Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2010

Supplementary Information

Fluoroglycoproteins: ready chemical site-selective incorporation of 2-fluorosugars into proteins

Omar Boutureira, François D'Hooge, Marta Fernández-González, Gonçalo J. L. Bernardes, Macarena Sánchez-Navarro, Julia R. Koeppe and Benjamin G. Davis*

Chemistry Research Laboratory
Department of Chemistry, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, UK

Ben.Davis@chem.ox.ac.uk

1. General Considerations

Melting points (m.p.) were recorded on a Leica Galen III hot stage microscope equipped with a Testo 720 thermocouple probe and are uncorrected. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on a Bruker AVII500 (500 MHz) spectrometer, as indicated. Carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded on a Bruker AVII500 (125 MHz) spectrometer, as indicated. Fluorine nuclear magnetic resonance (^{19}F NMR) spectra were recorded on a Bruker AVII500 (470.4 MHz) spectrometer, as indicated. NMR Spectra were fully assigned using COSY, HSQC, HMBC, and NOESY. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard (^{1}H NMR: $CDCl_3$ = 7.26, CD_3OD = 4.87; DMSO-d_6 = 2.50 and ^{13}C NMR: $CDCl_3$ = 77.0; CD_3OD = 49.0; DMSO-d_6 = 39.5) and $CFCl_3$ as external standard for ^{19}F NMR. Coupling constants (J) are reported in Hz with

the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, and a = apparent.

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using thin films on NaCl plates for liquids and oils and KBr discs for solids and crystals. Absorption maxima (υ_{max}) are reported in wavenumbers (cm^{-1}).

Low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer using electrospray ionization (ESI) and high resolution mass spectra (HRMS) were recorded on a Bruker MicroTOF ESI mass spectrometer. Nominal and exact *m/z* values are reported in Daltons.

Optical rotations were measured on a Perkin–Elmer 241 polarimeter with a path length of 1.0 dm and are reported with implied units of 10^{-1} deg cm^2 g^{-1}. Concentrations (c) are given in g/100 ml.

Thin layer chromatography (TLC) was carried out using Merck aluminium backed sheets coated with 60F254 silica gel. Visualization of the silica plates was achieved using a UV lamp (λ_{max} = 254 nm), and/or ammonium molybdate (5% in 2 M H_2SO_4), and/or potassium permanganate (5% $KMnO_4$ in 1 M NaOH with 5% potassium carbonate). Flash column chromatography was carried out using BDH 40–63 μm silica gel (VWR). Mobile phases are reported in relative composition (*e.g.* 1:2:4 H_2O/*i*PrOH/EtOAc).

SDS-PAGE electrophoresis was carried out using Invitrogen system (XCell-Cell, NuPAGE Novex Bis-Tris gel, NuPAGE MES running buffer). Protein molecular weights were approximated by comparison to a protein marker (Perfect Protein Markers 15–150 kDa from Novagen). Gels were visualised by coomassie staining (Instant Blue from Expedeon).

Anhydrous solvents were purchased from Fluka or Acros. Triethylamine was stored over NaOH pellets. All other solvents were used as supplied (Analytical or HPLC grade), without prior purification. Distilled water was used for chemical reactions and Milli–QR purified water for protein manipulations. Reagents were purchased from

Aldrich and used as supplied, unless otherwise indicated. 'Petrol' refers to the fraction of light petroleum ether boiling in the range 40–60 ºC. All reactions using anhydrous conditions were performed using flame-dried apparatus under an atmosphere of argon or nitrogen. Brine refers to a saturated solution of sodium chloride. Anhydrous magnesium sulfate ($MgSO_4$) was used as drying agents after reaction workup, as indicated.

DOWEX 50WX8 (H^+ form) was conditioned as follows: 100 g of the commercial resin was placed in a 500 mL sintered filter funnel and allowed to swell with 200 mL of acetone for 5 minutes. The solvent was removed by suction and the resin was washed successively with 800 mL of acetone, 500 mL methanol, 500 mL 5 M HCl, and then 1 L of water or until the pH of filtrate was ~ 7, as indicated by pH paper. The resin was partially dried on the filter and then stored and used as needed.

In addition to those specified above, the following abbreviations, designations, and formulas are used throughout the Supporting Information:

MeOH = methanol	H_2O = water	Et_2O = diethyl ether
EtOAc = ethyl acetate	CH_2Cl_2 = dichloromethane	DMF = dimethylformamide
*i*PrOH = isopropanol	PBu_3 = tributylphosphine	Et_3N = triethylamine
Boc = *tert*-butoxycarbonyl	K_2CO_3 = potassium carbonate	Cs_2CO_3 = caesium carbonate
$NaHCO_3$ = sodium bicarbonate	NaOH = sodium hydroxide	NH_4Cl = ammonium chloride
NH_4OH = ammonium hydroxide	aq. = aqueous	sat. = saturated
TFA = trifluoroacetic acid	TIS = triisopropyl silane	EDT = 1,2-ethanedithiol

Solid Phase Peptide Synthesis

Peptides were synthesized by means of Fmoc-chemistry on Rink amide MBHA-polystyrene resin [1% divinyl benzene, GL Biochem] using a microwave assisted Liberty CEM peptide synthesizer. Side-chain protection for the amino acids was *t*Bu for Asp, Ser, Thr; Trt for Asn, Cys, Gln, His and Boc for Lys.

Standard steps included:

Fmoc deprotection: resin treated with 20% piperidine in DMF (10 mL) and irradiated at 60 W for 35 s followed by a second portion of 20% piperidine/DMF (10 mL) and irradiate at 65 W for 180 s (max. temp. 77 ºC).

Coupling: resin treated with protected amino acid (5 mL of a 0.2 M solution in DMF), HBTU (2 mL of a 0.45 M solution in DMF), DIEA (1 mL of a 2 M solution in NMP) and irradiated at 25 W for 300 s (max. temp. 75 ºC). "His/Cys cycle" uses a modified method in which irradiation was carried out at 25 W for 11 min (max. temp. 53 ºC).

Capping: resin treated with 20% acetic anhydride in DMF (10 mL) and irradiated at 50 W for 60 s. (max. Temp. 65 ºC).

Cleavage and side-chain deprotection: resin treated with TFA/TIS/H_2O/EDT (94:1:2.5:2.5, 10 mL) and irradiated at 15 W for 18 min (max. temp. 38 ºC).

Work up: the cleavage peptide solution was precipitated with iced Et_2O (20 mL), centrifuged (3000 rpm, 5 min) and decanted. New washes with iced Et_2O (20 mL) were repeated twice. The pellet was then dissolved in water and lyophilized.

RP-HPLC analysis

RP-HPLC was performed on a Dionex UltiMate 3000 instrument coupled to a UltiMate 3000 variable wavelength detector, using Chromeleon software (version 6.80). Analytical analyses were carried out using a Phenomenex Synergi 4u Fusion-RP 80A (100 × 4.6mm) column. Preparative purifications were carried out using a Phenomenex Synergi 4u Fusion-RP 80A (100 × 21.20mm) column. Water (solvent A) and acetonitrile (solvent B), each containing 0.1% TFA, were used as the mobile phase in all the cases.

Protein Mass Spectrometry

Liquid chromatography-mass spectrometry (LC–MS) was performed on a Micromass LCT (ESITOF–MS) coupled to a Waters Alliance 2790 HPLC using a Phenomenex Jupiter C4 column (250 × 4.6 mm × 5μm). Water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as the mobile phase at a flow rate of 1.0 mL min^{-1}. The gradient was programmed as follows: 95% A (5 min isocratic) to 100% B after 15 min then isocratic for 5 min. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 35 V for β-galactosidase (*Ss*βG). Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L

hr^{-1}. Spectra were calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin obtained at a cone voltage of 25 V. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.0 from Waters) according to the manufacturer's instructions.

Materials

(*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)hex-5-ynoic acid **4** was prepared as previously described.[1] Precursors 3,4,6-tri-*O*-acetyl-2-deoxy-2-fluoro-β-D-gluco, galacto, and mannopyranosyl azides were prepared following minor modification of the methodology described by Albert and Dax.[2] *Ss*βG Hpg1 Hpg43 C344S 9Ile was cloned and expressed as previously described.[1] Tristriazole ligand triethyl 2,2',2''-(4,4',4''-nitrilotris(methylene)tris(1H-1,2,3-triazole-4,1-diyl))triacetate was prepared as previously described.[3]

Radiochemistry

[^{18}F]Fluoride was produced by PETNet Solutions at Mount Vernon Hospital (UK) in a tantalum target via the $^{18}O(p,n)^{18}F$ reaction and delivered as [^{18}F]fluoride in [^{18}O]water (1–4 GBq, 1–5 ml). Radio-high-performance liquid chromatography (HPLC) analyses were performed using a Gilson 322 or a Dionex UltiMate 3000 HPLC system equipped with a NaI/photomultiplier tube (PMT)-radiodetector and an ultraviolet (UV) detector. Radio-thin layer chromatography (TLC) was performed on Macherey-Nagel Polygram Silica Plates and analysed using a plastic scintillator/PMT detector. Radiochemical yields are reported non-decay corrected. Radiolabelling was performed with a Scintomics device. The [^{18}F]fluoride solution was passed through a strong anion exchanger column (QMA cartridge, Chromafix 30 PS HCO3, ^{18}F separation cartridge) and [^{18}F]Fluoride adsorbed on the resin was eluted into a reactor with a solution of K_2CO_3 (3.0 mg) and Kryptofix 222 (15.0 mg) in acetonitrile:water (1mL, 8:2). The complex was dried under a N_2 stream at 125°C and co-evaporated to dryness with acetonitrile (2x200 μL). The resulting dry complex of K^{18}F/Kryptofix 222 was used for further nucleophilic ^{18}F-fluorination.

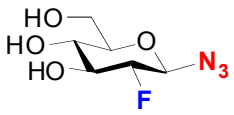
2. Experimental Section

2.1. Preparation of 2-Deoxy-2-fluoroglycosyl Azides 1–3

General procedure for the synthesis of 2-deoxy-2-fluoroglycosyl azides

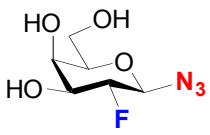
The corresponding glycopyranosyl azide (1 mmol) was treated with 0.1 M NaOMe in MeOH (7 mL/mmol) at room temperature. The reaction mixture was stirred at the same temperature for 15 minutes and neutralized with Dowex (H^+ 50WX8-200). The ion exchanger was filtered off and washed with MeOH. The resulting solution was concentrated under reduced pressure and the residue purified by chromatographic techniques.

2-Deoxy-2-fluoro-β-D-glucopyranosyl azide (1)



The title compound was prepared following the general procedure above, starting from 3,4,6-tri-*O*-acetyl-2-deoxy-2-fluoro-β-D-glucopyranosyl azide[4] (25 mg, 0.075 mmol) and 0.1 M NaOMe in MeOH (525 µL). After concentration under reduced pressure crude product (15.5 mg, 100%) was obtained as a white foam. R_f (7:1 EtOAc/MeOH): 0.43; $[\alpha]_D^{20}$: –9.5 (0.9, MeOH); ^{1}H NMR (CD_3OD, 500 MHz) δ in ppm: 4.83 (dd, 1H, $J_{1,2}$ = 8.8 Hz, $J_{1,F}$ = 2.3 Hz, H-1), 3.93 (ddd, 1H, $J_{2,F}$ = 51.6 Hz, $J_{1,2}$ = $J_{2,3}$ = 8.8 Hz, H-2), 3.88 (dd, $J_{6a,b}$ = 12 Hz, $J_{5,6a}$ = 2.2 Hz, H-6a), 3.71–3.60 (m, 2H, H-3,6b), 3.45–3.33 (m, 2H, H-4,5); ^{13}C NMR (CD_3OD, 125.8 MHz) δ in ppm: 93.3(d, $J_{2,F}$ = 186.9 Hz, C-2), 89.2 (d, $J_{1,F}$ = 22.9 Hz, C-1), 80.4 (C-5), 76.4 (d, $J_{3,F}$ = 16.2 Hz, C-3), 70.9 (d, $J_{4,F}$ = 8.6 Hz, C-4), 62.4 (C-6); ^{19}F NMR (CD_3OD, 470.4 MHz) δ in ppm: –199.4 (ddd, $J_{2,F}$ = 51.6 Hz, $J_{3,F}$ = 14.9 Hz, $J_{1,F}$ = 2.3 Hz, F-2); FT–IR (KBr) ν in cm^{-1}: 3358, 3075, 2733, 2687, 2105; HRMS (TOF ES–) for ($M–H^-$) $C_6H_9FN_3O_4$ (*m/z*): calc. 206.0583; found 206.0583; spectroscopic data was identical to that previously reported.[4]

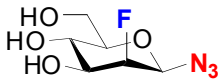
2-Deoxy-2-fluoro-β-D-galactopyranosyl azide (2)



The title compound was prepared following the general procedure above, starting from 3,4,6-tri-*O*-acetyl-2-deoxy-2-fluoro-β-D-galactopyranosyl azide[5] (27.5 mg, 0.083 mmol) and 0.1 M NaOMe in MeOH (578 µL). After concentration under reduced pressure crude product (17.1 mg, 100%) was obtained as a white solid. R_f (7:1 EtOAc/MeOH): 0.43; mp: 92–94 ºC; $[\alpha]_D^{20}$: +2.9 (0.53, MeOH); ^{1}H NMR (CD_3OD, 500 MHz) δ in ppm: 4.74 (dd, 1H, $J_{1,2}$ =

8.4 Hz, $J_{1,F}$ = 3.7 Hz, H-1), 4.29 (ddd, 1H, $J_{2,F}$ = 52.0 Hz, $J_{1,2}$ = $J_{2,3}$ = 8.4 Hz, H-2), 3.92 (m, H-4), 3.80–3.66 (m, 4H, H-3, 5, 6a,b); ^{13}C NMR (CD_3OD, 125.8 MHz) δ in ppm: 92.5 (d, $J_{2,F}$ = 183.1 Hz, C-2), 89.6 (d, $J_{1,F}$ = 23.8 Hz, C-1), 79.2 (C-5), 73.3 (d, $J_{3,F}$ = 17.2 Hz, C-3), 71.0 (d, $J_{4,F}$ = 8.6 Hz, C-4), 62.4 (C-6); ^{19}F NMR (CD_3OD, 470.4 MHz) δ in ppm: –207.9 (dddd, $J_{2,F}$ = 52.0 Hz, $J_{3,F}$ = 13.9 Hz, $J_{1,F}$ = 3.7 Hz, $J_{4,F}$ = 3.5 Hz, F-2); FT–IR (KBr) ν in cm^{-1}: 3356, 2925, 2122, 1373, 1251, 1071; HRMS (TOF ES–) for (M–H$^-$) $C_6H_9FN_3O_4$ (*m/z*): calc. 206.0583; found 206.0582.

2-Deoxy-2-fluoro-β-D-mannopyranosyl azide (3)



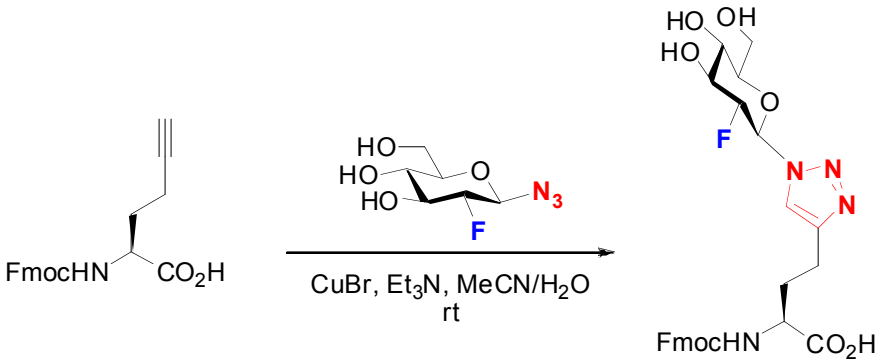
The title compound was prepared following the general procedure above, starting from 3,4,6-tri-*O*-acetyl-2-deoxy-2-fluoro-β-D-mannopyranosyl azide (89 mg, 0.267 mmol) and 0.1 M NaOMe in MeOH (1.9 mL). After concentration under reduced pressure the residue was purified by column chromatography (7:1 EtOAc/MeOH) to afford the title compound (19.9 mg, 36%) as a white hygroscopic solid. R_f (7:1 EtOAc/MeOH): 0.33; $[\alpha]_D^{20}$: +3.2 (0.60, MeOH); ^{1}H NMR (CD_3OD, 500 MHz) δ in ppm: 4.71 (d, 1H, $J_{2,F}$ = 51.9 Hz), 4.63 (d, 1H, $J_{1,F}$ = 22.2 Hz, H-1), 3.93 (d, $J_{6a,b}$ = 12.1 Hz, H-6a), 3.73 (dd, $J_{6a,b}$ = 12.1 Hz, $J_{5,6b}$ = 6.3 Hz, H-6b), 3.62–3.50 (m, 2H, H-3,4), 3.38 (m, 1H, H-5); ^{13}C NMR (CD_3OD, 125.8 MHz) δ in ppm: 93.2 (d, $J_{2,F}$ = 185.1 Hz, C-2), 86.6 (d, $J_{1,F}$ = 15.9 Hz, C-1), 80.8 (C-5), 74.0 (d, $J_{3,F}$ = 17.5 Hz, C-3), 68.2 (d, $J_{4,F}$ = 1.9 Hz, C-4), 62.7 (C-6); ^{19}F NMR (CD_3OD, 470.4 MHz) δ in ppm: –220.2 (m, F-2); FT–IR (KBr) ν in cm^{-1}: 3357, 2920, 2122, 1071; HRMS (TOF ES–) for (M–H$^-$) $C_6H_9FN_3O_4$ (*m/z*): calc. 206.0583; found 206.0584.

2.2. Preparation of Glycoamino Acids 5–7

General procedure for the synthesis of triazole-linked 2-deoxy-2-fluoro-glycoamino acids

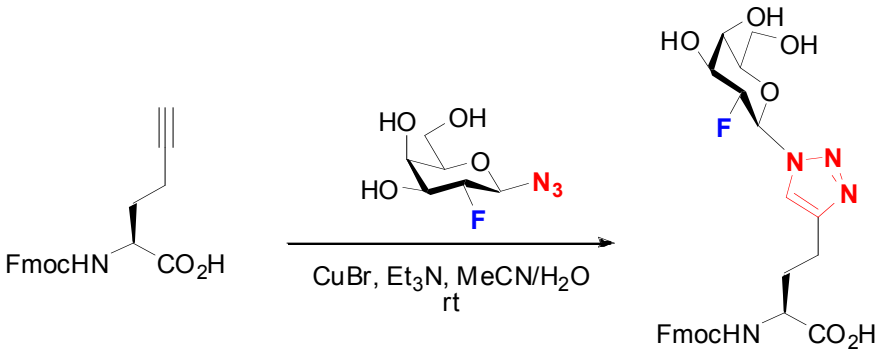
To a deoxygenated (argon was bubbled for 20 min) solution of 2-deoxy-2-fluoro-β-D-glycopyranosyl azides **1**–**3** (1.1 mmol) and (*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)hex-5-ynoic acid **4** (1 mmol) in CH_3CN (35 mL), copper(I) bromide (0.2 mL of 10 mg/mL deoxygenated CH_3CN, 0.25 mmol) was added, followed by Et_3N (1.8 mmol) and water (2.8 mL). The reaction was monitored by TLC. The mixture was concentrated *in vacuo* and purified by chromatographic techniques.

(*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)-4-{1-(2-deoxy-2-fluoro-β-D-glucopyranosyl)[1,2,3]-triazol-4-yl}butanoic acid (5)



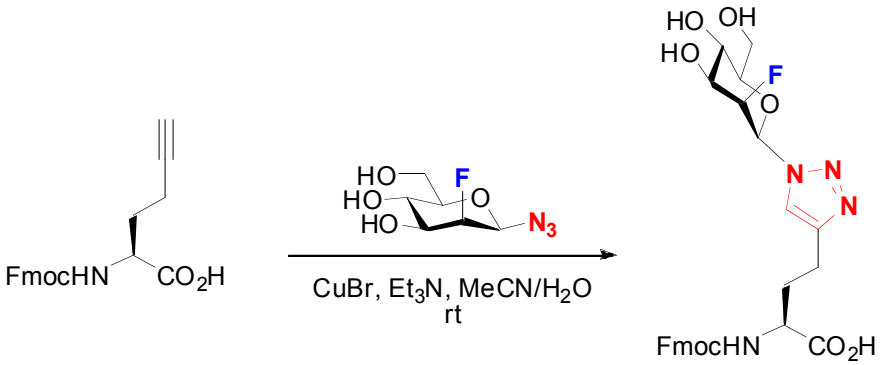
The title compound was prepared following the general procedure above, starting from 2-deoxy-2-fluoro-β-D-glucopyranosyl azide **1** (13 mg, 0.063 mmol), (*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)hex-5-ynoic acid **4** (20 mg, 0.057 mmol) in CH_3CN (2 mL), copper(I) bromide (0.2 mL of 10 mg/mL deoxygenated CH_3CN, 0.014 mmol), Et_3N (14 μL, 0.100 mmol) and water (0.16 mL). After 18 h, TLC (65:35 EtOAc/MeOH) indicated the formation of a product (R_f 0.19) and complete consumption of the amino acid starting material (Rf 0.44). >95% conversion calculated by ^{19}F NMR analysis of the crude. The mixture was concentrated *in vacuo* and purified by flash column chromatography (7:3 to 3:2 EtOAc/MeOH), and the isolated compound was dissolved in water and lyophilized to afford the triazole **5** (14 mg, 44%) as a white solid. $[\alpha]_D^{20}$: +5.8 (0.03, MeOH); ^1H NMR (CD_3OD, 500 MHz) δ in ppm: 8.05 (s, 1H, $ArH_{triazole}$), 7.79–7.66 (m, 4H, ArH), 7.39–7.28 (m, 4H, ArH), 5.87 (m, 1H, H-1), 4.84–4.73 (m, 1H, H-2), 4.39–4.32 (m, 2H, $CH_{2\ Fmoc}$), 4.22 (t, J = 6.7 Hz, 1H, CH_{Fmoc}), 4.08 (m, 1H, Hα), 3.88–3.80 (m, 2H, H-3, H-6b), 3.69 (dd, $J_{6a,b}$ = 12.3 Hz, $J_{5,6}$ = 5.3 Hz, 1H, H-6a), 3.63–3.59 (m, 1H, H-5), 3.54 (t, J = 9.5 Hz, 1H, H-4), 2.78 (t, J = 7.6 Hz, 2H, Hγ, Hγ'), 2.19 (m, 1H, Hβ), 2.01 (m, 1H, Hβ'); ^{13}C NMR (CD_3OD, 125 MHz) δ in ppm: 158.3 (OC(O)N), 149.1, 145.5, 145.3, 145.65, 142.63 (4×Ar-C, $C{=}C_{triazole}$), 128.8, 128.2, 126.3, 120.9 (8×Ar-H), 122.7 ($CH_{triazole}$), 92.2 (d, $J_{2,F}$ = 188.1 Hz, C-2), 86.5 (d, $J_{1,F}$ = 24.2 Hz, C-1), 81.2 (C-5), 76.5 (d, $J_{3,F}$ = 16.7 Hz, C-3), 70.7 (d, $J_{4,F}$ = 7.9 Hz, C-4), 67.8 ($CH_{2\ Fmoc}$), 62.2 (C-6), 57.4 (Cα), 48.3 (CH_{Fmoc}), 33.8 (Cβ), 22.9 (Cγ); $^{19}F\{^1H\}$ NMR (CD_3OD, 470.4 MHz) δ in ppm: –200.5 (s, F-2).; HRMS (ES+) for ($M+Na^+$) $C_{27}H_{29}N_4NaO_8$ (*m/z*): calc. 579.1862; found 579.1860.

