Synthesis of a well-defined glycoprotein vaccine by a tyrosine-selective conjugation strategy

Qi-Ying Hu,*[a] Martin Allan,[a] Roberto Adamo,[b] Doug Quinn,[a] Huili Zhai,[a] Guangxiang Wu,[a] Kirk Clark,[a] Jing Zhou,[a] Sonia Ortiz,[a] Bing Wang,[a] Elisa Danieli,[b] Stefano Crotti,[b] Marta Tontini,[b] Giulia Brogioni [b] and Francesco Berti[b]

Supporting Information

1. General methods	S2
2.1 Modification of thymopentin in PBS with PTAD	**S4**
2.2 Modification of proteins/peptides with PTAD in Tris buffer	**S6**
2.3 Modification of CRM_{197} with alkyne reagent 3	**S7**
2.4 Synthesis of alkyne reagent 3	**S9**
2.5 Calculation of surface exposed area of tyrosines on CRM_{197}	**S13**
2.6 NMR and IR spectra for 2a and 2b	**S14**
2.7 NMR and IR spectra of linkers for modified CRM_{197}	**S31**
2.8 ESI MS spectra of modified proteins	**S44**
2.9 NMR study of PTAD stability	**S50**
2.10 MS/MS analysis of proteolytic digests (Myoglobin conjugate)	**S54**
3.1 Synthesis of hexasaccharide 5	**S54**
3.2 General procedure for click chemistry with modCRM_{197} 4	**S62**
3.3 NMR spectra of carbohydrate compounds	**S64**

1. General methods

Chemicals, solvents and buffers were acquired from commercial suppliers, and used as is. Thymopentin, myoglobin and α-chymotrypsinogen A were obtained from Sigma-Aldrich. CRM_{197} was obtained from Manufacturing of Novartis Vaccines and Diagnostics. NMR spectra was recorded on a Bruker AVANCE-II NMR spectrometer equipped with a 5 mm Broadband BBFO probe with z-gradient operating at a frequency of 400.13 MHz for ^{1}H NMR, 100.61 MHz for ^{13}C NMR. Chemical shifts were referenced by setting internal tetramethylsilane (TMS) to 0 ppm. In some examples, NMR data were acquired at a temperature of 300 K on a Bruker AVANCE-II NMR spectrometer equipped with a 5 mm TCI CryoProbe with z-gradient operating at a frequency of 600.13 MHz for ^{1}H, 150.92 MHz for ^{13}C, and 60.82 MHz for ^{15}N. Chemical shifts for the ^{1}H and ^{13}C spectra were referenced by setting internal tetramethylsilane (TMS) to 0 ppm. FT-IR was recorded using dry powder or oil on crystal of Thermo Nicolet380 FT-IR. LC/MS was ran on a Waters Acquity with a XBridge C18 column (3x30 mm, 3.5 µm) with 5 mM ammonium formate, 2 min gradient elution (acetonitrile –water, 0-30% in 1.2 min, 30-95% in 0.6 min, 95% for 0.15 min), and at 2 mL/min flow rate. Column used for thymopentin reactions: Acquity CSH column (21 x 50mm, 1.7 µm) with 3.75 mM ammonium acetate and 0.05% formic acid. HRMS was recorded on an Agilent 6220 mass spectrometer with electrospray ionization source and Agilent 1200 liquid chromatography on Inertsil ODS-4 C18 column (3 x100 mm, 3 µm), or on a Waters Xevo G2 QTof with electrospray ionization source on Acquity EBH column (2.1x50 mm, 1.7um). Preparative-HPLC was run on a Waters auto purification system with Sunfire C18 column (30 x 50 mm, 5 µm). Melting point was recorded on a Uni-melt Thomas Hoover capillary melting point apparatus. Protein ESI-MS was recorded on a Waters Xevo G2 QTof with Proswift Monolith column (4.6 x 50 mm), 0.1% formic acid, 2 min gradient elution (3-80% MeCN (0.04% formic acid)), and 3 mL/min flow rate, at 40 °C. SDS-PAGE was run using an Invitrogen Novex Mini-Cell with NuPAGE 4-12% Bis-Tris gel

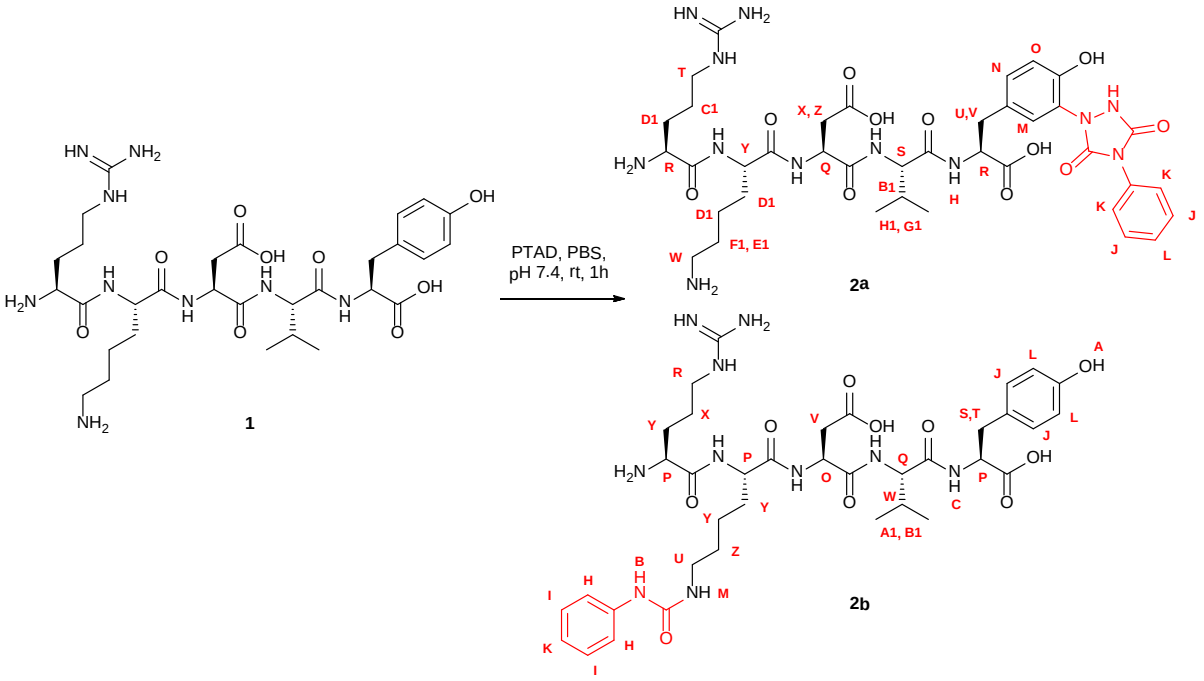
and MOPS running buffer. Invitrogen SeeBlue Plus2 pre-stained standard was used for molecular mass markers.

