Electronic Supplementary Material (ESI) for Chemical Communications.

Supplementary Information for

Preparation of iminosugars from aminopolyols via selective oxidation using galactose oxidase

Kathryn Yeow[1], Marianne B. Haarr[1*], Jimmy Muldoon[1], and Elaine O'Reilly[1*]

[1]School of Chemistry, University College Dublin, Belfield, Dublin 4, Ireland.

*Corresponding author. Email: elaine.oreilly@ucd.ie, marianne.haarr@ucd.ie

Table of Contents

General

NMR analysis: NMR analysis was performed using Varian VnmrS 400, 500 and 600 MHz, and Agilent DD2 500 MHz instruments operated by the School of Chemistry NMR Centre at University College Dublin. NMR spectra were processed using MestReNova 14.0 software (Mestrelab Research SL). Deuterated methanol (δ 3.31 ppm / δ 49.00 ppm) and deuterated water (ref. δ 4.79 ppm) were used as solvents and served as internal references. All δ values are reported in ppm. All *J* values are reported in Hz. Data for ^{1}H NMR are reported as follows: chemical shift (δ/ ppm) (multiplicity, coupling constant (Hz), integration, identity). Multiplicities are reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. Data for ^{13}C NMR are reported in terms of chemical shift (δ/ ppm). COSY, HSQC and HMBC NMR experiments were used in the structural assignment of the aminopolyols.

HRMS analysis: Accurate mass data was obtained using the Agilent 1260 Infinity Prime II LC coupled to Agilent 6546 Quadrupole Time-Of-Flight (QTOF) MS system operated by the School of Chemistry Mass Spectrometry Facility at University College Dublin.

UV-Vis spectroscopic analysis: OD_{600} measurements and protein concentration were determined using the Multiskan Sky spectrophotometer in disposable cuvettes with a diameter of 10 mm (ThermoFisher Scientific). DNA concentration was measured on the µDrop™ plate (Thermo Scientific) with elution buffer as the blank.

I. Chemicals

Chemicals and solvents were purchased from Sigma Aldrich, Fisher Scientific or Fluorochem and used without further purification. TLC analyses were performed on Merck silica gel 60 F254 plates using $KMnO_4$ for visualization. Silica gel NORMASIL 60® 40-63 µm was used for flash column chromatography. The synthesis of *N*-Cbz-protected aminopolyols was carried out according to methods described in our previous work.[1]

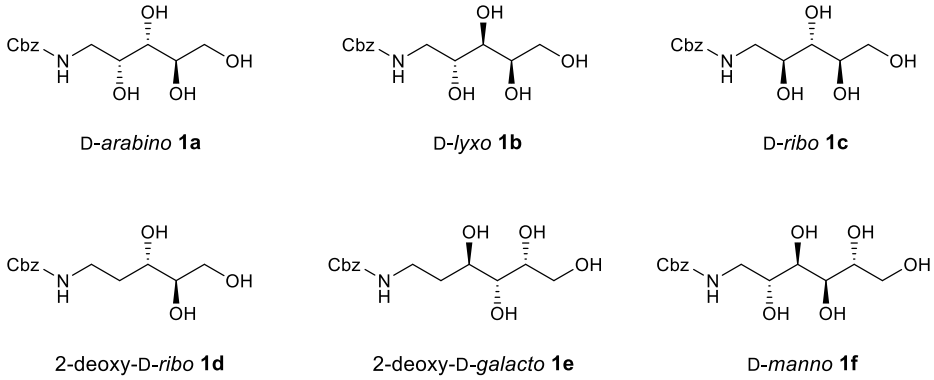


Figure S1. Chemically synthesized *N*-Cbz protected aminopolyols **1a-f** screened in this work.

II. Enzymes

The expression construct for GOase variant F_2 in pET-30b was provided by Prof. Sabine Flitsch (Manchester Institute of Biotechnology, UK).

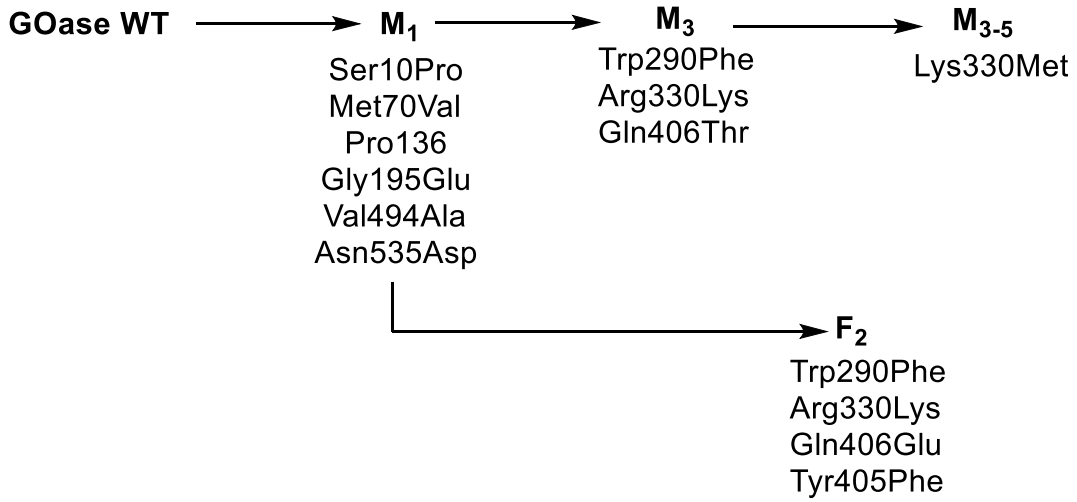


Figure S2. An overview of the key mutations in the evolution of GOase.

GOase F_2 sequence

```
ASAPIGSAIPRNNWAVTCDSAQSGNECNKAIDGNKDTFWHTFYGANGDPKPPHTYTIDMKTTQNVNGL
SVLPRQDGNQNGWIGRHEVYLSSDGTNWGSPVASGSWFADSTTKYSNFETRPARYVRLVAITEANGQP
WTSIAEINVFQASSYTAPQPGLGRWGPTIDLPIVPAAAAIEPTSGRVLMWSSYRNDAFEGSPGGITLT
SSWDPSTGIVSDRTVTVTKHDMFCPGISMDGNGQIVVTGGNDAKKTSLYDSSSDSWIPGPDMQVARGY
QSSATMSDGRVFTIGGSFSGGVFEKNGEVYSPSSKTWTSLPNAKVNPMLTADKQGLYKSDNHAWLFGW
KKGSVFQAGPSTAMNWYYTSGSGDVKSAGKRQSNRGVAPDAMCGNAVMYDAVKGKILTFGGSPDFEDS
DATTNAHIITLGEPGTSPNTVFASNGLYFARTFHTSVVLPDGSTFITGGQRRGIPFEDSTPVFTPEIY
VPEQDTFYKQNPNSIVRAYHSISLLLPDGRVFNGGGGLCGDCTTNHFDAQIFTPNYLYDSNGNLATRP
KITRTSTQSVKVGGRITISTDSSISKASLIRYGTATHTVNTDQRRIPLTLTNNGGNSYSFQVPSDSGV
ALPGYWMLFVMNSAGVPSVASTIRVTQ
```

Protein expression and purification

Table 1. Buffers used for protein production.

Buffers	
Resuspension	NaPi (50 mM), NaCl (300 mM), pH 8
Elution	NaPi (50 mM), NaCl (300 mM), desthiobiotin (2.5 mM), pH 8
Reaction	NaPi (100 mM), pH 8

Transformation To a 1.5 mL Eppendorf tube containing chemocompetent BL21(DE3) cells (50 µL) thawed on ice, plasmid (4 µL) is added and the contents is gently mixed by flicking. Cells are incubated on ice for 30 minutes, and then heat shock (42°C, 30-45 s), then incubated for 2 minutes. LB medium (250 µL) is added, and the cells are incubated (37°C, 1 h). An aliquot of culture (150 µL) is spread onto LB-Kanamycin plates and incubated overnight (37°C).

Expression An overnight culture is prepared by picking a single colony transformant and inoculating LB medium (5 mL) supplemented with kanamycin (50 µg mL^{-1}, 5 µL). The overnight culture is used to inoculate LB/K medium (culture volumes between 3-5 litres) and the cultures are grown for 2-3 hours at 37°C with constant shaking at 225 rpm. Expression is induced by the addition of IPTG (0.5 mM) at OD_{600} 0.6. The cultures

are grown at 20°C overnight with constant shaking at 225 rpm. Cells are harvested by centrifugation (4°C, 4000 rpm, 40 minutes) and the cell pellets are resuspended in the binding buffer (in 20% volume of the shake culture), and sonicated (35% amplitude) for 15 minutes (10 s on, 15 s off) or until a reduction in viscosity and colour change is observed. The cell free extract is centrifuged at 4°C, 4000 rpm, 30 minutes followed by 9500 rpm for 30 minutes and clarified using 0.2 or 0.45 µm syringe filters.

Protein Purification The clarified cell-free lysate is applied to *Strep*-Tactin® Superflow Plus cartridge (5 mL) (Qiagen). The column is washed with binding buffer *via* peristaltic pump P-1 (GE Healthcare) until protein concentration in the flow-through reaches < 0.001 mg mL^{-1} according to the spectrophotometer. Binding buffer containing desthiobiotin (2.5 mM) is used for elution of Strep-tagged protein. Fractions (1.5 mL) are collected in cuvettes and evaluated on the spectrophotometer. The fractions containing protein concentration > 0.1 mg mL^{-1} are pooled and transferred to a Vivaspin® 6 centrifugal spin column (20 mL, 10 kDa MWCO). The protein is dialyzed at room temperature against the desired storage buffer (rt, 4000 rpm, t=varies), concentrated to >1 mg mL^{-1} and flash frozen in liquid nitrogen for storage at -80°C until further use. Yields of approximately 12 mg L^{-1} of culture were obtained.[a]

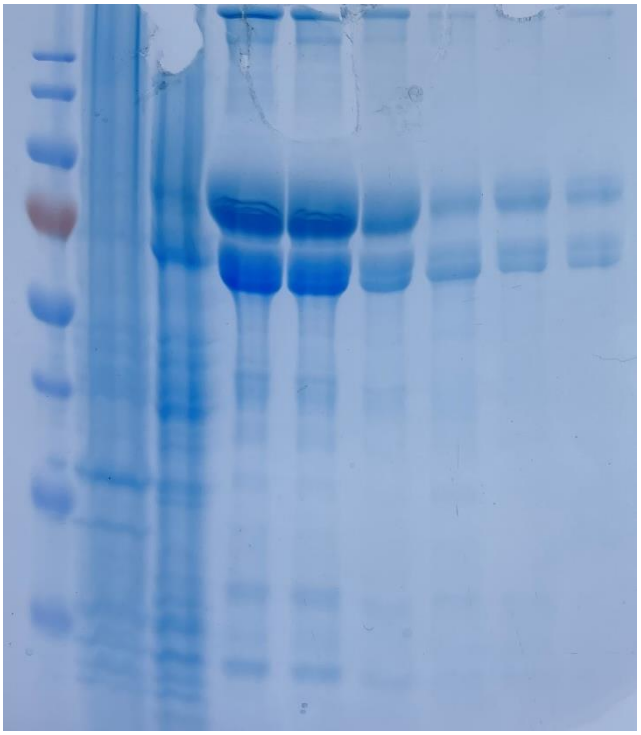


Figure S3. GOase F_2. L→R: EZ Fisher Bioreagents Ladder (**1**), pET-28(a) lysate (**2**), F_2 lysate (**3**), F_2 purified fractions (**4-9**).

[a]Some protein degradation is evident in the purified fractions from the bands corresponding to a lower Mw.

III. Biocatalytic oxidations with GOase F_2

Analytical scale biotransformations

Substrate stocks of **1a-f** (50-200 mM) were prepared in reaction buffer depending on substrate solubility. In a 2 mL Eppendorf tube, substrate (20 mM), HRP Type I (0.1 mg mL^{-1}), catalase (0.1 mg mL^{-1}) and purified GOase F_2 (7.25 µM, 0.5 mg mL^{-1}) were added to reaction buffer (50 mM, pH 8). The 1 mL reaction mix was incubated at 25°C in an orbital shaker at 250 rpm. Reaction progress was monitored by thin layer chromatography (TLC) (Figure S4).

Figure S4. GOase F_2 Substrate screening using thin layer chromatography (TLC, silica). MeCN/0.1% formic acid 99:1. In all cases, the oxidised products **2a-f** appear as the top spot, with higher Rf value than the starting materials **1a-f**.

Preparative scale biotransformations

To a 50 mL round-bottom flask containing substrate[b] (0.2 mmol, 20 mM) in NaPi buffer (100 mM, pH 8), $CuSO_4$ (50 µM, 100 µL), HRP Type I (10.9 U mL^{-1}, 175 µL), catalase (0.05 mg mL^{-1}, 50 µL) and purified GOase F_2 (0.25 mg mL^{-1}) was added to a total volume of 10 mL. The reaction was stirred at ambient temperature for the appropriate amount of time and monitored by TLC analysis. Additional enzyme and auxiliary enzymes are added over the course of the reaction to push completion and maximise yield for full characterisation and further analysis. The reaction was quenched by flash freezing in liquid nitrogen, the mixture was lyophilised and reconstituted in methanol. The majority of precipitated phosphate buffer salts and proteinaceous matter were filtered through cotton wool, and the filtrate was concentrated. The residue was purified by flash column chromatography (silica, acetonitrile: water with 0.1% formic acid 95:5 v/v).

[b] Due to the changing reaction volume, 20 mM is given as the starting concentration which subsequently decreases upon addition of enzyme solutions.

IV. Time course analysis and quantification of reaction conversion

Analytical scale biotransformations and sample preparation

Substrate stocks of **1a-f** (50-200 mM) were prepared in reaction buffer depending on substrate solubility. In a 2 mL Eppendorf tube, substrate (20 mM), HRP Type I (0.1 mg mL^{-1}), catalase (0.1 mg mL^{-1}), $CuSO_4$ (50 µM) and purified GOase F_2 (2 mg mL^{-1}) were added to reaction buffer (100 mM, pH 8). The 1 mL reaction mix was incubated at 20°C in an orbital shaker at 250 rpm, reactions are performed in triplicate. At t = 1, 2, 4, 6, 8, 24 and 48 hours, samples (100 µL) were taken and made up to 1 mL by adding LC/MS grade MeOH (900 µL). The solution was filtered through a syringe filter into a vial for analysis. In some cases, a further ten-fold dilution was carried out to reach quantifiable amounts of sample.

Quantification

Stock solutions were prepared by dissolving standards of **2a-f** in LC/MS grade methanol at concentrations of 100 mM and stored at -20°C. Working standard solutions (5–300 µM) were prepared by diluting the stock solutions in LC/MS grade methanol. Then 1 µL of standard solutions were subjected to HILIC–MS, and standard curves were constructed.

LC/MS parameters

The LC experiments were carried out using a Waters XBridge Amide column (3.5 µm, BEH Amide, 100 Å, 4.6 × 250 mm). Injection volume was 1 µL and column temperature was maintained at 35°C. Solvent A = deionised water (MilliQ) containing 0.1% formic acid pH 2; solvent B = LC/MS grade acetonitrile (Optima, Fisherbrand). The column was eluted with an isocratic method consisting of 40% (solvent A) and 60% (solvent B). The flow rate was set at 0.5 mL min^{-1} and the run time was 7 minutes.

Analytical LC and nominal mass data was obtained with using the Agilent 1260 Infinity II LC/6125 XT MSD system. The Agilent 1260 Infinity II vial sampler (G7129A) and Agilent 1260 Infinity II Binary pump system (G7112B) is coupled *via* an electrospray ionization (API-ES) interface working under positive polarity to a single quadrupole LC/MSD XT (G6125B) mass spectrometer. The parameters were set as follows: Gas temp, 350 °C, drying gas 12 L min^{-1}, nebuliser pressure, 35 psig, quad temperature, 100°C, Capillary voltage (Positive), 3000 V and fragmentor voltage 110 V. The molecular ion adducts for each species were first determined under the specified analytical method and MS parameters. The most prevalent adduct was set as the target mass to be recorded in selected ion monitoring (SIM) mode. The SIM parameters were set as follows: Fragmentor 110 V, Gain EMV 1.0, Actual Dwell 590. Data acquisition and processing were performed using Agilent MassHunter software.

Standard curves

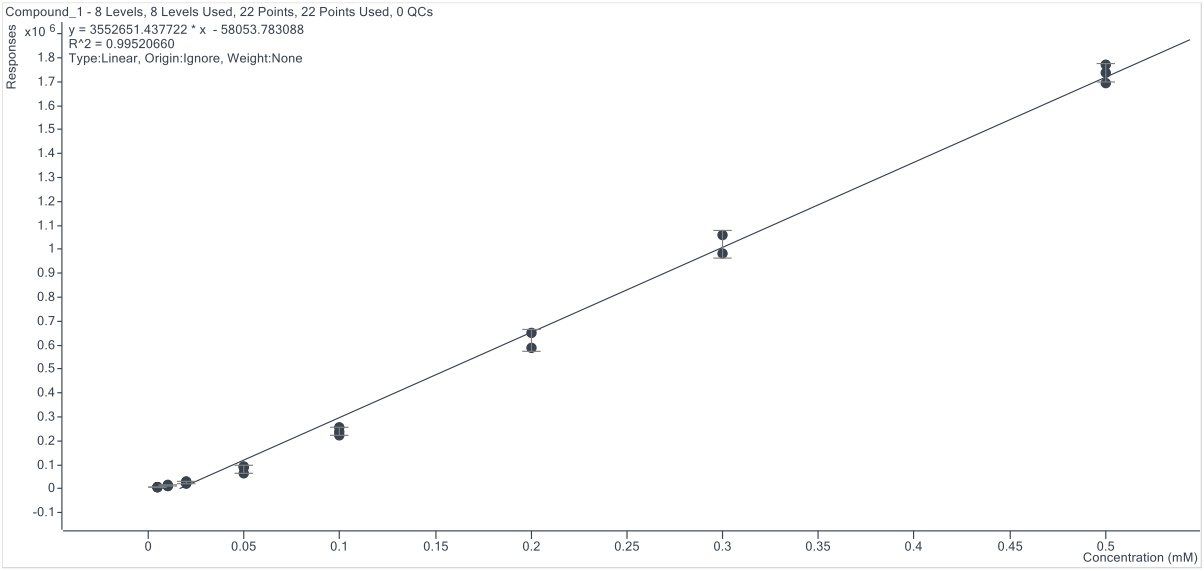


Figure S5. Calibration curve for compound **2a** over eight levels (0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, 0.5 mM) with linearity $R^2 = 0.995$.

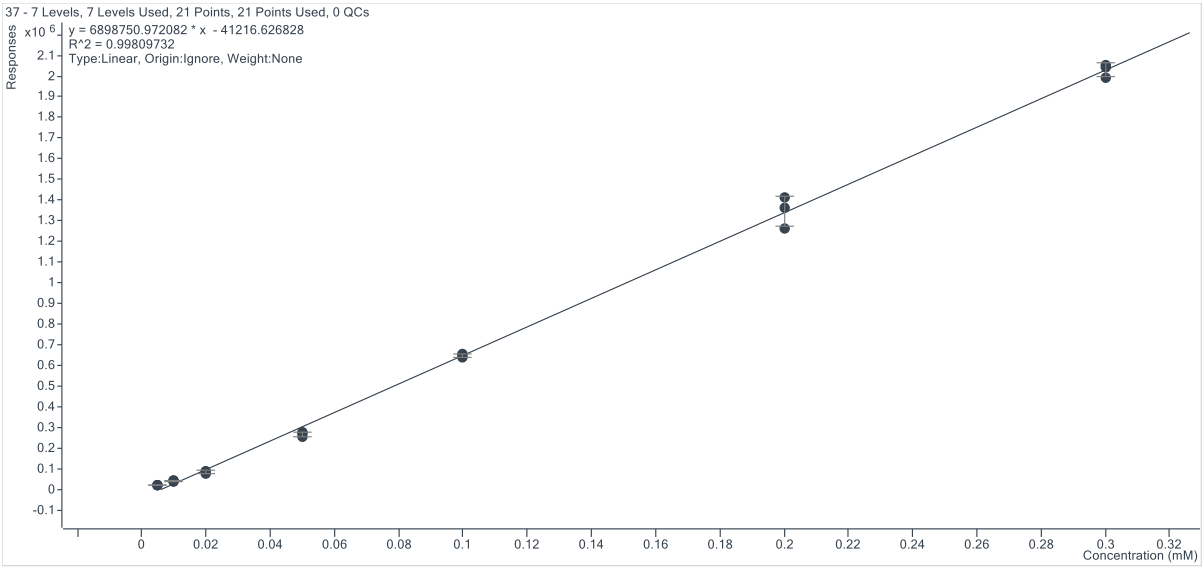


Figure S6. Calibration curve for compound **2b** over seven levels (0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3 mM) with linearity $R^2 = 0.998$.

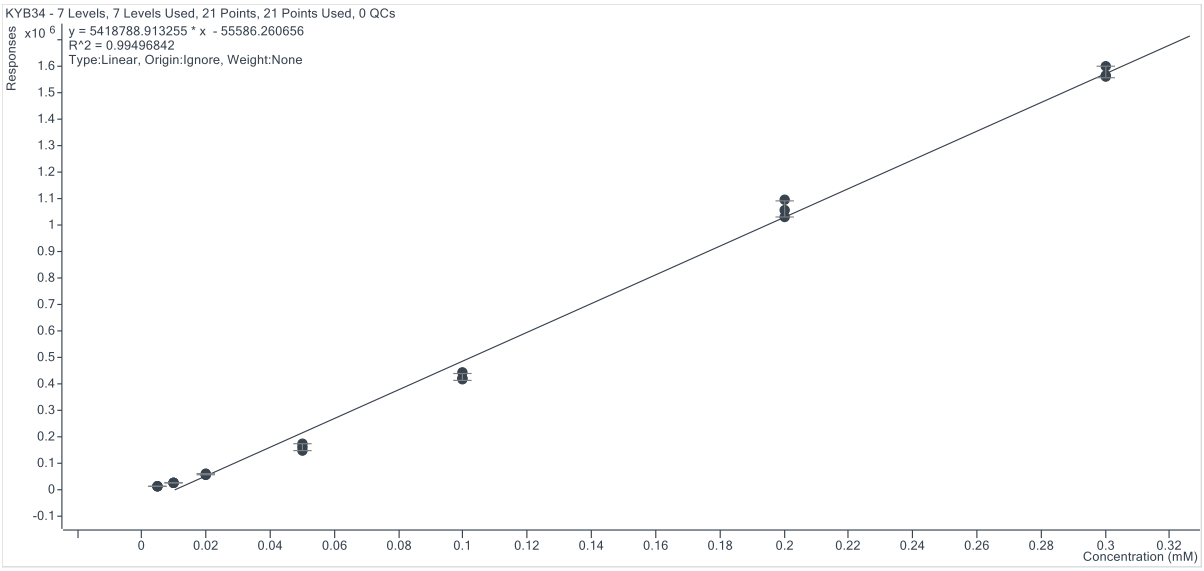


Figure S7. Calibration curve for compound **2c** over seven levels (0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3 mM) with linearity R^2 = 0.995.

