

Design and Synthesis of Active Heparan Sulfate-based Probes

Supplementary Information

Wen Zhou^{1,2,4}, Po-Hung Hsieh^{1,4}, Yongmei Xu¹, Timothy R. O’Leary, Xuefei Huang³ and

Jian Liu^{1*}

1. Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA
2. School of Pharmacy, Shanghai Jiaotong University, Shanghai, 200240, China
3. Department of Chemistry, Michigan State University, East Lansing, MI, USA
4. Both authors contributed equally to this work.

This work is supported in part by National Institutes of Health Grants HL094463 (to J.L.), GM102137 (to J.L.) and GM072667 (to X.H.). W.Z. was partially supported by a fellowship from the China Scholarship Council. T.R.O. is supported by Ruth L. Kirschstein Individual National Research Service Award from National Institutes of Health (HL120598).

*To whom correspondence should be addressed: Rm 1044, Genetic Medicine Building, University of North Carolina, Chapel Hill, NC 27599. Tel: 919-843-6511; E-mail: jian_liu@unc.edu.

Email contact information for each author:

Wen Zhou, wzhou60@sjtu.edu.cn; Po-Hung Hsieh, phsieh@email.unc.edu; Timothy R. O’Leary, troleary@email.unc.edu; Yongmei Xu, yongmeix@email.unc.edu; Xuefei Huang, xuefei@chemistry.msu.edu; Jian Liu, jian_liu@unc.edu.

Experimental Procedures

Synthesis of Compound 4 and 5- To a solution containing 3.7 g (0.05 mole) of glyoxylic acid in 100 ml of water at 50 °C, a prewarmed (about 50 °C) solution containing 9.3 g (0.05 mole) of *p*-toluenesulfonylhydrazid in 20% hydrochloric acid was added dropwise. The mixture was heated in a 60 °C water-bath until all the hydrazone had solidified. The reaction had been then allowed to cool down to room temperature and then kept in a 4°C freezer overnight. The crude product containing *p*-toluenesulfonyl-hydrazone was collected on a filter and dried at room temperature for two days. The crude product was dissolved in 100 mL of boiling ethyl acetate, filtered to get rid of any insoluble material and cooled down to room temperature. After the filtrate was kept overnight at 4°C, 8.6 g (71%) of *p*-toluenesulfonylhydrazone was collected in the form of white crystals.

Both 242 mg (1 mmol) of *p*-toluenesulfonylhydrazone and 120 mg (1.04 mmol) of *N*-hydroxylsuccinimide were dissolved in 50 mL of ice-cold tetrahydrofuran (THF). At 0 °C, dicyclohexylcarbodiimide (DCC, 216 mg, 1.05 mmol) in 5 mL (THF) was added drop wise in 30 min, and the reaction solution was stirring for 5 hours at room temperature under the atmosphere of nitrogen. Precipitates were removed by filtration and the filtrate was evaporated to dryness under reduced pressure. The crude product was purified by silica-gel chromatography with ethyl acetate/acetone (V:V=2:1) as eluent to give 125 mg (67.9%) succinimidyl diazoacetate **5**. ¹H NMR (400 MHz, CDCl₃): δ 2.78 (s, 4H, CH₂ (succinimide)), 5.05 (s, 1H, CHN₂) (**Supplementary Fig S1**). The data was consistent with the spectrum previously reported by Ouhia²¹.

To a suspension of NaH (100 mg, 60% dispersion in oil) in 6 mL of anhydrous THF, *p*-nitrophenol (278 mg, 2 mmol) dissolved in 6 mL of anhydrous THF was added

dropwise at 0 °C. Compound **5** (368 mg, 2 mmol) in 6 mL of THF was then added slowly at room temperature, and stirring was maintained for 20 min. The precipitates were removed by filtration, and the crude product was purified by silica gel with ethyl acetate/acetone (V:V=2:1) as eluent, affording 336.9 mg (81%) of the *p*-nitrophenyl diazoacetate **4**. ¹H NMR (400 MHz, CDCl₃): δ 4.99 (s, 1H, CHN₂), 7.28 (d, 2H, *J*=1.2, H_{Ar}); 8.21(d, 2H,*J*=1.2,H_{Ar}); ¹³C NMR (100 MHz, CDCl₃): δ 46.3, 121.4,124.3, 144.2, 154.3 (**Supplementary Fig S2** and **Supplementary Fig S3**). The data was in agreement with the previous report. ¹

Synthesis of compound 1, 2 and 3- HS oligosaccharides was synthesized from commercially available GlcA-PNP (1-*O*-(*para*-nitrophenyl)-glucuronide) by using *P. multocida* heparosan synthase 2 (pmHS2), *N*-sulfotransferase, C₅-epimerase/2-*O*-sulfotransferase, 6-*O*-sulfotransferases and UDP-GlcNTFA (uridine diphosphate *N*-trifluoroacetyl glucosamine) and UDP-GlcNAc (uridine diphosphate *N*-acetyl glucosamine), UDP-GlcA (uridine diphosphate glucuronic acid) and PAPS (3'-phosphoadenosine 5'-phosphosulfate) as described previously ². Subsequently, UDP-GlcNTFA was introduced to the non-reducing end of HS oligosaccharides at the presence of pmHS2 and then *N*-unsubstituted glucosamine HS oligosaccharides, which were key intermediates for synthesis of compounds **1** and **2**, were easily produced with the incubation with 0.1 M LiOH at 4 °C. *p*-Nitrophenyl diazoacetate **4** (6.3 mg, 31 μmol) was added to a solution of unsubstituted glucosamine hexasaccharide (4 mg, 3.1 μmol) dissolved in 0.5 mL of 2:1 dioxane/H₂O in the presence of 20 μL triethylamine. The reaction solution was stirred overnight. After dried, the crude residue was subjected to DEAE-HPLC to get 4.2 mg (99%) of compound **1** as white powders. Compound **2**

essentially followed the same procedure of compound **1** when unsubstituted glucosamine hexasaccharide was replaced with unsubstituted glucosamine octasaccharide. The synthesis of Compound **3** has been accomplished according to the procedure proposed by our group ².

NMR spectra were recorded for these compounds. ¹H-NMR for anomeric protons of compound **1** (850 MHz, D₂O): δ 5.63 (d, 1H), 5.60 (d, 1H); 5.41 (d, 1H); 5.27 (d, 1H); 4.51 (d, 1H); 4.47 (d, 1H) (**Supplementary Fig S4**). ¹³C-NMR for anomeric carbons of Compound **1** (212.5 MHz, D₂O): δ 105.0, 104.9, 102.0, 99.8, 99.8, and 99.8 (**Supplementary Fig S5**). ¹H-NMR for anomeric protons of compound **2** (850 MHz, D₂O): δ 5.59 (d, 1H), 5.41 (m, 3H); 5.40 (d, 1H); 5.39 (d, 1H); 5.27 (d, 1H); 5.21 (d, 1H); 4.56 (m, 2H) (**Supplementary Fig S7**). ¹³C-NMR for anomeric carbons of Compound **2** (212.5 MHz, D₂O): δ 104.6, 104.6, 102.0, 101.7, 100.3, 100.0, 99.5, and 99.2 (**Supplementary Fig S8**). ¹H-NMR for anomeric protons of compound **3** (700 MHz, D₂O): δ 5.63 (d, 1H), 5.60 (d, 1H); 5.39 (d, 1H); 5.27 (d, 1H); 4.51 (d, 1H); 4.47 (d, 1H) (**Supplementary Fig S10**). ¹³C-NMR for anomeric carbons of Compound **3** (175 MHz, D₂O): δ 105.0, 104.9, 102.0, 99.9, 99.8, and 99.6 (**Supplementary Fig S11**). ¹H-NMR for anomeric protons of compound **7** (850 MHz, D₂O): δ 5.69 (d, 1H), 5.58 (d, 1H); 5.40 (d, 1H); 5.39 (d, 1H); 5.27 (d, 1H); 5.21 (d, 1H); 4.56 (d, 1H); and 4.56 (d, 1H) (**Supplementary Fig S13**). ¹³C-NMR for anomeric carbons of Compound **7** (212.5 MHz, D₂O): δ 104.6, 104.6, 102.0, 101.8, 100.3, 99.5, 99.2, and 98.0 (**Supplementary Fig S14**).

Synthesis of compound 12, 13 and 14- Unsubstituted glucosamine disaccharides were prepared from GlcA-PNP by using pmHS2 and UDP-GlcNTFA for elongation

followed by de-*N*-trifluoroacetylation under LiOH condition. 2-Bromoacetic acid (3.1 mg, 21.0 μmol) dissolved in 0.5 mL of 1:1 dioxane/ H₂O was added dropwise to a solution of unsubstituted glucosamine disaccharide (5 mg, 10.5 μmol) and potassium carbonate (14.5 mg, 105.0 μmol) in 0.5 mL of H₂O. The reaction was completed at room temperature overnight. The resulting mixture was purified by C₁₈ column followed by a DEAE (diethylaminoethyl) column purification to give 2.8 mg (49.9 %) compound **12**. ¹H-NMR for anomeric protons of compound **12**: (850 MHz, D₂O) δ 5.83 (d, 1H) and 5.26 (d, 1H) (**Supplementary Fig S21**). Similarly, to a solution of unsubstituted glucosamine disaccharide (5 mg, 10.5 μmol) and triethylamine (14.7 μL , 105.0 μmol) in 0.5 mL of 2:1 dioxane/H₂O, methyl 2-hydroxyacetate (4.95 mg, 52.5 μmol) was added. Following the same purification as described for compound **12**, 3.5 mg (63%) of compound **13** was obtained as white powders. ¹H-NMR for anomeric protons of compound **13**: (850 MHz, D₂O) δ 5.41 (d, 1H) and 5.30 (d, 1H) (**Supplementary Fig S22**). 2-hydroxyacetic acid was activated with *N*-hydroxylsuccinimide at the presence of DCC to produce succinimidyl 2-hydroxyacetate to prepare compound **14** as described previously²¹. Treatment of unsubstituted glucosamine hexasaccharide (5 mg, 3.8 μmol) with succinimidyl 2-hydroxyacetate (6.5 mg, 38 μmol) gave 3.9 mg (74.7%) of compound **14** as white powders. ¹H-NMR for anomeric protons of compound **14** (850 MHz, D₂O): δ 5.63 (d, 1H), 5.60 (d, 1H); 5.41 (d, 1H); 5.27 (d, 1H); 4.51 (d, 1H); and 4.47 (d, 1H) (**Supplementary Fig S24**). ¹³C-NMR for anomeric carbons of Compound **14** (212.5 MHz, D₂O): δ 105.0, 104.9, 102.0, 99.9, 99.8 and 99.6 (**Supplementary Fig S25**).

Expression of 2-OST and 3-OST –HS 2-*O*-sulfotransferase (2-OST) and 3-*O*-sulfotransferase isoform 1 (3-OST) were expressed in *E. Coli* and purified by appropriate

affinity chromatography as described previously². The proteins were further purified by preparative GPC column to reach homogeneity as determined by 11% SDS-PAGE analysis.

Determination of the activities of 2-OST and 3-OST- Activity measurement for 2-OST using the completely desulfated *N*-sulfated heparin (CDSNS, purchased from Neoparin) substrate was determined by incubating 0.5 µg of 2-OST with 1 µL of CDSNS (10 mg/mL) and 1~5 ×10⁵ cpm of [³⁵S]PAPS in 200 µL of buffer containing 50 mM MES(2-(*N*-morpholino)ethanesulfonic acid) (pH 7.2), 10 mM CaCl₂ at 37 °C for 1 h. Reactions were quenched by the addition of 6 M urea, 1mM EDTA (ethylenediaminetetraacetic acid), 0.01% Triton X-100, then subjected to 200-µL DEAE-Sephrose (from Sigma) chromatography to purify the ³⁵S-labeled polysaccharide product. The quantity of [³⁵S]sulfated polysaccharide was determined by liquid scintillation counting. The negative control contained all of the components with the exception of 2-OST enzyme.

Activity measurement for 3-OST using the heparan substrate (isolated from bovine kidney) was determined by incubating 0.5 µg of 3-OST with 1 µL of HS (10 mg/mL) and 1~5 ×10⁵ cpm of [³⁵S]PAPS in 200 µL of buffer containing 50 mM MES (pH 7.2), 10 mM MnCl₂ and 5 mM MgCl₂ at 37 °C for 1 h. Work-up of reaction mixture and activity measurement followed essentially the same procedure as described for 2-OST.

Coupling 2-OST and compound 1- 2-OST (30 µL, 2.9 mg/mL) was incubated with 10 µg of compound **1** in 100 µL of buffer containing 25 mM MES (pH 5.5), 10 mM CaCl₂ at 4 °C for 4 h. The reaction mixture was directly subjected to GPC-HPLC for analysis.

Determination of the inhibition effect of compound 1 on 2-OST- Prior to analyzing the inhibition effect of compound **1** on the activity of 2-OST enzyme, 1 µg of 2-OST was deactivated by incubating with various amount (0, 1, 2, 4, 6 and 10 µL) of compound **1** (5 µM) at pH 5.5 to initiate the cross-linking reaction in 100 µL of 25 mM MES buffer. After 1 h incubation, the inhibition effect of compound **1** on 2-OST was determined by incubating 10 µL of the deactivated 2-OST protein above, 10 µg of CDSNS heparin and $1\sim 5 \times 10^5$ cpm of [³⁵S]PAPS in 100 µL of a buffer containing 50 mM MES (pH 7.2), 10 mM CaCl₂. All the reactions were incubated at 37 °C for 1h and quenched by the addition of a buffer containing 6 M urea, 1mM EDTA, and 0.01% Triton X-100. The samples were then loaded onto a 200-µL DEAE-Sepharose (Sigma-Aldrich) column to purify the [³⁵S] CDSNS heparin, which were eluted with 1M NaCl and 0.01% Triton X-100. The quantity of [³⁵S] HS was determined by liquid scintillation counting. The inhibition activity of compound **1** against 2-OST was showed by IC₅₀ (half maximal inhibitory concentration) value that was calculated by linear regression analysis of the concentration-response curves by means of Sigmaplot 12.5.

Determination of the inhibition effect of compound 2 on 3-OST- 3-OST (1 µg) was deactivated by incubating it with various amount (0, 0.5, 1.0, 1.5 and 2.0 µL) of compound **2** (2.0 µM) at pH 5.5 to initiate the cross-linking reaction in 100 µL of 25 mM MES buffer. After 1 h incubation, the inhibition effect of compound **2** on 3-OST was determined by incubating 10 µL of the deactivated 3-OST protein above, 10 µg of HS and $1\sim 5 \times 10^5$ cpm of [³⁵S]PAPS in 100 µL of a buffer containing 50 mM MES (pH 7.2), 10 mM MnCl₂, 5 mM MgCl₂. All the reactions were incubated at 37 °C for 1h and quenched by the addition of a buffer containing 3 M urea, 1mM EDTA, and 0.01%

Triton X-100. The samples were then loaded onto a 200- μ L DEAE-Sepharose (Sigma-Aldrich) column to purify the [35 S]HS, which were eluted with 1M NaCl and 0.01% Triton X-100. The quantity of [35 S] HS was determined by liquid scintillation counting. The inhibition activity of compound **2** against 3-OST was showed by IC₅₀ value that was calculated by linear regression analysis of the concentration-response curves by means of Sigmaplot 12.5.