For the synthesis of carbohydrate derivatives:

All chemicals were of reagent grade, and were used without further purification. Reactions were monitored by thin-layer chromatography (TLC) on Silica Gel 60 F_{254} (Sigma Aldrich); after examination under UV light, compounds were visualized by heating with 10% (v/v) ethanolic H_2SO_4. In the work up procedures, organic solutions were washed with the amounts of the indicated aqueous solutions, then dried with anhydrous Na_2SO_4, and concentrated under reduced pressure at 30–50 ºC on a water bath. Column chromatography was performed on pre-packed silica cartridges RediSep (Teledyne-Isco, 0.040–0.063 nm) or SiliaSep HP (Silicycle, 0.015–0.040 nm). Unless otherwise specified, a gradient 0 → 100% of the elution mixture was applied in a Combiflash R_f (Teledyne-Isco) or Spot II (Armen) instrument. Solvent mixtures less polar than those used for TLC were used at the onset of separation. ^{1}H NMR spectra were measured at 400 MHz and 298 K with a Bruker Avance^{III} spectrometer; δ_H values were reported in ppm, relative to the internal standard Me_4Si (δ_H = 0.00, $CDCl_3$); water signal was used as reference at 4.79 ppm. ^{13}C NMR spectra were measured at 100 MHz and 298 K with a Bruker Avance^{III} spectrometer; δ_C values are reported in ppm relative to the signal of $CDCl_3$ (δ_C = 77.0, $CDCl_3$). Assignments of All NMR signals were assigned by homonuclear and heteronuclear 2-dimensional correlation spectroscopy. In addition ^{13}C NMR spectra of some compounds were aided by comparison with spectra of related substances reported previously from this laboratory or elsewhere. When reporting assignments of NMR signals, sugar residues in oligosaccharides are indicated with capital letters, uncertain attributions are denoted "/". Nuclei associated with the linker are denoted with a prime. Exact masses were measured by electron spray ionization cut-off spectroscopy, using a Q-Tof *micro* Macromass (Waters) instrument. Structures of these compounds follow unequivocally from the mode of synthesis, NMR data and *m/z* values found in their mass spectra. Optical rotation was measured with a P-2000 Jasco polarimeter at 25 ºC.

2.1 Modification of thymopentin in PBS with PTAD

Scheme 1



To thymopentin (**1**, 51 mg, 75 µmol) in 250 mL PBS pH 7.4 (300 µM) at room temperature was added a solution of PTAD in acetonitrile (83 µmol, 1.65 mL, 50 mM). The reaction was stirred for 16 h at this temperature, and then lyophilized. The residue was purified on HPLC (10-30% CH_3CN / water (0.1% TFA). **2a** (6 mg, 9%) was isolated as a colorless solid after concentration.

^{1}H NMR (600 MHz, DMSO-d6) δ 12.75 (br s, 1H), 12.44 (br s, 1H), 11.14 (br s, 1H), 10.03 (br s, 1H), 8.62 (**E**, d, *J* = 8.1 Hz, 1H), 8.52 (d, *J* = 7.5 Hz, 1H), 8.32 (H, d, *J* = 7.5 Hz, 1H), 8.20 (br s, 3H), 7.76 (s, 3H), 7.64 (s, 1H), 7.57 – 7.47 (**J**, **K**, m, 4H), 7.44 (**L**, tt, *J* = 7.1, 1.7 Hz, 1H), 7.28 (**M**, d, *J* = 2.2 Hz, 1H), 7.17 (**N**, dd, *J* = 8.6, 2.2 Hz, 1H), 6.90 (**O**, d, *J* = 8.3 Hz, 1H), 6.57 (s, 1H), 4.58 (**Q**, q, *J* = 7.2 Hz, 1H), 4.36 (**R**, td, *J* = 8.2, 5.5 Hz, 2H), 4.24 (**S**, dd, *J* = 9.0, 5.7 Hz, 1H), 3.83 (**Y**, br s, 1H), 3.12 (**T**, q, *J* = 6.6 Hz, 2H), 2.99 (**U**, dd, *J* = 14.2, 5.5 Hz, 1H), 2.85 (**V**, dd, *J* = 14.2, 8.6 Hz, 1H), 2.74 (**S**, m, 2H), 2.69 (**X**, dd, *J* = 16.9, 5.9 Hz, 1H), 2.54 (**Z**, m, 1H),

2.00 (**B1**, dq, *J* = 13.4, 6.7 Hz, 1H), 1.74-1.64 (**C1**, m, 3H), 1.57-1.42 (**D1**, m, 5H), 1.36 (**E1**, m, 1H), 1.34 – 1.26 (**F1**, m, 1H), 0.83 (**H1**, d, *J* = 6.8 Hz, 3H), 0.78 (**G1**, d, *J* = 6.8 Hz, 3H).

^{13}C NMR (151 MHz, DMSO-d6) δ 173.2, 172.2, 171.6, 171.2, 170.7, 168.7, 157.2, 153.1, 152.7, 151.9, 132.4, 131.7, 130.2, 129.4, 128.7, 128.3, 126.6, 123.5, 117.2, 57.3, 54.2, 52.6, 52.2, 50.0, 40.6, 39.1, 36.1, 35.9, 32.4, 31.6, 28.9, 27.2, 24.6, 22.6, 19.7, 17.9.

LC-MS (MH^+) 855.8, t = 0.69 min. HRMS: calculated for $C_{38}H_{55}N_{12}O_{11}$ (MH^+) 855.4113, observed 855.4099. IR: 3077, 2363, 2163, 1667, 1519, 1434. m.p.: 240 °C. **2b** (10 mg, 17%) was also isolated as a colorless solid after concentration.

