Supplementary Information

Effect of *O*-Glycosylation and Tyrosine Sulfation of Leech-Derived Peptides on Binding and Inhibitory Activity against Thrombin.

Yves S. Y. Hsieh,[a] Deni Taleski,[a] Brendan L. Wilkinson,[a] Lakshmi C.Wijeyewickrema,[b] Ty E. Adams,[c] Robert N. Pike,[b] and Richard J. Payne*[a]

[a] *School of Chemistry, The University of Sydney, NSW 2006, AUSTRALIA. Fax：+61 2 9351 3329; Tel: +61 2 9351 3329; E-mail: richard.payne@sydney.edu.au*
[b] *Department of Biochemistry and Molecular Biology, Monash University, VIC , AUSTRALIA.*
[c] *Department of Haematology, University of Cambridge, Cambridge Institute for Medical Research, Cambridge CB2 0XY, UNITED KINGDOM.*

Contents

1. General Procedures

Liquid chromatography-mass spectrometry (LC-MS) was performed on a Shimadzu LCMS 2020 system using a LC-20AD Pump and a SPD-20A detector. Separation was achieved using a Waters SunFire™ series C18 column (150 x 2.1 mm, 5 µm particle size), at a flow rate of 0.2 mL/min over a linear gradient from 0% to 50% solvent B over 30 minutes (solvent A: 100:0.1 v/v Milli-Q water/formic acid, solvent B: 100:0.1 v/v acetonitrile/formic acid). Analytical HPLC was performed on Waters (Waters 2695 Separation Module) instrument using an analytical column (Waters SunFire™ series C18 column, 150 x 2.1 mm, 5 µm particle size), at flow rate 0.2 mL/min using two standard gradients: ***Method A***: linear gradient from 0% to 100% solvent B over 40 minutes (solvent A: 100:0.1 v/v Milli-Q water/trifluoroacetic acid, solvent B: 100:0.1 v/v acetonitrile/trifluoroacetic acid). ***Method B*:** linear gradient 0% to 50% Solvent D over 40 min (Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile). Preparative reverse-phase HPLC was performed on a Waters 2535 Quanternary gradient module, Waters 2489 UV detector operating at 230 and 254 nm and results analysed by Waters Empower 2 software. Separation was achieved on a Sunfire™ PrepC_{18} OBD™ column (5 µm, 150 × 19 mm ID) at a flow rate of 7.0 mL/min. Semi-preparative HPLC was performed on a Sunfire™ SemiPrepC_{18} OBD™ column (5 µm, 250 × 10 mm ID) at a flow rate of 4.0 mL/min. ***Method C***: Separation of non-sulfated or TCE-protected sulfopeptides utilised a mobile phase of 0.1% trifluoroacetic acid in water (Solvent A) and 0.1% trifluoroacetic acid in acetonitrile (Solvent B) using a linear gradient 0% to 60% Solvent B over 60 minutes. ***Method D***: Separation of sulfated peptides utilised a mobile phase of 0.1 M NH_4OAc (Solvent C) and 100% acetonitrile (Solvent D) using the linear gradient 0% to 50% solvent D over 60 minutes. Purified peptides in ammonium acetate eluent were recovered by repeated lyophilisation until a constant weight was achieved.

^{1}H nuclear magnetic resonance (NMR) spectroscopy of peptides were recorded using a DPX 400 spectrometer at a frequency of 400.61 MHz. ^{1}H-NMR spectroscopy of modified amino acids and reagents were recorded using a Bruker Avance DPX 300 spectrometer at a frequency of 300.13 MHz. The data is reported as chemical shift (δ), multiplicity (br = broad, s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublets, ddd = doublet of doublet of doublets, m = multiplet), coupling constant (*J* Hz) and relative integral where possible. ^{13}C nuclear magnetic resonance spectra were recorded on a Bruker Avance DPX 400 spectrometer at a frequency of 100.61 MHz, or a Bruker Avance DPX 300 spectrometer at a frequency of 75.47 MHz and are reported in parts per million (ppm) relative to the residual solvent peak.

Materials

Solid-phase synthesis was carried out in polypropylene syringes equipped with Teflon filters and were purchased from Torviq. Commercial materials were used as received unless otherwise noted. Amino acids, coupling reagents and resins were obtained from Novabiochem or GL Biochem. Dichloromethane (DCM) and methanol (MeOH) were distilled from calcium hydride. *N,N*-dimethylformamide (DMF) was obtained as peptide synthesis grade from EMD Chemicals. Acetic anhydride and pyridine were obtained from Ajax. *N,N*-diisopropylethylamine (DIPEA) and *N*-methylmorpholine (NMM) were obtained from Aldrich. Piperidine was purchased from Alfa Aesar.

1.1 General procedures for Solid-Phase Peptide Synthesis (SPPS)

1.1.1 Amino acid loading onto Wang resin

Wang resin LL (100-200 mesh, Novabiochem, loading = 0.44 mmol/g, 150 μmol) was washed with DMF (5 x 5 mL), DCM (5 x 5 mL), DMF (5 x 5 mL), and allowed to swell in DMF (3 mL) for 1 h. Fmoc-Ser(*t*Bu)-OH (10 equiv, 1.5 mmol) was dissolved in dry DCM (10 mL) under N_2 and cooled on ice before addition of DIC (116 μL, 0.75 mmol) in DCM (5 mL). The reaction was stirred for 30 min before the solvent was removed *in vacuo*. The resulting residue was redissolved in DMF (4 mL) containing DMAP (2 mg, 15 μmol) and added to the resin. The resin was shaken for 1 h and the solution drained before washing with DMF (5 x 5 mL), DCM (5 x 5 mL) and DMF (5 x 5 mL). The resin was subsequently capped with acetic anhydride/pyridine (1:9 v/v, 4 mL) for 3 min before the resin was washed with DMF (5 x 5 mL), DCM (5 x 5 mL) and DMF (5 x 5 mL).

1.1.2 Determination of resin loading

The *N*-terminal Fmoc protecting group was removed according to the general procedure outlined below and the combined drained Fmoc-deprotection solutions diluted with a solution of piperidine/DMF (1:9 v/v) so that the maximum concentration of the fulvene-piperidine adduct was in the range of 2.5–7.5 × 10^{-5} M. A sample of this solution was transferred to two matched 1 cm quartz cuvettes and the UV-Vis absorbance measured at λ = 301 nm using a solution of piperidine/DMF (1:9 v/v) as a reference. An average of the two absorbance values was used to calculate the resin loading using the molar extinction coefficient for the piperidine-fulvene adduct (ε = 7800 M^{-1} cm^{-1}).

1.1.3 Iterative Fmoc-strategy SPPS (25 µmol scale)

***N*-terminal Fmoc deprotection:** A solution of piperidine/DMF (2 × 5 mL, 1:9 v/v) was added to the resin and agitated for 5 min. The resin was subsequently drained and washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL). The resulting resin-bound amine was used immediately in the next peptide coupling step. The efficiency of previous amino acid couplings were determined by measurement of the resulting fulvene-piperidine adduct at λ = 301 nm as described above.

Amino acid coupling: A solution of Fmoc-protected amino acid (100 µmol), PyBOP (52 mg, 100 µmol) and NMM (22 µL, 200 µmol) in DMF (1 mL) was added to the resin and the resulting suspension gently agitated for 1 h. The resin was then drained and washed sequentially with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL). In the case of TBS-protected tyrosine amino acid **8**, coupling was performed using 1.3 equiv. of the amino acid, PyBOP (1.3 equiv.) and NMM (2.6 equiv.) in DMF (1 mL).

Capping: A mixture of acetic anhydride/pyridine (5 mL, 1:9 v/v) was added to the resin and agitated at rt for 5 min. At this time the resin was drained and washed with DMF (5 × 3 mL), DCM (5 × 3 mL) and DMF (5 × 3 mL).

1.2 General procedures for Microwave-assisted Fmoc-strategy SPPS (150 µmol scale)

Automated microwave-assisted solid-phase peptide synthesis was carried out using a Liberty 1 microwave peptide synthesiser (CEM Corporation) equipped with a 30 mL Teflon reaction vessel.

Fmoc-protected amino acids were made up as 0.2 M solutions in DMF. An activator solution was made up as 0.45 M HBTU in DMF. An activator base solution was made up as 2 M DIPEA in NMP. A capping solution was made up as 90/8/2 v/v/v DMF/Ac_2O/DIPEA.

1.2.1 Instrument settings for deprotection, coupling and capping

Methods for deprotection and coupling of each amino acid are described in Table 1. After each amino acid coupling, the capping method described in Table 2 was used. Microwave power, reaction time and temperature parameters for the deprotection and coupling of Fmoc-His(Trt)-OH and Fmoc-Arg(Pbf)-OH were modified and are described in 1.2.2.

Table 1. Standard deprotection and coupling steps on the Liberty 1.

Operation	Parameter	Volume (mL)	Drain
Wash-Top		10	Y
Add deprotection	Initial Deprotection	10	
Microwave		10	Y
Wash-Top		10	Y
Add deprotection	Standard Deprotection	10	
Microwave		10	Y
Wash-Top		10	Y
Add deprotection	Standard Deprotection	10	
Microwave			Y
Wash-Top		10	Y
Wash-Bottom		10	Y
Wash-Top		10	Y
Wash-Top		10	Y
Add Amino Acid		5	
Add Activator		2	
Add Activator Base		1	
Microwave	Standard Coupling		Y
Wash top		10	Y
Wash bottom		10	Y

Table 2: Standard capping steps on the Liberty 1.

Operation	Parameter	Volumn (mL)	Drain
Wash-Top		10	Y
Wash-Bottom		10	Y
Wash-Top		10	Y
Wash-Top		10	Y
Add capping reagent		10	
Microwave method	Standard coupling		Y
Wash-Top		10	Y
Wash-Bottom		10	Y
Wash-Top		10	Y

1.2.2 Parameter settings

Parameters for standard Fmoc-deprotection and amino acid coupling

Initial deprotection: 0 W, 60 s, 25 °C

Deprotection: 40 W, 180 s, 75 °C

Coupling: 45 W, 180 s, 75 °C

Parameters for deprotection and coupling of Fmoc-His(Trt)-OH

Initial deprotection: 0 W, 60 s, 25 °C

Deprotection: 30 W, 30 s, 50 °C

Coupling: Start with 0 W and 50 °C for 360 s, followed by 15 W, 50 °C for 180 s

Parameters for deprotection and coupling of Fmoc-Arg(Pbf)-OH

Initial deprotection: 0 W, 60 s, 25 °C

Deprotection: 30 W, 30 s, 50 °C

Coupling: Start with 0 W and 75 °C for 1500 s, followed by 30 W, 75 °C for 300s

1.3 General procedure for solid-phase sulfation (25 μmol scale)

1.3.1 TBS *O*-tyrosine ether deprotection (25 μmol scale)

A solution of tetrabutylammonium fluoride (0.4 mL, 0.40 mmol) and acetic acid (24 μL, 0.40 mmol) in DCM (2 mL) was added to the fully assembled resin-bound (glyco)peptides **12**-**14** (25 μmol) and gently agitated for 3 h. The resin was then drained and washed sequentially with dry DCM (10 × 3 mL), DMF (5 × 3 mL) and DCM (10 × 3 mL).

1.3.2 Solid-phase sulfation (25 μmol scale)

Tyrosine deprotected resin bound (glyco)peptides **12**-**14** were swelled in anhydrous DCM (3 mL) for 15 min. The resin was then drained and treated with a solution of triethylamine (28 μL, 0.20 mmol) and imidazolium-sulfating reagent **17** (79 mg, 0.20 mmol) in dry DCM (2 mL) and gently agitated for 16 h. This step was repeated to ensure complete sulfation. The resin was then drained and washed sequentially with DCM (5 × 3 mL), DMF (5 × 3 mL) and DCM (10 × 3 mL).

1.4 General procedure for acidolytic deprotection and cleavage of peptides and protected sulfopeptides from the resin

A mixture of TFA/tri*iso*propylsilane/water (3 mL, 90:5:5 v/v/v) was added to the resin and gently agitated for 2 h. The resin was then drained and washed with TFA (3 × 3 mL) and the combined cleavage/washing solutions were concentrated using a gentle nitrogen stream. Ice cold diethylether was added to the resulting residue and the peptide recovered by centrifugation followed by decanting off the diethylether. The resulting pellet was dissolved in water (2 mL) and the desired (glyco)peptides purified by preparative reverse-phase HPLC (Method C). Concentration of the appropriate fractions provided the desired (glyco)peptides **1**, **15** and **16** and protected sulfated (glyco)peptides **18**-**20**.

1.5 General procedure for deprotection of trichloroethyl(TCE)-protected sulfopeptides (25 μmol scale)

Deprotection of sulfated (glyco)peptides **18**-**20** was conducted using the method described by Ali *et al*.[1] Briefly, to a solution of the alkyl-protected sulfopeptide in 1:1 v/v water/methanol (approximately 4 mL per 30 mg peptide) was added triethylamine (520 μL, 3.75 mmol) and $Pd(OH)_2$ catalyst (30 wt%, 5 mg) and the mixture stirred under an atmosphere of hydrogen for 18 h. The reaction mixture was filtered through a pad of C_{18} silica which was subsequently washed with methanol (3 × 20 mL). The filtrates were combined and the solvent removed under reduced pressure. The resulting residue was purified by preparative reverse-phase HPLC (Method D). Repetitive lyophilization of the appropriate HPLC fractions until a constant mass was obtained providing the desired sulfated (glyco)peptides **4**-**6**.

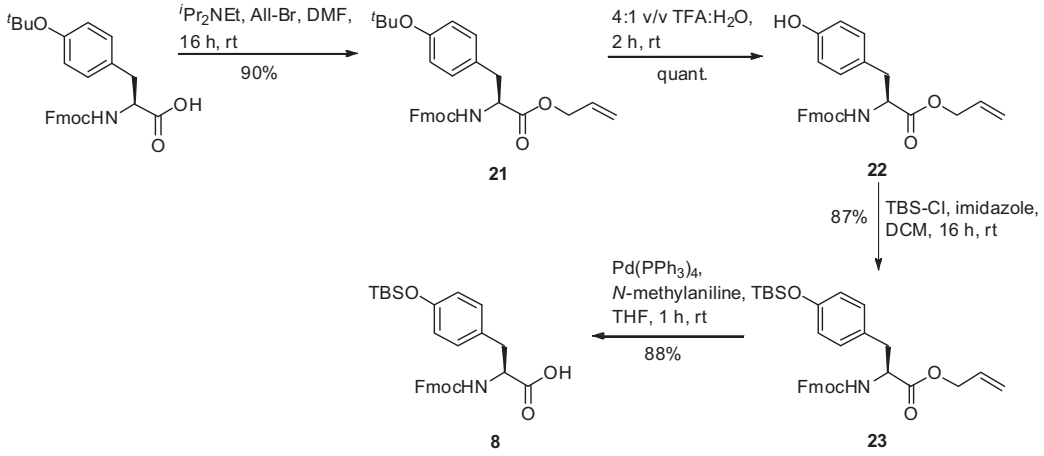
1.6 General procedure for deacetylation of glycopeptides

The removal of acetyl groups on glycopeptides **15, 16, 19** and **20** was performed by dissolving glycopeptides (2-5 mg) in 4:1 v/v NMP:6 M Gn.HCl, 1 M HEPES (4:1 v/v) (200 µL) in presence of hydrazine hydrate (40 µL). The resulting solution was incubated at rt for 16 h before purification by preparative HPLC (***Method C*** for **15** and **16** and ***Method D*** for **19** and **20**).

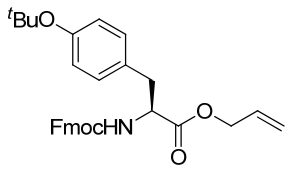
2. Synthesis of Compounds

2.1 Synthesis of Fmoc-Tyr(TBS)-OH (8)

Scheme 1. Synthesis of Fmoc-Tyr(OTBS)-OH (**8**)[1].



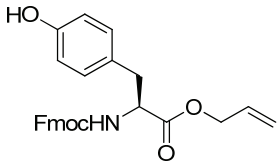
Allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-*tert*-butoxyphenyl)propanoate (21)



To a solution of Fmoc-L-Tyr(O^{t}Bu)-OH (2.50 g, 5.44 mmol) in DMF (5 mL) at 0 °C was added dropwise *N*,*N*-diisopropylethylamine (1.90 mL, 10.9 mmol) followed by allyl bromide (0.92 mL, 11 mmol). The mixture was stirred at rt for 16 h before diluting with EtOAc (50 mL). The solution was subsequently washed with water (5 × 50 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The residue was purified by column chromatography

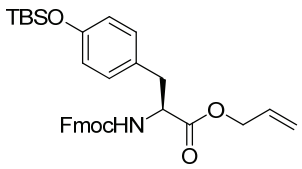
(eluent Hex/EtOAc, 4:1 v/v) to afford **21** (2.45 g, 90%) as a colourless oil. $[\alpha]_D^{25}$ = +3.1° (*c* 2.6, $CHCl_3$); IR (film) 2976, 1726, 1505, 1237, 1161 cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 7.62 (d, 2H, *J* = 7.4 Hz), 7.56 (d, 2H, *J* = 7.4 Hz), 7.40 (t, 2H, *J* = 7.4 Hz), 7.30 (t, 2H, *J* = 7.4 Hz), 7.00 (d, 2H, *J* = 7.7 Hz), 6.89 (d, 2H, *J* = 7.7 Hz), 5.85 (ddt, 1H, *J* =16.5, 10.8, 5.4 Hz), 5.32-5.22 (m, 3H), 4.69-4.31 (m, 5H), 4.20 (t, 1H, *J* = 6.8 Hz), 3.07 (app.s, 2H), 1.31 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.2, 155.5, 154.5, 143.7, 141.3, 131.4, 130.4, 129.8, 127.7, 127.0, 125.0, 124.1, 119.9, 119.0, 78.4, 66.9, 66.0, 54.9, 47.2, 37.7, 28.8; MS (ESI) *m/z* 499.93 $[M+H]^+$.

Allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-hydroxyphenyl) propanoate (22)



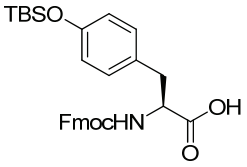
Water (2 mL) was added dropwise to a solution of **21** (0.92 g, 1.84 mmol) in TFA (8 mL). After 2 h the mixture was concentrated under reduced pressure and the residue purified by column chromatography (eluent Hex/EtOAc, 4:1 v/v) to give **22** (0.82 g, quant.) as a yellow solid. m.p. = 96-97 °C; $[\alpha]_D^{25}$ = +2.3° (c 2.5, $CHCl_3$); IR (film) 3431, 1734, 1515, 1448, 1249 cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 7.75 (d, 2H, *J* = 7.4 Hz), 7.54, (d, 2H, *J* = 7.4 Hz), 7.38 (t, 2H, *J* = 7.4 Hz), 7.29 (t, 2H, *J* = 7.4 Hz), 6.94 (d, 2H, *J* = 8.1 Hz), 6.71 (d, 2H, *J* = 8.1 Hz), 5.93-5.80 (m, 1H), 5.33-5.23 (m, 3H), 4.60-4.30 (m, 5H), 4.18 (t, 1H, *J* = 6.9 Hz), 3.10-2.97 (m, 2H), O-H signal not observed; ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.8, 156.1, 155.4, 144.1, 141.7, 131.8, 130.9, 128.1, 127.8, 127.5, 125.5, 120.4, 119.6, 115.9, 67.5, 66.6, 55.4, 47.6, 37.9; MS (ESI) MH^+, 443.87.

Allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-(*tert*-butyldimethylsilyloxy) phenyl)propanoate (23)



Imidazole (0.24 g, 3.5 mmol) was added to a solution of Fmoc-Tyr(OH)-OAll **22** (0.50 g, 1.1 mmol) in DCM (6 mL) and cooled to 0 °C before the addition of TBS-Cl (0.42 g, 2.8 mmol). The mixture was stirred for 16 h at rt before the addition of water (120 mL) and extraction with DCM (2 × 100 mL). The combined organic fractions were dried (Na_2SO_4), filtered and concentrated under reduced pressure. The resulting residue was purified by column chromatography (eluent Hex/EtOAc, 9:1 v/v) to afford **23** (0.55 g, 87%) as a colourless oil. $[\alpha]_D^{25}$ = +2.1° (c 1.8, $CHCl_3$); IR (film) 1719, 1509, 1255 cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 7.77 (d, 2H, *J* = 7.5 Hz), 7.57 (d, 2H, *J* = 4.5 Hz), 7.40 (t, 2H, *J* = 7.5 Hz), 7.31 (t, 2H, *J* = 7.5 Hz), 6.97 (d, 2H, *J* = 8.0 Hz), 6.75 (d, 2H, *J* = 8.6 Hz), 5.94-5.83 (m, 1H), 5.34-5.24 (m, 3H), 4.69-4.60 (m, 3H), 4.46-4.32 (m, 2H) 4.21 (t, 1H, *J* = 6.9 Hz) 3.07 (s, 2H), 0.98 (s, 9H), 0.18 (s, 6H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.4, 155.7, 154.9, 143.9, 141.5, 131.6, 130.5, 128.4, 127.8, 127.2, 125.2, 120.3, 120.1, 119.2, 67.1, 66.2, 55.1, 47.3, 37.7, 25.8, 18.3, -4.3; MS (ESI) 2*M*Na^+, 1136.91; HRMS (ESI) calcd. for $C_{33}H_{39}NO_5SiNa$ $[M+Na]^+$ 580.2495, found 580.2490.

2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-3-(4-(*tert*-butyldimethylsilyloxy) phenyl)propanoic acid (8)



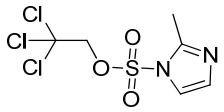
To a solution of **23** (0.47 g, 0.84 mmol) in THF (20 mL) was added *N*-methylaniline (0.92 mL, 8.5 mmol) and $Pd(PPh_3)_4$ (0.01 g, 0.01 mmol) and the resulting mixture was stirred at rt for 1 h. The reaction mixture was subsequently concentrated under reduced pressure and the resulting residue purified by column chromatography (eluent DCM/AcOH, 100:2 v/v) to afford **8** (0.39 g, 88%) as a white solid. m.p. = 92-93 °C; $[\alpha]_D^{25}$ = +5.5° (c 1.7, $CHCl_3$); IR

(film) 2969, 1724, 1510, 1256 cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 7.75 (d, 2H, *J* = 7.3 Hz), 7.54 (d, 2H, *J* = 4.8 Hz), 7.41-7.26 (m, 4H), 6.99 (d, 2H, *J* = 8.0 Hz), 6.76 (d, 2H, *J* = 8.0 Hz), 5.17 (d, 1H, *J* = 8.0 Hz), 4.65 (q, 1H, *J* = 5.9 Hz), 4.46-4.33 (m, 2H), 4.20 (t, 1H, *J* = 6.6 Hz), 3.17-3.00 (m, 2H), 0.96 (s, 9H), 0.17 (s, 6H), O-H signal not observed; ^{13}C NMR (300 MHz, $CDCl_3$) δ 176.1, 155.7, 154.9, 143.6, 141.3, 130.3, 128.0, 127.7, 127.1, 125.0, 120.2, 120.0, 67.1, 54.6, 47.1, 37.0, 25.7, 18.2, -4.4; MS (ESI) *m/z* 517.9 $[M+H]^+$.

2.2 Synthesis of trichlorovinyl (TCE) sulfating reagent (17)

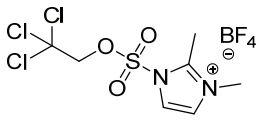
Imidazolium sulfating reagent **17** was prepared according to the method of Ali and Taylor.[2]

2,2,2-Trichloroethyl 2-methyl-1H-imidazole-1-sulfonate (24)



To a cooled suspension of 2-methylimidazole (4.22 g. 51.4 mmol) in THF (30 mL) at 0 °C was added dropwise a solution of trichloroethyl chlorosulfate[1] (3.54 g, 14.3 mmol) in THF (30 mL). The mixture was stirred at 0 °C for 1 h then at rt for 2 h before diluting with Et_2O (50 mL). The reaction mixture was filtered through a pad of Celite® and washed with water (3 x 50 mL), dried (Na_2SO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (eluent Hex/EtOAc, 3:2 v/v) to afford **24** (3.73g, 75%) as a white solid. m.p. = 49-52 °C; IR (film) 1423, 1204 cm^{-1};^1H NMR (300 MHz, $CDCl_3$) δ 7.33 (d, 1H, *J* =1.5 Hz) 6.95 (d, 1H, *J* = 1.5 Hz), 4.67 (s, 2H) 2.68 (s, 3H); ^{13}C NMR (75Hz, $CDCl_3$) δ 146.4, 128.1, 120.0, 91.7, 79.9, 14.8; MS (ESI) *m/z* 294.9 $[M+H]^+$. These data are in agreement with those previously reported by Ali and Taylor.[2]

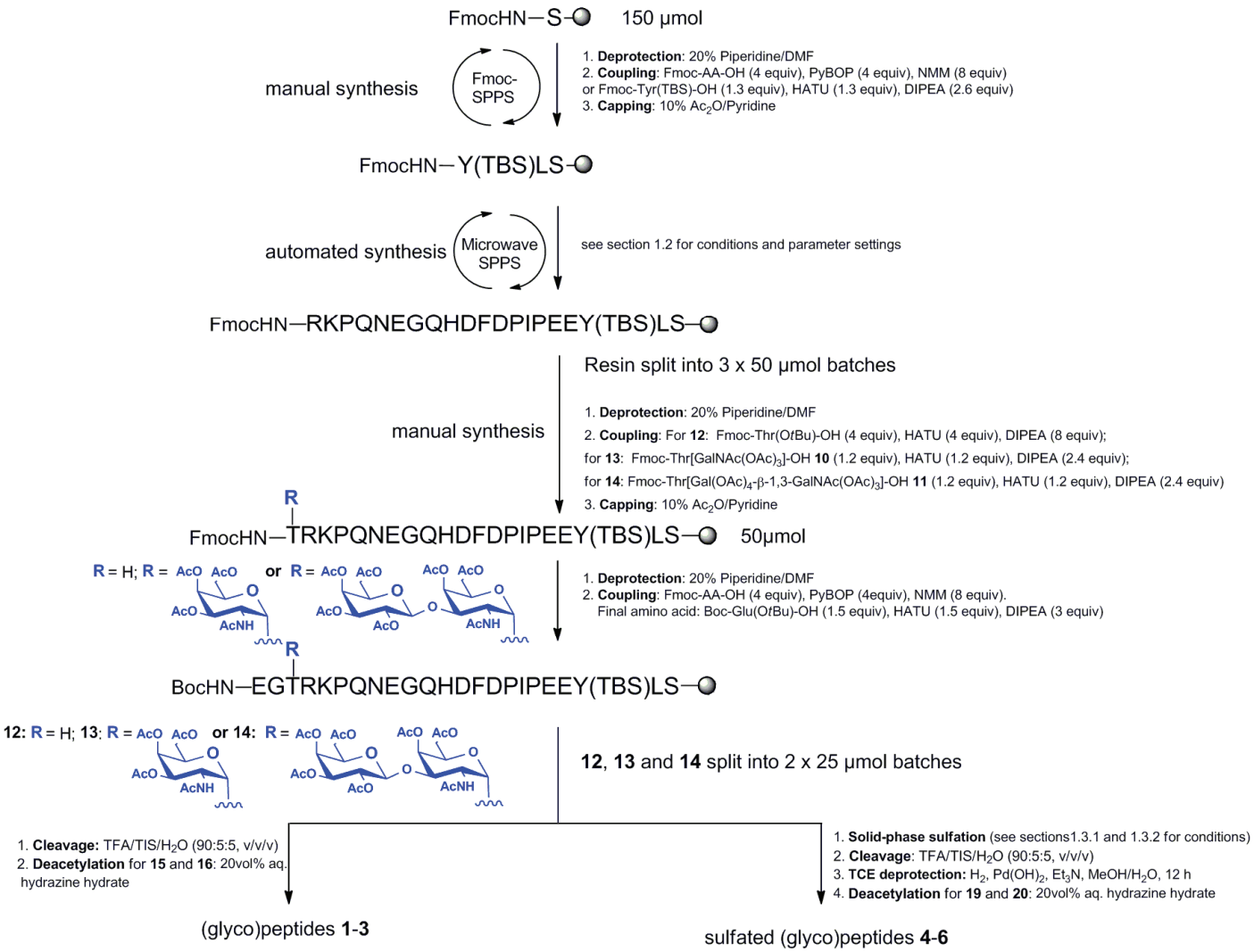
2,3-Dimethyl-1-(2,2,2-trichloroethoxysulfonyl)-1H-imidazol-3-ium tetrafluoroborate (17)



To a cooled suspension of $(Me)_3OBF_4$ (0.53 g, 3.6 mmol) in THF (10 mL) at 0 °C was added dropwise a solution of **24** (1.0 g, 3.4 mmol) in THF (10 mL) and the resulting mixture slowly warmed to rt. After 24 h the mixture was concentrated under reduced pressure before triturating with DCM/Et_2O (1:4 v/v) to obtain the desired imidazolium salt **17** (1.07 g, quant.)

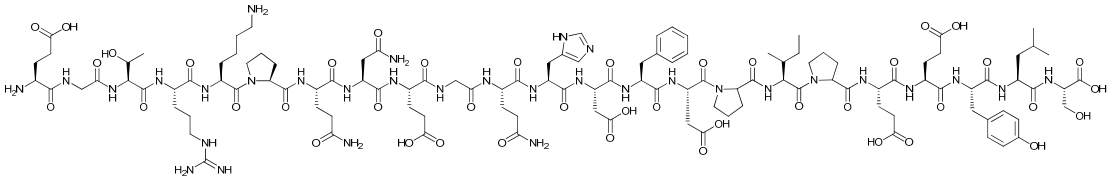
as a white solid. m.p. = 142-144 °C; IR (film) 1432, 1232, 1208, 1058 cm^{-1}; ^{1}H NMR (300 MHz, MeOD) δ 8.07 (d, 1H, *J* = 2.2 Hz), 7.73 (d, 1H, *J* = 2.2 Hz), 5.33 (s, 2H), 3.94 (s, 3H), 2.93 (s, 3H); ^{13}C NMR (75 MHz, MeOD) δ 148.6, 123.5, 120.7, 91.7, 82.0, 35.3, 10.4; MS (ESI) *m/z* 308.9 $[M\text{-}BF_4^-]^+$. These data are in agreement with those previously reported by Ali and Taylor.[2]

3. General Solid-Phase Protocols for the Preparation of HP6 Fragments 1-6



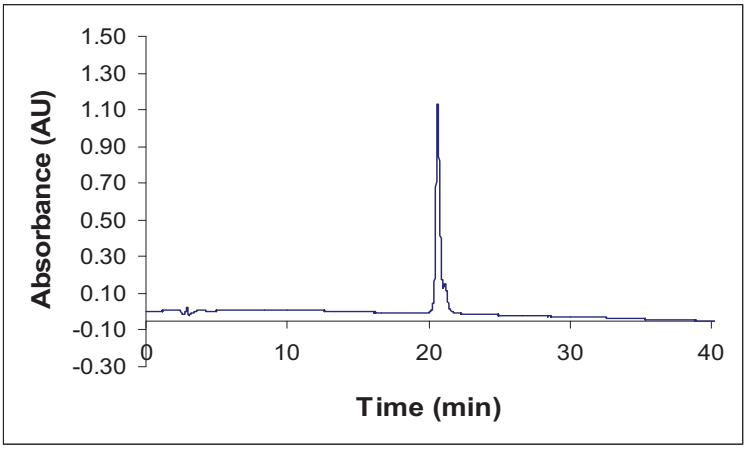
4. Analytical Data for HP6 Fragments 1-6

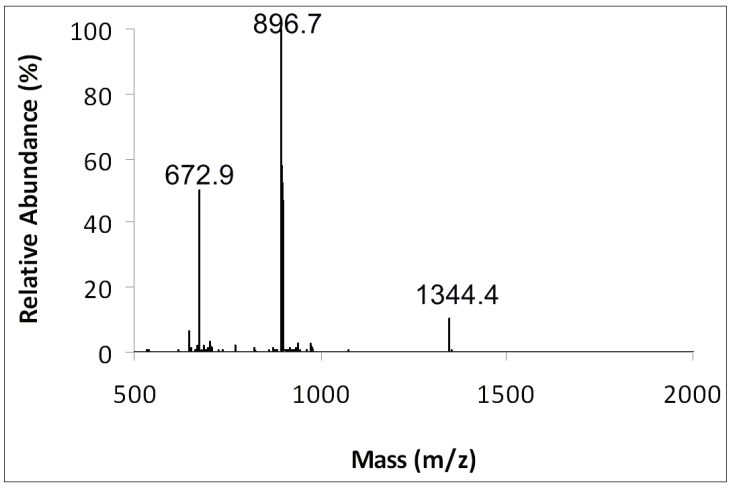
H_2N-EGTRKPQNEGQHDFDPIPEEYLS**-OH (1)**



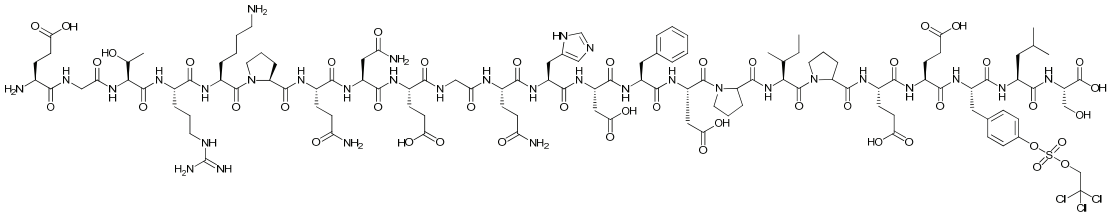
Peptide **1** was synthesised according to the general Fmoc-SPPS procedures described above in sections 1.1 and 1.2. The peptide was produced as a white solid after lyophilisation (14.9 mg, 22% isolated yield based on the original 25 μmol resin loading).

ESIMS: calcd M = 2685.23; found, 1344.4 $[M + 2H]^{2+}$, 896.7$[M + 3H]^{3+}$, 672.9 $[M + 4H]^{4+}$, Analytical HPLC: t_R: 20.6 min (Gradient 0 to 50% D over 40 min; Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile).



