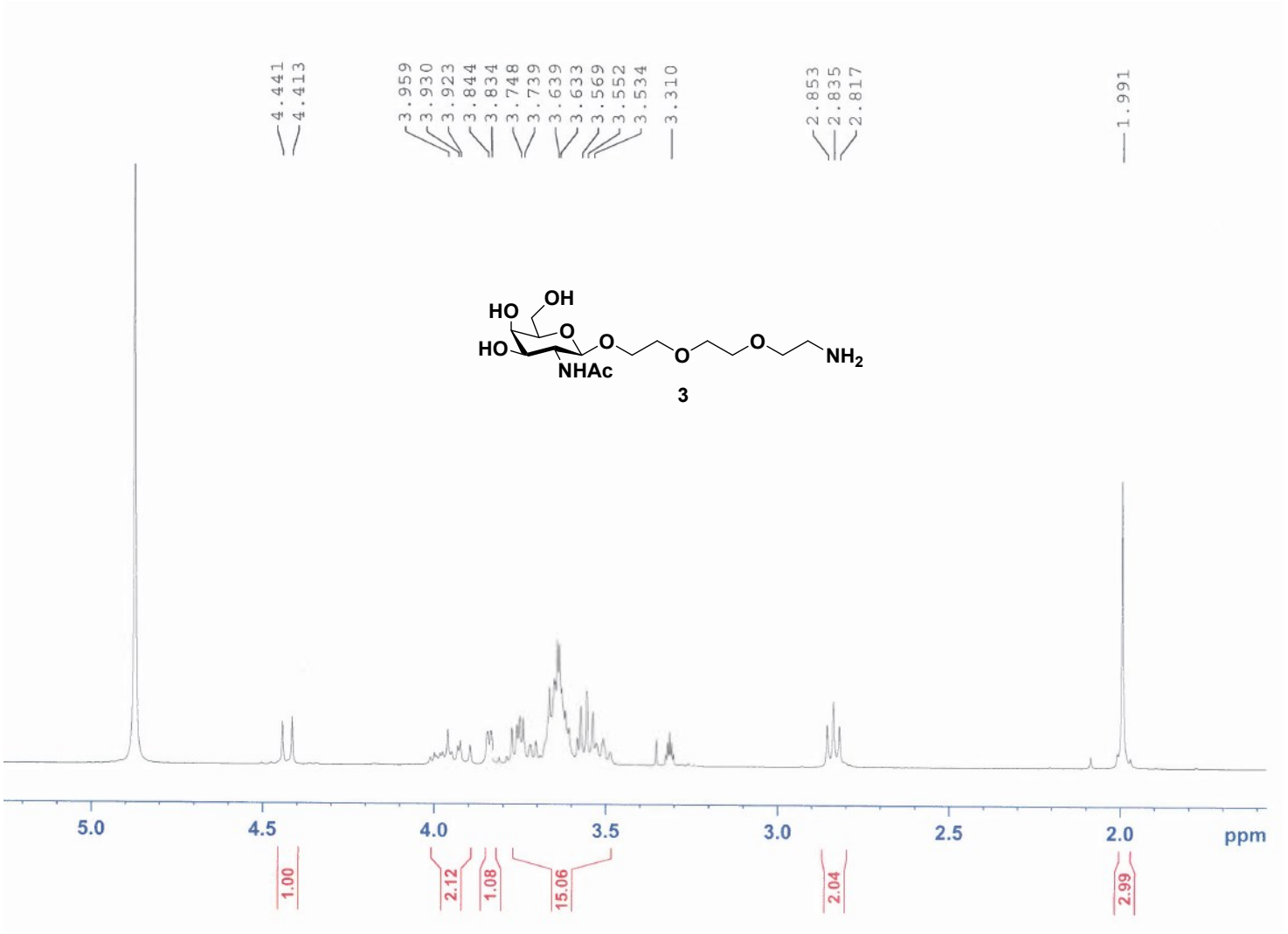


Fig. S4: ^{1}H NMR of compound **3**

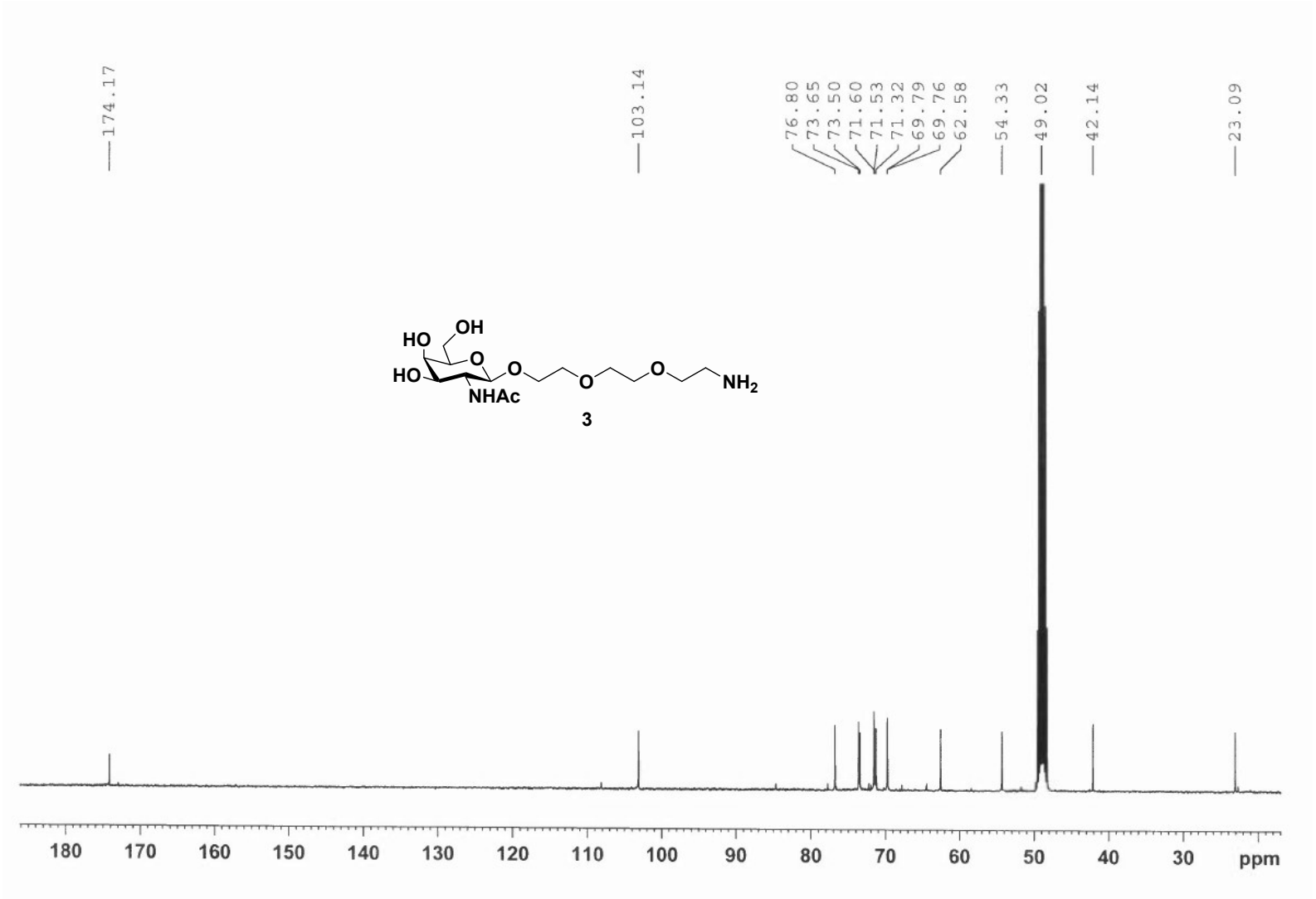


Fig. S5: ^{13}C NMR of compound **3**

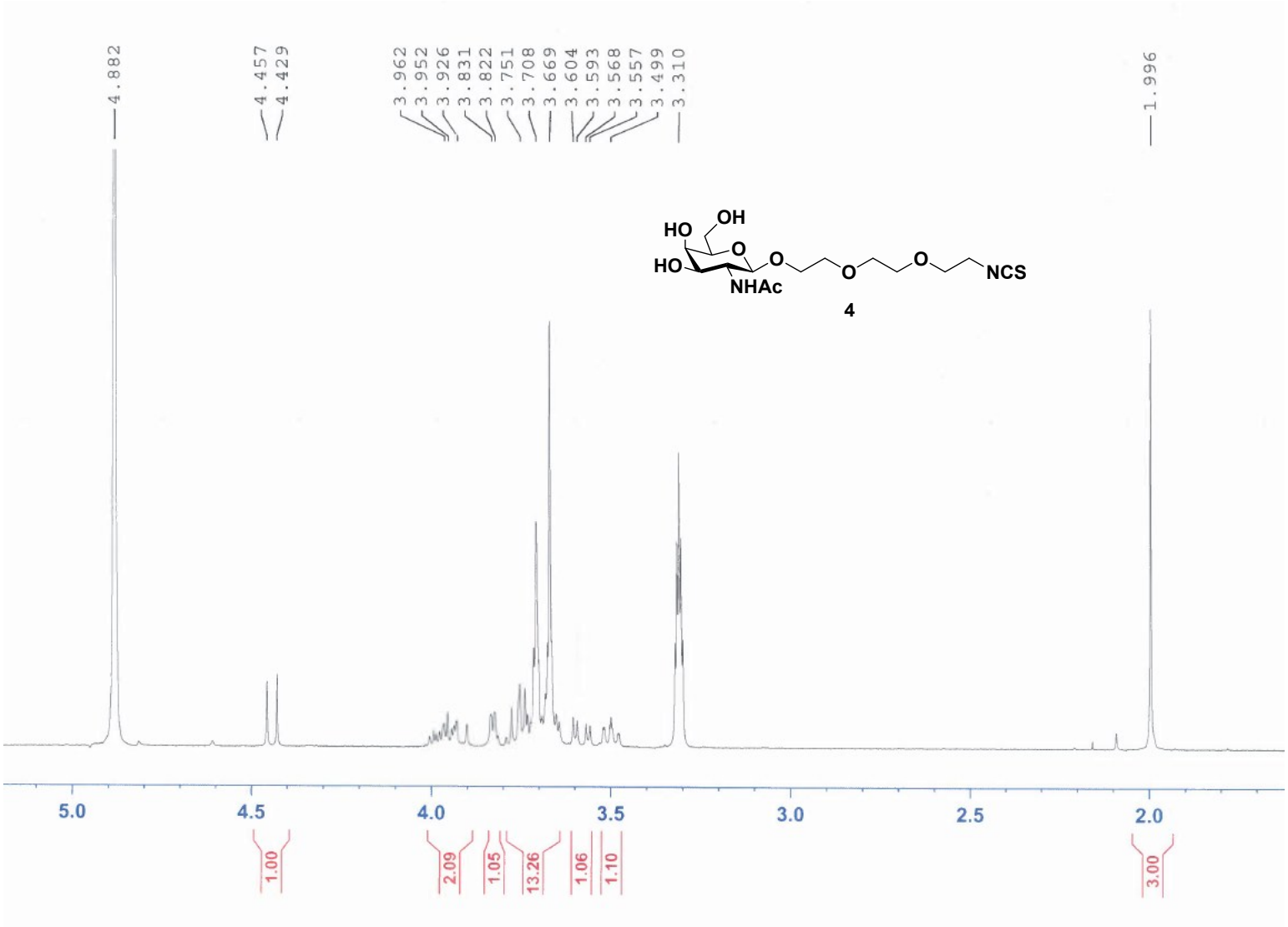


Fig. S6: ^{1}H NMR of compound **4**

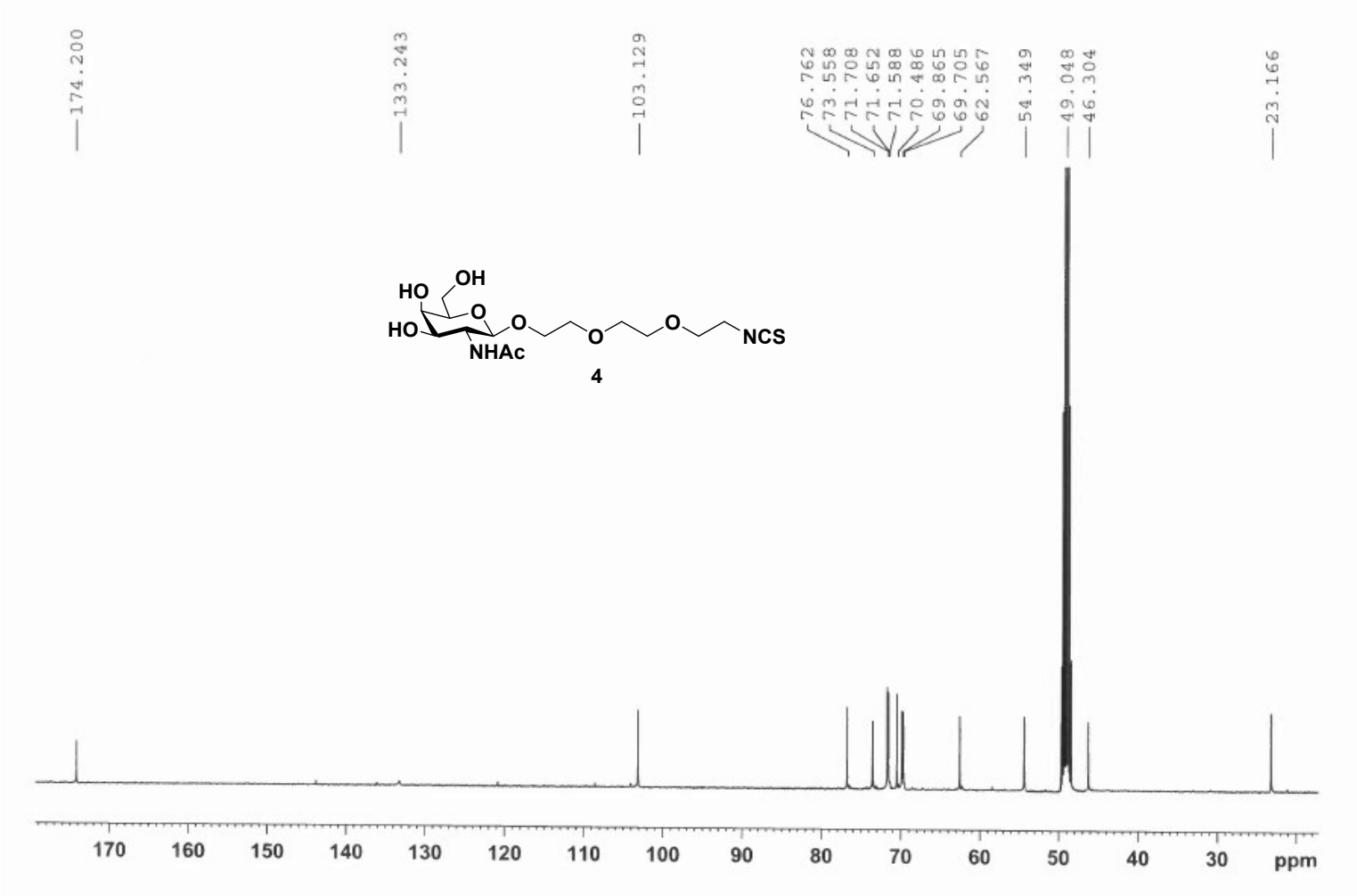


Fig. S7: ^{13}C NMR of compound **4**

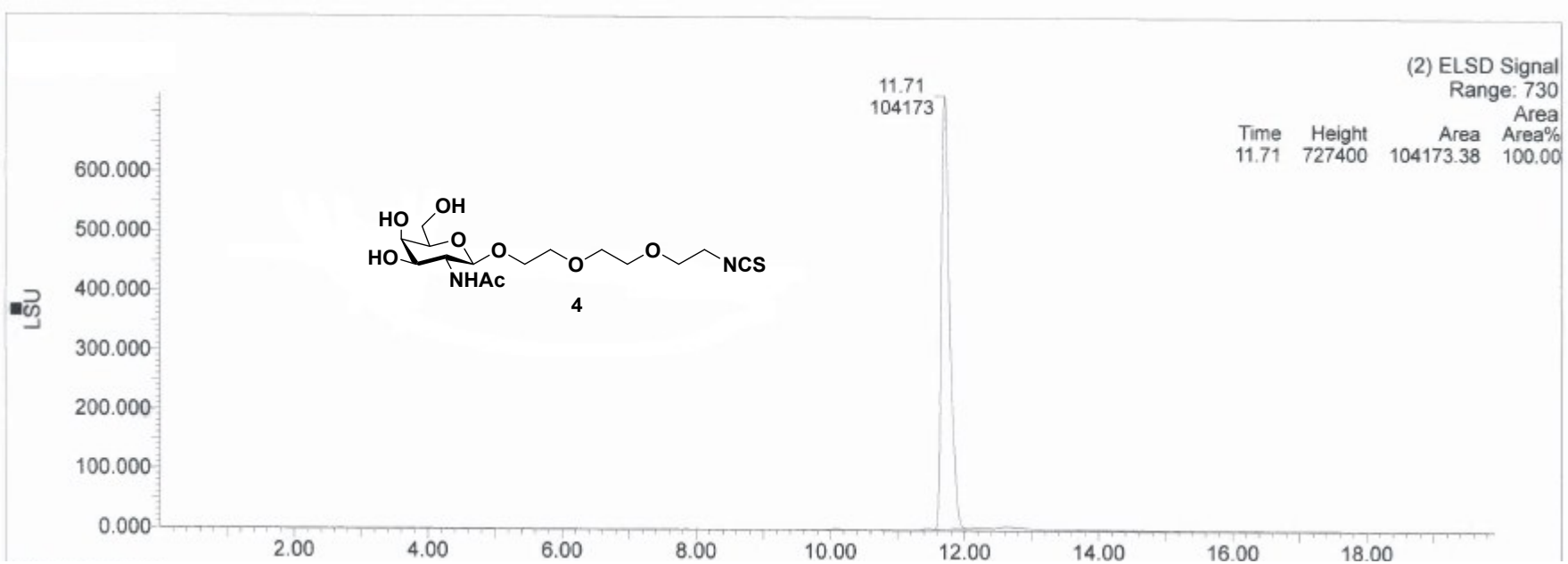


Fig. S8: Analytical HPLC of compound **4**