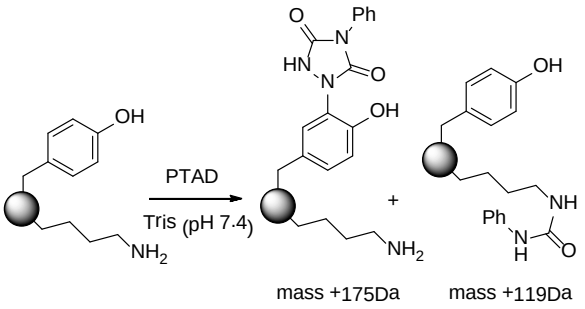
^{1}H NMR (600 MHz, DMSO-d6) δ 12.62 (br s, 1H), 12.45 (br s, 1H), 9.24 (**A**, s, 1H), 8.77 (**B**, s, 1H), 8.42 (br s, 1H), 8.31 (br s, 2H), 8.21 (**C**, d, *J* = 7.5 Hz, 1H), 7.67 (br s, 3H), 7.37 (**H**, d, *J* = 7.9 Hz, 2H), 7.23 (**I**, t, *J* = 7.9 Hz, 2H), 7.01 (**J**, d, *J* = 8.2 Hz, 2H), 6.91 (**K**, tt, *J* = 7.3, 1.2 Hz, 1H), 6.66 (**L**, d, *J* = 8.2 Hz, 2H), 6.47 (**M**, d, *J* = 8.1 Hz, 1H), 4.54 (**O**, br s, 1H), 4.32 (**P**, m, 3H), 4.22 (**Q**, dd, *J* = 9.1, 5.9 Hz, 1H), 3.12 (**R**, s, 2H), 2.91 (**S**, dd, *J* = 14.0, 5.7 Hz, 1H), 2.79 (**T**, dd, *J* = 14.0, 8.6 Hz, 1H), 2.77 – 2.70 (**U**, m, 2H), 2.66 (**V**, m, 2H), 1.95 (**W**, dq, *J* = 13.4, 6.7 Hz, 1H), 1.69 (**X**, br s, 2H), 1.52 (**Y**, m, 6H), 1.42 – 1.19 (**Z**, m, 2H), 0.81 (**A1**, d, *J* = 6.7 Hz, 3H), 0.76 (**B1**, d, *J* = 6.7 Hz, 3H).

^{13}C NMR (151 MHz, DMSO-d6) δ 173.3, 172.9, 172.1, 171.0, 170.6, 157.2, 156.4, 155.2, 140.7, 130.4, 129.2, 127.8, 121.7, 120.8, 119.6, 118.8, 117.9, 116.8, 115.5, 114.8, 57.3, 54.3, 52.5, 52.3, 50.0, 40.5, 39.1, 36.3, 36.2, 32. 1, 31.6, 31.4, 27.1, 25.3, 22.8, 19.6, 18.0.

LC-MS (MH^+) 799.8, t = 0.88 min. HRMS: calculated for $C_{37}H_{55}N_{10}O_{10}$ (MH^+) 799.4103, observed 799.4092. IR: 3271, 1630, 1548. m.p.: 225°C.

2.2 Modification of proteins/peptides with PTAD in Tris buffer

Table 1. Condition optimization on thymopentin and proteins



substrate	condition[a]		Conjugates (%)[b]		
	T °C	Tris (M)	unmodified	K	Y[c]
thymopentin	20	1.0	22	0	78
myoglobin	20	0.2	90	10	0
	4	1.0	87	0	13
α-chymotrypsinogen A	20	0.2	45	10	45
	4	0.2	16	7	77
	4	1.0	25	0	75
CRM_{197}	20	0.2	0	0	100

[a]: 1.1 eq. of PTAD was used for thympentin, and 30 eq. of PTAD was used for other cases. [b]: Determined according to peak intensity; [c]: W labeling was not excluded.

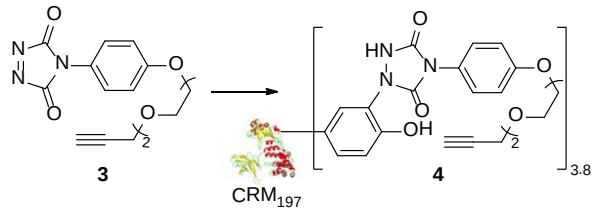
Protocol for proteins: To a solution of protein in Tris buffer (120 µL, 25 µM, 0.003 µmol) at the desired temperature was added 3 portions of solution of PTAD in acetonitrile (3 x 0.03 µmol, 3x1.2 µL, 25 mM) every minute, and vortexed after each addition. Each reaction was then agitated for 15 min, and analyzed by ESI-MS. The ratio of products was determined according to peak intensity.

* Protocol for thymopentin: To thymopentin (5.1 mg, 7.5 µmol) in 25 mL of buffer (300 µM) at room temperature was added a solution of PTAD in acetonitrile (165 µL, 50 mM, 8.25 µmol).

The reaction was stirred for 16 h, and then analyzed by LC-MS. The ratio of products was determined according to the area under the peak.

2.3 Modification of CRM_{197} with alkyne reagent 3

Scheme 2



To a solution of CRM_{197} (2.92 mg, 0.05 µmol) in Tris·HCl (2.0 mL, 0.2M, pH 7.4) was added 10 additions of freshly prepared reagent **3** (2.5 µL, 100 mM in CH_3CN, 0.25 µmol) every minute (50 eq. in total). The mixture was agitated at 20 ℃ for 30 min, and then desalted and buffer was exchanged to PBS pH 7.4 three times with Zeba 7K MWCO spin column (Pierce). 2 mL (1.2 mg/mL) of modified CRM_{197} **4** was recovered (82%, according to its UV absorption at 210 nm in comparison with standard). An average loading of 3.8 modified tyrosines per protein was derived from the peak intensity on ESI-MS.

Protein digestion and mapping of CRM_{197}-alkyne (4)

Sample Proteolysis:

Three aliquots of 100 picomol of CRM_{197}-alkyne (**4**) were diluted to a final volume of 50µl each in 3% dimethyl sulfoxide, 5 mM dithiothreitol (DTT) and 50 mM ammonia bicarbonate buffer. Proteolysis was then performed on each of the three aliquots with an additional 5 µl of either 0.25 micrograms of trypsin (Promega), 0.25 micrograms of chymotrypsin (Roche) or 0.5 micrograms of GluC (Roche). The digests took place at 23 ºC overnight.

