Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Thioglycoligation of aromatic thiols using natural glucuronide donor.

Martyna Kurdziel [a,b,†] Magdalena Kopeć [a,b,†], Arnaud Pâris[a], Krzysztof Lewiński[b], Pierre Lafite[a,*], and Richard Daniellou[a,*]

[a.] *Institut de Chimie Organique et Analytique (ICOA), Université d'Orléans/CNRS, UMR 7311, BP 6759, F-45067, Orléans Cedex 2, France*
[b.] *Jagiellonian University, Faculty of Chemistry, Department of Crystal Chemistry and Crystal Physics, Gronostajowa 2, 30-387, Kraków, Poland*
†Both authors contributed equally to this work.

Supplemental data

1- Detailed experimental procedures

a. Cloning, expression, and purification of *Dt*GlcA WT and *Dt*GlcA E396Q

Amplification of *ditch_1429* (accession number ACI18585.1) gene coding for *Dt*GlcA (Uniprot B5YFE0) was performed by PCR using genomic DNA from *Dictyoglomus thermophilum* strain H-6-12 DSM3690 (DSMZ, Leibniz Institute, Germany). Primers containing NdeI (upstream) and HindIII (downstream) restriction sites (underlined) were used to insert cloning sites : 5' AACATATGCTTTATCCTAAAGAGAGTGAAA-3' and 5'- TTAAGCTTCTATTTTTCCCAAAAATCA-3'. Following proper digestion, insert was ligated into pET-28a(+) expression vector (Merck, France) using both restriction sites to yield pET28-*DtGlcA* construct. The corresponding gene encode for *Dt*GlcA fused with a N-terminal polyhistidine tag. Site-directed mutagenesis was performed using Quick Change Lightening Site-Directed Mutagenesis Kit (Agilent Genomics, France), according to manufacturer procedures, and gave a mutated pET28-*Dt*GlcA-E396Q in which the E396 residue was mutated to glutamine. Sequencing of inserts confirmed proper cloning and mutation (Eurofins Genomics, France).

E. coli Rosetta(DE3) strain transformed with pET28-DtGlcA or pET28-DtGlcA in LB medium containing Chloramphenicol and Kanamycin was incubated for 2-3h at 37°C under 250rpm shaking until OD_{600} reached 0.6. Induction was carried out by addition of IPTG (1mM final concentration), and incubated at 25°C.After an overnight culture, cells were harvested, washed by Tris 20mM pH 8.0 buffer, and pellets were frozen at -20°C until subsequent purification.

Frozen pellets were resuspended in lysis buffer (Tris-HCl pH 8.0, 50 mM, NaCl 100 mM), incubated with PMSF (phenylmethylsulfonylfluoride) and lysozyme at 4°C for 20 minutes. Then cells were lyzed by several freeze-thaw cycles, and sonications. Soluble fraction was loaded onto a Nickel-containing IMAC column (HisPure, Thermo Scientific), washed with lysis buffere supplemented with 10mM imidazole pH 8.0, and elution was carried out by increasing imidazole concentration to 500mM. Protein purity was monitored by SDS-PAGE analysis (>99%) and was sufficient for enzymatic studies or chemoenzymatic synthesis.

For crystallogenesis assays, *Dt*GlcA was further purified by Size Exclusion Chromatography (SEC).

b. Crystallogenesis of *Dt*GlcA and structure solving.

Following SEC purification in Tris 20 mM / NaCl 100 mM buffer, *Dt*GlcA buffer was concentrated to 0.5 mg/ml, buffer was exchanged to Tris 20mM pH 8.0 and screening assays were performed using *JBScreen Classic HTS* I and *JBScreen Classic HTS II* screening kits (Jena Bioscience, Jena, Germany). Crystals suitable for diffraction were obtained in HTSII C2 conditions (20% w/v 2-methyl-2,4-pentanediol; 100mM Sodium (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) HEPES pH 7.5; 100mM Sodium citrate). Crystals were flash-frozen in liquid nitrogen, data was collected at Proxima 1 Beamline (SOLEIL synchrotron, Saint-Aubin, France), and reduced and scaled using X-ray Dectetor Software(XDS)[1]. Molecular replacement using PDB 1BHG as template was performed using Phaser[2], the resulting model was completed by iterative cycles of manual model building and real space refinement using the program Coot[4] and crystallographic refinement using phenix.refine[3]. The processing and final refinement statistics are presented in Supplemental table. Structure was deposited in the PDB with the accession code 6XXW.

c. Enzymatic assays

*General procedure : Dt*GlcA-WT and E396Q activity were assayed in triplicate in a 96 well-plate at 37°C in 200 µL reaction mixtures containing substrate, enzyme (0.2 µM final concentration) and phosphate citrate buffer (200 mM, pH 5.0). Residual spontaneous hydrolysis of the substrate was determined on sample containing dH_2O instead of enzyme. For para-nitrophenol (pNP) containing substrates, after 15 min reaction, 100 µL of sodium carbonate 1 M were added, and produced pNP was quantified by absorbance measurement at 405 nm.

Specificity : General procedure was performed for each substrate tested (1mM final concentration) (Carbosynth, UK):
4-nitrophenyl-α-D-glucopyranoside, 4-nitrophenyl-β-D-glucopyranoside, 4-nitrophenyl-α-D-galactopyranoside, 4-nitrophenyl-β-D-galactopyranoside, 4-nitrophenyl-α-D-mannopyranoside, 4-nitrophenyl-β-D-mannopyranoside, 4-nitrophenyl-α-D-xylopyranoside, 4-nitrophenyl-β-D-xylopyranoside, 4-nitrophenyl-α-L-fucopyranoside, 4-nitrophenyl-β-L-fucopyranoside, 4-nitrophenyl-β-D-fucopyranoside, 4-nitrophenyl-α-D-2-deoxy-2-*N*-Acetyl-glucosaminopyranoside, 4-nitrophenyl-β-D-2-deoxy-2-*N*-Acetyl-glucosaminopyranoside, 4-nitrophenyl-α-L-arabinopyranoside, 4-nitrophenyl-β-L-arabinopyranoside, 4-nitrophenyl-α-L-rhamnopyranoside, 4-nitrophenyl-α-L-arabinofuranoside, 4-nitrophenyl-β-D-ribofuranoside, 4-nitrophenyl-β-D-galactofuranoside, 4-nitrophenyl-β-D-glucuronide (*p*NP-GlcA).

Kinetics constants determination : General procedure was performed, with *p*NP-GlcA concentrations ranging from 10 to 7500 µM.

pH, temperature and co-solvent dependence : General procedure was done by modifying the pH of the phosphate-citrate buffer, temperature of the incubation, or by addition of the appropriate amount of organic solvent to the reaction mixture.

d. Thioglycosylation general procedure

Analytical scale
4-Nitrophenyl β-D-glucuronide or baicalin (1 mM), DTT (5 mM) and acceptor (1mM) were mixed in citric acid/sodium phosphate buffer (25mM, pH 5). DtGlcA E396Q was added (0.2 µM), for a total volume of 200 µl. After an overnight incubation at 37°C, reaction was stopped and acidified by addition of 100 µl of quenching solution (Acetonitrile:Formic acid/10:1) before analysis on HPLC. Injected solution (typically 5µl) were separated on a Zorbax Eclipse XDB-C18 column (4.6x150 mm, 3.5-Micron, Agilent Technologies) at a flow rate of 1 mL/min and maximum pressure limit of 300 bar. Chromatographic analysis was carried out on 1220 Infinity II LC system (Agilent Technologies, France) with the Infinity II Diode Array Detector (DAD). The composition of mobile phase was water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (solvent B) and it was used according to the following elution gradient: 0-14min: 90% A; 14-16min: 40% A; 16-21min 0% A; 21-25min 90% A.
Chromatograms are presented in Supplemental data part 4. Compounds were identified according to their retention time and absorbance maxima indicated.

Preparative scale

4-Nitrophenyl β-D-glucuronide (5mg, 1 eq.), DTT (200 mM) and acceptor (10 eq.) were mixed in citric acid/sodium phosphate buffer (25mM, pH 5). DtGlcA E396Q was added (0.2 µM), for a total volume of 1ml. After an overnight incubation at 37°C, reaction was stopped and acidified by addition of 500 µl of quenching solution (Acetonitrile:Formic acid/10:1).

Semi-preparative HPLC purification was done on the similar HPLC system (1220 Infinity II LC system) and buffer composition, on a Zorbax Eclipse XDB-C18 column (9.4x150 mm, 5 µ, Agilent Technologies) at a flow rate of 4 mL/min.

Fractions were further analyzed on HPLC as described above to assess purity. Corresponding chromatogram of purified products are presented below.

HRMS analysis of purified product was performed on a Bruker maXis UHR-Q-TOF spectrometer (Bremen, Germany), and 1H NMR and 13C NMR spectra were recorded on Bruker Avance II 400 and spectrometer.

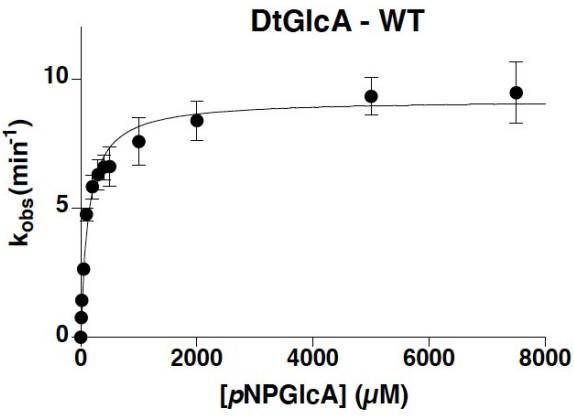
2- Data collection and refinement statistics of *Dt*GlcA crystal structure.