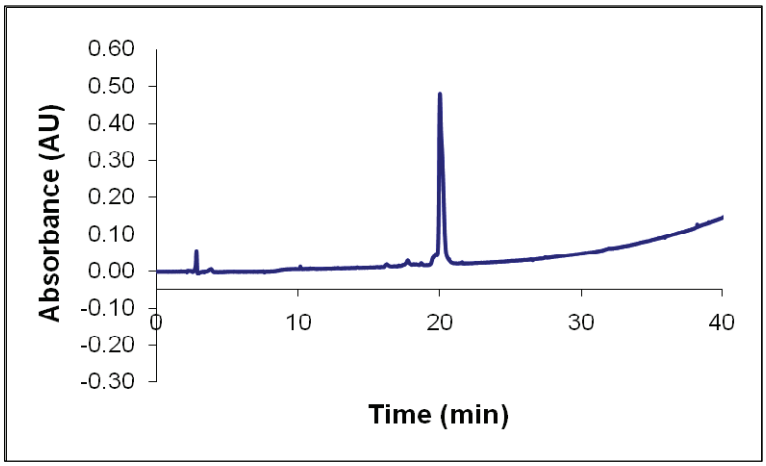


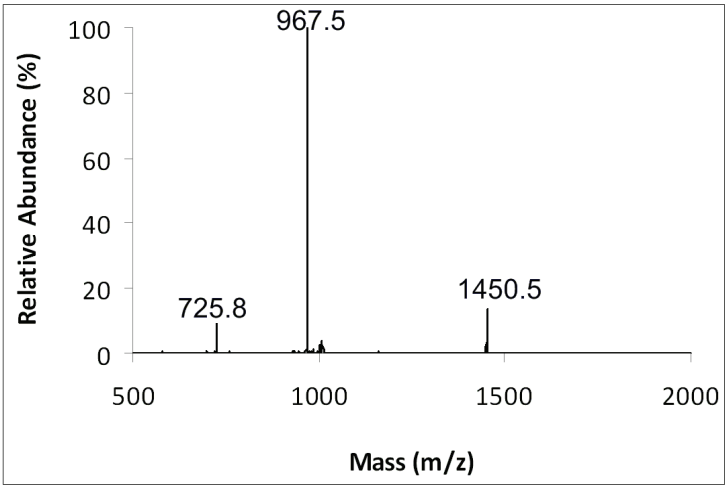
H_2N-EGTRKPQNEGQHDFDPIPEEY(SO_3TCE)LS-OH (18)



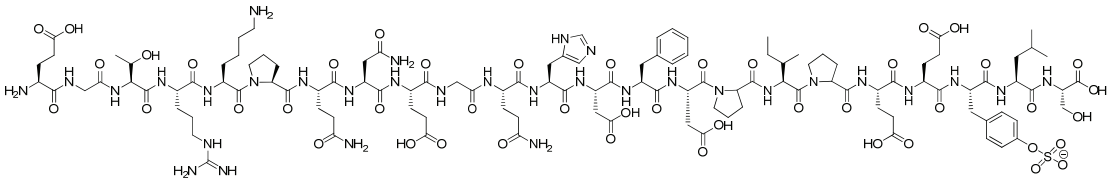
Resin bound peptide **12** (25 μmol) was TBS deprotected, sulfated and deprotected according to the general procedures outlined in sections 1.3 and 1.4 to afford protected sulfopeptide **18** as a white solid (11.9 mg, 17% isolated yield based on the 25 μmol resin loading).

ESI-MS: calcd $M = 2895.10$; found, 1450.5 $[M + 2H]^{2+}$, 967.5 $[M + 3H]^{3+}$, 725.8 $[M + 4H]^{4+}$, Analytical HPLC: t_R: 20.0 min (Gradient 0 to 100% B over 40 min).



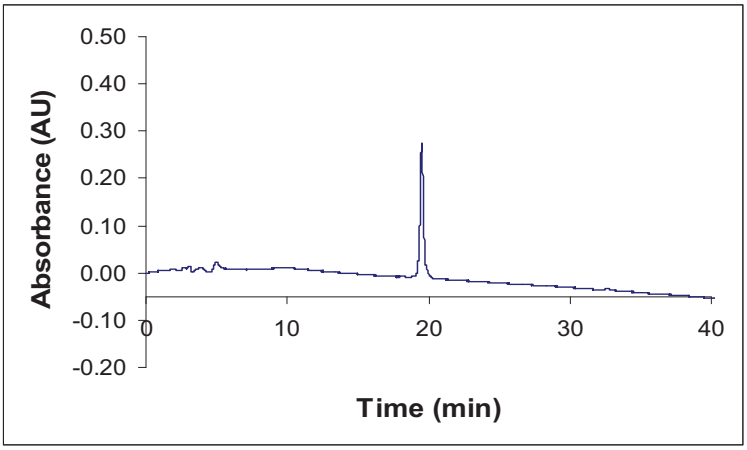


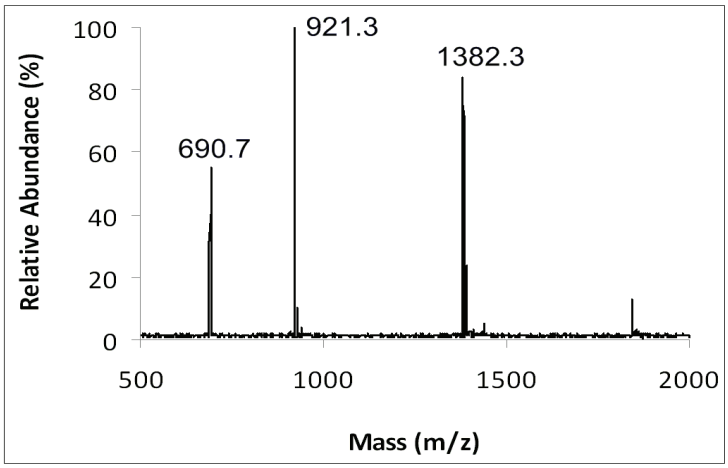
H_2N-EGTRKPQNEGQHDFDPIPEEY($SO_3^-NH_4^+$)LS-OH (4)



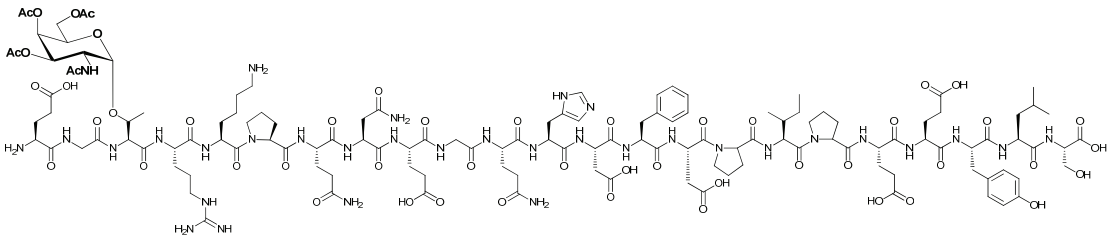
Protected sulfopeptide **18** (5.5 mg, 1.9 μmol) was deprotected and purified according to the general procedure outlined in section 1.5 to afford sulfopeptide **4** as a white solid (3.8 mg, 74% isolated yield).

ESI-MS: calcd M = 2766.20; found, 1382.3 $[M - 2H]^{2-}$, 921.3 $[M - 3H]^{3-}$, 690.7 $[M - 4H]^{4-}$, Analytical HPLC: t_R: 19.5min (Gradient 0 to 50% D over 40 min; Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile)



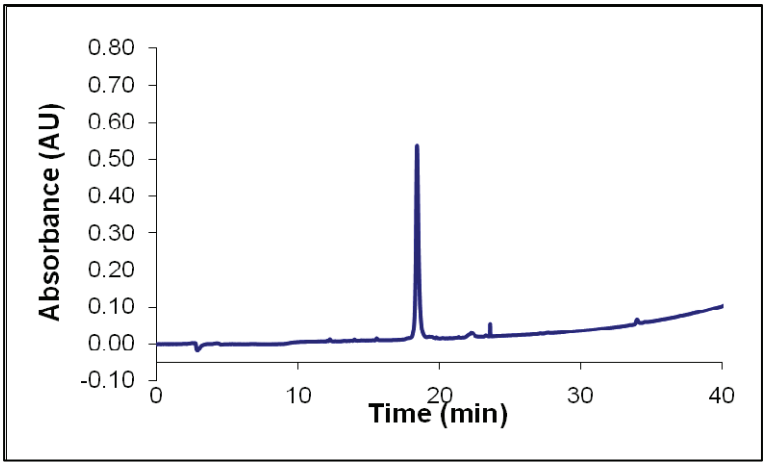


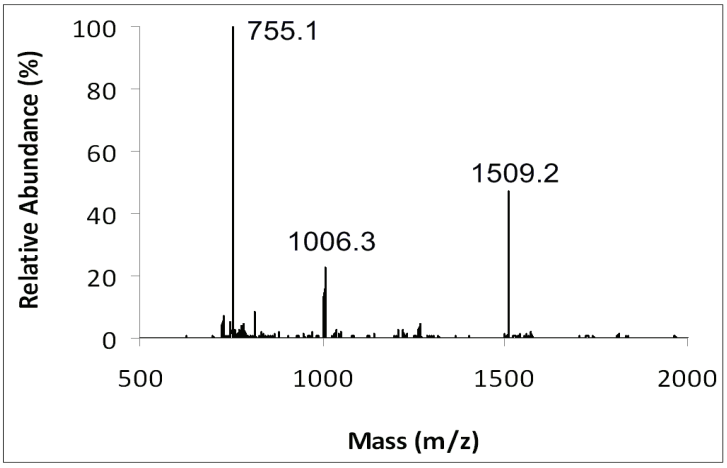
H_2N-EGT[α-GalNAc$(OAc)_3$]RKPQNEGQHDFDPIPEEYLS-OH (15)



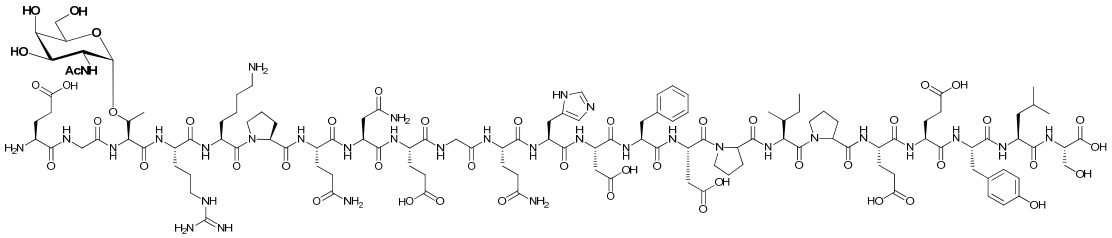
Protected glycopeptide **15** was synthesised according to the general Fmoc-SPPS procedures described above in sections 1.1 and 1.2. The glycopeptide was produced as a white solid after lyophilisation (14.5 mg, 19% isolated yield based on the original 25 μmol resin loading).

ESI-MS: calcd $M = 3014.34$; found, 1509.2 $[M + 2H]^{2+}$, 1006.3 $[M + 3H]^{3+}$, 755.1 $[M + 4H]^{4+}$, Analytical HPLC: t_R: 18.4 min (Gradient 0 to 100% B over 40 min).



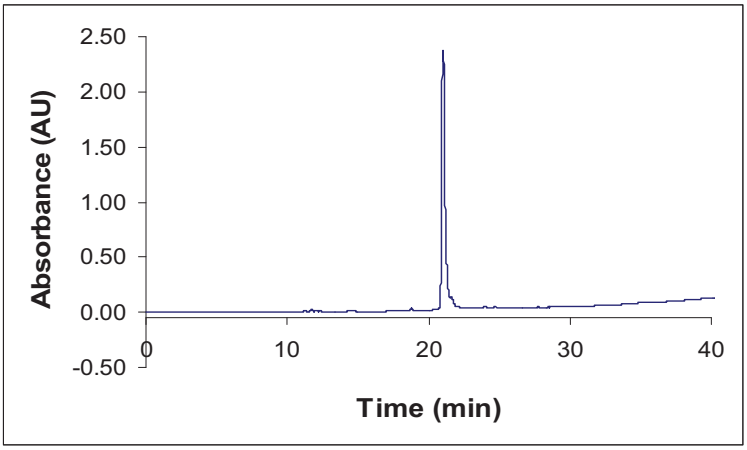


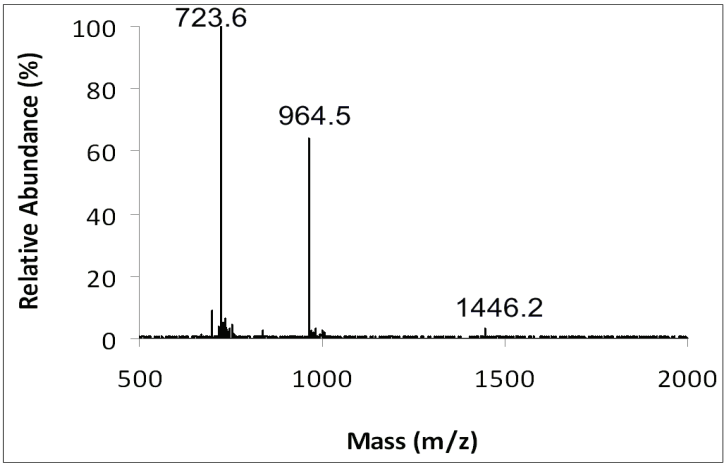
H_2N-EGT(α-GalNAc)RKPQNEGQHDFDPIPEEYLS-**OH (2)**



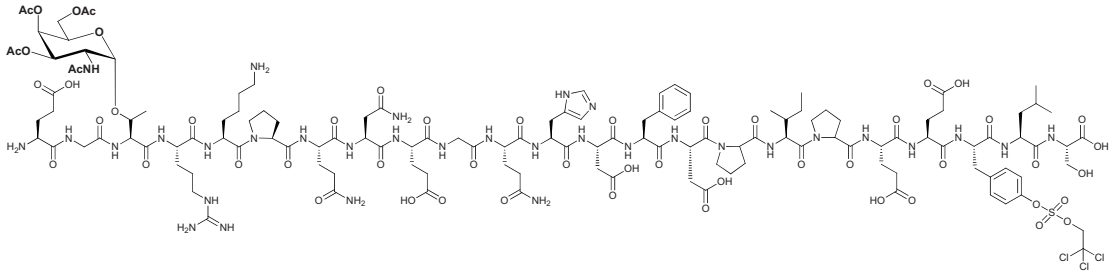
Glycopeptide **15** (5.5 mg, 1.8 μmol) was deacetylated according to the general procedure described in section 1.6. The glycopeptide was produced as a white solid after lyophilisation (4.6 mg, 89% isolated yield).

ESI-MS: calcd $M = 2890.33$; found, 1446.2 $[M + 2H]^{2+}$, 964.5 $[M + 3H]^{3+}$, 723.6 $[M + 4H]^{4+}$, Analytical HPLC: t_R: 20.5min (Gradient 0 to 50% D over 40 min; Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile).



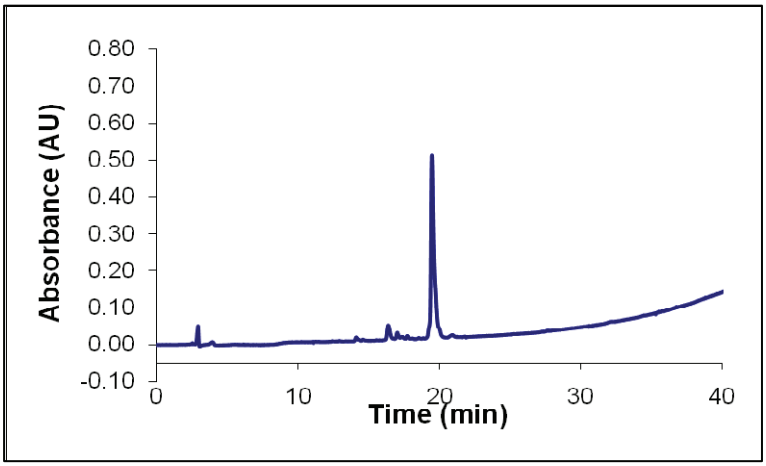


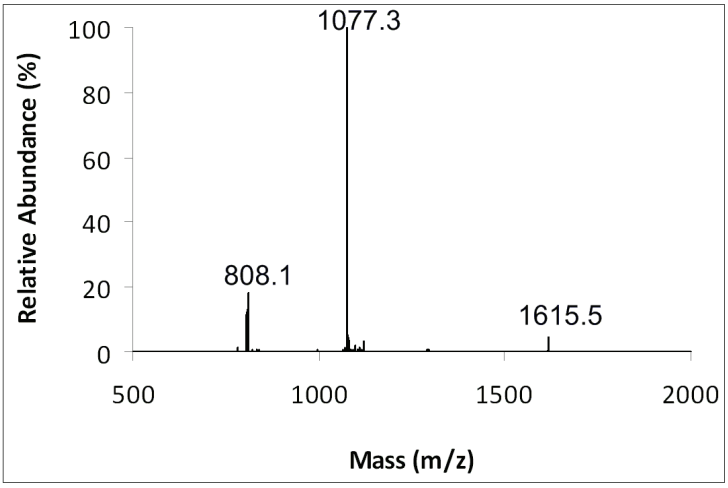
H_2N-EGT[α-GalNAc(OAc)$_3$]RKPQNEGQHDFDPIPEEY(SO_3TCE)LS**-OH (19)**



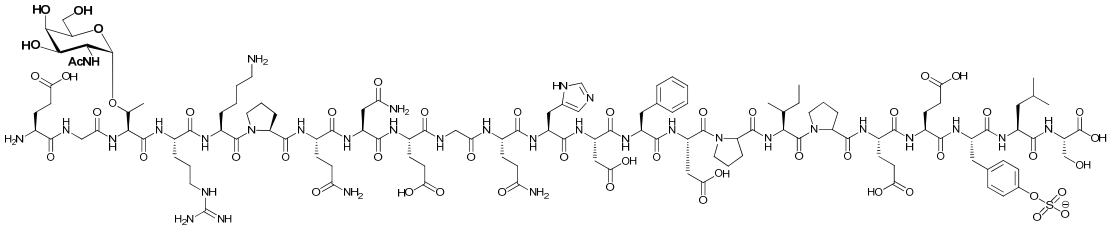
Resin bound glycopeptide **13** (25 μmol) was TBS deprotected, sulfated and deprotected according to the general procedures outlined in sections 1.3 and 1.4 to afford protected sulfopeptide **19** as a white solid (11.6 mg, 15% isolated yield based on the 25 μmol resin loading)

ESI-MS: calcd M = 3224.21; found, 1615.5 $[M + 2H]^{2+}$, 1077.3 $[M + 3H]^{3+}$, 808.1$[M + 4H]^{4+}$, Analytical HPLC: t_R: 19.4 min.