(*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)-4-{1-(2-deoxy-2-fluoro-β-D-galactopyranosyl)[1,2,3]-triazol-4-yl}butanoic acid (6)



The title compound was prepared following the general procedure above, starting from 2-deoxy-2-fluoro-β-D-galactopyranosyl azide **2** (11 mg, 0.053 mmol), (*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)hex-5-ynoic acid **4** (16 mg, 0.046 mmol) in CH_3CN (1 mL), copper(I) bromide (0.2 mL of 6 mg/mL deoxygenated CH_3CN, 0.008 mmol), Et_3N (12 μL, 0.086 mmol) and water (0.11 mL). After 2 h, fresh copper(I) bromide (0.2 mL, 0.008 mmol) was added. After 18 h, TLC (3:2 EtOAc/MeOH) indicated the formation of a product (R_f 0.18) and almost complete consumption of the amino acid starting material (R_f 0.44). >95% conversion calculated by ^{19}F NMR analysis of the crude. The mixture was concentrated *in vacuo* and purified by flash column chromatography (3:1 to 3:2 EtOAc/MeOH), and the isolated compound was dissolved in water and lyophilized to afford the triazole **6** (8.4 mg, 33%) as a white solid. $[\alpha]_D^{20}$: +2.0 (0.3, MeOH); ^1H NMR (CD_3OD, 500 MHz) δ in ppm: 8.55 (br s, 1H, NH), 8.06 (s, 1H, $ArH_{triazole}$), 7.79–7.67 (m, 4H, ArH), 7.39–7.29 (m, 4H, ArH), 5.82 (dd, $J_{1,2}$ = 8.9 Hz, $J_{1,F}$ = 3.2 Hz, 1H, H-1), 5.06 (dt, $J_{2,F}$ = 51.6 Hz, $J_{1,2}$ = 9.1 Hz, $J_{2,3}$ = 9.1 Hz, 1H, H-2), 4.40–4.32 (m, 2H, $CH_{2\ Fmoc}$), 4.23 (t, J = 6.7 Hz, 1H, CH_{Fmoc}), 4.09 (m, 1H, Hα), 4.03 (m, 1H, H-4), 3.99–3.96 (m, 1H, H-3), 3.89 (m, 1H, H-5), 3.78 (dd, $J_{6a,b}$ = 11.5 Hz, $J_{5,6}$ = 6.9 Hz, 1H, H-6a), 3.72 (dd, $J_{6a,b}$ = 11.6 Hz, $J_{5,6b}$ = 4.8 Hz, 1H, H-6b), 2.78 (t, J = 7.8 Hz, 2H, Hγ, Hγ'), 2.19 (m, 1H, Hβ), 2.01 (m, 1H, Hβ'); ^{13}C NMR (CD_3OD, 125 MHz) δ in ppm: 170.3 (C=O), 158.3 (OC(O)N), 149.3, 145.6, 145.3, 145.65, 142.63 (4×Ar-C, $C{=}C_{triazole}$), 128.8, 128.2, 126.3, 120.9 (8×Ar-H), 122.3 ($CH_{triazole}$), 91.6 (d, $J_{2,F}$ = 184.2 Hz, C-2), 87.1 (d, $J_{1,F}$ = 25.0 Hz, C-1), 80.1 (C-5), 73.5 (d, $J_{3,F}$ = 16.8 Hz, C-3), 71.0 (d, $J_{4,F}$ = 8.5 Hz, C-4), 67.8 ($CH_{2\ Fmoc}$), 62.2 (C-6), 57.5 (Cα), 48.4 (CH_{Fmoc}), 33.9 (Cβ), 22.9 (Cγ); ^{19}F NMR (CD_3OD, 470.4 MHz) δ in ppm: –208.9 (dddd, $J_{2,F}$ = 51.4 Hz, $J_{3,F}$ = 13.9 Hz, $J_{1,F}$ = 2.9 Hz, $J_{4,F}$ = 2.9 Hz, F-2); HRMS (ES+) for ($M+Na^+$) $C_{27}H_{29}N_4NaO_8$ (*m/z*): calc. 579.1862; found 579.1860.

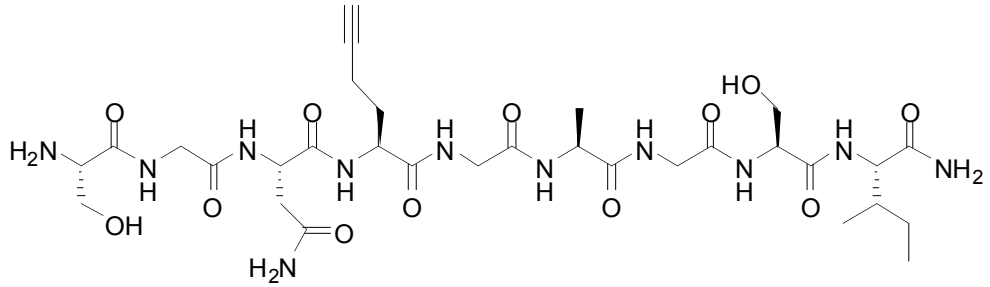
(*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)-4-{1-(2-deoxy-2-fluoro-β-D-mannopyranosyl)[1,2,3]-triazol-4-yl}butanoic acid (7)



The title compound was prepared following the general procedure above, starting from 2-deoxy-2-fluoro-β-D-mannopyranosyl azide **3** (11 mg, 0.053 mmol), (*S*)-2-(9H-fluoren-9-ylmethoxycarbonylamino)hex-5-ynoic acid **4** (17 mg, 0.048 mmol) in CH_3CN (2 mL), copper(I) bromide (0.35 mL of 5 mg/mL deoxygenated CH_3CN, 0.012 mmol), Et_3N (12 μL, 0.086 mmol) and water (0.13 mL). After 16.5 h, fresh copper(I) bromide (0.15 mL, 0.005 mmol) and Et_3N (4 μL, 0.028 mmol) were added. After 44 h, TLC (3:2 EtOAc/MeOH) indicated the formation of a product (R_f 0.18) and almost complete consumption of the amino acid starting material (R_f 0.44). >95% conversion calculated by ^{19}F NMR analysis of the crude. The mixture was concentrated *in vacuo* and purified by flash column chromatography (3:1 to 3:2 EtOAc/MeOH), and the isolated compound was dissolved in water and lyophilized to afford the triazole **7** (5.9 mg, 21%) as a white solid. $[\alpha]_D^{20}$: –1.4 (0.10, MeOH); ^1H NMR (CD_3OD, 500 MHz) δ in ppm: 7.97 (s, 1H, $ArH_{triazole}$), 7.77–7.64 (m, 4H, ArH), 7.37–7.27 (m, 4H, ArH), 6.06 (d, $J_{1,F}$ = 22.1 Hz, 1H, H-1), 4.92 (d, $J_{2,F}$ = 50.8 Hz, 1H, H-2), 4.32 (m, 2H, $CH_{2\ Fmoc}$), 4.20 (m, 1H, CH_{Fmoc}), 4.07 (m, 1H, Hα), 3.91–3.71 (m, 4H, H-3, H-4, H-6a,b), 3.54 (m, 1H, H-5), 2.74 (m, 2H, Hγ, Hγ'), 2.16 (m, 1H, Hβ), 1.99 (m, 1H, Hβ'); ^{13}C NMR (CD_3OD, 125 MHz) δ in ppm: 158.2 (OC(O)N), 148.8, 145.7, 145.4, 142.8 (4×Ar-C, $C{=}C_{triazole}$), 128.8, 128.3, 126.5, 126.4, 122.9, 121.0 (8×Ar-H, $CH_{triazole}$), 91.8 (d, $J_{2,F}$ = 186.5 Hz, C-2), 86.4 (d, $J_{1,F}$ = 15.8 Hz, C-1), 81.46, 81.44 (C-5), 73.7 (d, $J_{3,F}$ = 17.7 Hz, C-3), 67.9, 67.8 (C-4, $CH_{2\ Fmoc}$), 62.51 (C-6), 57.4 (Cα), 48.3 (CH_{Fmoc}), 34.2 (Cβ), 23.0 (Cγ); ^{19}F NMR (CD_3OD, 470.4 MHz) δ in ppm: –218.3 (m, F-2); HRMS (ES+) for ($M+Na^+$) $C_{27}H_{29}N_4NaO_8$ (*m/z*): calc. 579.1862; found 579.1857.

2.3. Preparation of Glycopeptides 9 and 10

H_2N-Ser-Gly-Asn-Hpg-Gly-Ala-Gly-Ser-Ile-C(O)NH_2 (8)

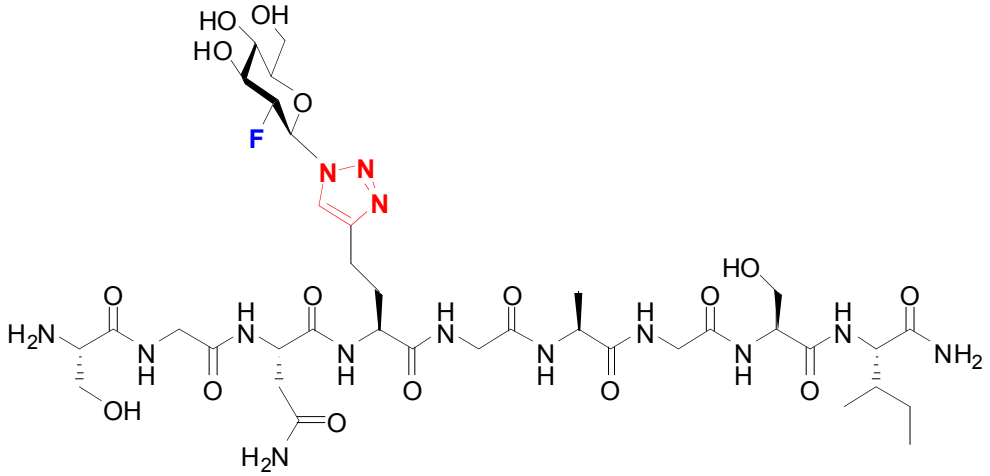


SBL peptide fragment (153–161; Ser-Gly-Asn-Hpg-Gly-Ala-Gly-Ser-Ile) was synthesized by means of Fmoc-chemistry on Rink amide MBHA-polystyrene resin [1% divinyl benzene, GL Biochem] using a microwave assisted Liberty CEM peptide synthesizer. HRMS (ES+) for ($M+Na^+$) $C_{31}H_{51}N_{11}NaO_{12}$ (*m/z*): calc. 792.3611; found 792.3592.

General procedure for the synthesis of triazole-linked 2-deoxy-2-fluoro-glycopeptides

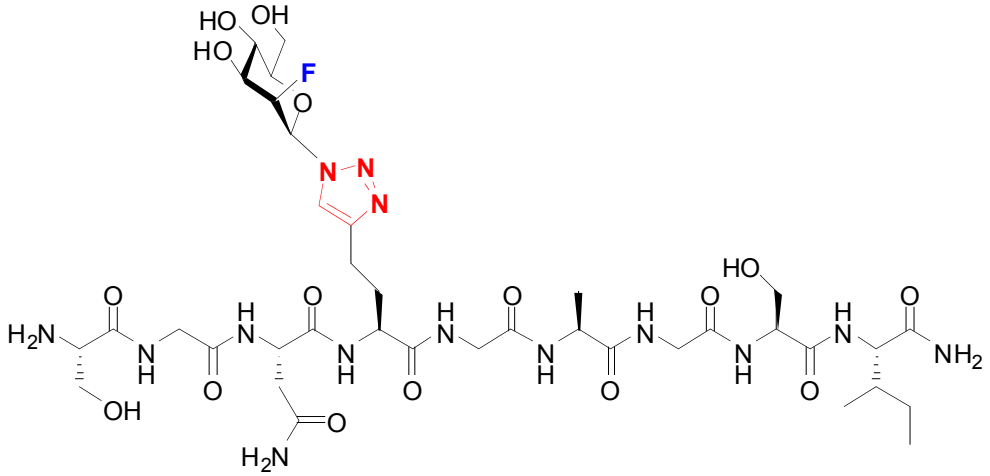
A freshly prepared solution of copper(I) bromide (99.999%) in acetonitrile (67 μL/μmol peptide of 10 mg/mL) was mixed with a solution of tris-triazolyl amine ligand 2,2',2''-(4,4',4''-nitrilotris(methylene)tris(1H-1,2,3-triazole-4,1-diyl))triacetate in acetonitrile (25 μL/μmol peptide of 127 mg/mL). The preformed Cu-complex solution (92 μL/μmol peptide) was added to a solution of peptide **8** (1 μmol) and 2-deoxy-2-fluoroglycosyl azides **1** and **3** (1.1 μmol) in 50 mM sodium phosphate buffer (229 μL/μmol peptide, pH 8.2) and stirred on a lab rotisserie for 2 h at room temperature. The mixture was then concentrated *in vacuo* and purified by RP-HPLC.

2FGlc-triazole-peptide (9)



The title compounds were prepared following the general procedure above, starting from peptide **8** (3.7 mg, 4.8 μmol), 2-deoxy-2-fluoro-β-D-glucopyranosyl azide **1** (1.1 mg, 5.3 μmol) and preformed Cu-complex solution (440 μL) in 50 mM sodium phosphate buffer (1.1 mL, pH 8.2). After 2 h, the crude product was concentrated (>95% conversion calculated by ^{19}F NMR) and purified by RP-HPLC [Phenomenex Synergi 4u Fusion-RP 80A (100 × 21.20mm) column; flow rate was 23 mL/min; gradient run, 0–1 min 0%B, 1–16 min linear gradient 0 to 60%B, 16–19 min linear gradient 60 to 100%B, 19–20 min linear gradient 100 to 0%B; t_R: 5.5 min; UV monitoring 214 nm], neutralized with $(NH_4)_2CO_3$, and lyophilized to afford **9** (0.62 mg, 13%) as a white hygroscopic solid. HRMS (ES+) for ($M+H^+$) $C_{37}H_{62}FN_{14}O_{16}$ (*m/z*): calc. 977.4447; found 977.4430.

2FMan-triazole-peptide (10)



The title compounds were prepared following the general procedure above, starting from peptide **8** (8.4 mg, 11 μmol), 2-deoxy-2-fluoro-β-D-mannopyranosyl azide **3** (2.5

mg, 12 μmol) and preformed Cu-complex solution (1 mL) in 50 mM sodium phosphate buffer (2.6 mL, pH 8.2). After 2 h, the crude product was concentrated (>95% conversion calculated by ^{19}F NMR) and purified by RP-HPLC [Phenomenex Synergi 4u Fusion-RP 80A (100 × 21.20mm) column; flow rate was 23 mL/min; gradient run, 0–1 min 0%B, 1–16 min linear gradient 0 to 60%B, 16–19 min linear gradient 60 to 100%B, 19–20 min linear gradient 100 to 0%B; t_R: 5.4 min; UV monitoring 214 nm], neutralized with $(NH_4)_2CO_3$, and lyophilized to afford **10** (3.6 mg, 77%) as a white hygroscopic solid. $^{19}F\{^1H\}$ NMR (D_2O, 470.4 MHz) δ in ppm: –216.80 (s, F-2); HRMS (ES+) for ($M+H^+$) $C_{37}H_{62}FN_{14}O_{16}$ (*m/z*): calc. 977.4447; found 977.4434.

2.4. Preparation of Glycoproteins 12–16

Protein expression using medium shift procedure for Met analogue incorporation: *Ss*βG Met43Hpg Cys344Ser 9Ile[1] (11)

An overnight culture of Escherichia coli B834(DE3), pET28d *Ss*βG C344S was grown in SelenoMet medium Base (200 mL) and nutrient mix (10 mL) (>16 hours) supplemented with kanamycin (50 μg/mL) and L-methionine (40 μg/mL). The overnight culture (4×50 mL) was used to inoculate pre-warmed (37 ºC) culture medium (4×800 mL, same composition as above) and the cells were grown for 4 hours (OD600 1.1). The medium shift was performed by centrifugation (8200 rpm, 12 min, 4 ºC), washed with new medium (without L-methionine, 600 mL) and centrifuged again. The resulting pellets were transferred to pre- warmed culture medium (4×800 mL) containing the amino acid L-Hpg (60 μg/mL). The culture was shaken for 30 min at 37 ºC and for 30 min at 30 ºC, and then induced by addition of IPTG to a final concentration of 1.0 mM. Expression was continued at 30 ºC for 17 hours. The culture was then centrifuged (8200 rpm, 12 min, 4 ºC) and the cell pellets stored at –20 ºC (23.5g frozen pellets). The protein was purified using His-tag nickel affinity purification: the cell pellets were combined and transferred into binding buffer (100 mL) and Lysozyme (1 mg/L) was added. The cells were broken up by sonication (3 x 30 seconds) and the mixture then centrifuged (20000 rpm, 30 min, 4 ºC). The supernatant was filtered (0.45 μm) and the protein was purified using His-tag nickel affinity purification eluting with an increasing concentration of imidazole. After SDS-PAGE analysis a new purification was necessary using Sepharose. The protein was stored at 4 ºC in sodium phosphate buffer (10 mM, pH 7.0). Final protein concentration (0.6

mg/mL; 6 mL) was determinated by BSA assay. Protein mass spectrometry (LC–MS) analysis: Expected 57323. Found 57321.

Sequence *N*-terminally-His7-tagged *Ss*βG-M21I-M43Hpg-M73I-M148IM204I-M236I-M275I- M280I-C344S-M383I-M439I

```
HpgGHHHHHHHS FPNSFRFGWS QAGFQSEIGT PGSEDPNTDW YKWVHDPENHpgAAGLVSGDLP    60
ENGPGYWGNY KTFHDNAQKI GLKIARLNVE WSRIFPNPLP RPQNFDESKQ DVTEVEINEN       120
ELKRLDEYAN KDALNHYREI FKDLKSRGLY FILNIYHWPL PLWLHDPIRV RRGDFTGPSG       180
WLSTRTVYEF ARFSAYIAWK FDDLVDEYST INEPNVVGGL GYVGVKSGFP PGYLSFELSR       240
RAIYNIIQAH ARAYDGIKSV SKKPVGIIYA NSSFQPLTDK DIEAVEIAEN DNRWWFFDAI       300
IRGEITRGNE KIVRDDLKGR LDWIGVNYYT RTVVKRTEKG YVSLGGYGHG SERNSVSLAG       360
LPTSDFGWEF FPEGLYDVLT KYWNRYHLYI YVTENGIADD ADYQRPYYLV SHVYQVHRAI       420
NSGADVRGYL HWSLADNYEW ASGFSIRFGL LKVDYNTKRL YWRPSALVYR EIATNGAITD       480
EIEHLNSVPP VKPLRH                                                       496
```

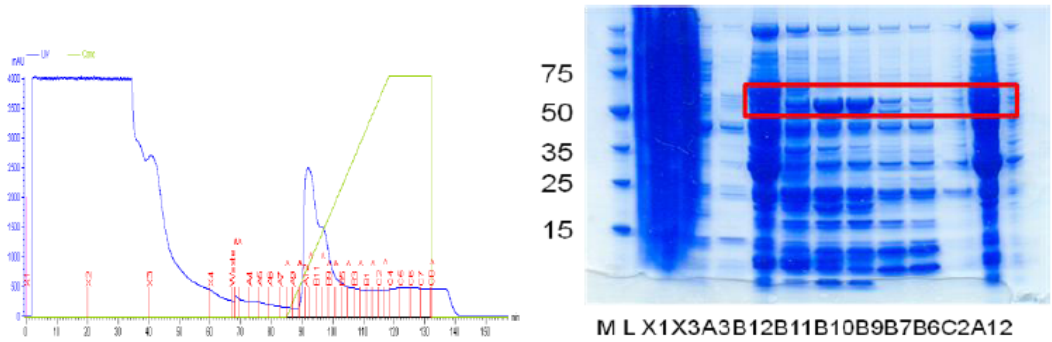


Figure ESI 1. FPLC purification of *Ss*βG-1Hpg43Hpg344Ser (Ni column).

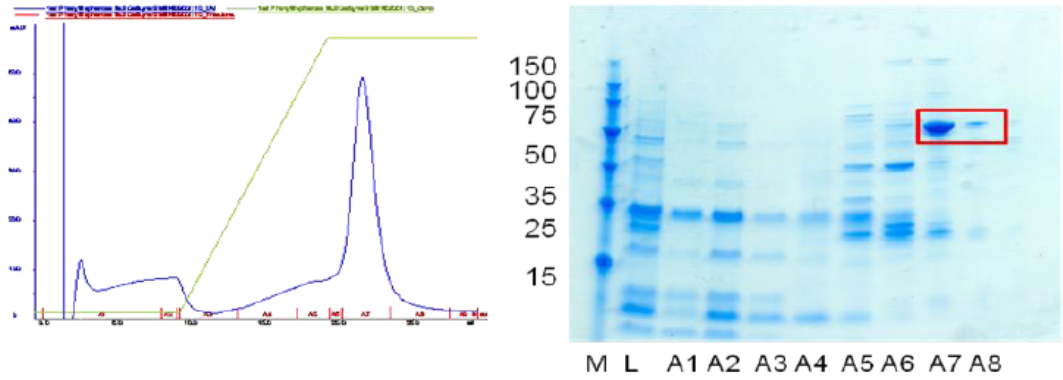


Figure ESI 2. FPLC purification of *Ss*βG-1Hpg43Hpg344Ser (Sepharose column).

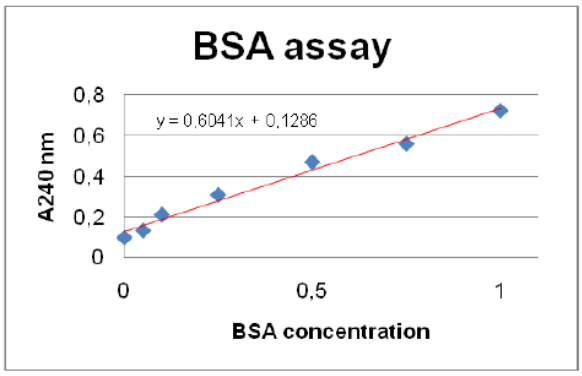


Figure ESI 3. Calibration curve: BSA assay.

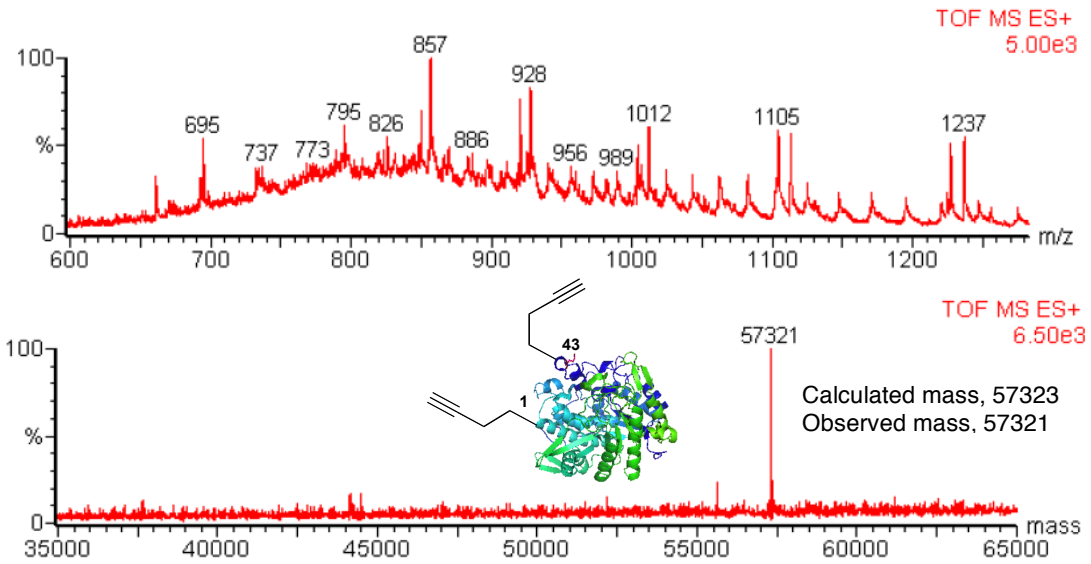
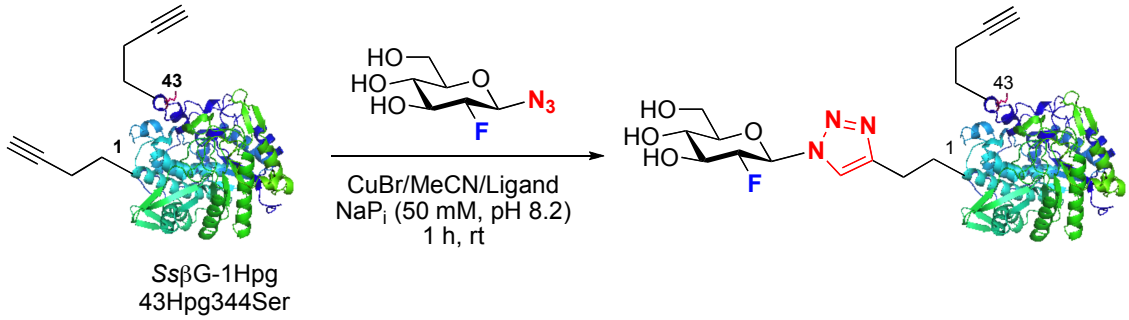


Figure ESI 4. ESI–MS spectrum of *Ss*βG-1Hpg43Hpg344Ser.