LC-MS/MS:

LC−MS/MS analyses were performed on a Thermo Scientific LTQ Orbitrap Discovery ion trap MS instrument equipped with a Agilent 1200 LC. The MS operated with Xcalibur version 2.0.7 system control and data analysis software, while the Agilent LC operated with Chemstation B.03.02-SR1 software. Analysis of samples was performed with an acetonitrile gradient and a Symmetry 300 column (Waters), with dimensions of 0.3 x 150 millimeters. The LC gradient

began at water:acetonitrile:formic acid (96:4:0.01, v/v), and at 1 min, a 54 min linear gradient increased acetonitrile to 45%. The slow gradient was followed by a rapid 8 minute linear gradient to 95% acetonitrile and a brief 3 minute hold period to wash the column. Total analysis time for the run was 80 min. The 5 microliter flow from the HPLC was introduced into the MS with a HESI source. The MS method performed a full scan followed by ten successive MS/MS experiments.

Analysis:

The data was acquired and analyzed using Thermo Scientific Excalibur software. Additional analysis was performed with Mascot software(Matrix Science) to identify mass modifications. All quantitative assessments were made using peak area derived from the Excalibur software.

Semi-quantization performed with peak intensity of modified vs unmodified.

CRM197 JAN 28 2011
Found in search of L:\Analytics\BasicResults\LTQ\2011\10\111010_CRM062_LYS_Glu.mgf

Nominal mass (M_r): **58377**; Calculated pI value: **5.85**
NCBI BLAST search of sp822 against nr
Unformatted sequence string for pasting into other applications

Variable modifications: CRM 062 TYR OCT 2011 (Y)
No enzyme cleavage specificity
Sequence Coverage: **70%**

Matched peptides shown in **Bold Red**

```
     1 GADDVVDSSK SFVMENFSSY HGTKPGYVDS IQKGIQKPKS GTQGNYDDDW
    51 KEFYSTDNKY DAAGYSVDNE NPLSGKAGGV VKVTYPGLTK VLALKVDNAE
   101 TIKKELGLSL TEPLMEQVGT EEFIKRFGDG ASRVVLSLPF AEGSSSVEYI
   151 NNWEQAKALS VELEINFETR GKRGQDAMYE YMAQACAGNR VRRSVGSSLS
   201 CINLDWDVIR DKTKTKIESL KEHGPIKNKM SESPNKTVSE EKAKQYLEEF
   251 HQTALEHPEL SELKTVTGTN PVFAGANYAA WAVNVAQVID SETADNLEKT
   301 TAALSILPGI GSVMGIADGA VHHNTEEIVA QSIALSSLMV AQAIPLVGEL
   351 VDIGFAAYNF VESIINLFQV VHNSYNRPAY SPGHKTQPFL HDGYAVSWNT
   401 VEDSIIRTGF QGESGHDIKI TAENTPLPIA GVLLPTIPGK LDVNKSKTHI
   451 SVNGRKIRMR CRAIDGDVTF CRPKSPVYVG NGVHANLHVA FHRSSSEKIH
   501 SNEISSDSIG VLGYQKTVDH TKVNSKLSLF FEIKS
```

CRM197 JAN 28 2011
Found in search of L:\Analytics\BasicResults\LTQ\2011\10\111010_CRM062_LYS.mgf

Nominal mass (M_r): **58377**; Calculated pI value: **5.85**
NCBI BLAST search of sp822 against nr
Unformatted sequence string for pasting into other applications

Variable modifications: CRM 062 TYR OCT 2011 (Y)
Cleavage by Lys-C: cuts C-term side of K unless next residue is P
Sequence Coverage: **50%**

Matched peptides shown in **Bold Red**

```
     1 GADDVVDSSK SFVMENFSSY HGTKPGYVDS IQKGIQKPKS GTQGNYDDDW
    51 KEFYSTDNKY DAAGYSVDNE NPLSGKAGGV VKVTYPGLTK VLALKVDNAE
   101 TIKKELGLSL TEPLMEQVGT EEFIKRFGDG ASRVVLSLPF AEGSSSVEYI
   151 NNWEQAKALS VELEINFETR GKRGQDAMYE YMAQACAGNR VRRSVGSSLS
   201 CINLDWDVIR DKTKTKIESL KEHGPIKNKM SESPNKTVSE EKAKQYLEEF
   251 HQTALEHPEL SELKTVTGTN PVFAGANYAA WAVNVAQVID SETADNLEKT
   301 TAALSILPGI GSVMGIADGA VHHNTEEIVA QSIALSSLMV AQAIPLVGEL
   351 VDIGFAAYNF VESIINLFQV VHNSYNRPAY SPGHKTQPFL HDGYAVSWNT
   401 VEDSIIRTGF QGESGHDIKI TAENTPLPIA GVLLPTIPGK LDVNKSKTHI
   451 SVNGRKIRMR CRAIDGDVTF CRPKSPVYVG NGVHANLHVA FHRSSSEKIH
   501 SNEISSDSIG VLGYQKTVDH TKVNSKLSLF FEIKS
```

above 80% modified; (Tyr 27, 46, 358, and 380)

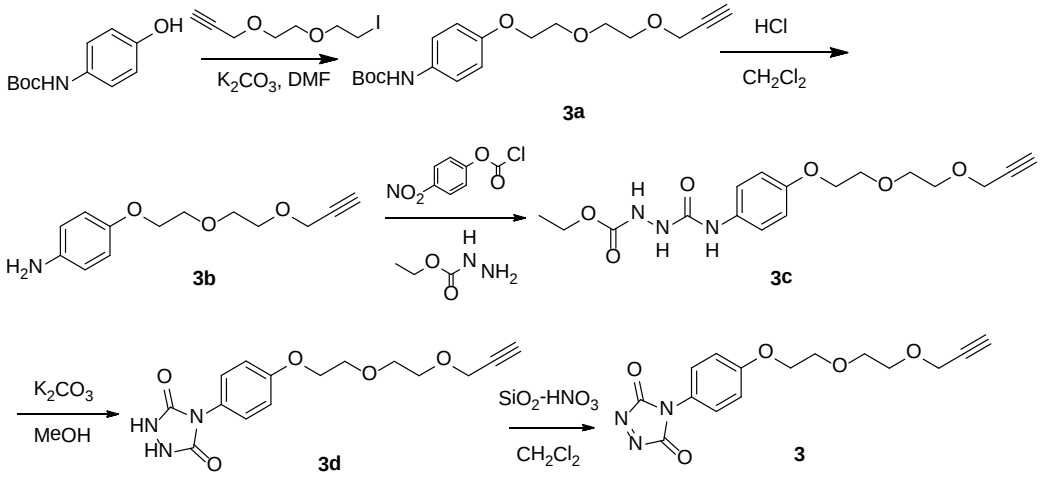
20-80% modified; (Tyr 60, 65, 85, 246)

below 20% modified; (Tyr 149, 514, and Trp 50)

Red means the sequence was detected in the study. Black means the sequence was missed in the study. The MS/MS digests mapping has been performed twice to ensure better sequence coverage.