HPLC analysis- Both DEAE-HPLC and polyamine-based anion exchange (PAMN)-HPLC were used to analyze and purify the oligosaccharides. DEAE-NPR column (0.46 \times 7.5 cm, Tosohaas) was used for DEAE-HPLC, and a silica-based polyamine column (0.45 \times 25 cm, YMC) was used for PAMN-HPLC. The elution conditions for the HPLC analysis were described elsewhere ³. GPC-HPLC was used to analyze and purify protein and protein-oligosaccharide conjugates. The GPC-HPLC conditions were as follows: the column G3000SW from Tosoh Bioscience (7.8 mm \times 300 mm, 5 μ m) the elution was isocratic; the buffer consisted of 25 mM MOPS and 400 mM NaCl; the dual wavelengths were set as 280 nm and 310 nm respectively, the flow rate was 0.2 mL/min.

MS analysis- The analyses were performed at a Thermo LCQ-Deca. The nonsulfated oligosaccharide (1 μ L) eluted from C₁₈ column was directly diluted in 200 μ L H₂O. A syringe pump was allowed to introduce the sample *via* direct infusion (50 μ L/min). Experiments were carried out in negative ionization mode with the electrospray source set 3 KV and 250 $^{\circ}$ C. Sulfated oligosaccharide (1 μ L) was diluted in a different working solution containing 200 μ L of 70% acetonitrile and 10 μ M imidazole. Experiments for sulfated oligosaccharides were carried out in negative ionization mode with the

electrospray source set to 2 KV and 175 °C. Nitrogen was used for both nebulizer (5 L/min) and drying gas (15 psi). The MS data were acquired and processed using Xcalibur 1.3.

NMR analysis- The structure of synthetic compounds were analyzed by NMR experiments, including one dimensional- (“zg” pulse sequence ¹H and “zgdc30” pulse sequence ¹³C), two dimensional- (¹H-¹H “cosygpmfphpp” pulse sequence COSY, ¹H-¹³C “hsqcgpph” pulse sequence HSQC (heteronuclear single quantum coherence), and “hmbcgpndqf” pulse sequence HMBC (heteronuclear multiple-bond correlation spectroscopy) NMR. NMR experiments were performed at 298 K on Bruker Avance 850 or 700 MHz spectrometer equipped with 5mm CryoProbe and processed by TopSpin 3.2 software. Samples (0.5 to 5.0 mg) were each dissolved in 0.5 mL D₂O (99.994%, Sigma, Co.) and lyophilized three times to evaporate the exchangeable protons. The samples were re-dissolved in 0.5 mL D₂O and transferred to NMR microtubes (O.D. 5 mm, Norrell). Compound **1** to **14** was dissolved in 0.5 mL D₂O containing 10 mM phosphate buffer at pH 7.4. Chemical shifts are referenced to external 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS, Sigma, Co.). Deuterated EDTA (10 mM) (Sigma, Co.) was added to resolve the coupling constants of signals. 1D ¹H-NMR experiments were performed with 64 scans and an acquisition time of 3.8 sec. 1D ¹³C-NMR experiments were performed with 30,000 scans, 1.5 sec relaxation delay, and an acquisition time of 1.0 sec. 2D (¹H-¹³C HSQC, ¹H-¹H COSY) spectra were recorded with GARP (Globally Optimized Alternating Phase Rectangular Pulse) carbon decoupling, a 1.5 sec relaxation delay, a 120 msec acquisition time with 512 increments for 48 scans. Sixteen dummy scans were used prior to the start of acquisition. 2,048 total points were collected in f2.

Gradients were used for coherence transfer selection in the HSQC experiment. ^{13}C transmitter offset was set at 90.0 ppm. The polarization transfer delay was set with a $^1J_{\text{C-H}}$ coupling value of 145 Hz. 2D ^1H - ^{13}C HMBC experiments were performed with 72 scans, 1.5 sec relaxation delay, and a 120 msec acquisition time. The delay for evolution of long-range couplings was set with J_{lr} of 7.6 Hz.

Reference

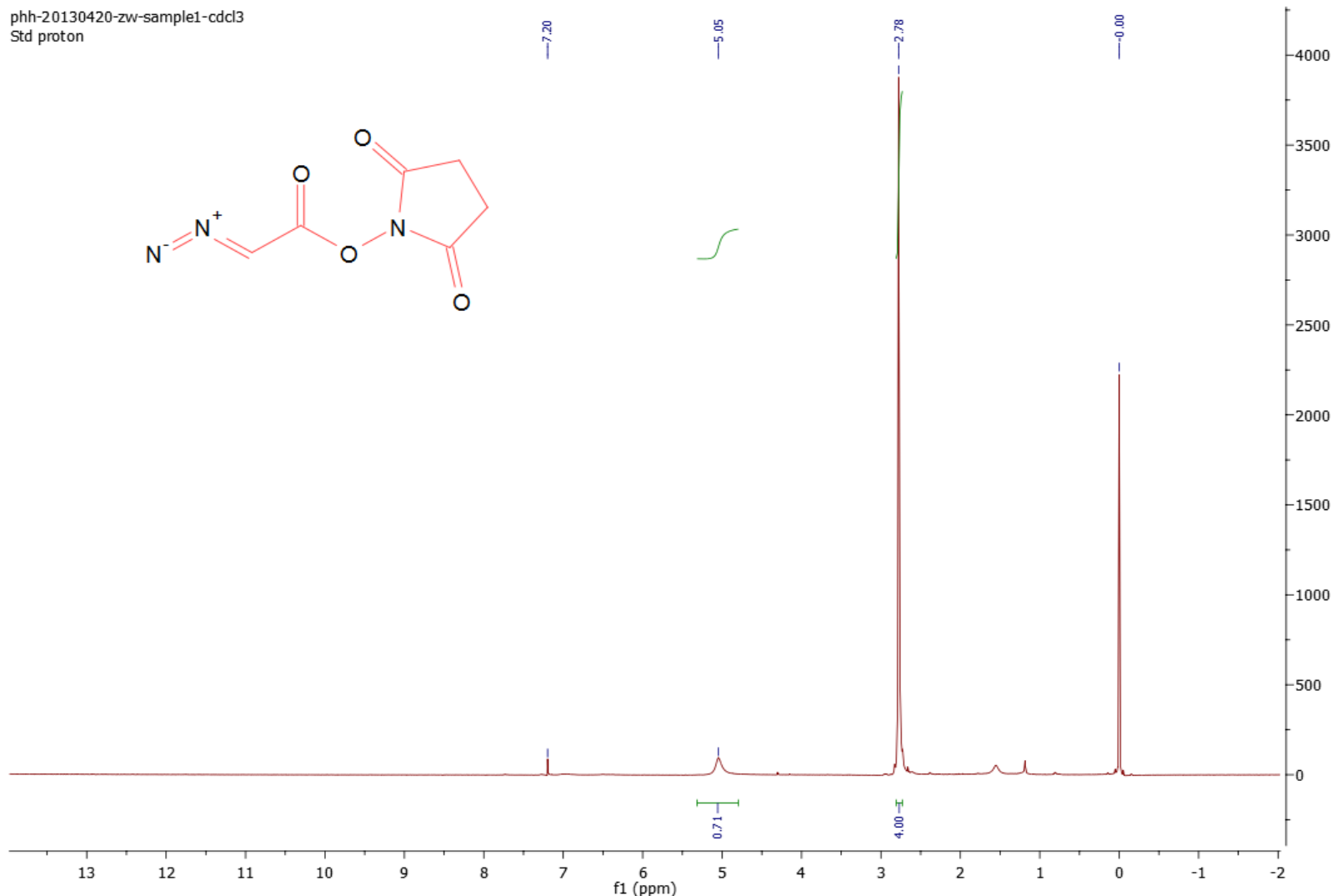
1. Shahrokh, M.; Paul, M., *Helvetica Chimica Acta* **1993**, 76, 2803-2813.
2. Liu, R.; Xu, Y.; Chen, M.; Weïwer, M.; Zhou, X.; Bridges, A. S.; DeAngelis, P. L.; Zhang, Q.; Linhardt, R. J.; Liu, J., *J Biol Chem* **2010**, 285, 34240-34249.
3. Xu, Y.; Wang, Z.; Liu, R.; Huang, X.; Liu, J., *Glycobiology* **2012**, 22, 96-106.

Supplementary Table S1. Optimization of reaction conditions to couple A GlcNDaz residue to hexasaccharide

Diazoacylating compound	Molar ratio with hexasaccharide substrate	1,4 dioxane/ water (v/v)	Incubation Time (hr)	Choice of Base	Coupling Yield ¹
4	5:1	1:1	4	LiOH	Not detectable
4	5:1	1:1	4	K ₂ CO ₃	Not detectable
4	5:1	1:1	4	N(C ₂ H ₅) ₃	Not detectable
5	5:1	1:1	4	LiOH	10%
5	5:1	1:1	4	K ₂ CO ₃	20%
5	5:1	1:1	4	N(C ₂ H ₅) ₃	90%
5	5:1	1:1	15	N(C ₂ H ₅) ₃	95%
5	5:1	2:1	15	N(C ₂ H ₅) ₃	96%
5	5:1	1:2	15	N(C ₂ H ₅) ₃	91%
5	5:1	3:1	15	K ₂ CO ₃	63%
5	10:1	2:1	15	N(C ₂ H ₅) ₃	99%

1. The coupling yield was determined by DEAE-HPLC method, which can resolve hexasaccharide substrate and the product (compound 1). The yield was confirmed by three independent experiments.

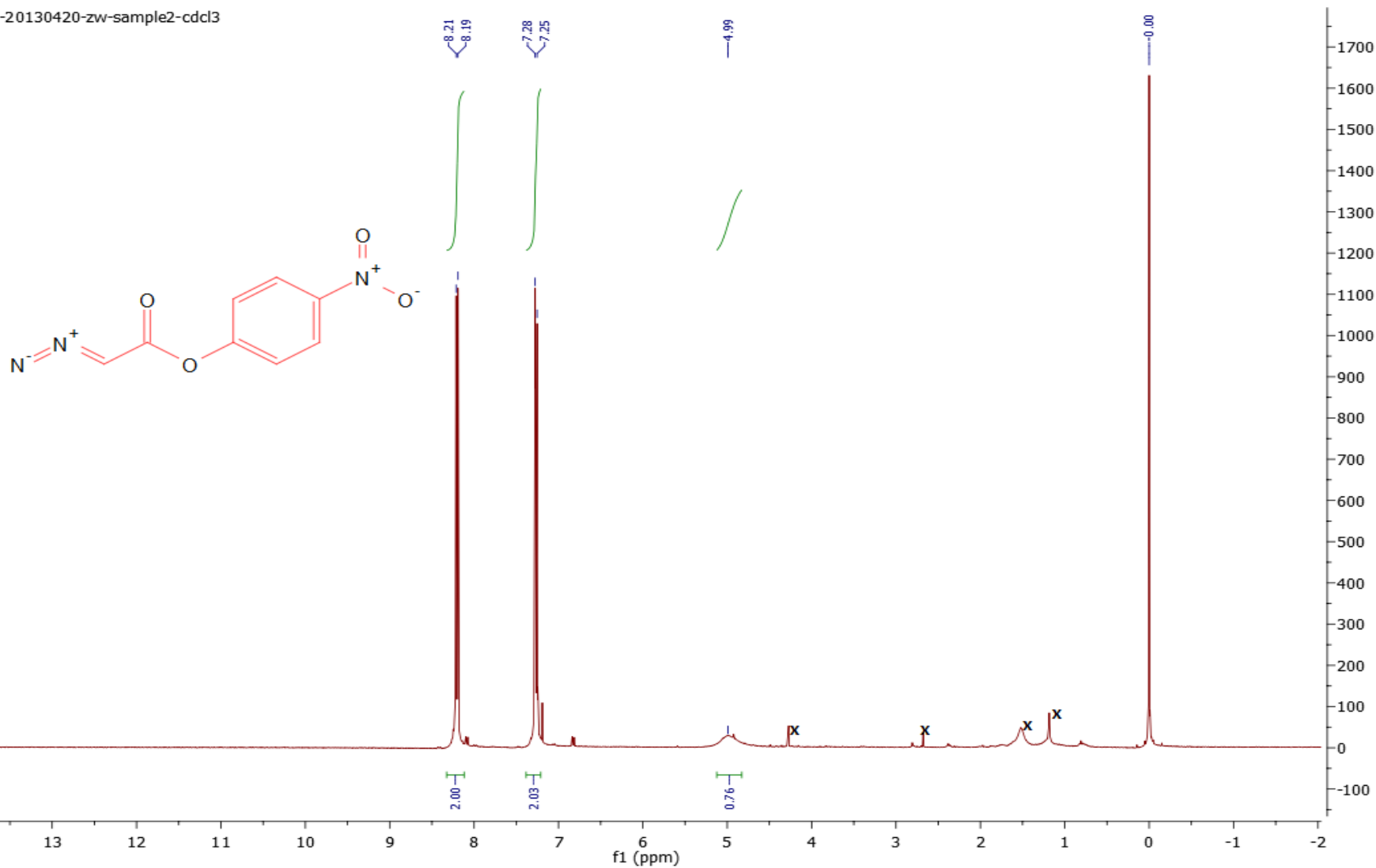
^1H NMR of succinimidyl diazoacetate (**5**)



Supplementary Fig S1. ^1H -NMR of compound **5** (400 MHz, CDCl_3). Proton signal at $\delta=5.05$ was the characteristic proton of $=\text{CHN}_2$