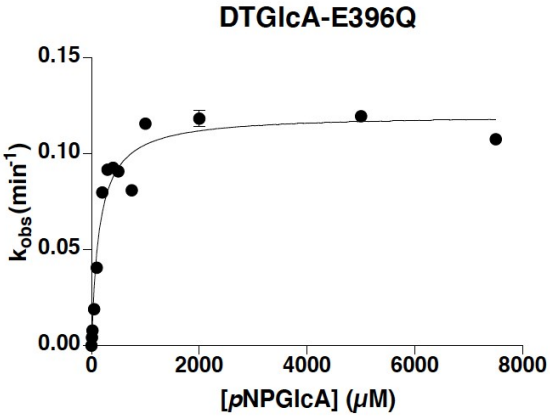
	***Dt*GlcA – PDB 6XXW**
Data collection	
Space Group	I222
Cell dimensions	
a,b,c (Å)	79.13, 115.82, 130.54
α,β,γ (°)	90, 90, 90
Resolution (Å)[a]	33.83-1.85 (1.96-1.85)
R_{merge}(%)[a]	11.1 (74.3)
R_{meas}(%)[a]	12.5 (83.7)
I/σ[a]	12.15 (2.46)
CC1/2 (%)[a]	99.7 (70.9)
Completeness (%)[a]	97.7 (97.0)
Redundancy[a]	4.68 (4.53)
Refinement	
Resolution	33.83-1.85
Number of reflections	50180
R_{work}/R_{free}	0.16/0.21
Number of atoms	5240
Water molecules	364
Ligand molecules	10
R.M.S. Bond lengths (Å)	0.007
R.M.S. Bond angles (°)	0.831
Overall B-factor (Å^2)	25.44

[a] Data in brackets correspond to highest resolution shell.

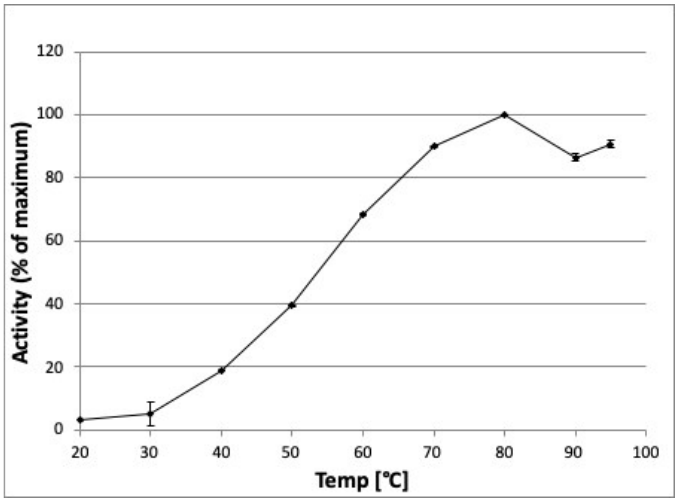
3- Kinetic data of DtGlcA WT and E396Q

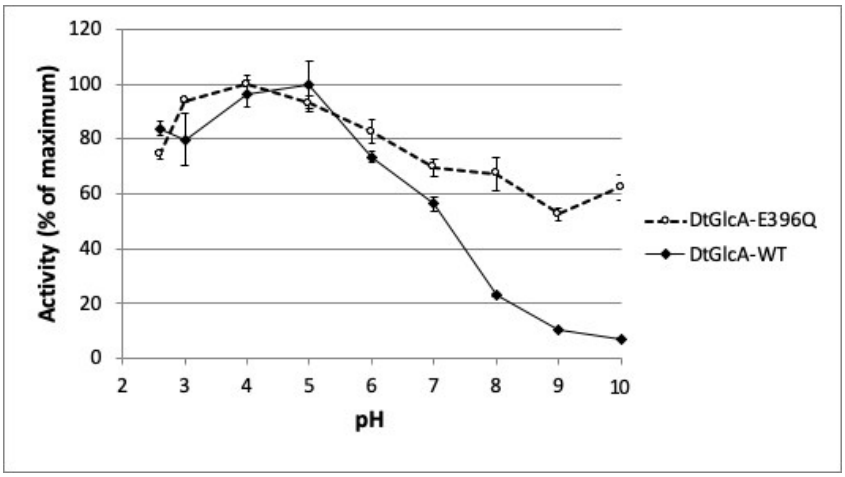
a. Michaelis Menten plots



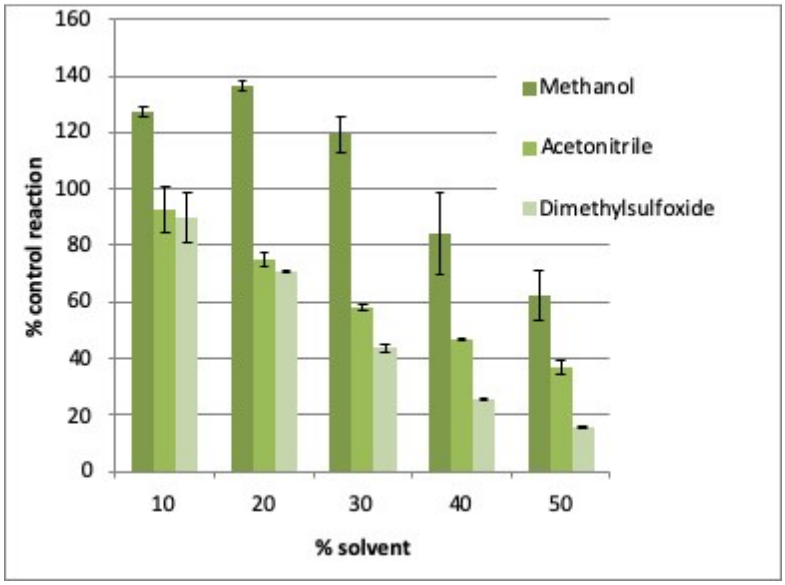


b. Temperature and pH dependence



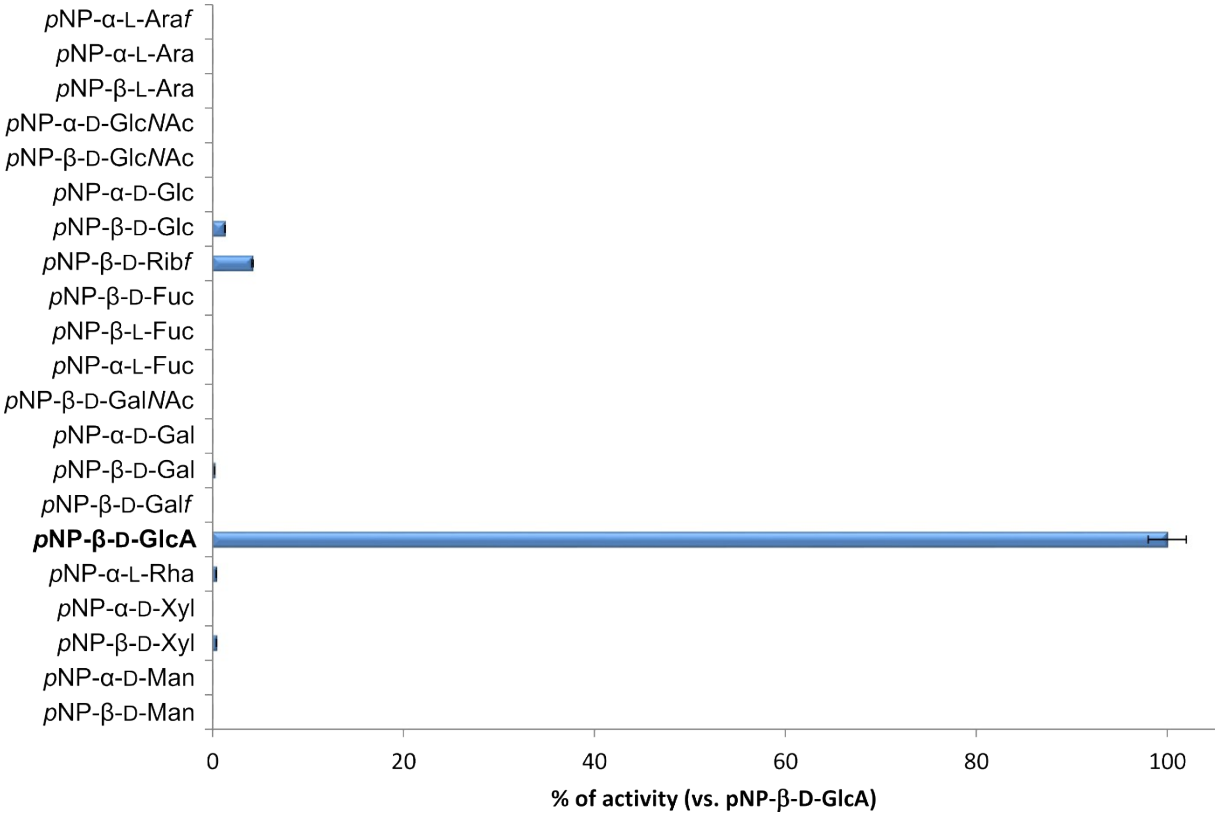


c. Co-solvent assays



Activity >100% observed for methanol below 30% can be attributed to transglycosylation reaction that increases pNP release rate compared to pure hydrolysis.