2.4 Synthesis of alkyne reagent 3

Scheme 4



Step 1:

tert-butyl 4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamate (3a)

To a solution of *tert*-butyl 4-hydroxyphenylcarbamate (1.675 g, 8.01 mmol) in DMF (53.4 mL) was added 3-(2-(2-iodoethoxy)ethoxy)prop-1-yne[1] (2.034 g, 8.01 mmol) and K_2CO_3 (3.32 g, 24.02 mmol). The mixture was stirred at 60 °C for 4 hours followed by the addition of saturated sodium bicarbonate solution. The mixture was then extracted with ethyl acetate twice, dried over anhydrous $MgSO_4$, filtered, and then concentrated in vacuo. The residue was purified by silica gel chromatography (0-40%, v/v, ethyl acetate/heptane) giving *tert*-butyl 4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamate (1.91 g, 71%) as colorless oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 1.50 (s, 9H, -C(C**H**$_3$)$_3$), 2.42 (t, *J* = 2.27 Hz, 1H, -C≡C**H**), 3.65-3.80 (m, 4H, -C**H**$_2$OC**H**$_2$-) 3.84 (t, *J* = 4.80 Hz, 2H, -OCH$_2$C**H**$_2$O-), 4.10 (t, *J* = 4.93 Hz, 2H, -OC**H**$_2$CH$_2$O-), 4.21 (d, *J* = 2.27 Hz, 2H, -OC**H**$_2$C≡CH), 6.33 (br. s., 1H, -CON**H**C-), 6.85 (d, *J* = 9.09 Hz, 2H, -NCCHC**H**C-), 7.17-7.33 (m, 2H, -NCC**H**CHC-). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 28.39, 58.47, 67.83, 69.71, 69.84, 70.66, 74.56, 79.65, 80.26, 115.12, 120.48, 131.66, 153.13, 154.87. HRMS: calculated for $C_{18}H_{26}NO_5$ (MH^+) 336.1806, observed 336.1798. FTIR: 3294, 2976, 2872, 1702.

Step 2:

4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)aniline (3b)

To a solution of *tert*-butyl 4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamate (1.91 g, 5.69 mmol) in CH_2Cl_2 (14.24 mL) was added HCl in dioxane (14.24 mL, 4M, 56.9 mmol). The mixture was stirred at room temperature for 2 hours, concentrated in vacuo, followed by the addition of aqueous saturated sodium bicarbonate. The mixture was then extracted twice with ethyl acetate, dried over anhydrous$MgSO_4$, filtered, and concentrated in vacuo. 4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)aniline was obtained as a colorless oil, and used "as is" for the next step. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.42 (t, *J* = 2.40 Hz, 1H, -C≡C**H**), 3.73 (m, 4H, -C**H**$_2$OC**H**$_2$-), 3.78-3.93 (m, 2H, -OCH$_2$C**H**$_2$O-), 3.99-4.17 (m, 2H, -OC**H**$_2$CH$_2$O-), 4.21 (d, *J* = 2.53 Hz, 2H, -OC**H**$_2$C≡CH), 6.63 (d, *J* = 8.84 Hz, 2H, -NCCHC**H**C-), 6.76 (d, *J* = 8.84 Hz, 2H, -NCC**H**CHC-). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 58.48, 68.19, 69.17, 69.96, 70.60, 74.51, 79.66, 115.93,

116.41, 140.04, 152.03. HRMS: calculated for $C_{13}H_{18}NO_3$ (MH^+) 236.1281, observed 236.1280. FTIR: 3354, 3269, 2870, 1626.

Step 3:

Ethyl 2-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamoyl)hydrazinecarboxylate (3c)

To a solution of 4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)aniline (1.339 g, 5.69 mmol) in THF (37.9 mL) at 0 ℃ was added 4-nitrophenyl chloroformate (2.064 g, 10.24 mmol) and triethylamine(1.420 mL, 10.24 mmol). The mixture was stirred at room temperature for 1 hour. Then ethyl carbazate (1.540 g, 14.79 mmol) and triethylamine (1.420 mL, 10.24 mmol) were added. The mixture was stirred at room temperature for 16 hours, followed by the addition of saturated sodium bicarbonate solution. The mixture was then extracted twice with ethyl acetate, dried over anhydrous $MgSO_4$, filtered, and then concentrated in vacuo. The crude product was purified by silica gel chromatography (40-100%, v/v, ethyl acetate/heptane) giving ethyl 2-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamoyl)hydrazinecarboxylate (1.31 g, 63% for 2 steps) as a foamy colorless oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 1.22 (t, *J* = 7.07 Hz, 3H, C**H**$_3$CH$_2$O-), 2.43 (t, *J* = 2.27 Hz, 1H, -C≡C**H**), 3.63-3.78 (m, 4H, -C**H**$_2$OC**H**$_2$-), 3.82 (t, *J* = 4.67 Hz, 2H, -OCH$_2$C**H**$_2$O-), 4.05 (t, *J* = 4.67 Hz, 2H, -OC**H**$_2$CH$_2$O-), 4.13 (q, *J* = 7.07 Hz, 2H, CH$_3$C**H**$_2$O-), 4.20 (d, *J* = 2.27 Hz, 2H, -OC**H**$_2$C≡CH), 6.76 (d, *J* = 9.09 Hz, 2H, -NCCHC**H**C-), 7.07-7.33 (m, 4H, -NCC**H**CHC-,-NHN**H**CON**H**C-), 7.56 (br. s., 1H, -CON**H**NHCO-). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 14.36, 58.42, 62.51, 67.69, 69.11, 69.80, 70.59, 74.68, 79.57, 114.96, 122.00, 131.07, 155.22, 156.51, 157.76. HRMS: calculated for $C_{17}H_{24}N_3O_6$ (MH^+) 366.1660, observed 366.1656. FTIR: 3256, 3083, 2899, 1738, 1682.