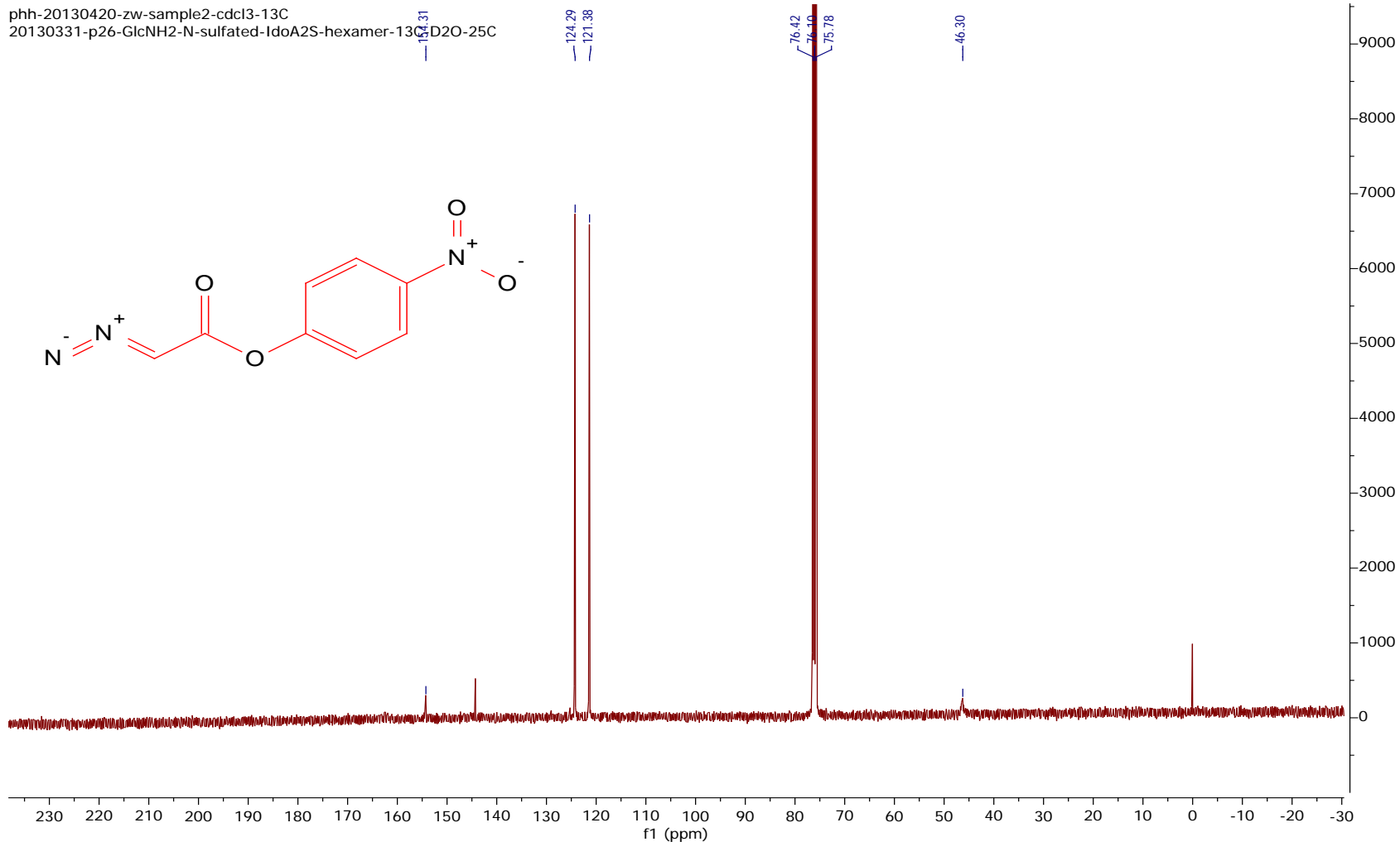
^1H NMR of *p*-nitrophenyl diazoacetate (**4**)

phh-20130420-zw-sample2-cdcl3



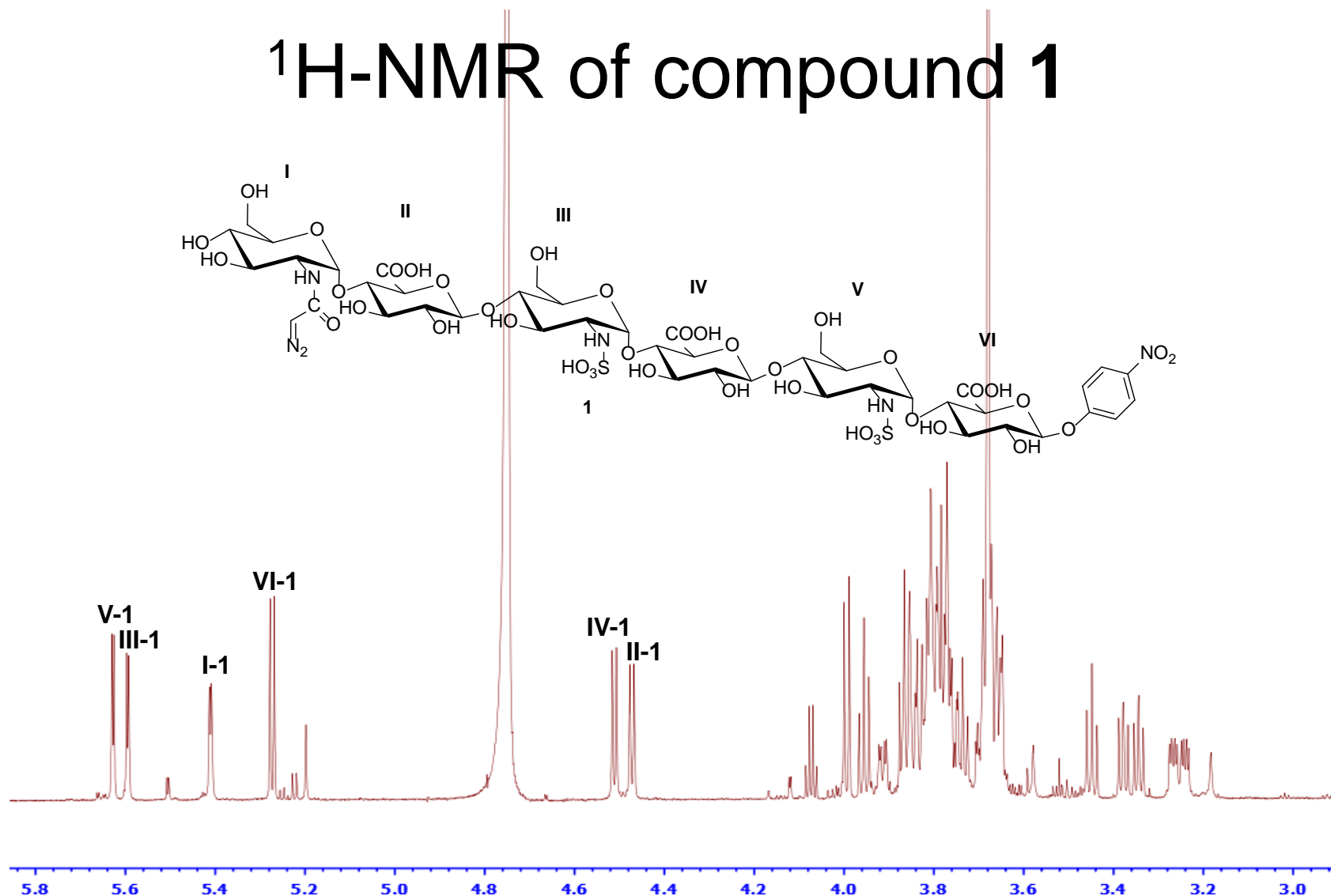
Supplementary Fig S2. ^1H -NMR of compound **4** (400 MHz, CDCl_3). Proton from $=\text{CHN}_2$ at $\delta=4.99$ was detected. Peaks labelled with "x" are from impurities. Peaks at $\delta=7.28$, 8.21 are from the aromatic protons, suggesting the successful introduction of *P*-nitrophenyl group.

^{13}C NMR of *p*-nitrophenyl diazoacetate (**4**)



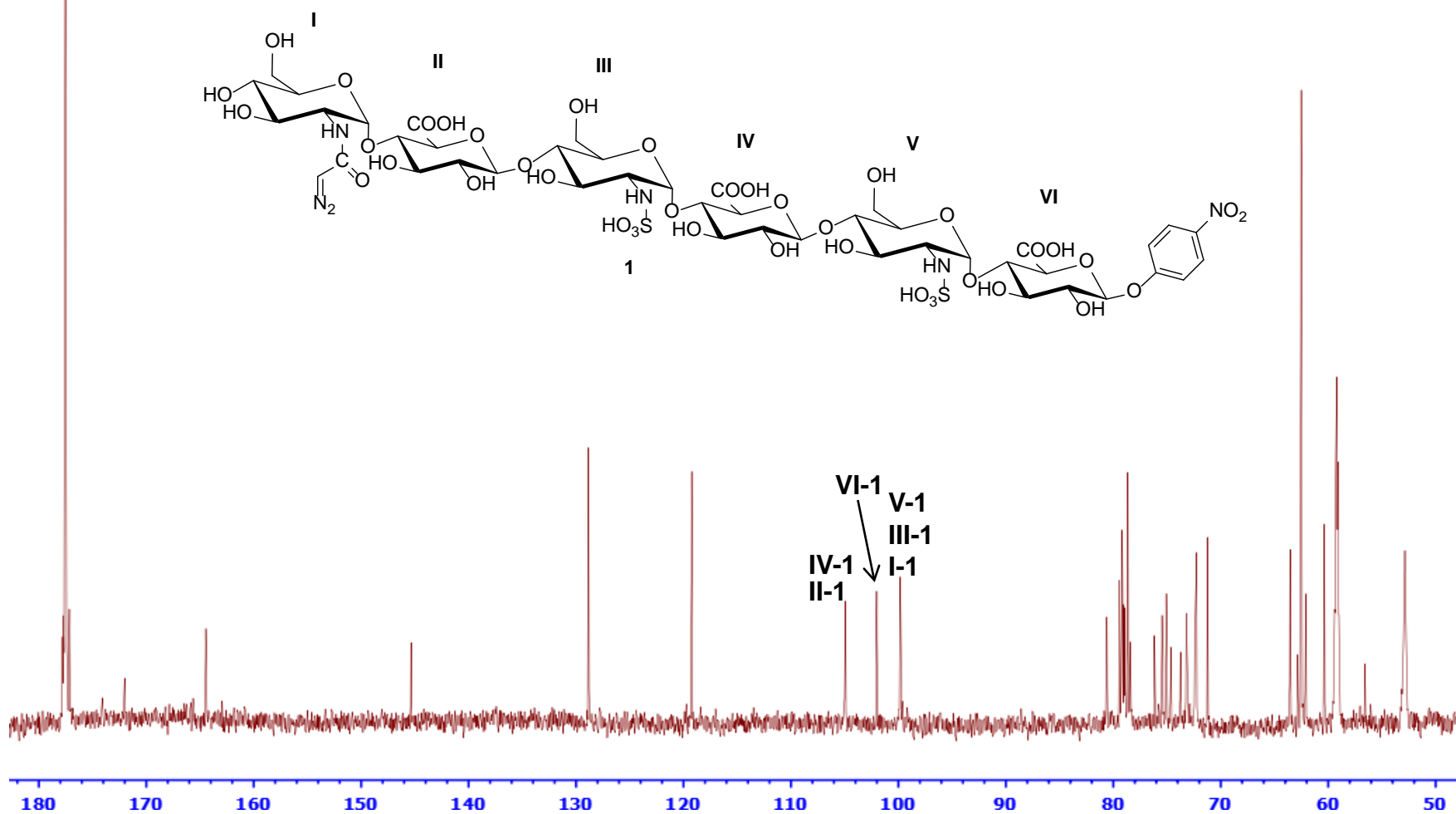
Supplementary Fig S3. ^{13}C -NMR of compound **4** (100 MHz, CDCl_3). A characteristic signal at $\delta=46.30$ was from the carbon of $=\text{CHN}_2$

$^1\text{H-NMR}$ of compound **1**



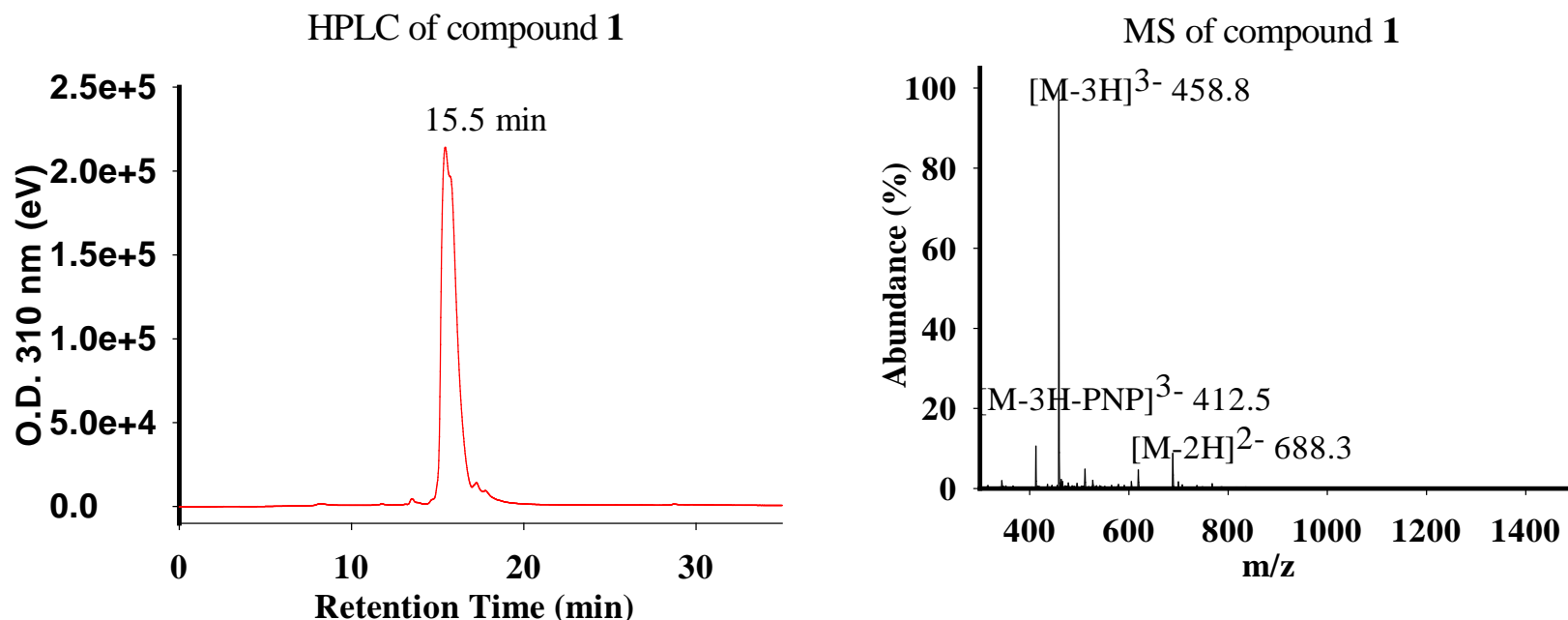
Supplementary Fig S4. $^1\text{H-NMR}$ of compound **1** (850 MHz, D₂O). The signals of anomeric protons are indicated. The anomeric protons resonate as doublet at δ 5.63, 5.60, 5.41, 5.27, 4.51, and 4.47 ppm. Chemical structure of compound **1** is shown on top of figure.