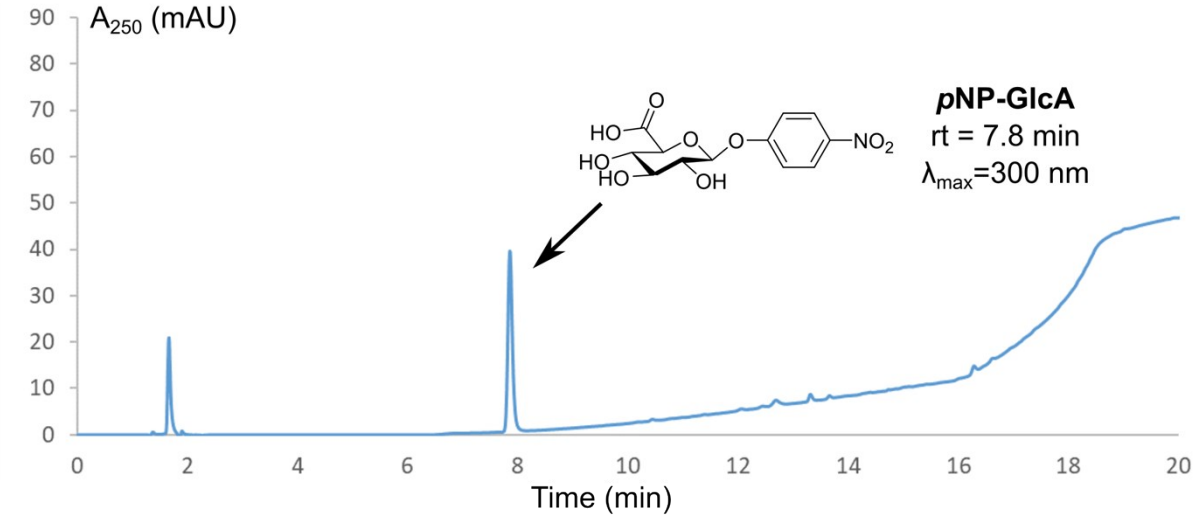
d. Substrate specificity

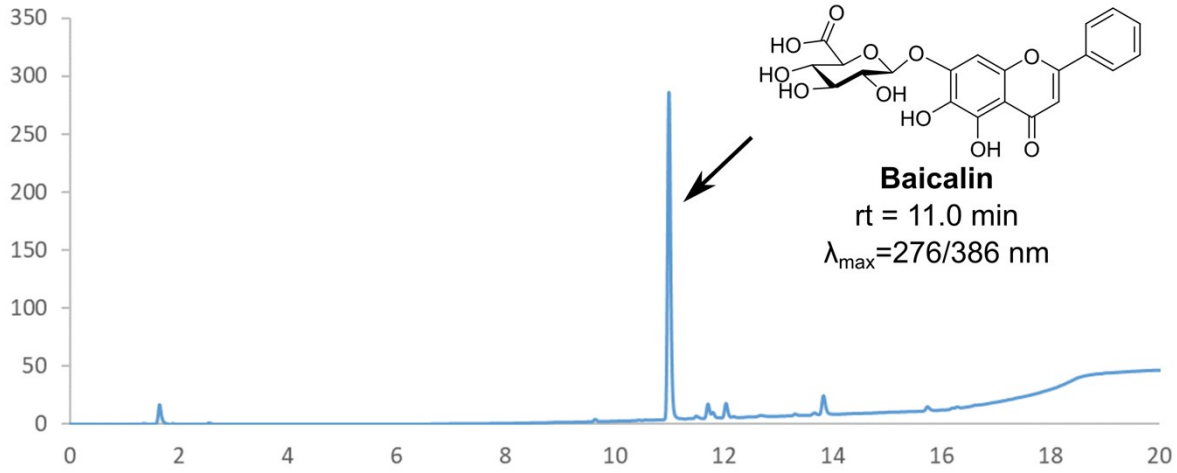


4- HPLC Analysis of reactions and purified products

a. Sugar donors standards

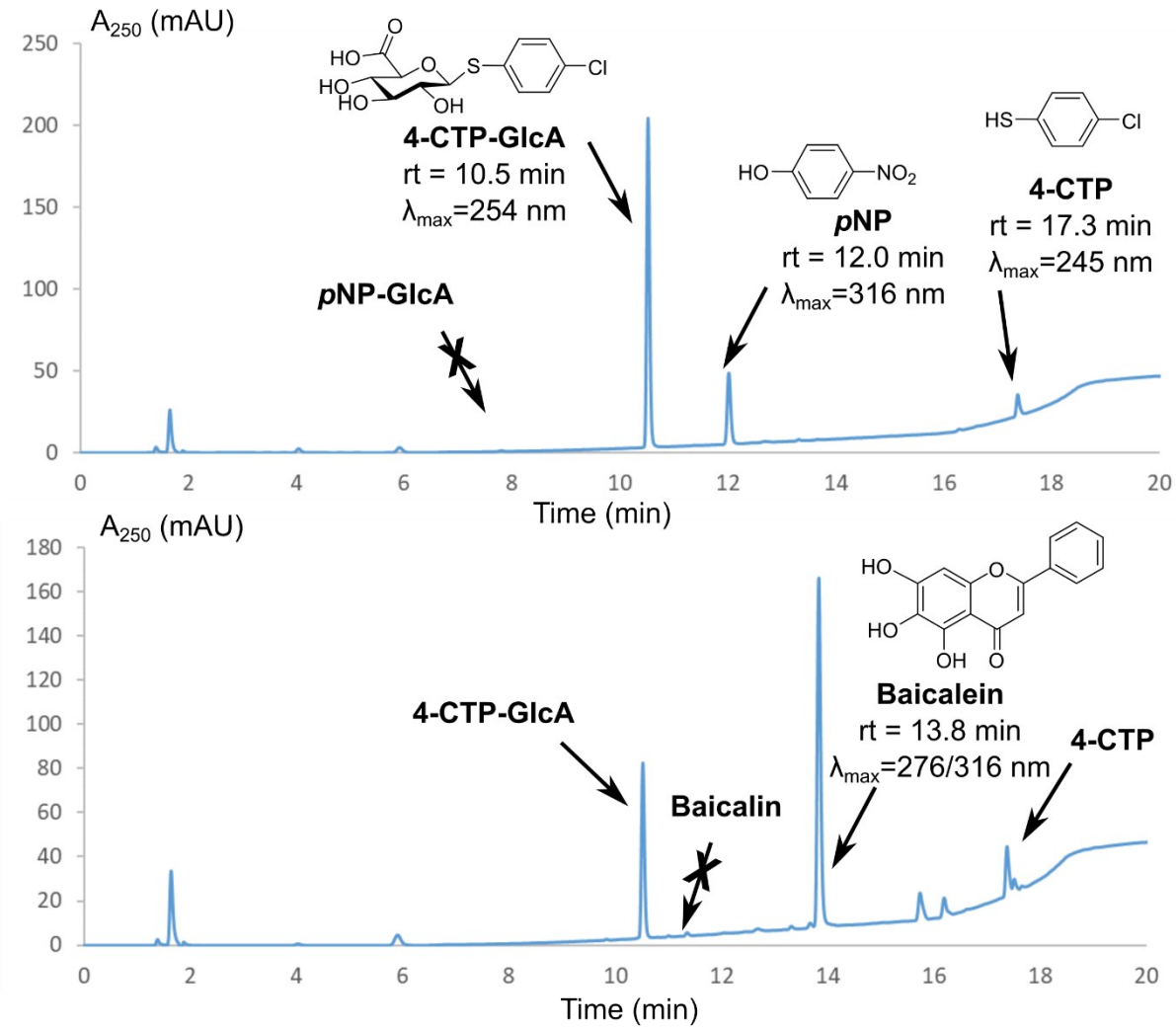
Substrates (sugar donors)



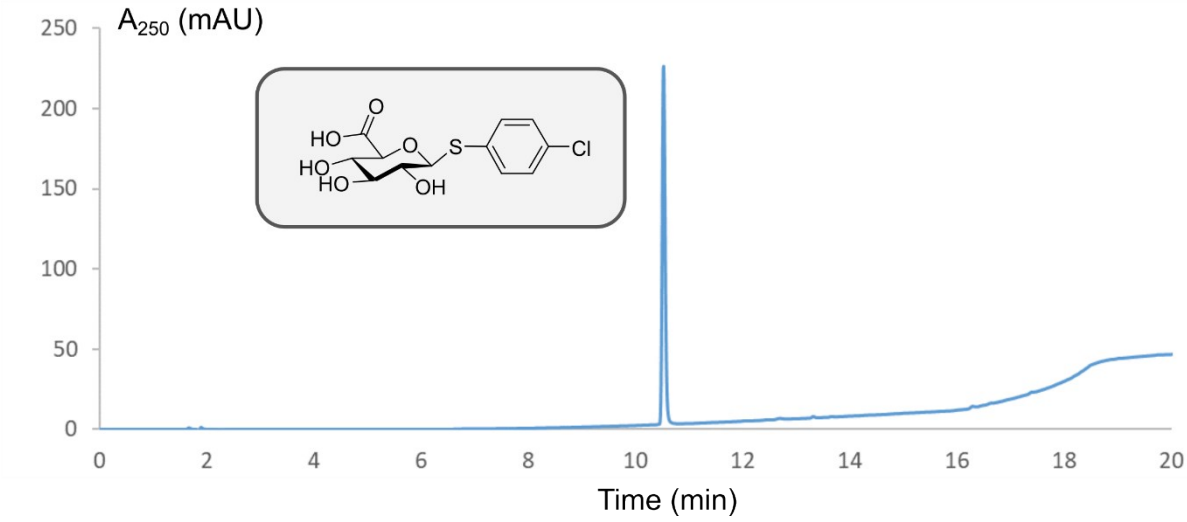


b. 4-Chlorophenyl-1-thio-β-D-glucopyranosiduronic acid

S-thioglycoligation (24h, 37°C, donor:acceptor 1:1)