2FGlc-triazole-1*Ss*βG (12)



2-Deoxy-2-fluoro-β-D-glucopyranosyl azide **1** (1.16 mg, 5.60 μmol) was dissolved in sodium phosphate buffer (50 μL, 50 mM, pH 8.2). A freshly prepared solution of copper(I) bromide (99.999%) in acetonitrile (163 μL, 10 mg/mL) was premixed with an acetonitrile solution of tris-triazolyl amine ligand (63 μL, 127 mg/mL). The preformed

Cu-complex solution (25 µL) was added to the above solution and mixed thoroughly. *Ss*βG-1Hpg43Hpg344Ser protein solution (100 µL, 0.6 mg/mL) was added to the mixture and the reaction was agitated on a rotator for 1 h at room temperature. 0.2 mL of High Affinity Ni-Charged resin was then added to the mixture and the reaction was agitated on a rotator for 1 h at 4 ºC. The sample was then placed in a syringe and eluted with 5 column volume of buffer A[6] (low imidazol concentration) and 5 column volume of buffer B[7] (high imidazol concentration). The eluted buffer B was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) and washed with sodium phosphate buffer (3 x 200 µL, 50 mM, pH 7.0). Finally, the solution was concentrated to 100 µL and the product was characterized by LC–MS (calc. 57530; found 57527). **Note:** The modified glycoproteins retained galactosidase activity as evidenced by X-Gal stain[8]. In addition, complete regioselective monoglycosylation only at position 1 was observed for dialkynic protein *Ss*βG-1Hpg43Hpg344Ser. This result is in full agreement with results previously reported[1]; for a detailed discussion on reactive accessibility of sites 1 and 43 on *Ss*βG-1Hpg43Hpg344Ser see reference [1] and corresponding supporting information.

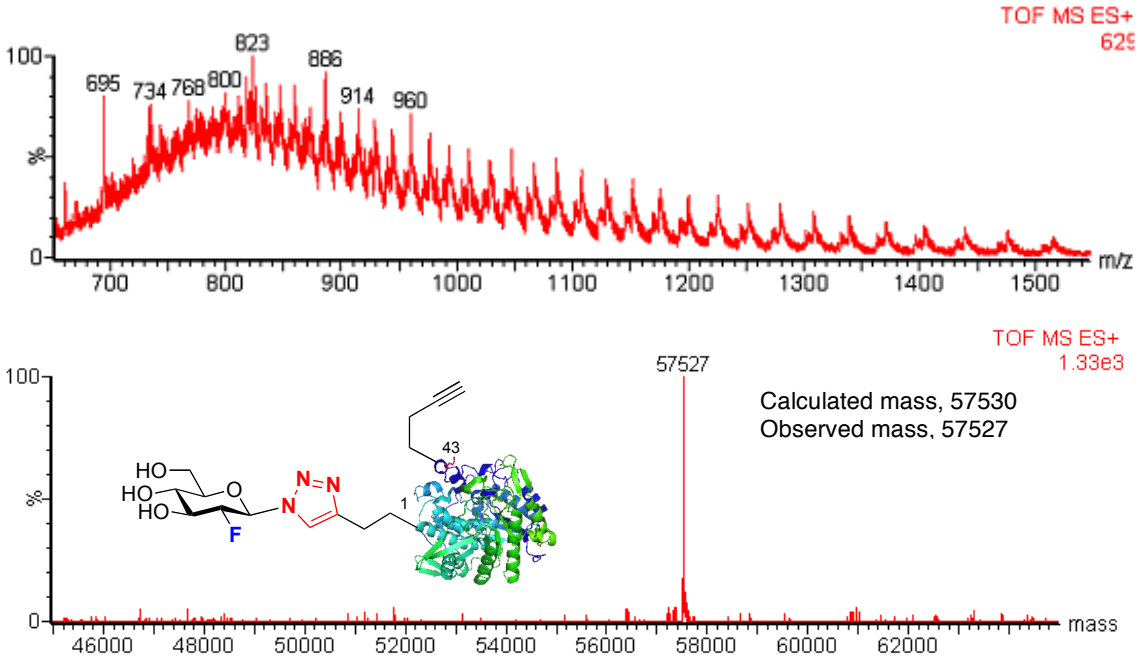
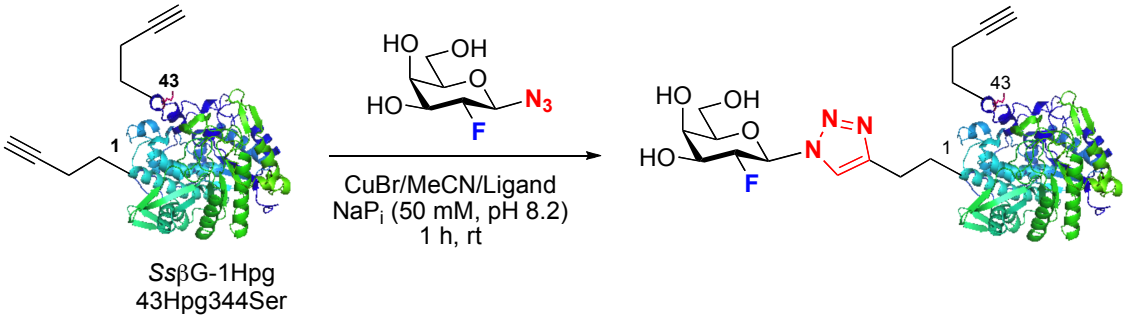


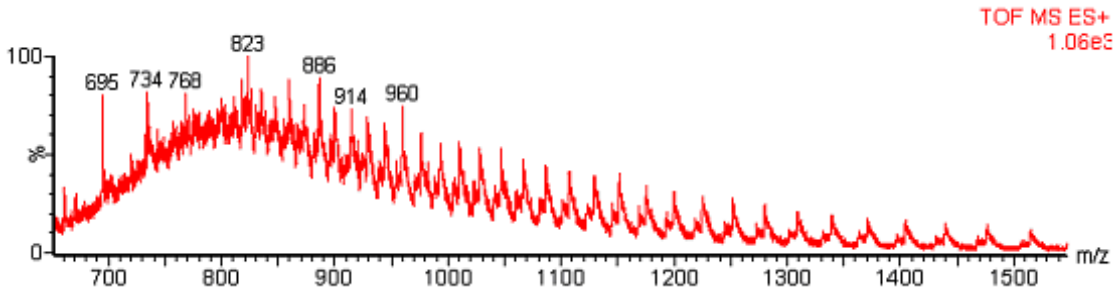
Figure ESI 5. ESI–MS spectrum of 2FGlc-triazole-1*Ss*βG **12**.

2FGal-triazole-1*Ss*βG (13)



*Ss*βG-1Hpg
43Hpg344Ser

2-Deoxy-2-fluoro-β-D-galactopyranosyl azide **2** (3.6 mg, 17.4 µmol) was dissolved in sodium phosphate buffer (50 µL, 50 mM, pH 8.2). A freshly prepared solution of copper(I) bromide (99.999%) in acetonitrile (163 µL, 10 mg/mL) was premixed with an acetonitrile solution of tris-triazolyl amine ligand (63 µL, 127 mg/mL). The preformed Cu-complex solution (75 µL) was added to the above solution and mixed thoroughly. *Ss*βG-1Hpg43Hpg344Ser protein solution (250 µL, 0.6 mg/mL) was added to the mixture and the reaction was agitated on a rotator for 1 h at room temperature. 0.5 mL of High Affinity Ni-Charged resin was then added to the mixture and the reaction was agitated on a rotator for 1 h at 4 ºC. The sample was then placed in a syringe and eluted with 5 column volume of buffer A[6] (low imidazol concentration) and 5 column volume of buffer B[7] (high imidazol concentration). The eluted buffer B was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) and washed with sodium phosphate buffer (3 x 200 µL, 50 mM, pH 7.0). Finally, the solution was concentrated to 100 µL and the product was characterized by LC–MS (calc. 57530; found 57526).



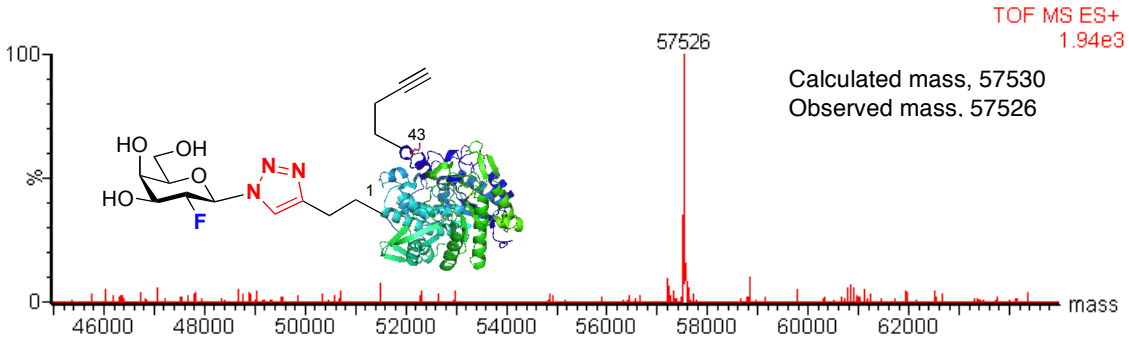
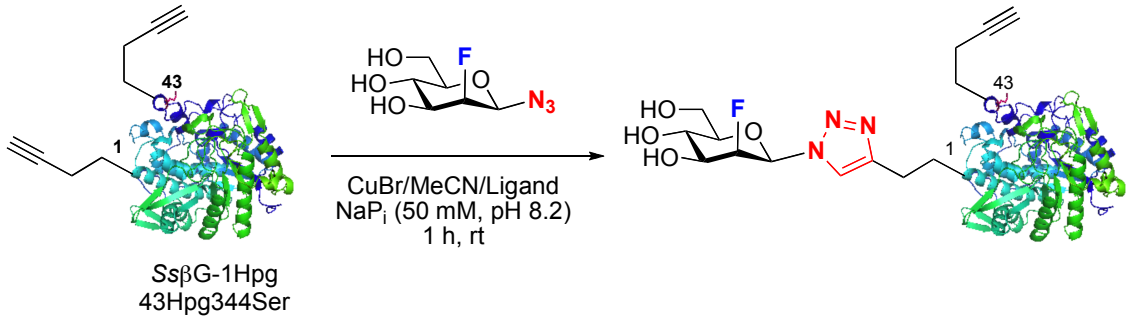


Figure ESI 6. ESI–MS spectrum of 2FGal-triazole-1*Ss*βG **13**.

2FMan-triazole-1*Ss*βG (14)



2-Deoxy-2-fluoro-β-D-mannopyranosyl azide **3** (1.13 mg, 5.45 µmol) was dissolved in sodium phosphate buffer (50 µL, 50 mM, pH 8.2). A freshly prepared solution of copper(I) bromide (99.999%) in acetonitrile (163 µL, 10 mg/mL) was premixed with an acetonitrile solution of tris-triazolyl amine ligand (63 µL, 127 mg/mL). The preformed Cu-complex solution (25 µL) was added to the above solution and mixed thoroughly. *Ss*ßG-1Hpg43Hpg344Ser protein solution (100 µL, 0.6 mg/mL) was added to the mixture and the reaction was agitated on a rotator for 1h at room temperature. 0.2 mL of High Affinity Ni-Charged resin was then added to the mixture and the reaction was agitated on a rotator for 1h at 4 ºC. The sample was then placed in a syringe and eluted with 5 column volume of buffer A[6] (low imidazole concentration) and 5 column volume of buffer B[7] (high imidazole concentration). The eluted buffer B was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) and washed with sodium phosphate buffer (3 x 200 µL, 50 mM, pH 7.0). Finally, the solution was concentrated to 100 µL and the product was characterized by LC–MS (calc. 57530; found 57531).

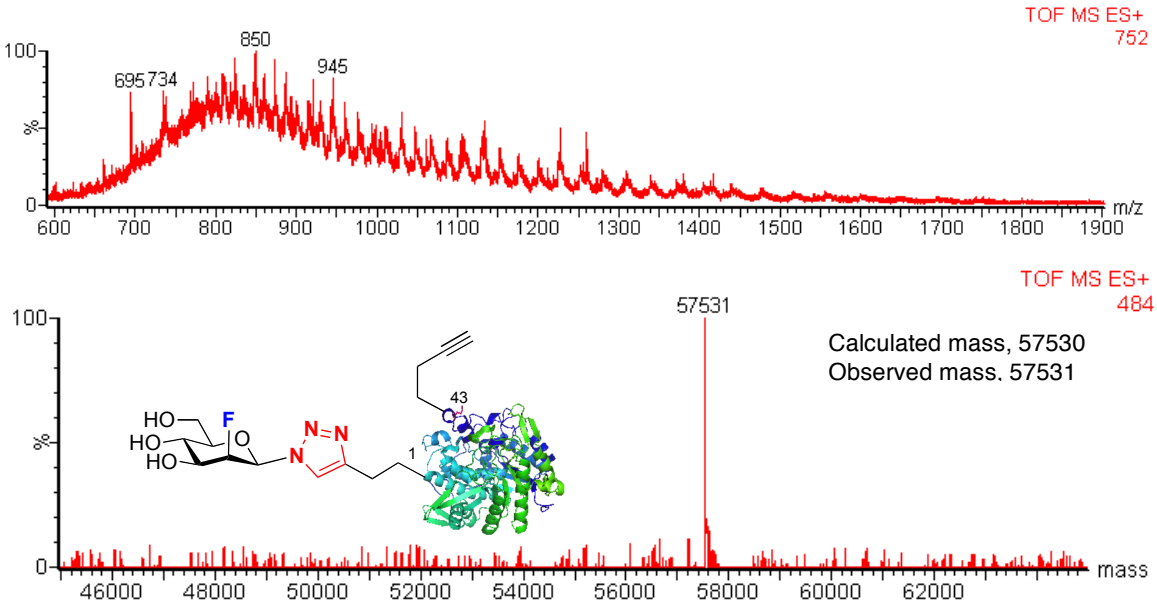


Figure ESI 7. ESI–MS spectrum of 2FMan-triazole-1*S*sβG **14**.

Expression of Qβ M16Hpg (15)

A glycerol stock of *E. coli* B834(DE3) containing p75M plasmid carrying the K16M Qβ gene, was used to inoculate SelenoMet media (10 mL) containing ampicillin (100 μg/mL) and supplemented with L-methionine (final concentration 40 mg/L). An overnight culture was grown for 16 hours at 37 °C. The overnight culture was used to inoculate SelenoMet media (625 mL) containing ampicillin (100 μg/mL) and supplemented with L-methionine (final concentration 40 mg/L). The cells were incubated at 37 °C overnight. The cells were then pelleted by centrifugation (8 minutes, 8000 rpm). The pellets were washed twice with SelenoMet media (2 x 400 mL) containing ampicillin (100 μg/mL) and then transferred to prewarmed SelenoMet media (625 mL) containing ampicillin (100 μg/mL) and supplemented with L-Hpg (final concentration 80 mg/L). The culture was incubated for 30 minutes at 37 °C, followed by 30 minutes at 30 °C. Expression was induced by addition of IPTG (1 mM) and was continued for 6 hours at 37 °C. The cells were harvested by centrifugation (8 minutes, 8000 rpm) to give pellets, which were collected and stored at –80 °C in buffer (20 mM Tris-HCl, 0.15 M NaCl, 2 mM EDTA, pH 7.5) (30 mL).

Purification of Qβ M16Hpg (15)

Cell lysis: The cell pellet was thawed and then sonicated (20% power in bursts of 30 seconds with a wait time of 1 minute between each burst). The cell debris was removed from the lysis mixture by centrifugation (30 minutes, 20,000 rpm).

Extraction of Qβ–Hpg virus-like particles: The cleared lysate was added to an equal volume of 1:1 butanol:chloroform and mixed by inversion. The layers were separated by centrifugation (20 minutes, 8000 rpm). The aqueous layer was removed and the VLPs were precipitated by addition of PEG8000 (10% w/v) and NaCl (100 mM). The mixture was vortexed and then incubated on a rocker at 4 ºC for 30 minutes. The precipitated VLPs were collected by centrifugation (20 minutes, 8000 rpm). The precipitate was resuspended in 10 mL of TBS (20 mM Tris, 150 mM NaCl, pH 7.5). The extraction step was then repeated a second time.

Sucrose gradient purification of VLPs: After the second extraction, the VLPs were purified on a 10-40% sucrose gradient. The gradients were spun in an ultracentrifuge (100,000 x g, 16 hours, Beckman Optima Ultracentrifuge, SW 41 Ti rotor). The particles were visible as a cloudy band in the gradient after spinning. This band was collected, and the VLPs were precipitated by addition of ammonium sulfate (2M). The mixture was allowed to rock at 4 ºC for 2 hours. The precipitate containing the VLPs was then isolated by centrifugation (20 minutes, 8000 rpm) and resuspended in 5 mL PBS (Sigma).

Preparation of VLPs for experiments: The Qβ-Hpg VLPs were exchanged into phosphate buffer (50mM NaH_2PO_4, pH 8) using a PD-10 column (GE Healthcare). The concentration was checked by BCA assay (Pierce) and the VLPs were diluted as necessary for reaction.

Protein detection

At all stages of expression and purification, protein was detected by SDS-PAGE. All gels used were precast 12% gels from Invitrogen and were run in MOPS buffer also from Invitrogen. Gels were stained using InstantBlue (Expedion) gel stain. A virus-like particle aliquot (5 μL) was mixed with 1 M DTT (dithiothreitol) in H_2O (5 μL), and

incubated at 60 ºC for 5 min to allow the protein to denature prior to analysis by LC–MS (m/z for monomer of Qβ M16Hpg: calculated mass, 14104; observed mass, 14104).

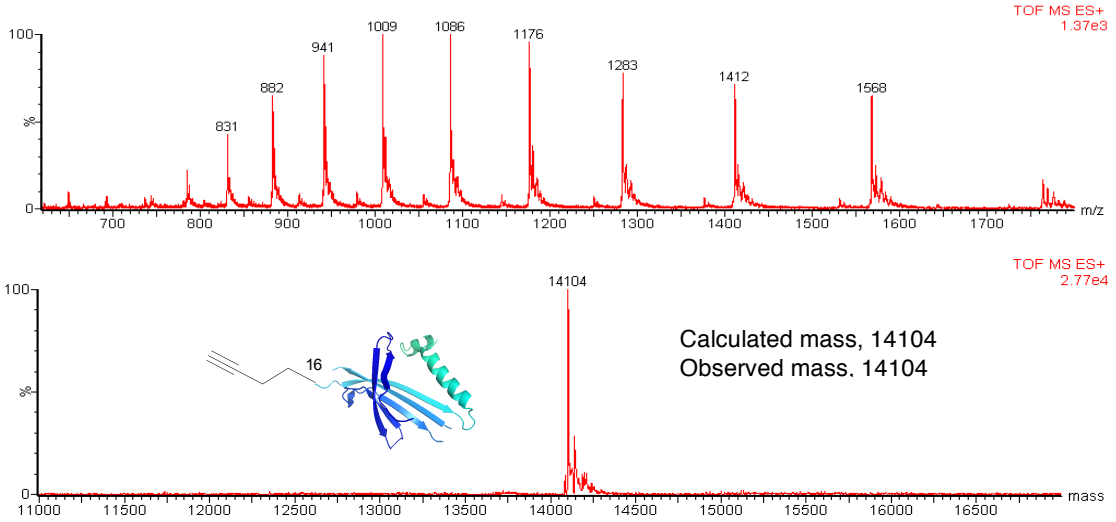
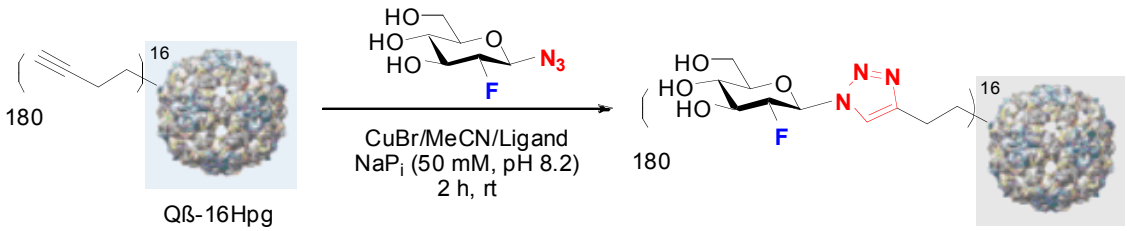


Figure ESI 8. ESI–MS spectrum for monomer of Qβ M16Hpg.

2FGlc-triazole-16Qβ (16)



2-Deoxy-2-fluoro-β-D-glucopyranosyl azide **1** (5.5 mg, 26.5 μmol) was dissolved in sodium phosphate buffer (100 μL, 50 mM, pH 8.2). A freshly prepared solution of copper(I) bromide (99.999%) in acetonitrile (89 μL, 10 mg/mL) was premixed with an acetonitrile solution of tris-triazolyl amine ligand (36 μL, 127 mg/mL). The preformed Cu-complex solution (125 μL) was added to the above solution and mixed thoroughly. Qβ M16Hpg protein solution (500 μL, 3 mg/mL) was added to the mixture and the reaction was agitated on a rotator for 2 h at room temperature. Small molecules were removed from the reaction mixture by loading the sample onto a PD10 desalting column (GE Healthcare) previously equilibrated with 10 column volumes of pH 8.0 sodium phosphate buffer (50 mM) and eluting with 1 mL of the same buffer. The collected sample was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) and washed with sodium phosphate buffer (5 x 100 μL, 50 mM, pH 8.2). A virus-like particle aliquot (5 μL) was mixed with 1 M DTT (dithiothreitol) in H_2O (5

μL), and incubated at 60 ºC for 5 min to allow the protein to denature prior to analysis by LC–MS (m/z for monomer of 2FGlc-triazole-16Qβ: calculated mass, 14311; observed mass, 14313). Finally, the sample was flash frozen with liquid nitrogen and stored at –80 ºC.