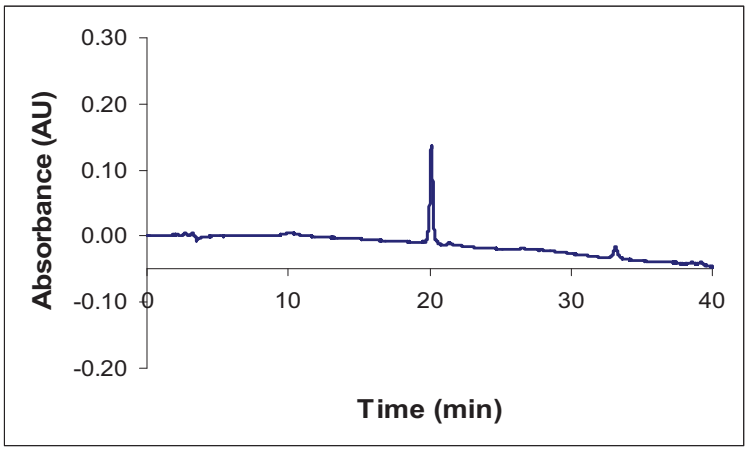


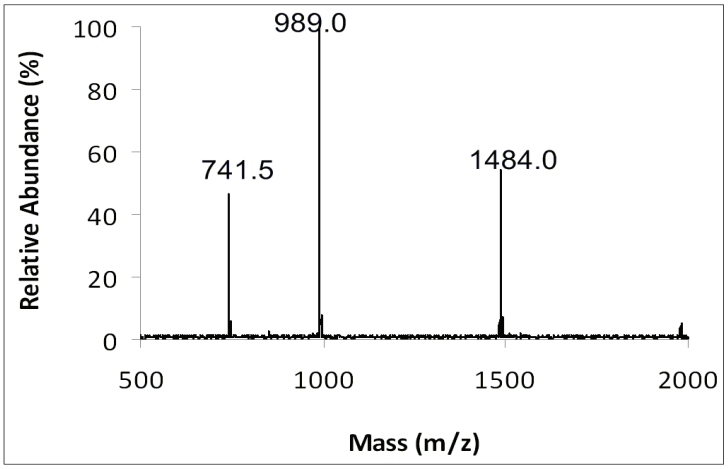
H_2N-EGT(α-GalNAc)RKPQNEGQHDFDPIPEEY($SO_3^-NH_4^+$)LS-**OH (5)**



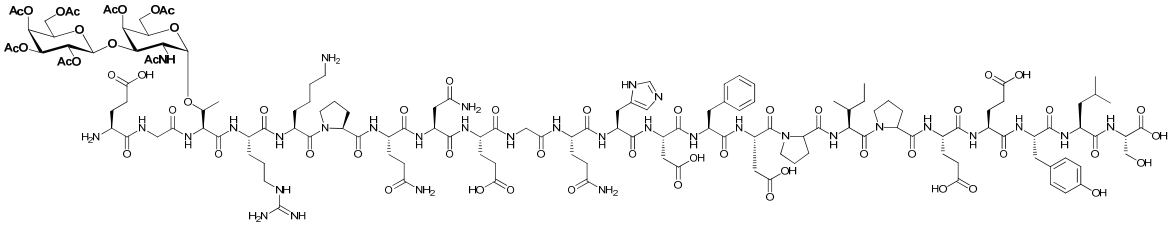
Protected sulfopeptide **19** (3.6 mg, 1.1 μmol) was deprotected and purified according to the general procedures outlined in sections 1.5 and 1.6 to afford sulfopeptide **5** as a white solid (2.3 mg, 73% isolated yield).

ESI-MS: calcd M = 2967.26; found, 1484.0 $[M - 2H]^{2-}$, 989.0 $[M - 3H]^{3-}$, 741.5 $[M - 4H]^{4-}$. Analytical HPLC: t_R: 20.0 min. (Gradient 0 to 50% D over 40 min; Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile).



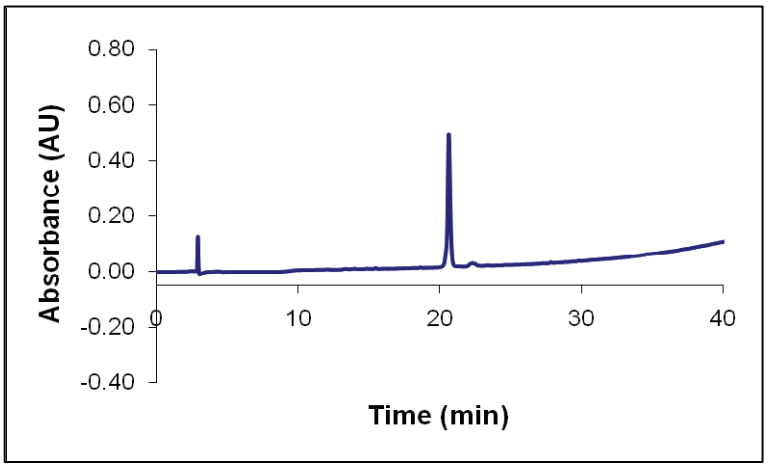


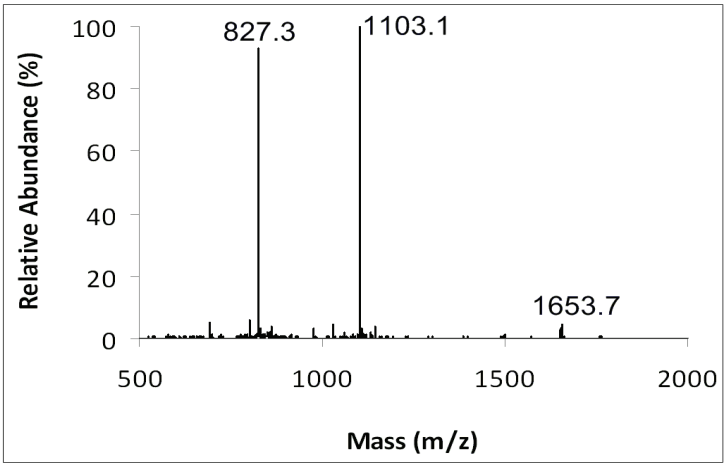
H_2N-EGT[Gal(OAc)$_4$-β-1,3-GalNAc(OAc)$_3$]RKPQNEGQHDFDPIPEEYLS-**OH (16)**



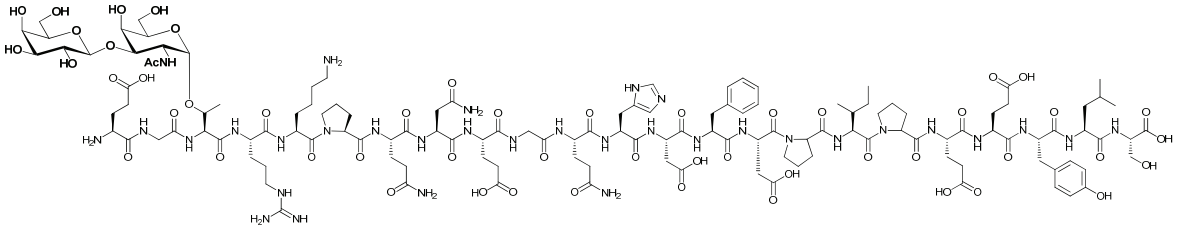
Protected glycopeptide **16** was synthesised according to the general Fmoc-SPPS procedures described above in sections 1.1 and 1.2. The glycopeptide was produced as a white solid after lyophilisation (14.2 mg, 17% isolated yield based on the original 25 μmol resin loading).

ESI-MS: calcd M = 3302.43; found, 1653.1 $[M + 2H]^{2+}$, 1103.1 $[M + 3H]^{3+}$, 827.3 $[M + 4H]^{4+}$. Analytical HPLC: t_R: 20.7 min. (Gradient 0 to 100% B over 40 min).



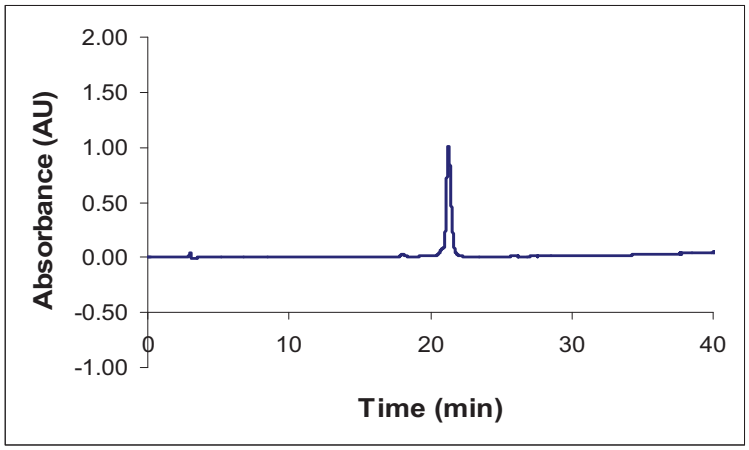


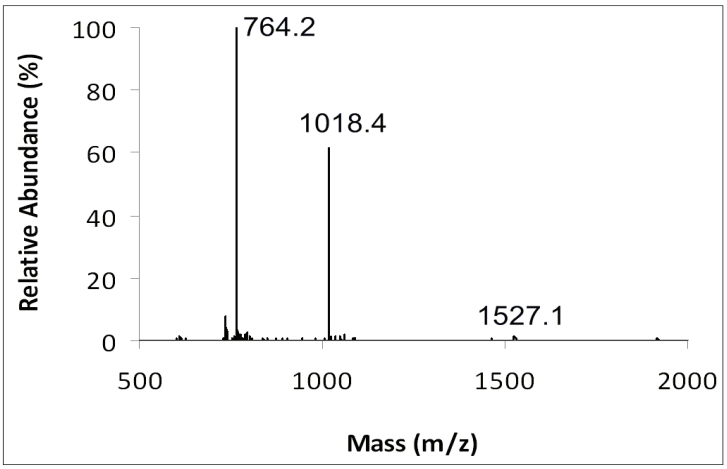
H_2N-EGT(Gal-β-1,3-GalNAc)RKPQNEGQHDFDPIPEEYLS-**OH (3)**



Glycopeptide **16** (3.7 mg, 1.6 μmol) was deacetylated according to the general procedure described in section 1.6. The glycopeptide was produced as a white solid after lyophilisation (3.1 mg, 91% isolated yield).

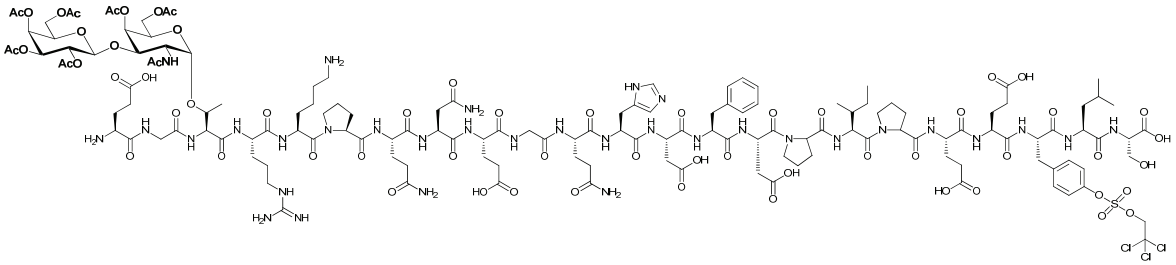
ESI-MS: calcd M = 3050.36; found, 1527.1 $[M + 2H]^{2+}$, 1018.4 $[M + 3H]^{3+}$, 746.2 $[M + 4H]^{4+}$. Analytical HPLC: t_R: 21.0 min. (Gradient 0 to 50% D over 40 min; Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile).





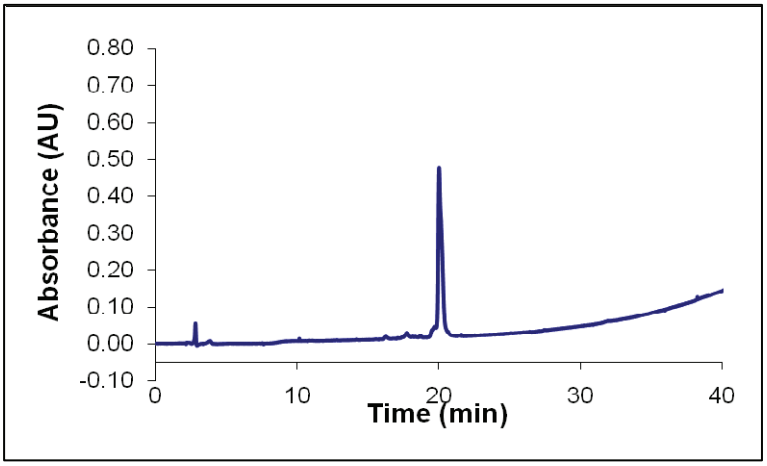
H_2N-EGT[Gal(OAc)$_4$-β-1,3-GalNAc(OAc)$_3$]RKPQNEGQHDFDPIPEEY(SO_3TCE)LS-**OH**

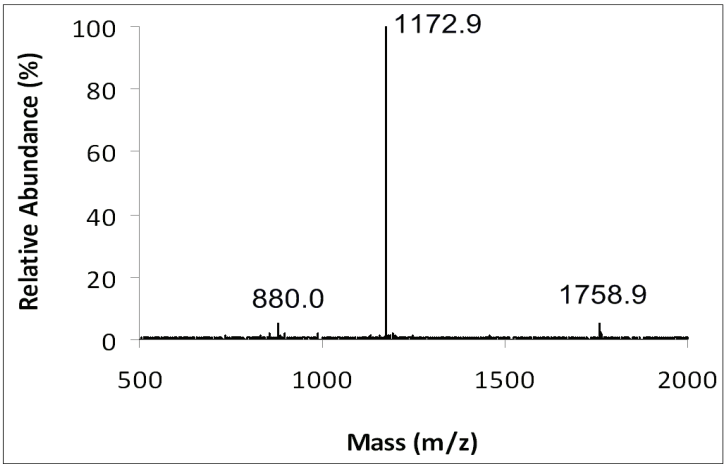
(20)



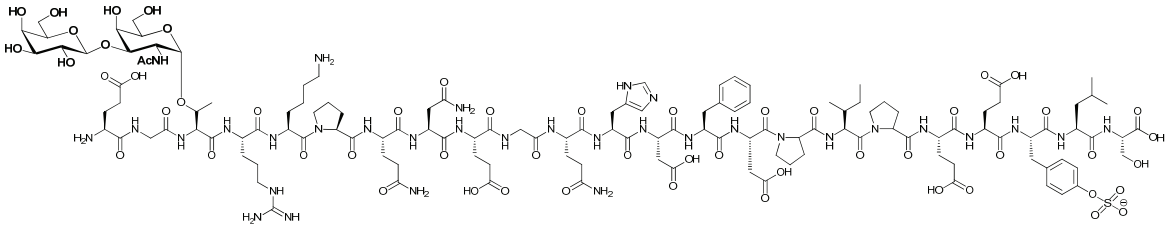
Resin bound peptide **14** (25 μmol) was TBS deprotected, sulfated and deprotected according to the general procedures outlined in sections 1.3 and 1.4 to afford protected sulfopeptide **20** as a white solid (11.5 mg, 14 % isolated yield based on the 25 μmol resin loading).

ESI-MS: calcd M = 3294.31; found, 1758.9 $[M + 2H]^{2+}$, 1172.9 $[M + 3H]^{3+}$, 880.0 $[M + 4H]^{4+}$. Analytical HPLC: t_R: 20.1 min.