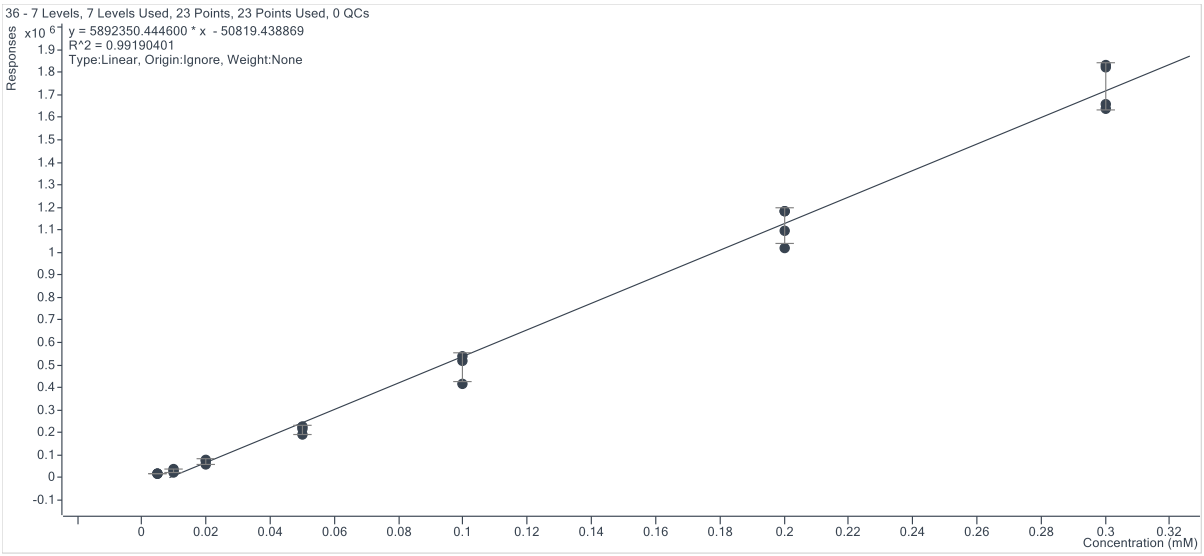


Figure S8. Calibration curve for compound **2d** over seven levels (0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3 mM) with linearity R^2 = 0.991.

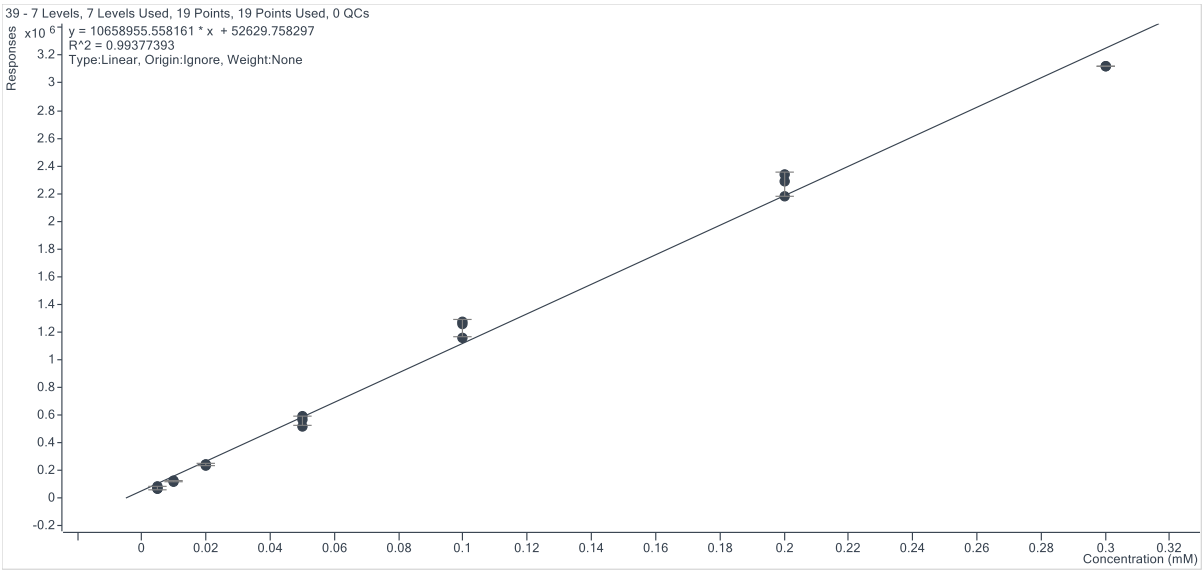


Figure S9. Calibration curve for compound **2e** over seven levels (0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3 mM) with linearity R^2 = 0.994.

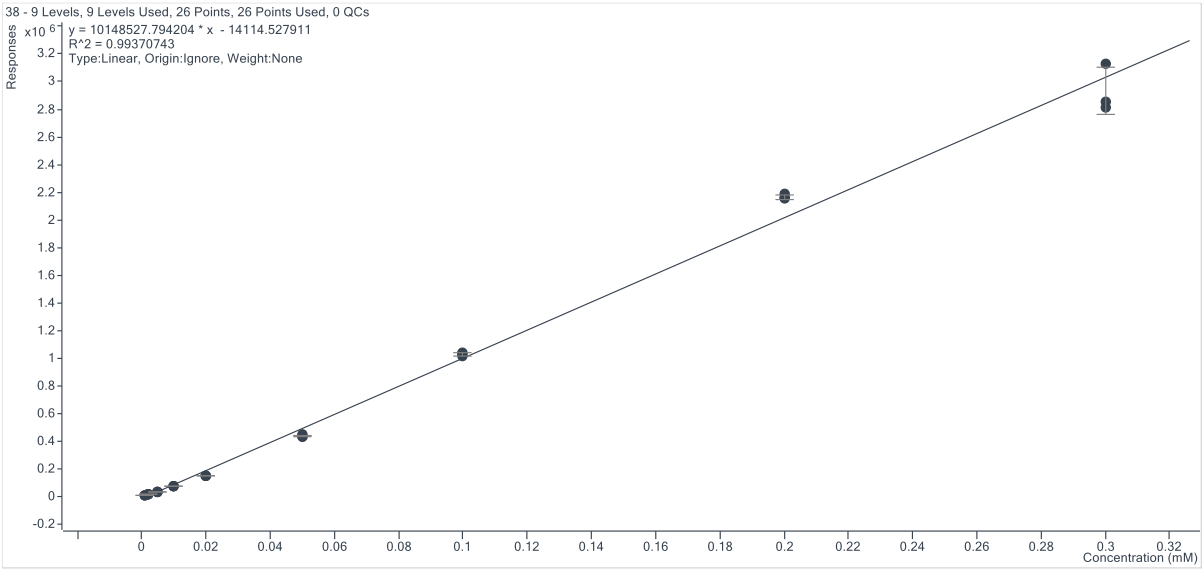


Figure S10. Calibration curve for compound **2f** over nine levels (0.0005, 0.001, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.3 mM) with linearity R^2 = 0.994.

Individual conversion data

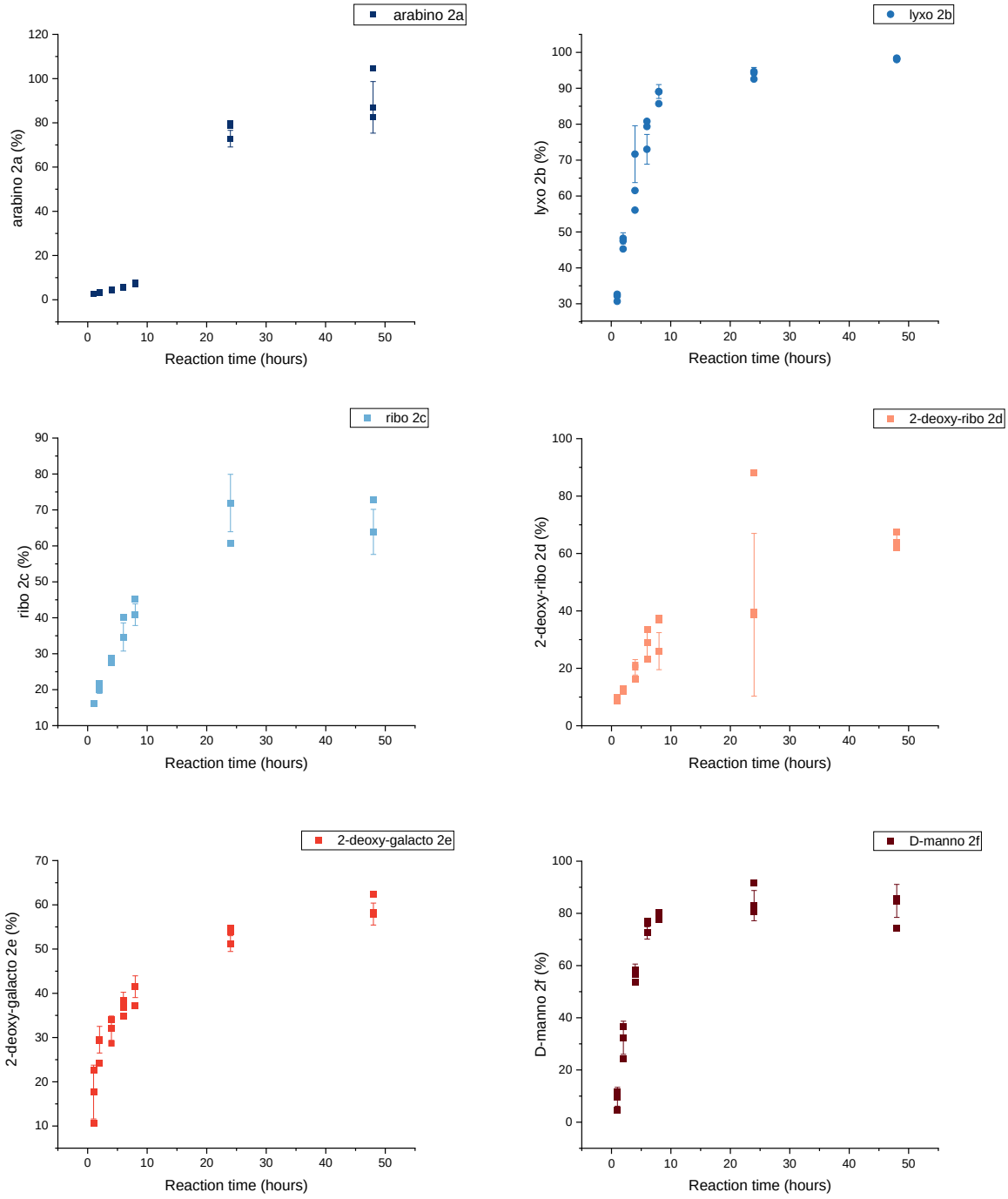
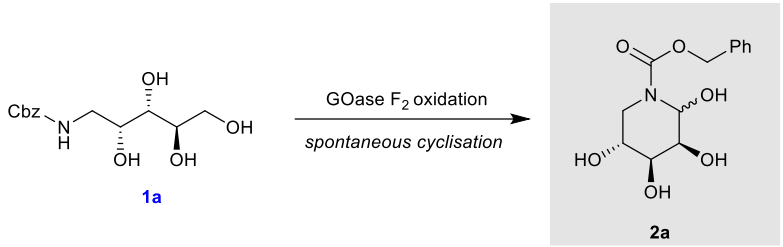


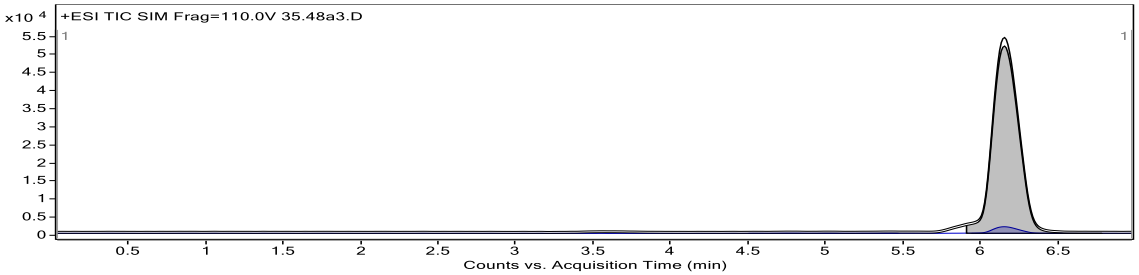
Figure S11. Reaction conversion to oxidised products **2a-f** over a 48-hour period, supplementary to Figure 1 in the main text. Reaction conditions: substrate (0.2 mmol, 20 mM), catalase (0.1 mg mL^{-1}), HRP Type I (0.1 mg mL^{-1}), $CuSO_4$ (50 µM), NaPi (100 mM, pH 8), purified GOase F2 (2 mg mL^{-1}) at 20°C. Reactions are carried out in triplicate. Conversion was determined by LC/MS analysis, monitoring product formation at timepoints in SIM mode. Conversions plotted in triplicate with standard deviation. Conversion to compound **2a** was calculated based on an altered substrate concentration of 23 mM. The outlier of **2d** at 24 hours is excluded in the calculation.

LC/MS traces for biocatalytic reactions

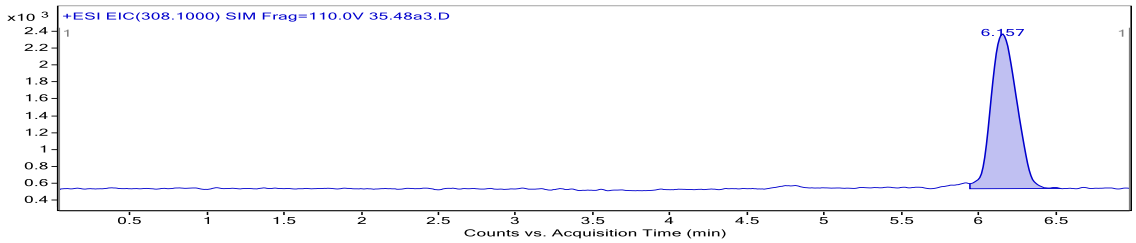


t=48 h

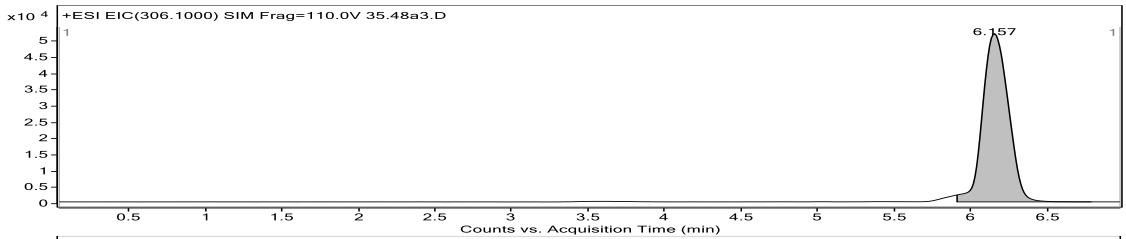
A Reaction



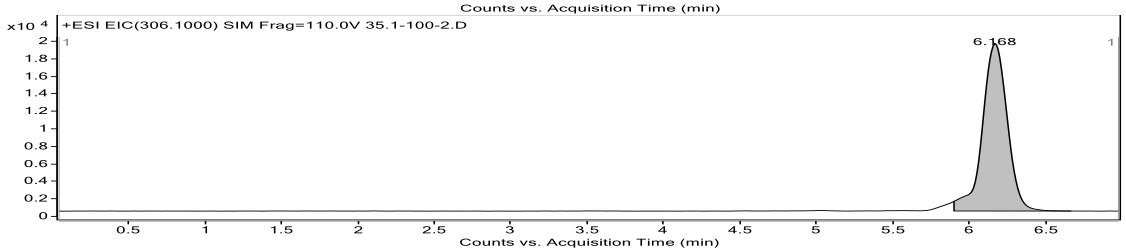
B Starting material **1a**



C Product **2a**



D Product standard **2a**



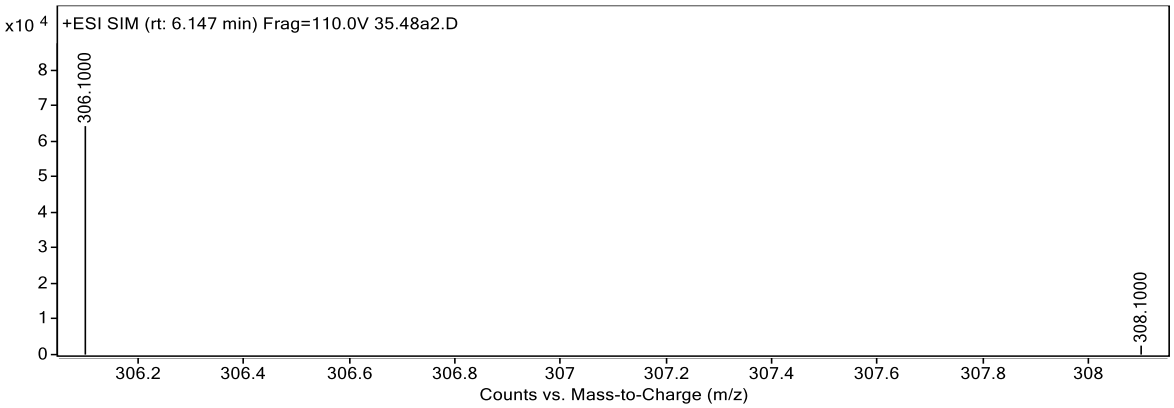
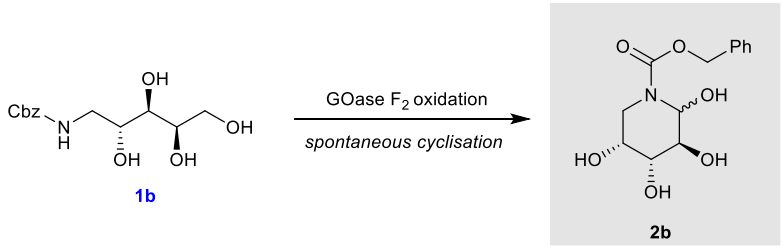


Figure S12. Oxidation of *D*-arabino substrate **1a** by GOase F_2. **A-D** LC traces of the reaction mixture after 48 hours. MS detection carried out in SIM mode, monitoring the formation of product **2a** ($[M+Na]^+$ m/z = 306.1) and disappearance of starting material **1a** ($[M+Na]^+$ m/z = 308.1).



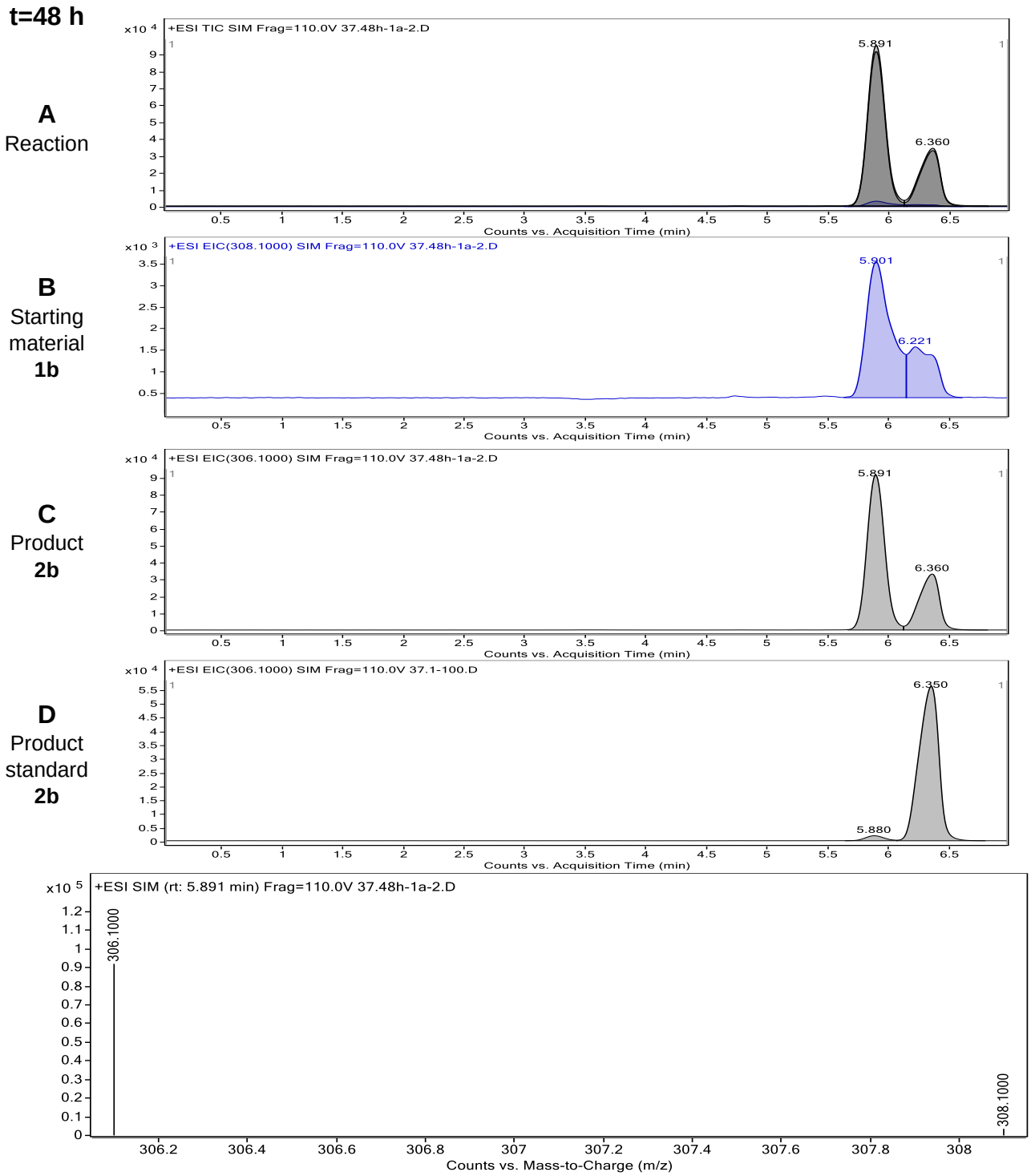


Figure S13. Oxidation of *D*-lyxo substrate **1b** by GOase F_2. **A-D** LC traces of the reaction mixture after 48 hours. MS detection carried out in SIM mode, monitoring the formation of product **2b** ($[M+Na]^+$ m/z = 306.1) and disappearance of starting material **1b** ($[M+Na]^+$ m/z = 308.1). **C-D** The product and product standard consist of an unequal ratio of anomers resolved by HPLC (rt = 5.89 min and rt= 6.36 min), the sum of integrated peaks is used for the calculation.

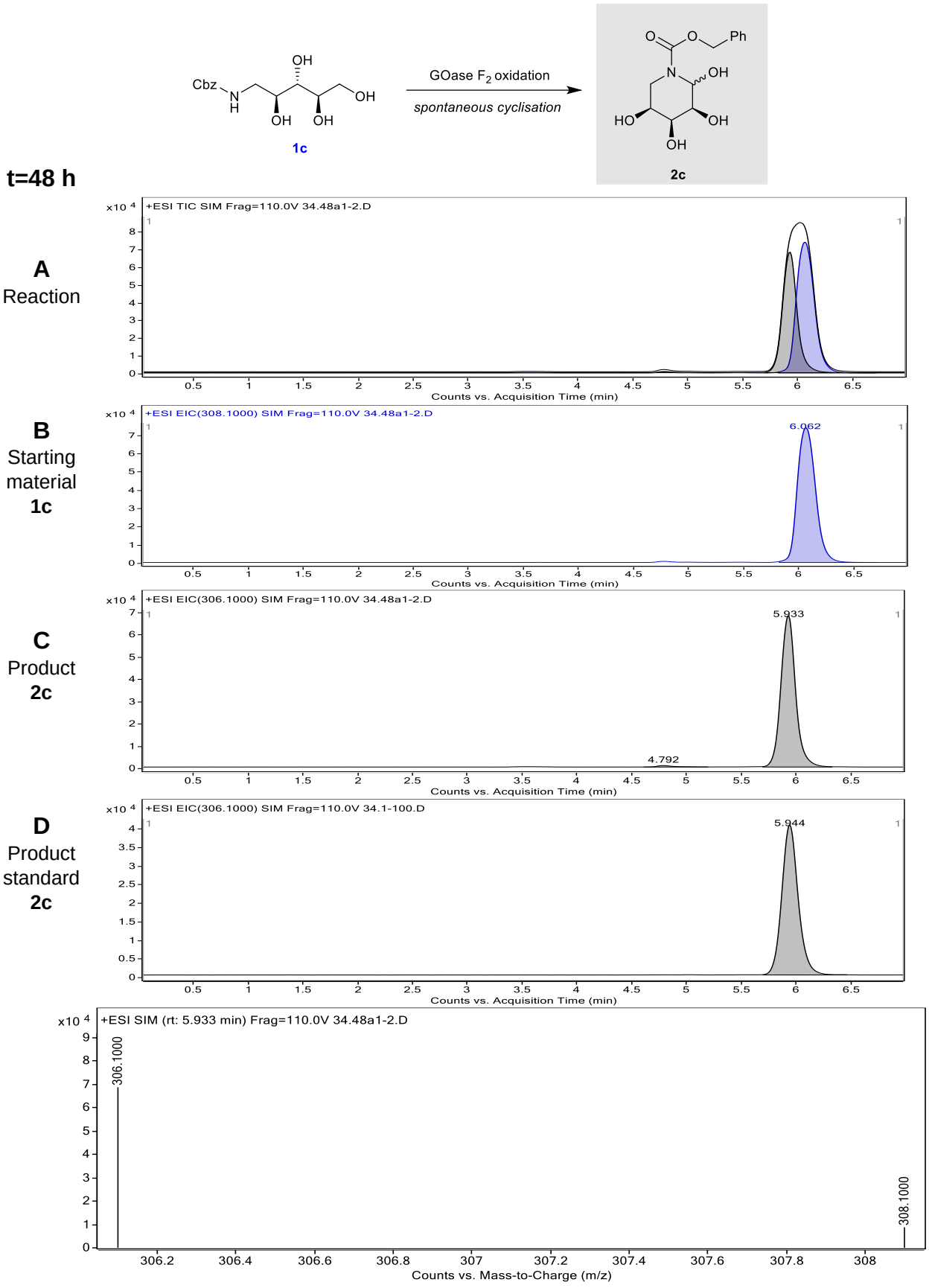


Figure S14. Oxidation of *D*-ribo substrate **1c** by GOase F_2. **A-D** LC traces of the reaction mixture after 48 hours. MS detection carried out in SIM mode, monitoring the formation of product **2c** ($[M+Na]^+$ m/z = 306.1) and disappearance of starting material **1c** ($[M+Na]^+$ m/z = 308.1).