^{13}C -NMR of compound 1



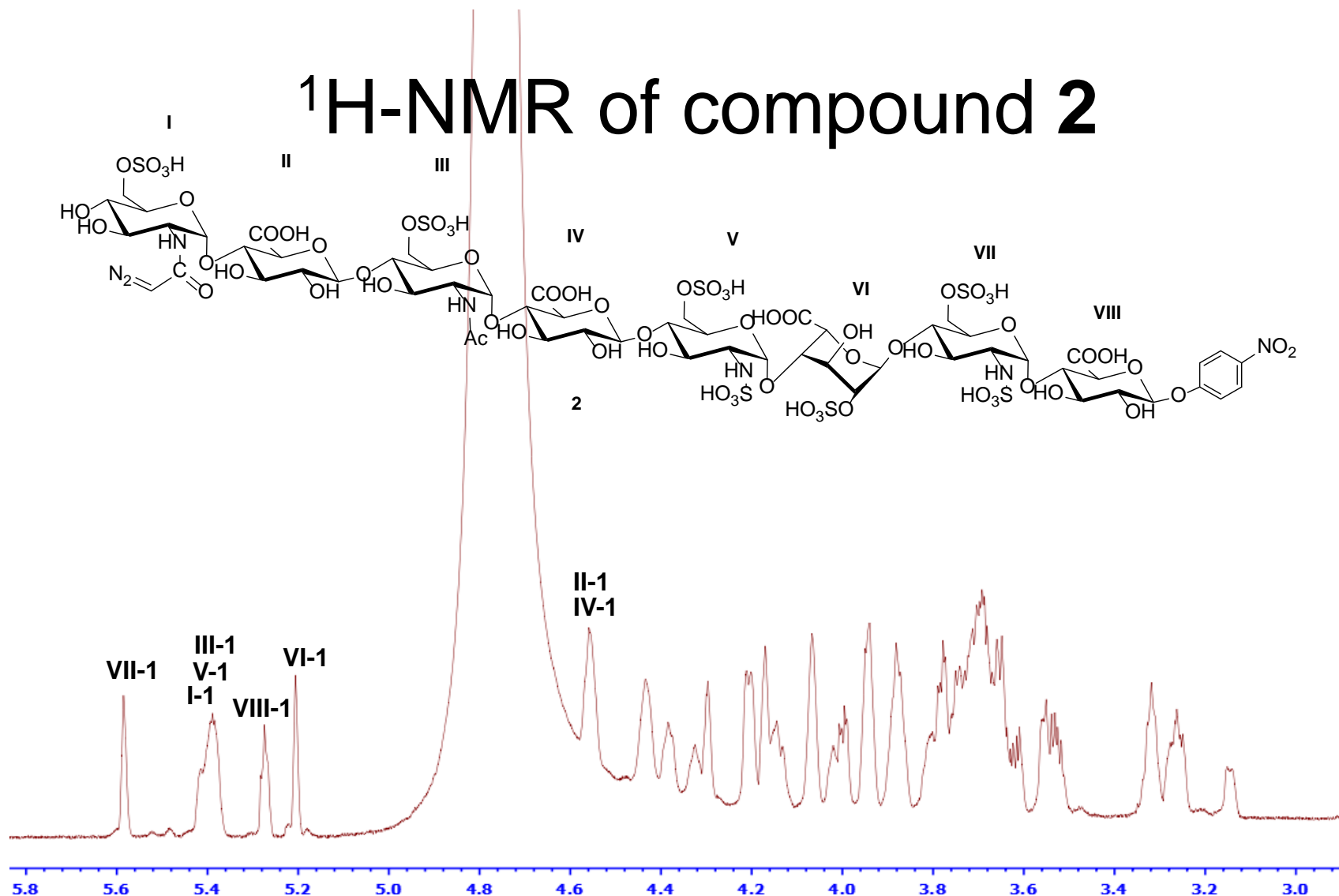
Supplementary Fig S5. ^{13}C -NMR of compound 1 (212.5 MHz, D₂O). The signals of anomeric carbons are indicated. The anomeric carbons resonate at δ 105.0, 104.9, 102.0, 99.8, 99.8, and 99.8 ppm. Chemical structure of compound 1 is shown on top of figure.

Purity and MS analysis of compound 1



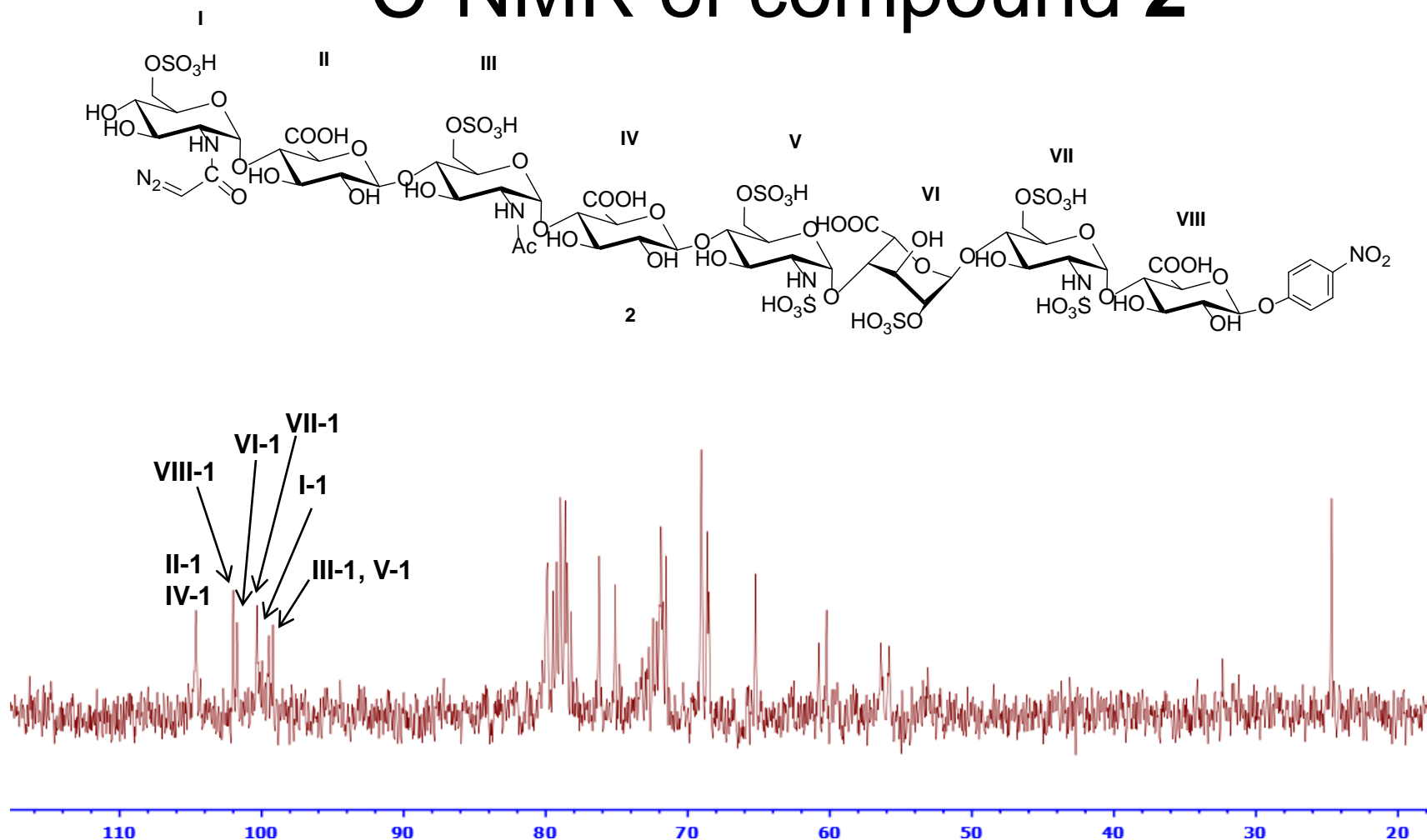
Supplementary Fig S6. Left panel shows the DEAE-HPLC chromatogram of compound 1. The compound was eluted as a single symmetric peak on DEAE-HPLC, suggesting that compound 1 was pure. Right panel shows the MS spectrum. The measured molecular mass was 1379.4, which is very close to the calculated molecular weight (1379.1). $[M-3H-PNP]^{3-}$ refers to the molecular ion that lost PNP (*p*-nitrophenyl) moiety.

¹H-NMR of compound 2



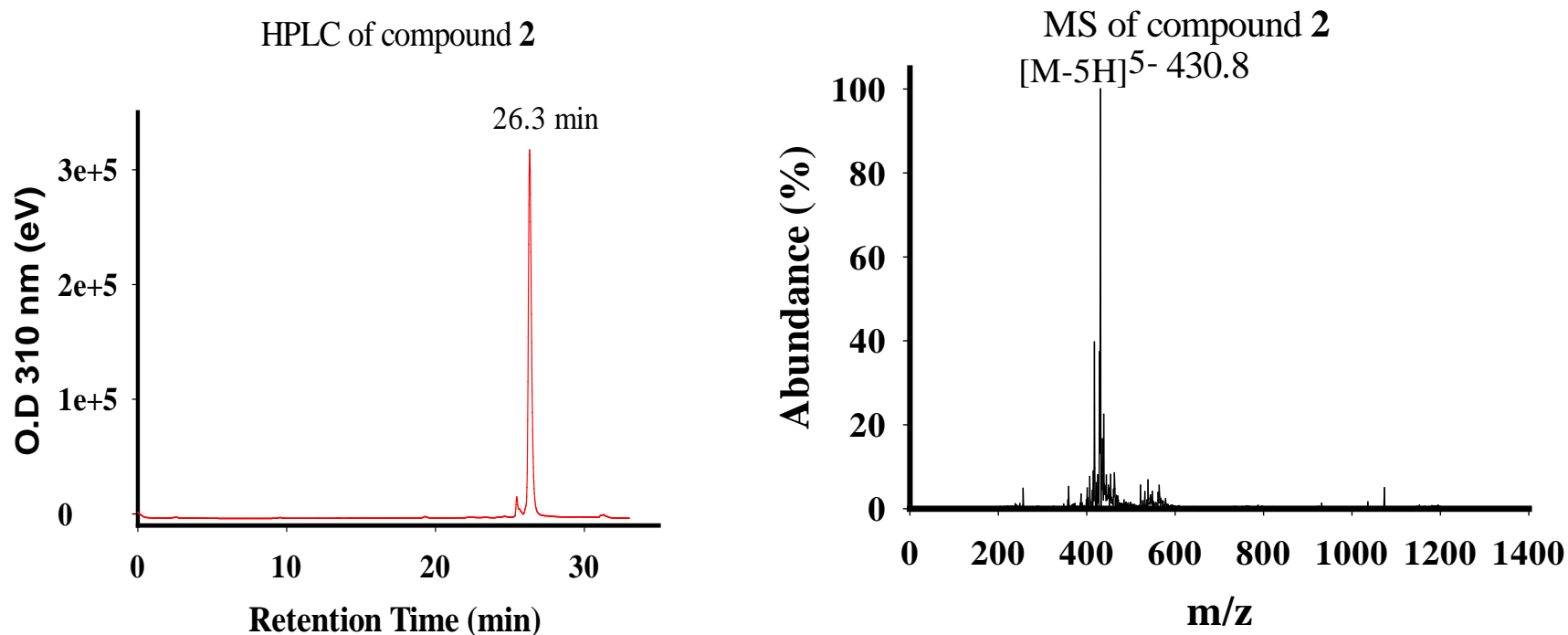
Supplementary Fig S7. ¹H-NMR of compound 2 (850 MHz, D₂O). The signals of anomeric protons are indicated. The anomeric protons resonate at δ 5.59, 5.41, 5.40, 5.39, 5.27, 5.21, 4.56, and 4.56 ppm. Chemical structure of compound 2 is shown on top of figure. The spectrum was acquired with 256 scans.

^{13}C -NMR of compound **2**



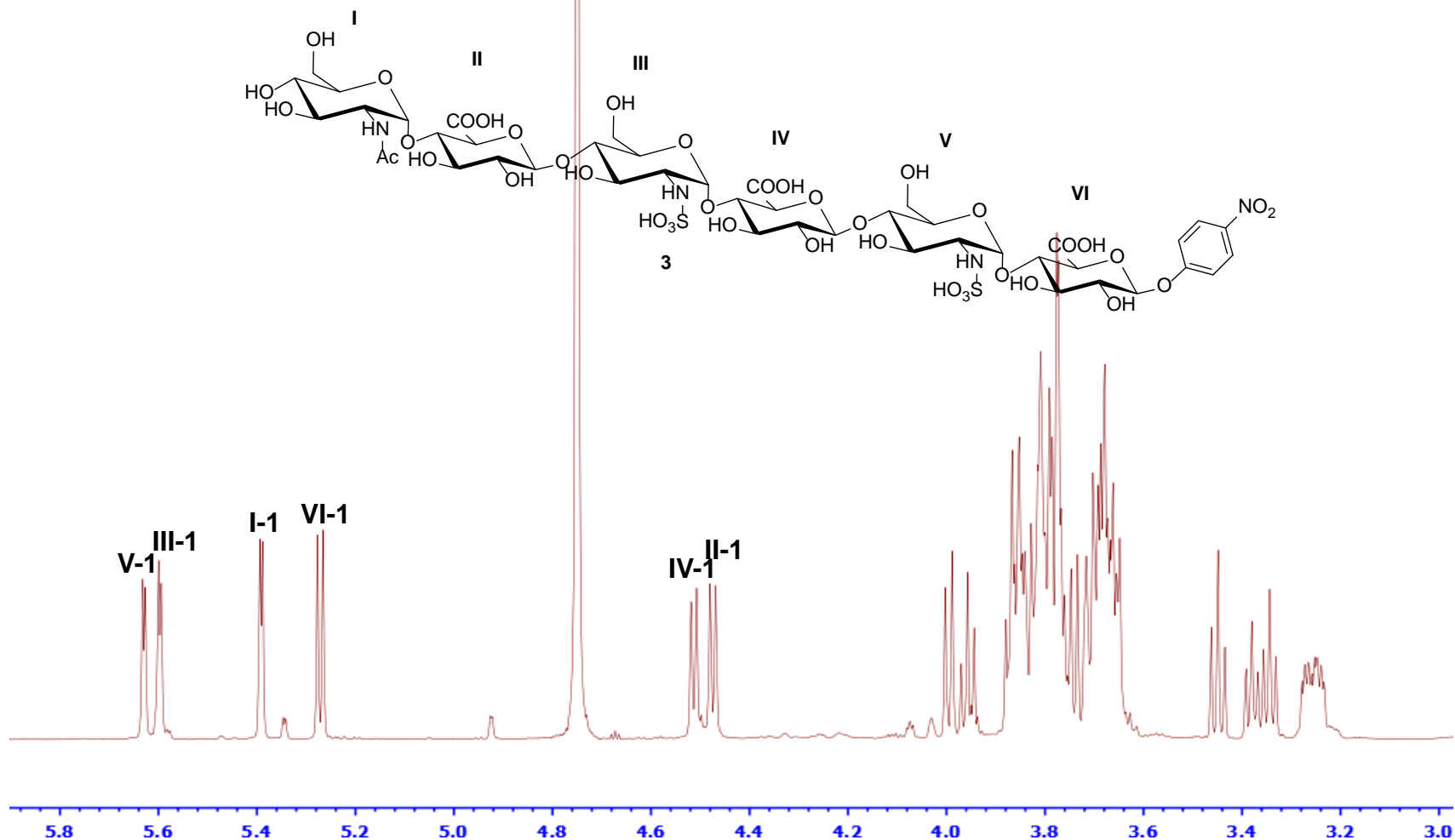
Supplementary Fig S8. ^{13}C -NMR of compound **2** (212.5 MHz, D_2O). The signals of anomeric carbons are indicated. The anomeric carbons resonate at δ 104.6, 104.6, 102.0, 101.7, 100.3, 100.0, 99.5, and 99.2 ppm. The structure of compound **2** is shown on top of figure. The spectrum was acquired using distortionless enhancement by polarization transfer with pulse width for ^1H at 45° Page 18 of 36