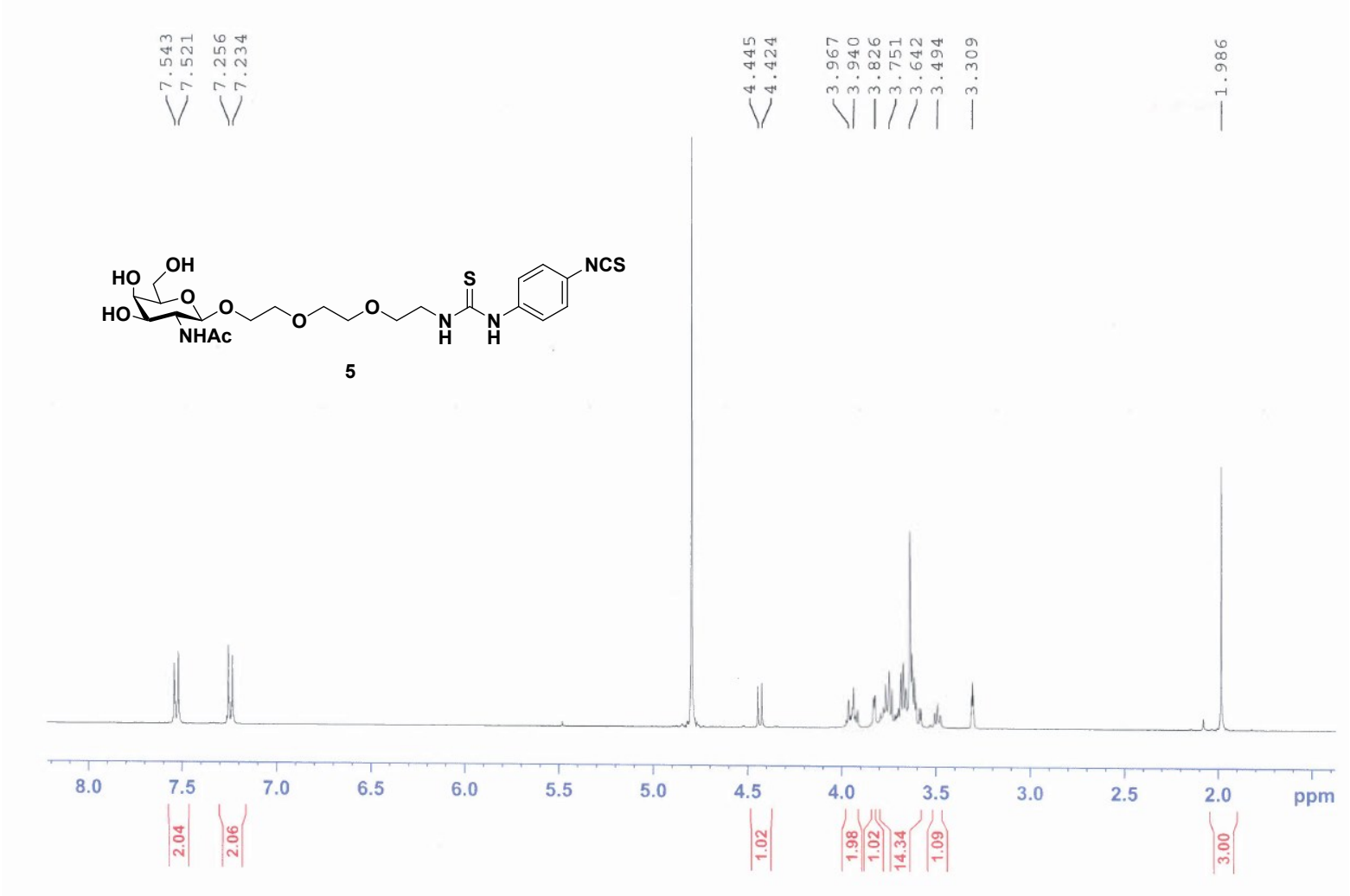


Fig. S9: ^{1}H NMR of compound **5**

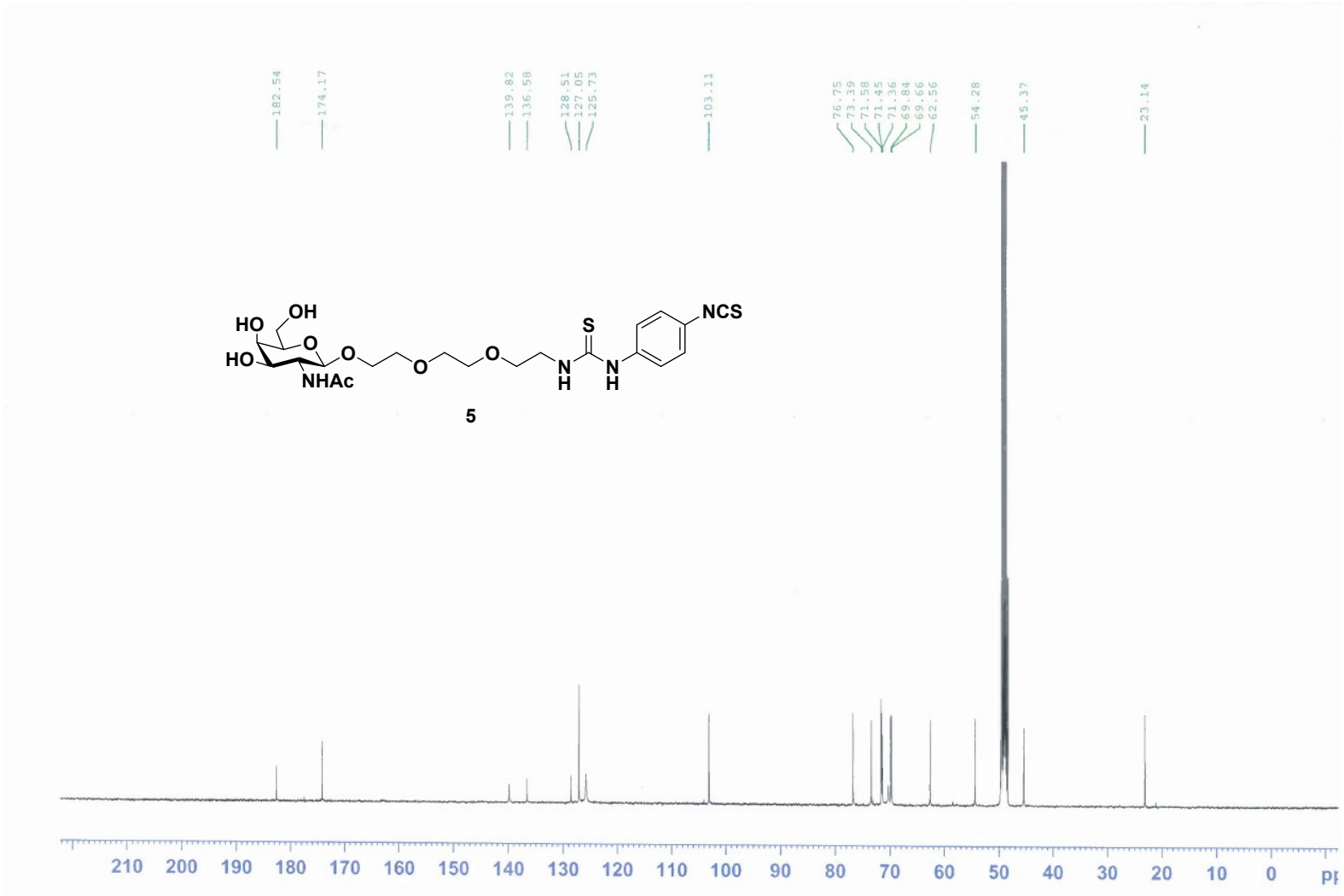


Fig. S10: ^{13}C NMR of compound **5**

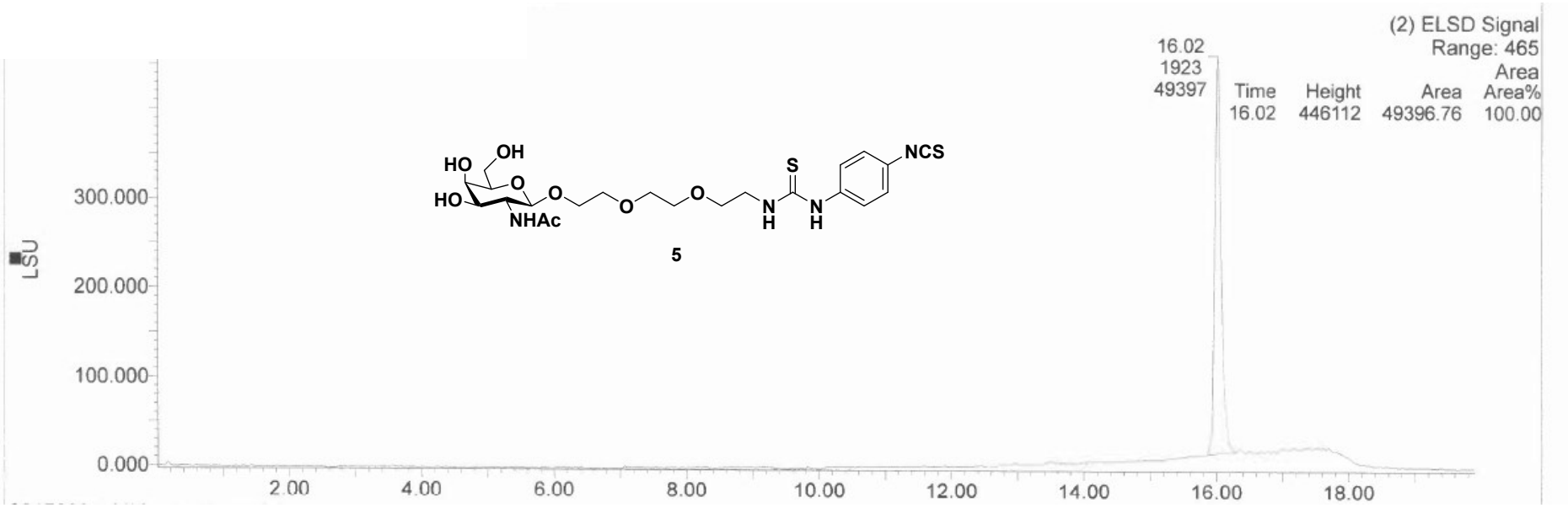


Fig. S11: Analytical HPLC of compound **5**

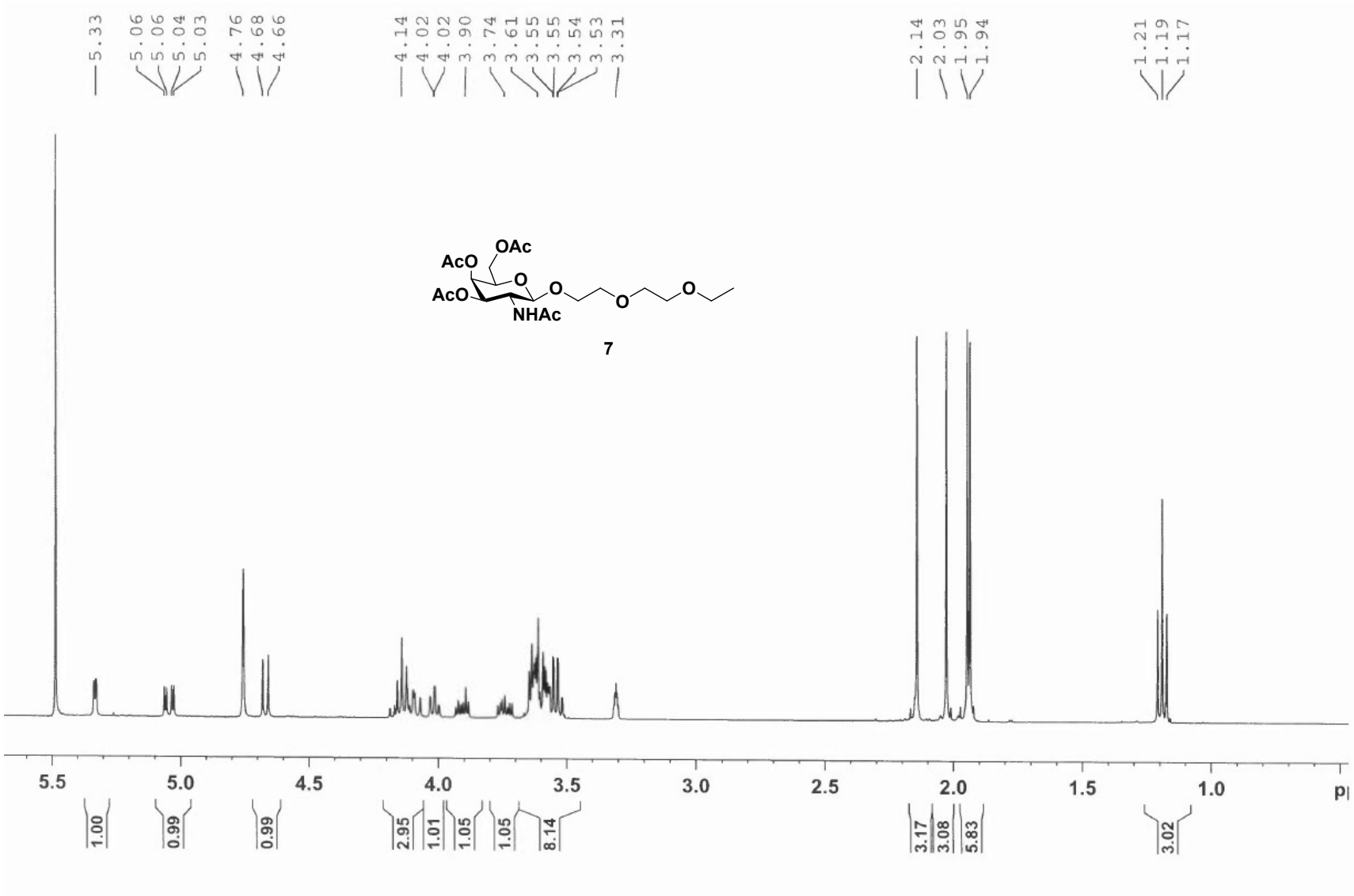


Fig. S12: ^{1}H NMR of compound **7**

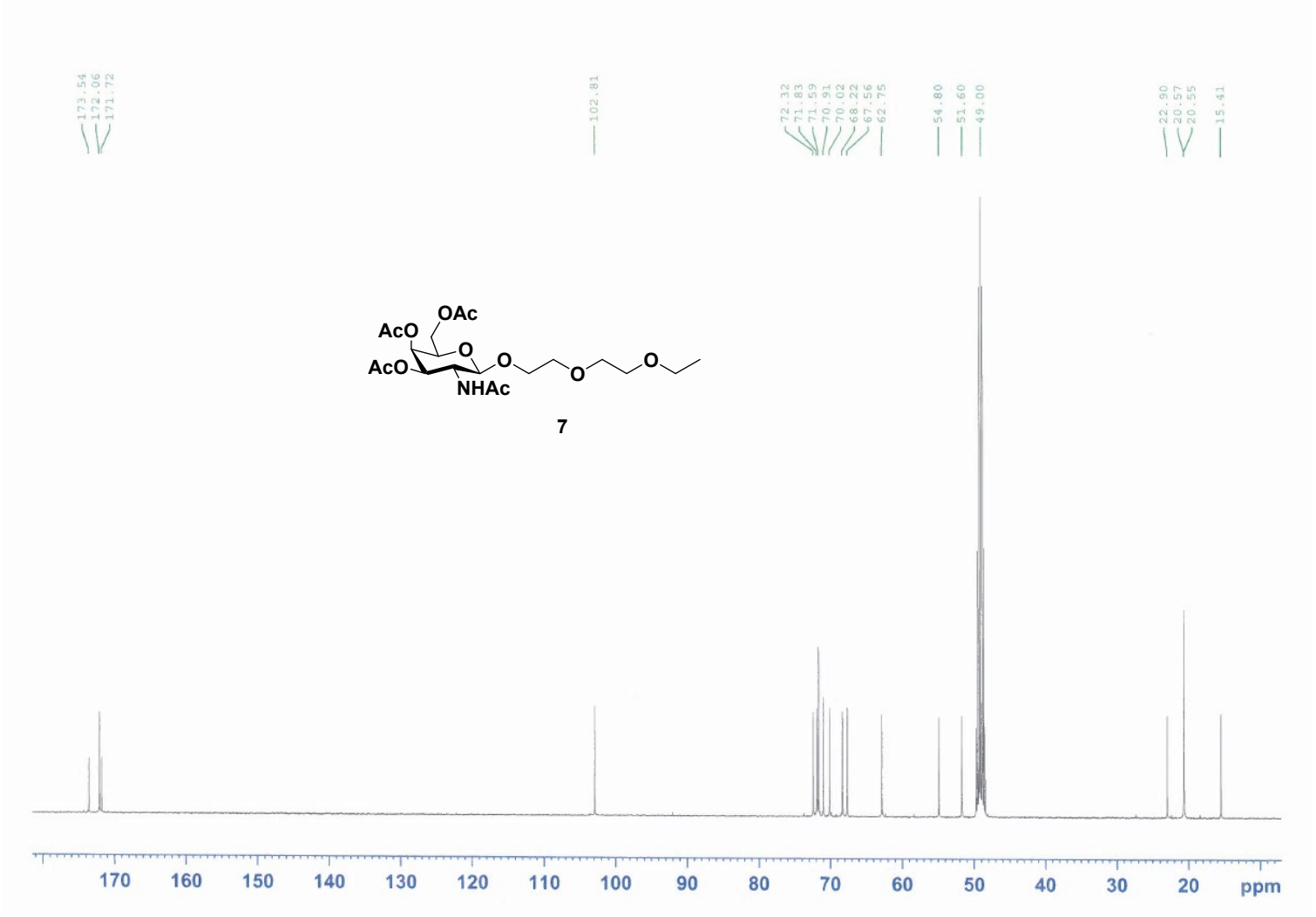


Fig. S13: ^{13}C NMR of compound **7**