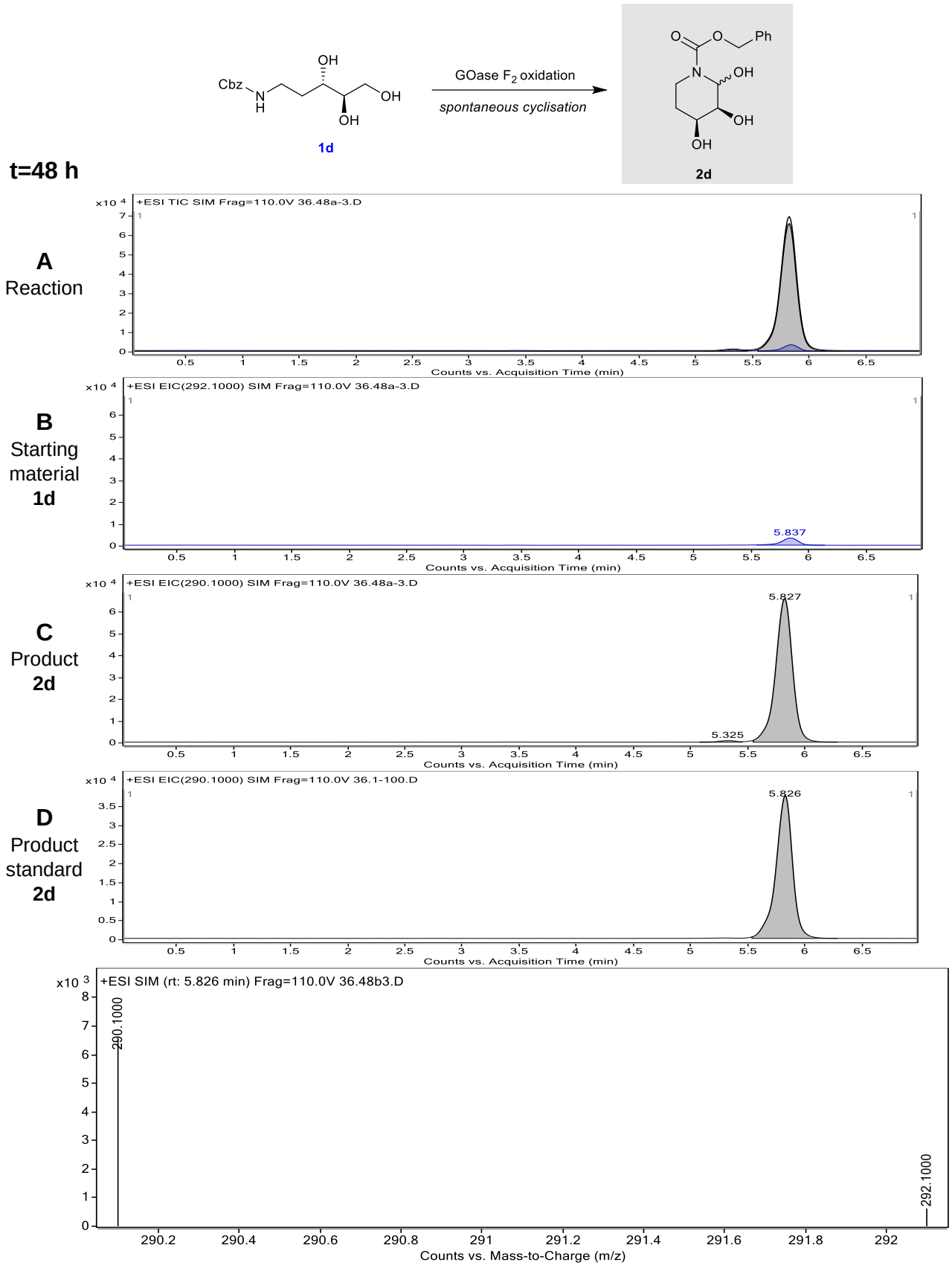


Figure S15. Oxidation of 2-deoxy-*D*-ribo substrate **1d** by GOase F_2. **A-D** LC traces of the reaction mixture after 48 hours. MS detection carried out in SIM mode, monitoring the formation of product **2d** ($[M+Na]^+$ m/z = 290.1) and disappearance of starting material **1d** ($[M+Na]^+$ m/z = 292.1).

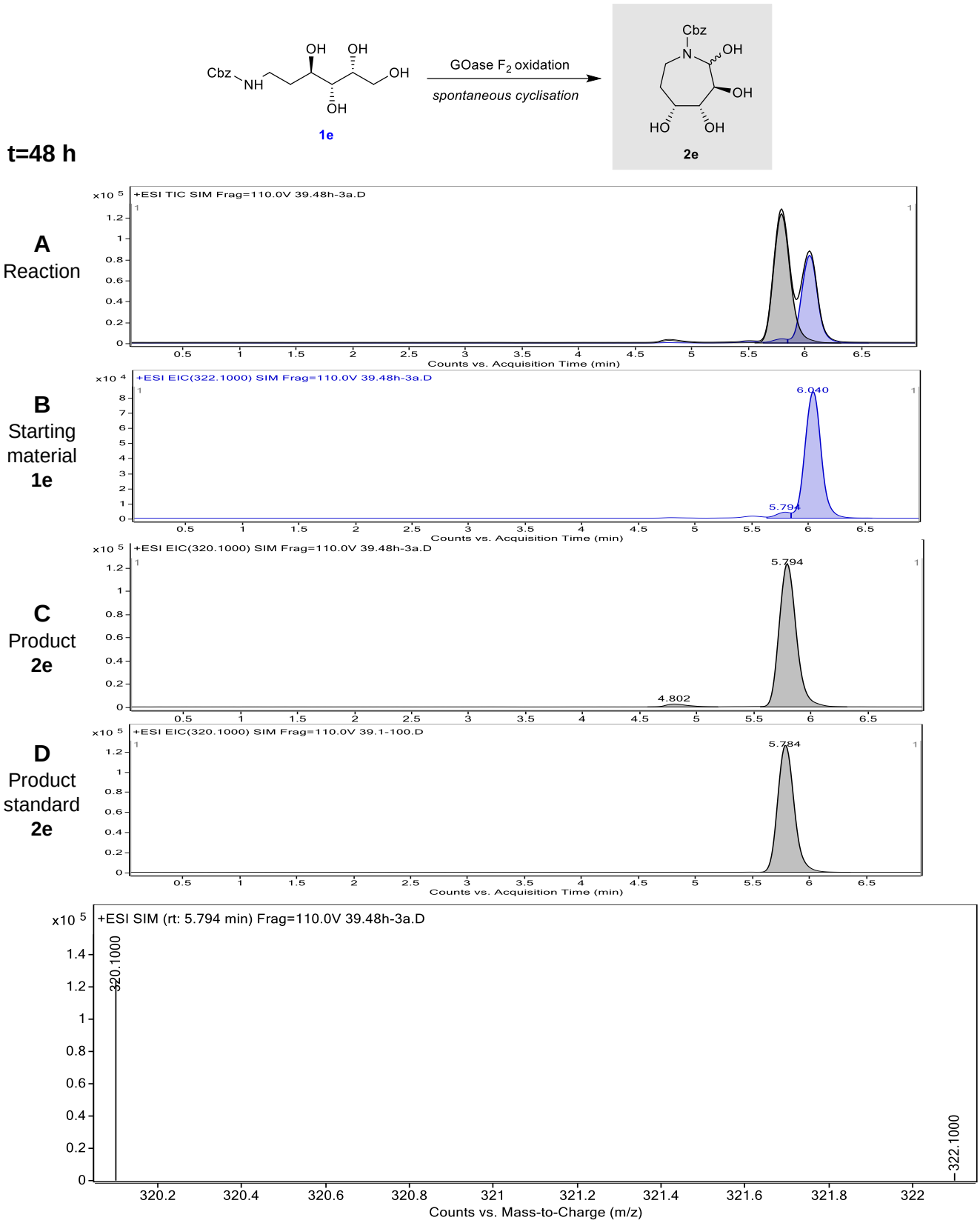
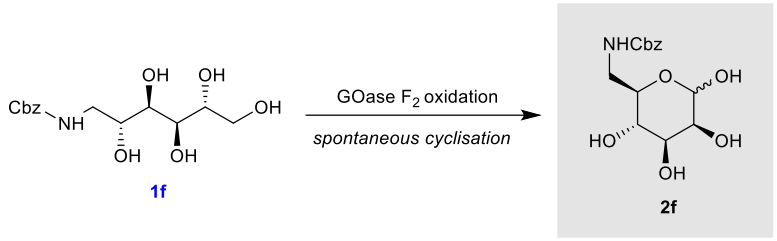
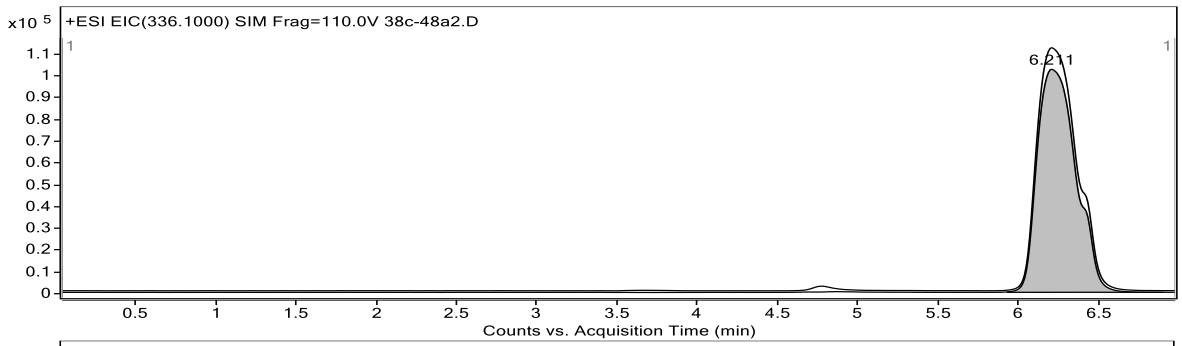


Figure S16. Oxidation of 2-deoxy-*D*-galacto substrate **1e** by GOase F_2. **A-D** LC traces of the reaction mixture after 48 hours. MS detection carried out in SIM mode, monitoring the formation of product **2e** ($[M+Na]^+$ m/z = 320.1) and disappearance of starting material **1e** ($[M+Na]^+$ m/z = 322.1).

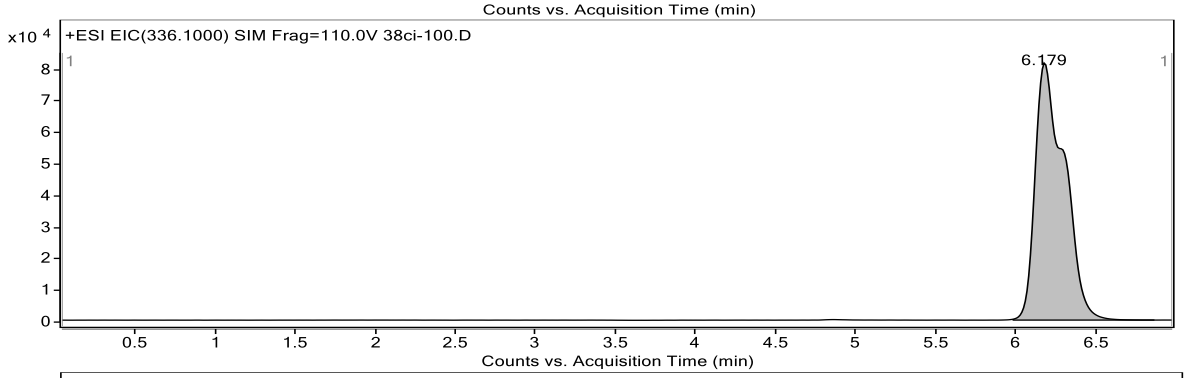


t=48 h

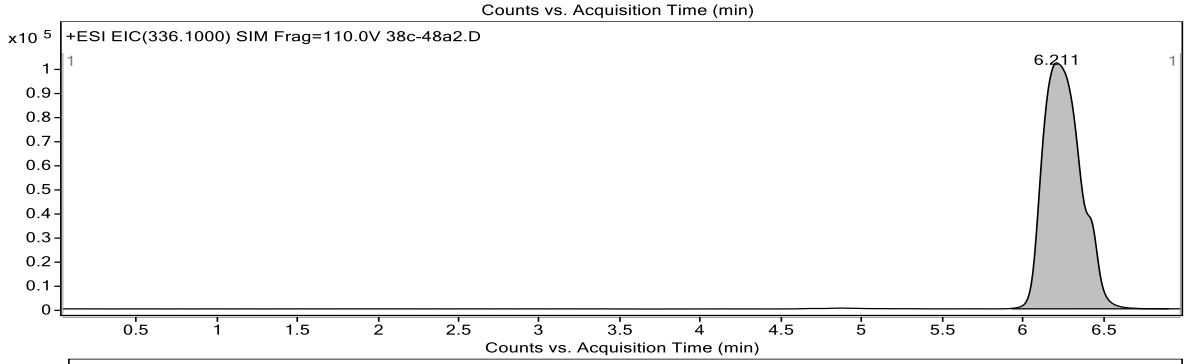
A Reaction



B Product **2f**



C Product standard **2f**



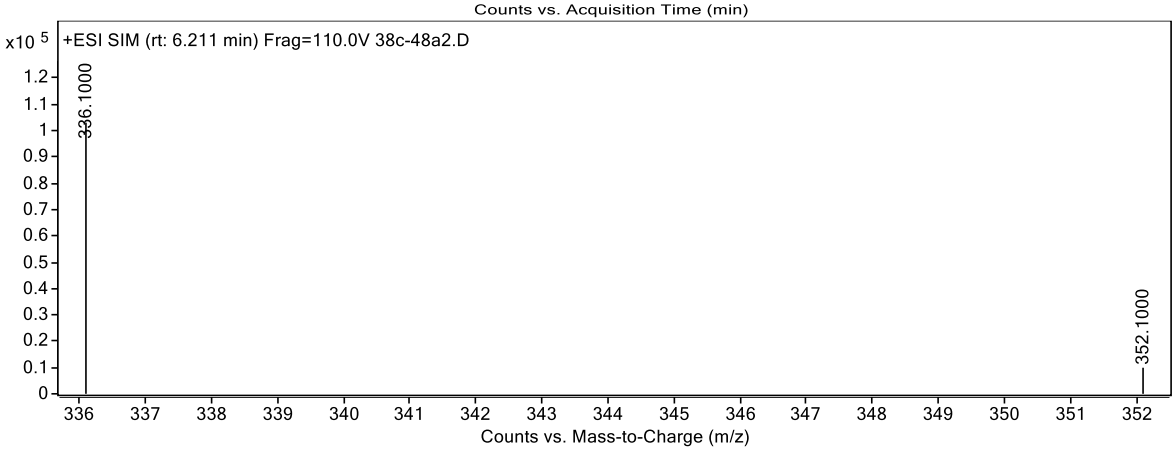


Figure S17. Oxidation of *D*-manno substrate **1f** by GOase F_2. **A-C** LC traces of the reaction mixture after 48 hours. MS detection carried out in SIM mode, monitoring the formation of product **2f** ($[M+Na]^+$ m/z = 336.1). Disappearance of starting material was not recorded in this case. Overlapping peaks correspond to the presence of anomeric mixtures, an unequal ratio of anomers is observed between the product and the product standards.

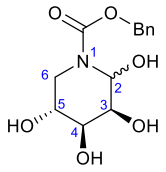
V. Chemical deprotection and reduction

General method for hydrogenation

A suspension of Pd/C (1 equiv.), acetic acid (1 drop), and the aldehyde **2a**-**2f** (20 mg) in MeOH (2 mL) was hydrogenated (1 atm) at room temperature for 16 h. The reaction mixture was filtered through Celite with the aid of MeOH. The filtrate was then cooled to 0°C and added 0.5 M methanolic HCl (2 mL). After stirring for 5 min volatiles were removed under reduced pressure to afford the iminosugar **3a**-**3f** as the HCl salt.

VI. Characterisation of compounds

GOase F_2-catalysed oxidation products



2a

benzyl (3*S*,4*S*,5*R*)-2,3,4,5-tetrahydroxypiperidine-1-carboxylate **2a**[2]

D-*arabino*

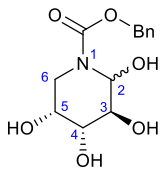
Isolated yield: 44 mg, 78%, white crystalline solid

Ratio anomers **2a'**/**2a''** ~ 9:1

^{1}H NMR (600 MHz) CD_3OD, δ = 7.39-7.30 (m, 5H, ArH), 5.70 (d, *J* = 3.5 Hz, 1H, H-2, **2a''**) 5.68 (d, *J* = 2.1 Hz, 1H, H-2, **2a'**), 5.17 (d, *J* = 6.7 Hz, 2H, CH_2, **2a''**), 5.14 (d, *J* = 3.3 Hz, 2H, CH_2, **2a'**), 4.06 (dd, *J* = 4.9, 12.4 Hz, 1H, H-6a, **2a'**), 3.89 (m, 1H, H-3, **2a'**, H-4 + H-5, **2a''**), 3.87 (m, 1H, H-6a, **2a''**), 3.80-3.68 (m, 2H, H-4 + H-5, **2a**, H-3, **2a''**) 3.52-3.49 (m, 1H, H-6b, **2a''**), 3.01-2.92 (m, 1H, H-6b, **2a'**).

^{13}C NMR (150 MHz) CD_3OD, δ = 157.4 (CO), 137.9 (ArC), 129.5-128.9 (ArCH), 80.3 (C-2), 73.3 (C-5, **2a''**), 72.9 (C-4), 72.7 (C-3), 70.3 (C-4, **2a''**), 68.5 (CH_2), 67.9 (C-5), 66.5 (C-3, **2a''**) 44.2 (C-6, **2a'**), 40.3 (C-6, **2a''**).

HRMS (ESI): *m/z* $[M+Na]^+$ calcd. For $C_{13}H_{17}NO_6$: 306.0948; found 306.0947.



2b

benzyl (3*S*,4*R*,5*R*)-2,3,4,5-tetrahydroxypiperidine-1-carboxylate **2b**[3]

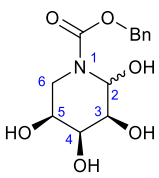
D-*lyxo*

Isolated yield: 10 mg, 50%, off-white solid

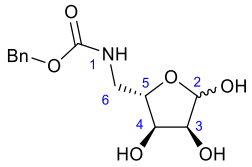
^{1}H NMR (400 MHz) CD_3OD, δ = 7.40-7.30 (m, 5H, ArH), 5.76 (d, *J* = 3.2 Hz, 1H, H-2), 5.14 (q, *J* = 12.4 Hz, 2H, CH_2Ph), 4.01 (dd, *J* = 13.9, 1.9 Hz, 1H, H-6a), 3.95-3.89 (m, 1H, H-5), 3.79 (dd, *J* = 10.0, 3.2 Hz, 1H, H-3), 3.71 (dd, *J* = 10.0, 3.7 Hz, 1H, H-4), 3.35-3.27 (m, 1H, H-6b).

^{13}C NMR (100 MHz) CD_3OD, δ = 161.2 (CO), 129.5 (ArC), 129.1-128.9 (ArC), 78.9 (C-2), 71.1 (C-3), 70.3 (C-4), 69.8 (C-5), 68.4 (CH_2), 44.4 (C-6).

HRMS (ESI): *m/z* $[M+Na]^+$ calcd. For $C_{13}H_{17}NO_6$: 306.0948; found 306.0946.



2c'/2c''



2c'''

benzyl (3*S*,4*S*,5*S*)-2,3,4,5-tetrahydroxypiperidine-1-carboxylate **2c'/2c''**

benzyl (((2*S*,3*R*,4*S*)-3,4,5-trihydroxytetrahydrofuran-2-yl)methyl)carbamate **2c'''**

D-*ribo*

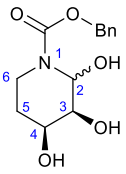
Isolated yield: 53 mg, 94%, yellow oil

Ratio anomers **2c'/2c''/2c'''** ~ 65:25:10

^{1}H NMR (400 MHz) CD_3OD, δ = 7.40-7.29 (m, 5H, ArH), 5.80 (m, 1H, H-2, **2c'**), 5.62 (m, 1H, H-2, **2c''**), 5.11 (m, 1H, H-2, **2c'''**), 5.20-5.08 (m, 2H, CH_2), 4.16 (ddd, *J* = 1.3, 2.8, 13.9 Hz, 1H, H-6a, **2c'**), 4.07 (dd, *J* = 5.0, 6.8 Hz, 1H, H-4, **2c'''**), 4.03 (m, 1H, H-4, **2c''**), 3.96-3.85 (m, 3H, H-5, **2c'**, H-6, **2c''**, H-5, **2c'''**), 3.83 (m, 3H, H-3 + H-4, **2c'** + H-3, **2c'''**), 3.59 (ddd, *J* = 2.6, 5.4, 11.3 Hz, 1H, H-5, **2c''**), 3.47 (dd, *J* = 2.8, 3.6 Hz, 1H, H-3, **2c''**), 3.39-3.32 (m, 3H, H-6b, **2c'**, H-6, **2c'''**), 3.24-3.16 (m, 1H, H-6b, **2c''**).

^{13}C NMR (100 MHz) CD_3OD, δ = 158.9 (CO, **2c'''**), 158.2 (CO, **2c'**), 156.8 (CO, **2c''**), 137.9 (ArC), 137.7 (ArC), 129.6-128.8 (ArCH), 103.0 (C-2, **2c'''**), 82.3 (C-5, **2c'''**), 80.7 (C-2, **2c'**), 79.3 (C-2, **2c''**), 77.2 (C-3, **2c'''**), 74.5 (C-4, **2c''**), 73.7 (C-3, **2c'**), 73.2 (C-4, **2c'''**), 70.9 (C-5, **2c'**), 69.1 (C-3, **2c''**), 68.7 (CH_2, **2c''**), 68.5 (CH_2, **2c'**), 67.5 (CH_2, **2c'''**), 66.7 (C-4, **2c'**), 45.3 (C-6, **2c'**), 44.7 (C-6, **2c'''**), 39.1 (C-6, **2c''**).

HRMS (ESI): *m/z* $[M+Na]^+$ calcd. For $C_{13}H_{17}NO_6$: 306.0948; found 306.0950.



2d

benzyl (3*S*,4*S*)-2,3,4-trihydroxypiperidine-1-carboxylate **2d**[4]

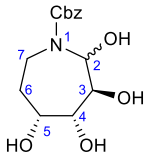
deoxy-D-*ribo*

Isolated yield: 34 mg, 64%, pale yellow oil

^{1}H NMR (600 MHz) CD_3OD, δ = 7.39-7.30 (m, 5H, ArH), 5.71 (d, *J* = 2.7 Hz, 1H, H-2), 5.16 (d, *J* = 12.4 Hz, 1H, CH_2), 5.12 (d, *J* = 12.4 Hz, 1H, CH_2), 4.00-3.95 (m, 2H, H-6a + H-4), 3.80 (br.s, 1H, H-3), 3.21-3.14 (m, 1H, H-6b), 1.84 (ddd, *J* = 25.5, 12.5, 4.9 Hz, 1H, H-5a), 1.63-1.57 (m, 1H, H-5b).