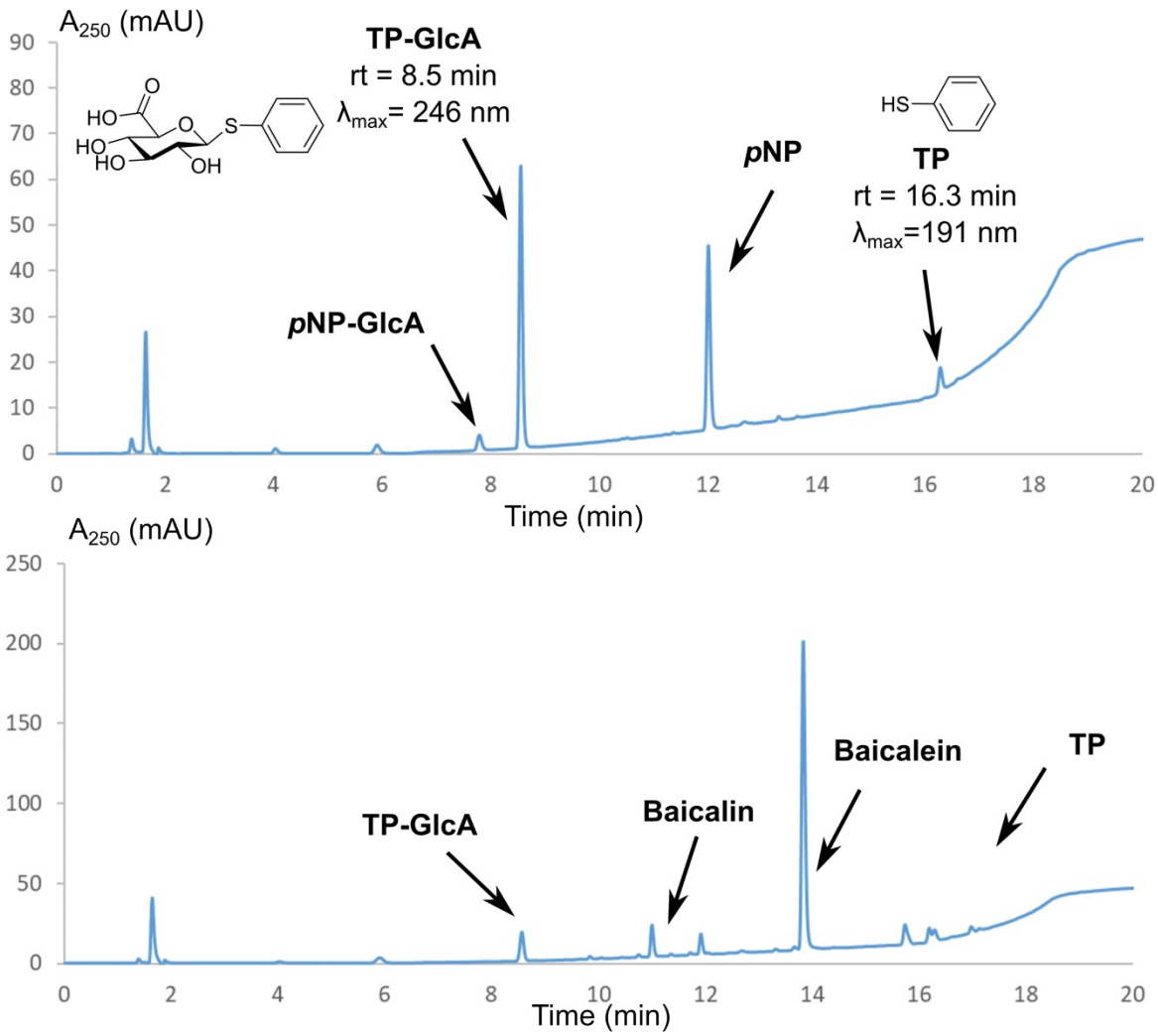


Purified *S*-thioglycoside (24h, 37°C, donor:acceptor 1:10)

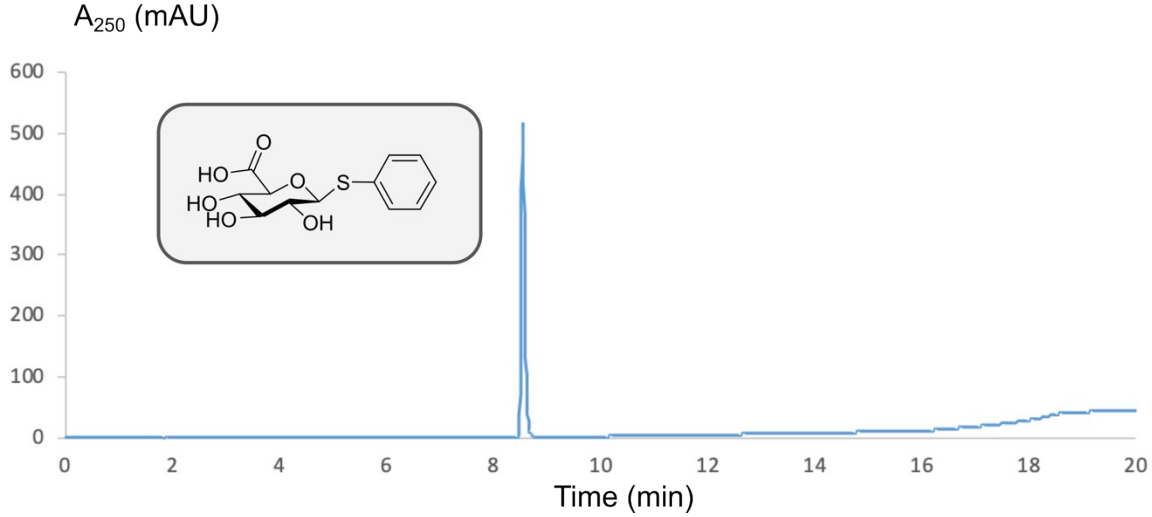


c. Phenyl-1-thio-β-D-glucopyranosiduronic acid

S-thioglycoligation (24h, 37°C, donor:acceptor 1:1)

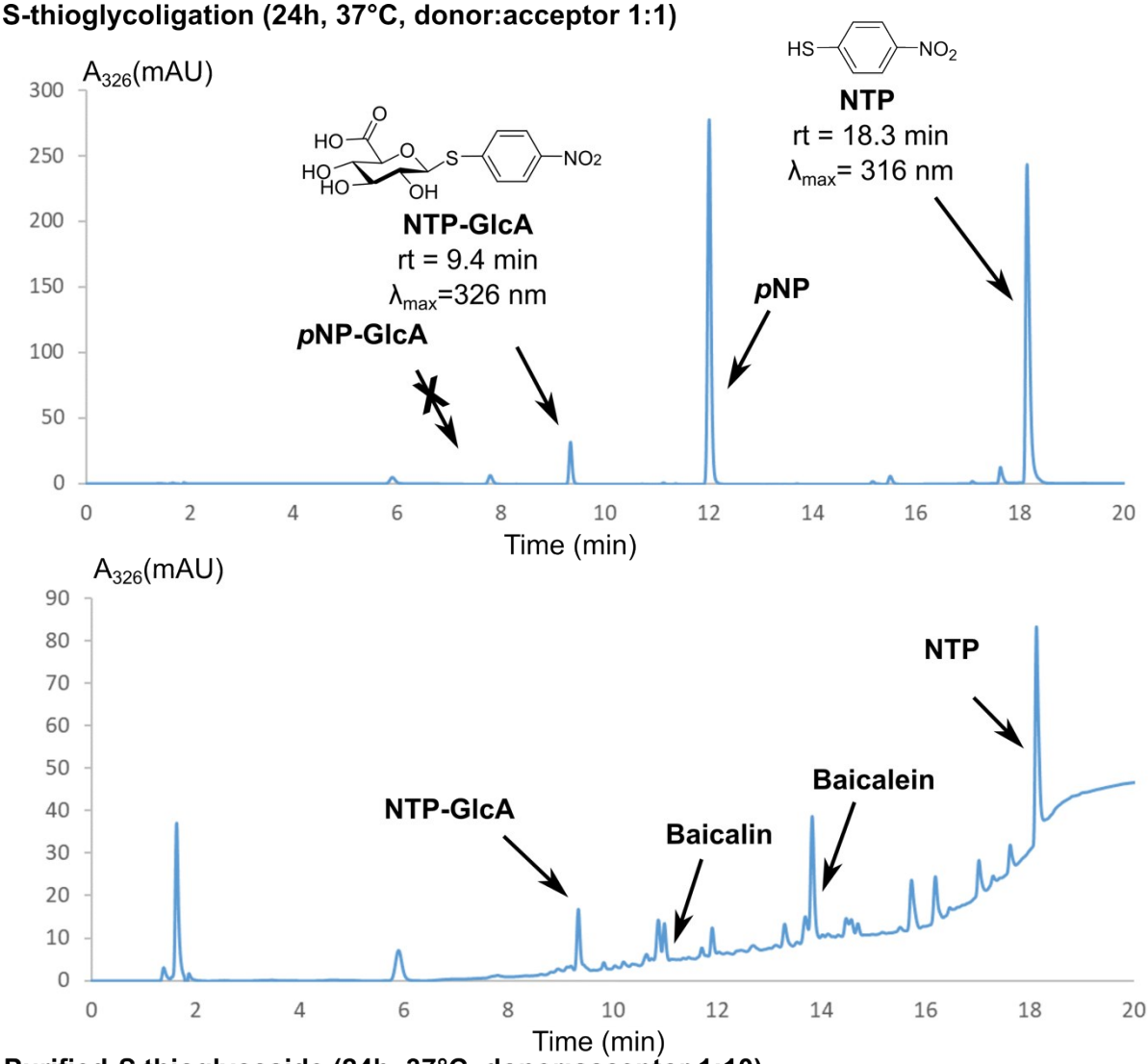


Purified *S*-thioglycoside (24h, 37°C, donor:acceptor 1:10)