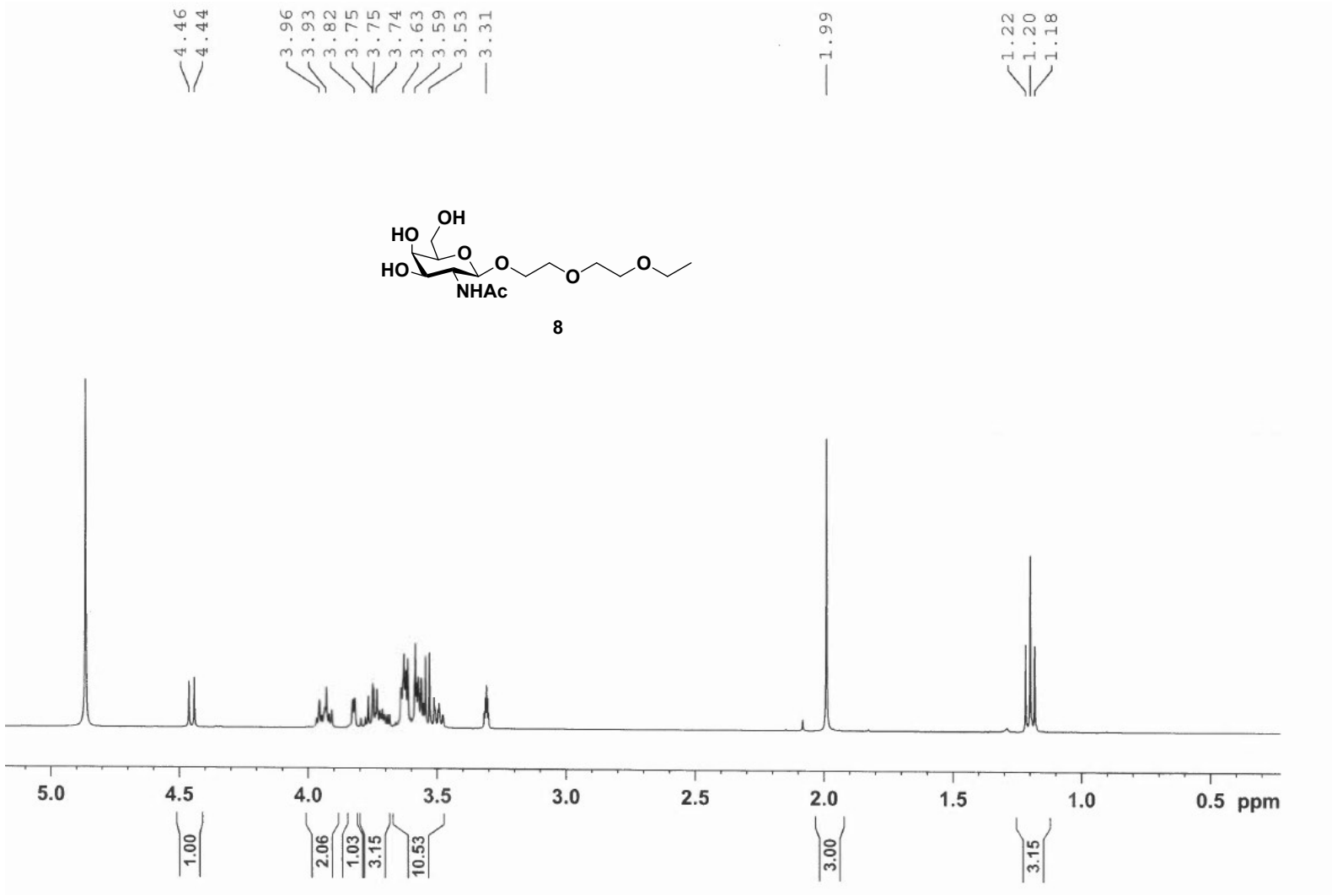


Fig. S14: ^{1}H NMR of compound **8**

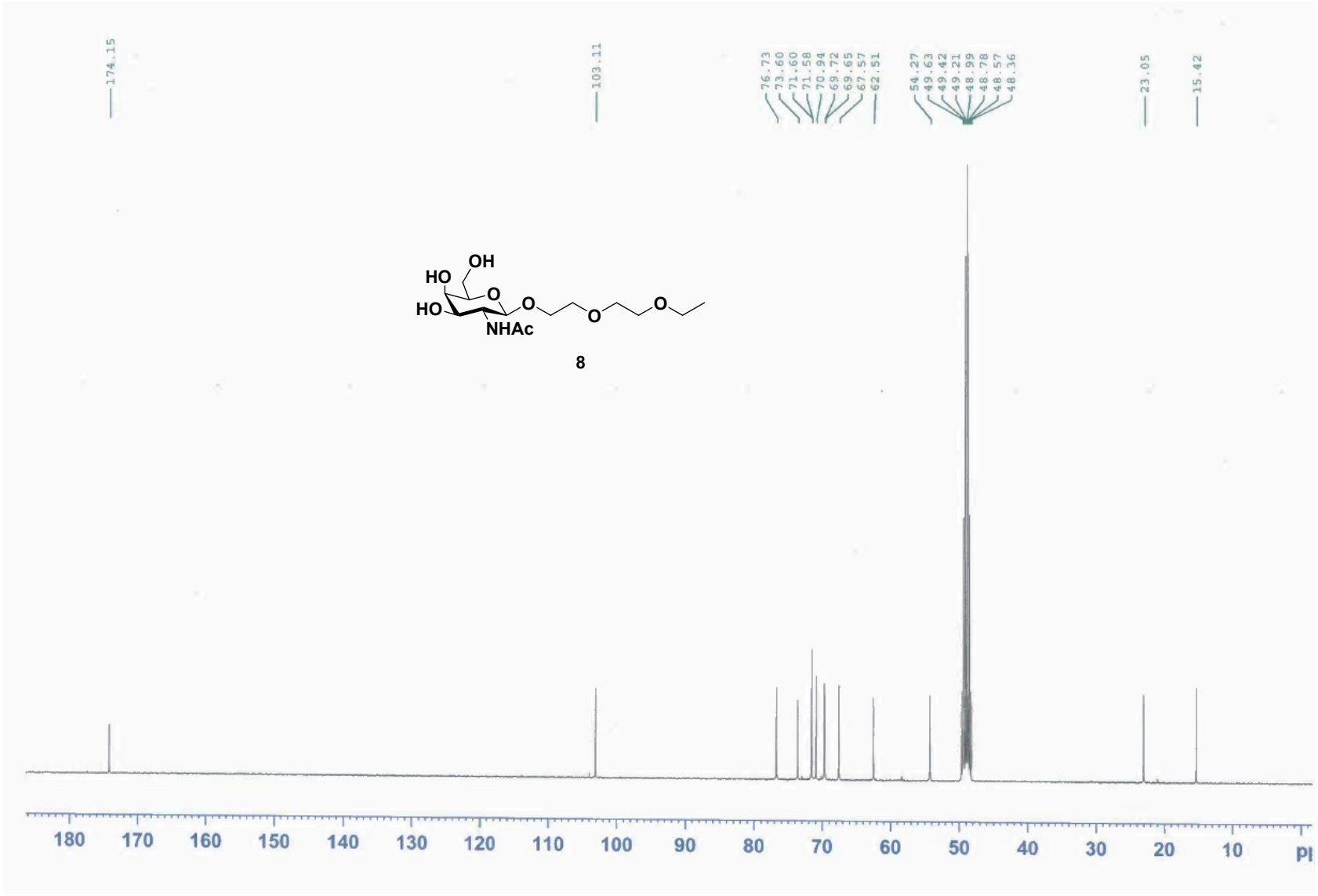


Fig. S15: ^{13}C NMR of compound **8**

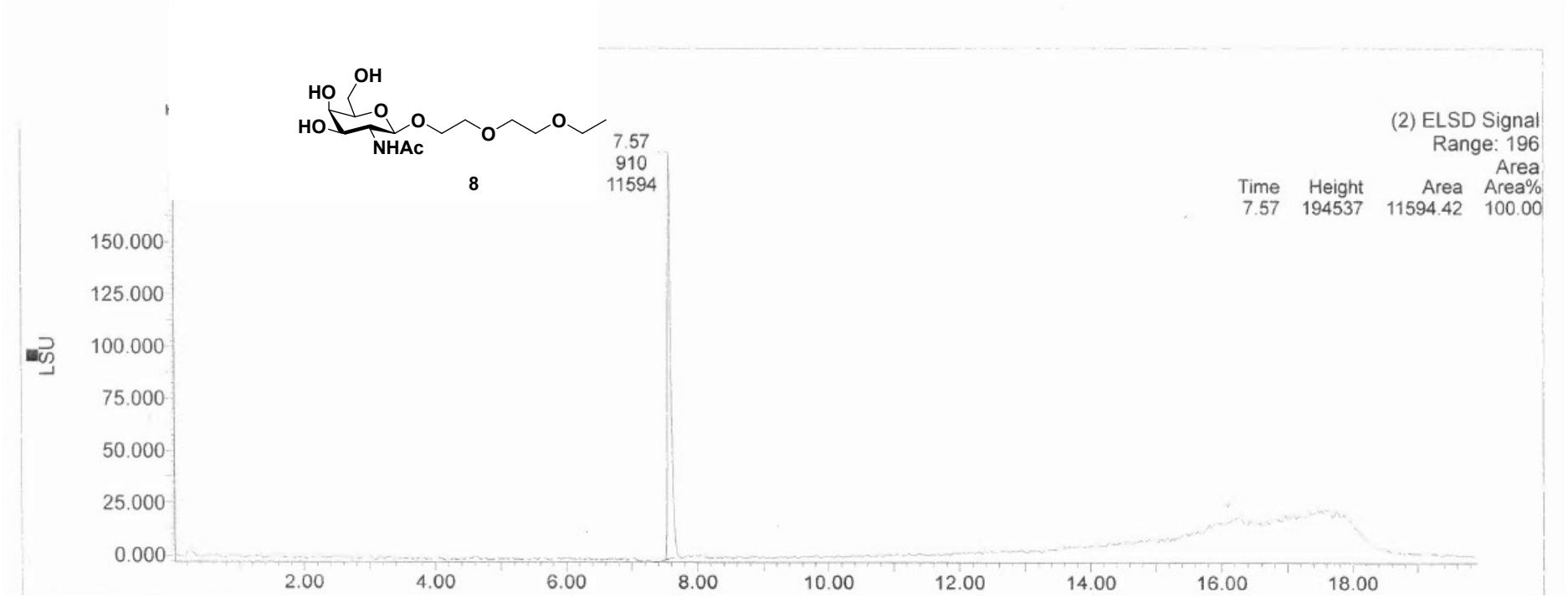


Fig. S16: Analytical HPLC of compound **8**

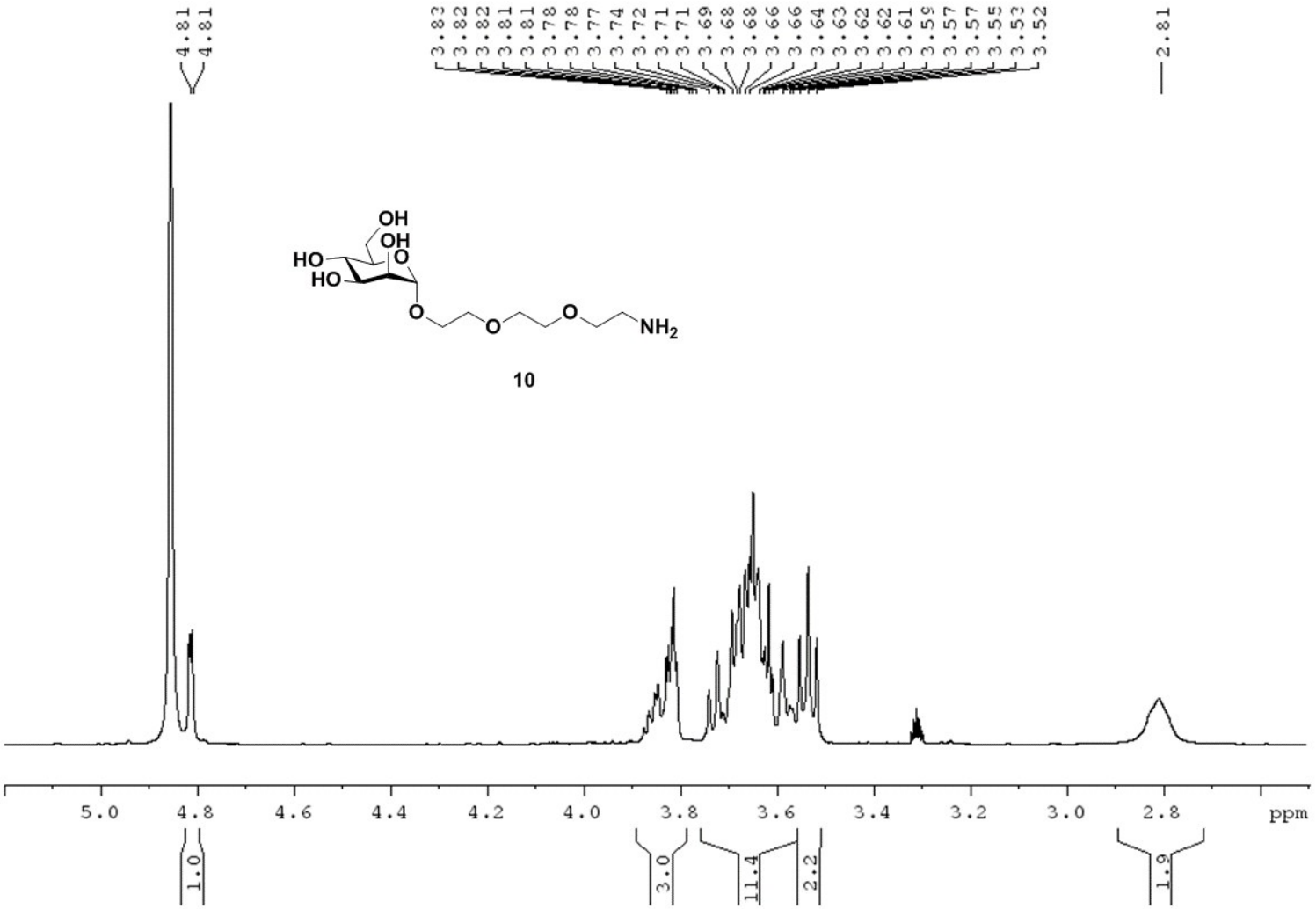


Fig. S17: ^{1}H NMR of compound **10**

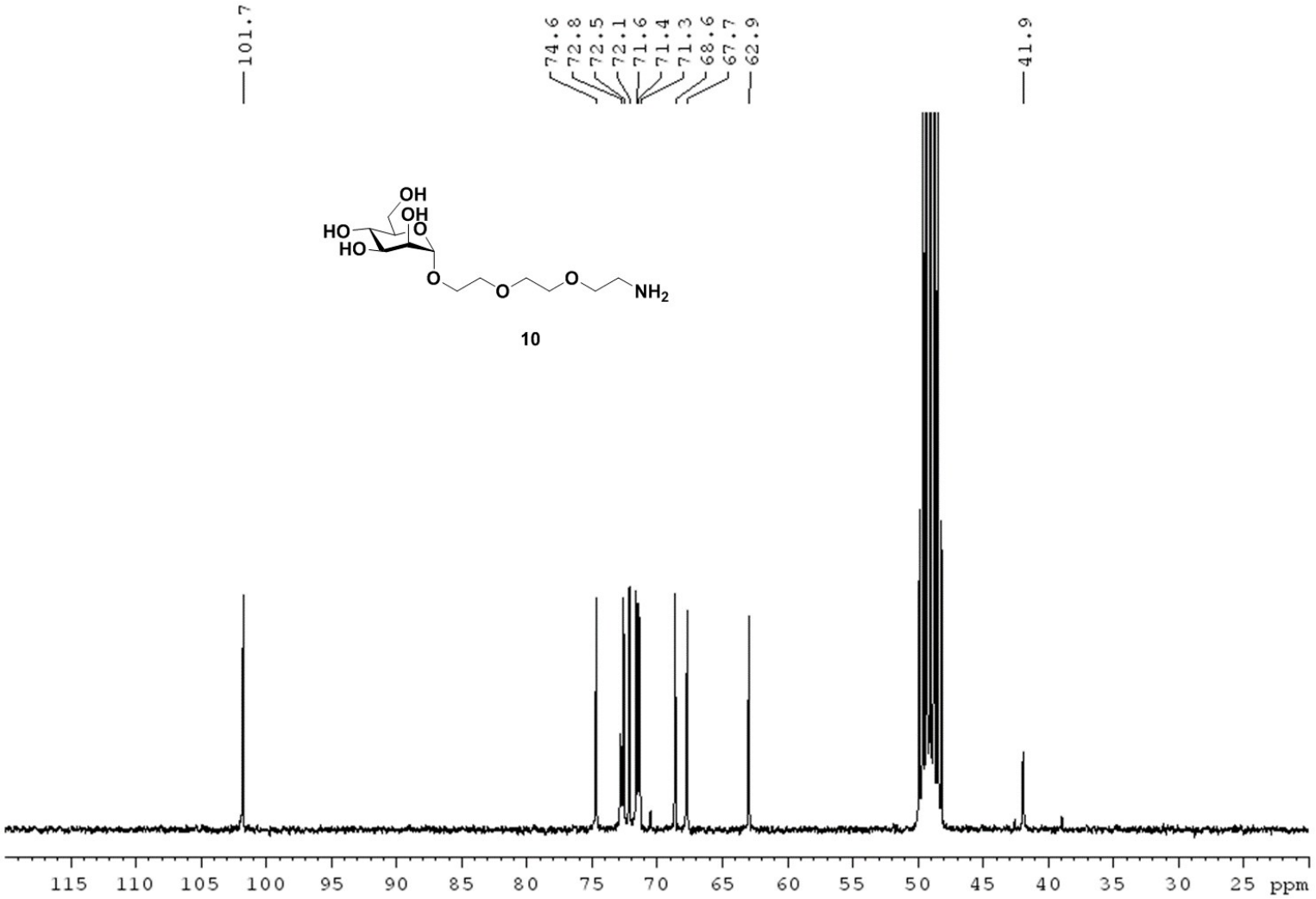