Purity and structural analysis of compound 2



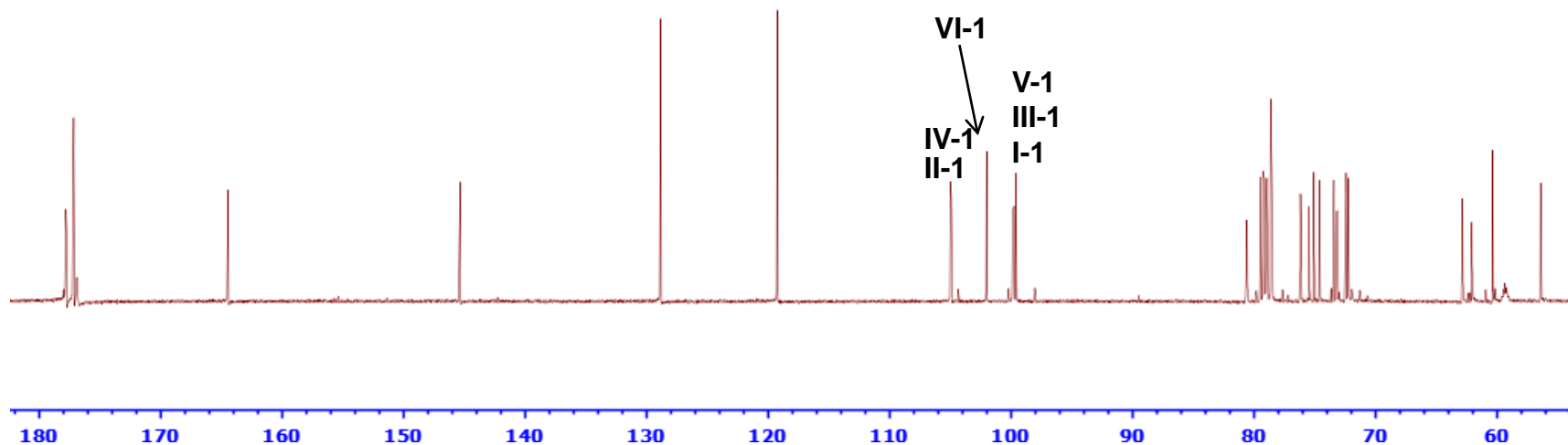
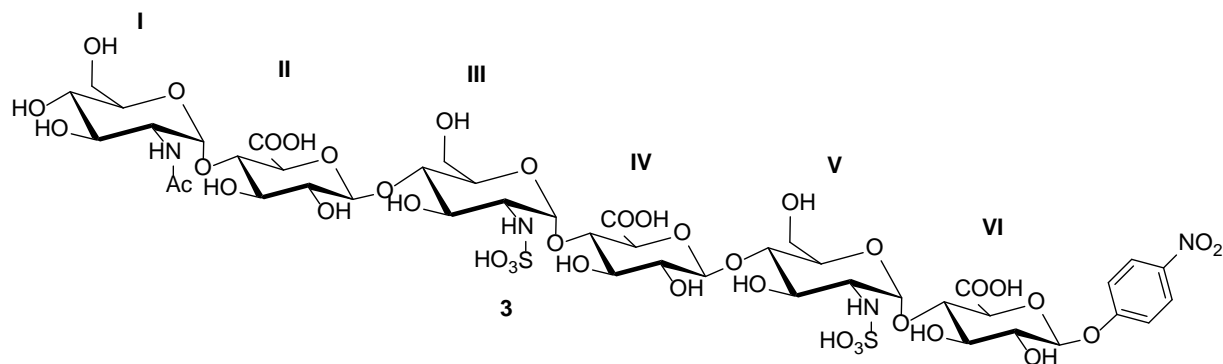
Supplementary Fig S9. Left panel shows the DEAE-HPLC chromatogram of compound 2. The compound was eluted as a major single symmetric peak on DEAE-HPLC, suggesting that compound 2 was pure. Right panel shows the MS spectrum. The measured molecular mass was 2159.0, which is very close to the calculated molecular weight (2158.7).

$^1\text{H-NMR}$ of compound **3**



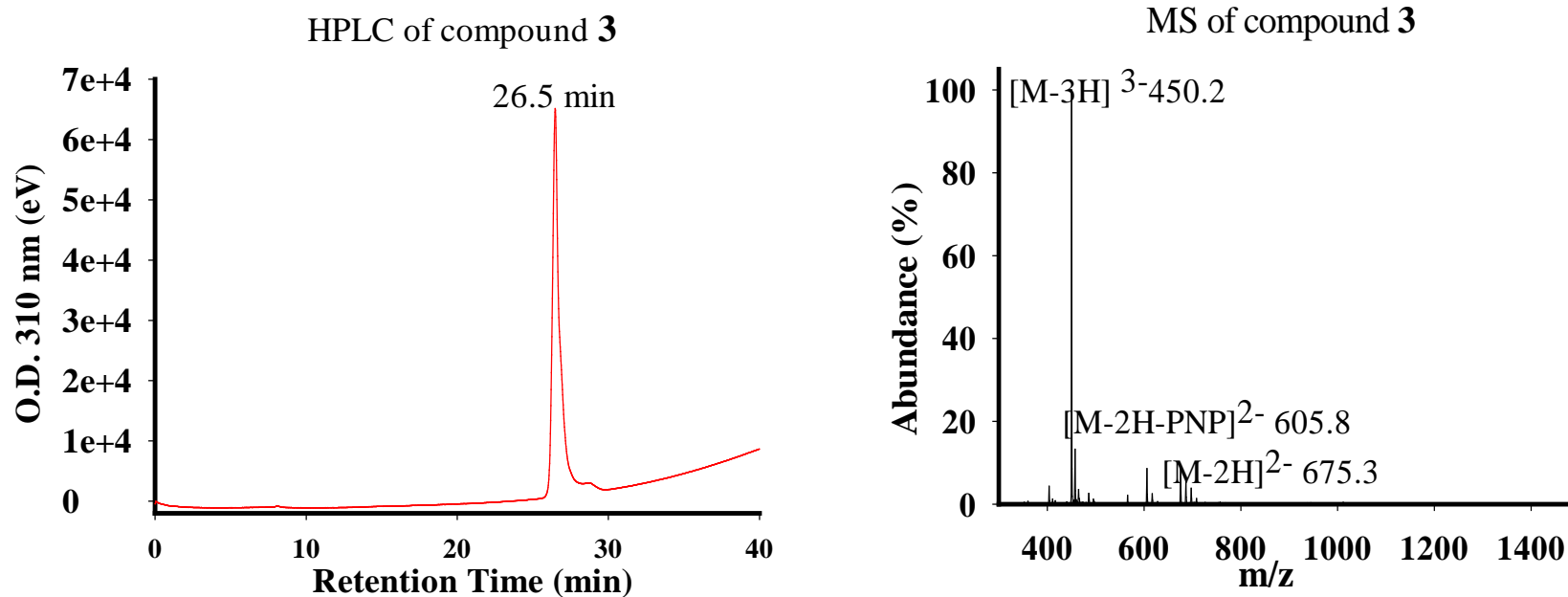
Supplementary Fig S10. $^1\text{H-NMR}$ of compound **3** (700 MHz, D_2O). The signals of anomeric protons are indicated. The anomeric protons resonate as doublet at δ 5.63, 5.60, 5.39, 5.27, 4.51, and 4.47 ppm. Chemical structure of compound **3** is shown on top of figure.

^{13}C -NMR of compound **3**



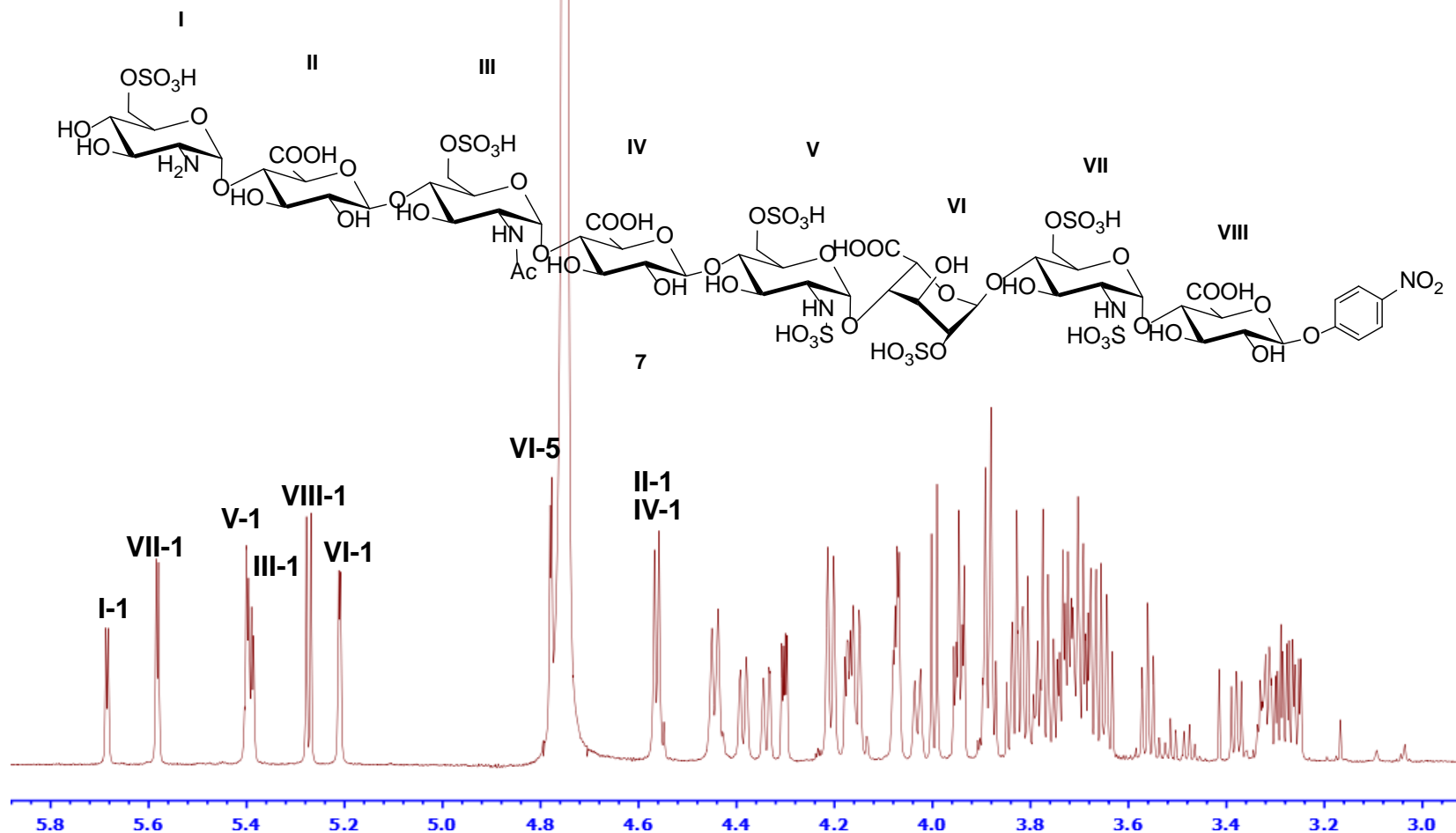
Supplementary Fig S11. ^{13}C -NMR of compound **3** (175 MHz, D₂O). The signals of anomeric carbons are indicated. The anomeric carbons resonate at δ 105.0, 104.9, 102.0, 99.9, 99.8, and 99.6 ppm. Chemical structure of compound **3** is shown on top of figure.

Purity and MS analysis of compound 3



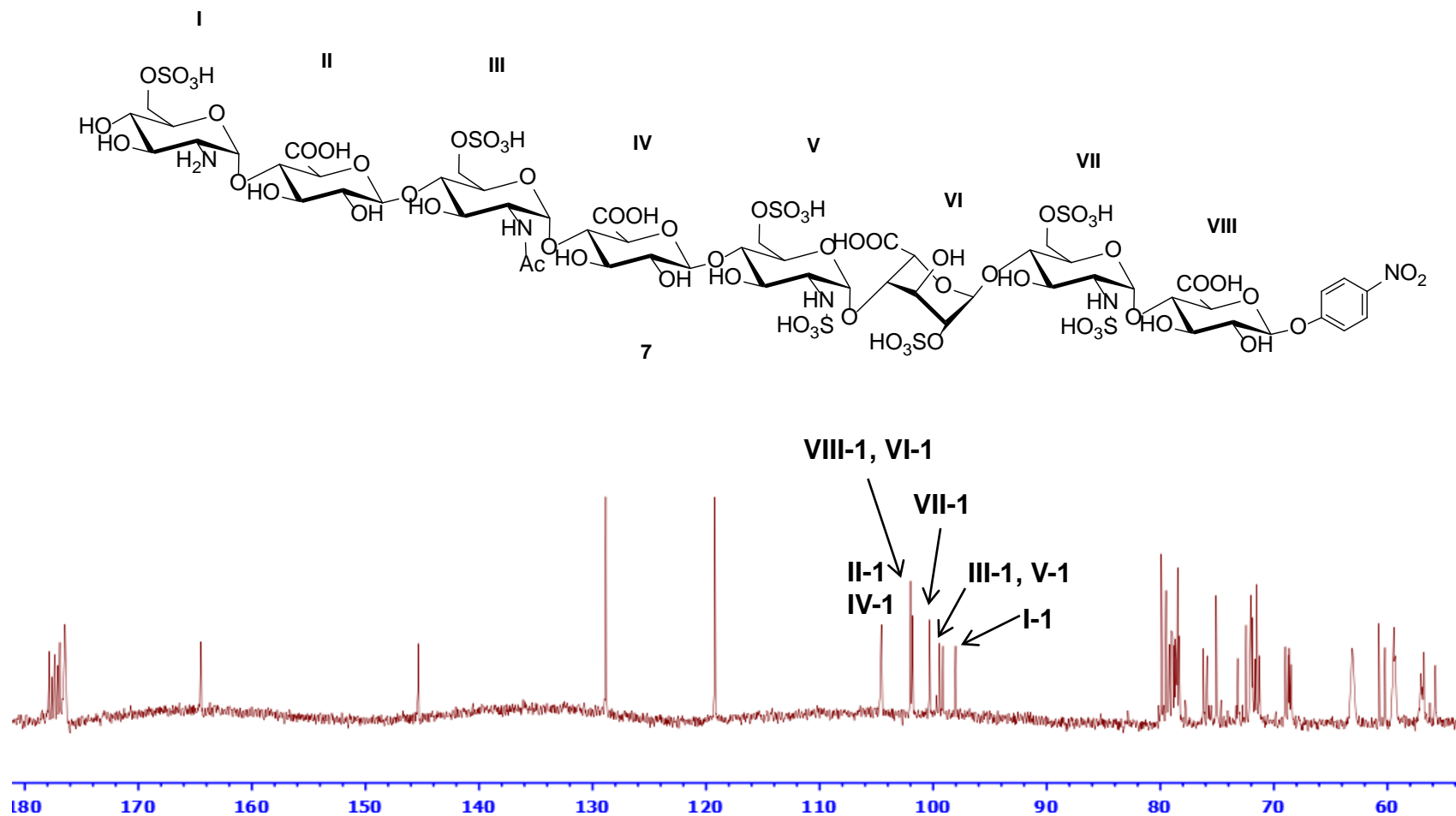
Supplementary Fig S12. Left panel shows the PAMN-HPLC chromatogram of compound **3**. The compound was eluted as a single symmetric peak on PAMN-HPLC, suggesting that compound **3** was pure. Right panel shows the MS spectrum. The measured molecular mass was 1353.6, which is very close to the calculated molecular weight (1353.1). [M-2H-PNP]²⁻ refers to the molecular ion that lost PNP (*p*-nitrophenyl) moiety.