Step 4:

4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-1,2,4-triazolidine-3,5-dione (3d)

To ethyl 2-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamoyl)-hydrazinecarboxylate (1.31 g, 3.59 mmol) in methanol (17.93 mL) was added K_2CO_3 (1.239 g, 8.96 mmol). The mixture was stirred at 55 ℃ for 30 minutes. After cooling to room temperature, HCl in dioxane

(3.59 mL, 4M, 14.34 mmol) was added. The precipitate was filtered, and the filtrate was concentrated in vacuo. The residue was then purified by silica gel chromatography (40-100% ethyl acetate/heptane, v/v) giving 4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-1,2,4-triazolidine-3,5-dione (318 mg, 28%) as a colorless solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 2.73 (t, *J* = 2.27 Hz, 1H, -C≡C**H**), 3.56-3.73 (m, 4H, -C**H**$_2$OC**H**$_2$-), 3.75-3.88 (m, 2H, -OCH$_2$C**H**$_2$O-), 4.09-4.24 (m, 4H, -OCH$_2$C**H**$_2$O-, -OC**H**$_2$C≡CH), 7.05 (d, *J* = 9.09 Hz, 2H, -NCCHC**H**C-), 7.35 (d, *J* = 8.84 Hz, 2H, -NCC**H**CHC-), 7.81 (br. s., 2H, -CON**H**N**H**CO-). ^{13}C NMR (100 MHz, CD_3CN) δ ppm 57.37, 67.36, 68.58, 68.84, 69.75, 74.35, 79.56, 114.43, 124.16, 127.53, 154.87, 158.16. HRMS: calculated for $C_{15}H_{18}N_3O_5$ (MH^+) 320.1241, observed 320.1230. FTIR: 3238, 2903, 1773, 1694. m.p.: 110 °C.

Step 5:

4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-3H-1,2,4-triazole- 3,5(4H)-dione (3)

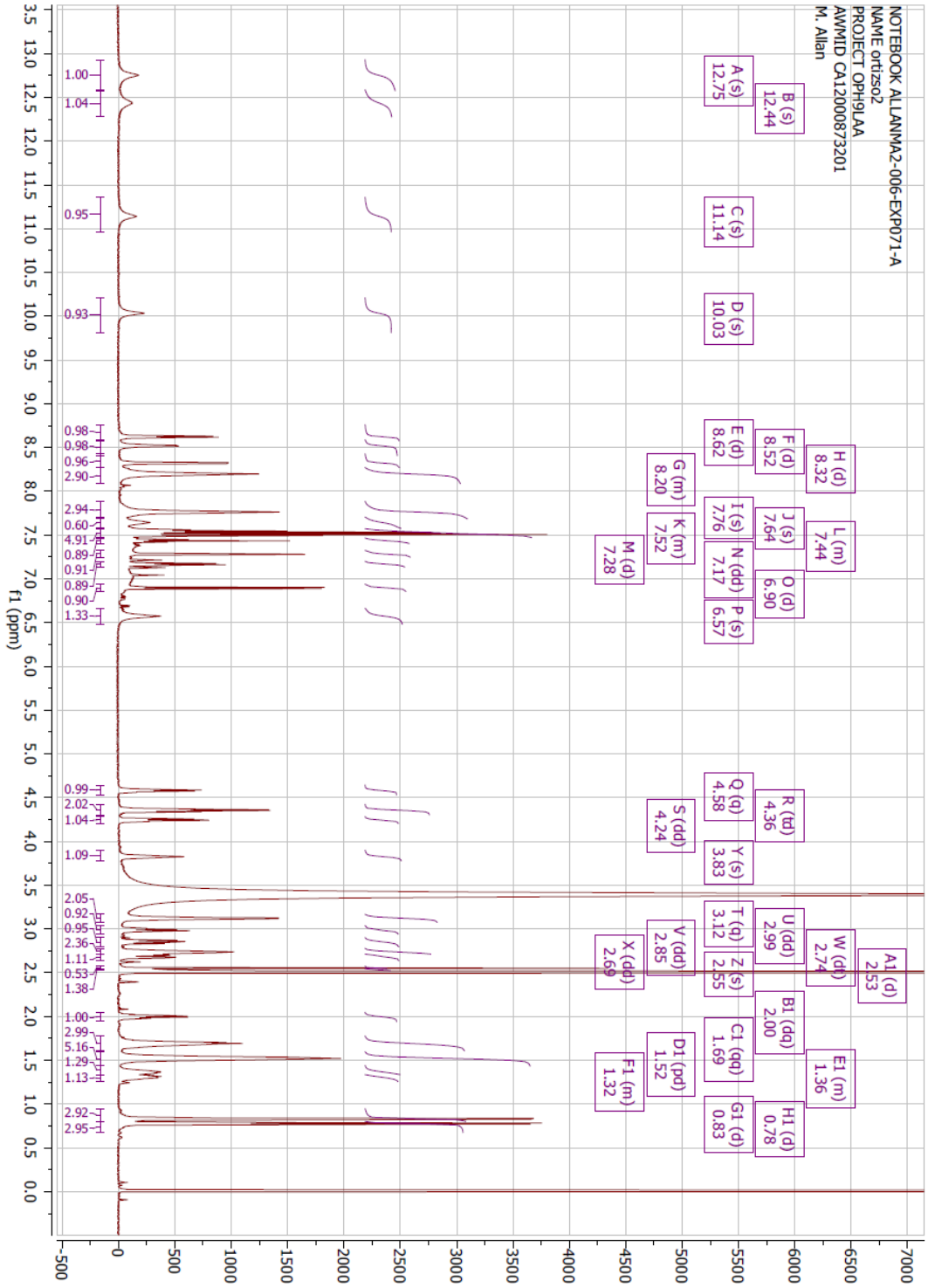
To a solution of 4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-1,2,4-triazolidine-3,5-dione (153 mg, 0.479 mmol) in CH_2Cl_2 (4.792 mL) was added SiO_2-HNO_3[2] (300 mg). The mixture was stirred at room temperature for 15 minutes, then filtered, concentrated in vacuo and dried giving 4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-3H-1,2,4-triazole-3,5(4H)-dione (112 mg, 74%) as a red oil. ^{1}H NMR (400 MHz,CD_3CN) δ ppm 2.60 (t, *J* = 2.40 Hz, 1H, -C≡C**H**), 3.44-3.62 (m, 4H, -C**H**$_2$OC**H**$_2$-), 3.68-3.75 (m, 2H, -OCH$_2$C**H**$_2$O-), 4.02-4.12 (m, 4H, -OCH$_2$C**H**$_2$O-, -OC**H**$_2$C≡CH), 7.02 (d, *J* = 9.09 Hz, 2H, -NCCHC**H**C-), 7.23 (d, *J* = 8.84 Hz, 2H, -NCC**H**CHC-).