^{13}C NMR (150 MHz) CD_3OD, δ = 157.6 (CO), 138.0 (ArC), 129.5-128.9 (ArCH), 80.5 (C-2), 71.4 (C-3), 68.4 (CH_2), 67.1 (C-4), 38.6 (C-6), 28.3 (C-5).

HRMS (ESI): *m/z* $[M+Na]^+$ calcd. For $C_{13}H_{17}NO_5$: 290.0999; found 290.0998.



2e

benzyl (3*S*,4*R*,5*R*)-2,3,4,5-tetrahydroxyazepane-1-carboxylate **2e**

deoxy-D-*galacto*

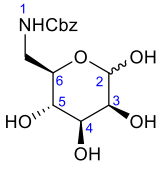
Isolated yield: 53 mg, 89%, colourless crystalline solid

Ratio anomers **2e'**:**2e''** ~ 6:4

^{1}H NMR (400 MHz) CD_3OD, δ = 7.34-7.27 (m, 5H, ArH), 5.15 (d, *J* = 3.5 Hz, 1H, H-2, **2e''**), 5.10 (d, *J* = 2.2 Hz, 1H, H-2, **2e'**), 5.06 (br.s, 2H, CH_2), 3.97 (ddd, *J* = 4.2, 6.8, j8.6 Hz, 1H, H-5, **2e'**), 3.90 (dd, *J* = 2.3, 4.5 Hz, 1H, H-3, **2e'**), 3.87-3.86 (m, 2H, H-3, H-4, **2e''**), 3.68-3.64 (m, 2H, H-4, **2e'**, H-5, **2e''**), 3.31-3.21 (m, 2H, H-7), 1.95-1.86 (m, 1H, H-6a), 1.84-1.77 (m, 1H, 6b, **2e''**), 1.77-1.68 (m, 1H, 6b, **2e'**)

^{13}C NMR (100 MHz) CD_3OD, δ = 158.8 (CO), 138.4 (ArC), 129.4-128.8 (ArCH), 103.3 (C-2, **2e'**), 97.0 (C-2, **2e''**), 84.4 (C-3, **2e'**), 82.3 (C-4, **2e'**), 81.7 (C-5, **2e'**), 80.6 (C-5, **2e''**), 80.1 (C-3, **2e''**), 78.6 (C-4, **2e''**), 67.4 (CH_2), 39.01 (C-7, **2e''**), 38.95 (C-7, **2e'**), 36.1 (C-6, **2e''**), 34.6 (C-6, **2e'**).

HRMS (ESI): *m/z* $[M+Na]^+$ calcd. For $C_{14}H_{19}NO_6$: 320.1105; found 320.1104.



2f

benzyl (((2R,3S,4S,5S)-3,4,5,6-tetrahydroxytetrahydro-2H-pyran-2-yl)methyl)carbamate **2f**

D-*manno*

Isolated yield: 30 mg, 48%, colourless oil

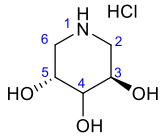
Ratio anomers **2f'**:**2f''** ~ 7:3

^{1}H NMR (400 MHz) CD_3OD, δ = 7.36-7.27 (m, 5H, ArH), 5.07 (s, 2H, CH_2), 5.04 (d, 1H, *J* = 1.2 Hz, 1H, H-2, **2f'**), 4.72 (br.s, 1H, H-2, **2f''**), 3.80-3.73 (m, 4H, H-3, H-6, H-4, **2f'**, H-3, **2f''**), 3.59-3.32 (m, 4H, H-5, **2f'**, 3CH, **2f''**)), 3.25-3.21 (m, 1H, CH, **2f''**).

^{13}C NMR (100 MHz) CD_3OD, δ = 159.1 (CO), 138.3 (ArC), 129.5-128.9 (ArCH), 95.9 (C-2, **2f'**), 95.7 (C-2, **2f''**), 76.2 (CH, **2f''**), 75.0 (CH, **2f''**), 73.1 (C-3, **2f''**), 72.9 (C-3, **2f'**), 72.2 (C-6, **2f'**), 72.0 (C-4, **2f'**), 69.7 (C-5, **2f'**), 69.4 (CH, **2f''**), 67.5 (CH_2), 43.1 (C-7).

HRMS (ESI): *m/z* $[M+Na]^+$ calcd. For $C_{14}H_{19}NO_7$: 336.1054; found 336.1052.

Hydrogenation products



3a/3b

(3*R*,5*R*)-piperidine-3,4,5-triol hydrochloride **3a/3b**[3]

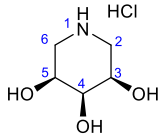
D-*arabino* / D-*lyxo*

Isolated yield: 6 mg, 64%, colourless oil

^{1}H NMR (500 MHz) D_2O, δ = 4.26 (dt, *J* = 5.8, 2.8 Hz, 1H, H-3), 4.12 (td, *J* = 8.1, 4.1 Hz, 1H, H-5), 3.80 (dd, *J* = 7.8, 2.9 Hz, 1H, H-4), 3.44 (dd, *J* = 12.9, 4.0 Hz, 1H, H-6a), 3.32 (dd, *J* = 12.9, 6.0 Hz, 1H, H-2a), 3.23 (dd, *J* = 13.0, 2.7 Hz, 1H, H-2b), 2.98 (dd, *J* = 12.7, 8.4 Hz, 1H, H-6b).

^{13}C NMR (150 MHz) D_2O, δ = 70.8 (C-4), 65.0 (C-5), 64.6 (C-3), 46.0 (C-2), 45.5 (C-6).

HRMS (ESI): *m/z* $[M+H]^+$ calcd. For $C_5H_{11}NO_3$: 134.0812; found 134.0812.



3c

(3*S*,4*S*,5*R*)-piperidine-3,4,5-triol hydrochloride **3c**[5]

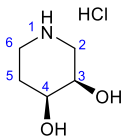
D-*ribo*

Isolated yield: 8 mg, 85%, pale yellow oil

^{1}H NMR (500 MHz) D_2O, δ = 4.10 (dt, *J* = 7.1, 3.7 Hz, 1H, H-4), 4.05 (t, *J* = 2.8 Hz, 2H, H-3/5), 3.29 – 3.19 (m, 4H, H2/6).

^{13}C NMR (125 MHz) D_2O, δ = 68.1 (C-4), 65.4 (C-3/5), 44.2 (C-/6).

HRMS (ESI): *m/z* $[M+H]^+$ calcd. For $C_5H_{11}NO_3$: 134.0812; found 134.0812.



3d

(3*R*,4*S*)-piperidine-3,4-diol hydrochloride **3d**[6]

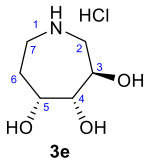
2-deoxy-D-*ribo*

Isolated yield: 9 mg, 78%, colourless oil

^{1}H NMR (500 MHz) D_2O, δ = 4.13 – 4.08 (m, 1H, H-3), 3.99 (dt, *J* = 9.6, 3.6 Hz, 1H, H-4), 3.38 – 3.32 (m, 2H, H-2a+H-6a), 3.20 (dd, *J* = 13.2, 1.9 Hz, 1H, H-2b), 3.13 – 3.07 (m, 1H, H-6b), 2.05 (ddt, *J* = 14.3, 10.0, 5.1 Hz, 1H, H-5a), 1.94 (dq, *J* = 14.0, 4.4 Hz, 1H, H-5b).

^{13}C NMR (125 MHz) D_2O, δ = 66.2 (C-4), 65.0 (C-3), 46.4 (C-2), 41.2 (C-6), 25.0 (C-5).

HRMS (ESI): *m/z* $[M+H]^+$ calcd. For $C_5H_{11}NO_2$: 118.0863; found 118.0864.



3e

(3*R*,4*R*,5*R*)-azepane-3,4,5-triol hydrochloride **3e**[7]

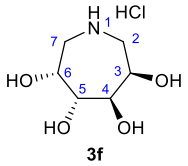
2-deoxy-D-*galacto*

Isolated yield: 7 mg, 57%, pale yellow oil

^{1}H NMR (500 MHz) D_2O, δ = 4.14 (dt, $J_{5,6}$ = 9.8 Hz, $J_{5,4}$ = 2.2 Hz, 1H, H-5), 4.07 (td, *J* = 6.6, 1.9 Hz, 1H, H-3), 3.91 (dd, *J* = 6.3, 2.2 Hz, 1H, H-4), 3.45 (dd, *J* = 14.1, 1.3 Hz, 1H, H-2a), 3.38-3.32 (m, 1H, H-7a), 3.30-3.26 (m, 1H, H-7b), 3.24 (dd, *J* = 14.1, 6.8 Hz, 1H, H-2b), 2.26 (ddt, *J* = 20.6, 9.2, 4.9 Hz, 1H, H-6a), 2.00 – 1.77 (m, 1H, H-6b).

^{13}C NMR (125 MHz) D_2O, δ = 74.8 (C-5), 68.5 (C-4), 67.0 (C-3), 44.3 (C-2), 40.9 (C-7), 25.1 (C-6).

HRMS (ESI): *m/z* $[M+H]^+$ calcd. For $C_6H_{13}NO_3$: 148.0968; found, 148.0969.



3f

(3*R*,4*R*,5*R,6R*)-azepane-3,4,5,6-tetrol hydrochloride **3f**[8]

D-*manno*

Isolated yield: 8 mg, 78%, colourless oil

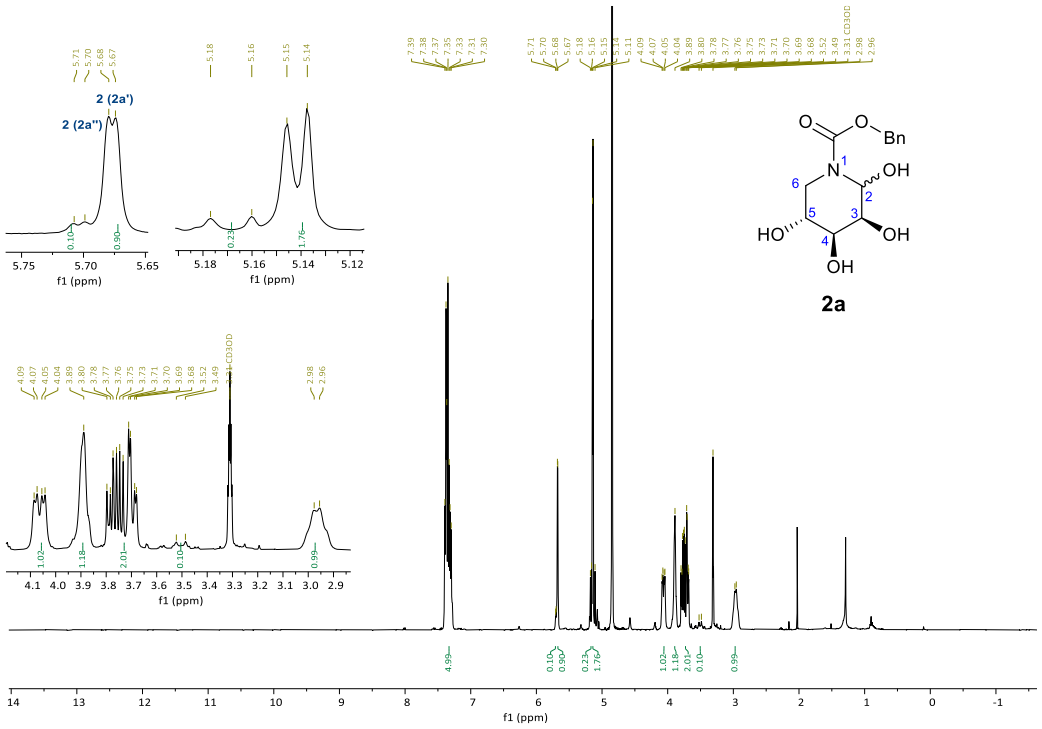
^{1}H NMR (500 MHz) D_2O, δ = 4.33 (d, *J* = 6.0 Hz, 1H, H-3/6), 3.88 (s, 1H, H-4/5), 3.47 – 3.34 (m, 2H, H-2/7).

^{13}C NMR (125 MHz) D_2O, δ = 73.1 (C-4/5), 66.9 (C-3/6), 45.0 (C-2/7).

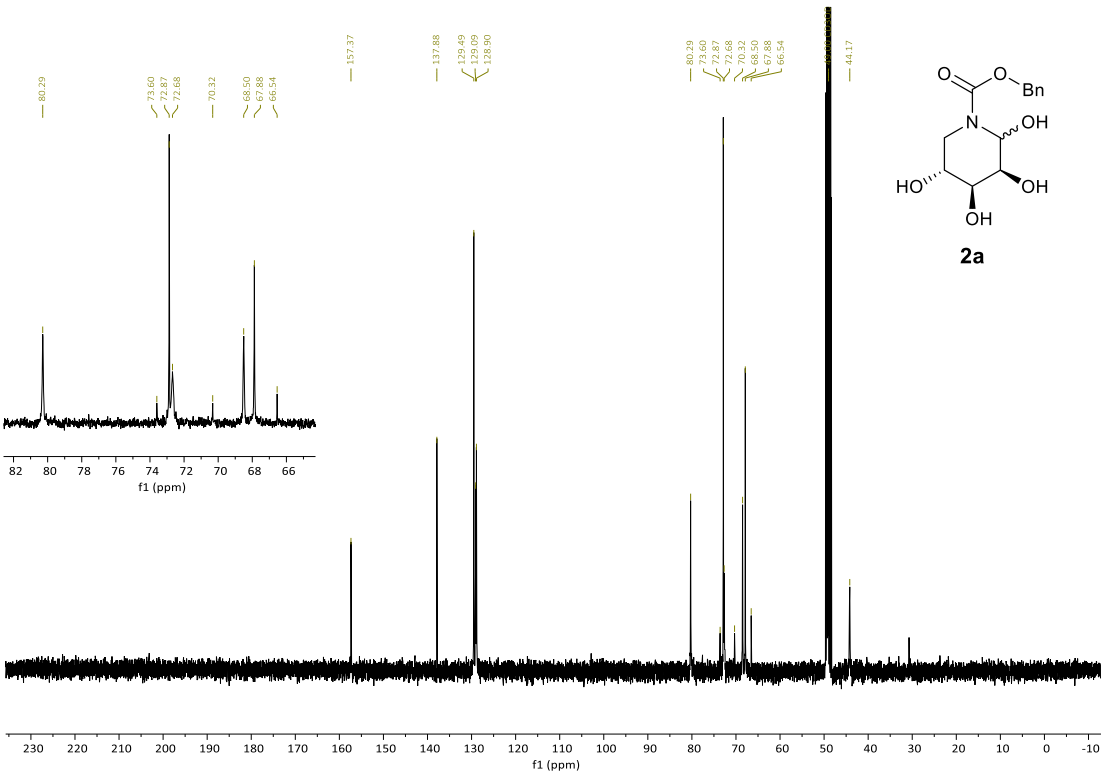
HRMS (ESI): *m/z* $[M+H]^+$ calcd. For $C_6H_{13}NO_4$: 164.0917; found 164.0917.

VII. NMR spectra

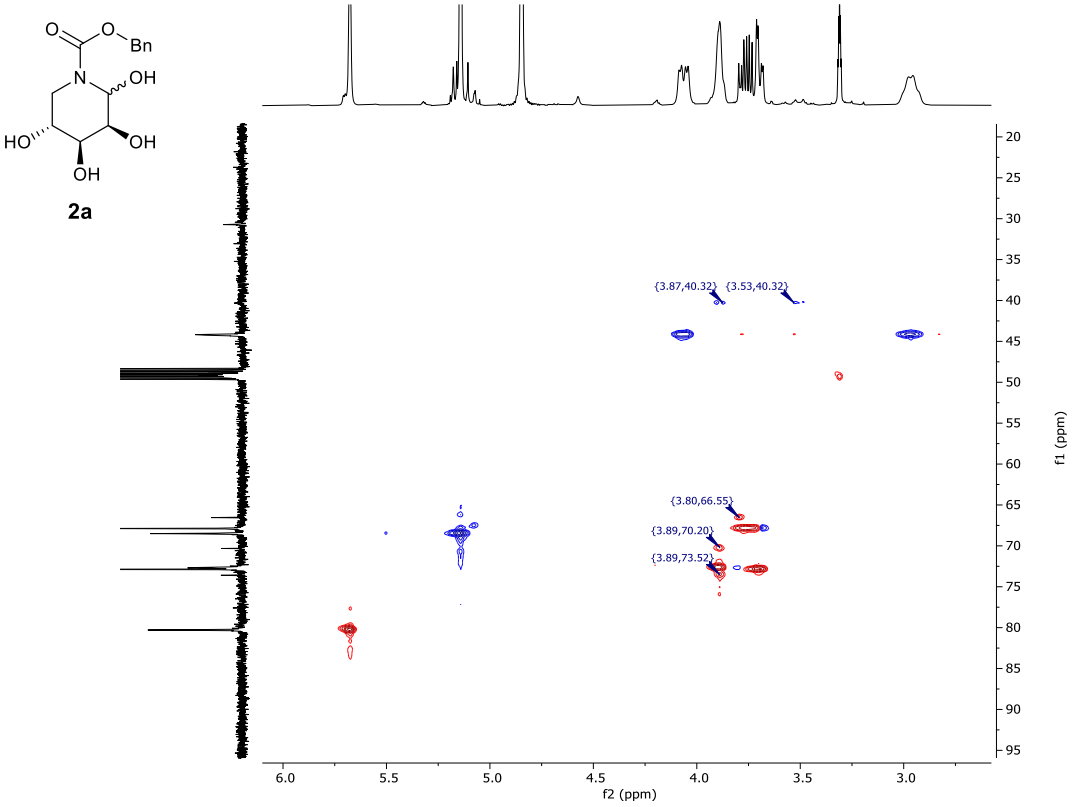
^{1}H NMR (600 MHz, CD_3OD) D-*arabino* **2a**



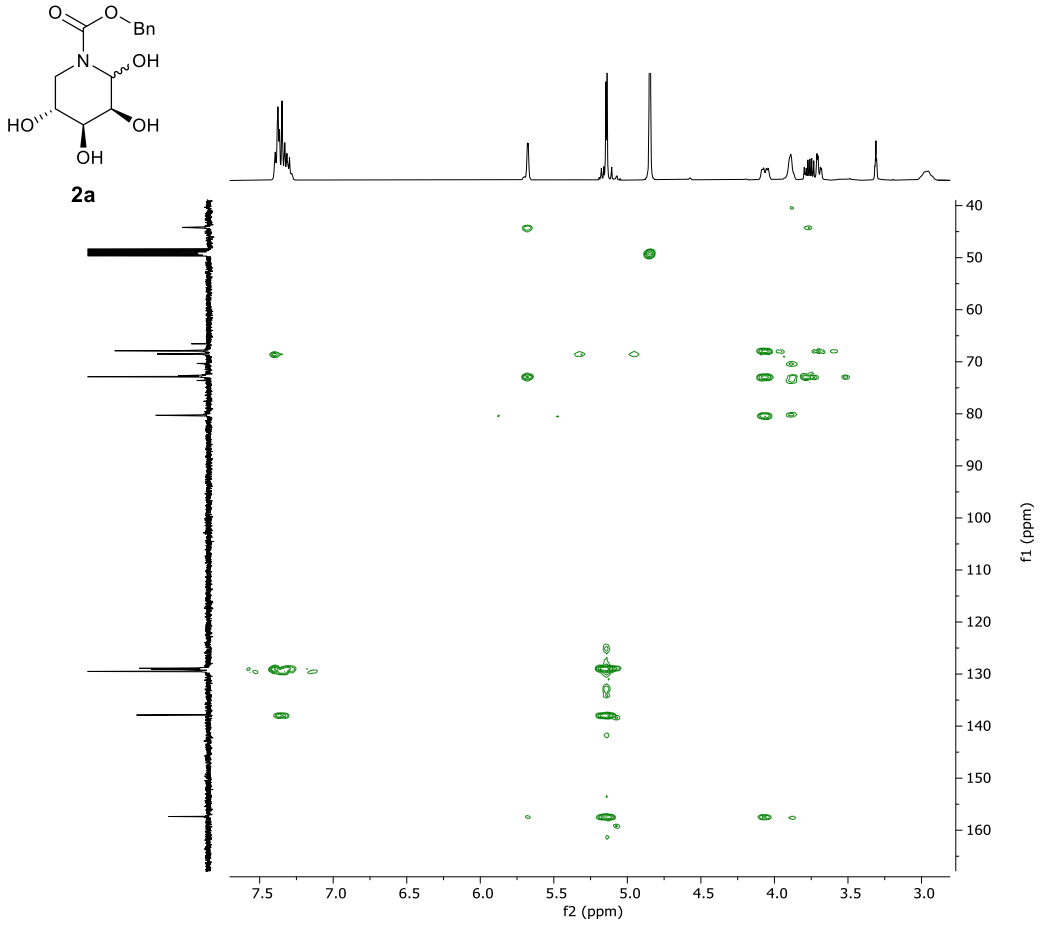
^{13}C NMR (150 MHz, CD_3OD) D-*arabino* **2a**



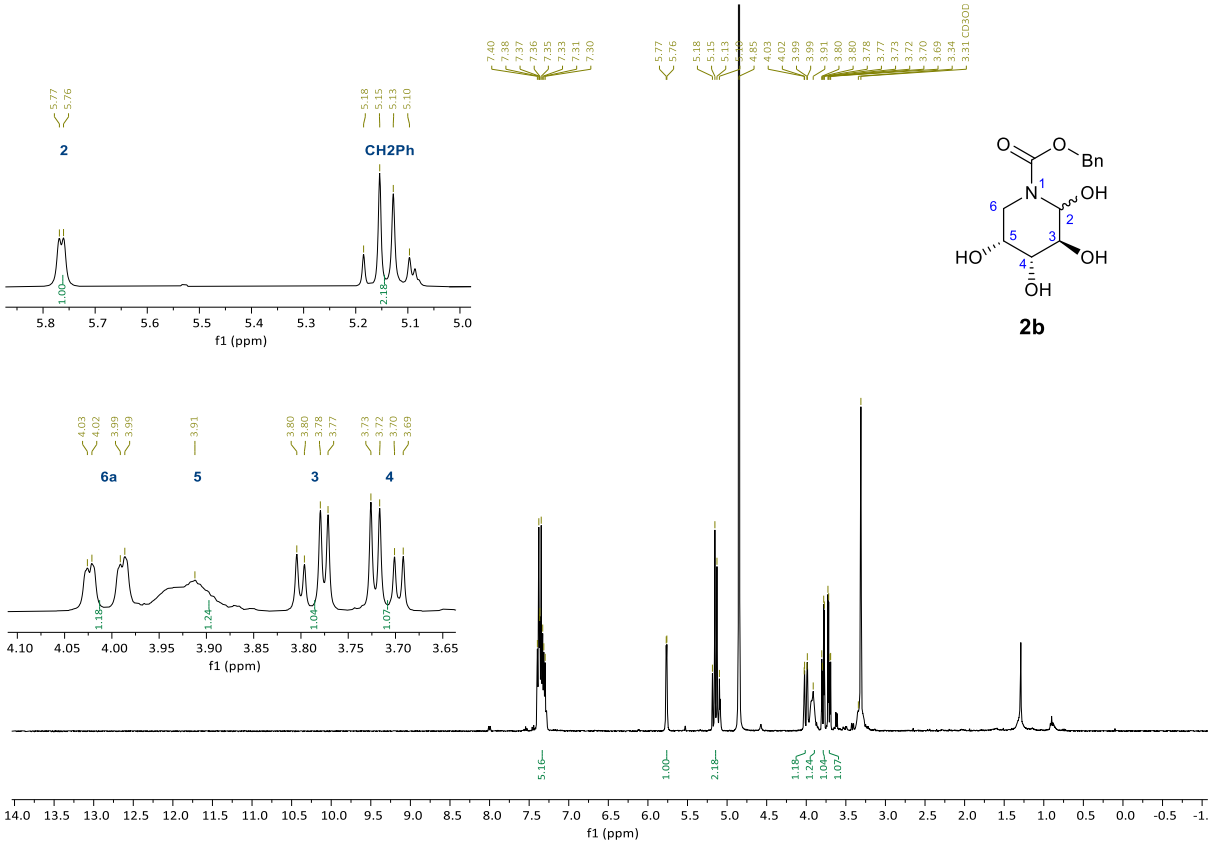
^{1}H-^{13}C HSQC NMR (CD_3OD) D-*arabino* **2a**