$^1\text{H-NMR}$ of compound **7**



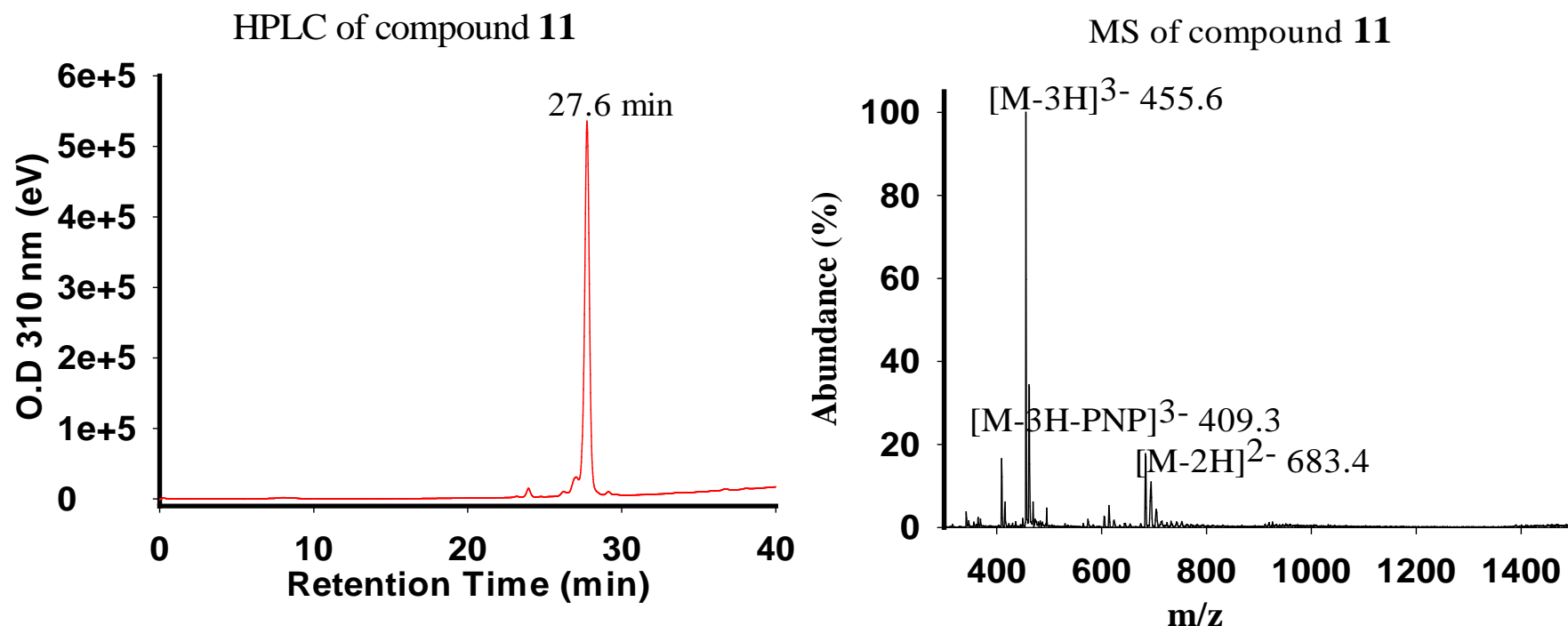
Supplementary Fig S13. $^1\text{H-NMR}$ of compound **7** (850 MHz, D_2O). The signals of anomeric protons are indicated. The anomeric protons resonate as doublet at δ 5.69, 5.58, 5.40, 5.39, 5.27, 5.21, 4.56, and 4.56 ppm. Chemical structure of compound **7** is shown on top of figure.

^{13}C -NMR of compound 7

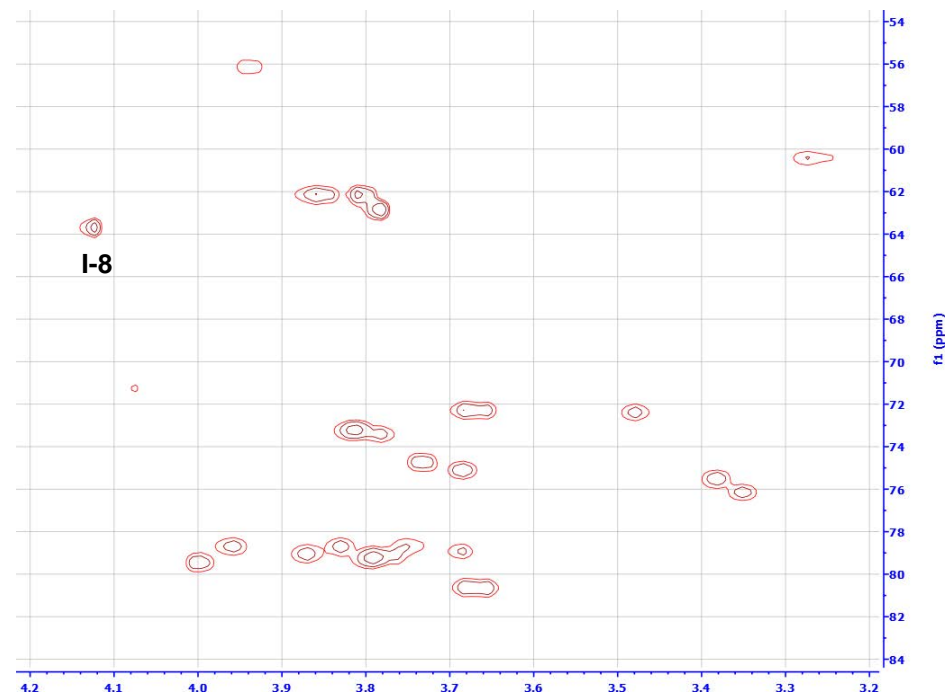
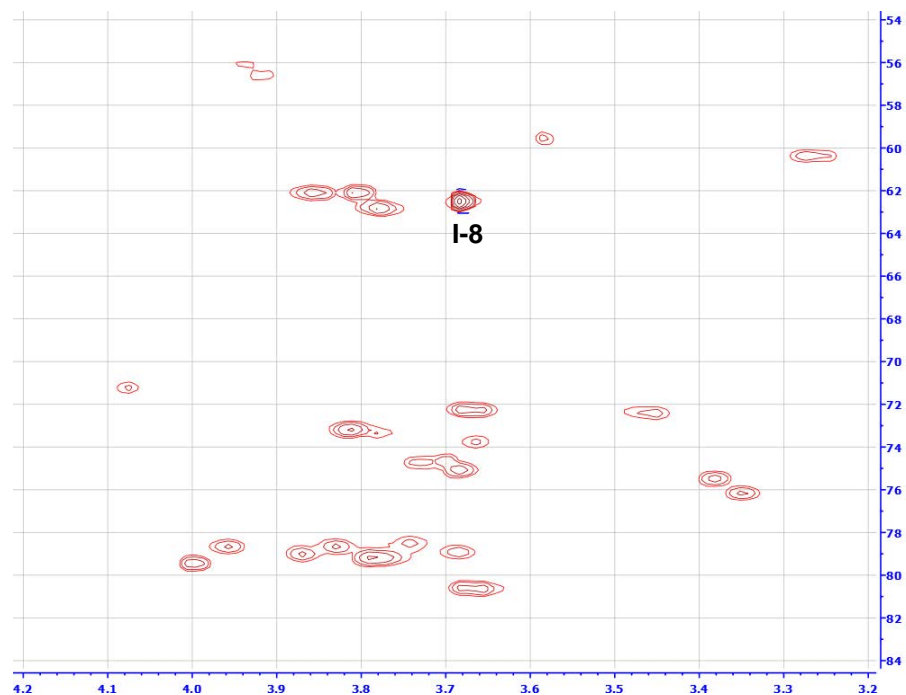
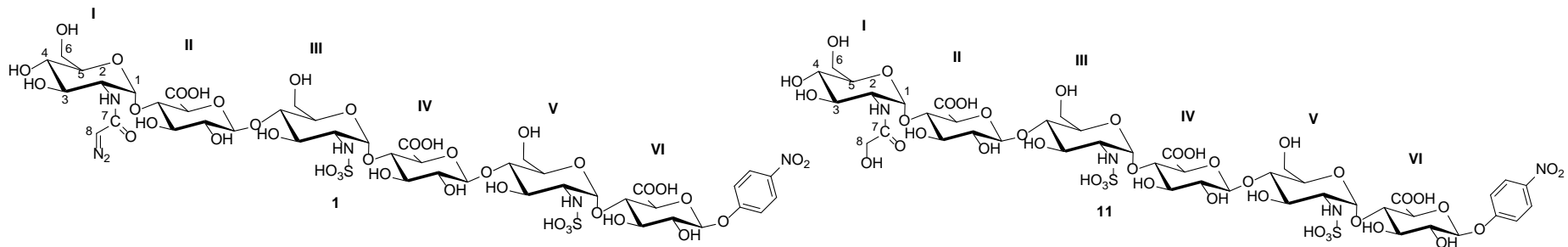


Supplementary Fig S14. ^{13}C -NMR of compound 7 (212.5 MHz, D₂O). The signals of anomeric carbons are indicated. The anomeric carbons resonate at δ 104.6, 104.6, 102.0, 101.8, 100.3, 99.5, 99.2, and 98.0 ppm. Chemical structure of compound 7 is shown on top of figure.

Purity and structural analysis of compound **11**

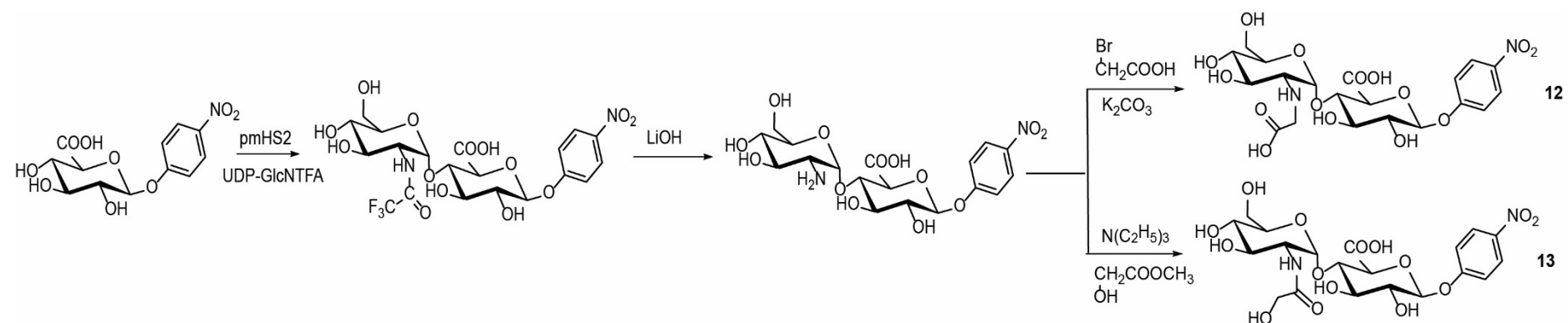


Supplementary Fig S15. Left panel shows the PMAN-HPLC chromatogram of compound **11**. The compound was eluted as a single symmetric peak on PMAN-HPLC, suggesting that **11** was pure. Right panel shows the MS spectrum. The measured molecular mass was 1369.3, which is very close to the calculated molecular weight (1369.1). [M-3H-PNP]³⁻ refers to the molecular ion of compound **11** that lost PNP moiety.



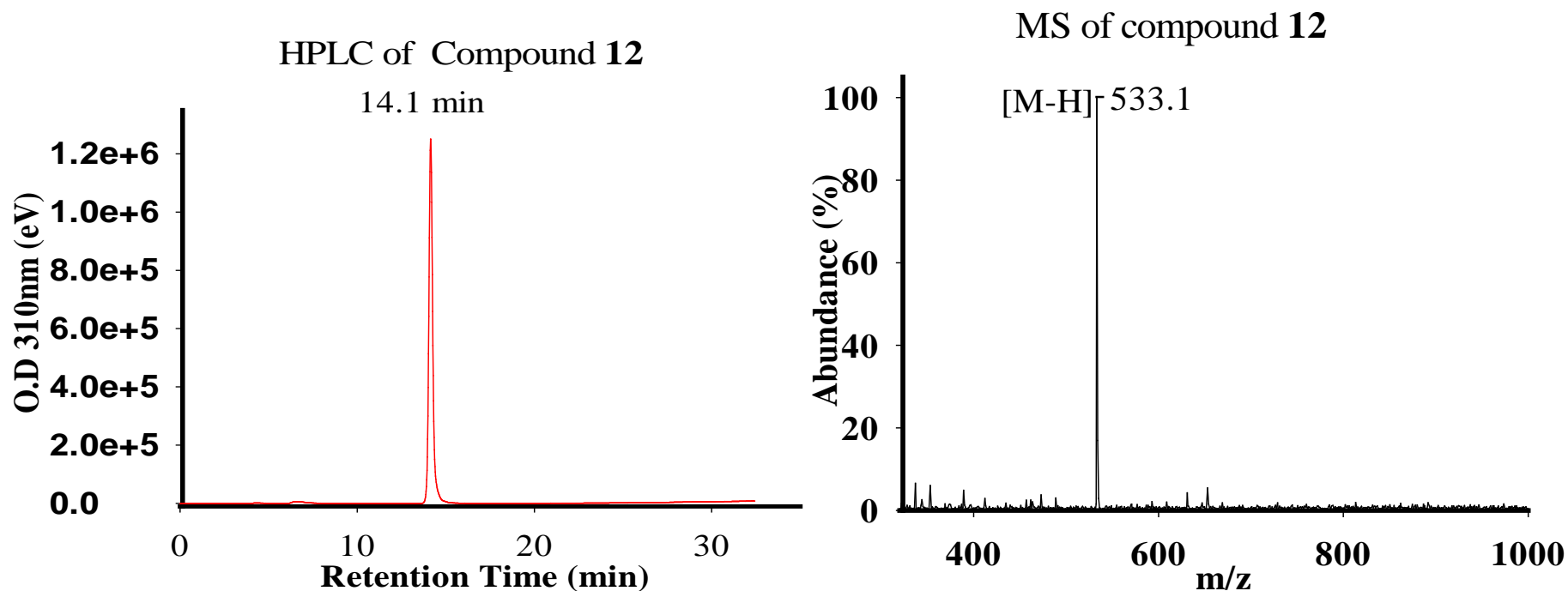
Supplementary Fig S16. ^1H - ^{13}C HSQC NMR spectra of **1** and **11**. Compound **11** is the product after **1** being exposed to water under an acidic condition. Spectra were recorded in D_2O containing 10 mM phosphate buffer at pD7.4 and 25°C . Under this condition, signals of compound **1** were detectable before the diazoacetyl group underwent decomposition reactions. The distinct signals are indicated in spectra.