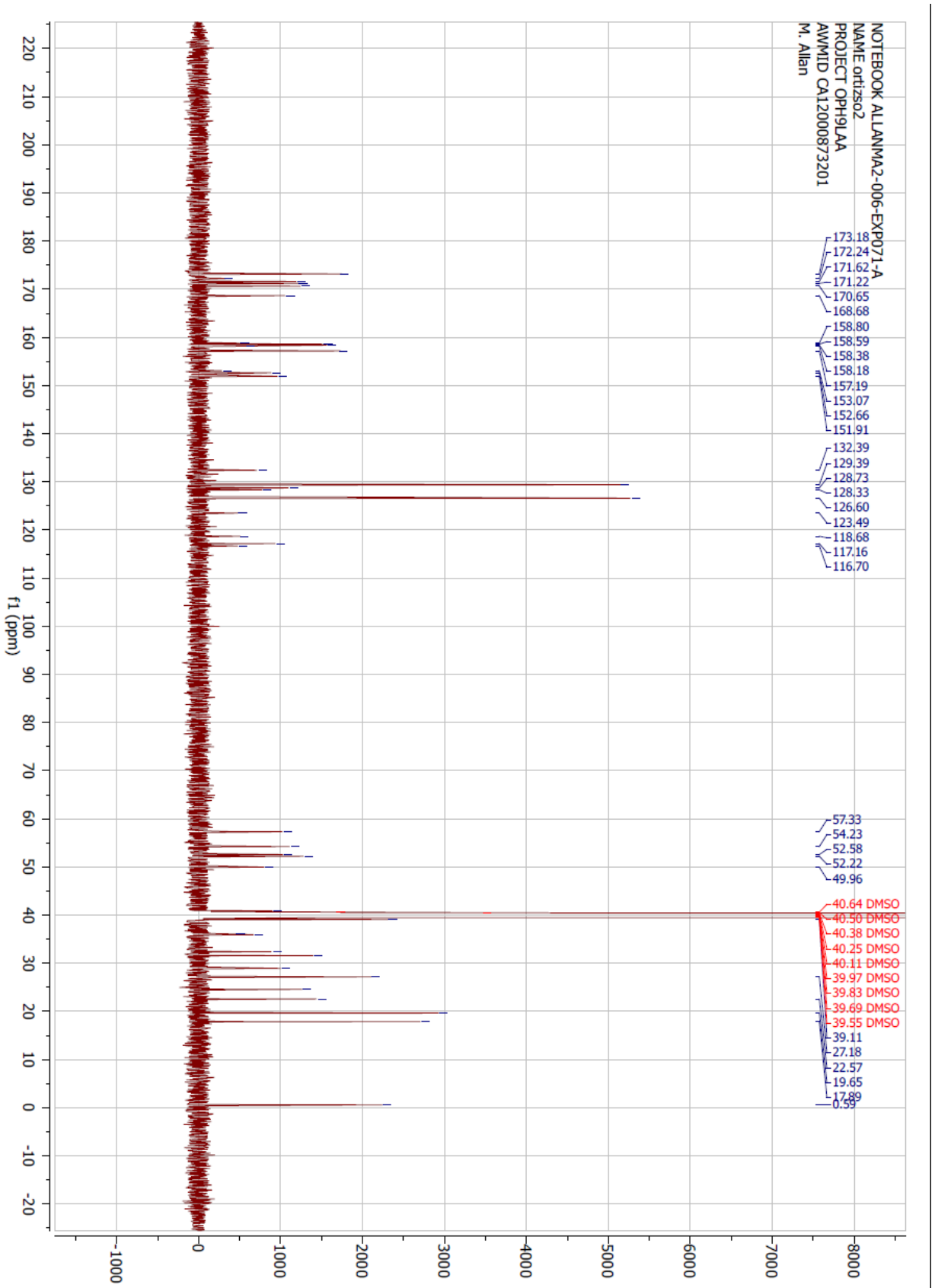
2.5 Calculation of surface exposed area of tyrosines on CRM_{197}