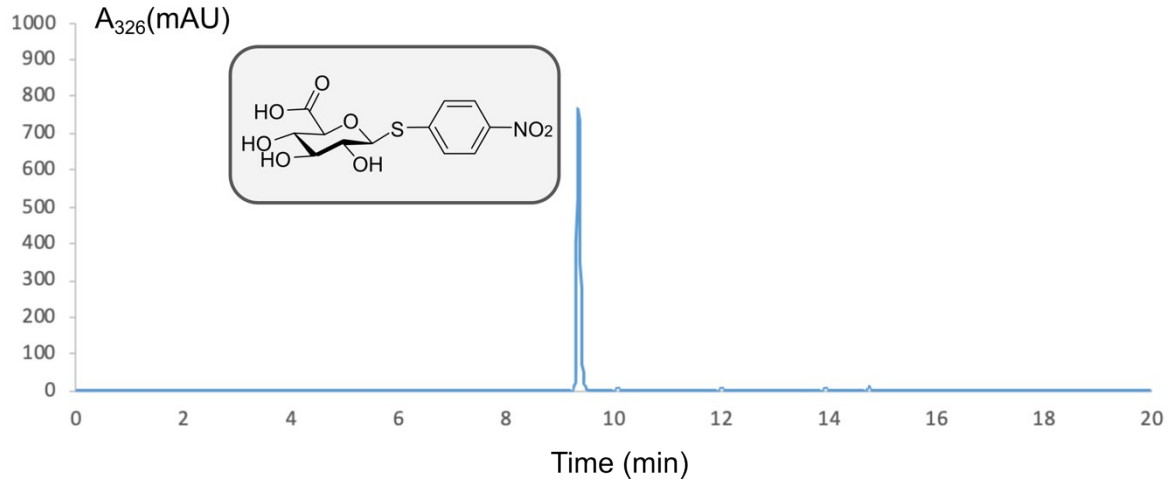


d. 4-Nitrophenyl-1-thio-β-D-glucopyranosiduronic acid

S-thioglycoligation (24h, 37°C, donor:acceptor 1:1)

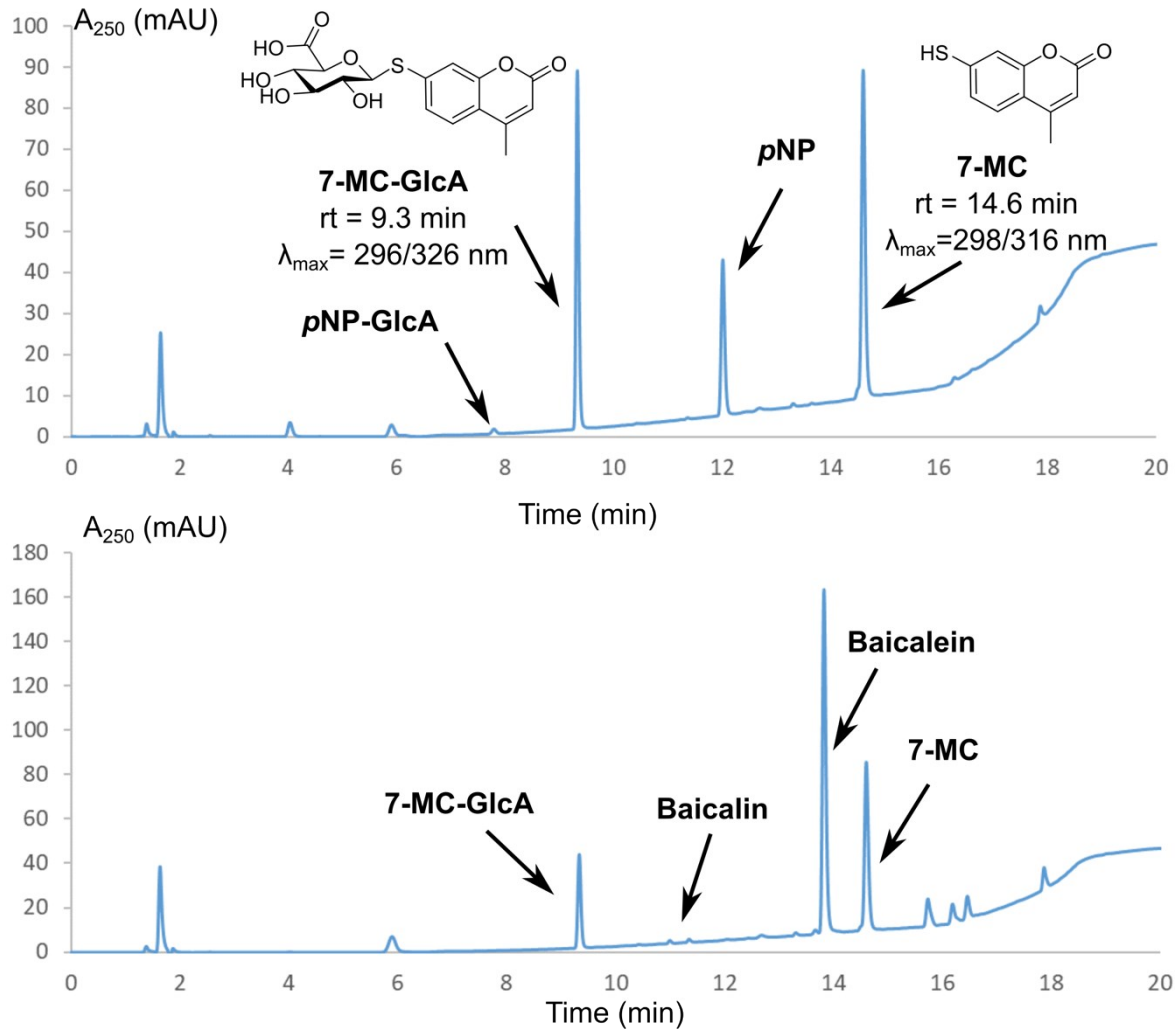


Purified *S*-thioglycoside (24h, 37°C, donor:acceptor 1:10)

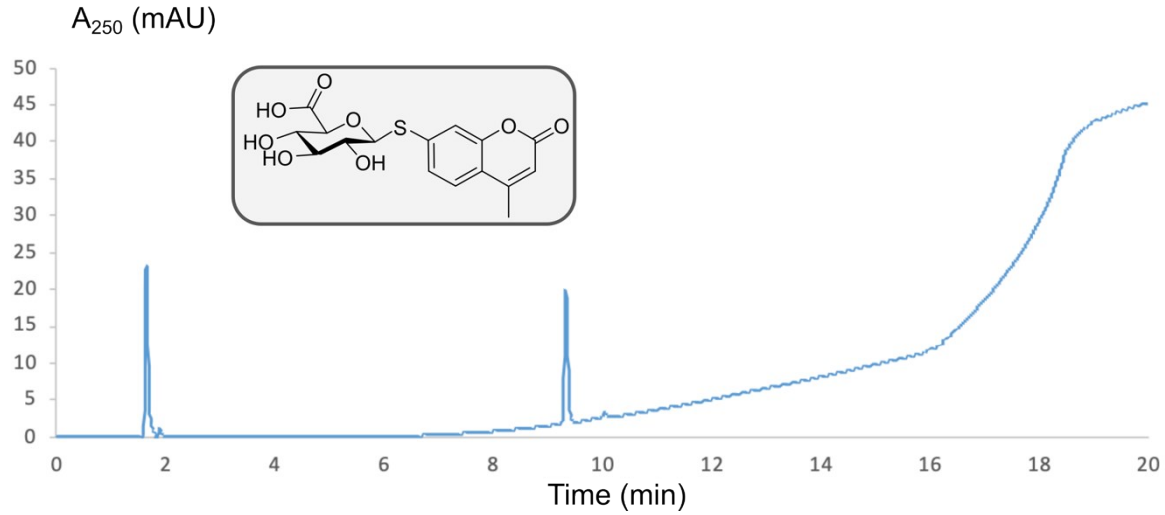


e. 4-Methylumbellifer-7-yl-1-thio-β-D-glucopyranosiduronic acid

S-thioglycoligation (24h, 37°C, donor:acceptor 1:1)

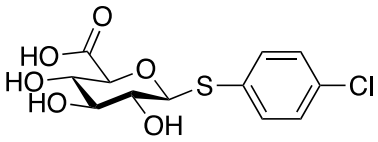


Purified *S*-thioglycoside (24h, 37°C, donor:acceptor 1:10)



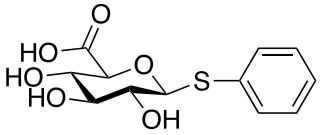
5- NMR Spectra

4-Chlorophenyl-1-thio-β-D-glucopyranosiduronic acid



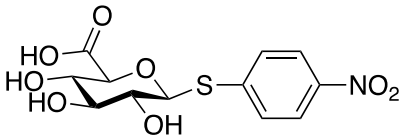
^{1}H-NMR (400 MHz, DMSO-d_6 + Deuterium Oxide) δ 7.46 (d, *J* = 8.7 Hz, 2H, Ar), 7.33(d, 2H, Ar), 4.59 (d, *J* = 9.8 Hz, 1H, H-1,), 3.46 (t, H-5), 3.23 (m, 2H, H-3, H-4), 3.03 (t, , *J* = 8.7 Hz, 1H, H-2).

Phenyl-1-thio-β-D-glucopyranosiduronic acid



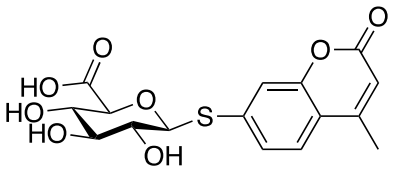
^{1}H-NMR (250 MHz, DMSO-d_6 + Deuterium Oxide) δ 7.4-7.2 (m, 5H, Ar.), 4.62 (d, *J* = 9.7 Hz, 1H, H-1), 3.61 (t, 1H, H-5), 3.29-3.23 (m, 2H, H-3, H-4), 3.07 (t, *J* = 8.8 Hz, 1H, H-2)

4-Nitrophenyl-1-thio-β-D-glucopyranosiduronic acid[1]



^{1}H-NMR (250 MHz, DMSO-d_6 + Deuterium Oxide) δ 8.08 (d, *J* = 8.0 Hz, 2H, Ar), 7,54 (d, 2H, Ar), 4.86 (d, *J* = 9.5 Hz, 1H, H-1), 3.55 (t, 1H, H-5), 3.34 (m, 2H, H-3, H-4), 3.21 (t, 1H, H-2).