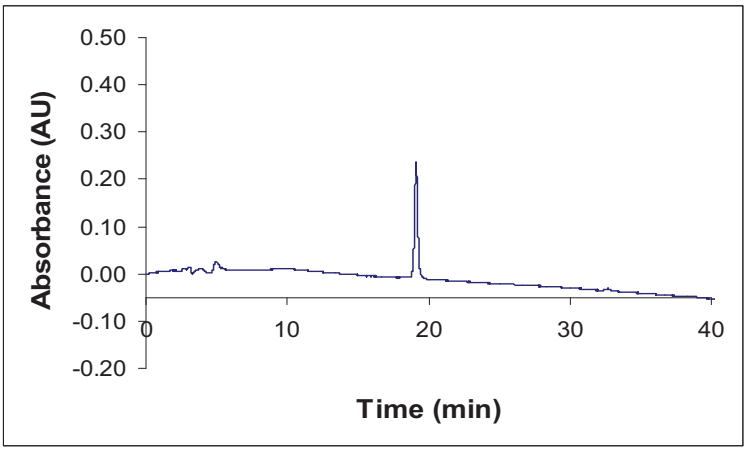


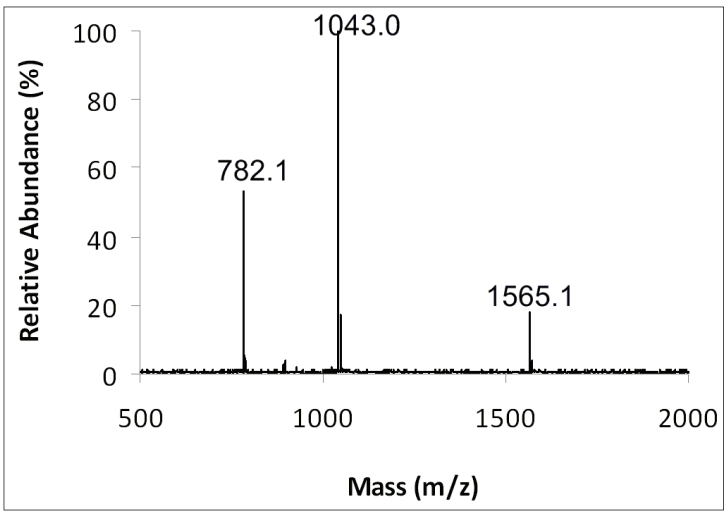
H_2N- EGT(Gal-β-1,3-GalNAc)RKPQNEGQHDFDPIPEEY($SO_3^-NH_4^+$)LS **-OH (6)**



Protected sulfopeptide **20** (3.7 mg, 1.1 μmol) was deprotected and purified according to the general procedures outlined in sections 1.5 and 1.6 to afford sulfopeptide **6** as a white solid (3.1 mg, 91% isolated yield).

ESI-MS: calcd M = 3129.31; found, 1565.1 $[M - 2H]^{2-}$, 1043.0 $[M - 3H]^{3-}$, 782.1 $[M - 4H]^{4-}$. Analytical HPLC: t_R: 19.1 min. (Gradient 0 to 50% D over 40 min; Solvent C: 0.1 M ammonium acetate, solvent D: acetonitrile).





5. Experimental and Raw Data for Thrombin Assays

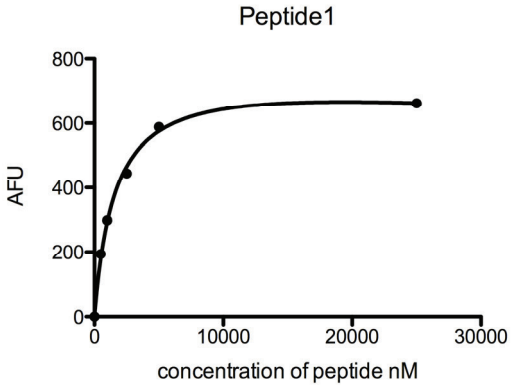
Measurement of affinity of hirudin analogue peptides for thrombin

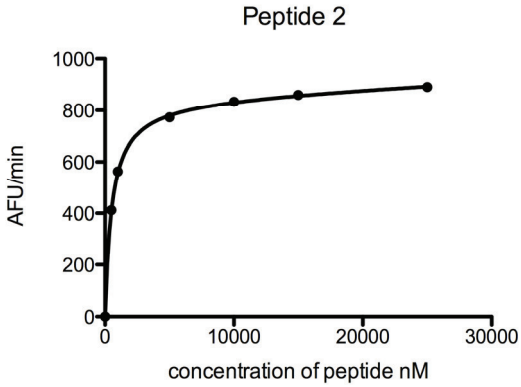
The K_d values for binding of HP6 peptides **1**-**6** to thrombin were determined by analysing their effect on the cleavage of a peptide substrate of the enzyme. Human thrombin at a final concentration of 20 pM was incubated with 0-40 μM of **1**-**6** for 15 min at 37 ºC. The activity of the enzyme was measured by determining the rate of cleavage of the substrate, tosyl-Gly-Pro-Arg-NHMec (NHMec = 7-amido-4-methylcoumarin), using a BMG Fluorostar plate reader with an excitation filter of 370 nm and an emission filter of 460 nm. The initial velocity for thrombin alone was subtracted from the values for thrombin with the addition of the various peptides and the resulting values were plotted against the concentration of the analogue peptides. These data were fitted using non-linear regression on GraphPad Prism to the following equation for one site – total binding: $Y = B_{max} \times X/(K_d+X)$. The B_{max} value represents the maximal stimulation of thrombin activity.

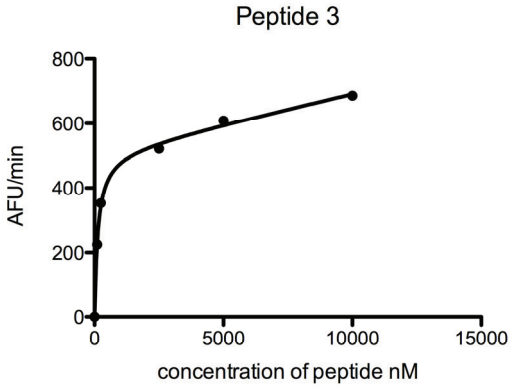
Competition assay to assess inhibitory effects of hirudin analogues

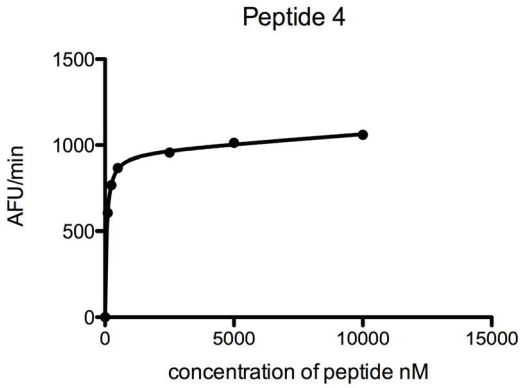
Competition for the secondary fibrinogen-binding site (Exosite I) on thrombin was measured by a turbidimetric clotting assay. Human thrombin (Haematologic Technologies, Essex Junction, VT, USA) [5 pM] was incubated at 37 ºC with HP6 peptides **1**-**6** in 0.05 M Tris, pH 7.8, 0.1 M NaCl, 0.1% (w/v) PEG 6000 for 30 min prior to addition of 5 μM fibrinogen (Sigma, St Louis, MI, USA). Fibrin formation was followed by measuring the optical turbidity at 405 nm for 30 min. The values for absorbance at 405 nm were then plotted against the concentration of peptide and these data were fitted by non-linear regression on GraphPad Prism to the following equation: $Y = \text{Bottom} + (\text{Top-Bottom})/(1+10^{[X-\text{LogIC50}]})$, where “Bottom” and “Top” refer to the highest and lowest values obtained, respectively.

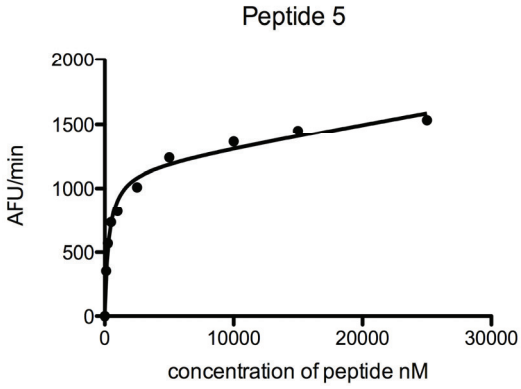
Shown below are the graphs displaying thrombin activity (AFU/min) against tosyl-Gly-Pro-Arg-NHMec in the presence of increasing concentrations of peptides **1**-**6**. The activity for thrombin alone hás been substracted for each data point.

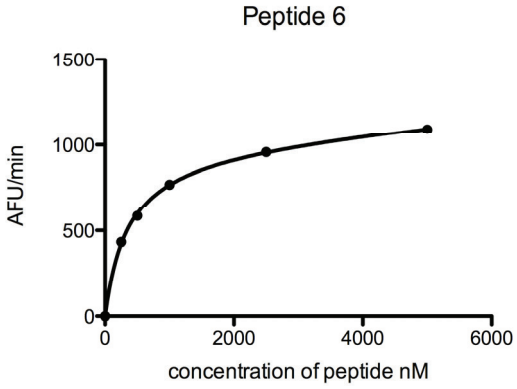




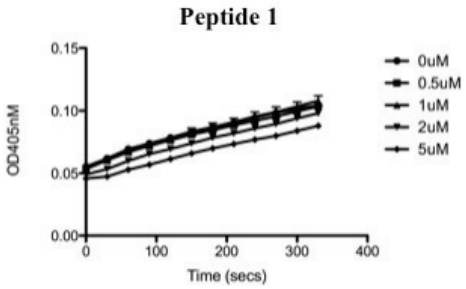


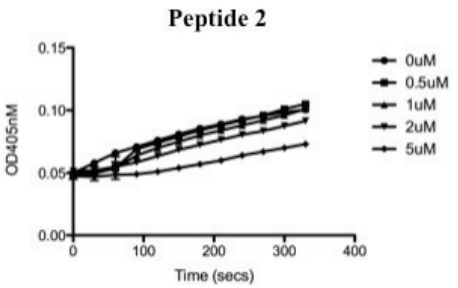


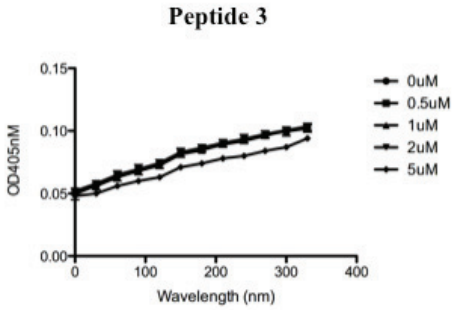


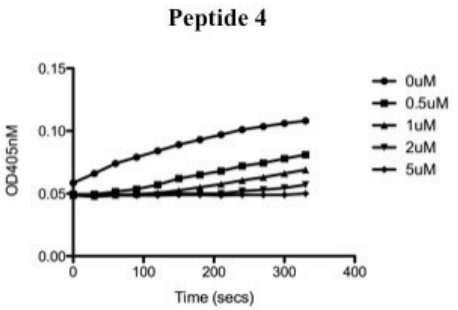


The graphs below indicate the increase in turbidity, as indicated by the OD at 405 nm, of solutions containing fibrinogen in the presence of thrombin or thrombin together with the indicated concentrations of each of the HP6 peptides (**1**-**6**). The values obtained after 300 sec in each case were used to plot activity versus the concentration of the peptide and fitted using non-linear regression to obtain IC_{50} values in each case.

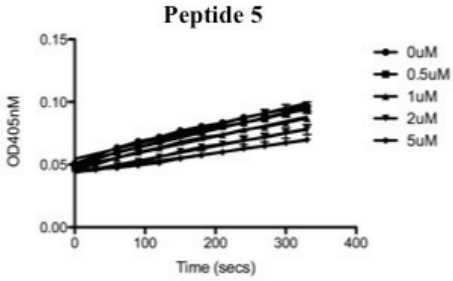




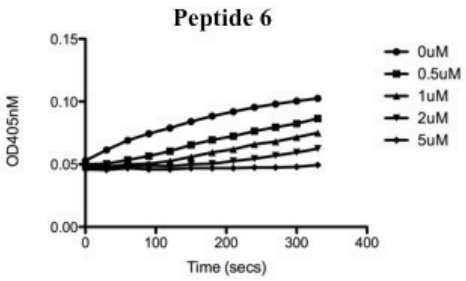




$R^2 = 0.999$

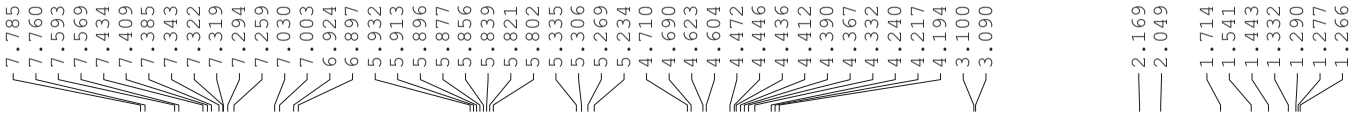


$R^2 = 0.998$



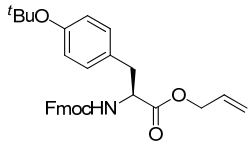
$R^2 = 0.999$

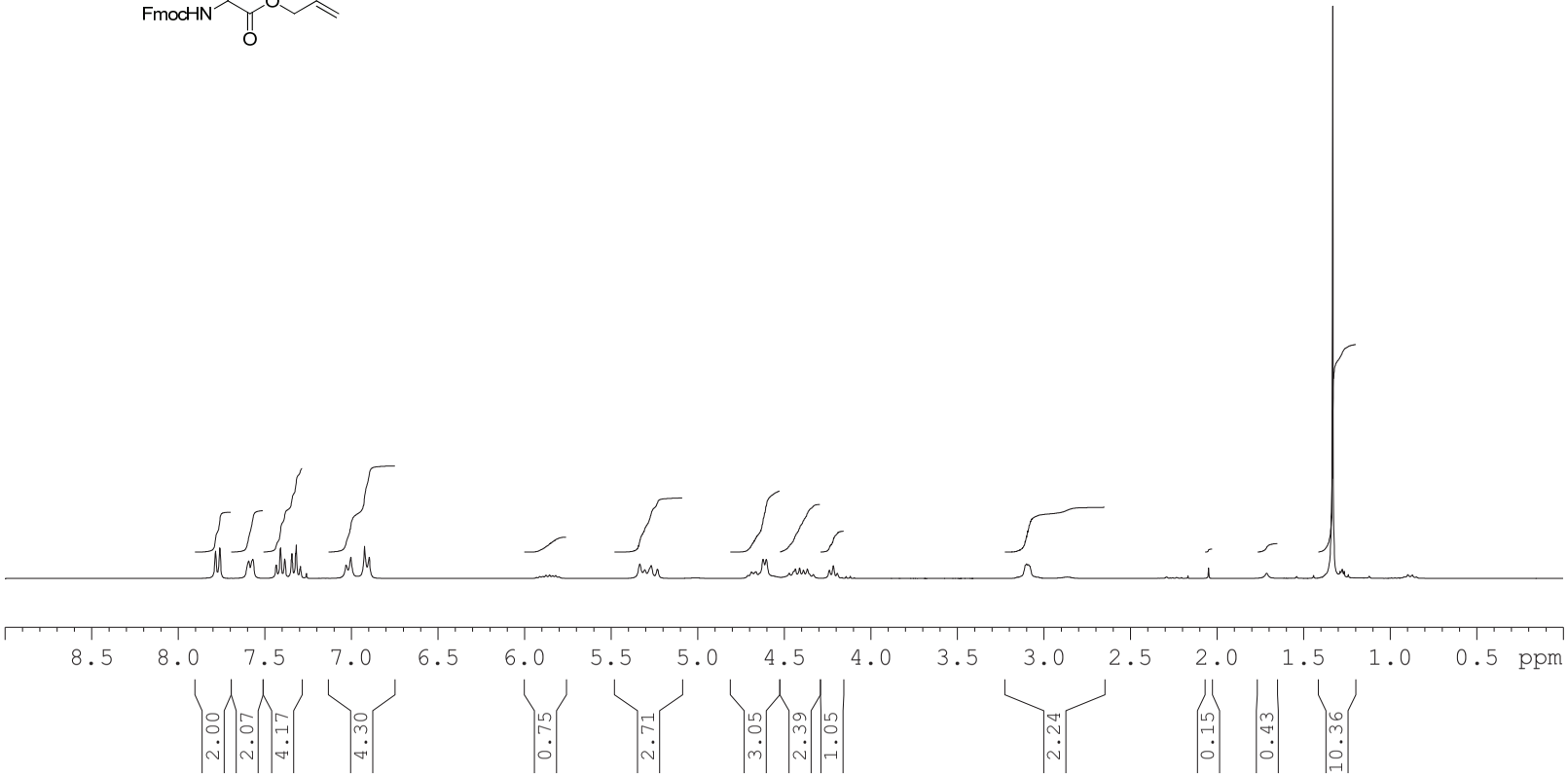
6. ^{1}H and ^{13}C-NMR Spectra of Modified Amino Acids, Reagents, and ^{1}H-NMR Spectra of 1-6