Fig. S18: ^{13}C NMR of compound **10**

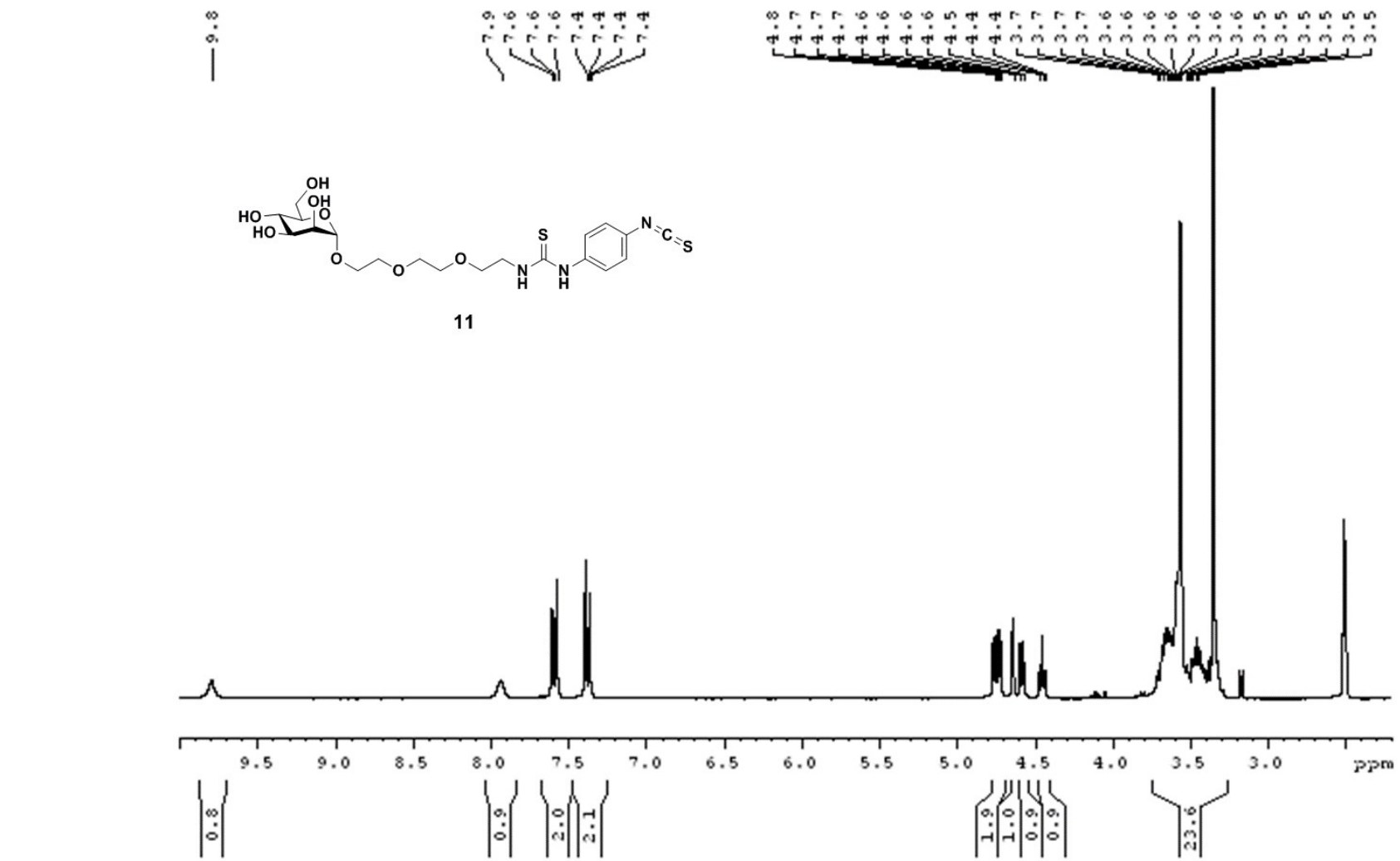


Fig. S19: ^{1}H NMR of compound **11**

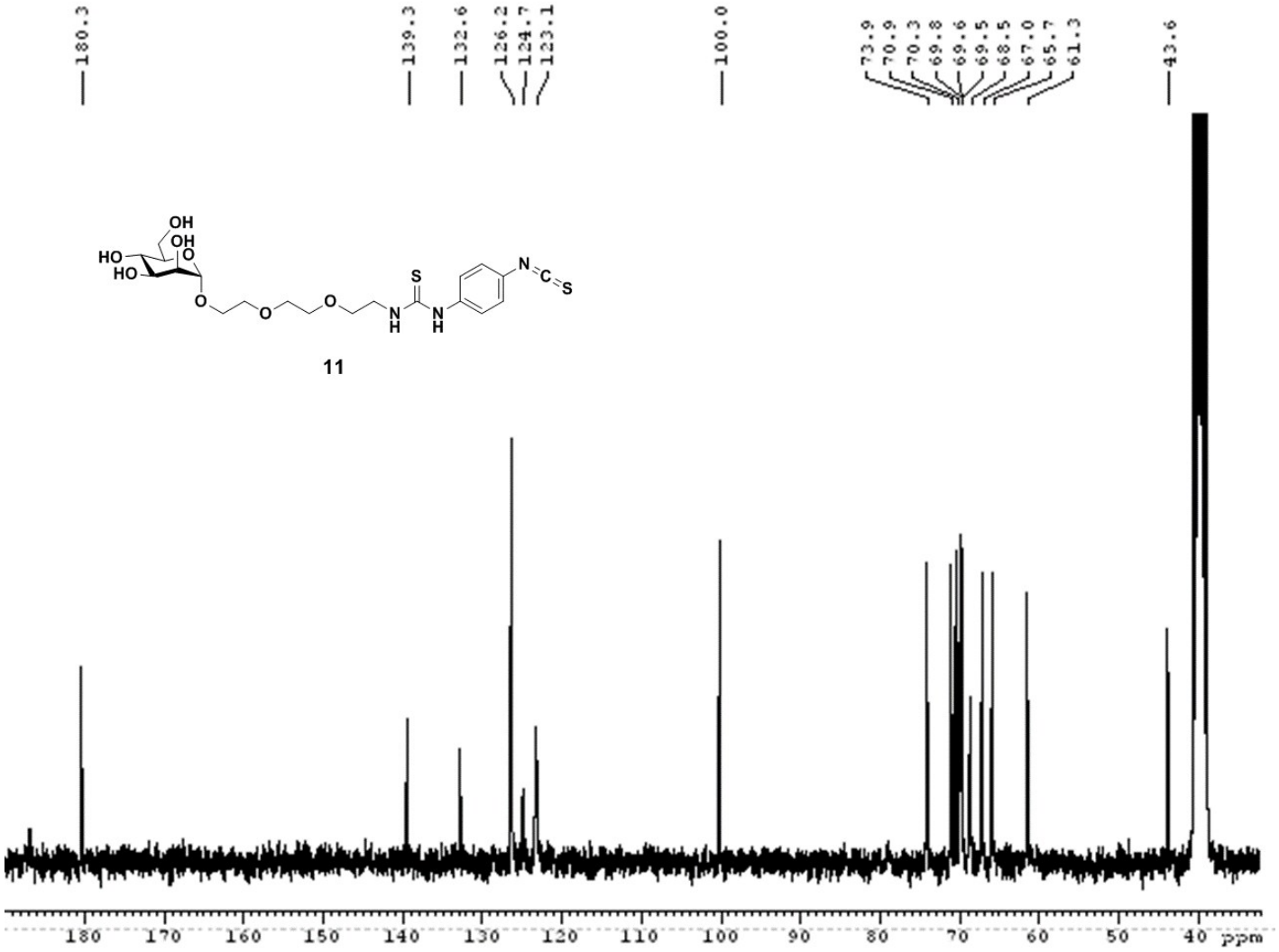


Fig. S20: ^{13}C NMR of compound **11**

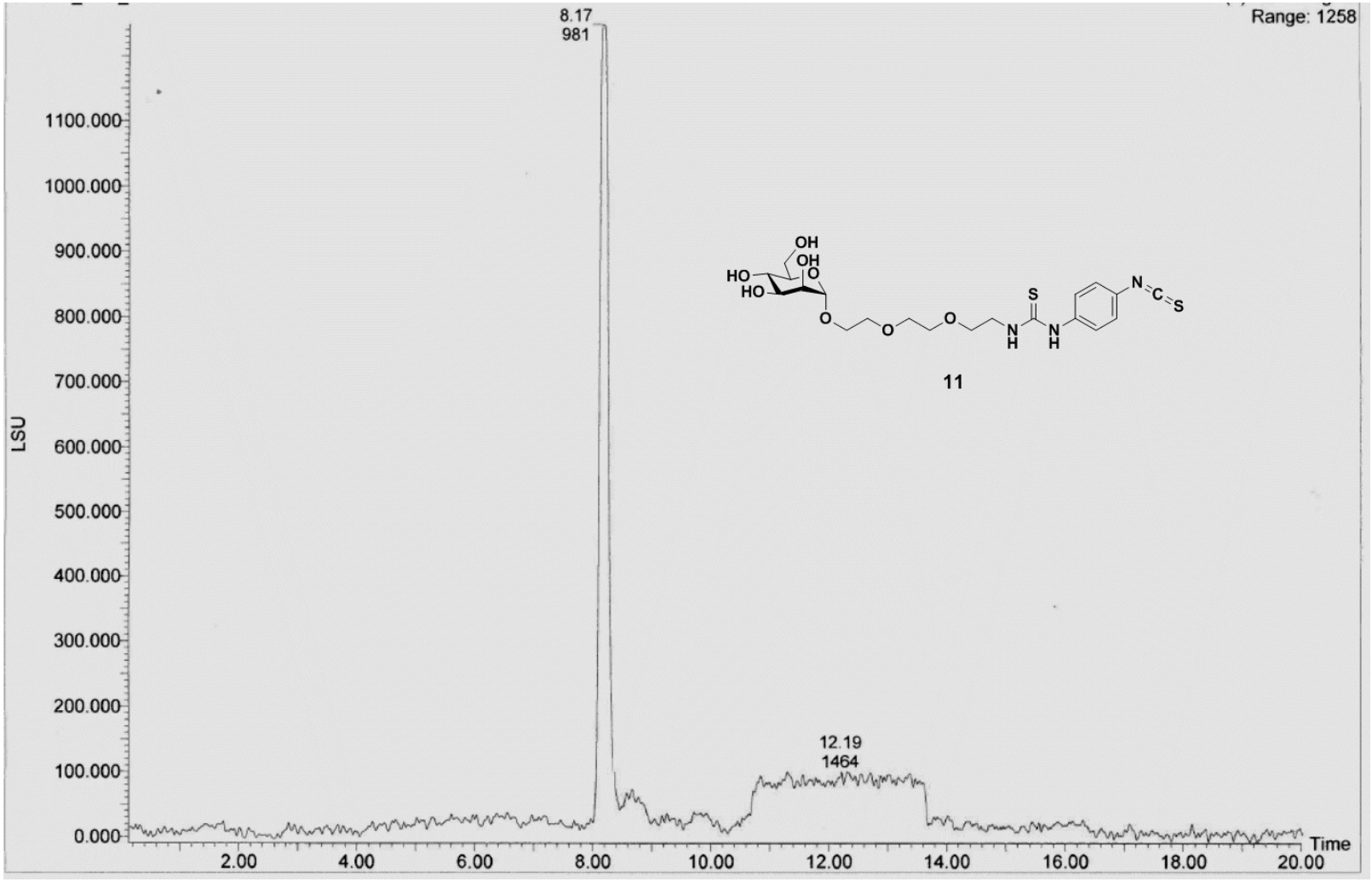


Fig. S21: Analytical HPLC of compound **11**

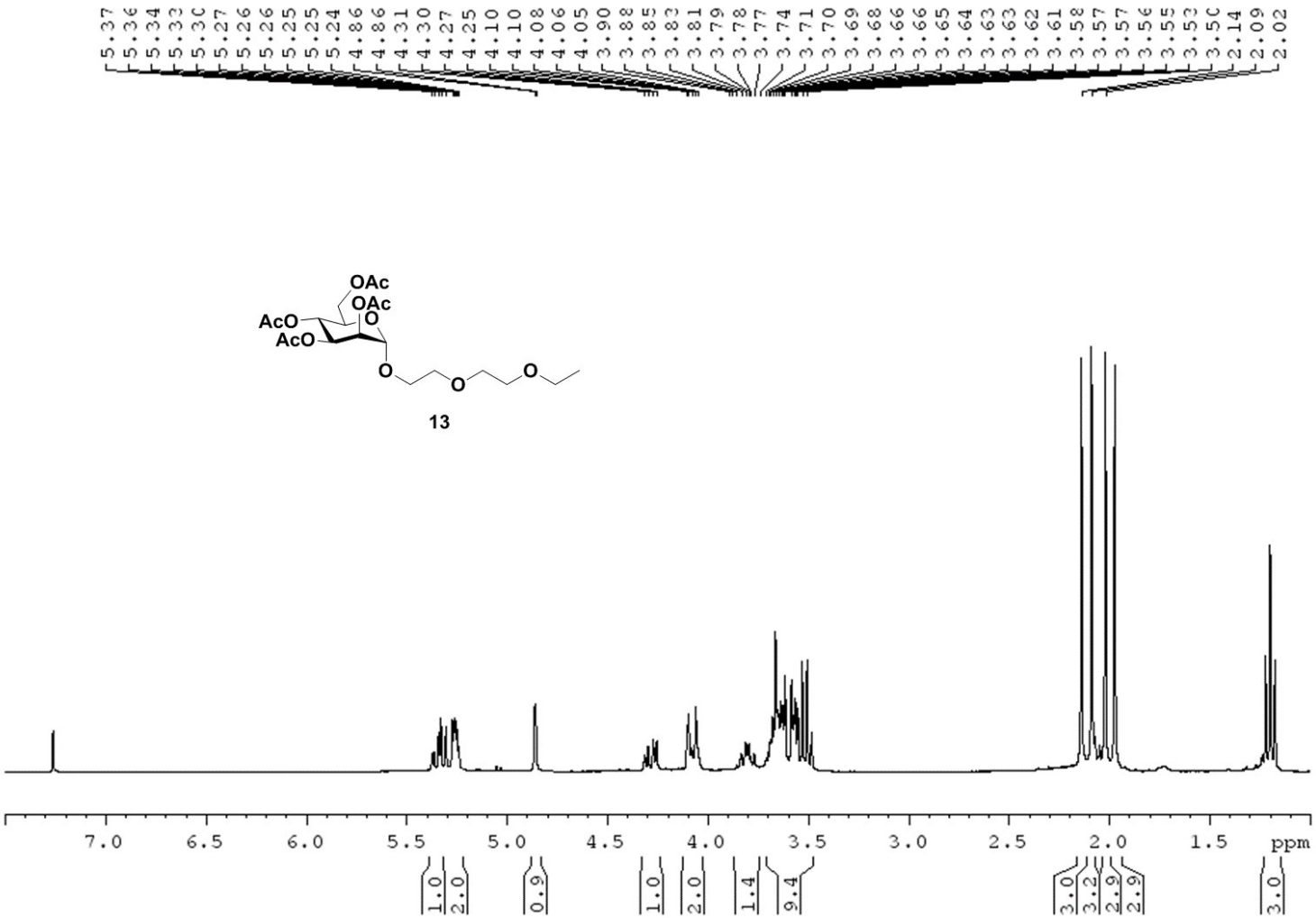


Fig. S22: ^{1}H NMR of compound **13**

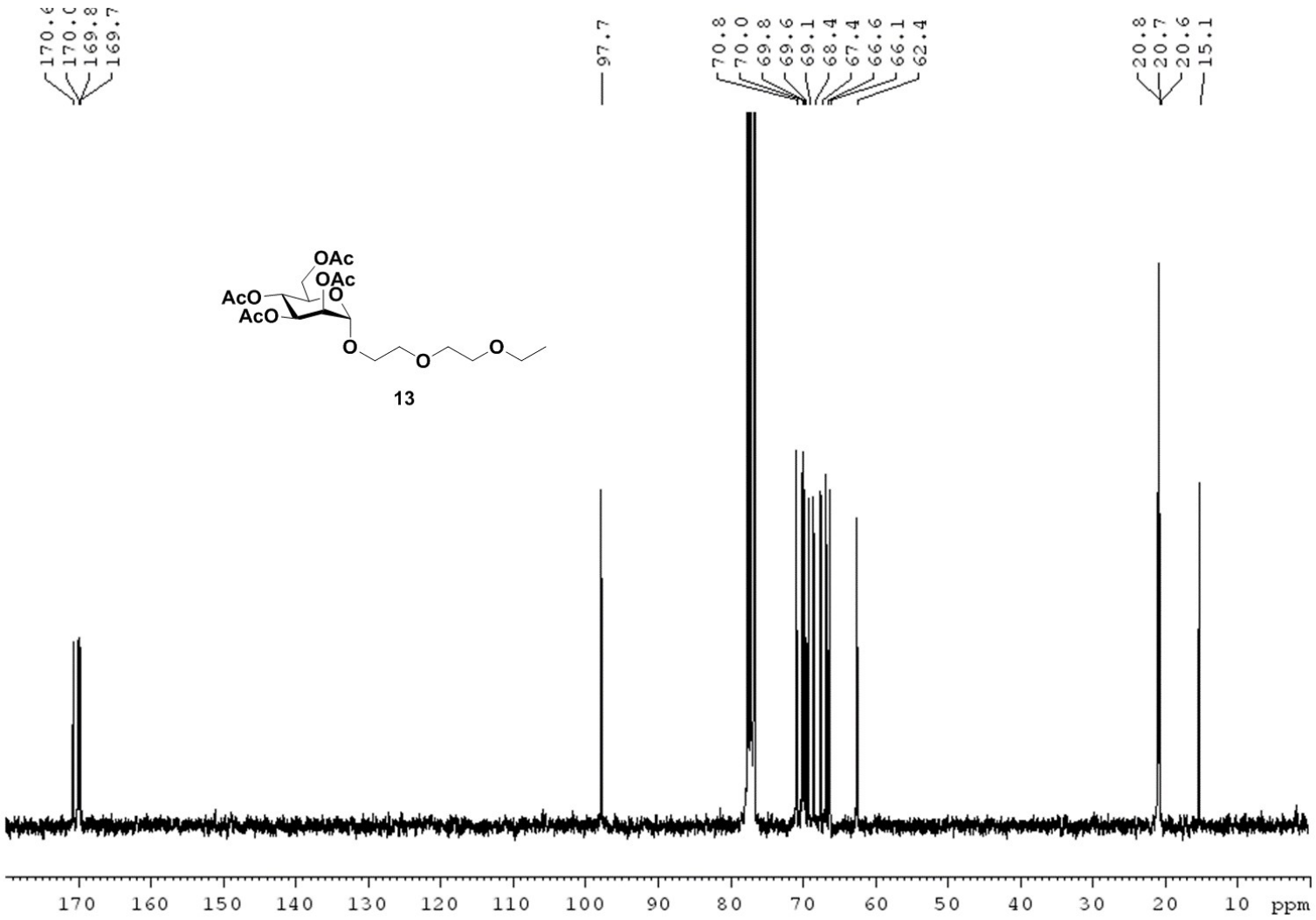