4-Methylumbellifer-7-yl-1-thio-β-D-glucopyranosiduronic acid[2]

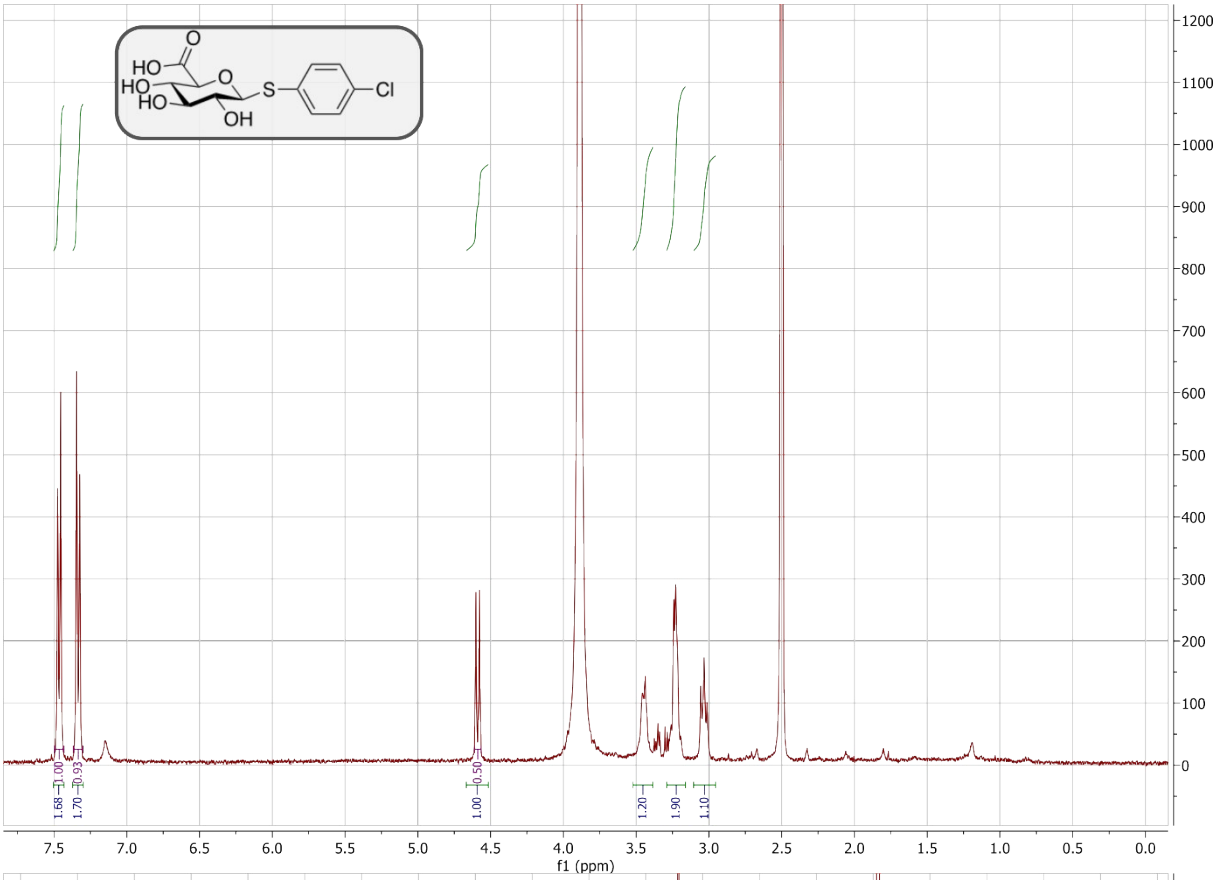


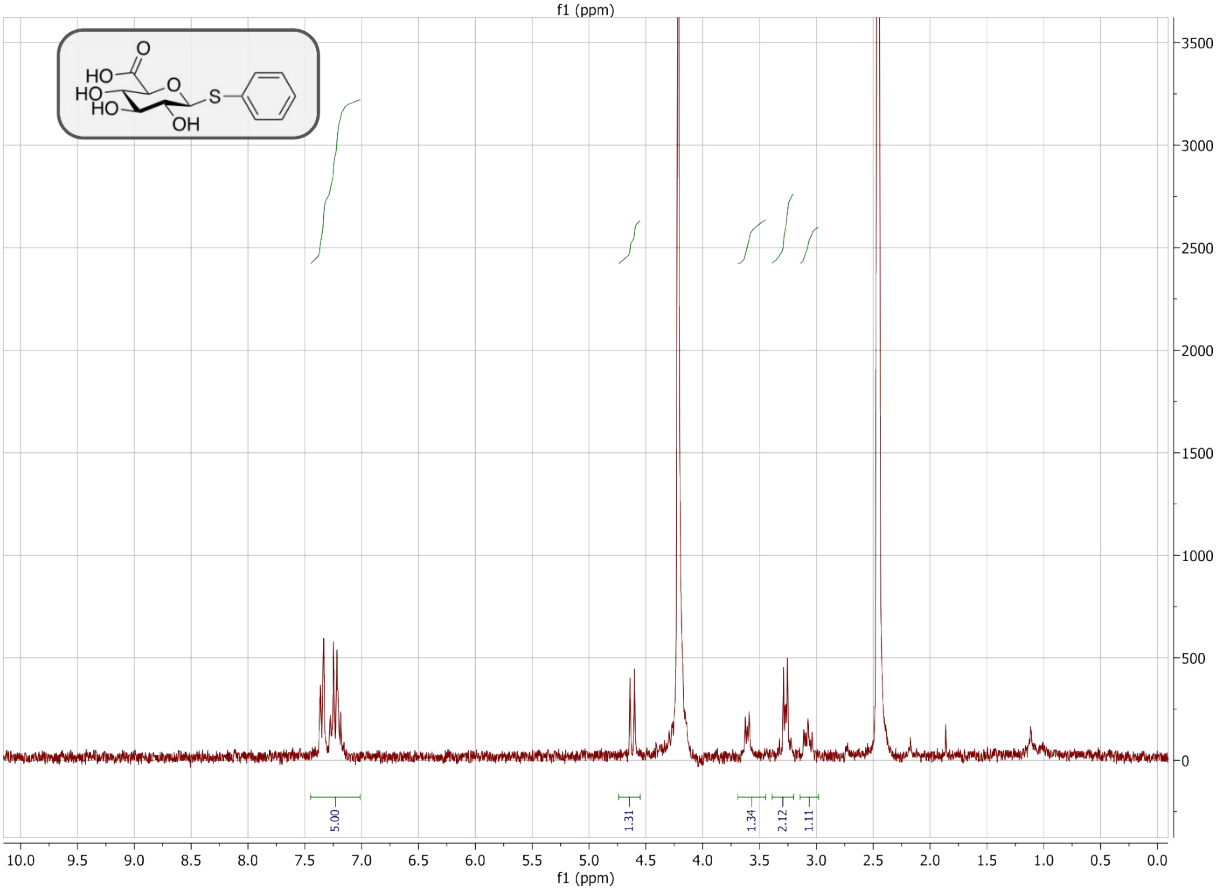
^{1}H-NMR (250 MHz, DMSO-d_6 + Deuterium Oxide) δ 7.67 (dd, *J* = 8.3, 1.5 Hz, 1H, Ar), 7.42-7.36 (m, 2H, Ar), 6.27 (d, *J* = 1.5 Hz, H-3'), 4.82 (d, *J* = 9.7 Hz Hz, 1H, H-1), 3.51 (d, *J* = 9.2 Hz, 1H, H-5), 3.37-3.21 (m, 2H, H-3, H-4), 3.16 (t, *J* = 8.7 Hz, 1H, H-2), 2.36 (s, 3H, Me).