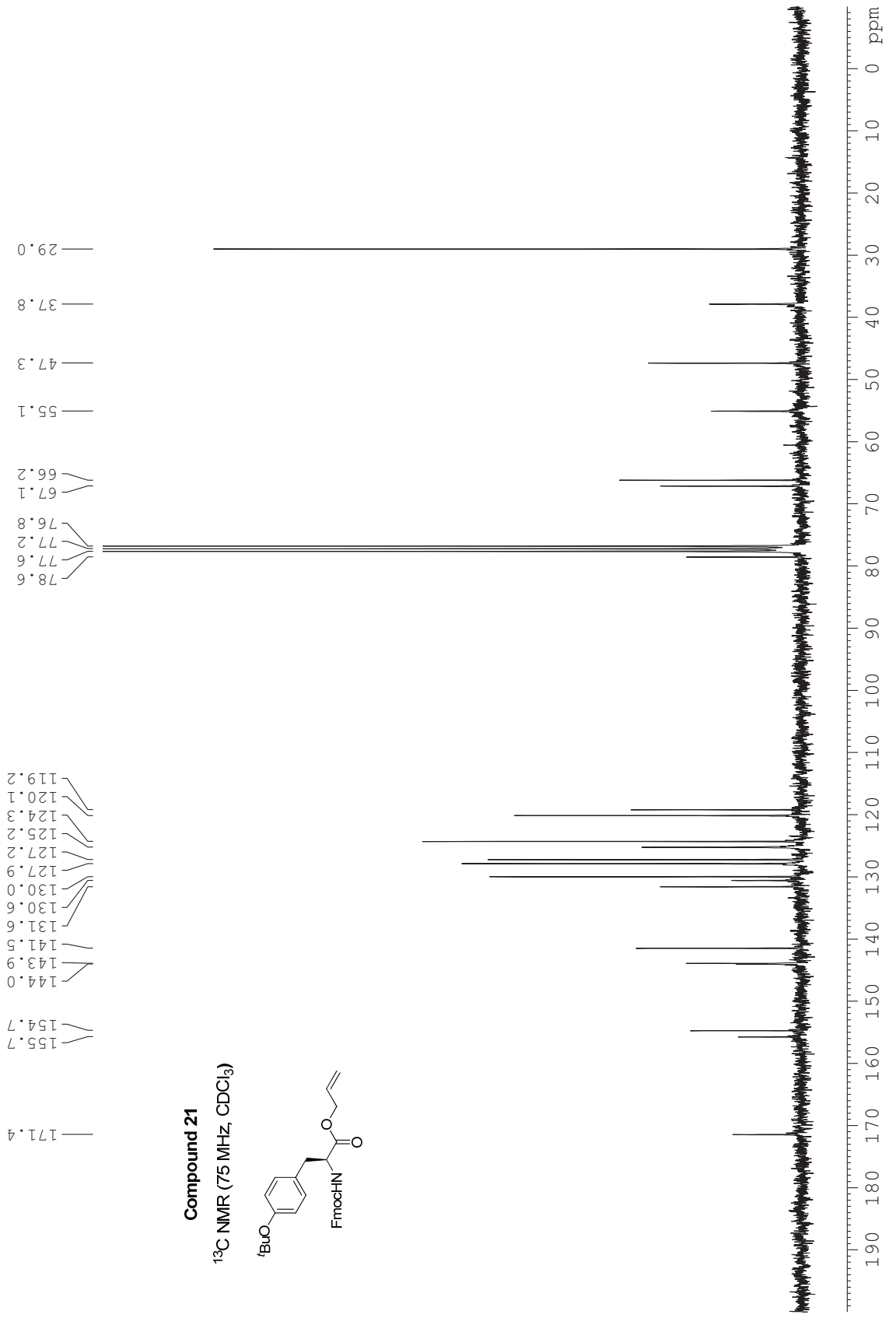


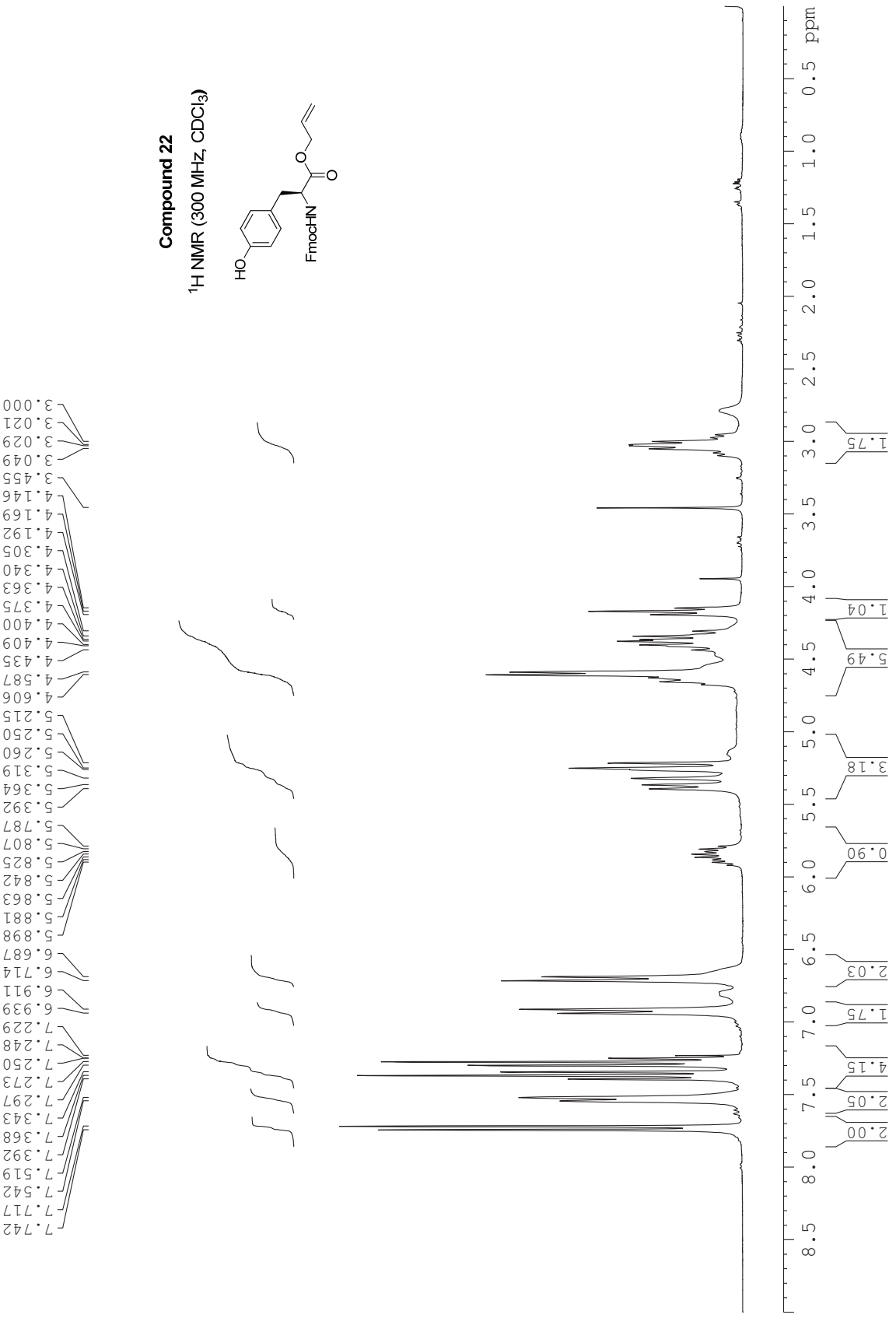
Compound 21

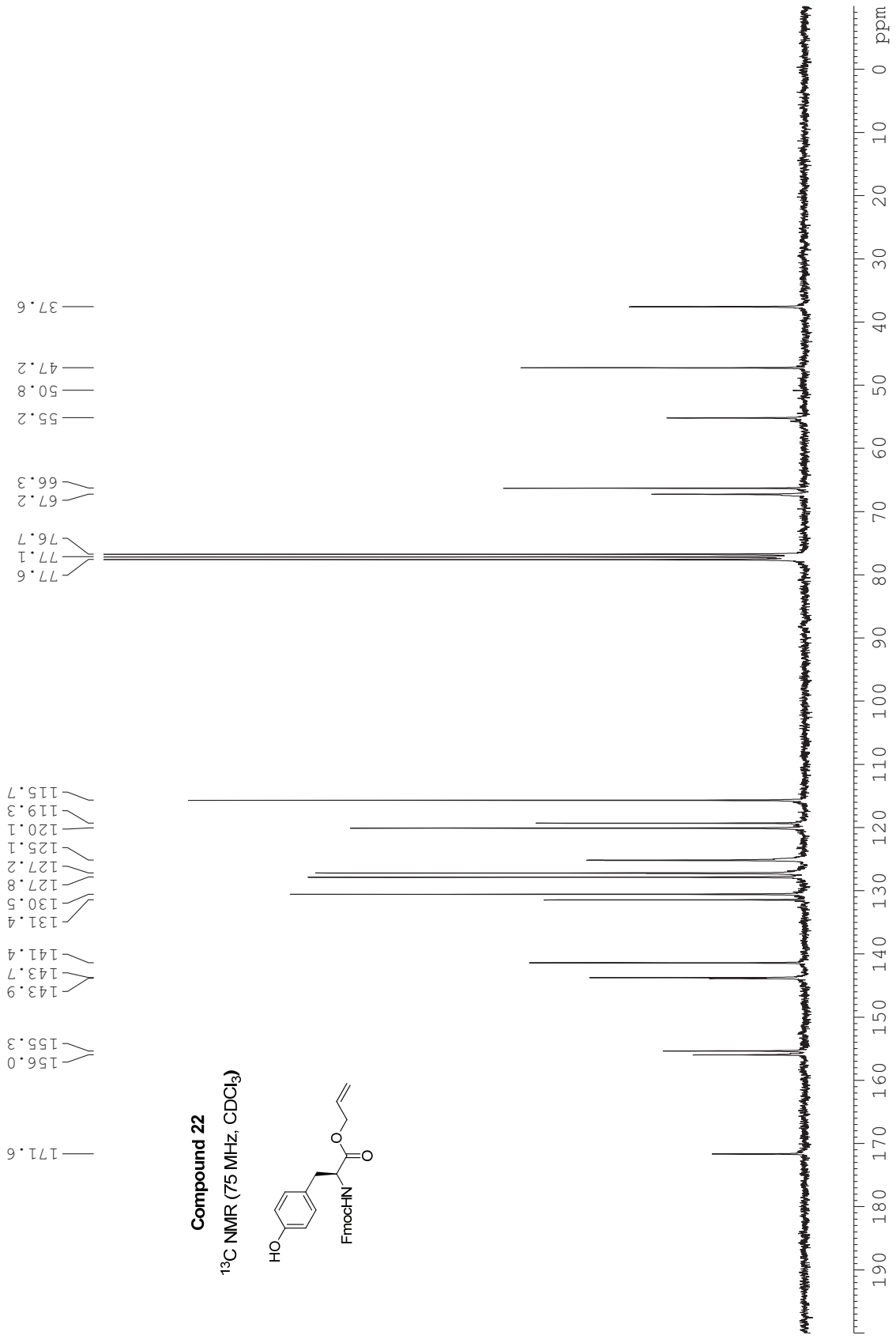
^{1}H NMR (300 MHz, $CDCl_3$)