Synthetic scheme for compound **12** and **13**



Supplementary Fig S17. Synthetic scheme for compound **12** and **13**

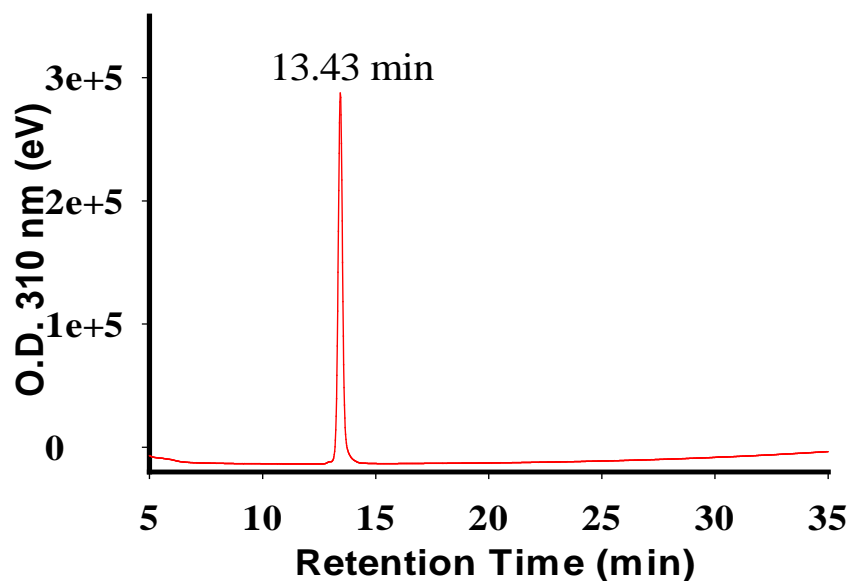
Purity and MS analysis of compound 12



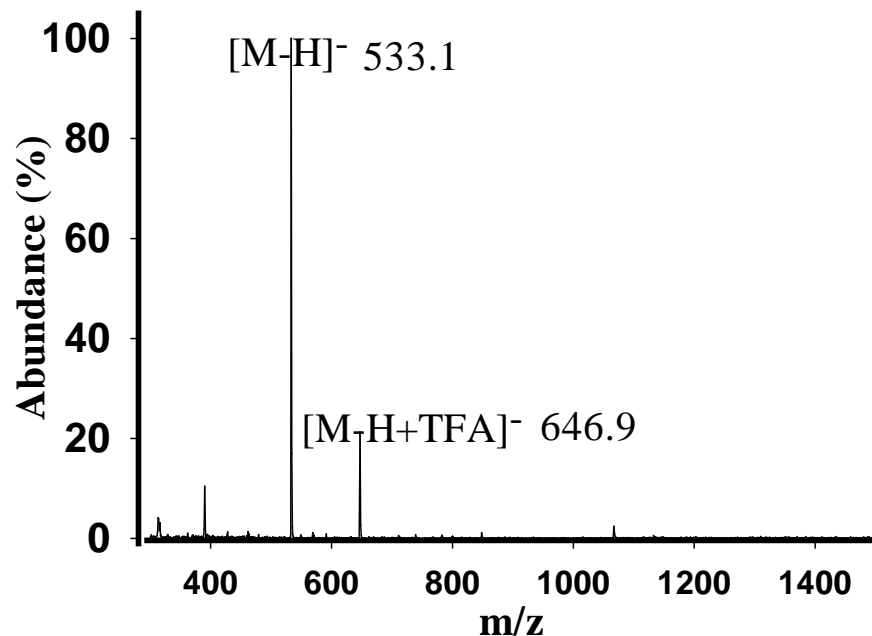
Supplementary Fig S18. Left panel shows the PMAN-HPLC chromatogram of compound 12. The compound was eluted as a single symmetric peak on PMAN-HPLC, suggesting that compound 12 was pure. Right panel shows the MS spectrum. The measured molecular mass was 534.1, which is very close to the calculated molecular weight (534.4).

Purity and MS analysis of **13**

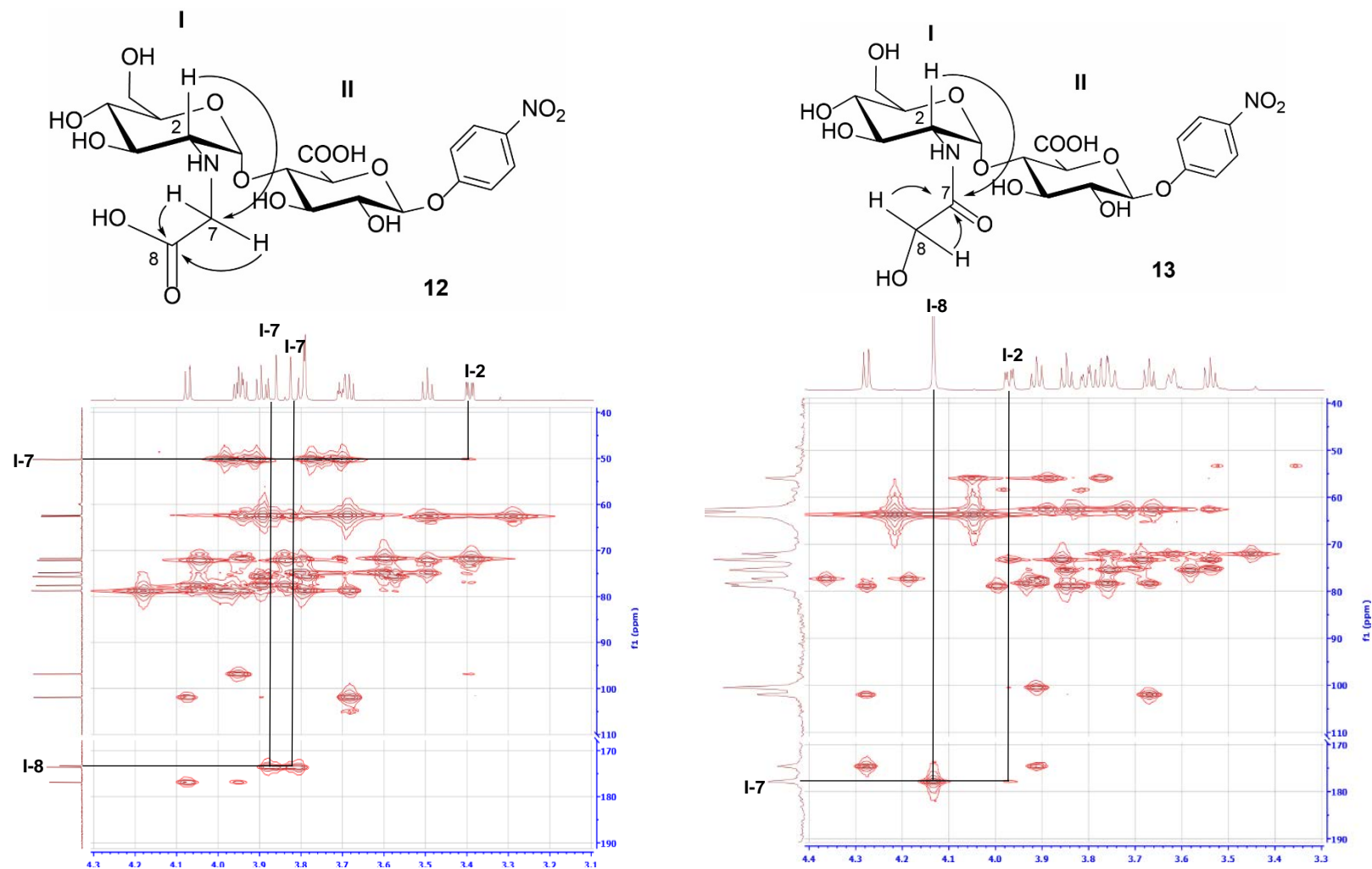
HPLC of compound **13**



MS of compound **13**

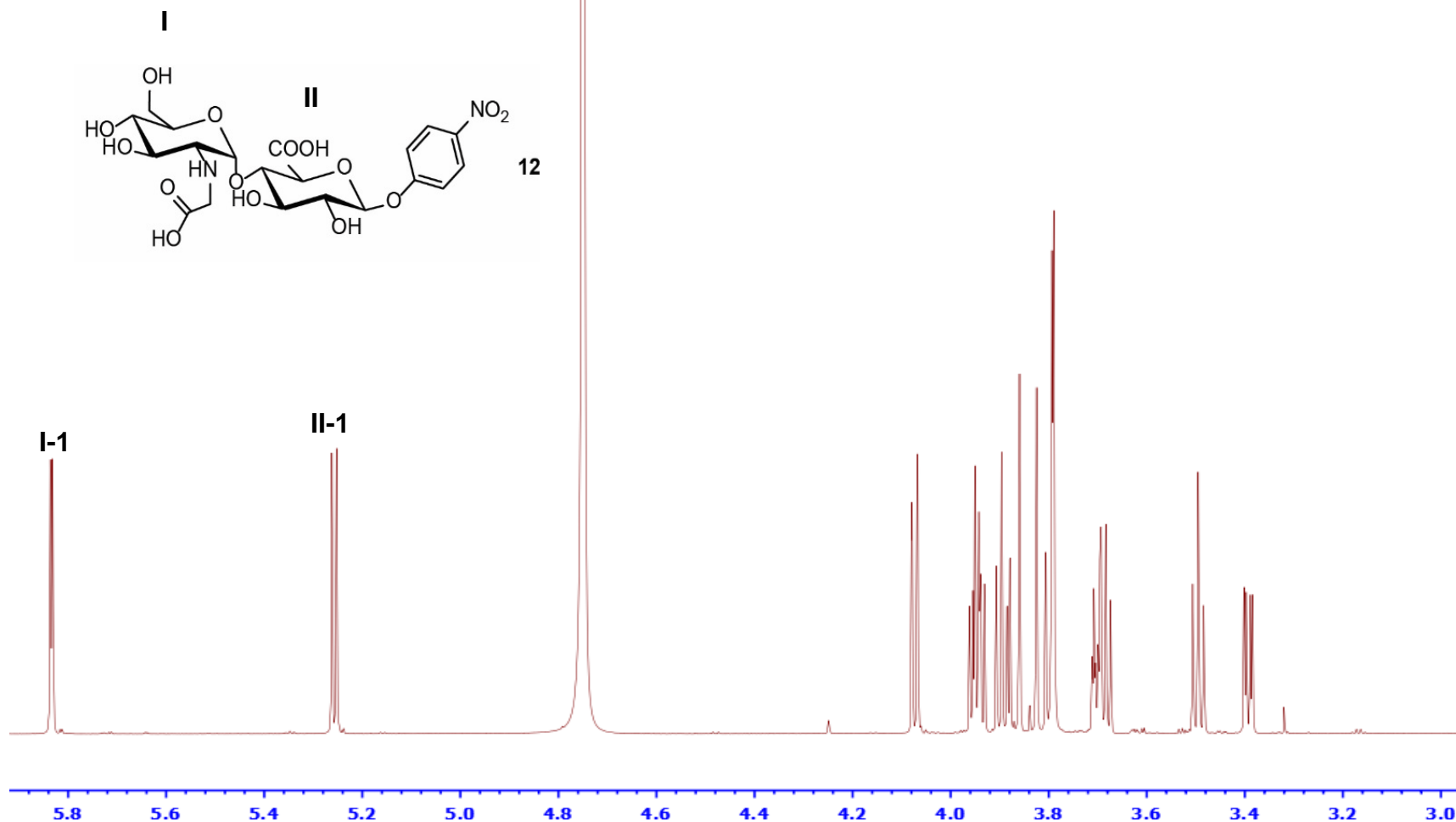


Supplementary Fig S19. Left panel shows the PMAN-HPLC chromatogram of compound **13**. The compound was eluted as a single symmetric peak on PMAN-HPLC, suggesting that compound **13** was pure. Right panel shows the MS spectrum. The measured molecular mass was 534.1, which is very close to the calculated molecular weight (534.4). [M-H+TFA]⁻ refers to the molecular ion of compound **13** and trifluoroacetic acid adduct.



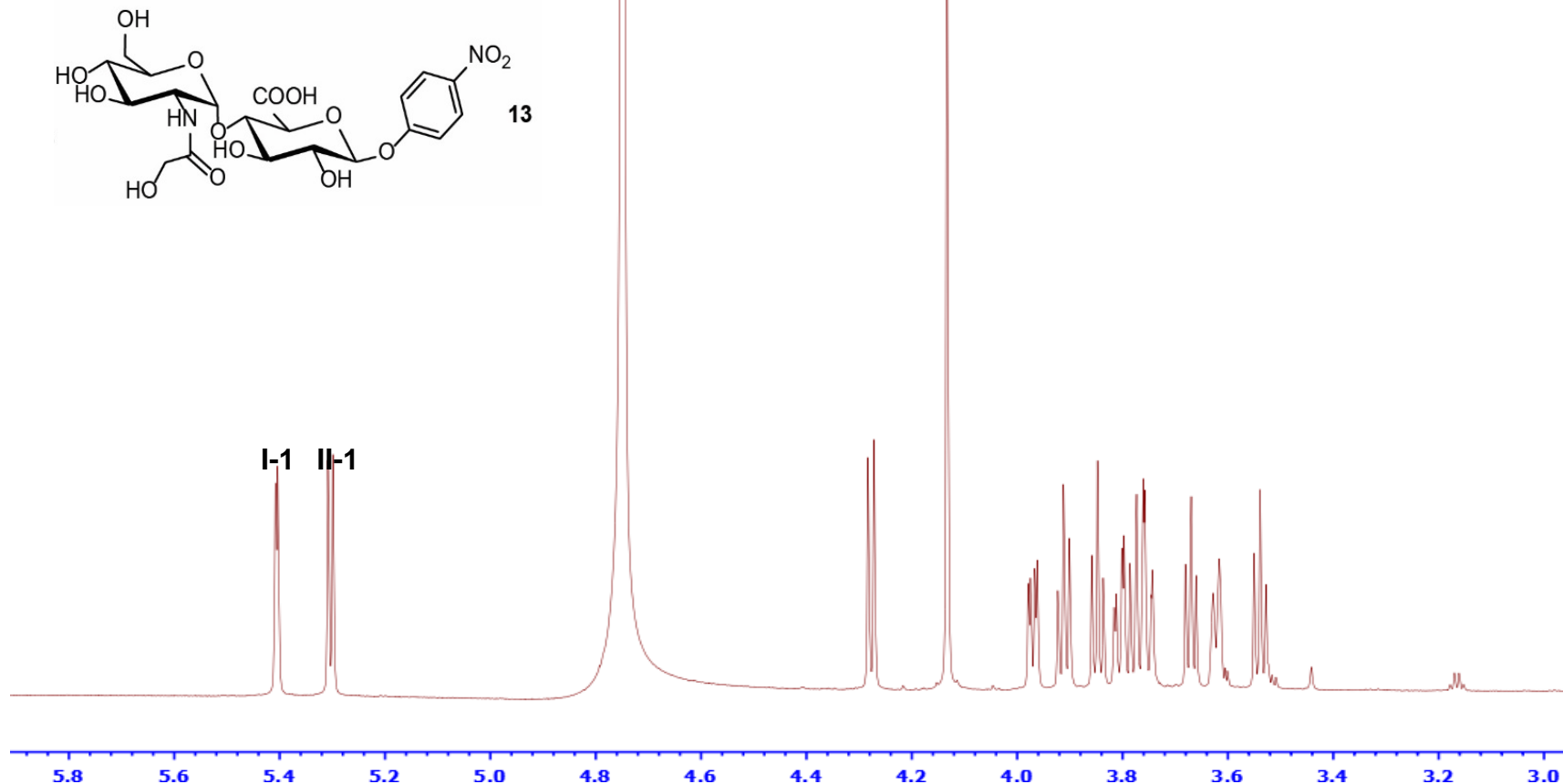
Supplementary Fig S20. ^1H - ^{13}C HMBC key correlations of **12** and **13**. The signals of featured groups are indicated. The cross peaks are connected with arrowed lines identifying correlated carbon and proton. The corresponding chemical structure is shown on top of each spectrum. The anticipated correlations are indicated in the structures by curved arrows.