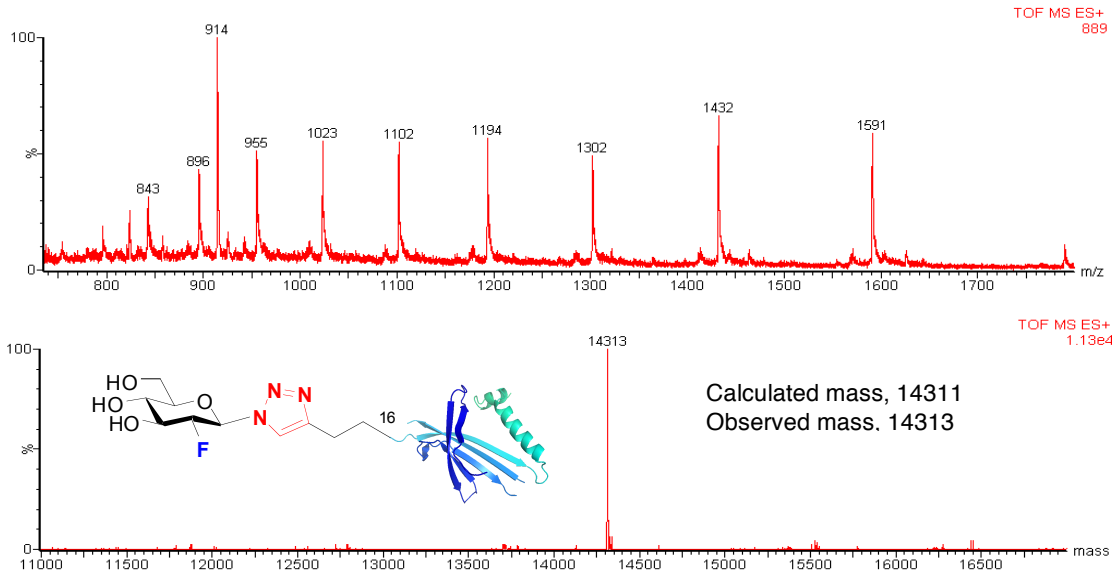


Figure ESI 9. ESI–MS spectrum for monomer of 2FGlc-triazole-16Qβ **16**.

2.5. SDS-PAGE analysis of Glycoproteins 12–14

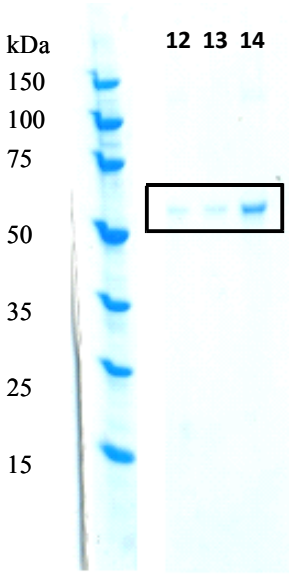
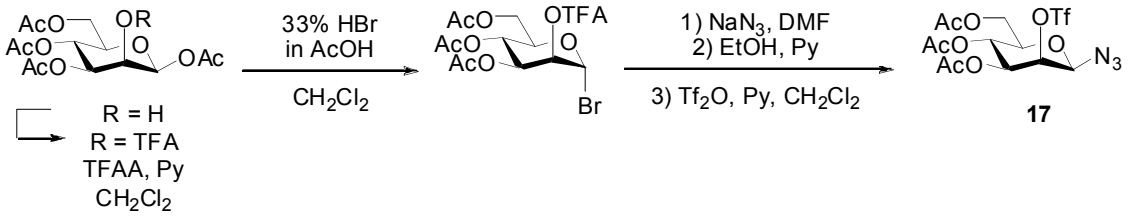


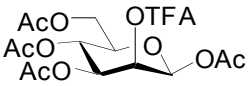
Figure ESI 9. SDS-PAGE analysis of glycoproteins **12–14** (MES running buffer and coomassie blue stain).

2.6. Preparation of 2-Deoxy-2-[^{18}F]fluoroglycosyl Azides 18 and 19

3,4,6-Tri-*O*-acetyl-2-*O*-trifluoromethanesulfonyl-β-D-mannopyranosyl azide (17)

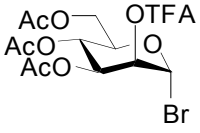


1,3,4,6-Tetra-*O*-acetyl-2-*O*-trifluoroacetyl-β-D-mannopyranose



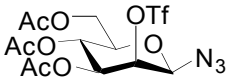
1,3,4,6-Tetra-*O*-acetyl-β-D-mannopyranose (1.20 g; 3.45 mmol) was dissolved in dry dichloromethane (30 mL) and pyridine (502 μL, 6.20 mmol, 1.8 eq) was added under a nitrogen atmosphere. Trifluoroacetic acid anhydride (TFAA, 781 μL, 5.62 mmol, 1.63 eq) was added dropwise at 0°C and the reaction was allowed to warm to room temperature and stir overnight (15h). The reaction was slowly neutralised by addition of ice, then ethyl acetate was added. The organic layers were washed twice with ice-cooled water, once with 0.2 N HCl and twice with brine before being dried over magnesium sulphate. The organic layers were evaporated to dryness to give 1.56 g of crude oil, which solidify on standing. This crude oil was purified by column chromatography (7:3 petroleum ether: ethyl acetate) to give 1.21 g of colourless oil (2.72 mmol, 79%). $[\alpha]_D^{25}$ –10.9 (1.0, $CHCl_3$); IR (thin film) υ in cm^{-1}: 1798 (OTFA), 1749 (OAc), 1371, 1224; TLC Rf (1:1 EtOAc/petrol): 0.67; product degrades on TLC plate into 0.33 and 0.17); ^{1}H NMR ($CDCl_3$, 400 MHz) δ in ppm: 5.94 (s, 1H, H-1), 5.56 (d, 1H, $J_{2,3}$ = 3.0 Hz, H-2), 5.30 (appt, 1H, $J_{3,4} = J_{4,5}$ = 10.0 Hz, H-4), 5.23 (dd, 1H, $J_{3,2}$ = 3.0 Hz, $J_{3,4}$ = 10.0 Hz, H-3), 4.27 (dd, 1H, $J_{6a,b}$ = 12.5 Hz, $J_{6a,5}$ = 4.5 Hz, H-6a), 4.16 (dd, 1H, $J_{6a,b}$ = 12.5 Hz, $J_{6a,5}$ = 2.2 Hz, H-6b), 3.85 (ddd, 1H, $J_{5,4}$ = 10.0 Hz, $J_{5,6a}$ = 4.5 Hz, $J_{5,6b}$ = 2.2 Hz, H-5), 2.12, 2.10, 2.07, 2.03 (s, 4 x 3H, 4 OAc CH_3); ^{13}C-DEPT NMR ($CDCl_3$, 126 MHz) δ in ppm: 89.5 (C-1), 73.3 (C-5), 72.4 (C-2), 70.0 (C-3), 64.9 (C-4), 61.6 (C-6); ^{19}F NMR ($CDCl_3$, 376.5 MHz) δ in ppm: –74.6 (CF_3); LRMS (ES+) for ($M+Na^+$) $C_{16}H_{19}F_3O_{11}Na$ (*m/z*): calc. 467.08; found 467.11; HRMS (FI+) for (M^+) $C_{16}H_{19}F_3O_{11}$ (*m/z*): calc. 444.0880; found 444.0754.

3,4,6-Tri-*O*-acetyl-2-*O*-trifluoroacetyl-α-D-mannopyranosyl bromide



1,3,4,6-Tetra-*O*-acetyl-2-*O*-trifluoroacetyl-β-D-mannopyranose (1.21g, 2.72 mmol) was dissolved in dry dichloromethane (30 mL) and the solution was cooled down to 0°C under nitrogen. 6 mL of HBr in AcOH (33% w/w, 30 mmol) were added slowly at 0°C and the mixture was allowed to warm to room temperature and stirred overnight. The solution was diluted with dichloromethane and washed with saturated $NaHCO_3$ followed by water. The organic layer was dried on sodium sulphate and evaporated to dryness to give the product as a crude clear oil (1.14 g, 2.45 mmol, 90%). $[\alpha]_D^{25}$ +118.8 (1.0, $CHCl_3$); IR (thin film) υ in cm^{-1}: 1799 (OTFA), 1754 (OAc), 1369, 1227, 1148; TLC Rf (1:1 EtOAc/petrol): 0.82 streaking; ^1H NMR ($CDCl_3$, 400 MHz) δ in ppm: 6.39 (d, 1H, $J_{1,2}$ = 1.5 Hz, H-1), 5.81 (dd, 1H, $J_{3,2}$ = 3.2 Hz, $J_{3,4}$ = 10.2 Hz, H-3), 5.61 (dd, 1H, $J_{2,3}$ = 3.2 Hz, $J_{2,1}$ = 1.5 Hz, H-2), 5.31(appt, 1H, $J_{3,4}$ = $J_{4,5}$ = 10.0 Hz, H-4), 4.24 (m, 3H, H-6a,b, H-5), 2.11, 2.09, 2.02 (s, 3 x 3H, 3 OAc CH_3); ^{13}C NMR ($CDCl_3$, 126 MHz) δ in ppm: 170.5, 169.5, 169.2 (3x C=O OAc), 80.8 (C-1), 75.3 (C-2), 72.9 (C-5), 67.6 (C-3), 64.6 (C-4), 61.0 (C-6), 25.6, 20.3, 20.3 (3x CH_3 OAc); ^{19}F NMR ($CDCl_3$, 376.5 MHz) δ in ppm: –74.6 (CF_3).

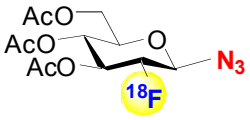
3,4,6-Tri-*O*-acetyl-2-*O*-trifluoromethanesulfonyl-β-D-mannopyranosyl azide (17)



3,4,6-Tri-*O*-acetyl-2-*O*-trifluoroacetyl-α-D-mannopyranosyl bromide (1.14 g, 2.45 mmol) was dissolved in 15 mL of dry DMF. Sodium azide (207 mg, 3.19 mmol, 1.3 eq) was added and the mixture was stirred overnight under nitrogen. TLC showed two new product spots along with remaining starting material. Sodium azide was added again (100 mg, 1.53 mmol, 0.6 eq) and the reaction was stirred 2.5 h more until TLC didn't detect any starting material left. 15 mL of dry ethanol and 1.5 mL of dry pyridine were added and the reaction was stirred for 2 h, then was evaporated to dryness to give 1.7 g of crude solid. This solid was dissolved in dichloromethane, washed with water twice and dried over sodium sulphate to give 792.3 mg of crude oil. This oil was purified by column chromatography (2:3 EtOAc/petrol) to give 385 mg of 3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl azide as a mixture of isomers (45% yield) that was used in the next step. 365.3 mg of the previous mixture of isomers of 3,4,6-tri-*O*-acetyl-D-mannopyranosyl azide (1.10 mmol) was dissolved in 20 mL of dry dichloromethane and pyridine (205 μL, 2.54 mmol, 2. 3 eq) was added. The mixture was cooled down to 0°C and trifluoromethanesulphonyl

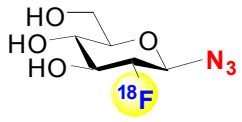
anhydride (Tf_2O, 205 μL, 1.21 mmol, 1. 1 eq) was added slowly under nitrogen. The mixture was stirred 40 minutes after which TLC showed complete conversion to two new compounds. The reaction mixture was then diluted in dichloromethane and ice-cold water was added. The organic layer was separated and dried over sodium sulphate, evaporated to dryness to give 483.6 mg of yellow oil. This oil was purified by silica column chromatography (1:4 EtOAc/petrol to 3:7 to 2:3), giving 239.8 mg of mixed alpha triflate mannopyranosyl azide (0.517 mmol, 21% over two steps); Rf (1: EtOAc/petrol): 0.78; followed by 176.9 mg of pure 2-*O*-trifluoromethanesulfonyl-3,4,6-tri-*O*-acetyl-β-D-mannopyranosyl azide (0.382 mmol, 16% over two steps) Rf (1:1 EtOAc/petrol): 0.56; ^{1}H NMR ($CDCl_3$, 400 MHz) δ in ppm: 5.28 (appt, 1H, $J_{4,3} = J_{4,5} =$ 10.0 Hz, H-4), 5.11 (dd, 1H, $J_{3,4} = 10.0$ Hz, $J_{3,2} = 3$ Hz, H-3), 5.04 (appd, 1H, $J_{2,3} = 3.0$ Hz, H-2), 5.03 (apps, 1H, H-1), 4.27–4.24 (m, 2H, H-6a,b), 3.83 (ddd, 1H, $J_{5,4} = 10.0$ Hz, $J_{5,6a} = 4.9$ Hz, $J_{5,6b} = 3.1$ Hz, H-5), 2.12 (s, 6H, 2xCH_3 OAc), 2.08 (s, 3H, CH_3 OAc); LRMS (ES+) for (M+Na^+) $C_{13}H_{16}F_3N_3O_{10}SNa$ (*m/z*): calc. 486.02; found 486.04; spectroscopic data was identical to that previously reported.[9]

3,4,6-Tri-*O*-acetyl-2-deoxy-2-[^{18}F]fluoro-β-D-glucopyranosyl azide[9] (18)



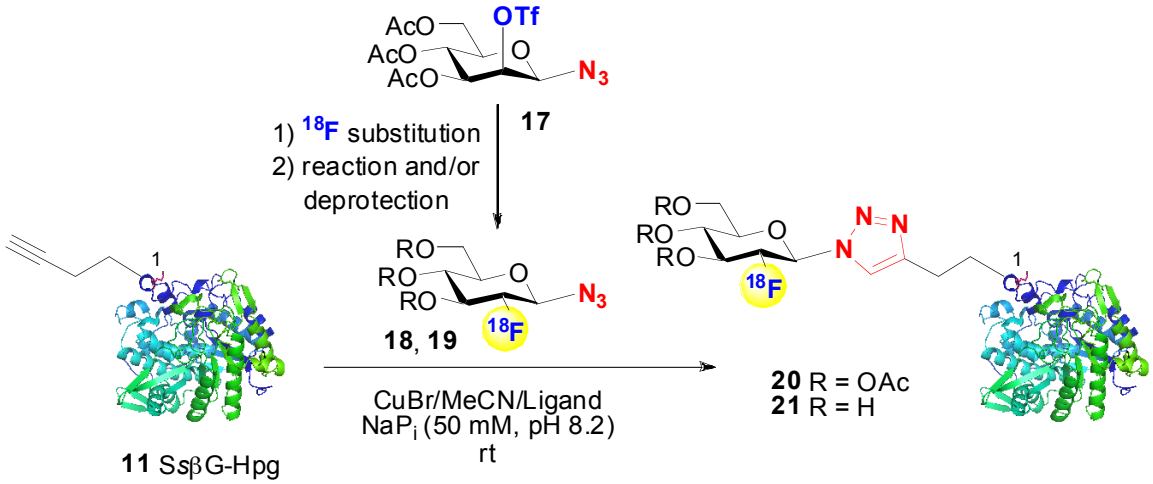
In a Scintomics radiochemistry setup, a QMA cartridge (Chromafix 30 PS HCO3, ^{18}F separation cartridge) with [^{18}F] fluoride (2.92–3.89 GBq) was eluted with a solution of 15 mg Kryptofix 2.2.2 and 3 mg of K_2CO_3 in 1mL of acetonitrile/water (8/2). The solution was evaporated using a stream of nitrogen at 125 °C and co-evaporated to dryness with acetonitrile (2x200 μL). 3,4,6-Tri-*O*-acetyl-2-*O*-trifluoromethanesulfonyl-β-D-mannopyranosyl azide **17** (7.2 mg in 450 μL) was added and the mixture was heated at 85 °C for 5 minutes. 1.5 mL of acetonitrile/water (30/70) was added and the mixture was injected into a semipreparative HPLC (Phenomenex Luna 5u C18 150 x 10.0 mm, eluting with 4mL/min isocratic acetonitrile 30% / water 70%, UV and radioactive detection). The product-containing fraction (t_R: 21 min) was diluted in water (40mL) and passed through a C18-cartridge (Waters, 100mg). The cartridge was washed with water (2mL), dried in a stream of nitrogen and eluted with 1mL of ethanol. (Activity: 50.7–171.3 MBq , RCY = 1.3–4.7% in 80–100 min, n=5).

2-Deoxy-2-[^{18}F]fluoro-β-D-glucopyranosyl azide[9] (19)



To the previous solution of **18** (39.0–88.8 MBq) was added NaOH (60 mM in water/ethanol 90/10, 0.25 mL) and the mixture was heated at 65 °C for 10 minutes. The solution was passed through a short plug of DOWEX 50WX-8 (H+ form), the plug was washed with 1mL of ethanol and the combined ethanol solution was evaporated to dryness under a stream of nitrogen at 65 °C and used for protein conjugation assay giving 26.2–56.4 MBq (RCY = 64–77% in 18–50 min, n=5)

2.7. Preparation of [^{18}F]-labelled Glyproteins 20 and 21



3,4,6-Tri-O-acetyl-2[^{18}F]FGlc-triazole-1*Ss*βG (20)

3,4,6-Tri-O-acetyl-2-deoxy-2-[^{18}F]fluoro-β-D-glucopyranosyl azide **18** (10.2 MBq) was dissolved in sodium phosphate buffer (25 µL, 50 mM, pH 8.2). A freshly prepared solution of copper(I) bromide (99.999%) in acetonitrile (163 µL, 10 mg/mL) was premixed with an acetonitrile solution of tris-triazolyl amine ligand (63 µL, 127 mg/mL). The preformed Cu-complex solution (12.5 µL) was added to the above solution and mixed thoroughly. The above mixture was added to the *Ss*βG-1Hpg43Hpg344Ser protein solution (50µL of 0.6 mg/mL, 0.52 nmol). After 3 h at room temperature the reaction was analyzed by radio-RP-HPLC (RCY = 60%, n=3).

2[^{18}F]FGlc-triazole-1*Ss*βG (21)

2-Deoxy-2-[^{18}F]fluoro-β-D-glucopyranosyl azide **19** (9.7 MBq) was dissolved in sodium phosphate buffer (25 µL, 50 mM, pH 8.2). A freshly prepared solution of

copper(I) bromide (99.999%) in acetonitrile (163 µL, 10 mg/mL) was premixed with an acetonitrile solution of tris-triazolyl amine ligand (63 µL, 127 mg/mL). The preformed Cu-complex solution (12.5 µL) was added to the above solution and mixed thoroughly. The above mixture was added to the *Ss*βG-1Hpg43Hpg344Ser protein solution (50µL of 0.6 mg/mL, 0.52 nmol). After 45 minutes at room temperature the reaction was analyzed by radio-RP-HPLC (RCY = 4.1%, n=1).

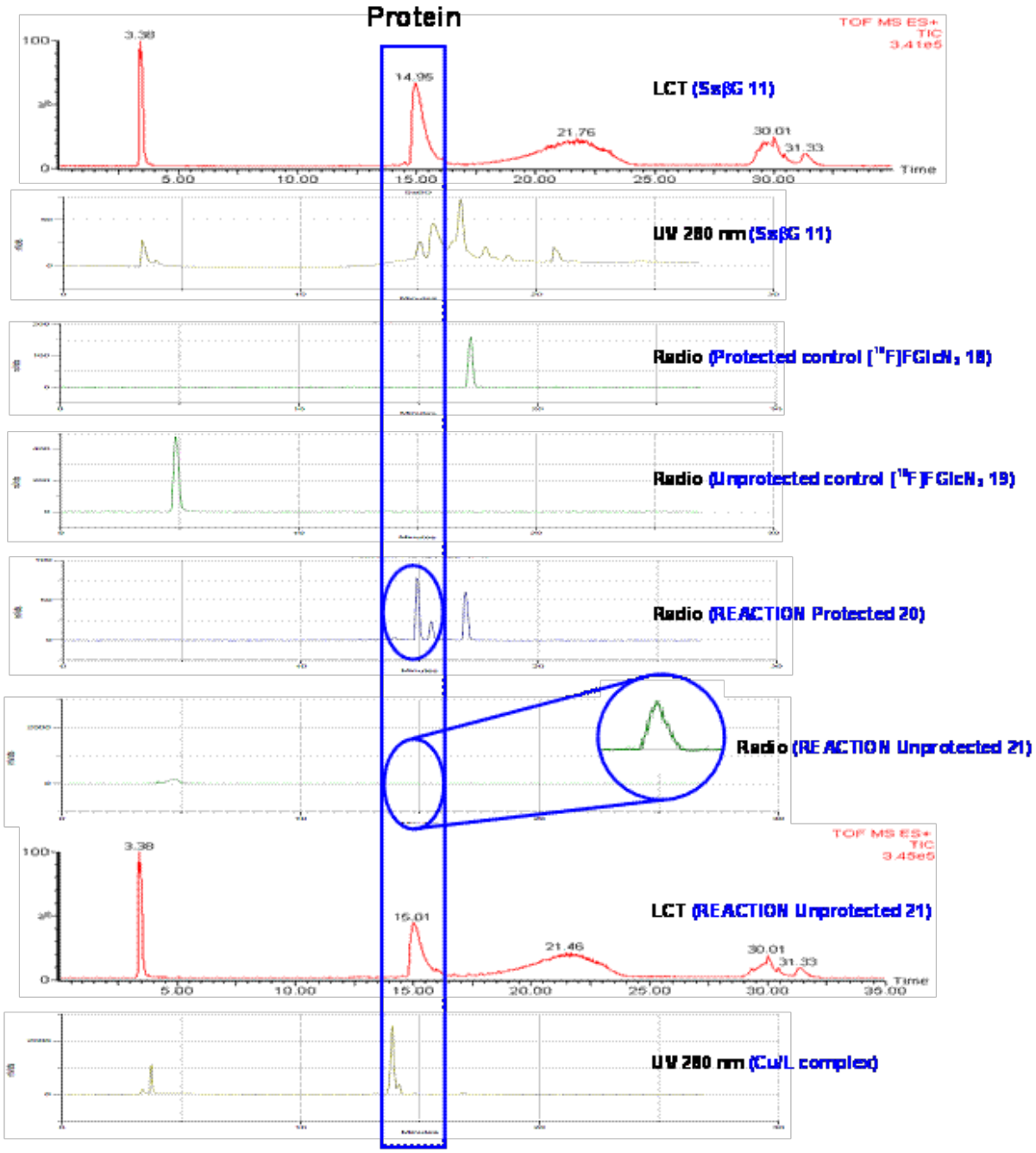


Figure ESI 10. LC–MS and RP-HPLC traces (radio and UV at 280 nm) including control experiments for the reparation of [^{18}F]-labelled Glyproteins **20** and **21**.

Double peaks in RP-HPLC analysis

RP-HPLC of a chemically pure protein leads in most cases to a single peak. However, proteins are sometimes eluted from HPLC columns as multiple or/and irregular-shaped peaks. This could lead to chromatograms with poor resolution or to wrong analysis. Proteins are stabilized by a combination of hydrogen bonding, electrostatic interactions, and hydrophobic interactions. The same forces are also involved in the chromatographic processes, thus conformational changes may occur induced by mobile or stationary phase or both, leading to loss of bioactivity and denaturation. The native and denaturated forms can be resolved if the kinetic process of conformational change is slow relative to the time scale of migration through the column, otherwise, a single broad peak will be obtained, which is the weighted average of the forms in equilibrium.

Such a chromatographic behaviour can be due to chemical transformations, conformational changes (denaturation, renaturation), or other structural changes (aggregation, chain disruption, *etc.*) undergone during protein elution. These changes in the analyte can be promoted by mobile phase and/or stationary phase mediated by temperature. A conformational change of the solute that takes place during its migration through the column can result in broadened peaks or multiple peaks if the kinetics of this change is on the time scale of the chromatographic experiment and the products have a retention time different from that of the original solutes. In other cases, multiple peaks are not caused by the chromatographic system, but rather due to chemical transformations, conformational changes or other structural changes taking place in the sample during the isolation, purification or other processing steps prior to its elution through the column.[10]

In particular, peak shapes for globular proteins above 20 KDa are very responsive to mobile phase mediated phenomena which affect the diffusion coefficients of these solutes as well as the kinetic resistance to migration of the solutes onto and off the alkylsilica support.[11]

On the basis of the SDS PAGE and LCMS analysis (single monoglycosylated protein), it is reasonable to conclude that the phenomenon being observed in the radiotrace of the protected [^{18}F]FGlc-labelled SsβG **20** (double and/or broad peak) is a consequence of

the denaturation or conformational variations due to multiple charge states of the resulting glycoprotein.