Fig. S23: ^{13}C NMR of compound **13**

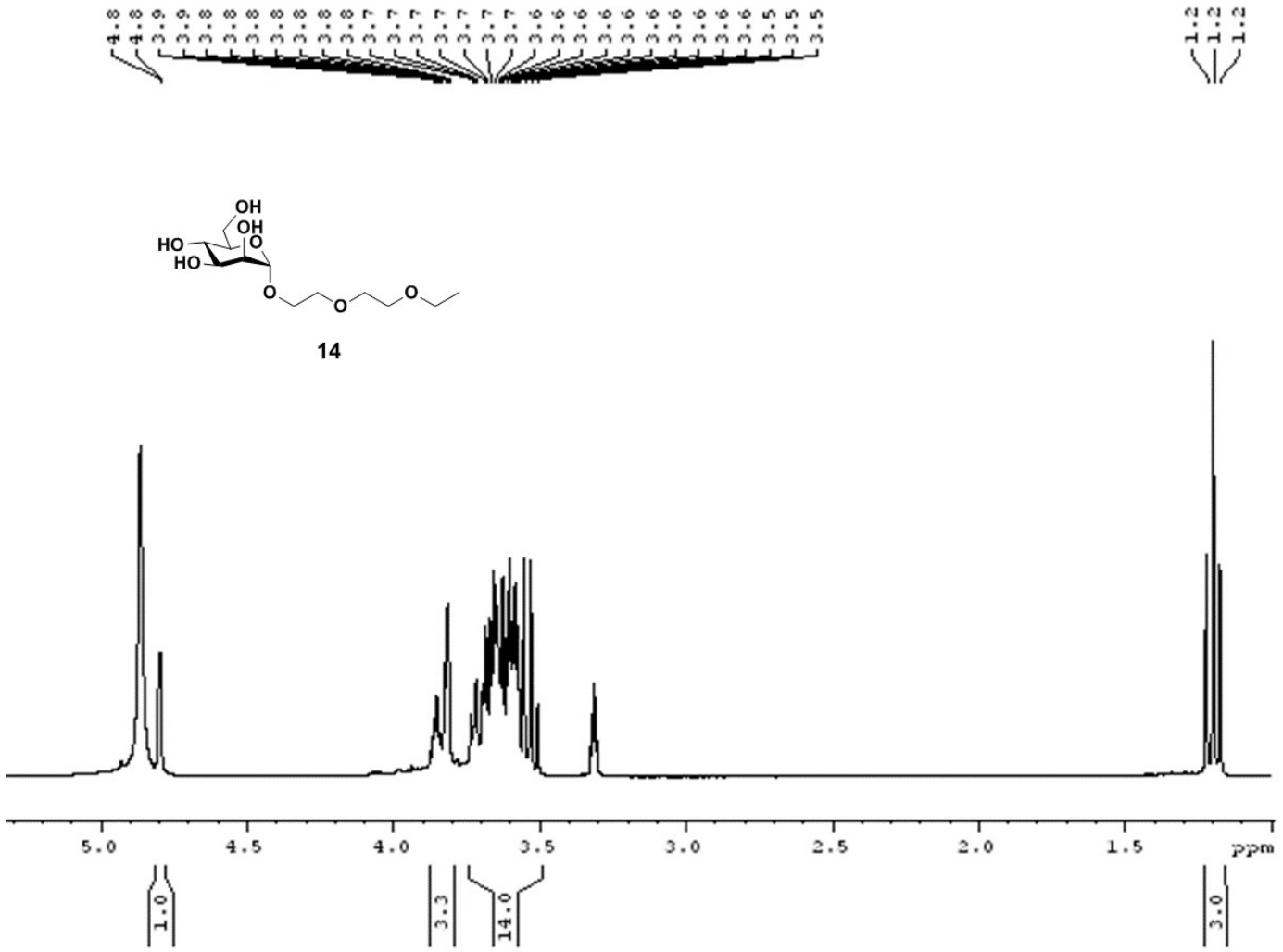


Fig. S24: ^{1}H NMR of compound **14**

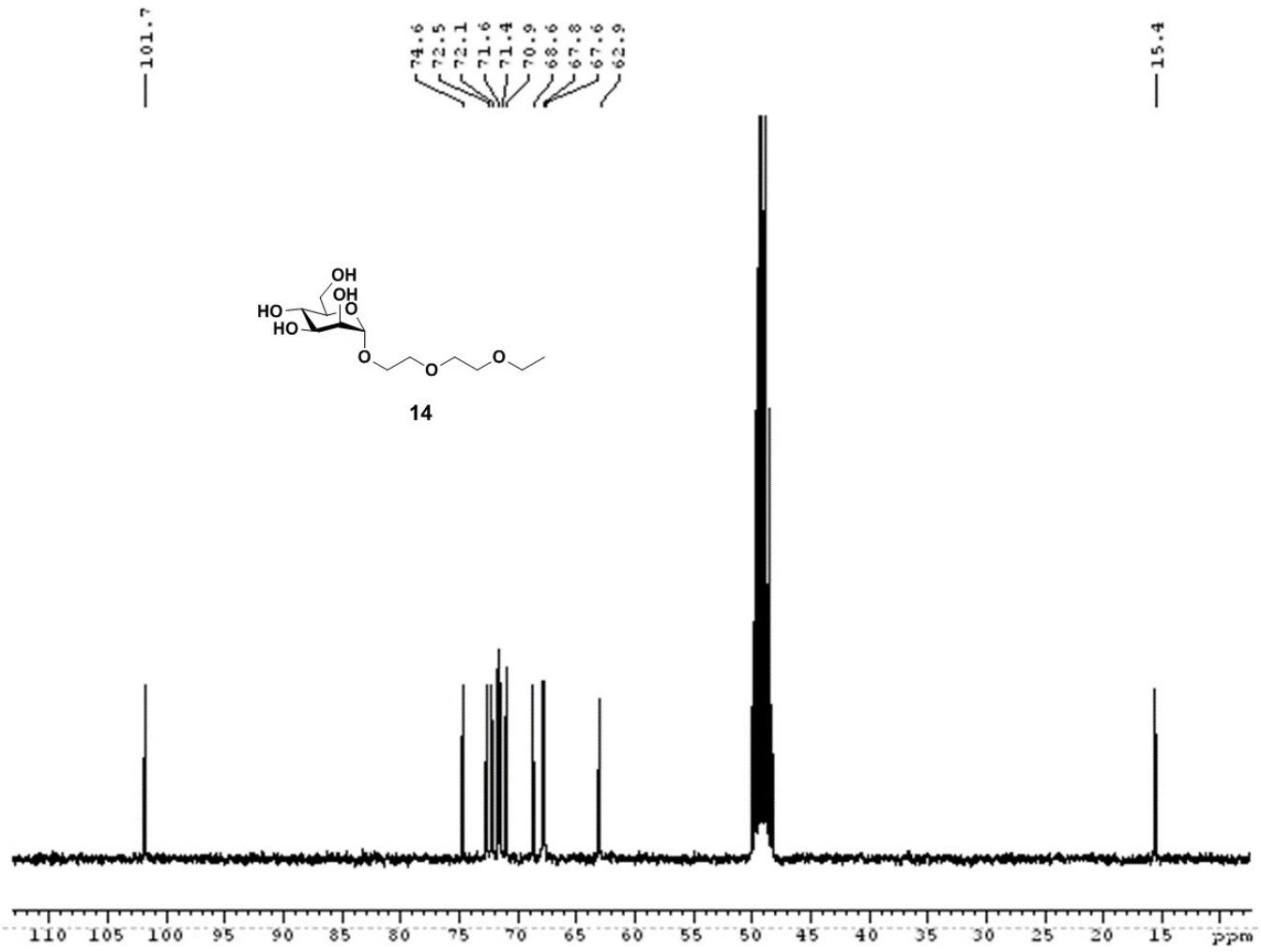


Fig. S25: ^{13}C NMR of compound **14**

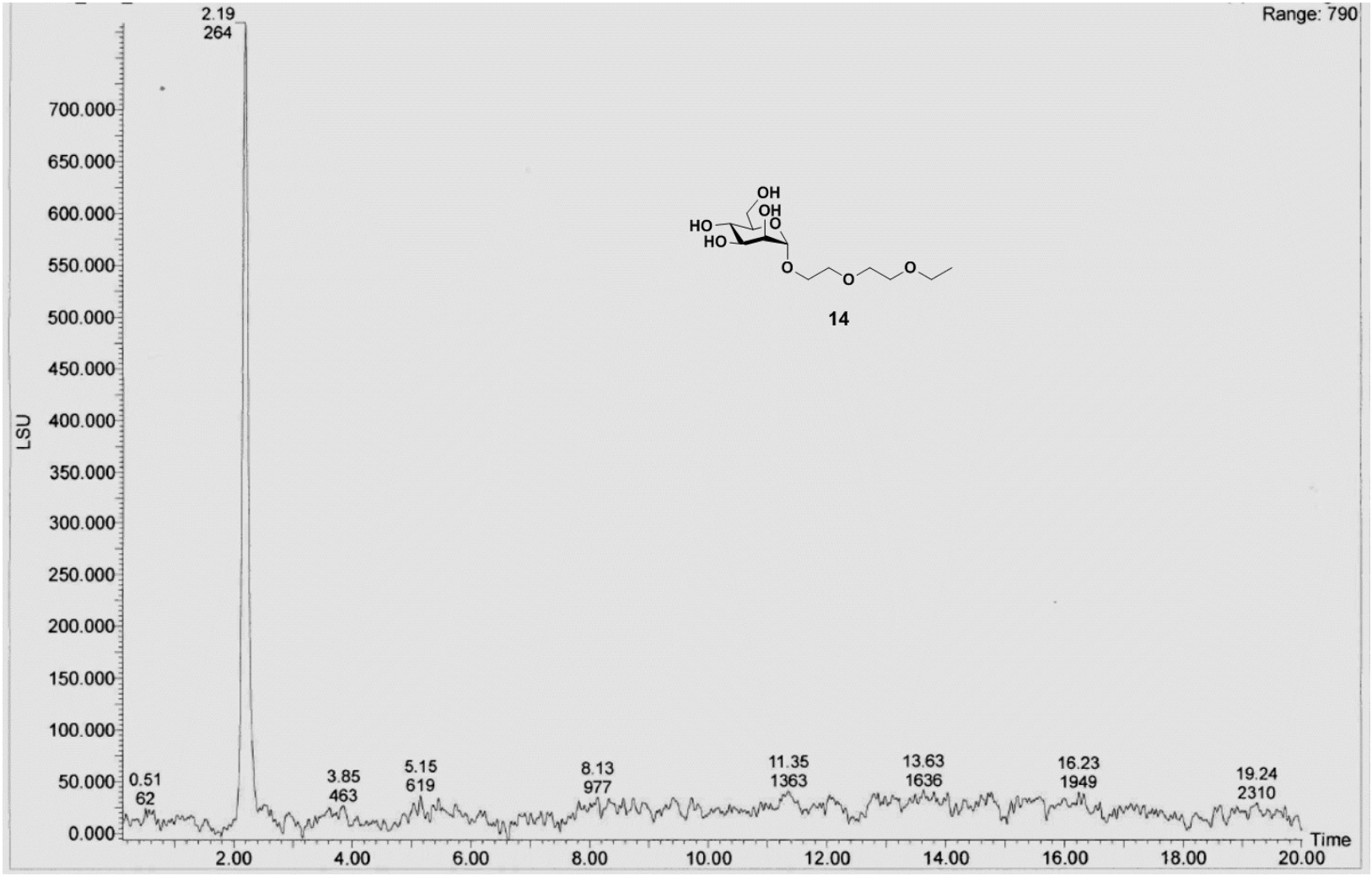


Fig. S26: Analytical HPLC of compound **14**

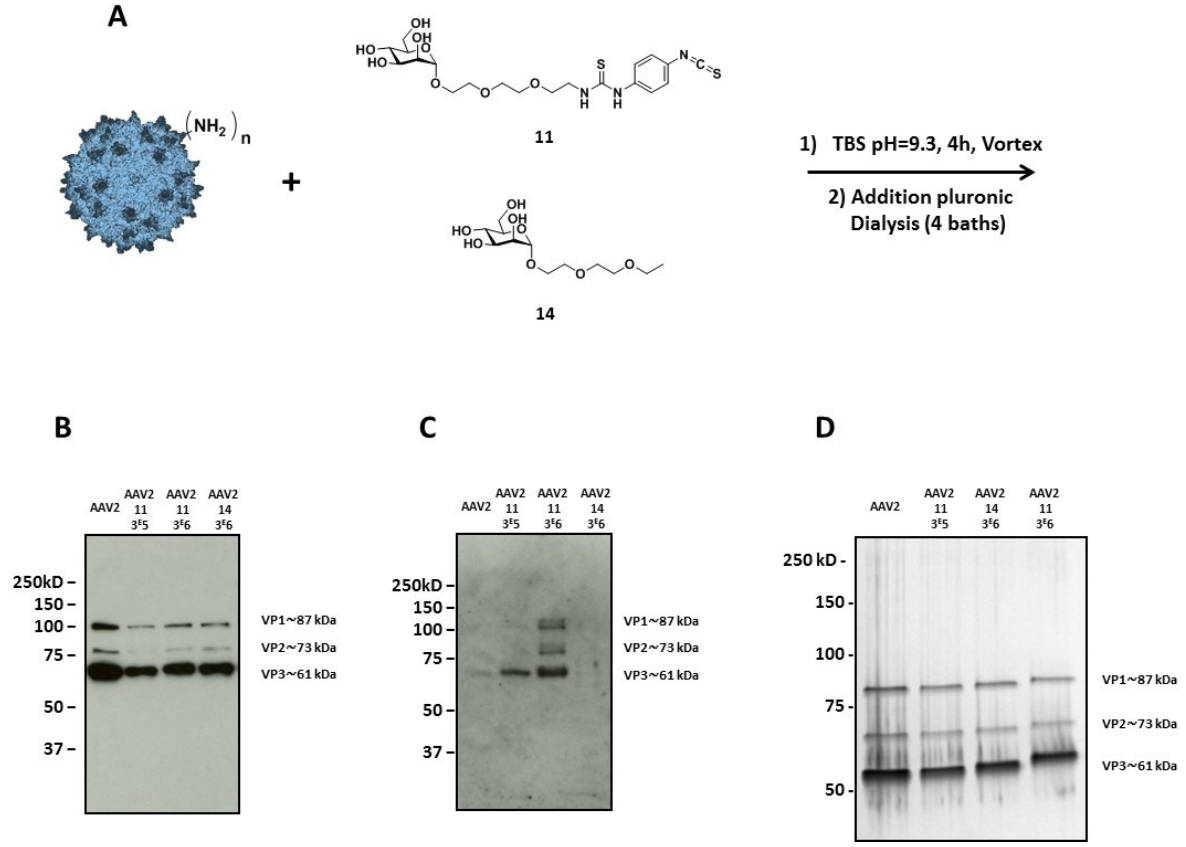


Fig. S27. Covalent coupling of 11 onto the capsid of AAV2 via primary amino groups. (**A**) 10^{12} vg of AAV2-GFP vectors were added to a solution of compound **11** (3E5 or 3E6 eq) in TBS buffer (pH 9.3) and incubated for 4 h at RT. The same experimental procedure was followed with compound **14** (3E6 eq) in TBS at pH 9.3 as a control. (**B, C**). Note that compound 14 is a negative control lacking the phi-NCS reactive function. 5E8 vg of the samples were analyzed by Western blot using a polyclonal antibody against the capsid to detect VP proteins **(B)** or using an FITC-concanavalin A lectin **(C).** (**D**) 10^{10} vg of each condition were analyzed by silver nitrate staining. VP1, VP2 and VP3 are the three proteins constituting the AAV capsid. Capsid protein molecular weight is indicated at the right of the images according to a protein ladder.