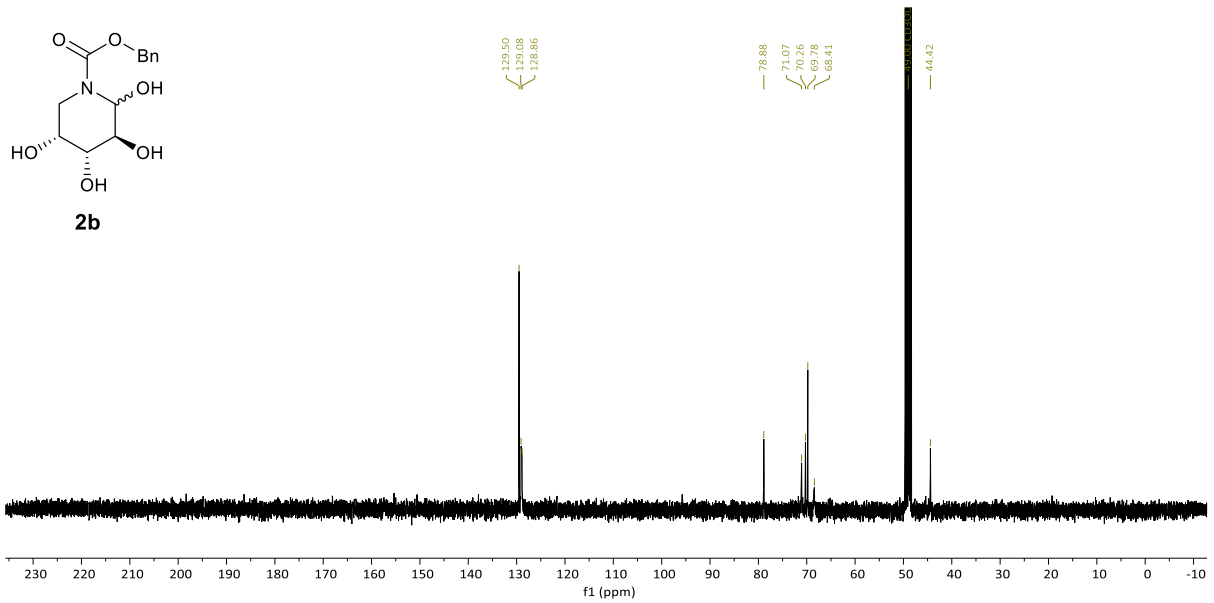
^{1}H-^{13}C HMBC NMR (CD_3OD) D-*arabino* **2a**



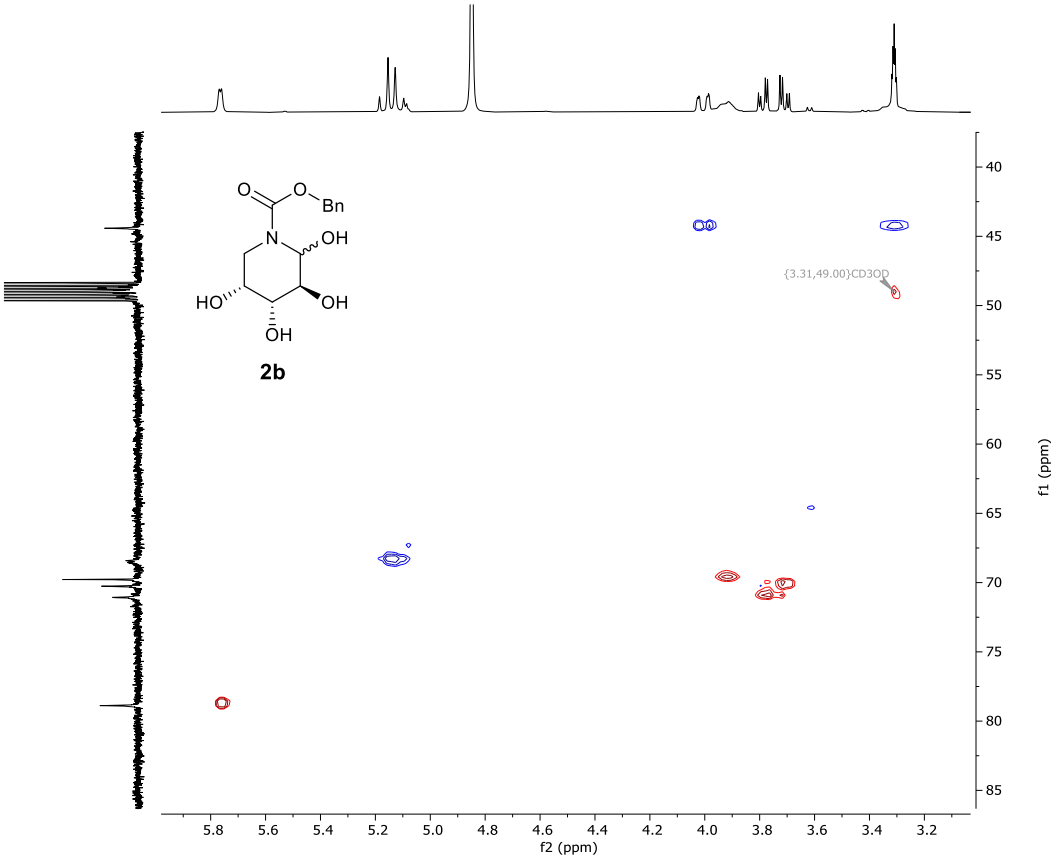
^{1}H NMR (400 MHz, CD_3OD) D-*lyxo* **2b**



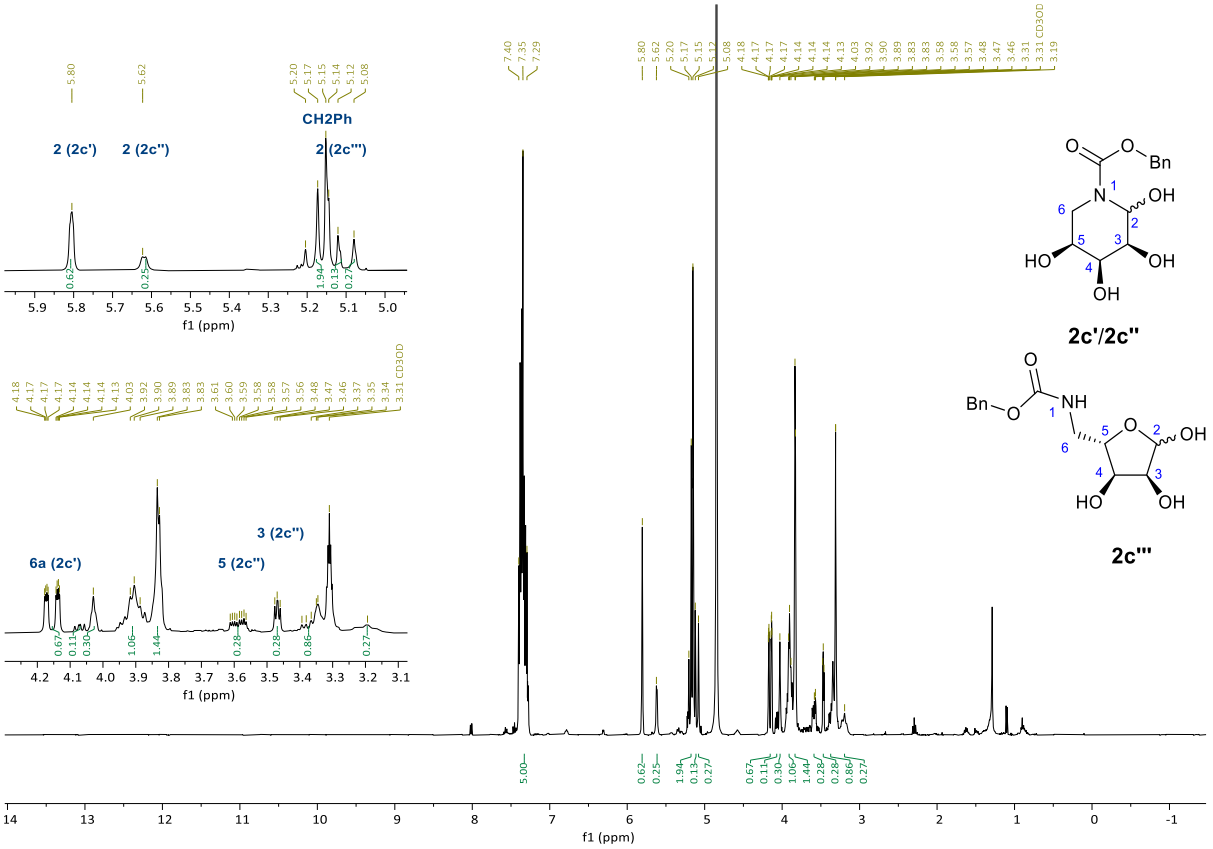
^{13}C NMR (100 MHz, CD_3OD) D-*lyxo* **2b**



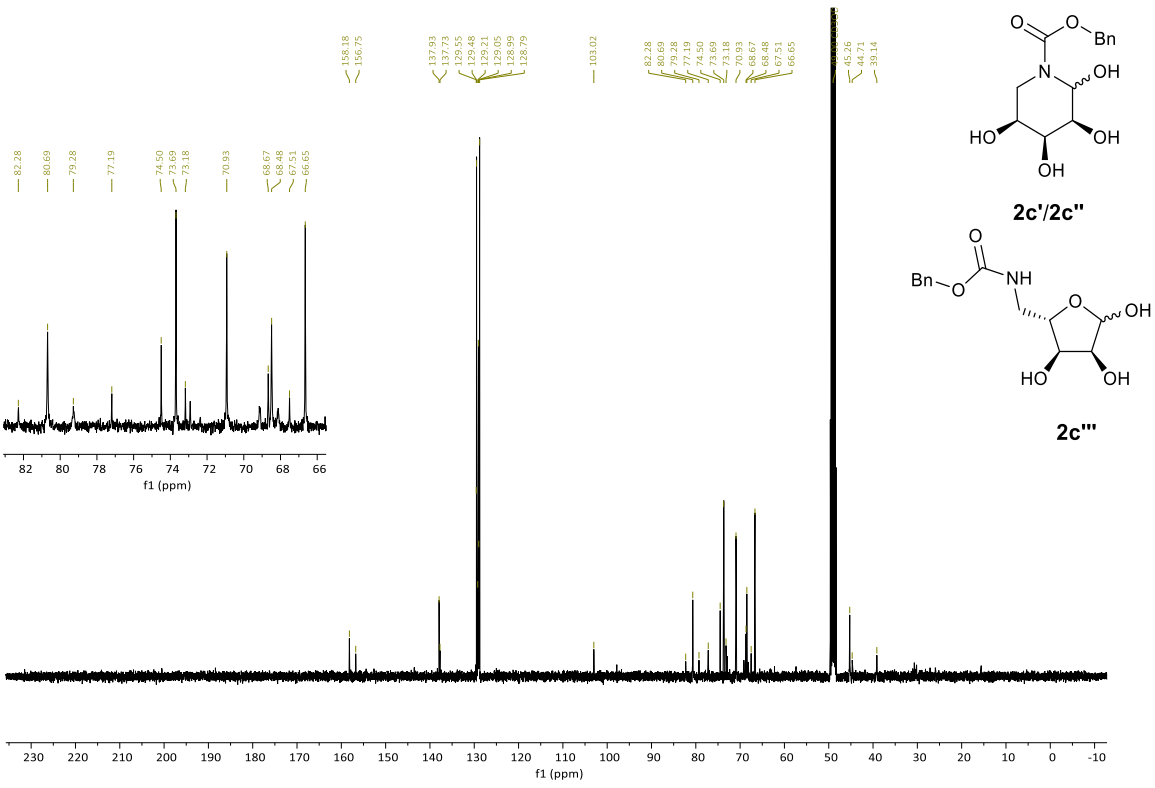
^{1}H-^{13}C HSQC NMR (CD_3OD) D-*lyxo* **2b**



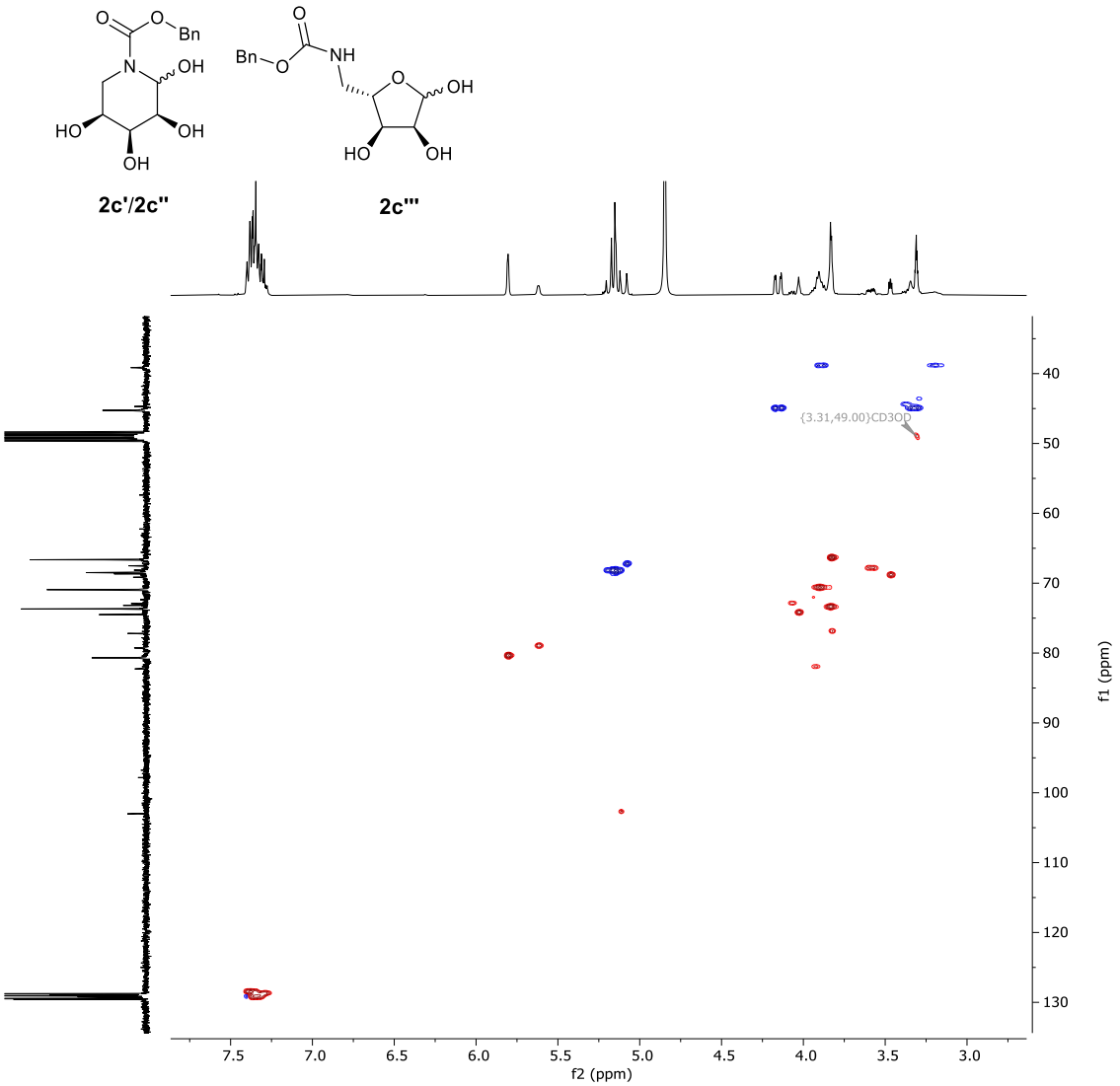
^{1}H NMR (400 MHz, CD_3OD) D-*ribo* **2c**



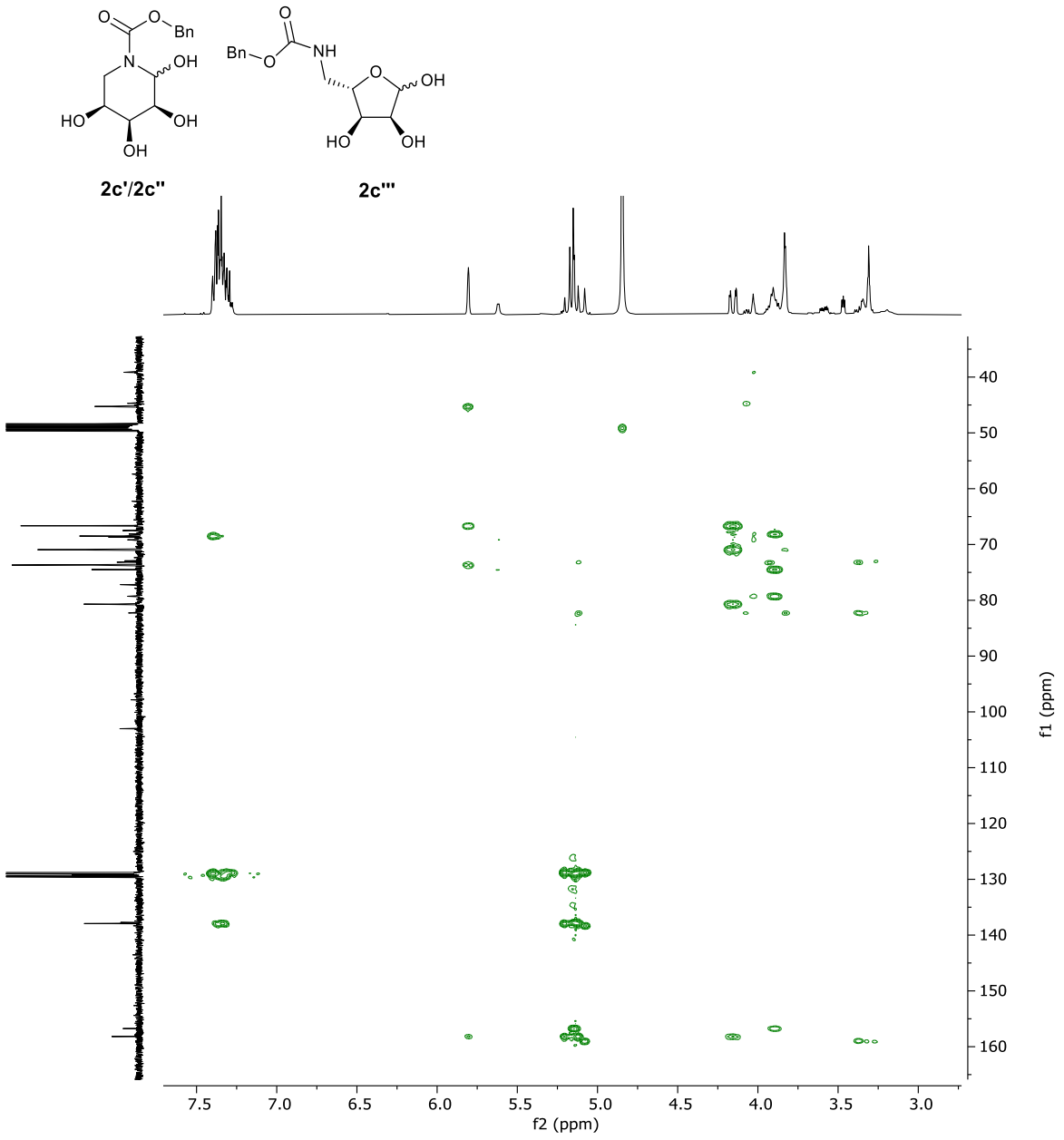
^{13}C NMR (100 MHz, CD_3OD) D-ribo **2c**



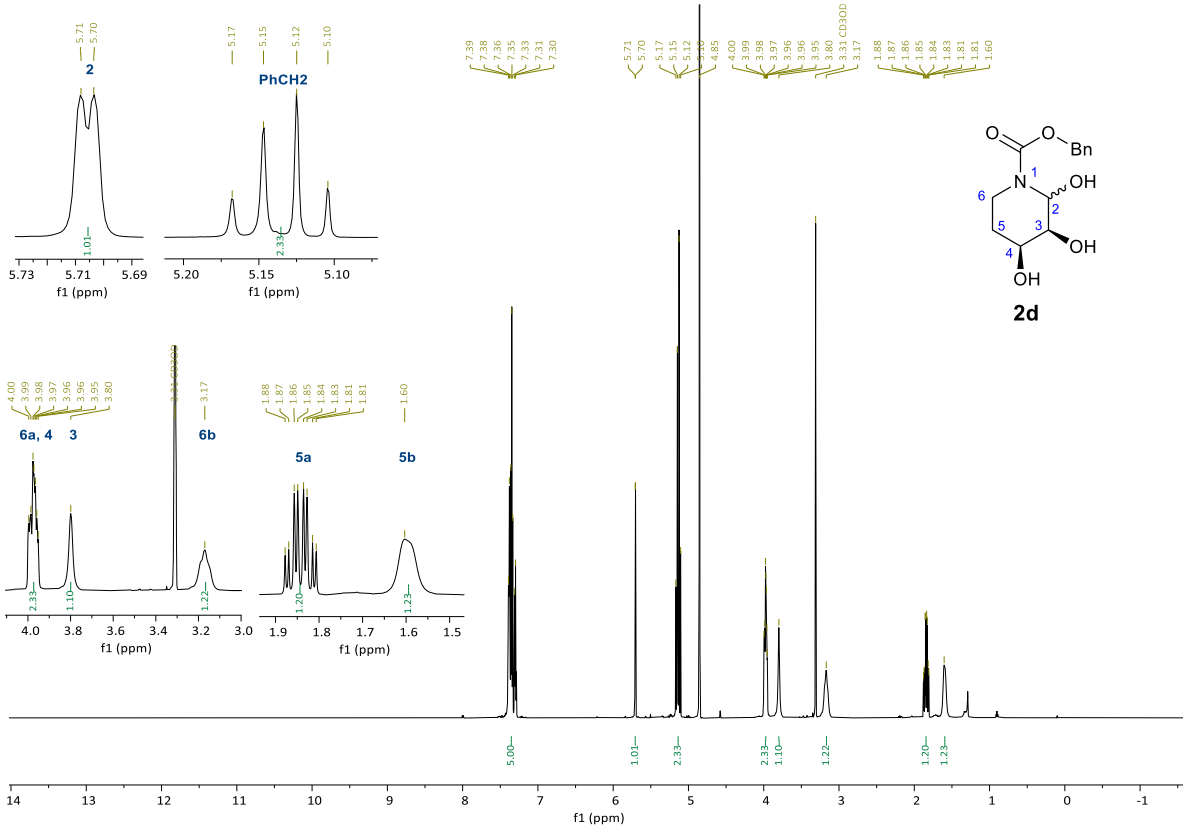
^{1}H-^{13}C HSQC NMR (CD_3OD) D-*ribo* **2c**