Importantly, the following references on multiple peaks found in HPLC analysis of radioactive peptides[12] show that multiple molecular configurations of identical molecular weight but differing in shape, charge, isomerism or lipophilicity such that they are resolved under the conditions of certain analyses. **One conclusion from this investigation is that the appearance of a single peak by any HPLC analysis offers no assurance that multiple peaks would not appear on alternative HPLC analyses. It will always be difficult by HPLC to prove that a single radiochemical species is present. Evidence that each species is due to radiolabelled active peptide and not to radiocontaminants is therefore potentially more important than evidence of a single peak.**

Supplementary Information

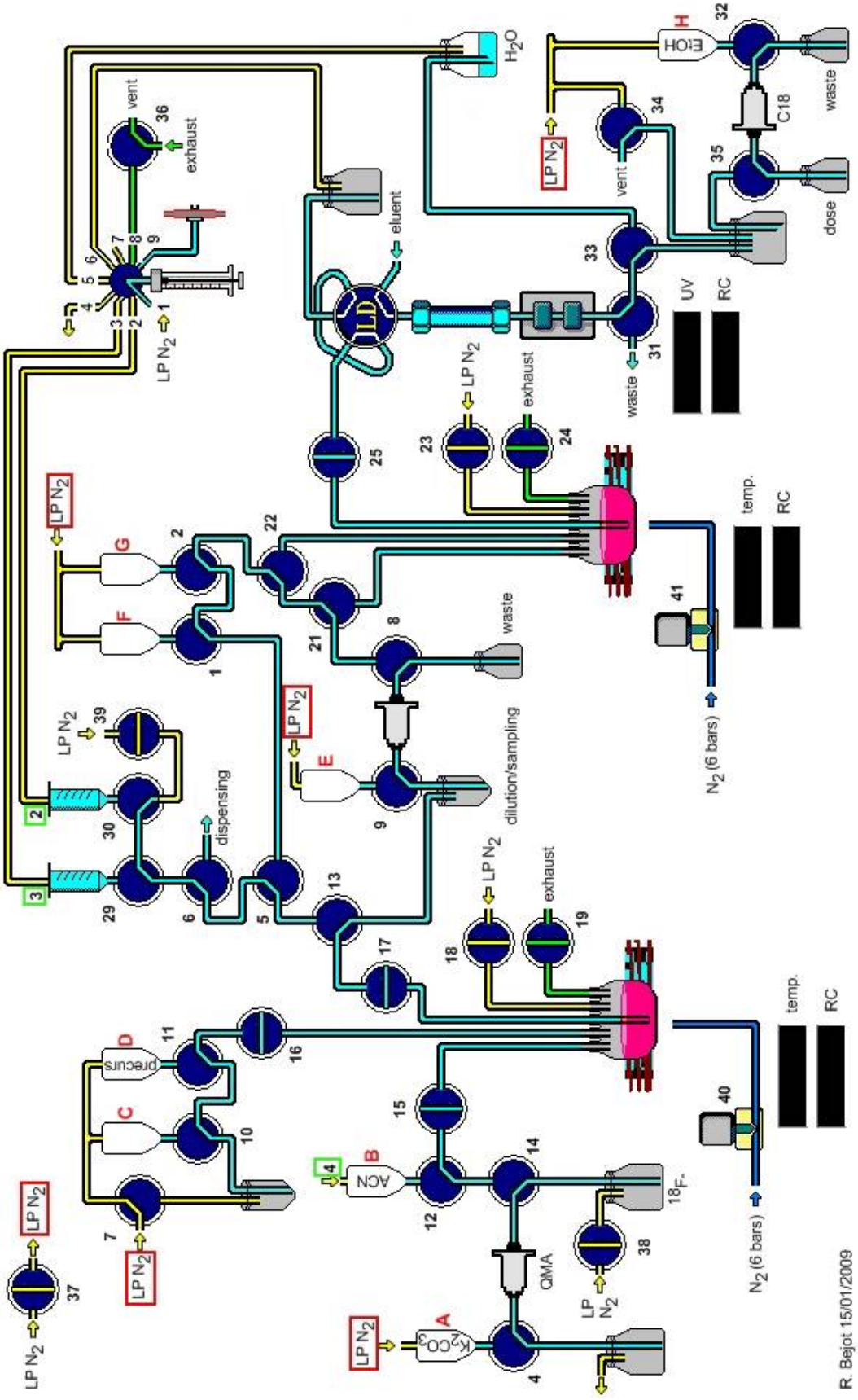


Figure ESI 11. Scintomics schematic map.

3. References

[1] S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony, B. G. Davis, *Nature* **2007**, *446*, 1105–1109.

[2] M. Albert, B. J. Paul, K. Dax, *Synlett* **1999**, 1483–1485.

[3] Z. Zhou, C. J. Fahrni, *J. Am. Chem. Soc.* **2004**, *126*, 8862–8863.

[4] S. Maschauer, O. Prante, *Carbohydr. Res.* **2009**, *344*, 753–761.

[5] M. Albert, K. Dax, J. Ortner, *Tetrahedron* **1998**, *54*, 4839–4848.

[6] Tris HCl 20 mM, NaCl 500 mM, imidazol 5 mM, pH 7.8.

[7] Tris HCl 20 mM, NaCl 500 mM, imidazol 15 mM, pH 7.8.

[8] β-Galactosidase activity screening; X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (40 mg) dissolved in DMSO (1 mL) was added to PBS buffer (40 mL) containing K_2FeCN_6 (5 mM) and K_3FeCN_6 (5 mM). β-Galactosidase activity was measured qualitatively. X-Gal solution (100 μL) was added to protein sample (50 μL) and the mixture was incubated at 37 ºC for *ca.* 1 hour. A solution colour change from yellow to blue was considered a positive result.

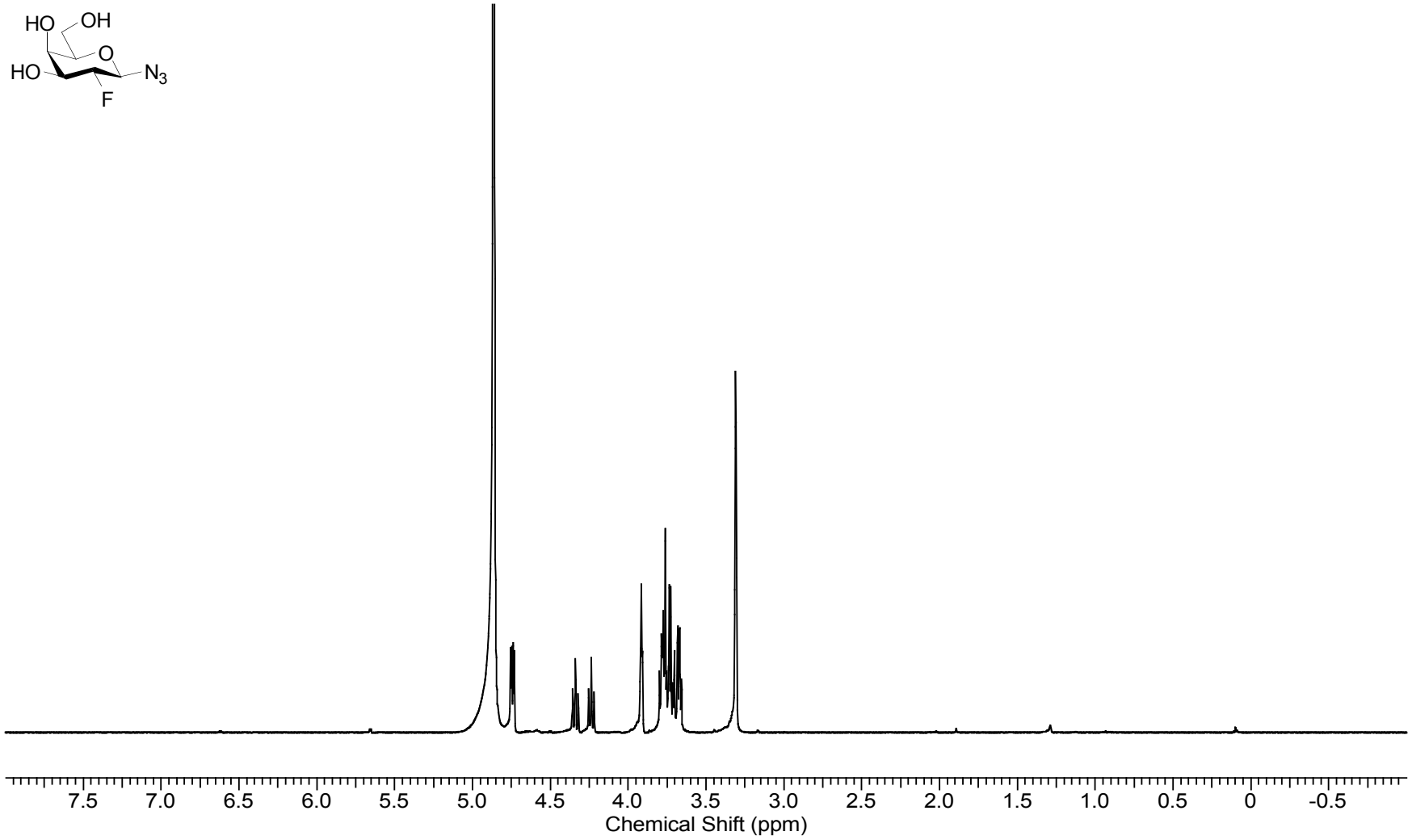
[9] S. Maschauer and O. Prante, *Carbohydr. Res.*, **2009**, *344*, 753–1761.

[10] (a) G. W. Welling, R. van der Zee, S. Welling-Wester, *Trends Anal. Chem.* **1986**, *5*, 225–230; (b) D. Corradini, D. Capitani, L. Cellai, *J. Liquid Chrom.* **1994**, *17*, 4335–4347 and references therein; (c) M. de Frutos, A. Cifuentes, J. C. Díez-Masa, E. Camafeita, E. Méndez, *J. High Resol. Chromatogr.* **1998**, *21*, 18–24 and references therein; (d) M. de Frutos, A. Cifuentes, J. C. Díez-Masa, *J. Chromatogr. A* **1997**, *778*, 43–52; (e) J. Zhang, E. Gonzalez, T. Hestilow, W. Haskins, Y. Huang, *Curr. Genom.* **2009**, *10*, 388–401; (f) G. K. C. Low, A. M. Duffield, P. R. Haddad, *Chromatographia* **1982**, *15*, 289–296; (g) S. A. Cohen, K. P. Benedek, S. Dong, Y. Tapuhi, B. L. Karger, *Anal. Chem.* 1984, 56, 217–221.

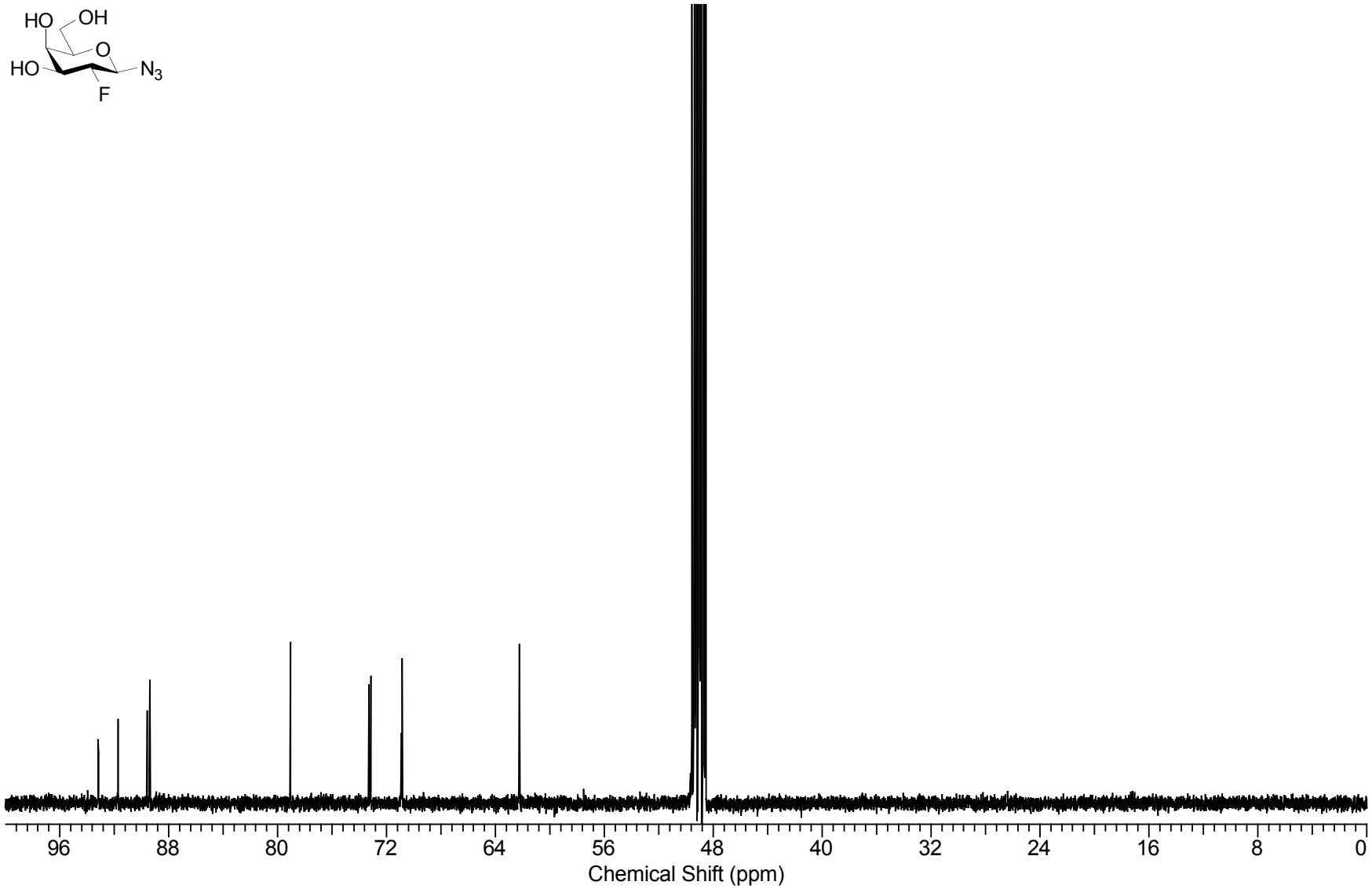
[11] B. Grego, M. T. W. Hearn, *J. Chromatrogr.* **1984**, *336*, 25–40.

[12] (a) D. J. Hnatowich, F. Chang, T. Qu, M. Rusckowski, *App. Radiat. Isot.* **1999**, *50*, 911-921.; (b) F. Wuest, C. Hultsch, M. Berndt, R. Bergmann, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5426–5428.