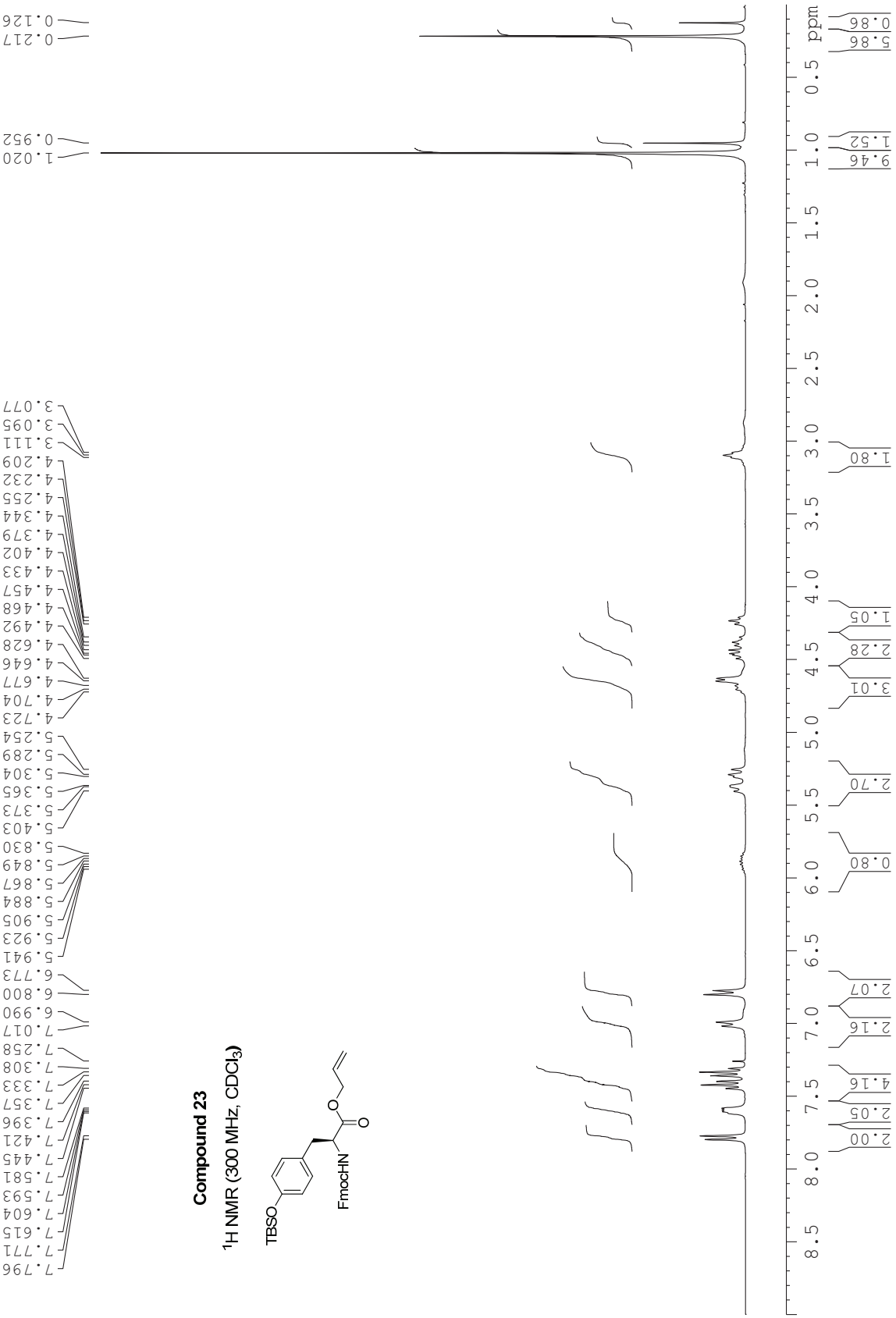


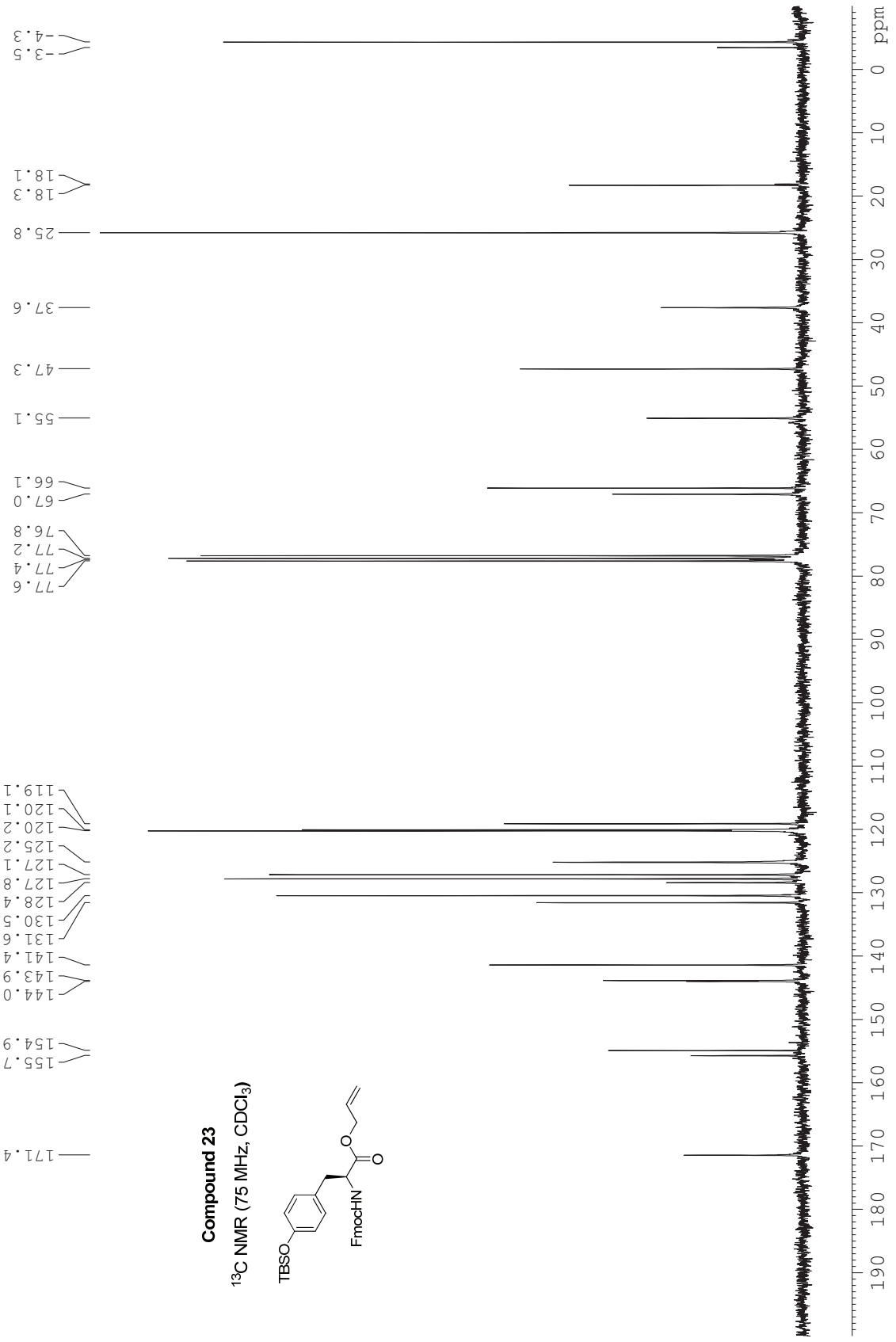


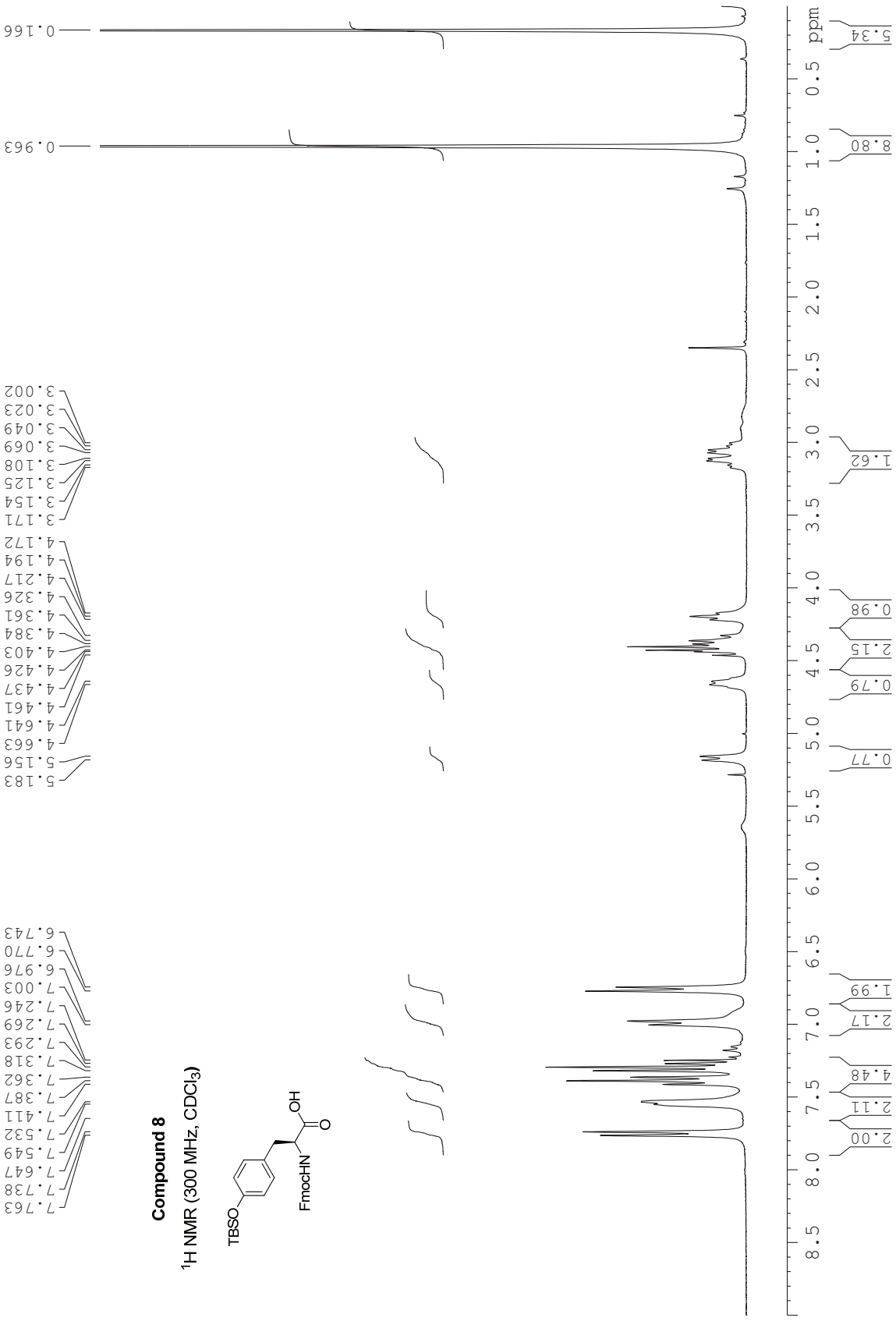

Compound 21
^{13}C NMR (75 MHz, $CDCl_3$)
^{t}BuO
FmocHN
O
O
O
171.4
155.7
154.7
144.0
143.9
141.5
131.6
130.6
130.0
127.9
127.2
125.2
124.3
120.1
119.2
78.6
77.6
77.2
76.8
67.1
66.2
55.1
47.3
37.8
29.0
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 ppm

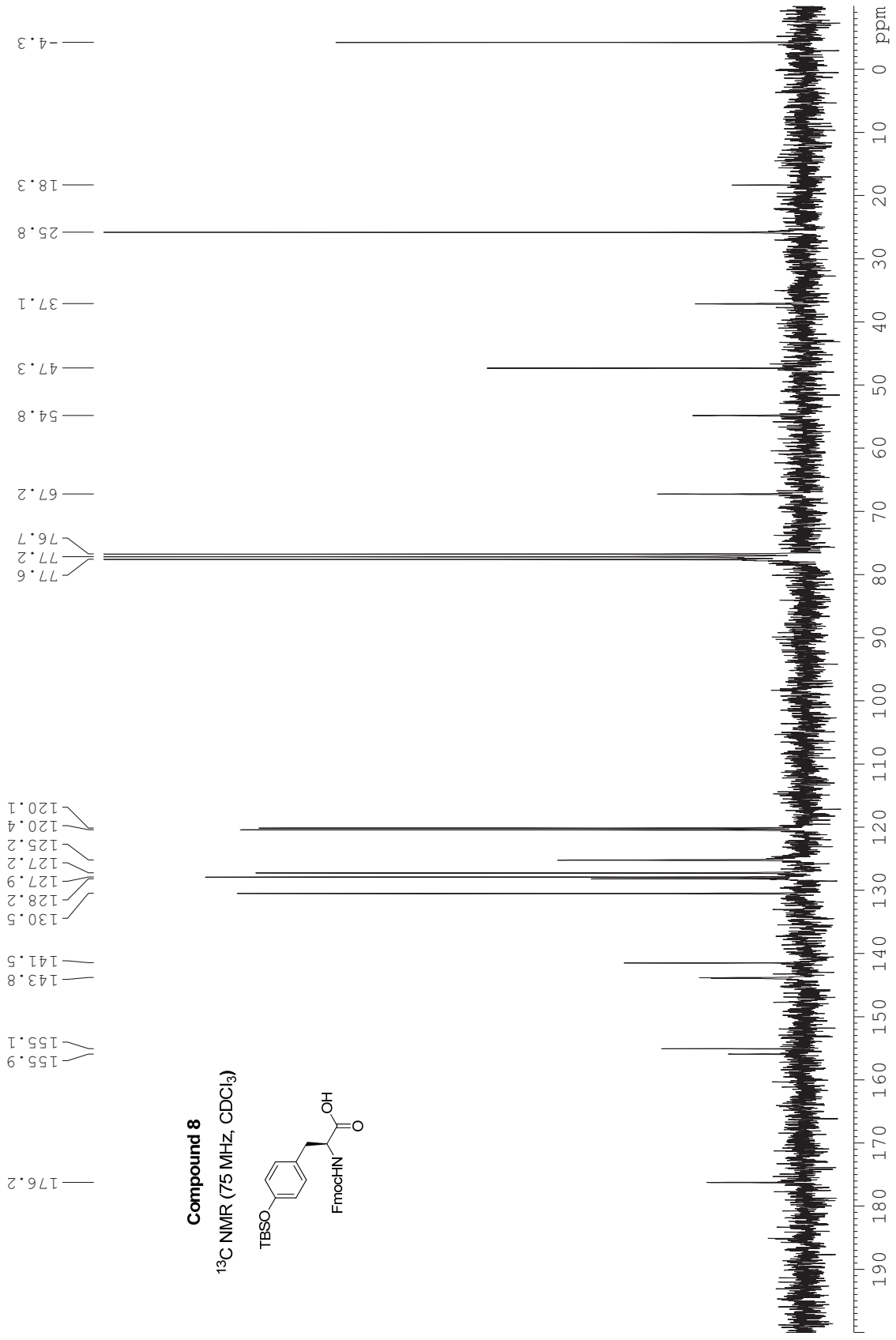

Compound 22
^{1}H NMR (300 MHz, $CDCl_3$)
HO
FmocHN
O
O
O
ppm
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
2.00
2.05
4.15
1.75
2.03
0.90
3.18
5.49
1.04
1.75
7.742
7.717
7.542
7.519
7.392
7.368
7.343
7.297
7.273
7.250
7.248
7.229
6.939
6.911
6.714
6.687
5.898
5.881
5.863
5.842
5.825
5.807
5.787
5.392
5.364
5.319
5.260
5.250
5.215
4.606
4.587
4.435
4.409
4.400
4.375
4.363
4.340
4.305
4.192
4.169
4.146
3.455
3.049
3.029
3.021
3.000


Compound 22
^{13}C NMR (75 MHz, $CDCl_3$)
HO
FmocHN
O
O
O
171.6
156.0
155.3
143.9
143.7
141.4
131.4
130.5
127.8
127.2
125.1
120.1
119.3
115.7
77.6
77.1
76.7
67.2
66.3
55.2
50.8
47.2
37.6
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
ppm

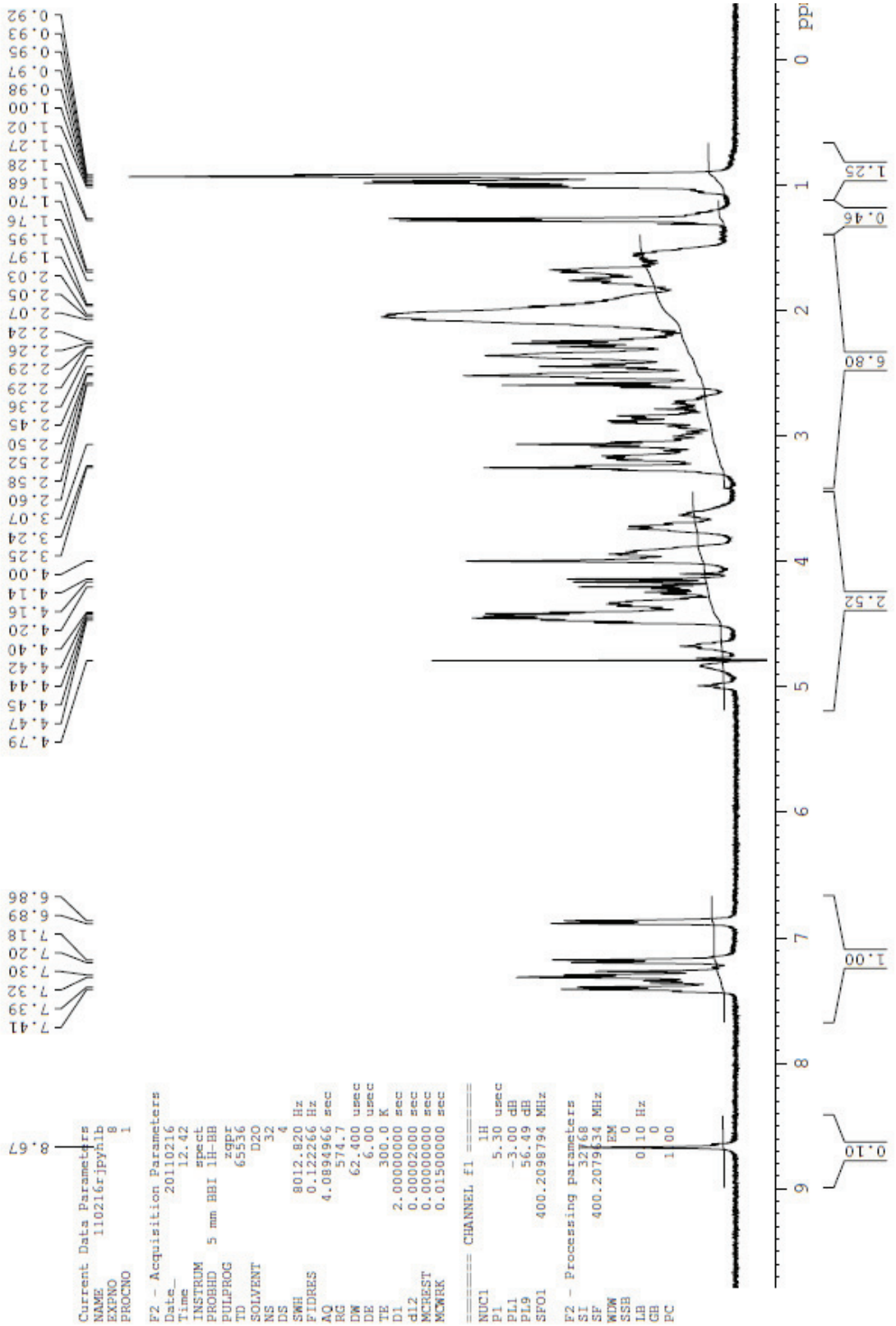

Compound 23
^{1}H NMR (300 MHz, CDCl$_3$)
TBSO
FmocHN
O
O
ppm


Compound 23
^{13}C NMR (75 MHz, CDCl$_3$)
TBSO
FmocHN
O
O
O
-4.3
-3.5
18.1
18.3
25.8
37.6
47.3
55.1
66.1
67.0
76.8
77.2
77.4
77.6
119.1
120.1
120.2
125.2
127.1
127.8
128.4
130.5
131.6
141.4
143.9
144.0
154.9
155.7
171.4
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 ppm

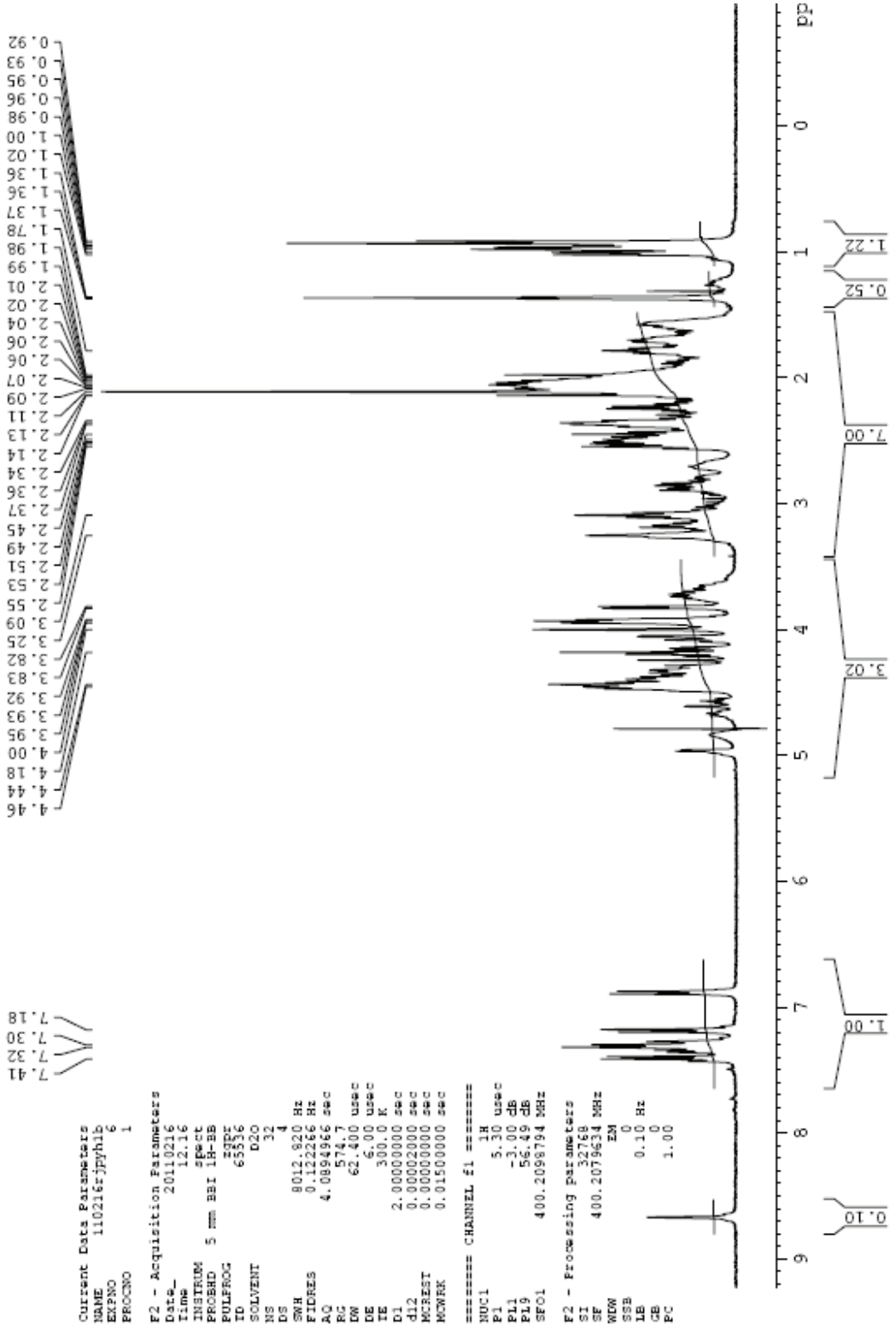

Compound 8
^{1}H NMR (300 MHz, CDCl$_3$)
TBSO
FmocHN
OH
O
ppm


Compound 8
^{13}C NMR (75 MHz, $CDCl_3$)
TBSO
FmocHN
OH
O
-4.3
18.3
25.8
37.1
47.3
54.8
67.2
76.7
77.2
77.6
120.1
120.4
125.2
127.2
127.9
128.2
130.5
141.5
143.8
155.1
155.9
176.2
ppm