$^1\text{H-NMR}$ of compound **12**



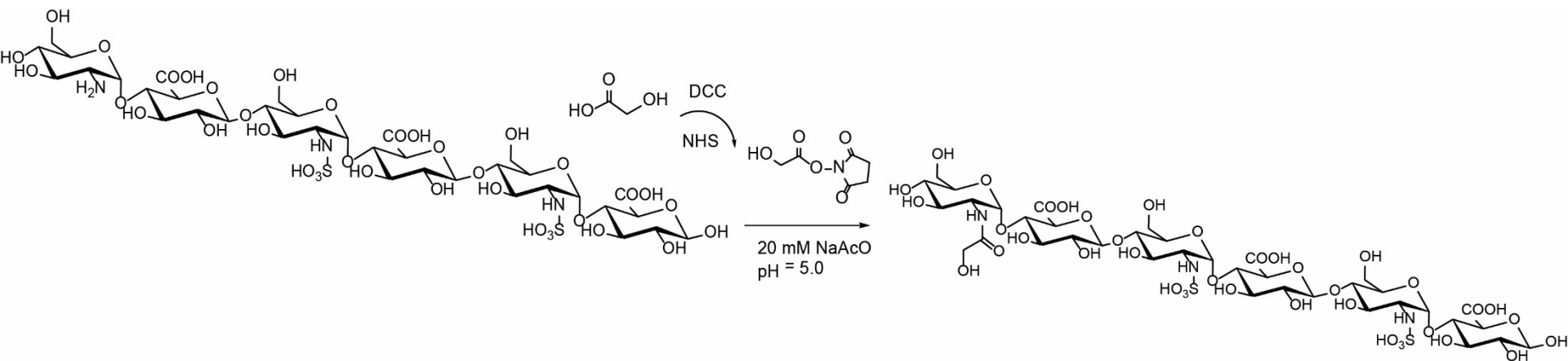
Supplementary Fig S21. $^1\text{H-NMR}$ of compound **12** (850 MHz, D_2O). The signals of anomeric protons are indicated. The anomeric protons resonate as doublet at δ 5.83 and 5.26 ppm. The chemical structure of compound **12** is shown on top of figure.

$^1\text{H-NMR}$ of compound **13**



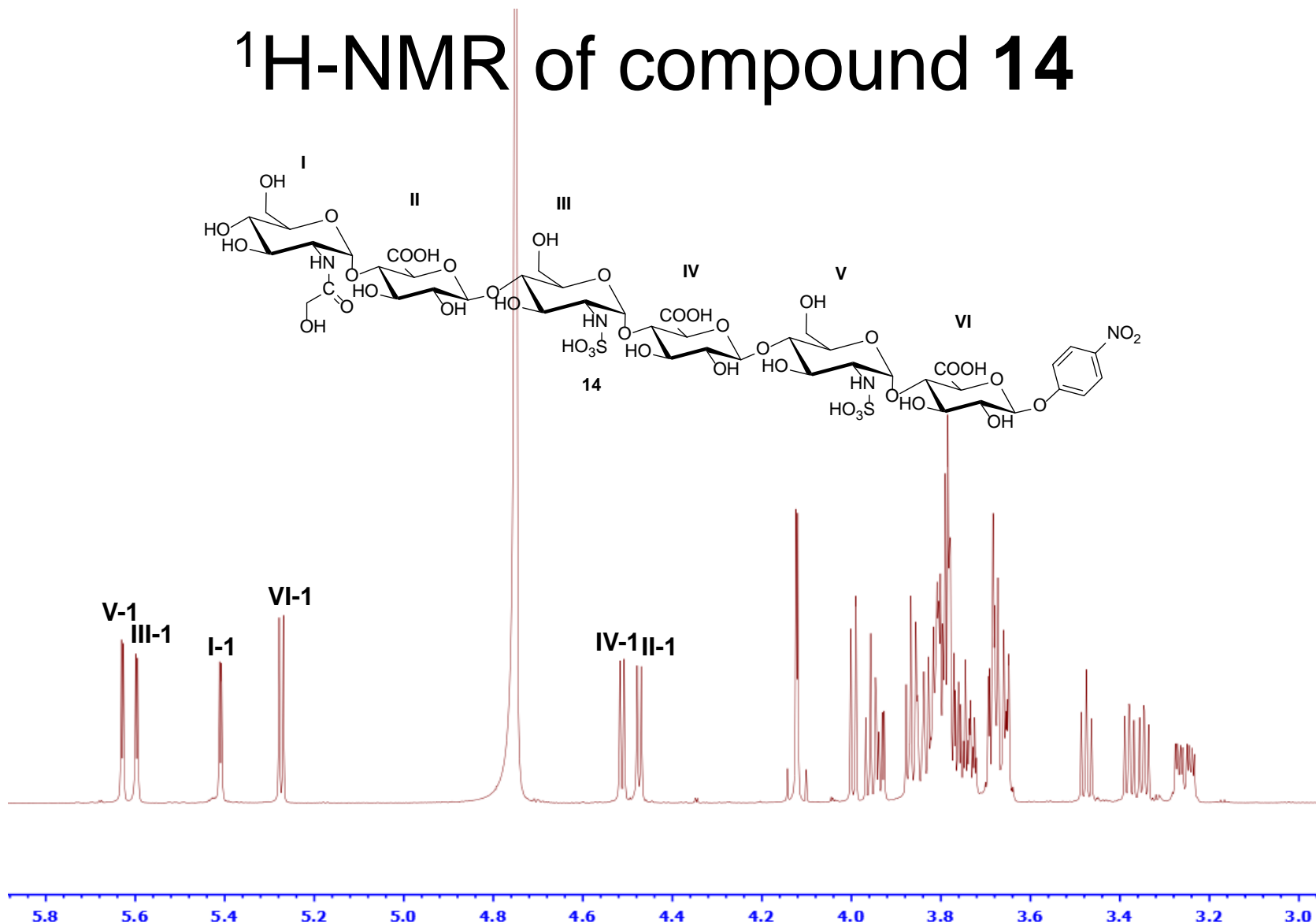
Supplementary Fig S22. $^1\text{H-NMR}$ of compound **13** (850 MHz, D_2O). The signals of anomeric protons are indicated. The anomeric protons resonate as doublet at δ 5.41 and 5.30 ppm. The anomeric proton signals from **13** were similar to the comparable ones observed from **14** (I-1, 5.41 ppm/VI-1, 5.27 ppm). The chemical structure of compound **13** is shown on top of figure.

Synthesis of 14



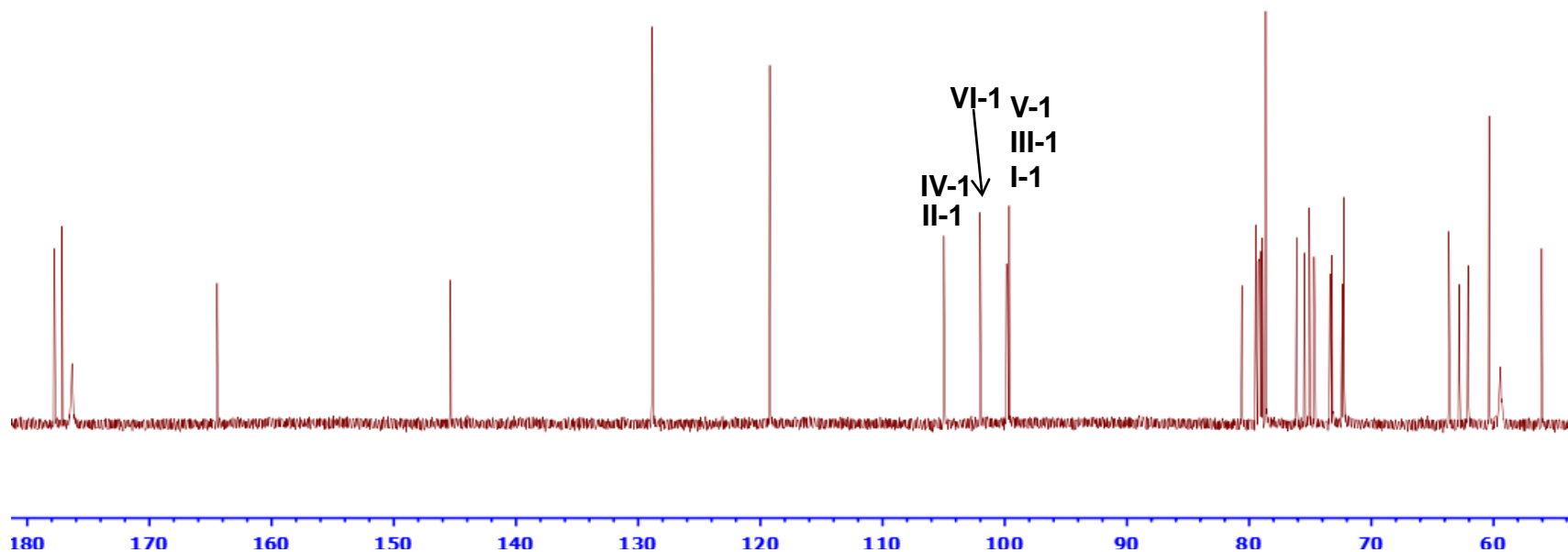
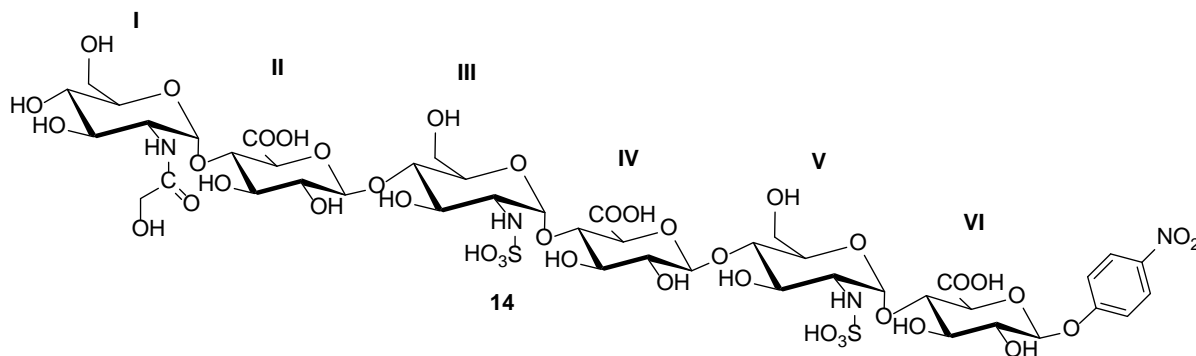
Supplementary Fig. S23. The scheme for the synthesis compound **14**. The starting hexasaccharide was an intermediate from the synthesis of compound **1**.

$^1\text{H-NMR}$ of compound **14**



Supplementary Fig S24. $^1\text{H-NMR}$ of compound **14** (850 MHz, D₂O). The signals of anomeric protons are indicated. The anomeric protons resonate as doublet at δ 5.63, 5.60, 5.41, 5.27, 4.51, and 4.47 ppm. Chemical structure of compound **14** is shown on top of figure.

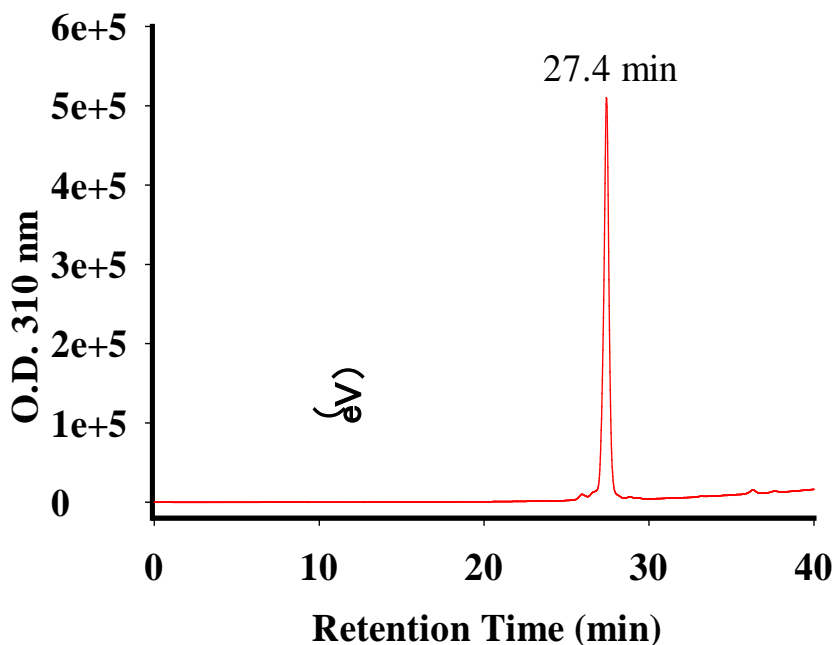
^{13}C -NMR of compound **14**



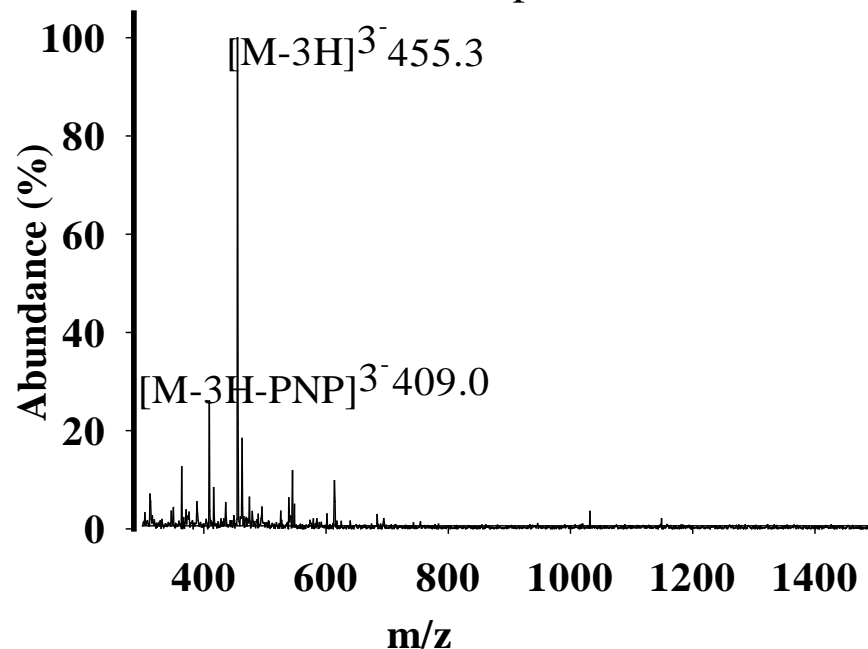
Supplementary Fig S25. ^{13}C -NMR of compound **14** (212.5 MHz, D_2O). The signals of anomeric carbons are indicated. The anomeric carbons resonate at δ 105.0, 104.9, 102.0, 99.9, 99.8, and 99.6 ppm. Chemical structure of compound **14** is shown on top of figure.

Purity and MS analysis of compound **14**

HPLC of compound **14**



MS of compound **14**



Supplementary Fig. S26. Left panel shows the PMAN-HPLC chromatogram of compound **14**. The compound was eluted as a single symmetric peak on PMAN-HPLC, suggesting that compound **14** was pure. Right panel shows the MS spectrum. The measured molecular mass was 1368.9, which is very close to the calculated molecular weight (1369.1). [M-3H-PNP]³⁻ refers to the molecular ion of compound **14** that lost PNP moiety.