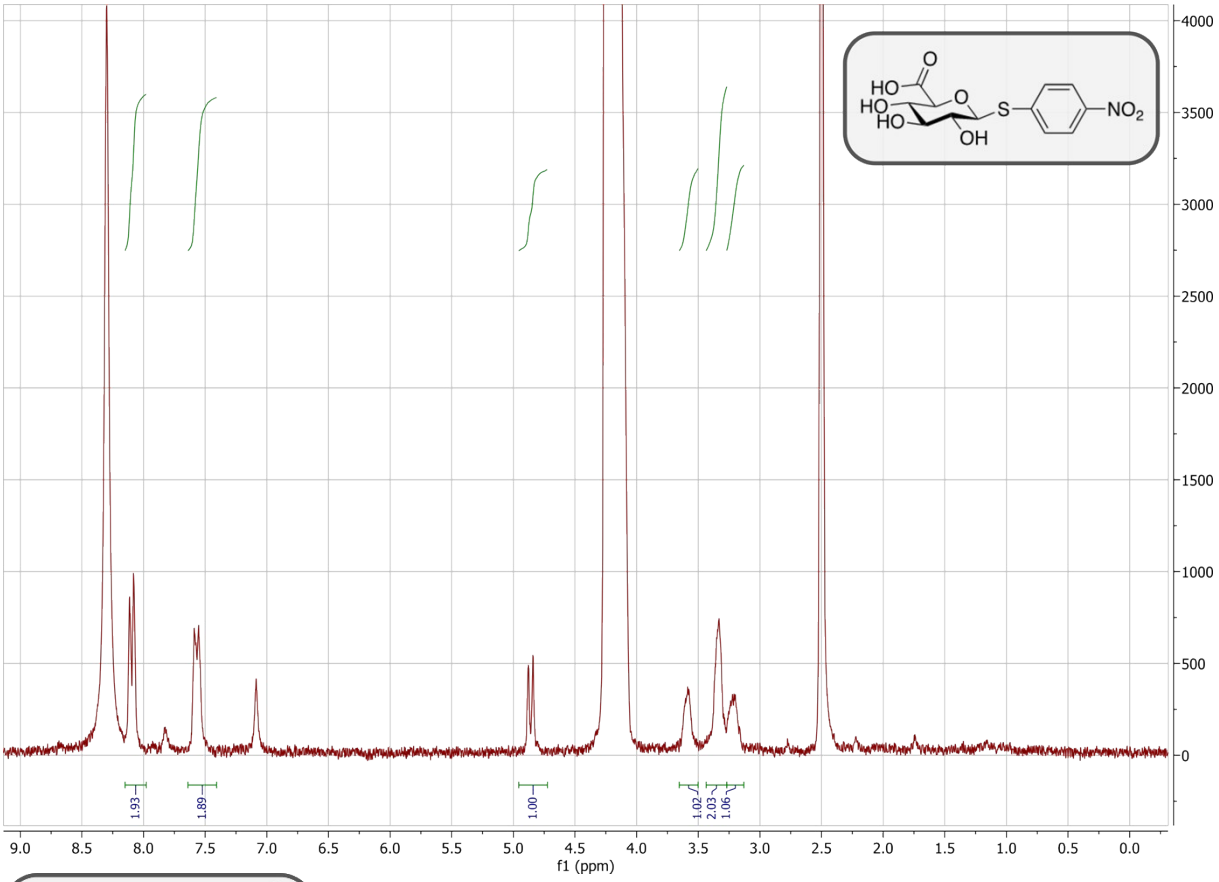
---

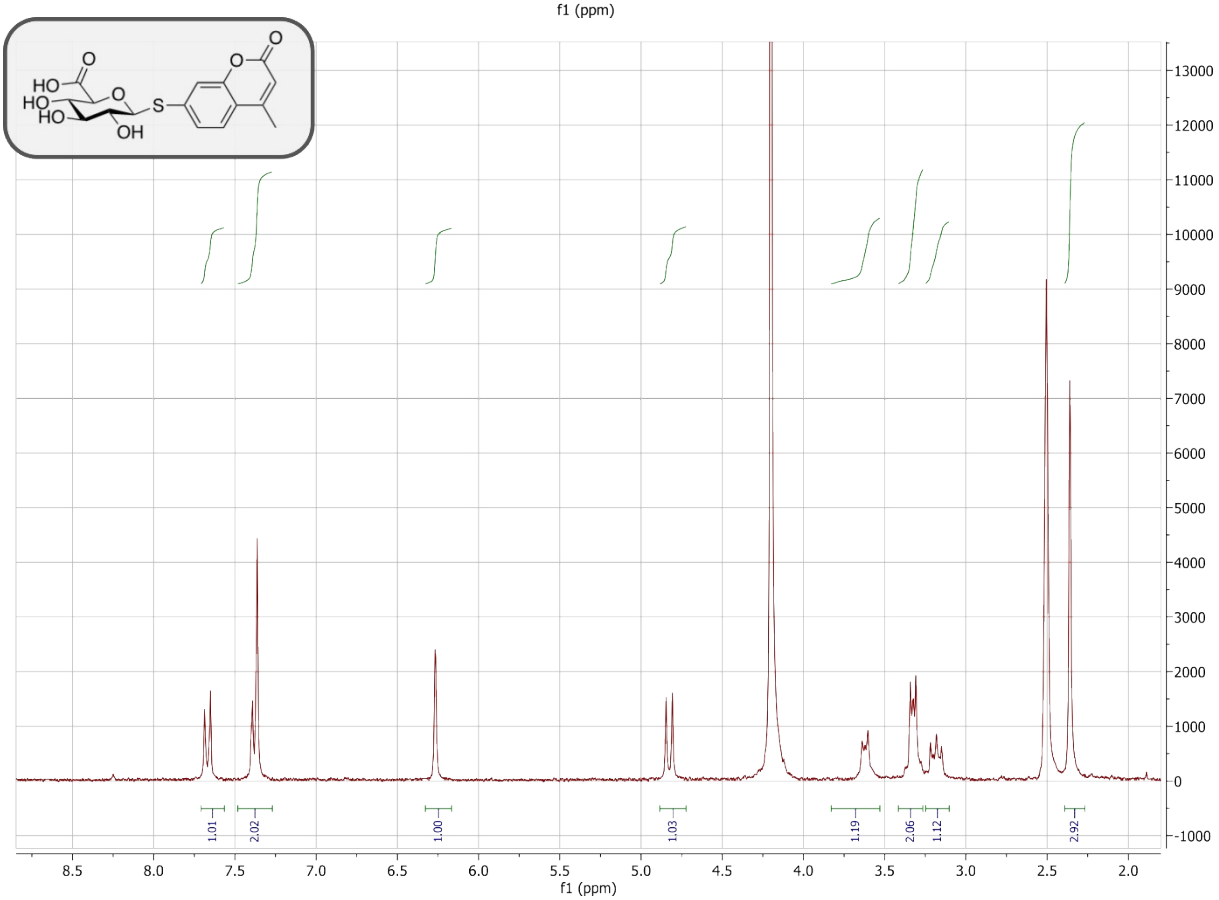
[1] Smith, J.E., Ross, D., Graham, A.B., Skellern, G.G. J, *J. Pharmaceut. Biomed.,* **1992**, 10, 461-463.

[2] Nasseri, S.A., Betschart, L., Opaleva, D., Rahfeld, P., Withers, S.G., *Angew. Chem. Int. Ed. Engl*. **2018**, 57(35):11359-11364


HO
HO
HO
O
O
S
Cl
OH
1.68
1.00
1.70
0.93
1.00
0.50
1.20
1.90
1.10
f1 (ppm)


HO
HO
HO
O
O
S
OH
5.00
1.31
1.34
2.12
1.11
f1 (ppm)


HO
HO
HO
O
O
S
OH
NO2
1.93
1.89
1.00
1.02
2.03
1.06
f1 (ppm)


HO
HO
HO
O
O
S
OH
O
1.01
2.02
1.00
1.03
1.19
2.06
1.12
2.92
f1 (ppm)

6- HRMS Spectra

Fédération de Recherche Physique et Chimie du Vivant (FR2708 : CBM/ICOA)
Plate-forme de Spectrométrie de Masse Haute Résolution

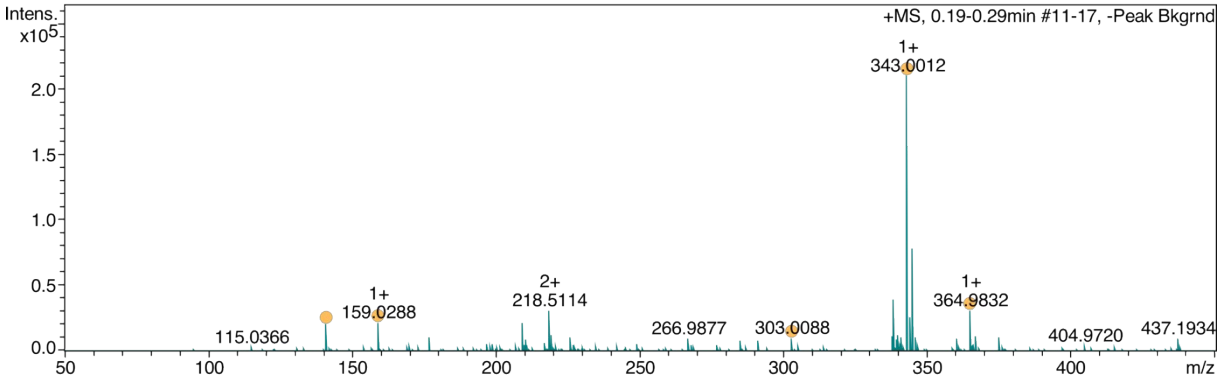
HRAM

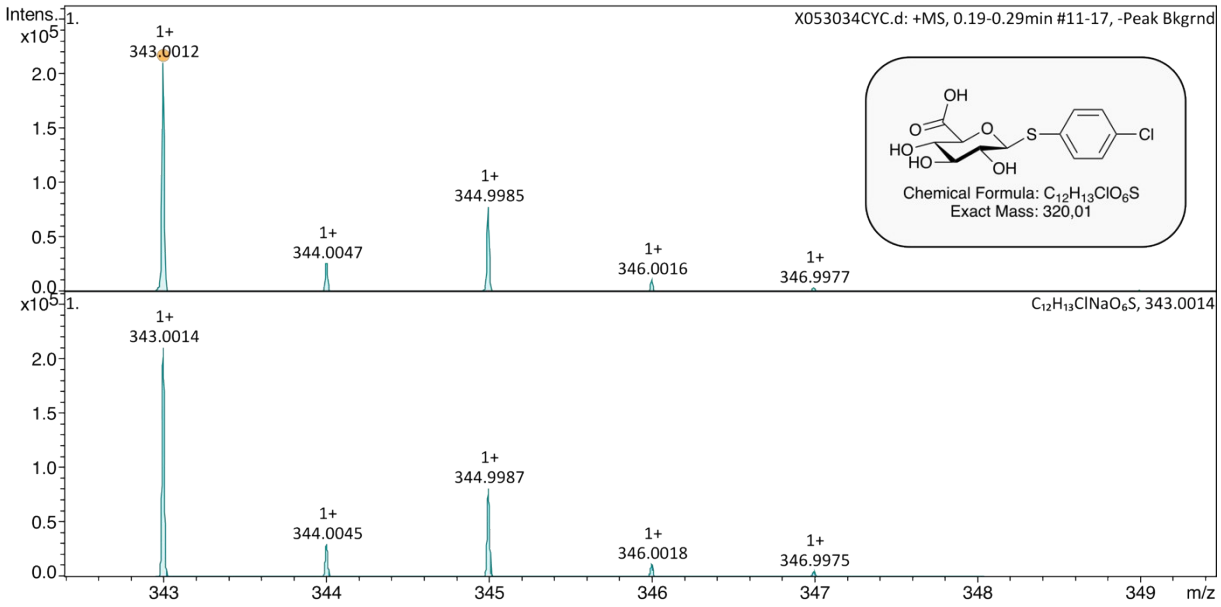
Analysis Info

Sample Name	**R1**	Acquisition Date	06/11/2019 14:24:47
Analysis Name	X053034CYC.d	Instrument / Ser#	maXis 255552.00086
		Method	Positif.m

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.6 Bar
Scan Begin	50 m/z	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan End	2500 m/z	Set Collision Cell RF	1800.0 Vpp	Set Dry Gas	7.0 l/min