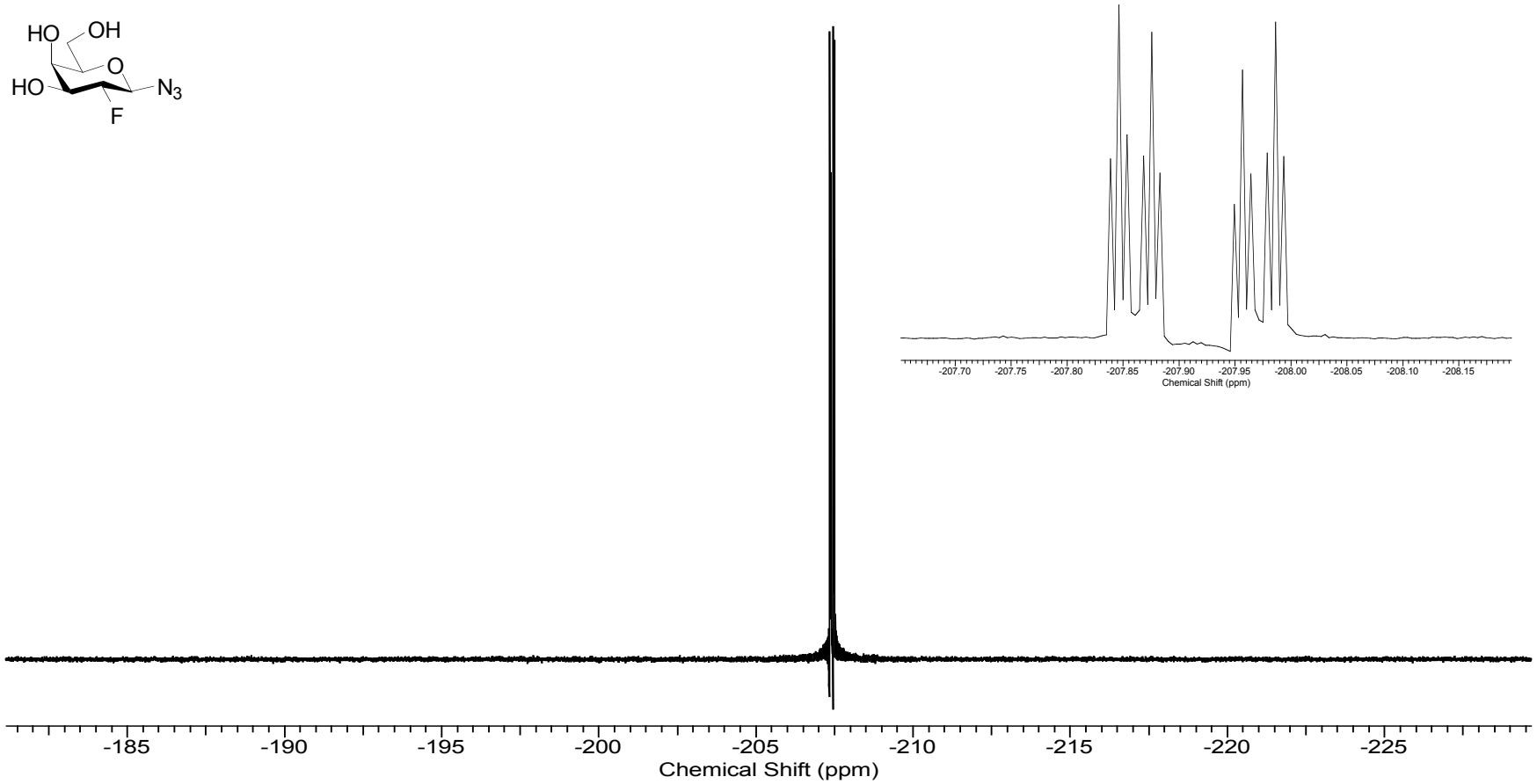
4. NMR Spectra



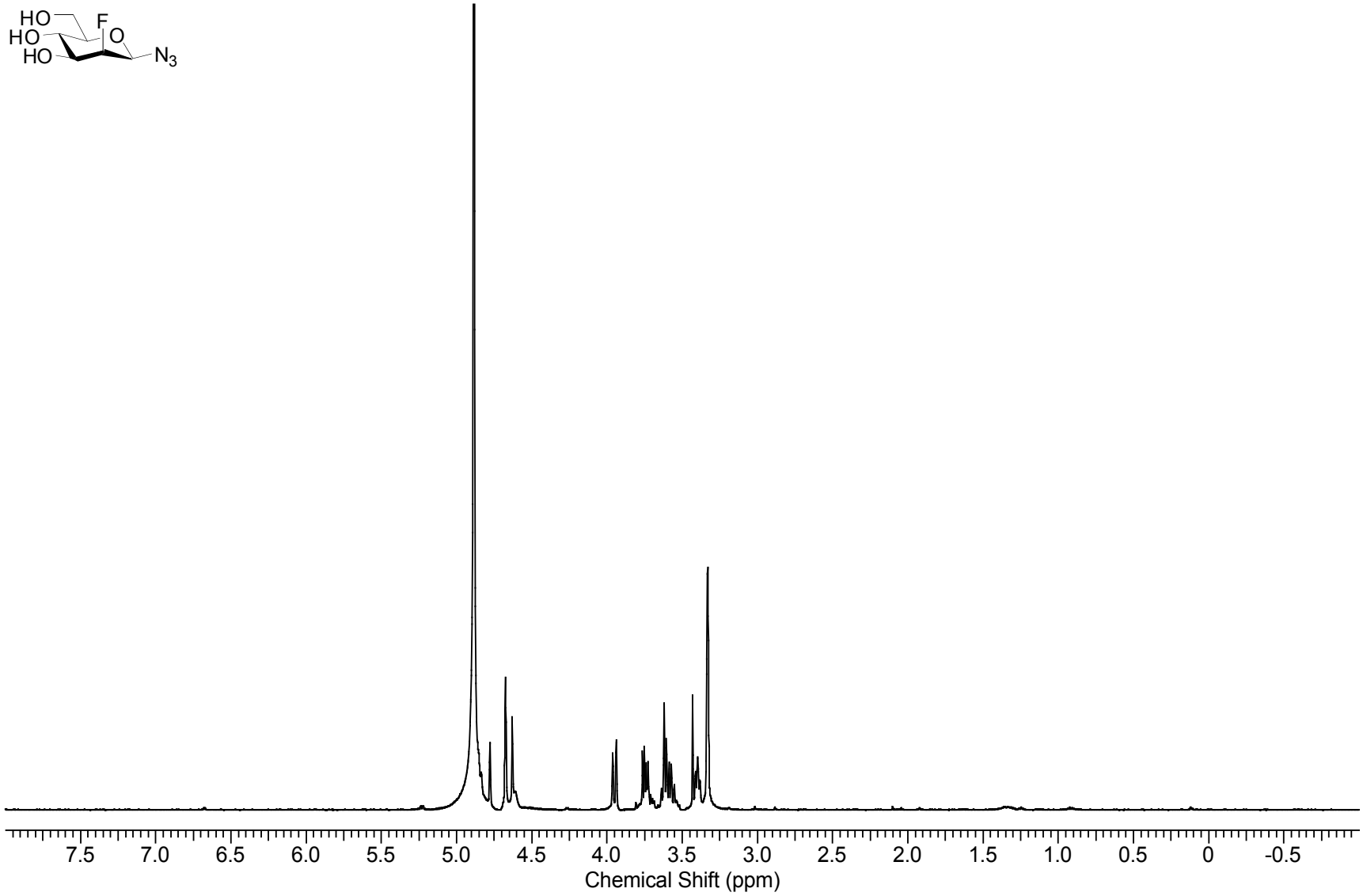
Supplementary Information



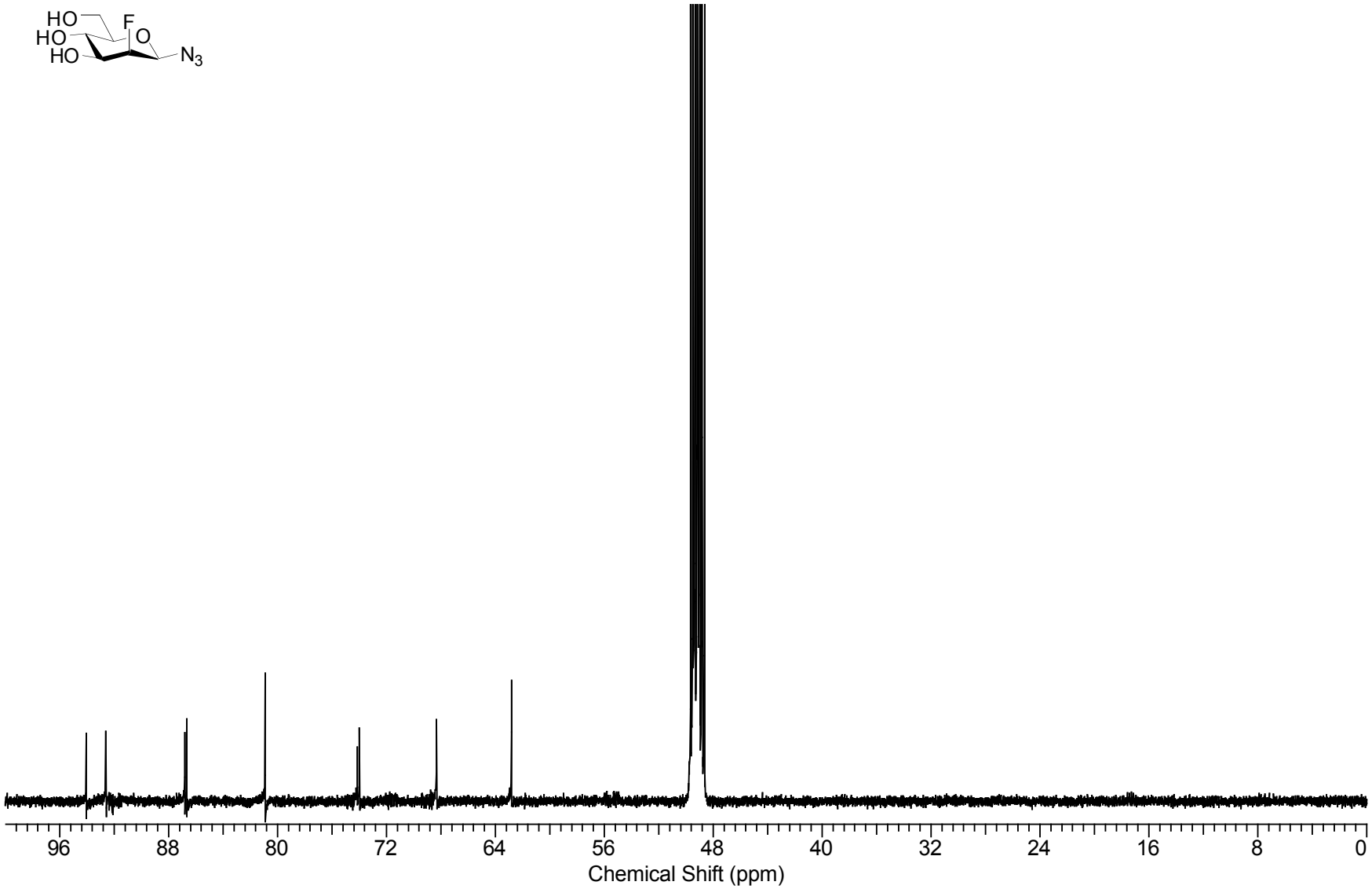
Supplementary Information



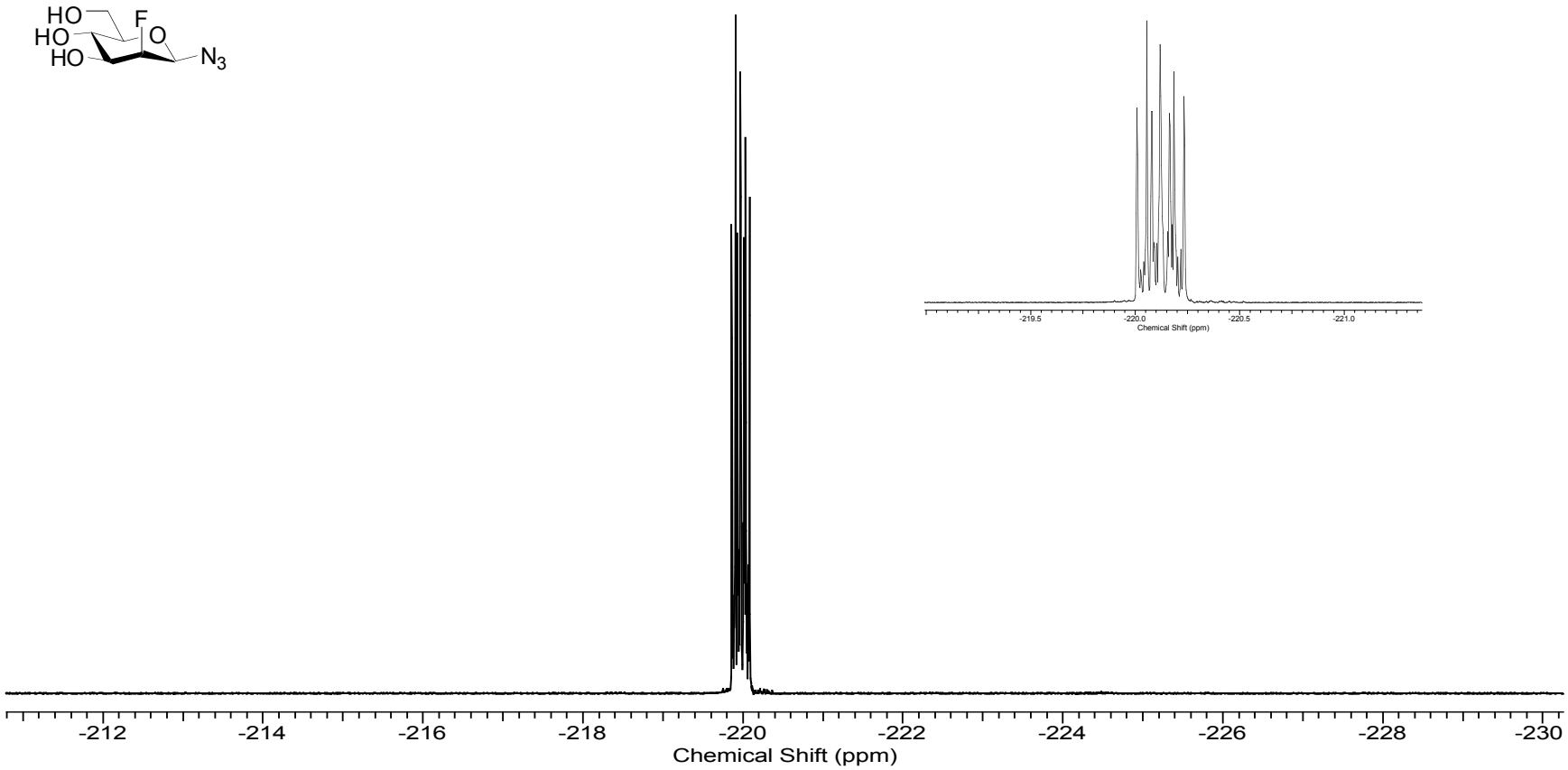
Supplementary Information



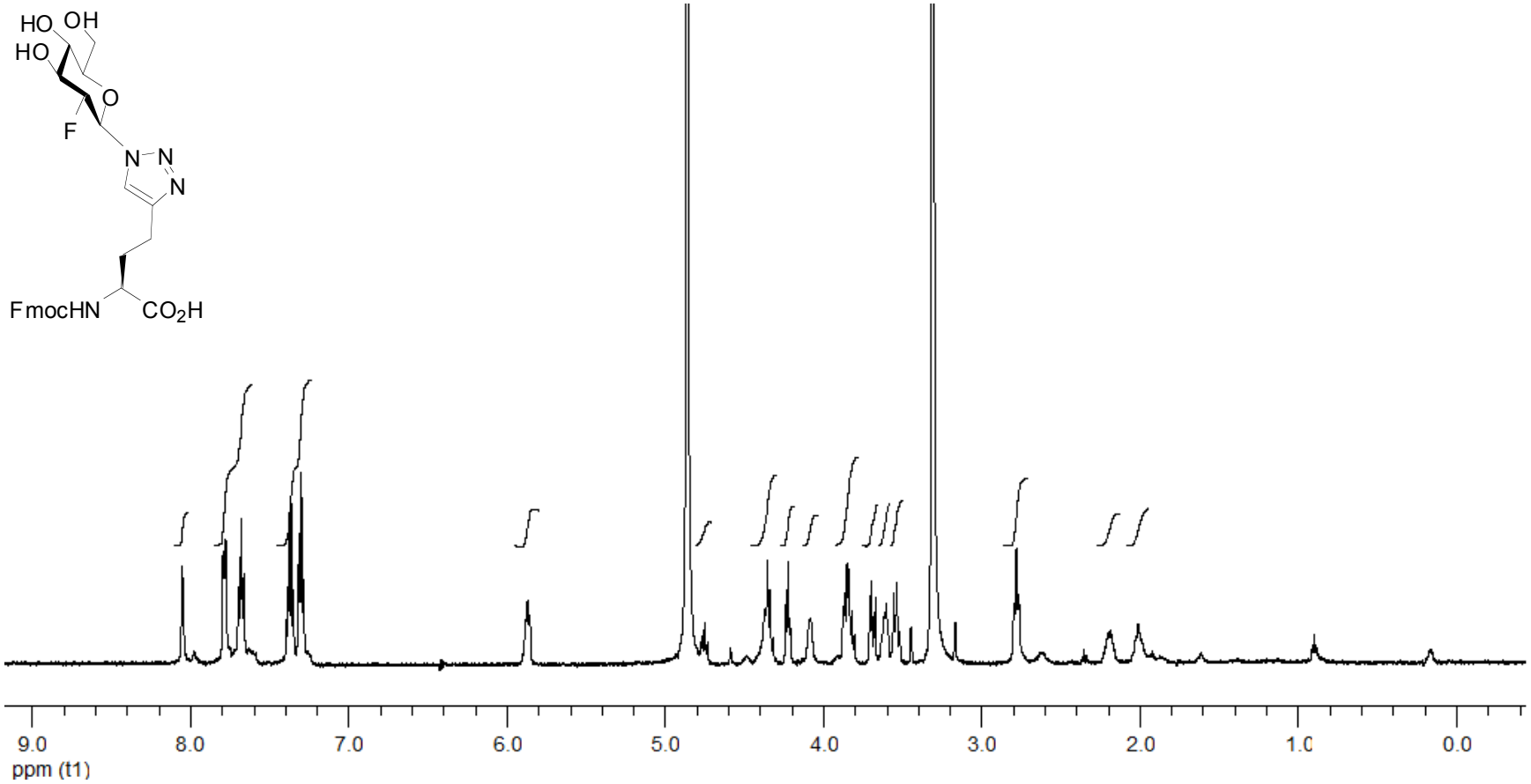
Supplementary Information



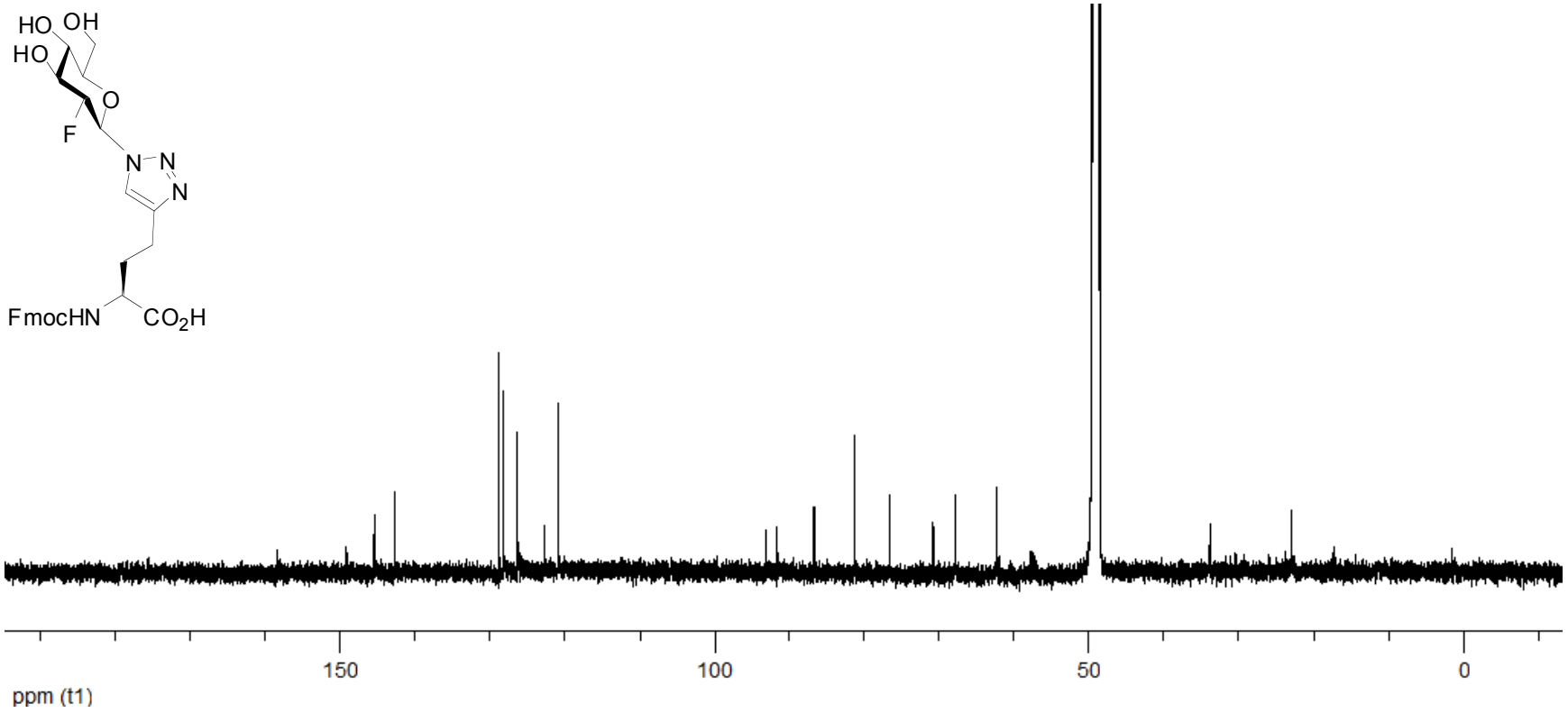
Supplementary Information



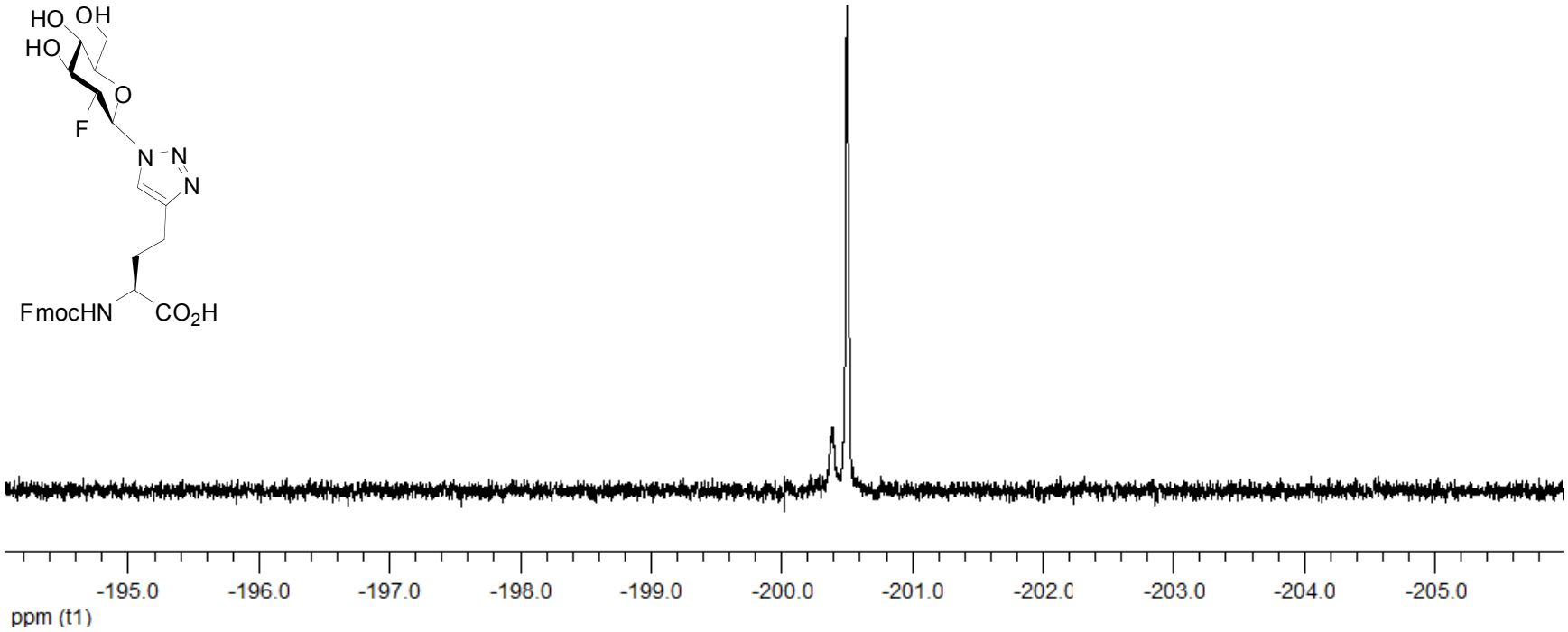
Supplementary Information



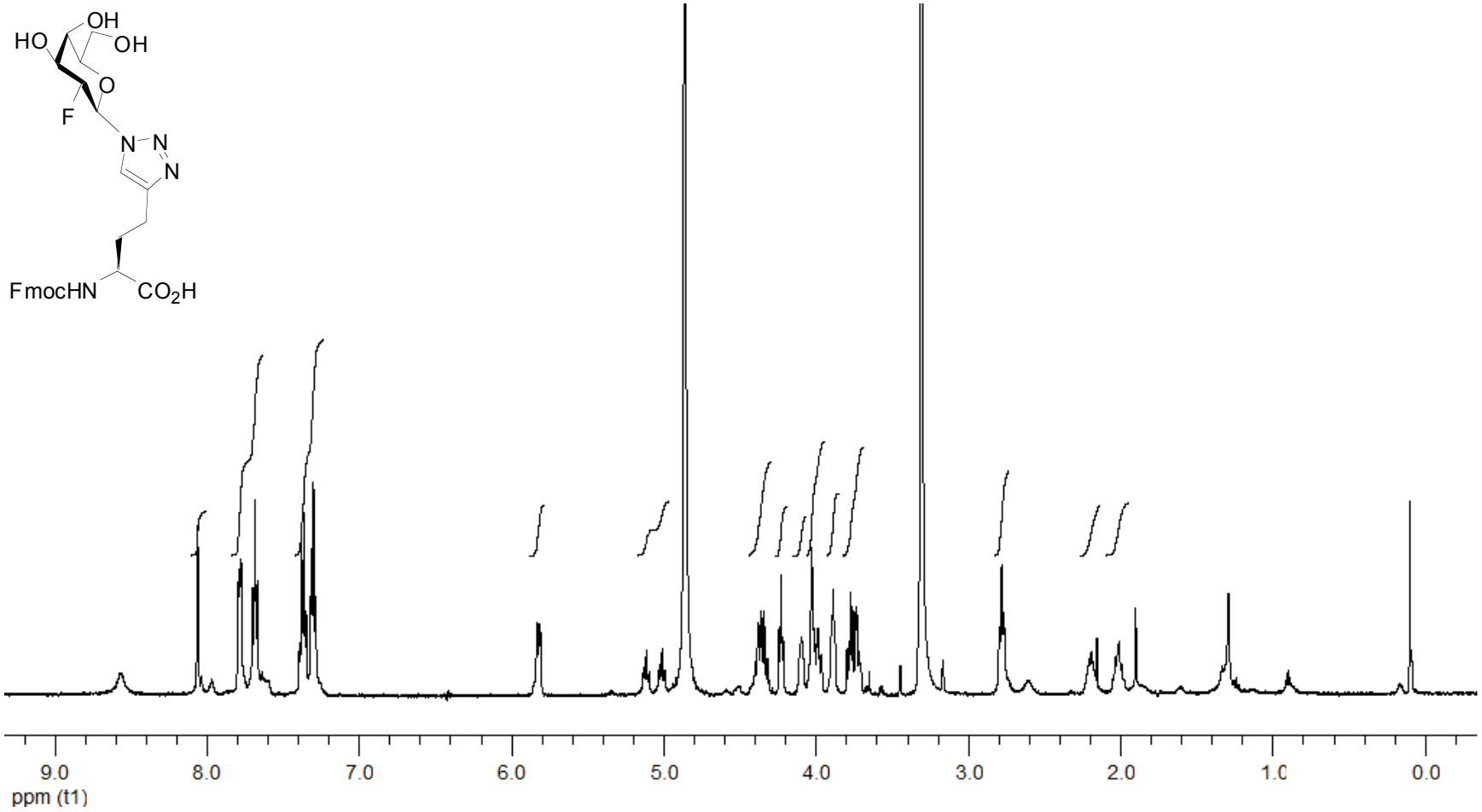
Supplementary Information



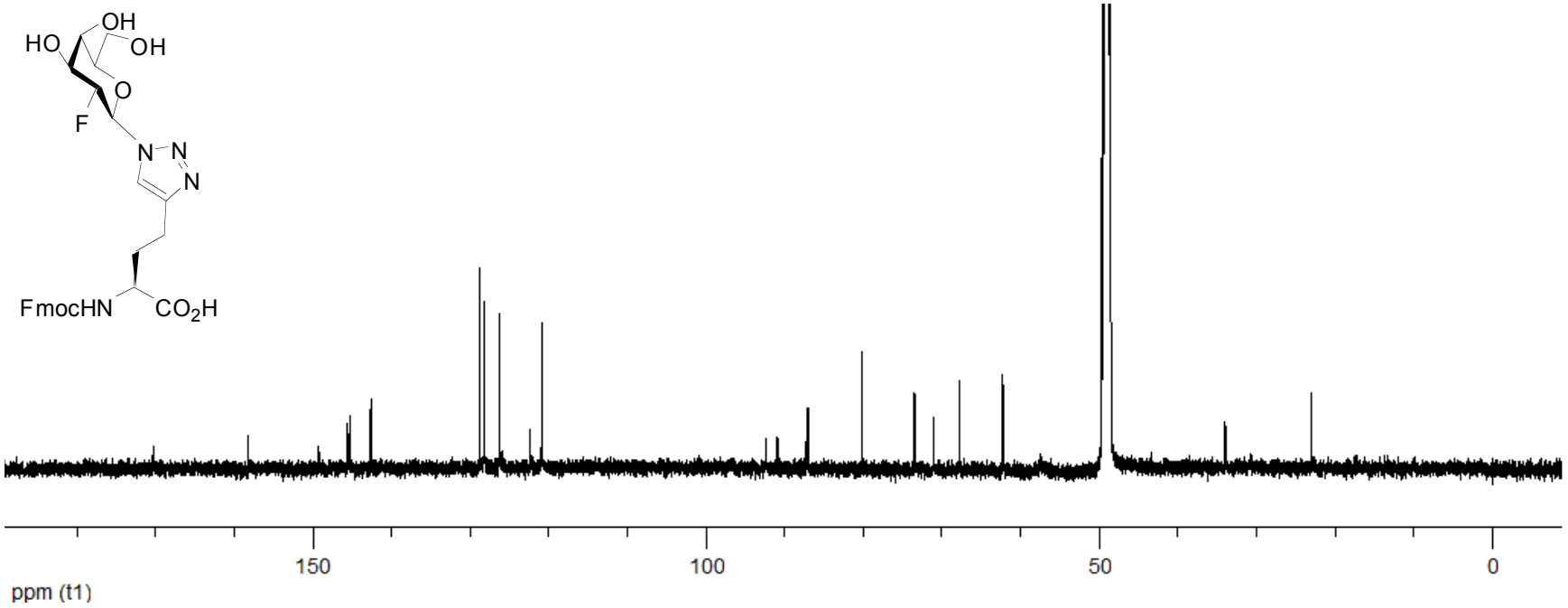
Supplementary Information