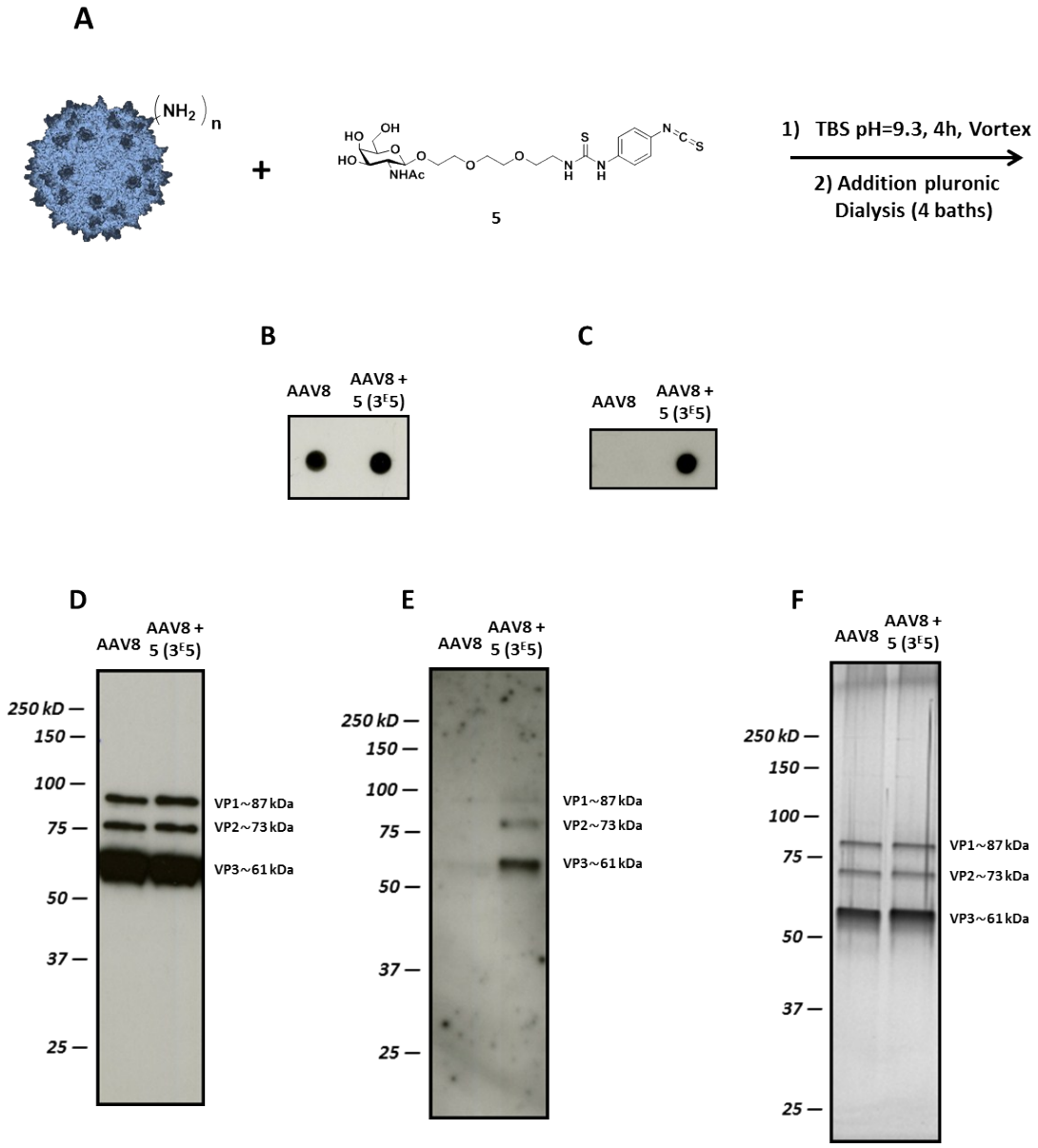


Fig. S28. Chemical modification of the AAV8 serotype with GalNAc molecules. (A) A dose of 10^{12} vg of AAV8-GFP vectors was added to a solution of **5** (3E5 eq) in TBS buffer (pH 9.3) and incubated for 4 h at RT. 10^{9} vg of each condition was analyzed by dot blot using the A20 antibody that recognizes the assembled capsid **(B)**, or using the soybean-FITC lectin that recognizes GalNAc sugar (**C**). 5E8 vg of the samples were analyzed by Western blot using a polyclonal antibody against the capsid to detect VP proteins **(D)** or using an FITC-Soybean lectin that recognizes GalNAc sugar **(E).** (**F**) 10^{10} vg of each condition were analyzed by silver nitrate staining. VP1, VP2 and VP3 are the three proteins constituting the AAV capsid. Capsid protein molecular weight is indicated at the right of the images according to a protein ladder.

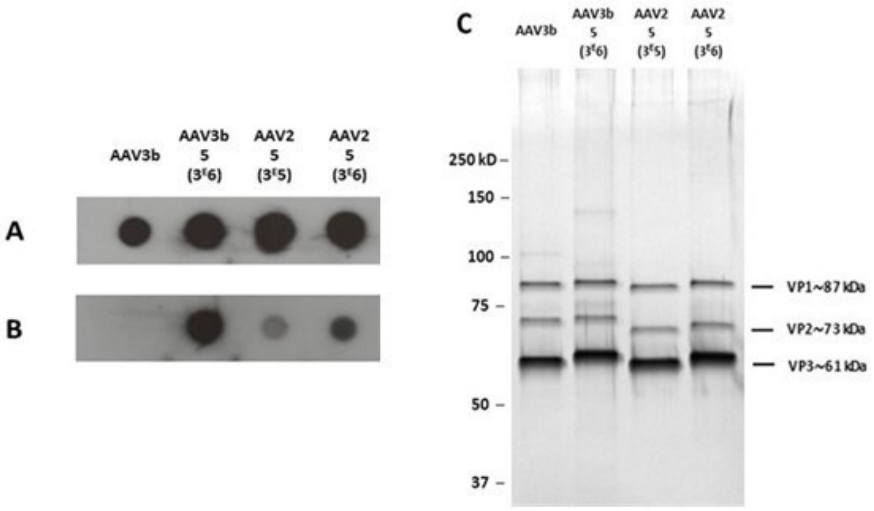


Fig. S29. Chemical modification of the AAV3b serotype with GalNAc molecules. A dose of 10^{12} vg of AAV3b was added to a solution of **5** (3E6 or 3E5 eq) in TBS buffer (pH 9.3) and incubated for 4 h at RT. **(A,B)** 10^{10} vg of each condition were used for dot blot analysis and labeled using the A20 antibody that recognizes the entire capsid (**A**) or using the soybean-FITC lectin that recognizes GalNAc sugar (**B**). (**C**) 10^{10} vg were analyzed by silver nitrate staining. Capsid protein molecular weight is indicated at the right of the images according to a protein ladder.

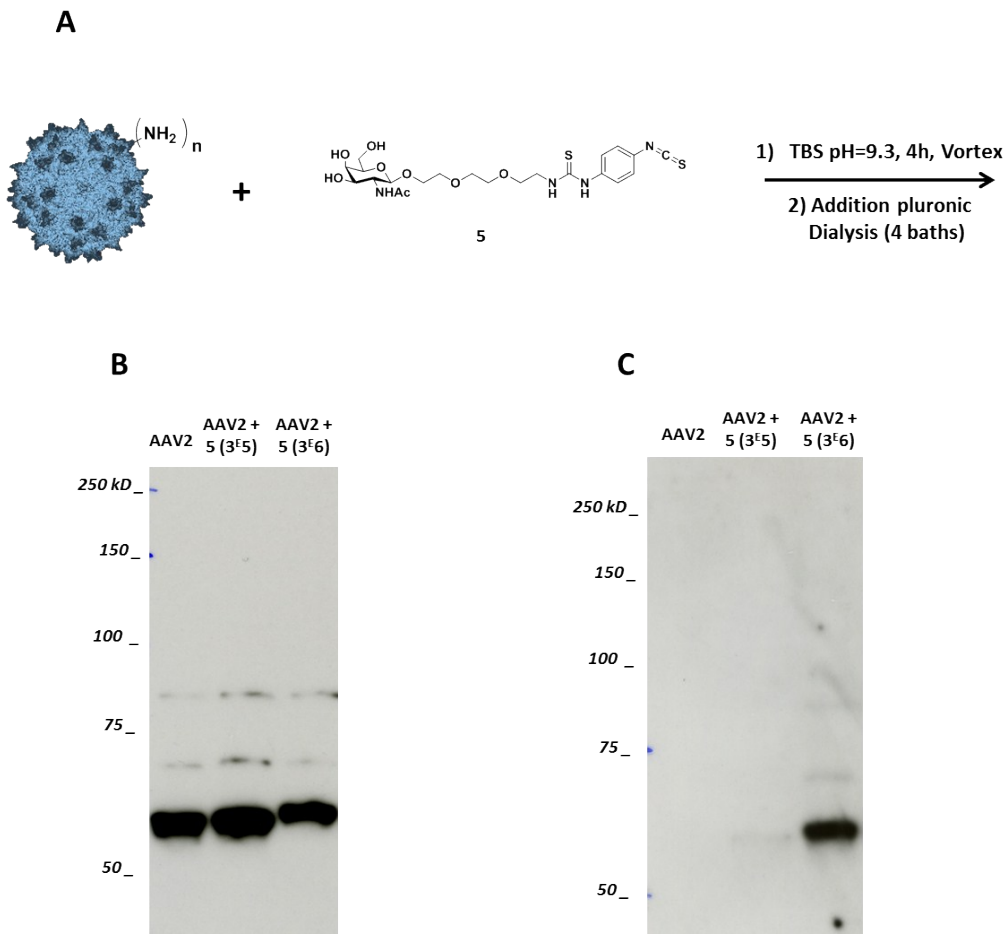


Fig. S30. Covalent coupling of 5 onto the capsid of AAV2-LacZ via primary amino groups. (**A**) 10^{12} vg of AAV2-LacZ vectors were added to a solution of compound **5** (3E5 or 3E6 eq) in TBS buffer (pH 9.3) and incubated for 4 h at RT. (**B, C**). 5E8 vg of the samples was analyzed by Western blot using a polyclonal antibody against the capsid to detect VP proteins **(B)** or using an FITC-soybean lectin **(C).** Capsid protein molecular weight is indicated at the right of the images according to a protein ladder.

Notes and References

1. P. C. Rensen, S. H. van Leeuwen, L. A. Sliedregt, T. J. van Berkel and E. A. Biessen, *J Med Chem*, 2004, **47**, 5798-5808.

2. S. D'Costa, V. Blouin, F. Broucque, M. Penaud-Budloo, A. Francois, I. C. Perez, C. Le Bec, P. Moullier, R. O. Snyder and E. Ayuso, *Molecular therapy. Methods & clinical development*, 2016, **5**, 16019.

3. Q. Xie, W. Bu, S. Bhatia, J. Hare, T. Somasundaram, A. Azzi and M. S. Chapman, *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**, 10405-10410.

4. M. Guilbaud, M. Devaux, C. Couzinie, J. Le Duff, A. Toromanoff, C. Vandamme, N. Jaulin, G. Gernoux, T. Larcher, P. Moullier, C. Le Guiner and O. Adjali, *Hum Gene Ther*, 2019, **30**, 802-813.