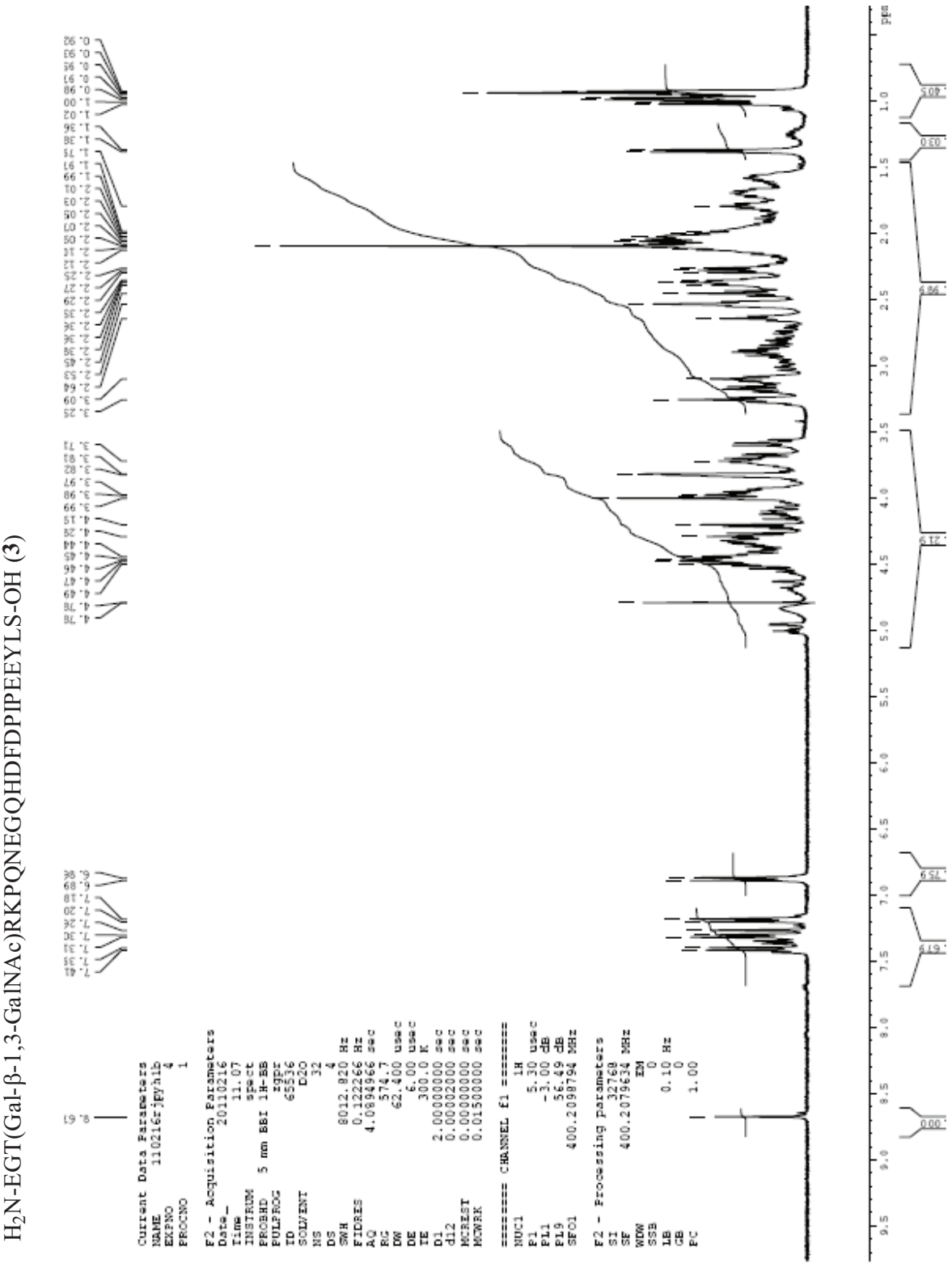
H_2N-EGTRKPQNEGQHDFDPIPEEYLS-OH (**1**)



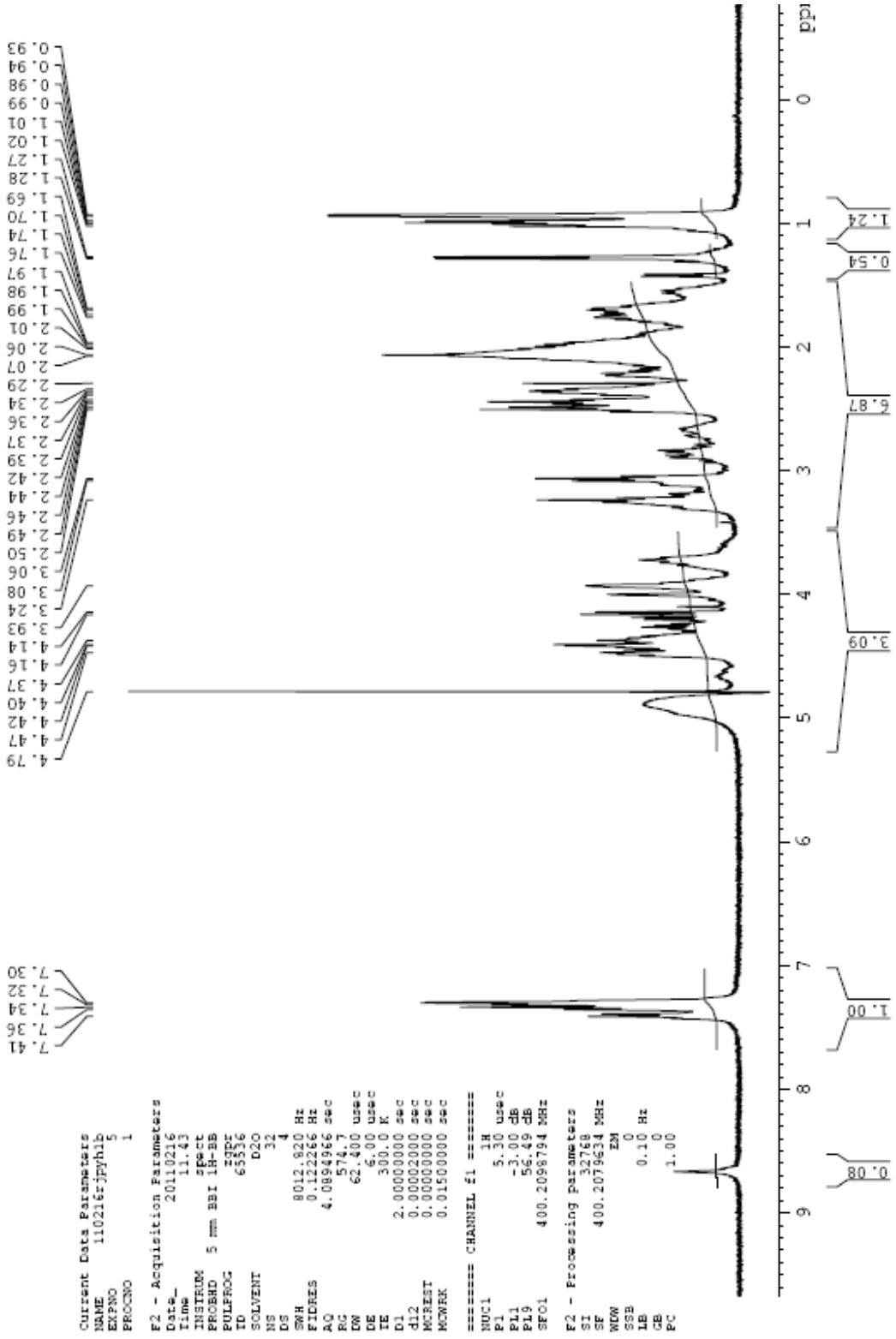
H_2N-EGT(GalNAc)RKPQNEGQHDFDPIPEEYLS-OH (**2**)



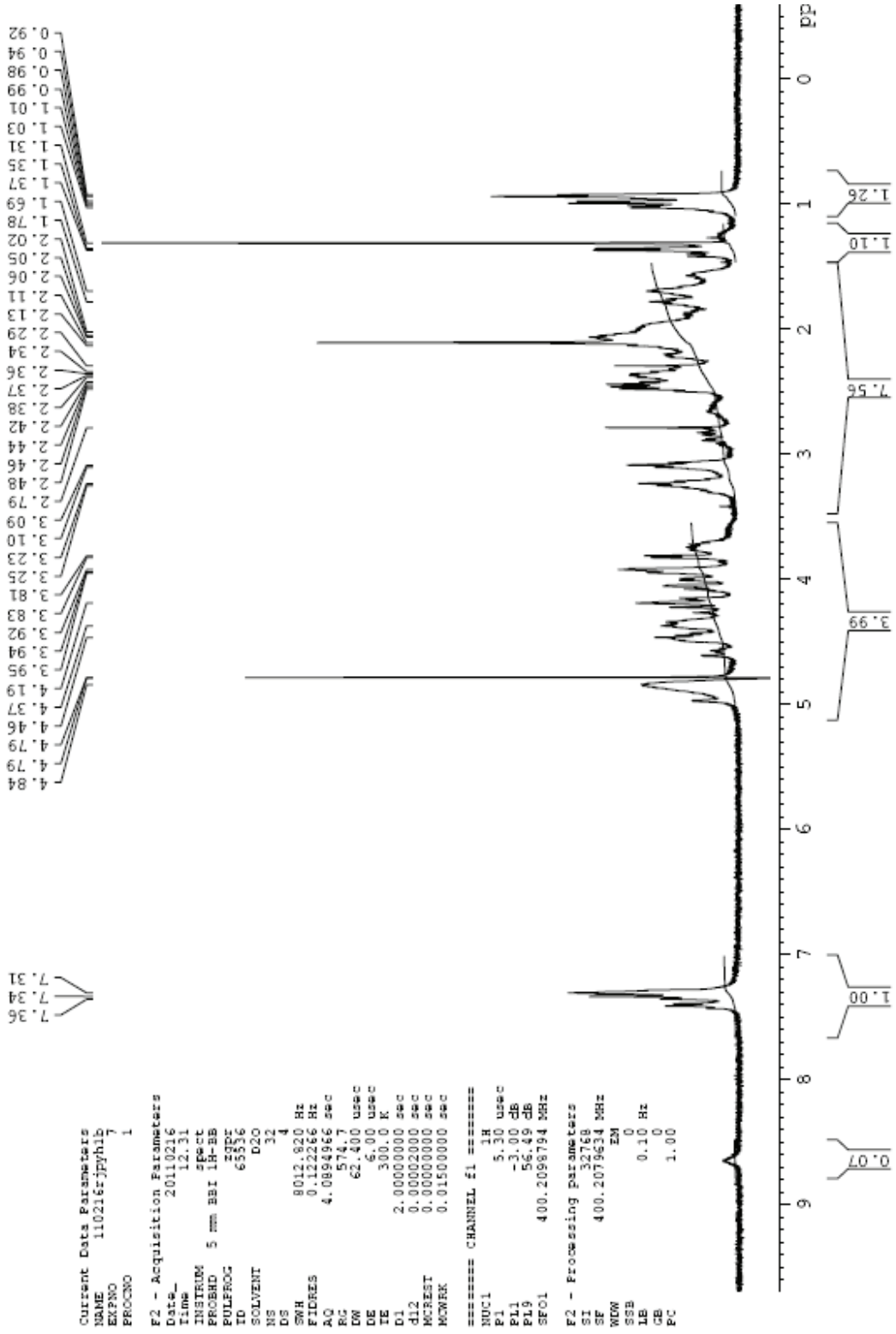
H_2N-EGT(Gal-β-1,3-GalNAc)RKPQNEGQHDFDPIPEEYLS-OH (**3**)



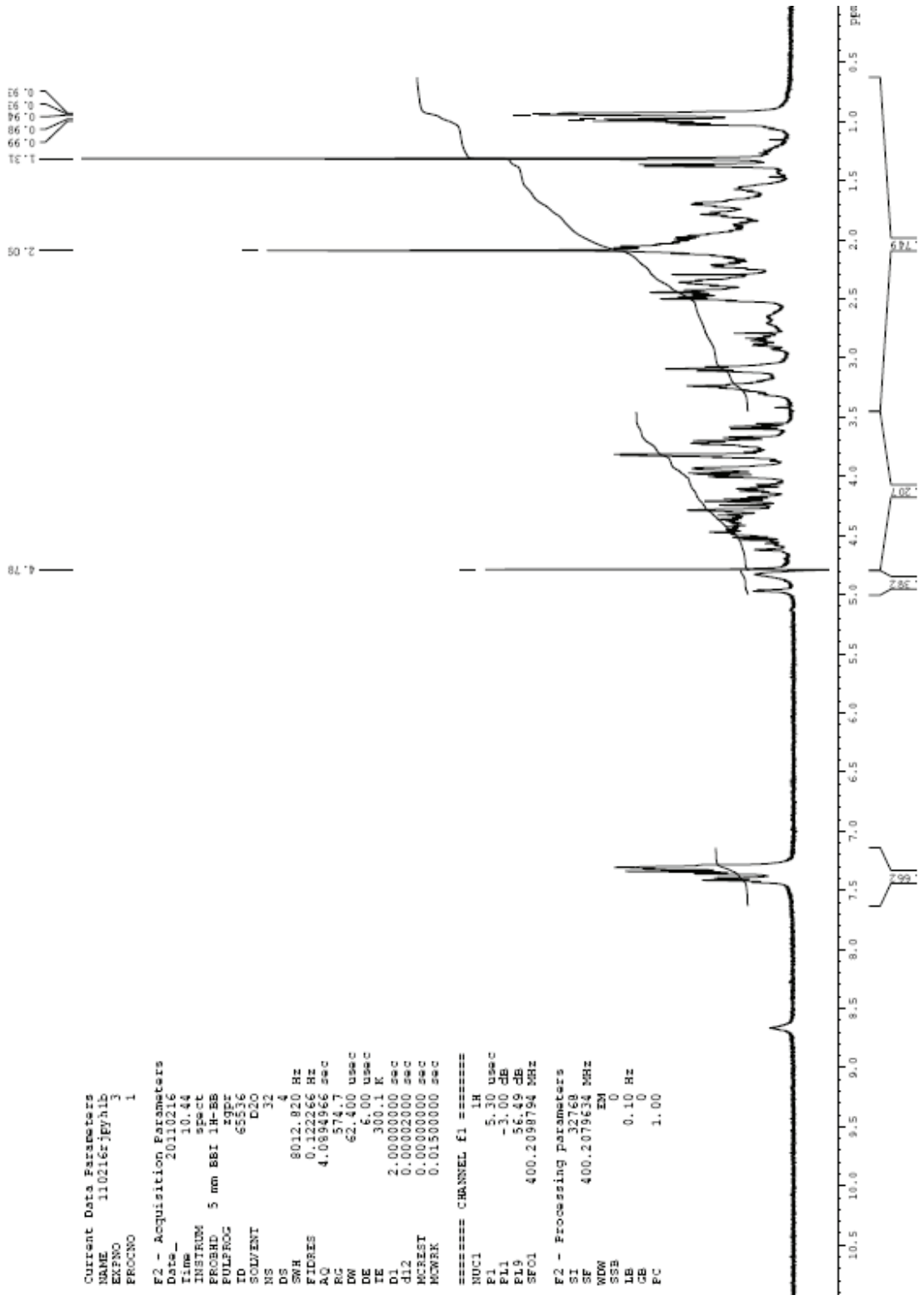
H_2N-EGTRKPQNEGQHDFDPIPEEY($SO_3^-NH_4^+$)LS-OH (**4**)



H_2N-EGT(GalNAc)RKPQNEGQHDFDPIPEEY($SO_3^-NH_4^+$)LS-OH (**5**)



H_2N-EGT(Gal-β-1,3-GalNAc)RKPQNEGQHDFDPIPEEY($SO_3^-NH_4^+$)LS-OH (**6**)



7. References

1. Taleski, D.; Butler, S. J.; Stone, M. J.; Payne, R. J. *Chem. Asian J.* **2011**, *6*, 1316.
2. Ali, A. M.; Taylor, S.D. *Angew. Chem. Int. Ed.* **2009**, *48* 2024.