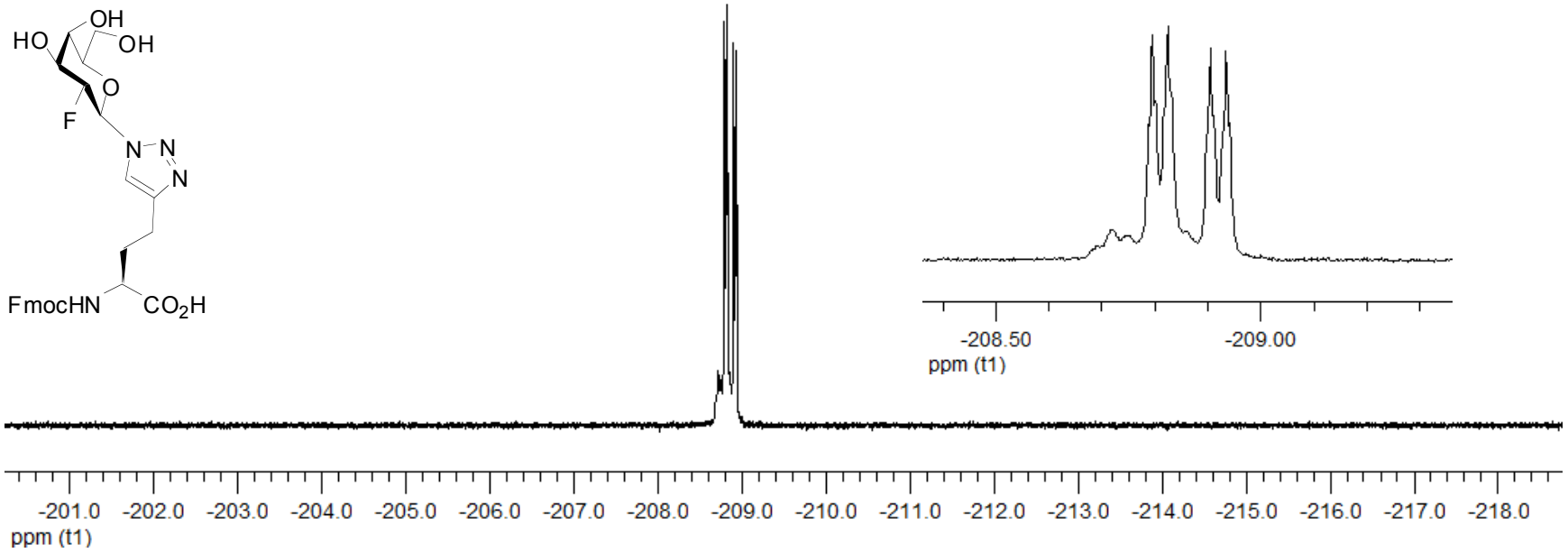
Supplementary Information



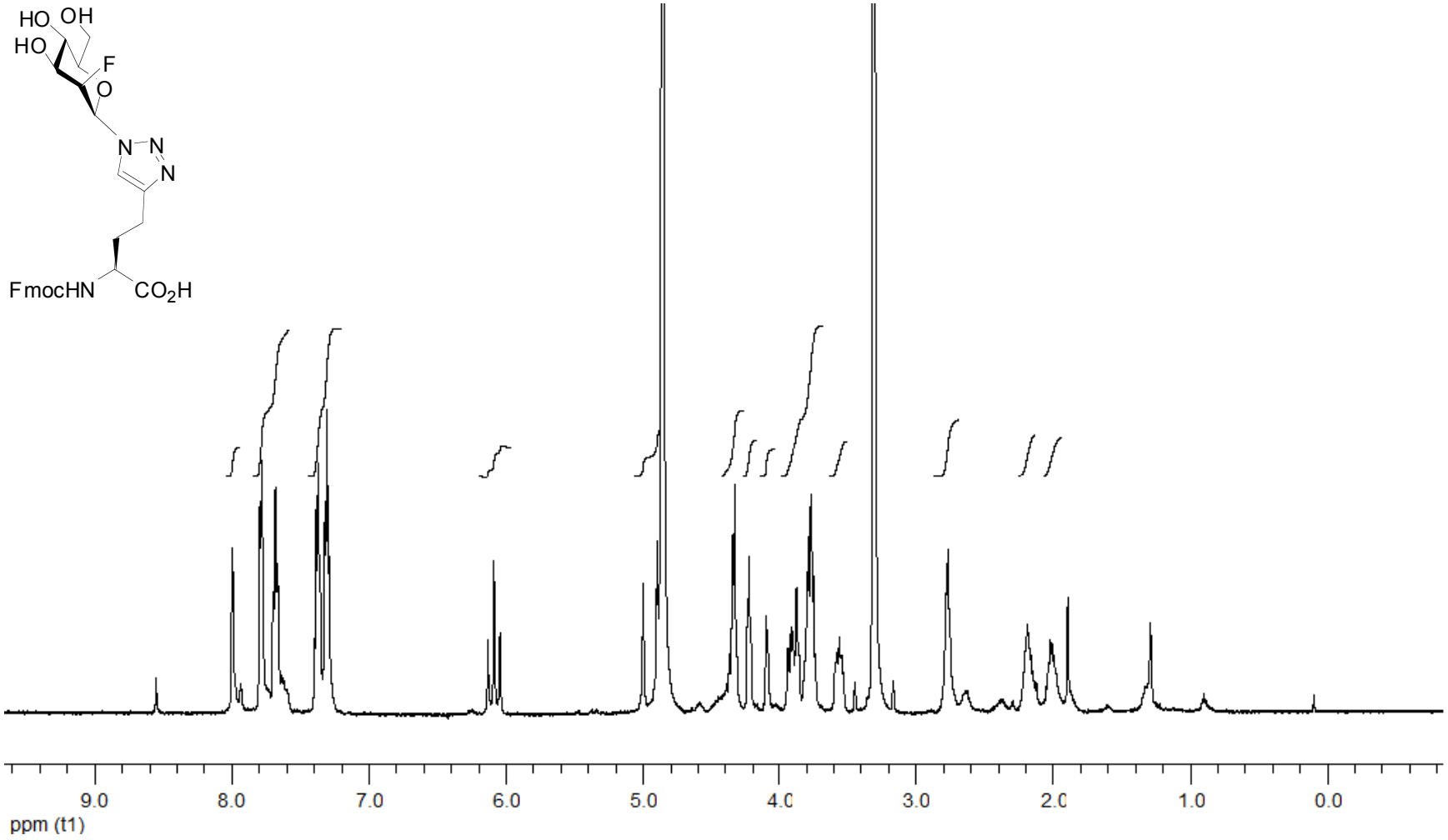
Supplementary Information



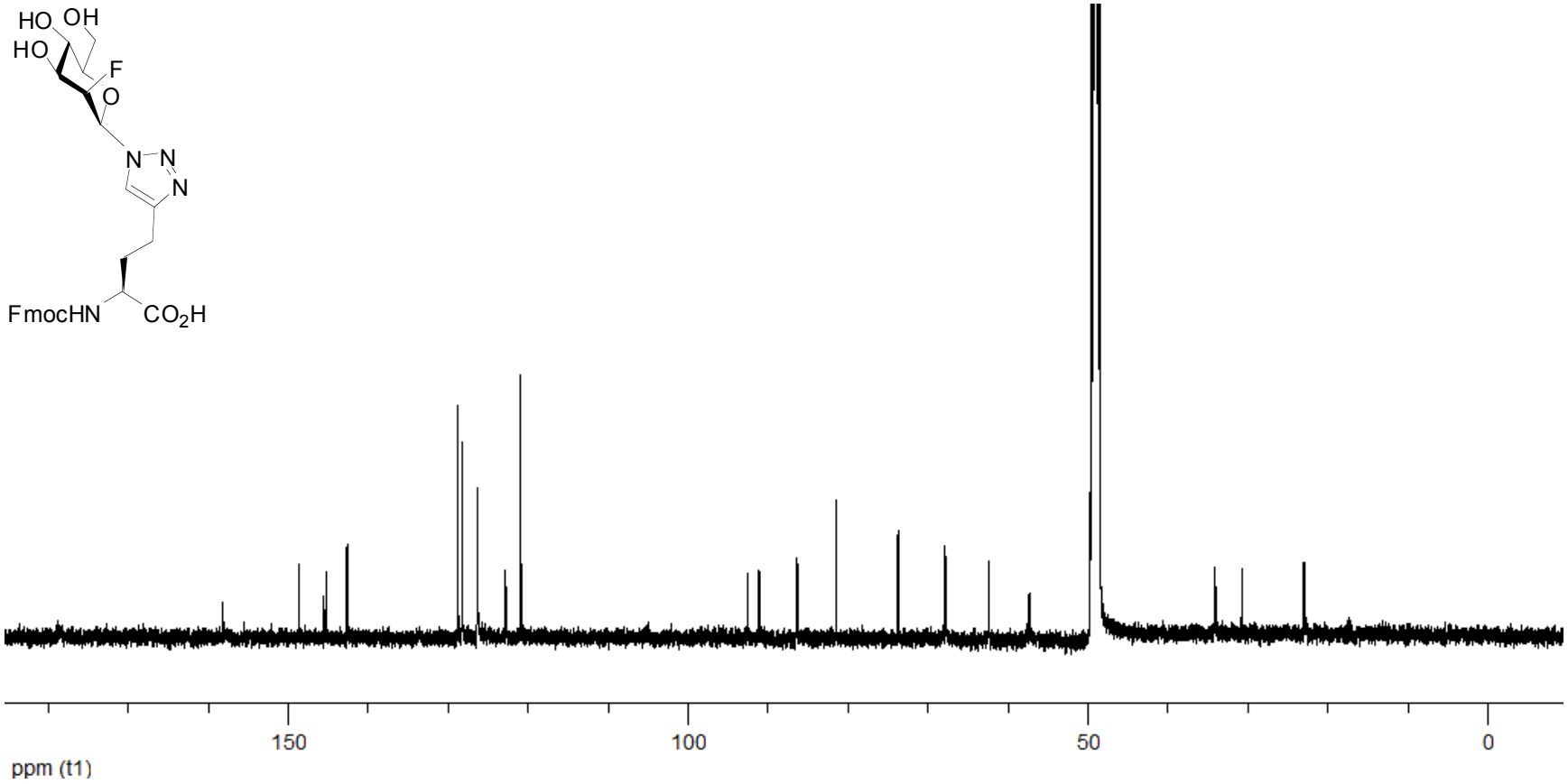
Supplementary Information



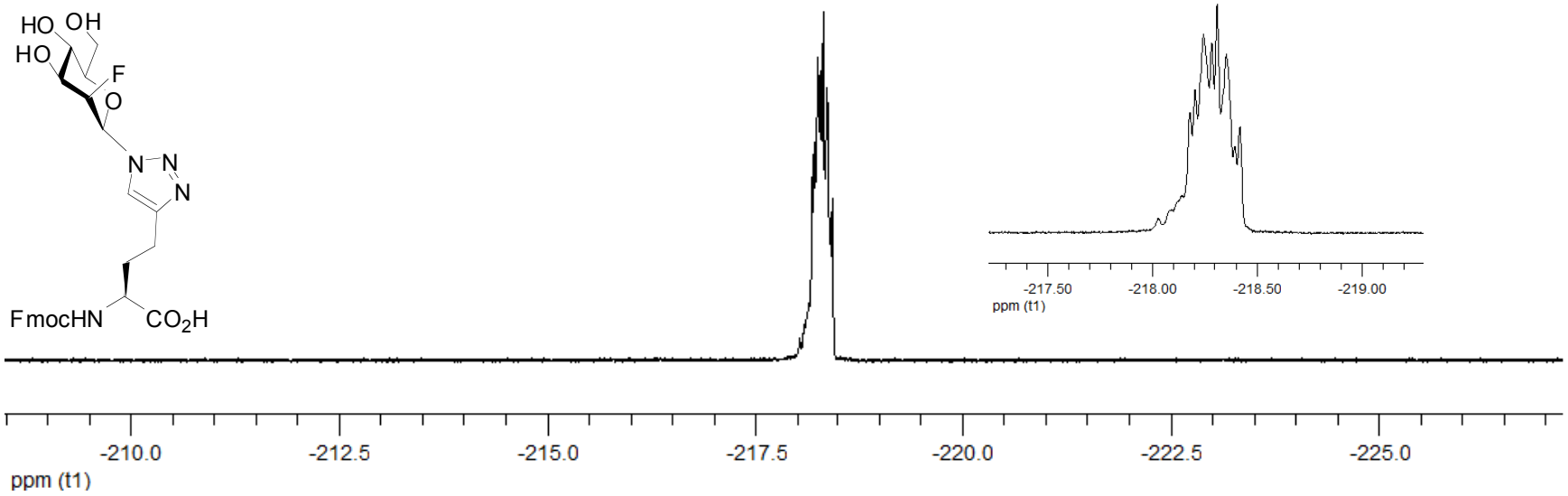
Supplementary Information



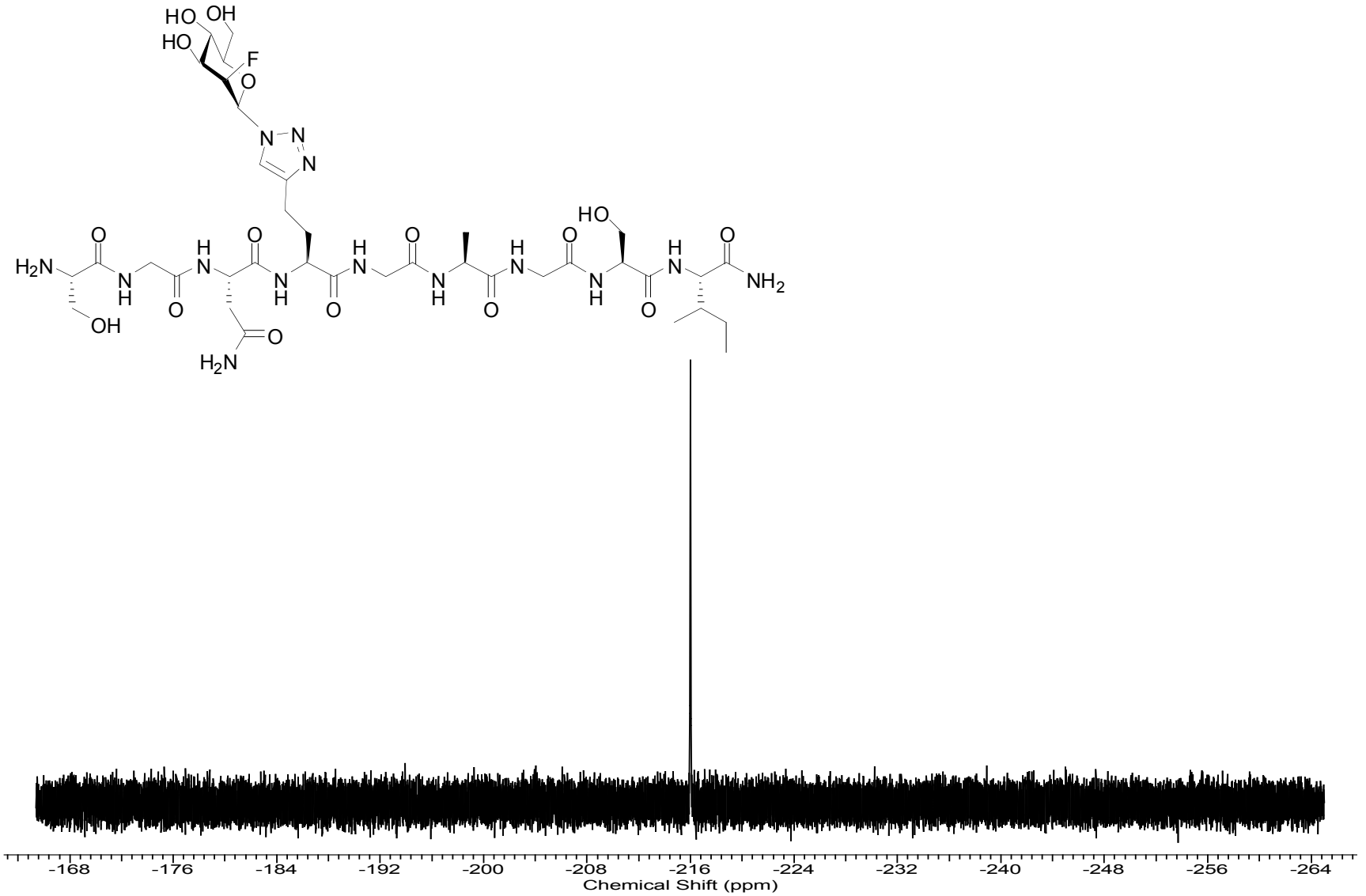
Supplementary Information



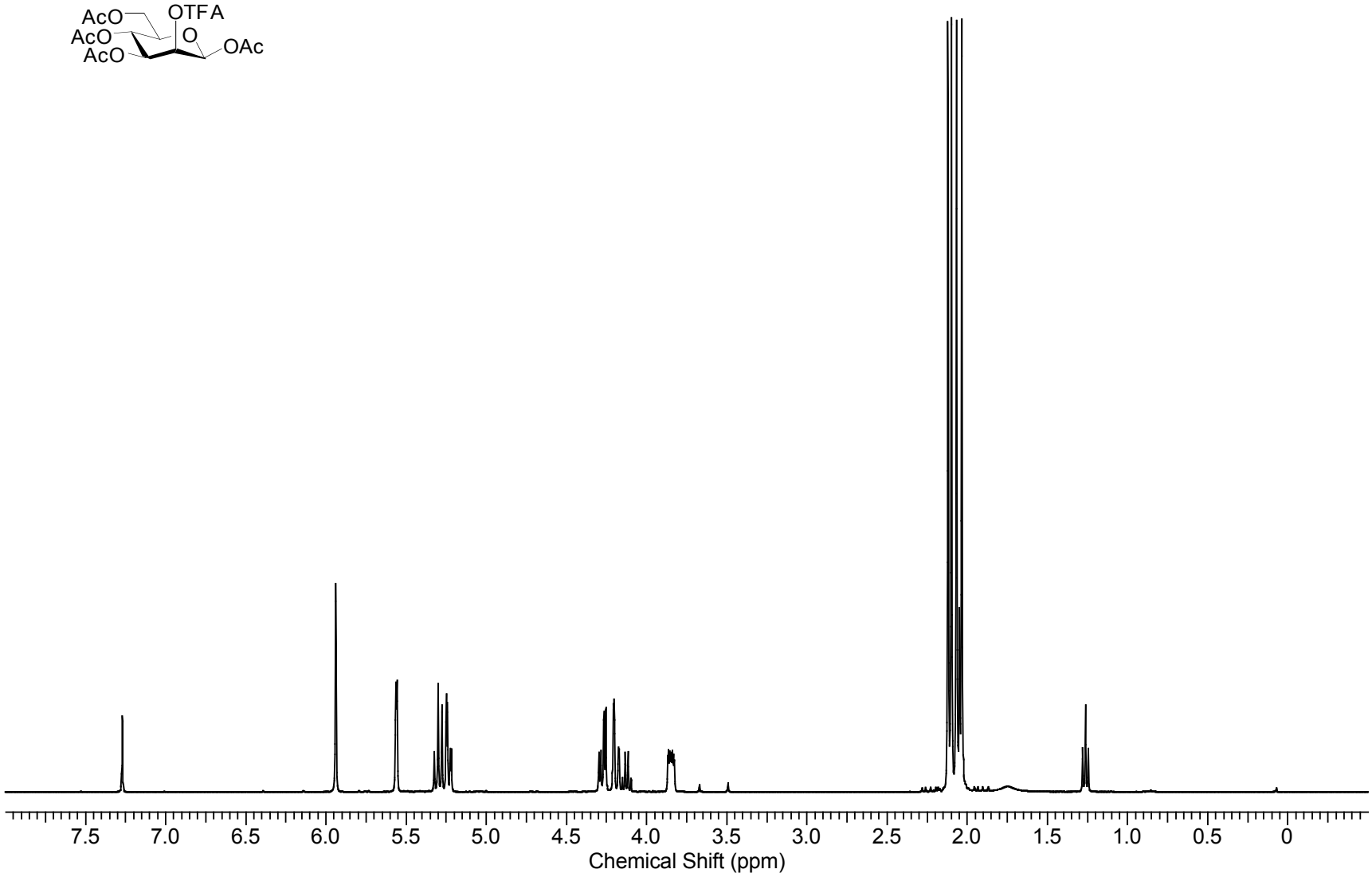
Supplementary Information



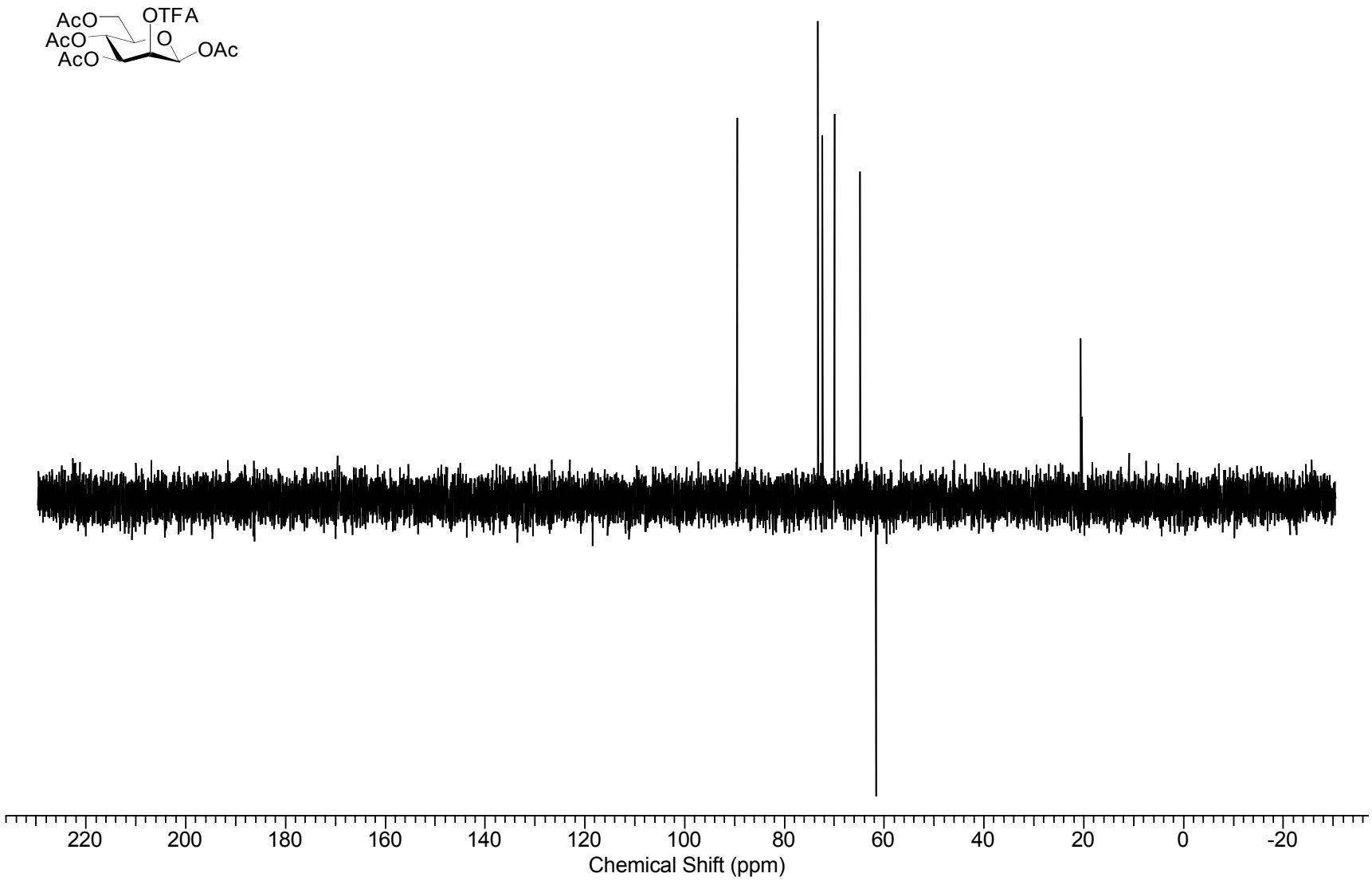
Supplementary Information



Supplementary Information

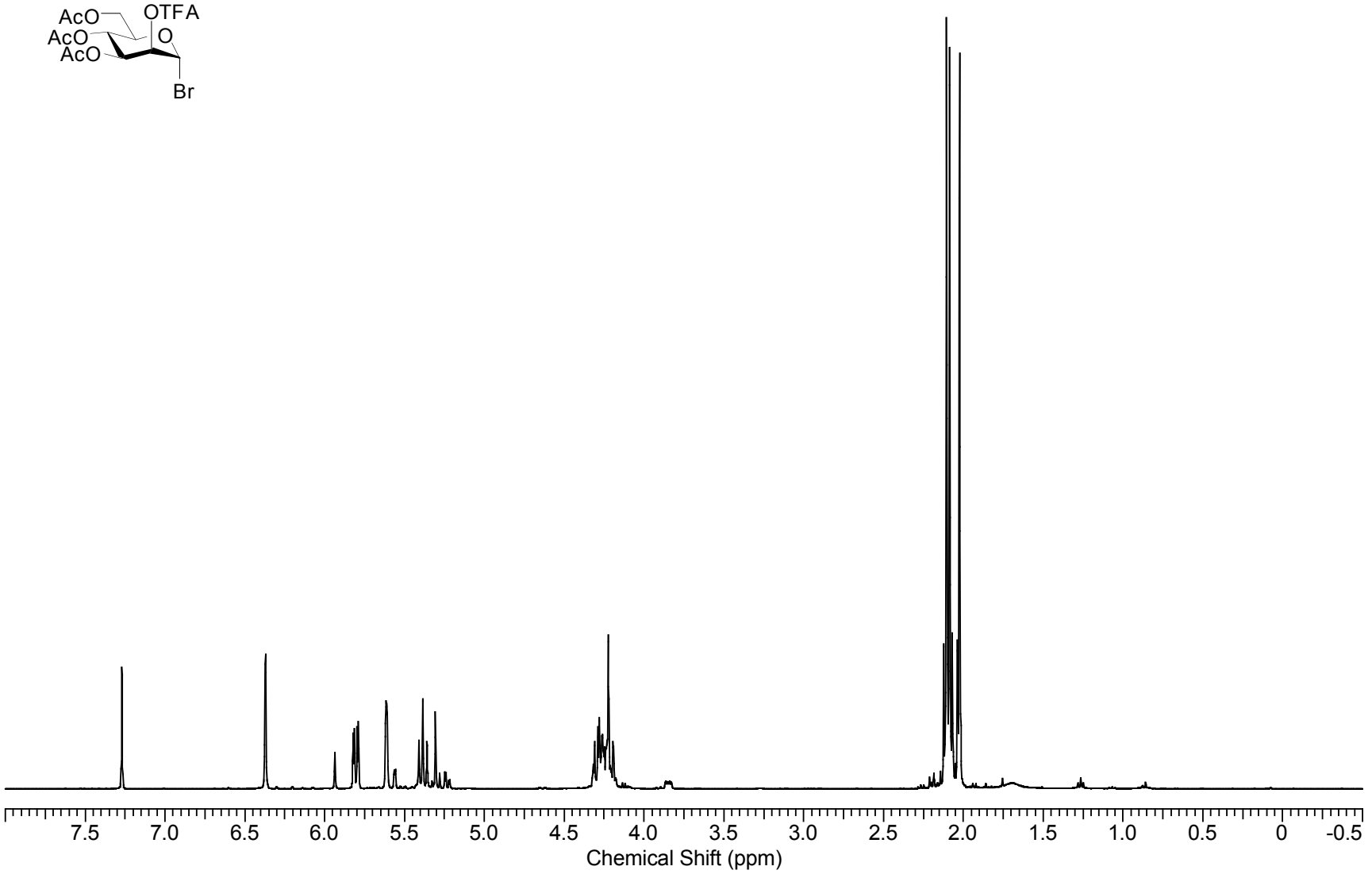


Supplementary Information