Meas. m/z	z	#	Ion Formula	m/z	err [ppm]	mSigma	rdb	e⁻ Conf
141.018169	1+	1	C6H5O4	141.018235	0.5	14.4	5.0	even
159.028840	1+	1	C6H7O5	159.028800	-0.3	23.0	4.0	even
303.008782	1+	1	C12H12ClO5S	303.008849	0.2	18.1	7.0	even
343.001244	1+	1	C12H13ClNaO6S	343.001358	0.3	10.8	6.0	even
364.983153	1+	1	C12H12ClNa2O6S	364.983302	0.4	9.8	6.0	even

Fédération de Recherche Physique et Chimie du Vivant (FR2708 : CBM/ICOA)
Plate-forme de Spectrométrie de Masse Haute Résolution

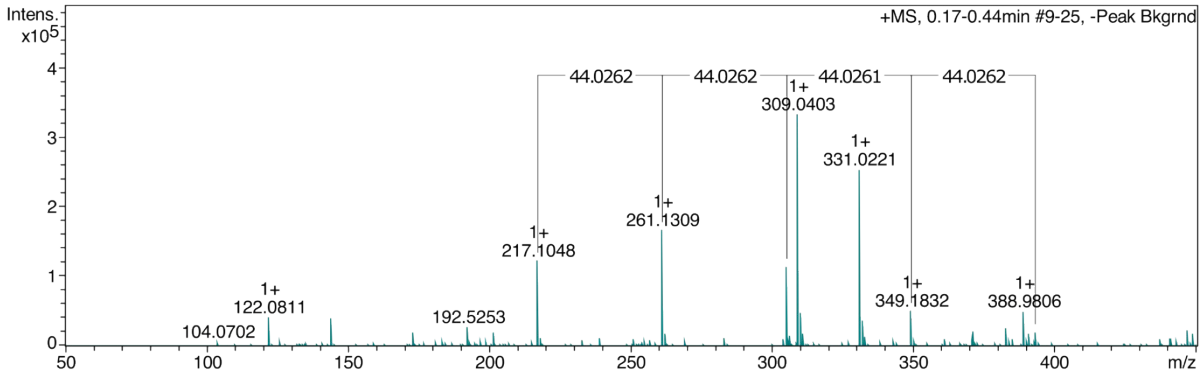
HRAM

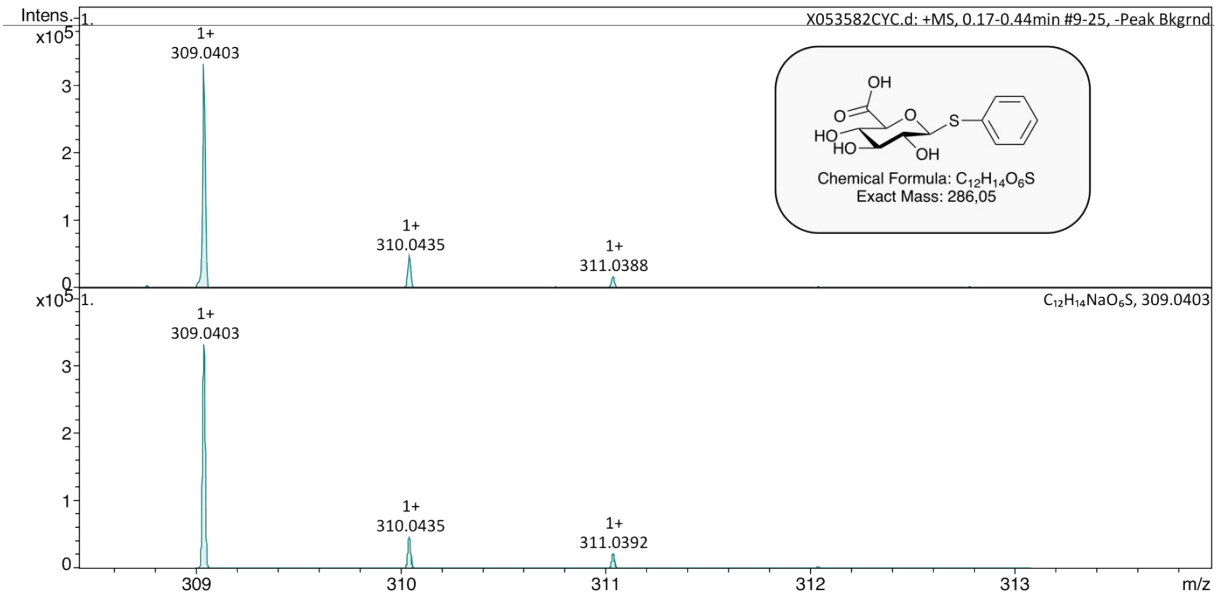
Analysis Info

Sample Name	**R4_F3**	Acquisition Date	04/12/2019 12:30:49
Analysis Name	X053582CYC.d	Instrument / Ser#	maXis 255552.00086
		Method	positif.m

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.6 Bar
Scan Begin	50 m/z	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan End	2500 m/z	Set Collision Cell RF	1800.0 Vpp	Set Dry Gas	7.0 l/min





Meas. m/z	z	#	Ion Formula	m/z	err [ppm]	mSigma	rdb	e⁻ Conf
309.040329	1+	1	C12H14NaO6S	309.040330	0.0	6.9	6.0	even
331.022105	1+	1	C12H13Na2O6S	331.022274	0.5	7.7	6.0	even

Fédération de Recherche Physique et Chimie du Vivant (FR2708 : CBM/ICOA)
Plate-forme de Spectrométrie de Masse Haute Résolution

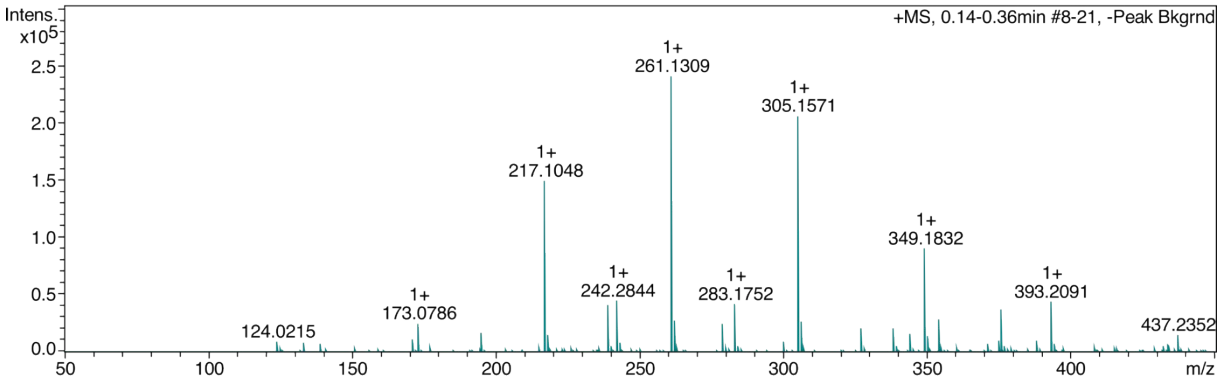
HRAM

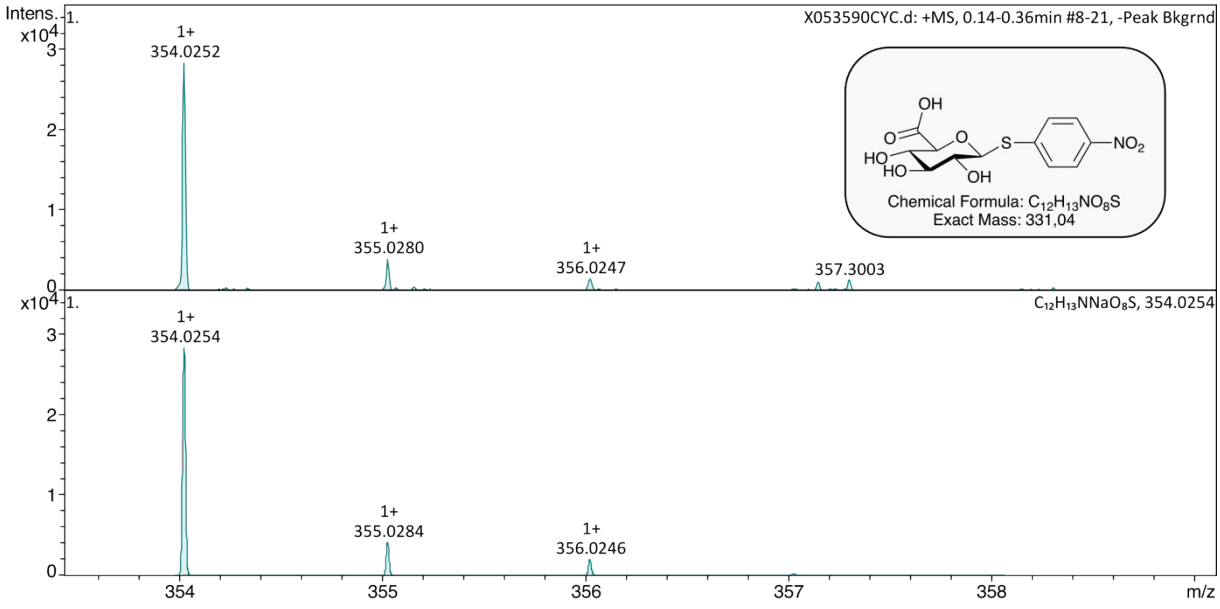
Analysis Info

Sample Name	**R6**	Acquisition Date	04/12/2019 12:36:41
Analysis Name	X053590CYC.d	Instrument / Ser#	maXis 255552.00086
		Method	positif.m

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.6 Bar
Scan Begin	50 m/z	Set Capillary	4500 V	Set Dry Heater	200 °C
Scan End	2500 m/z	Set Collision Cell RF	1800.0 Vpp	Set Dry Gas	7.0 l/min





Meas. m/z	z	#	Ion Formula	m/z	err [ppm]	mSigma	rdb	e⁻ Conf
354.025190	1+	1	C12H13NNaO8S	354.025408	0.6	10.8	7.0	even
376.007044	1+	1	C12H12NNa2O8S	376.007352	0.8	6.6	7.0	even