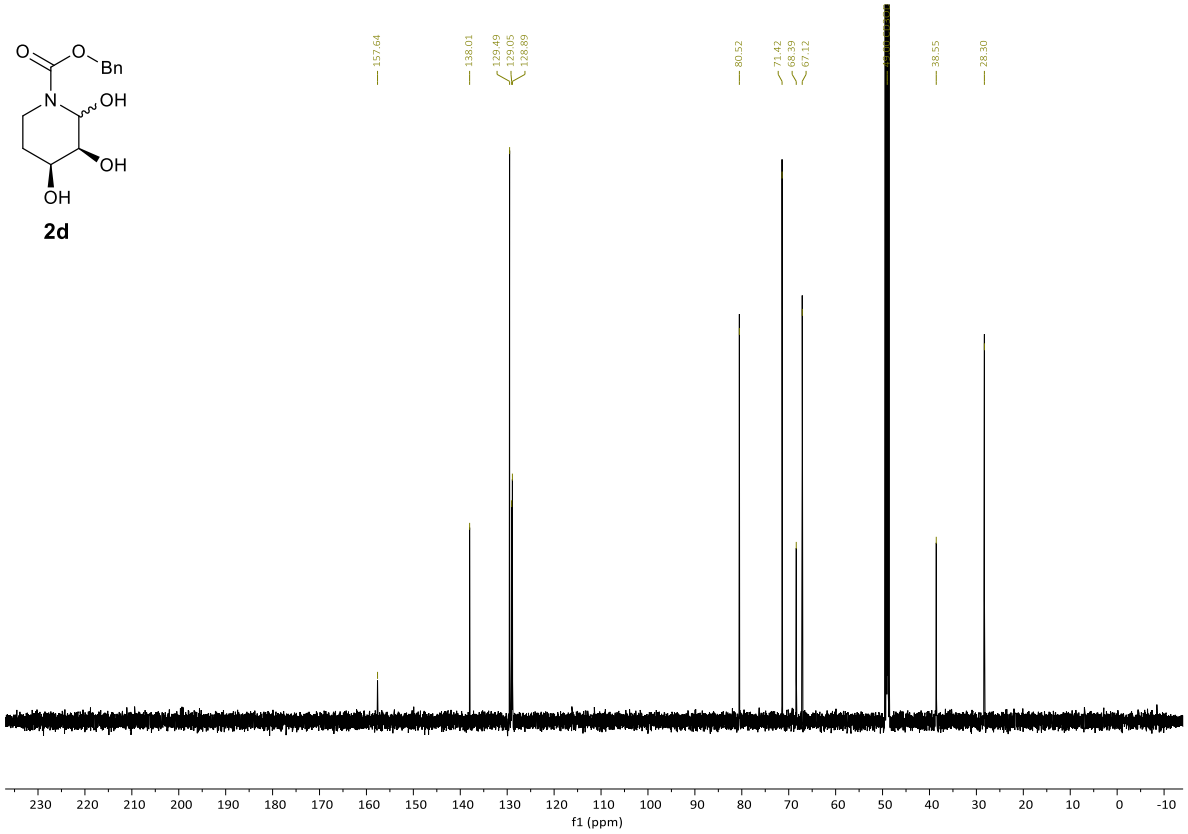
^{1}H-^{13}C HMBC NMR (CD_3OD) D-*ribo* **2c**



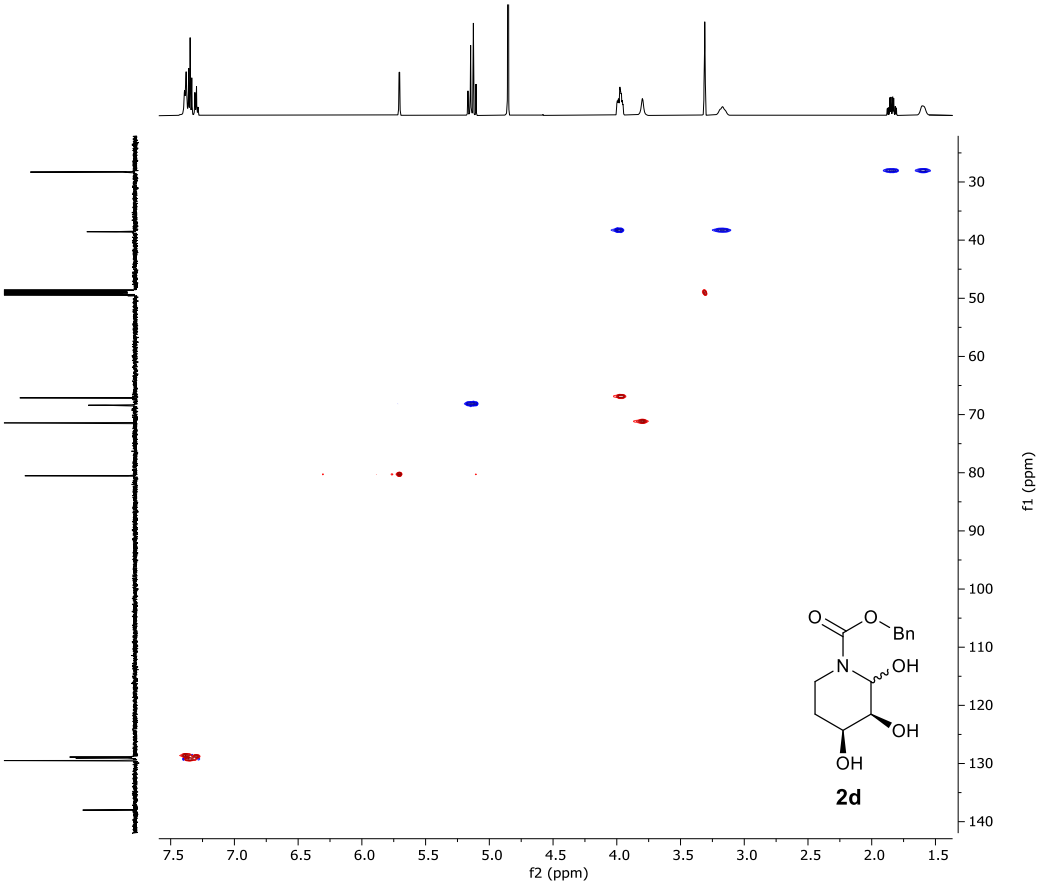
^{1}H NMR (600 MHz, CD_3OD) deoxy-D-*ribo* **2d**



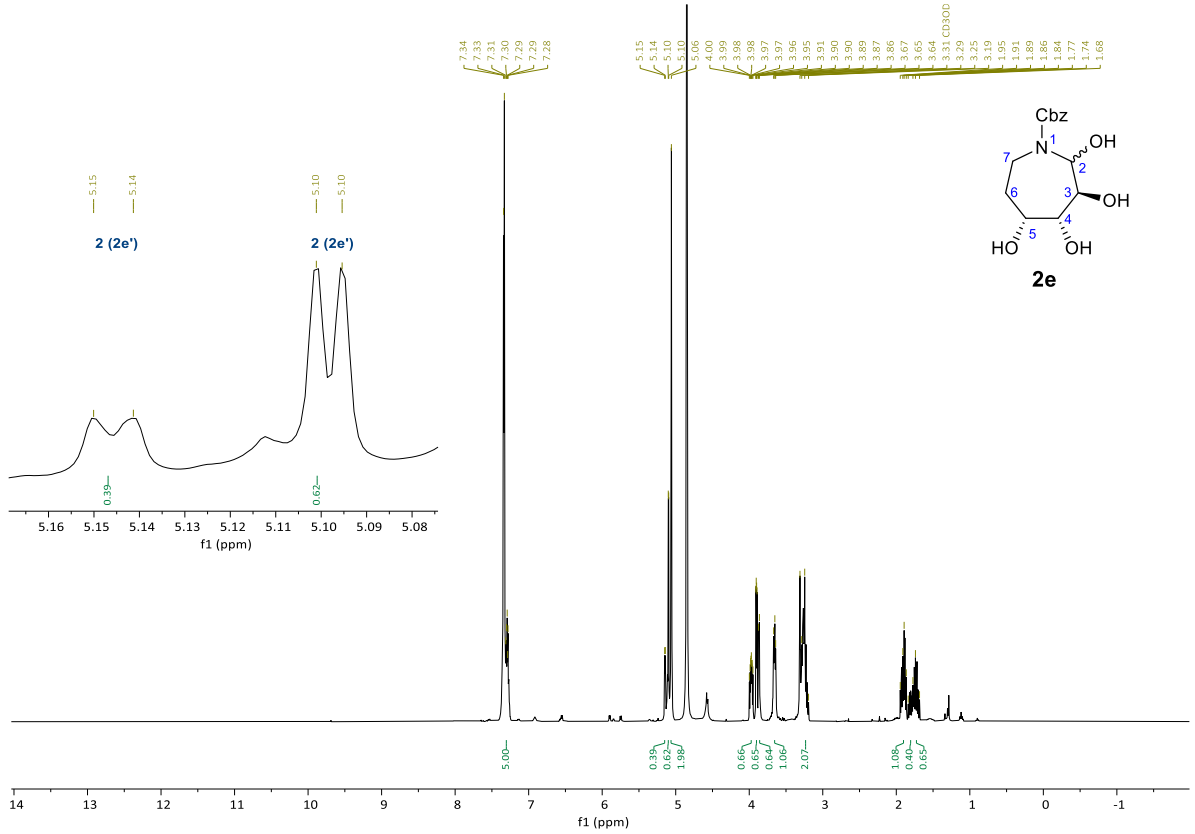
^{13}C NMR (100 MHz, CD_3OD) deoxy-D-*ribo* **2d**



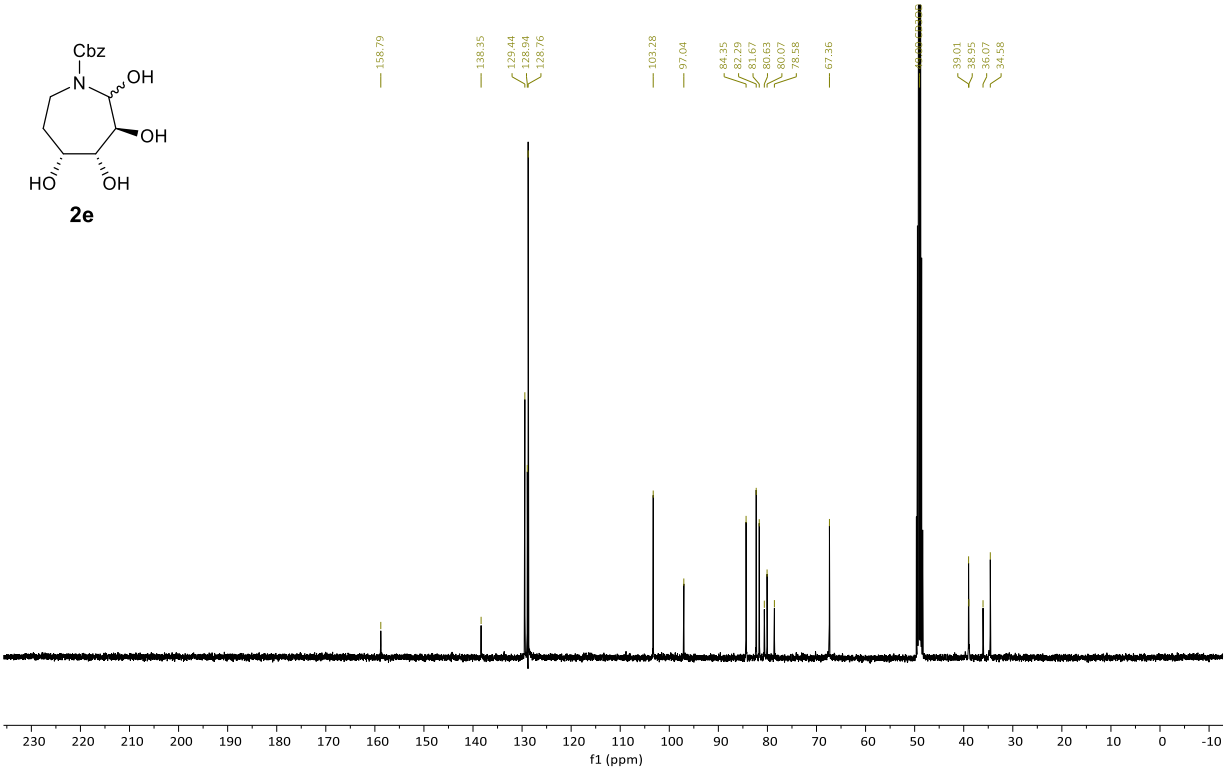
^{1}H-^{13}C HSQC NMR (CD_3OD) deoxy-D-*ribo* **2d**



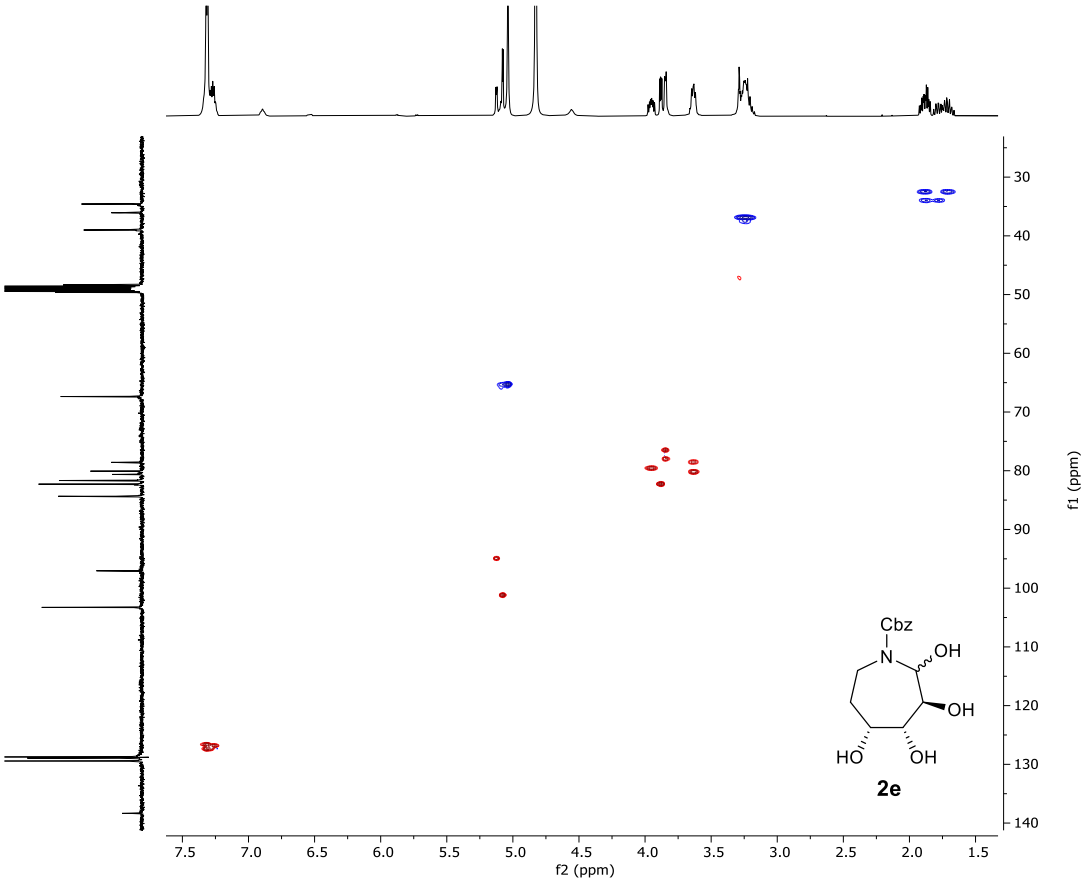
^{1}H NMR (400 MHz, CD_3OD) deoxy-D-*galacto* **2e**



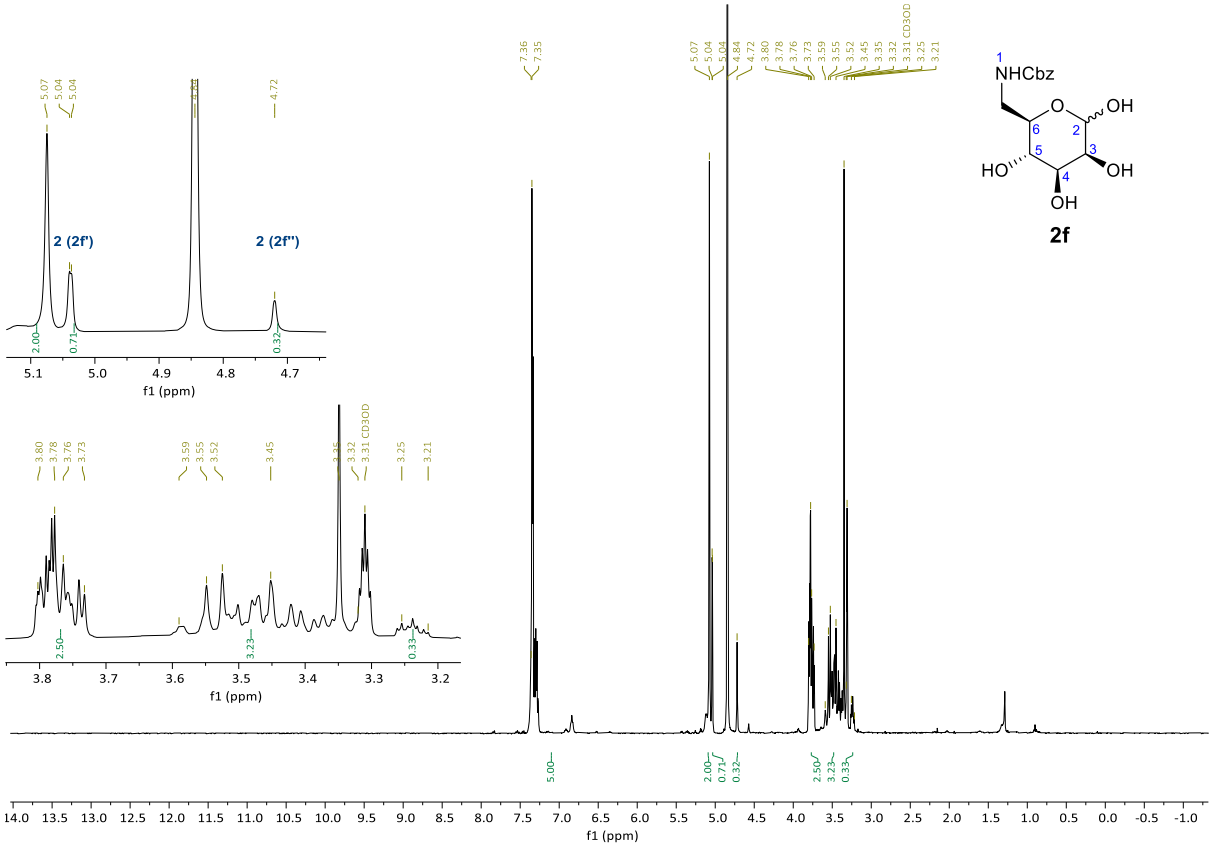
^{13}C NMR (100 MHz, CD_3OD) deoxy-D-*galacto* **2e**



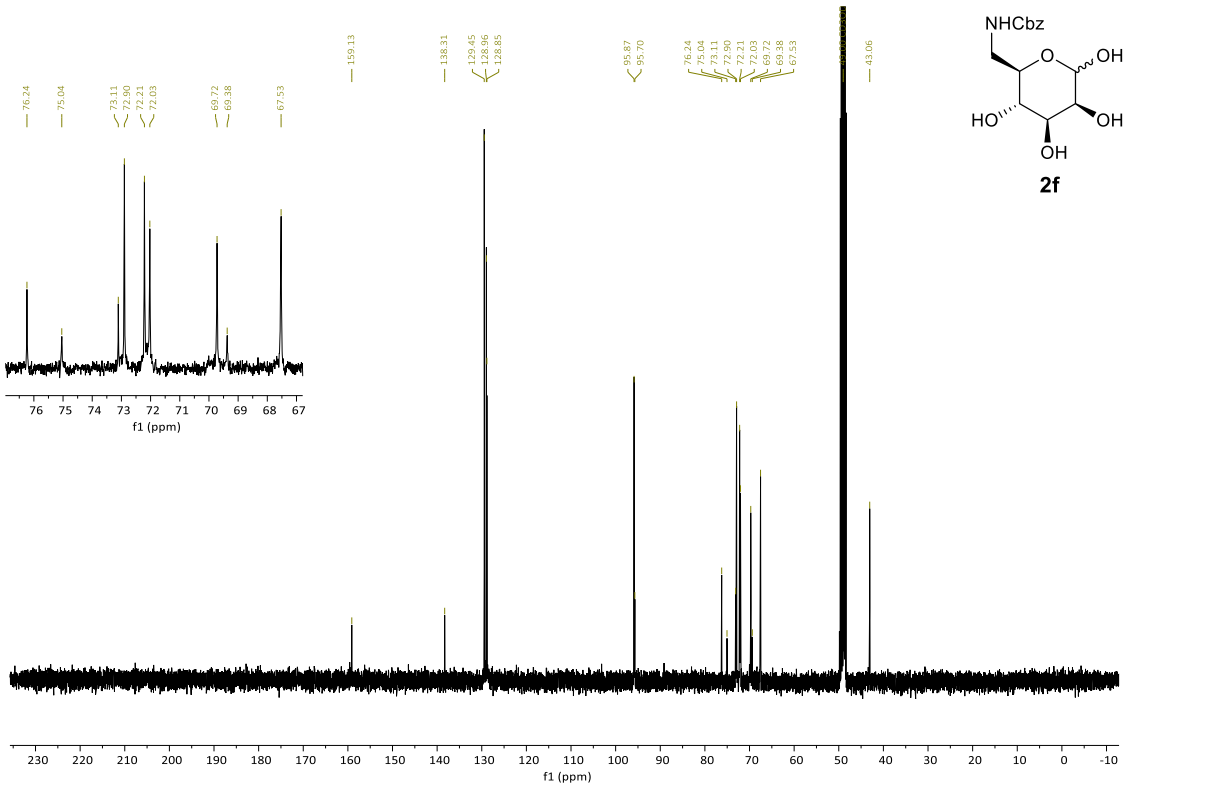
^{1}H-^{13}C HSQC NMR (CD_3OD) deoxy-D-*galacto* **2e**



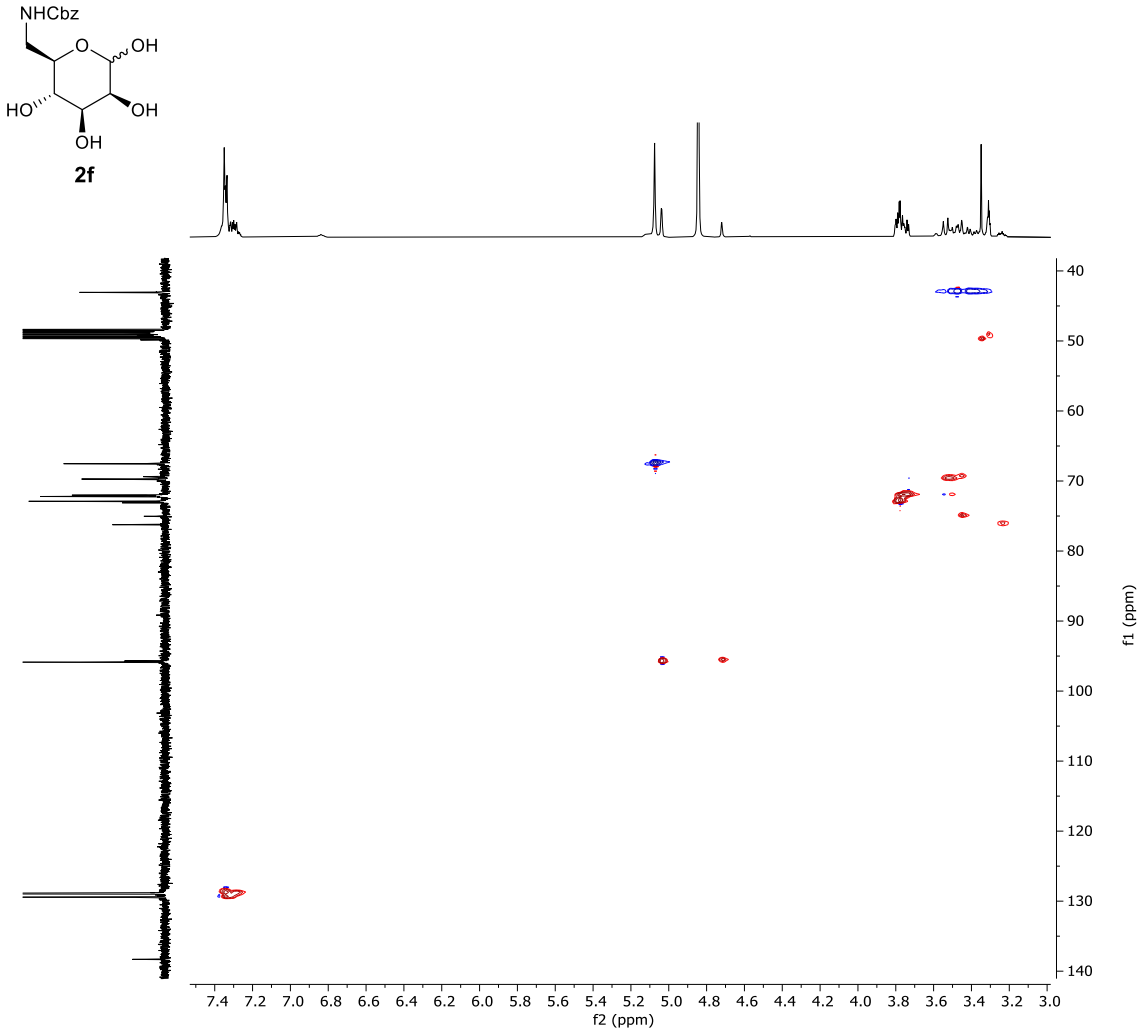
^{1}H NMR (400 MHz, CD_3OD) D-*manno* **2f**