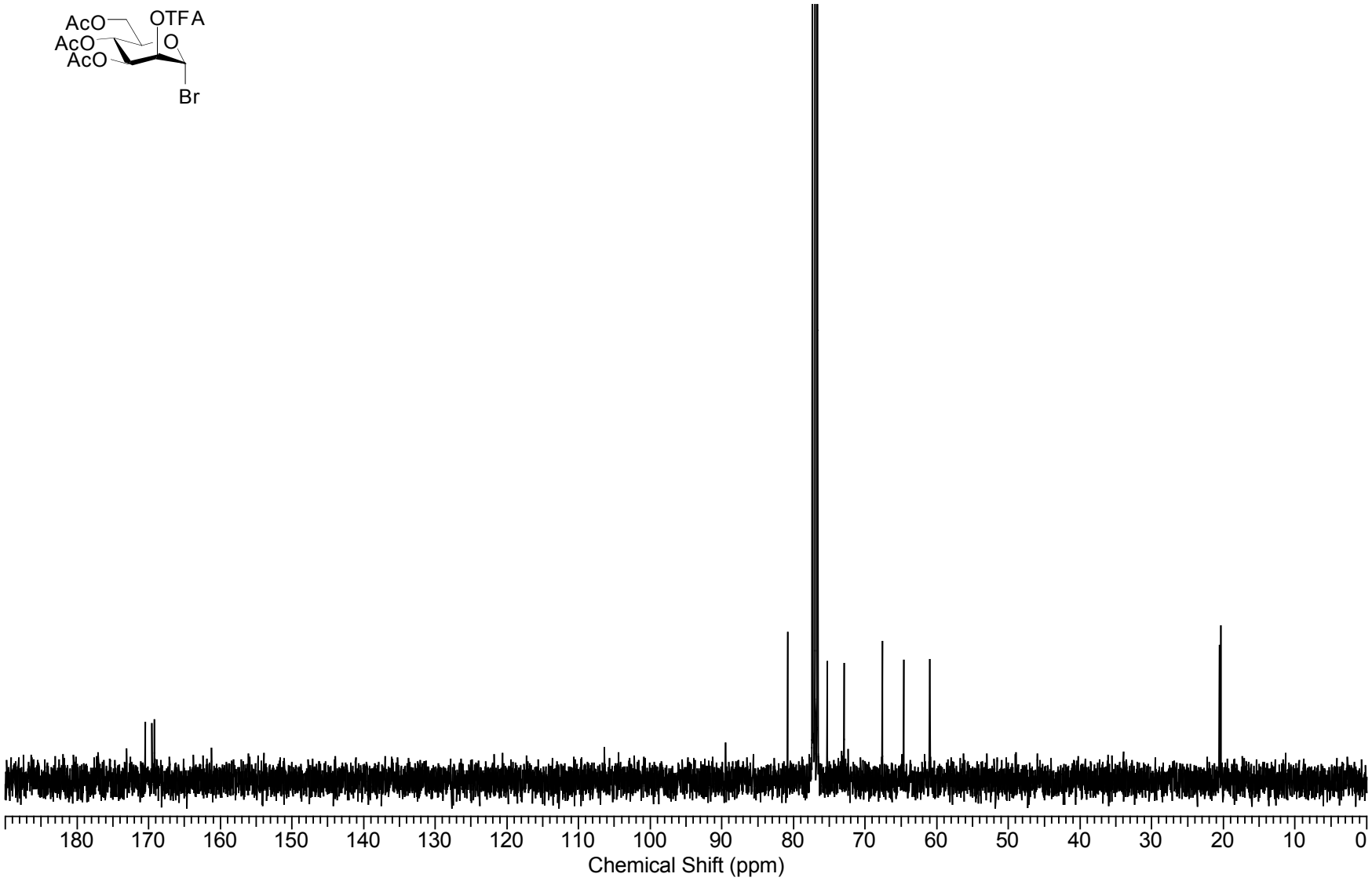


Supplementary Information

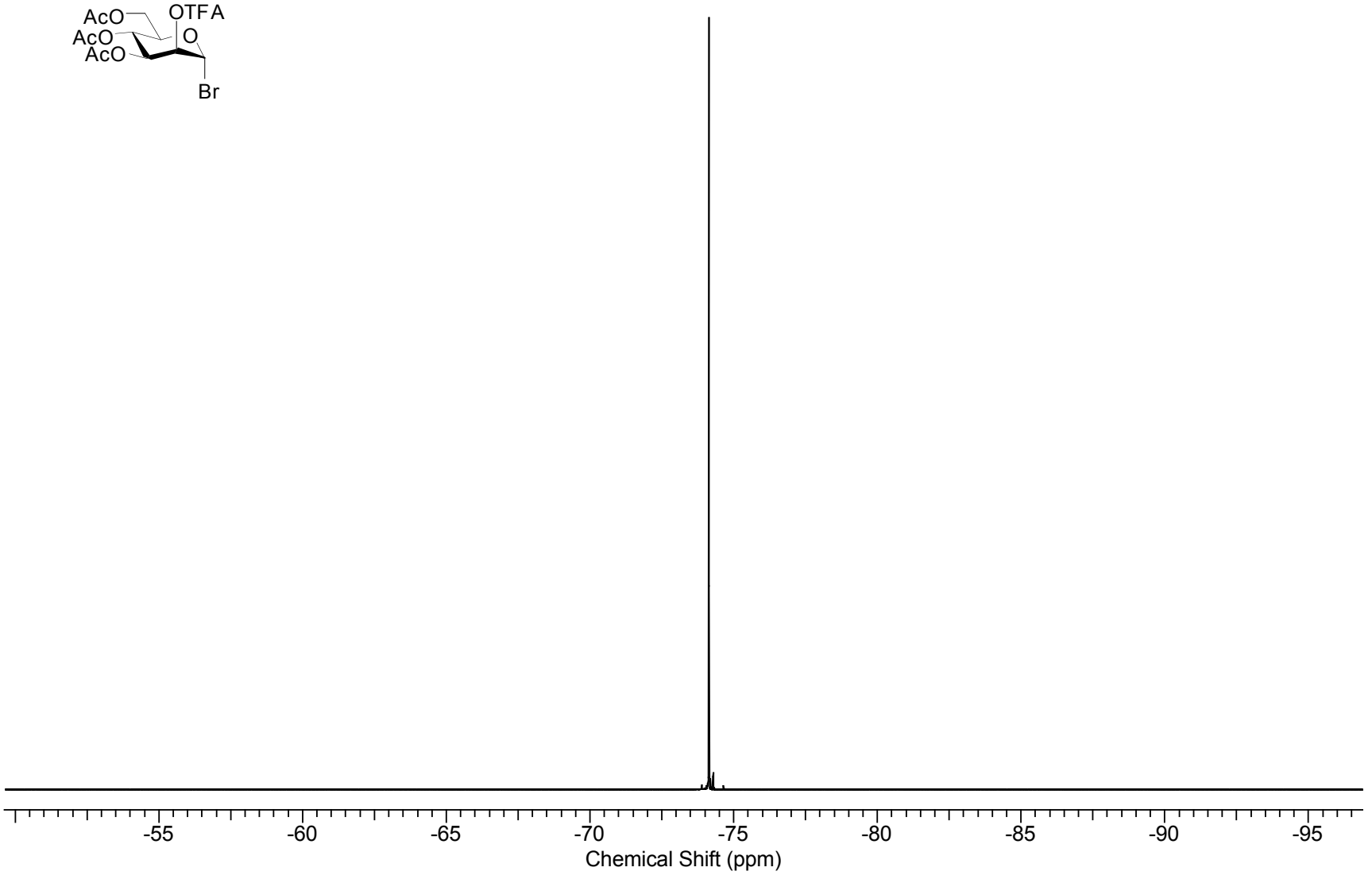
Supplementary Information



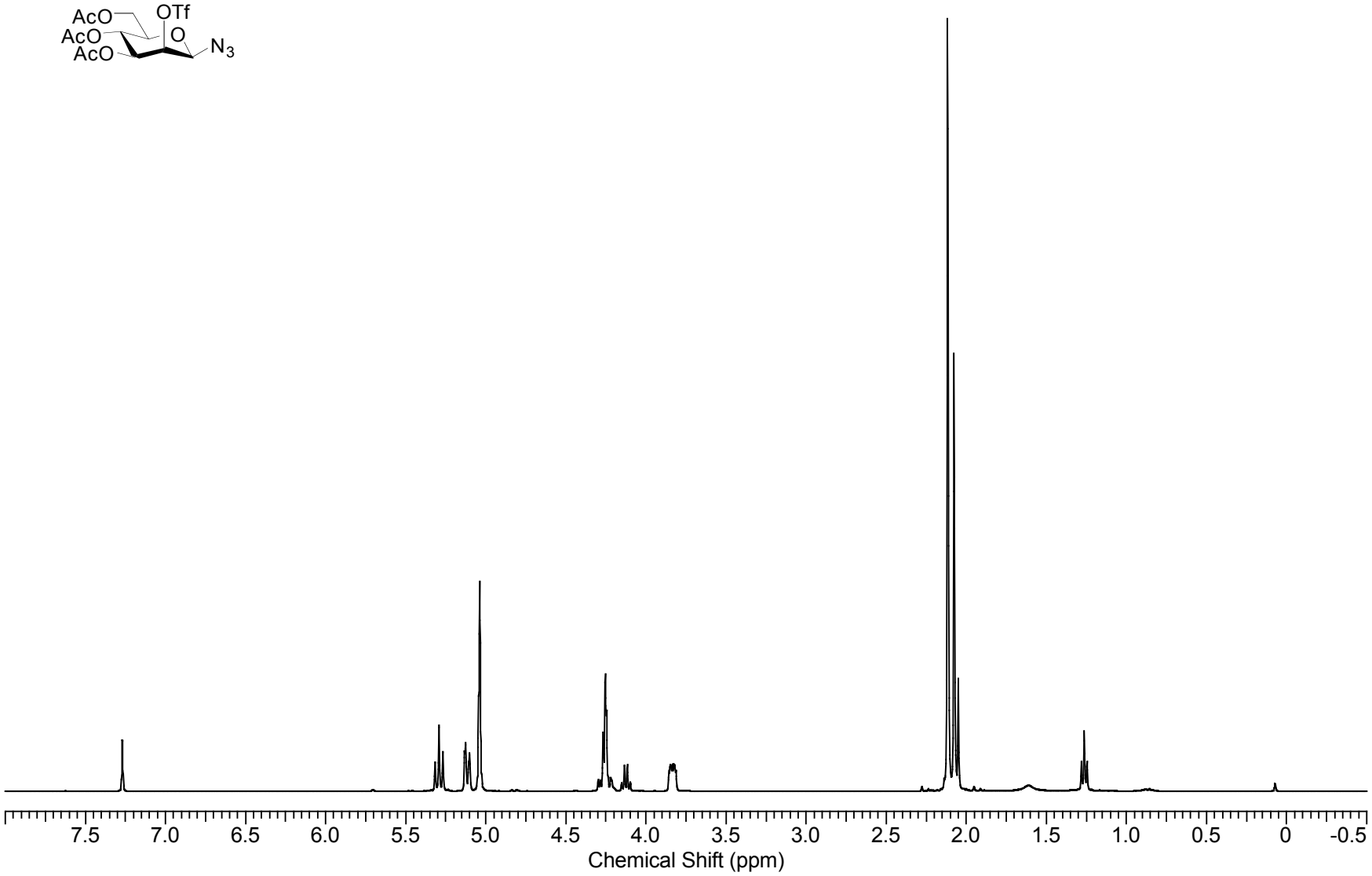
Supplementary Information



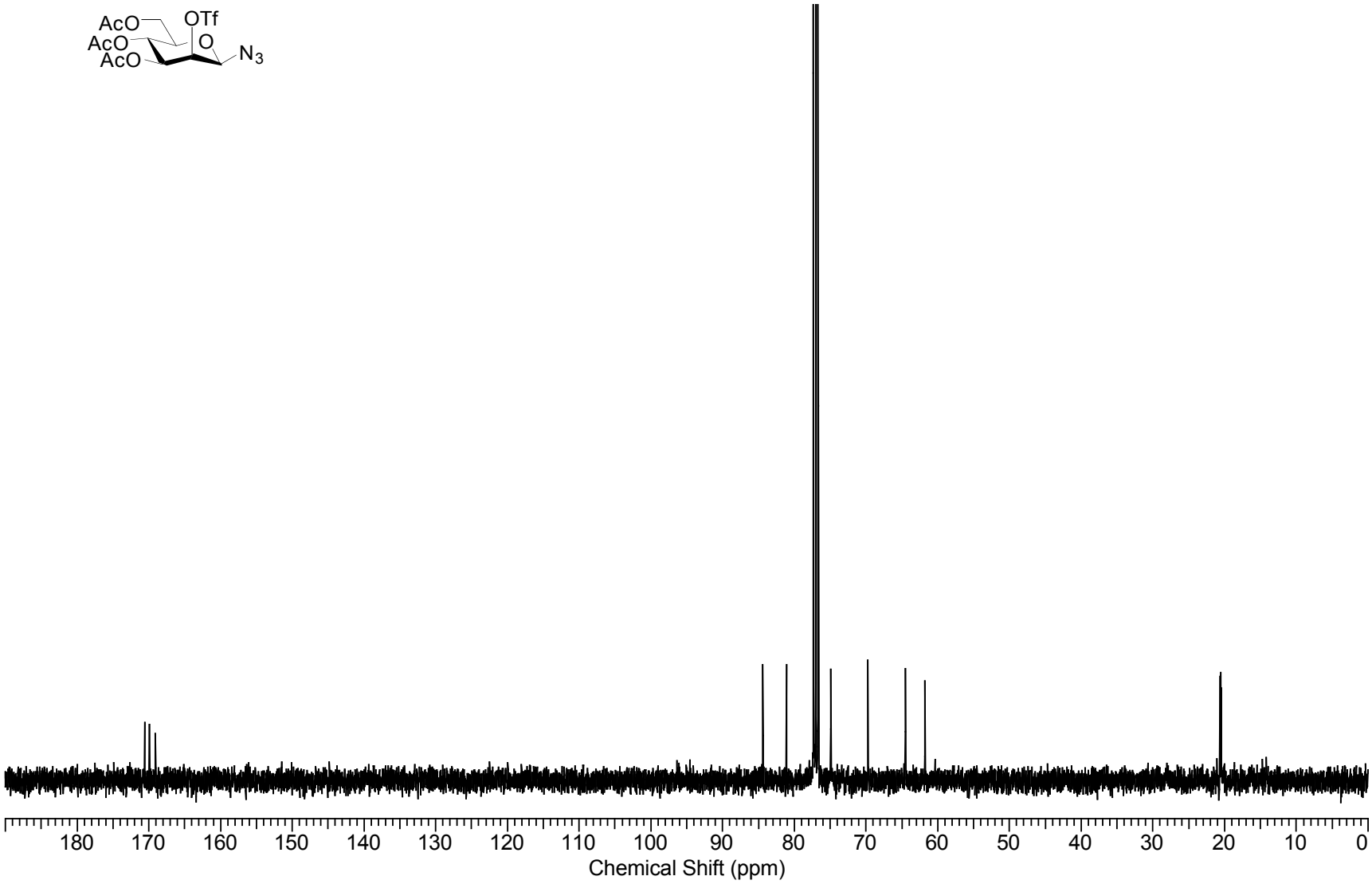
Supplementary Information



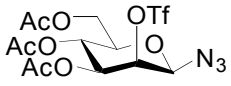
Supplementary Information



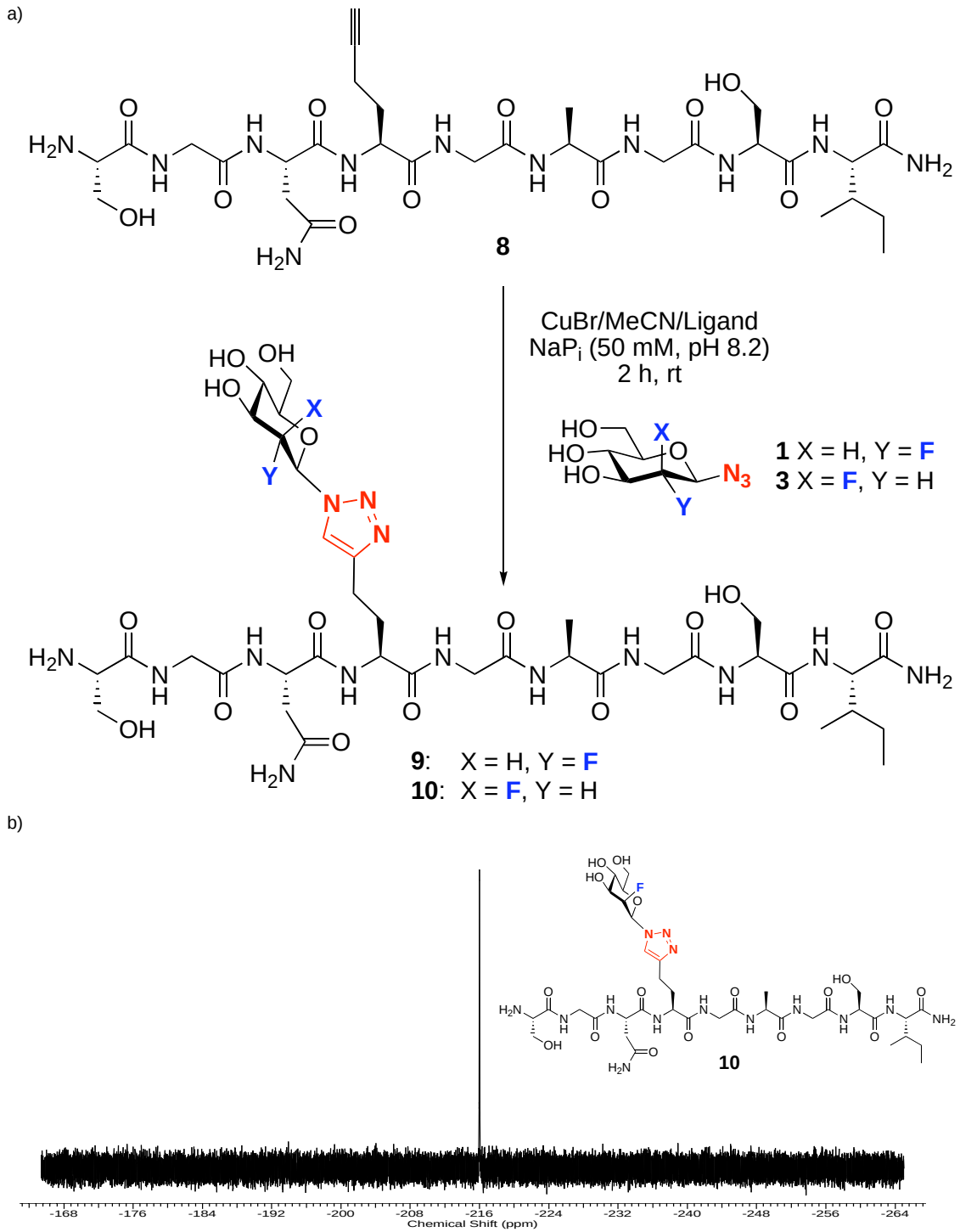
Supplementary Information



Supplementary Information



Supporting Information



Scheme S1: (a) Peptide FGlyco-CCHCs of 2-deoxy-2-fluoro-sugars **1** and **3**.
(b) $^{19}F\{^{1}H\}$ NMR experiment for **10** in D_2O (470.4 MHz)