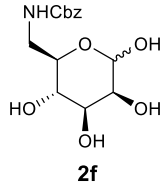
^{13}C NMR (100 MHz, CD_3OD) D-*manno* **2f**



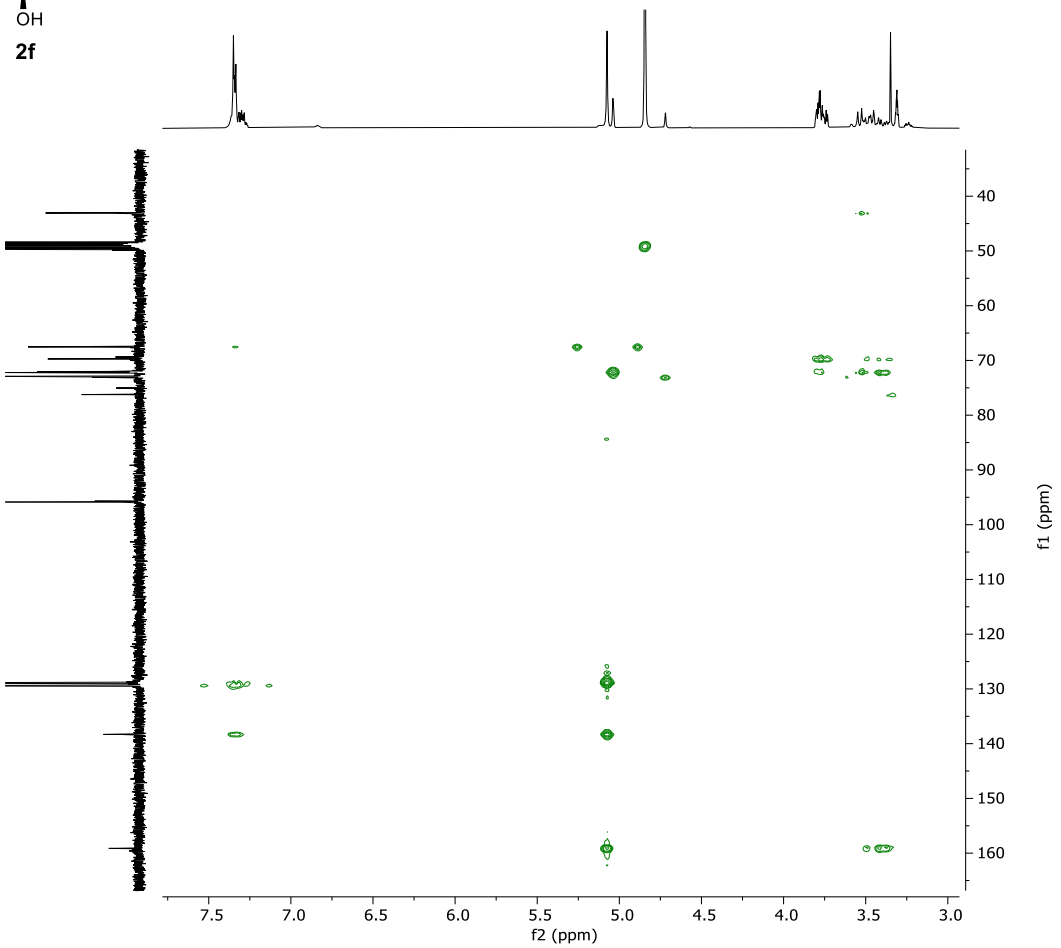
^{1}H-^{13}C HSQC NMR (CD_3OD) D-*manno* **2f**



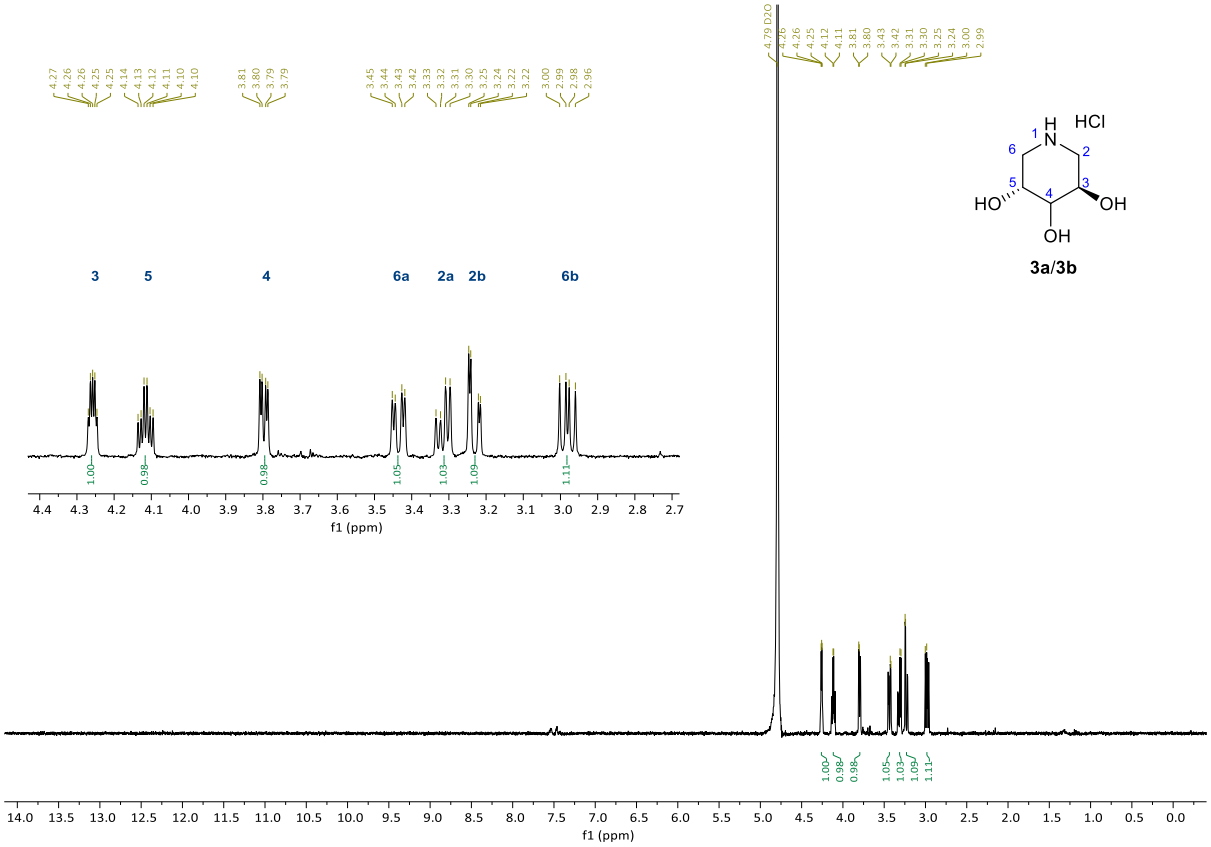
^{1}H-^{13}C HMBC NMR (CD_3OD) D-*manno* **2f**



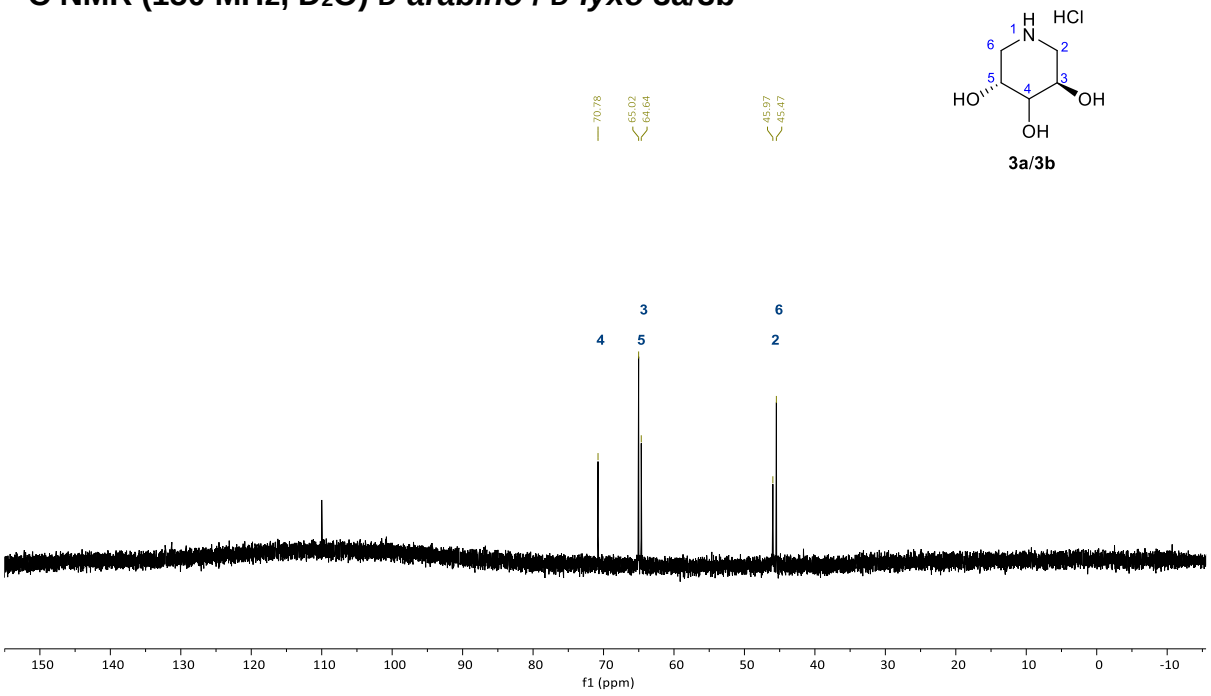
2f



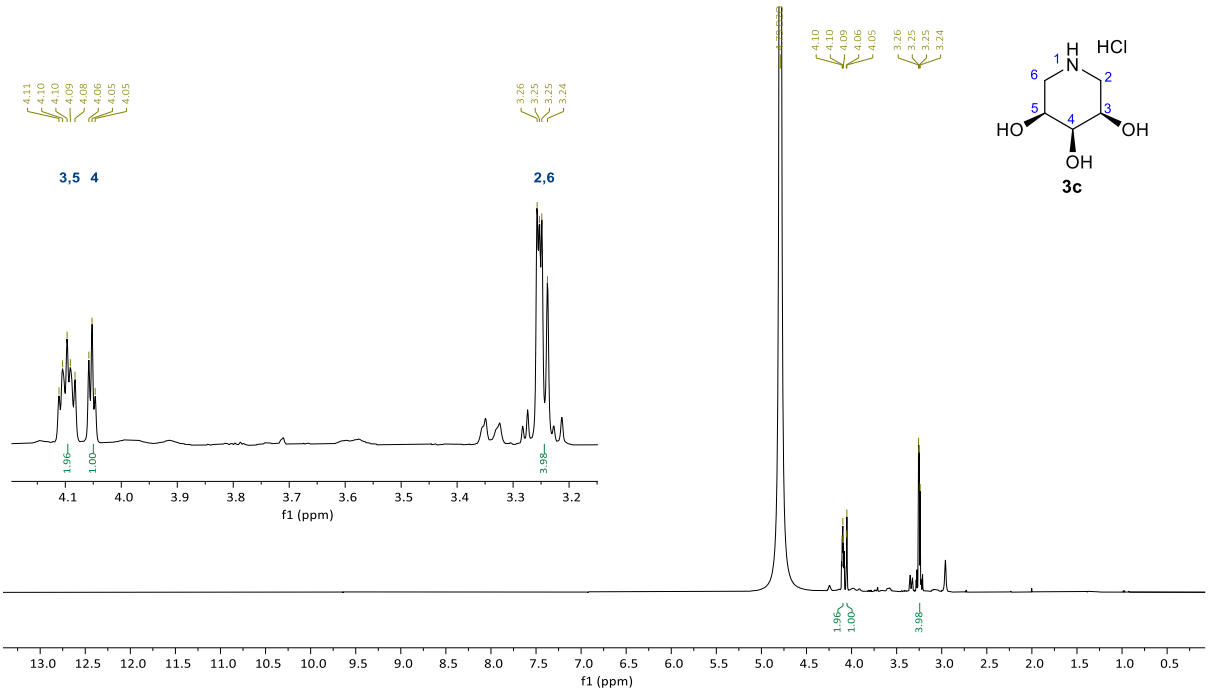
^{1}H NMR (500 MHz, D_2O) D-*arabino* / D-*lyxo* **3a/3b**



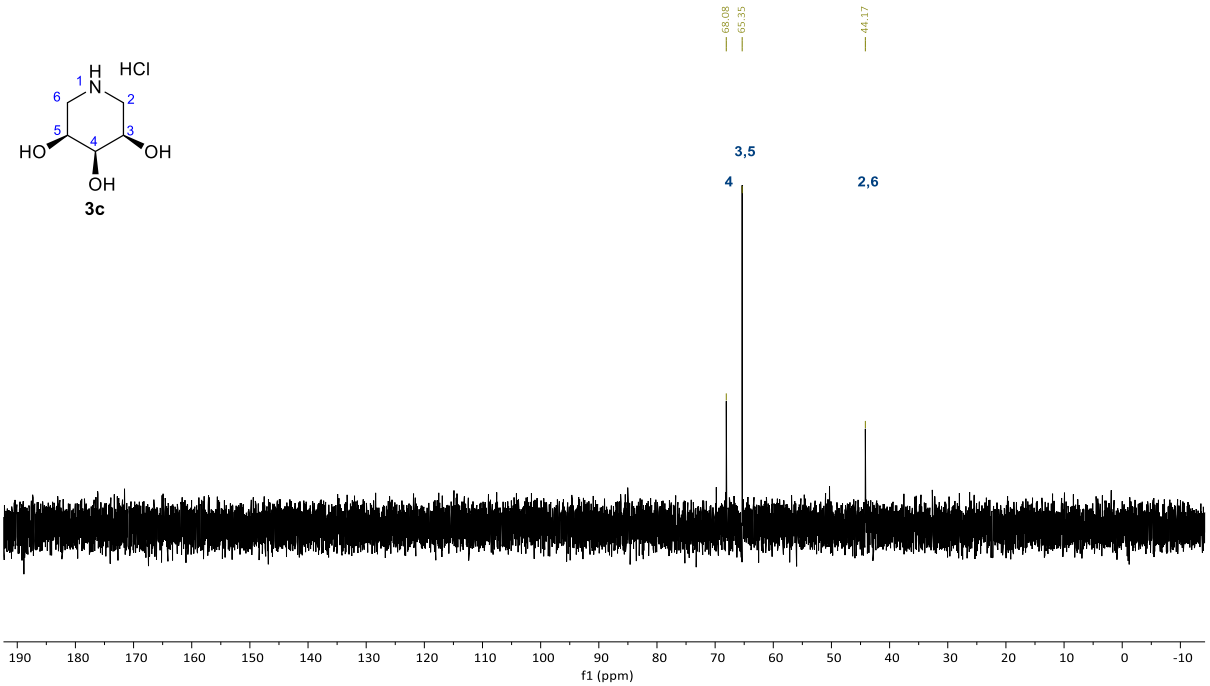
^{13}C NMR (150 MHz, D_2O) D-*arabino* / D-*lyxo* 3a/3b



^{1}H NMR (500 MHz, D_2O) D-*ribo* **3c**[c]

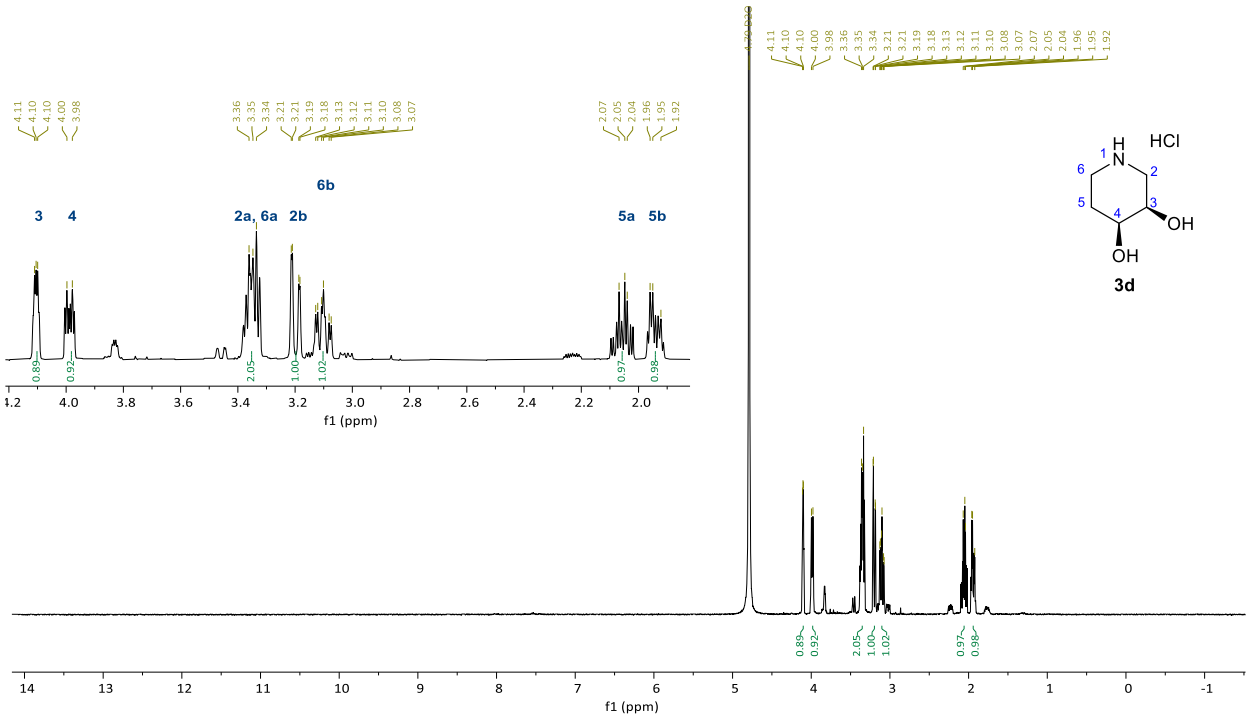


^{13}C NMR (125 MHz, D_2O) D-*ribo* **3c**

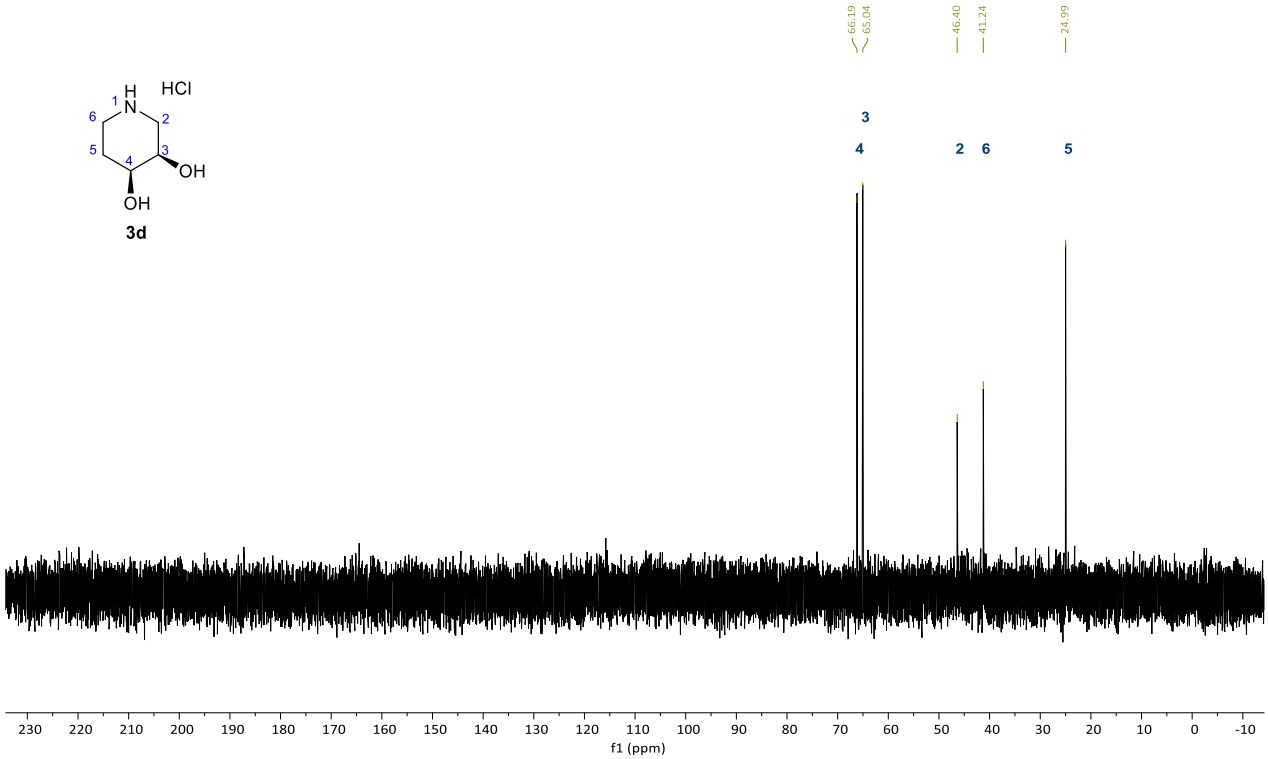


[c] A discussion of potential minor impurities is included in Section VIII.

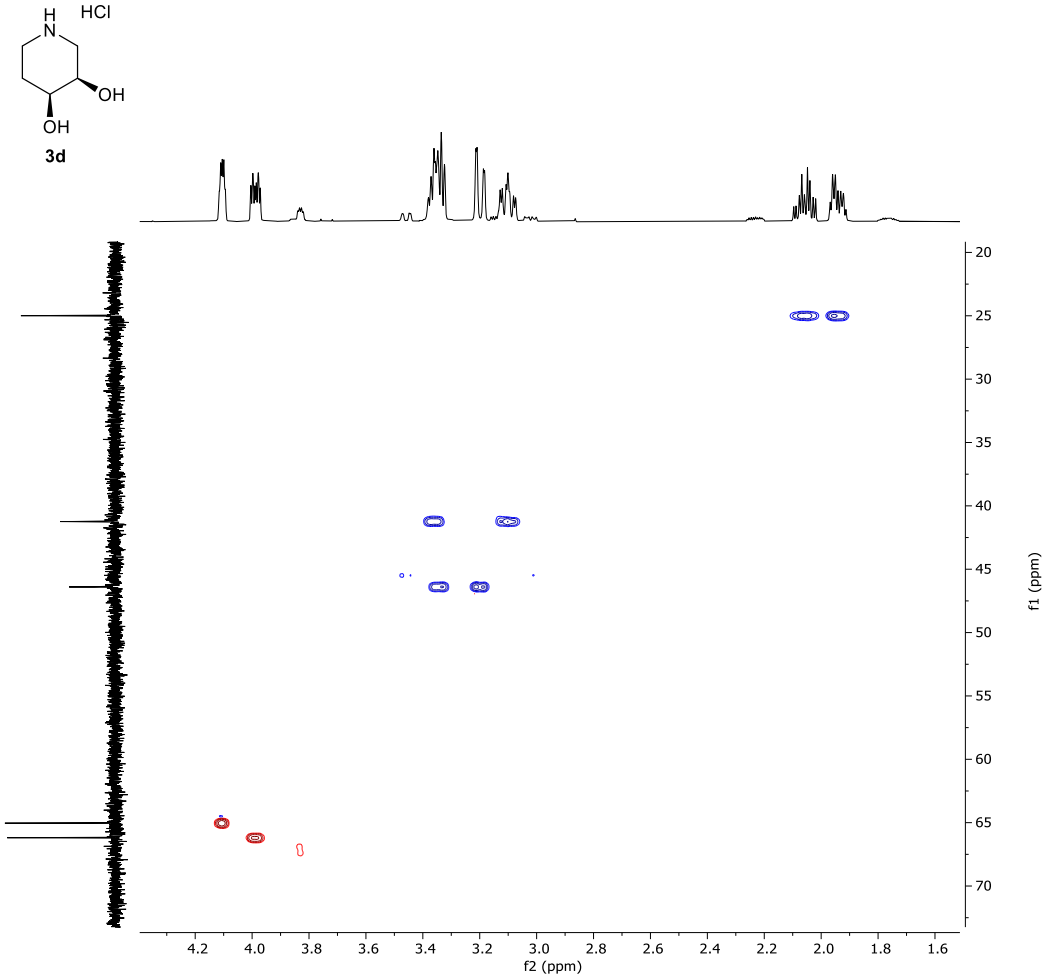
^{1}H NMR (500 MHz, D_2O) 2-deoxy-D-*ribo* **3d**



^{13}C NMR (125 MHz, D_2O) 2-deoxy-D-*ribo* **3d**

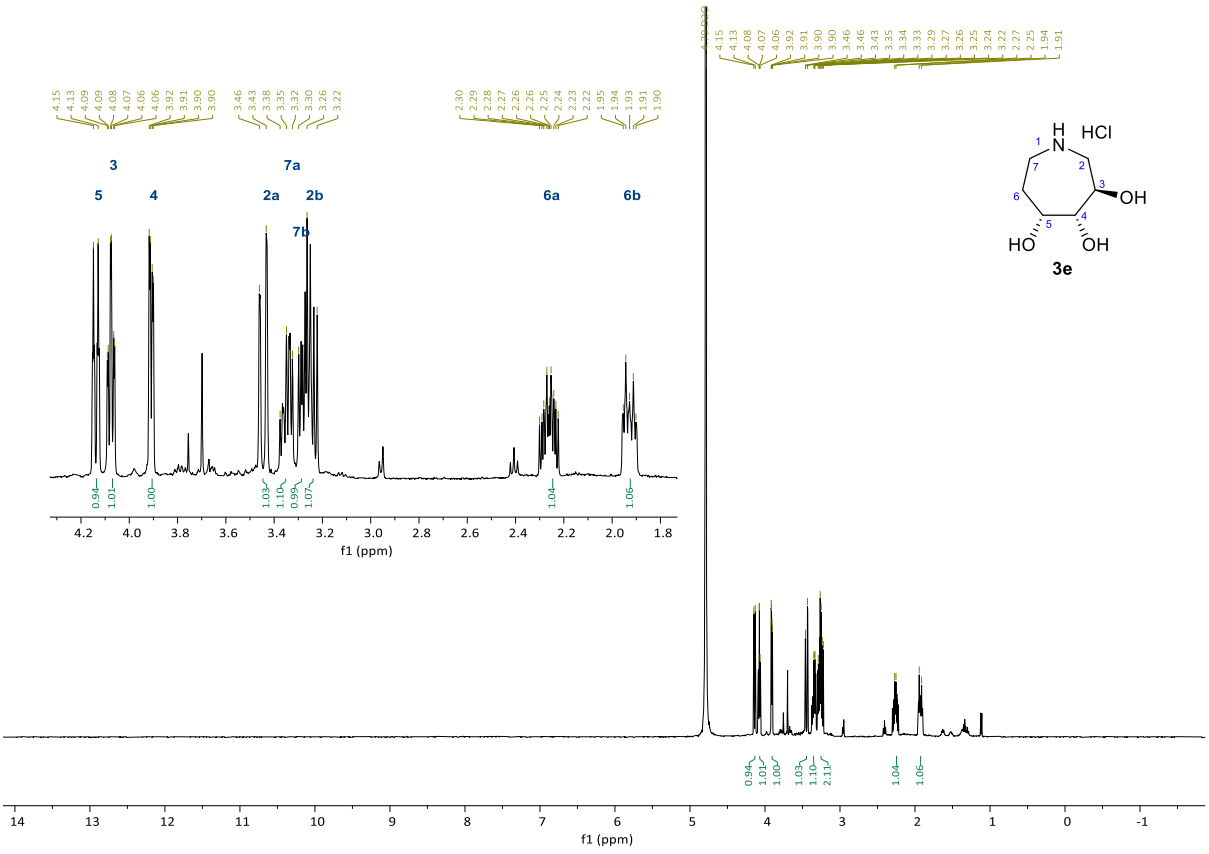


^{1}H-^{13}C HSQC NMR (D_2O) 2-deoxy-D-*ribo* **3d**[d]

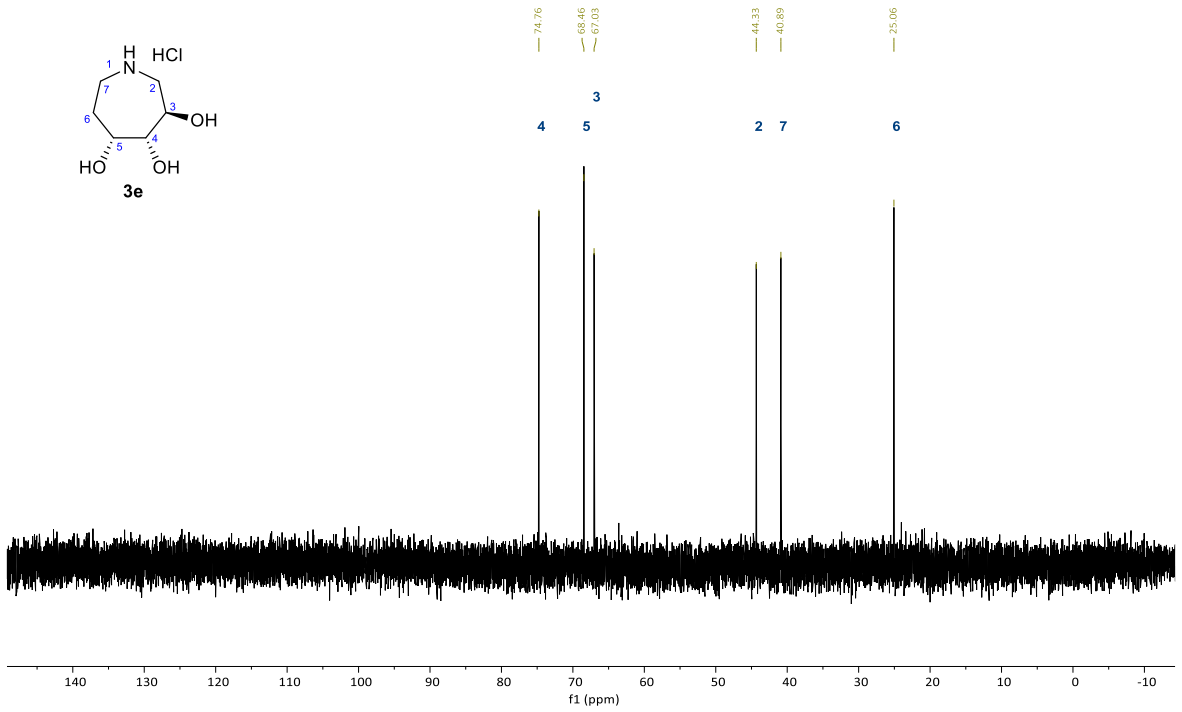


[d] The minor product observed is explained in Section VIII.

^{1}H NMR (500 MHz, D_2O) 2-deoxy-D-*galacto* **3e**[e]

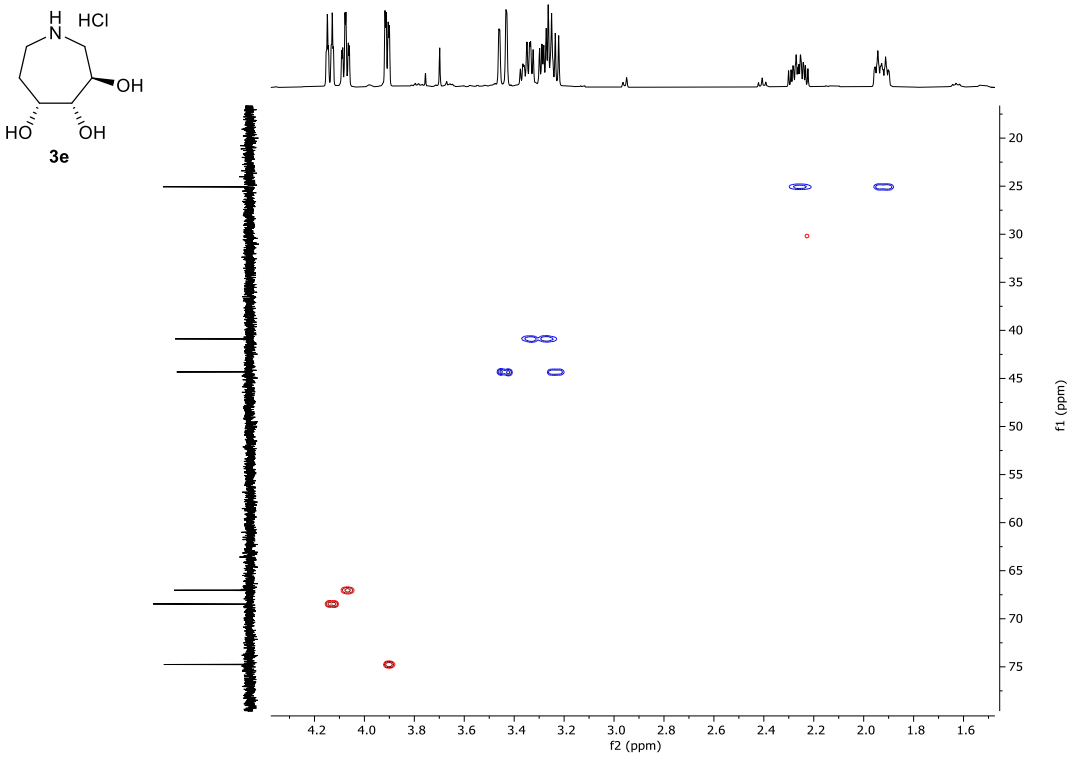


^{13}C NMR (125 MHz, D_2O) 2-deoxy-D-*galacto* **3e**

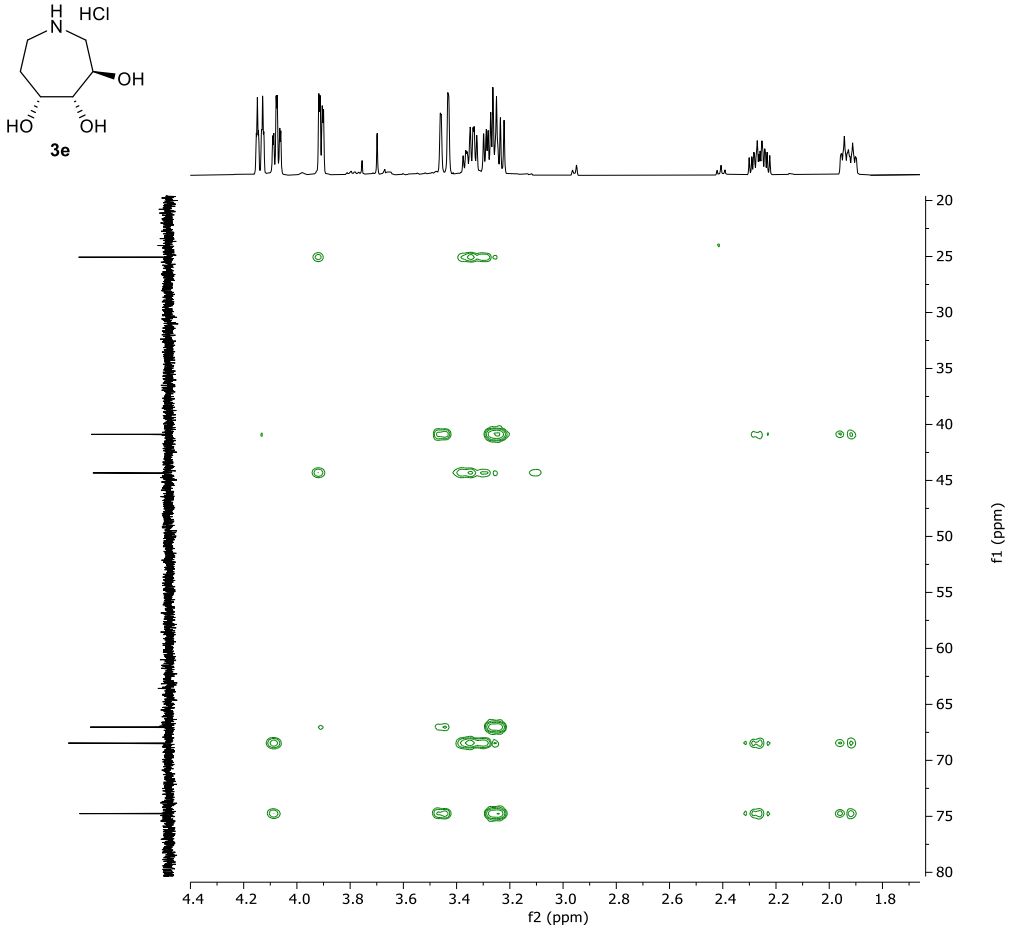


[e] There is a low concentration of desthiobiotin from protein purification which contaminates this sample.

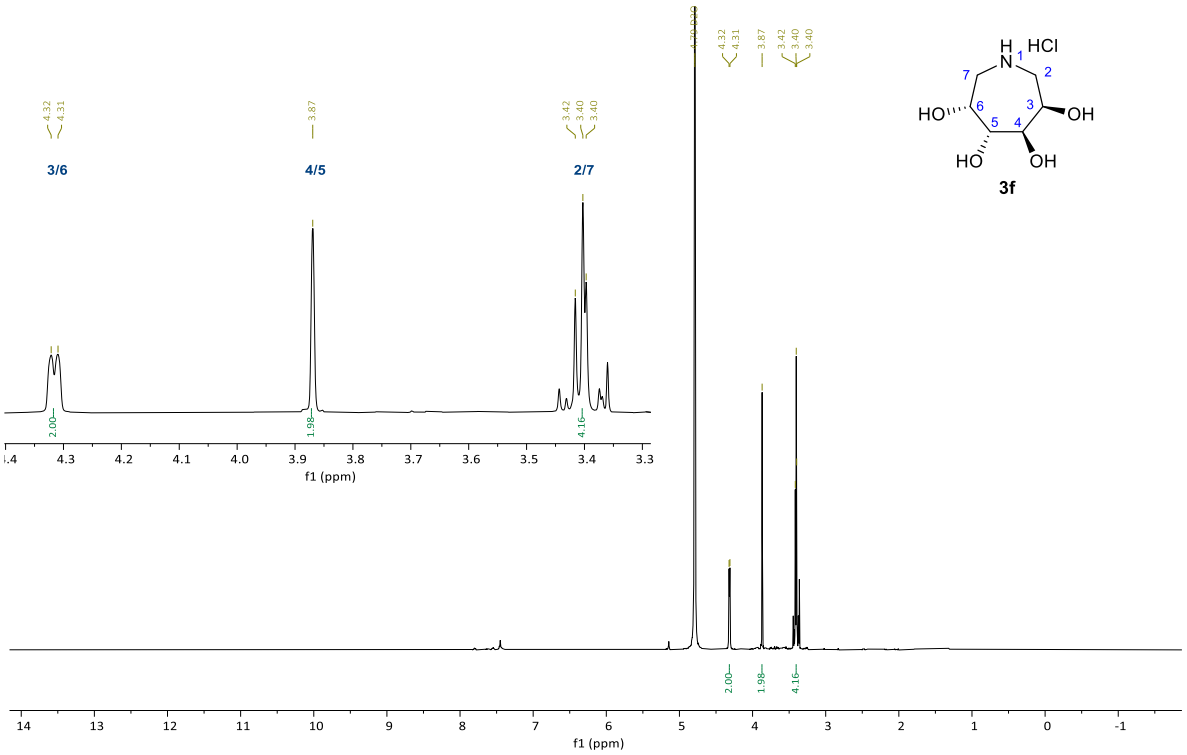
^{1}H-^{13}C HSQC NMR (D_2O) 2-deoxy-D-*galacto* **3e**



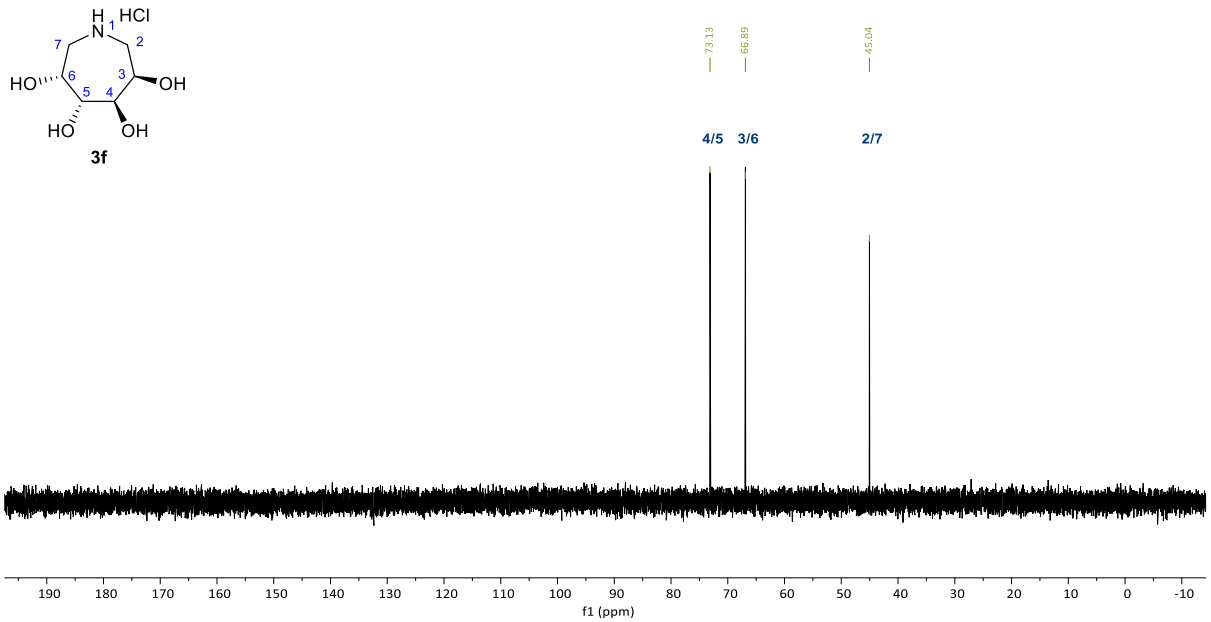
^{1}H-^{13}C HMBC NMR (D_2O) 2-deoxy-D-*galacto* **3e**



^{1}H NMR (500 MHz, D_2O) D-*manno* **3f**

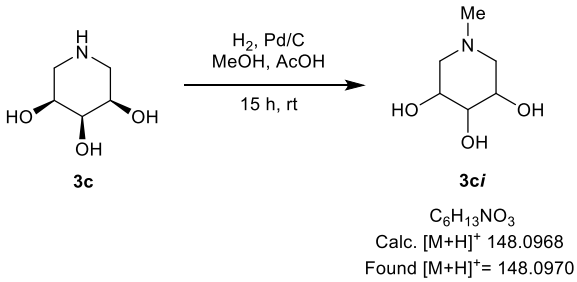


^{13}C NMR (125 MHz, D_2O) D-*manno* **3f**



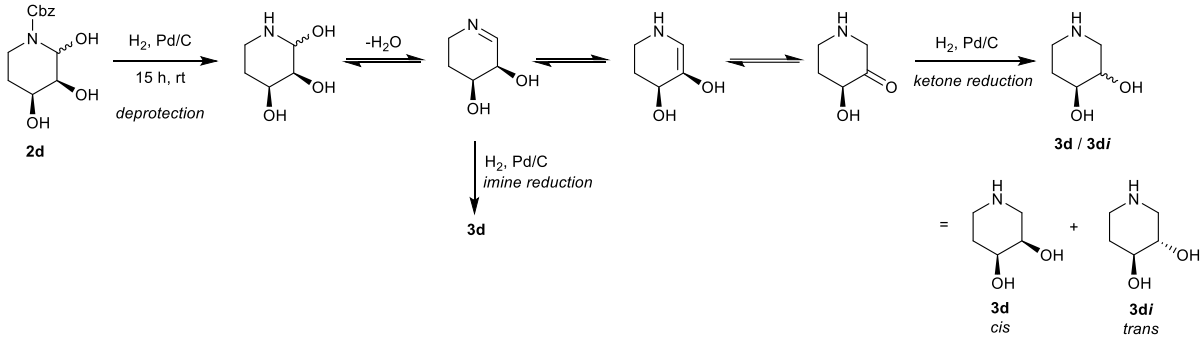
VIII. An additional discussion relating to side product formation

The *N*-methylated iminosugar is the proposed side product of catalytic hydrogenation to deprotect and reduce **2c** to **3c.** HRMS data for the compound was obtained. Chemical shifts and integration are characteristic of **3c*i***, but the absolute stereochemistry could not be ascertained.[3,10]



Scheme S1. The proposed side product arising from *N*-methylation of **3c**.

There is a side product **3d*i*** in the final NMR spectra of **3d,** the 2D NMR characterisation indicated a compound with resonances extremely similar to **3d** yet possessing distinct chemical shifts. The LC/MS data did not reveal the identity of **3d*i*.** Through surveying the literature, we propose that the imine-enaminol tautomerisation pathway first elucidated by Hans Paulsen in 1966[11] occurs to racemise the C3-OH of **3d.** Under the reducing conditions of catalytic hydrogenation, the trans diastereomer (3*R*,4*R*)-piperidine-3,4-diol **3d*i*** is formed as a minor product by non-stereoselective ketone reduction. This is outlined in Scheme S2, and characterisation of **3d*i*** by NMR analysis is confirmed by literature.[12]



Scheme S2. The imine-enaminol tautomerization pathway.

IX. References

1. J. Kuska, F. Taday, K. Yeow, J. Ryan and E. O'Reilly, *Catal. Sci. Technol.*, 2021, **11**, 4327–4331.
2. X. Garrabou, J. A. Castillo, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, and P. Clapés, *Angew. Chem. Int. Ed.*, 2009, **48**, 5521-5525.
3. T. Tschamber, F. Backenstrass, M. Neuburger, M. Zehnder, and J. Strieth, *Tetrahedron*, 1994, **50**, 1135-1152.

4. A. Szekrenyi, A. Soler, X. Garrabou, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, J. Bujons and P. Clapés, *Chem. – Eur. J.*, 2014, **20**, 12572-12583.
5. Y. Igarashi, M. Ichikawa, and Y. Ichikawa, *Bioorg. Med. Chem. Lett.*, 1996, **6**, 553-558.
6. L. Chen, D. P. Dumas, and C. H. Wong, *J. Am. Chem. Soc.*, 1992, **114**, 741-748.
7. P. R. Andreana, T. Sanders, A. Janczuk, J. I. Warrick, and P. G. Wang, *Tetrahedron lett.*, 2002, **43**, 6525-6528.
8. B. B. Lohray, Y. Jayamma, and M. Chatterjee, *J. Org. Chem.*, 1995, **60**, 5958-5960.
9. L. Chen, D. P. Dumas, and C. H. Wong, *J. Am. Chem. Soc,* 1992, **114**, 741-748.
10. S. Front, E. Gallienne, J. Charollais-Thoenig, S. Demotz, and O. R. Martin, *ChemMedChem*, 2016, **11**, 133.
11. a) H. Paulsen, *Angew. Chem. Int. Ed.*, 1966, **5**, 495-510; b) D. Dhara, A. Dhara, J. Bennett, and P. V. Murphy, *The Chemical Record,* 2021, **21**, 11, 2958-2979.
12. Y. Ichikawa, Y. Igarashi, M. Ichikawa, and Y. Suhara, *J. Am. Chem. Soc.*,1998, **120**, 13, 3007-3018.