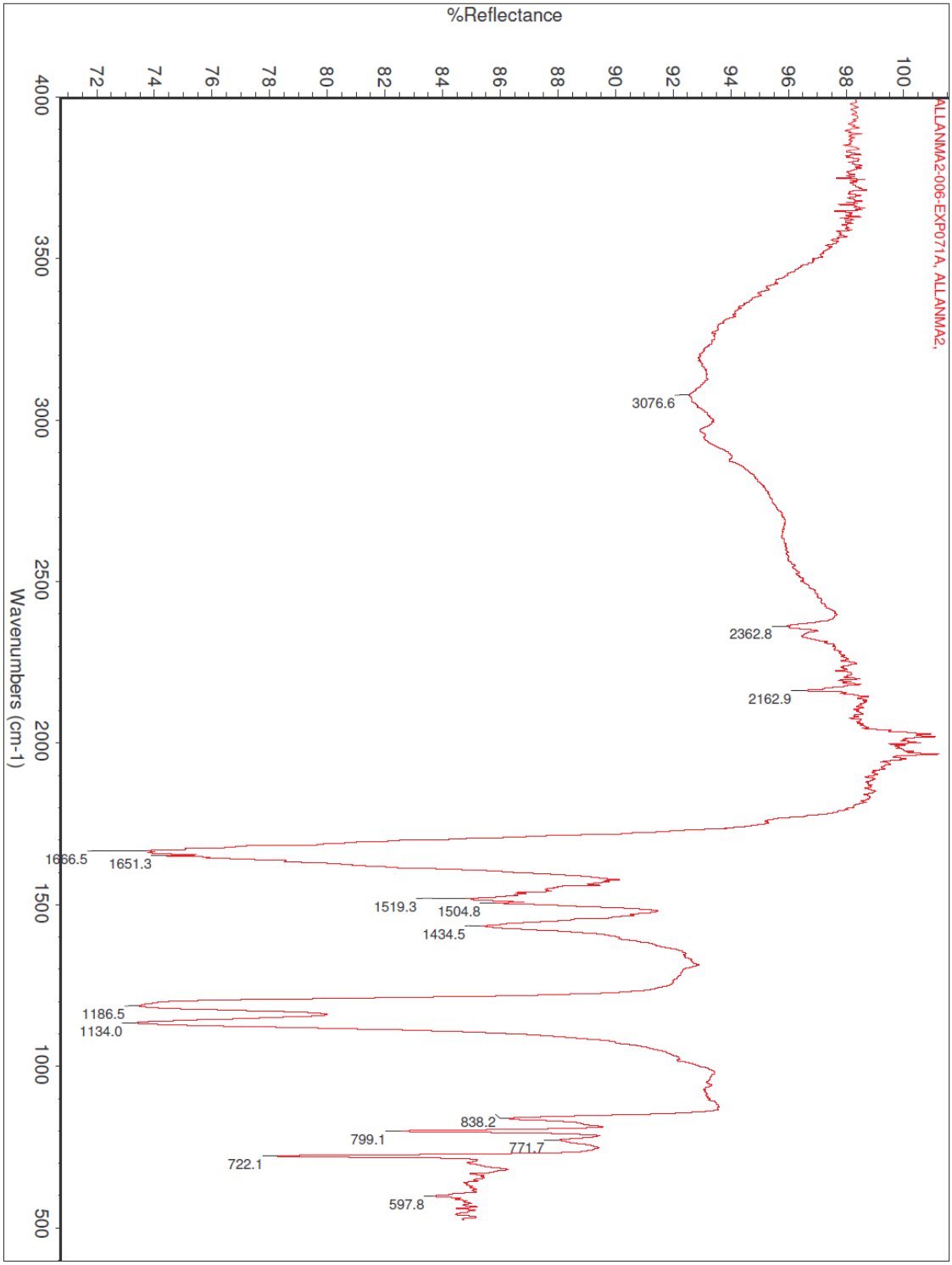
Residue	number	Exposed area	
TYR	278	0.6	None
TYR	85	0.7	Low
TYR	246	1.1	None
TYR	394	4.2	None
TYR	478	5.0	None
TYR	181	11.3	None
TYR	179	15.2	None
TYR	149	15.7	None
TYR	514	36.5	Low
TYR	20	37.5	None
TYR	54	39.4	Medium
TYR	375	40.3	None
TYR	60	42.2	Low
TYR	65	58.9	Medium
TYR	46	92.6	High
TYR	380	115.6	High
TYR	27	137.0	High
TYR	358	149.6	High

2.6 NMR and FT-IR spectra for 2a and 2b.

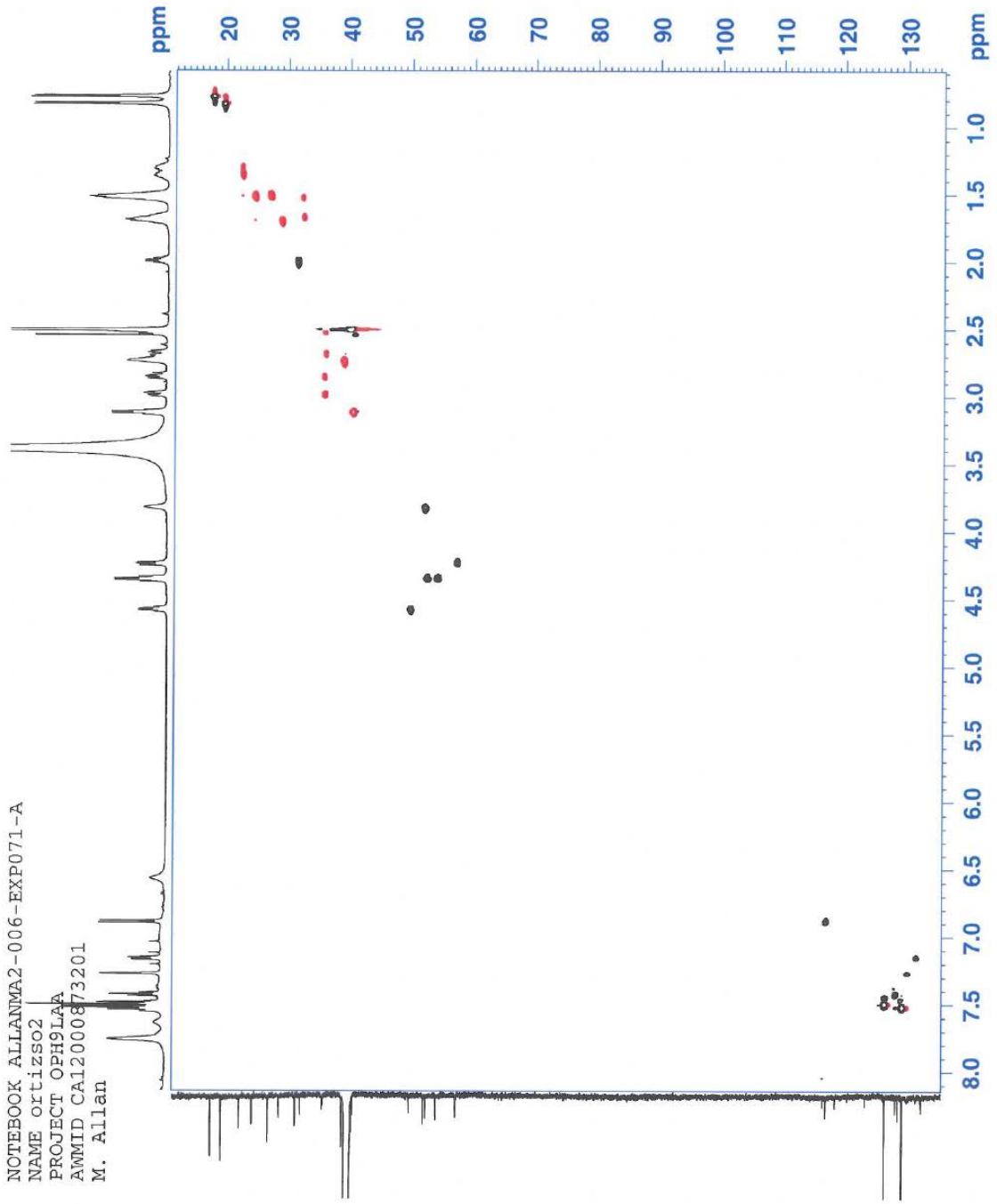
Tyrosine-mediated thymopentin conjugate (**2a**)



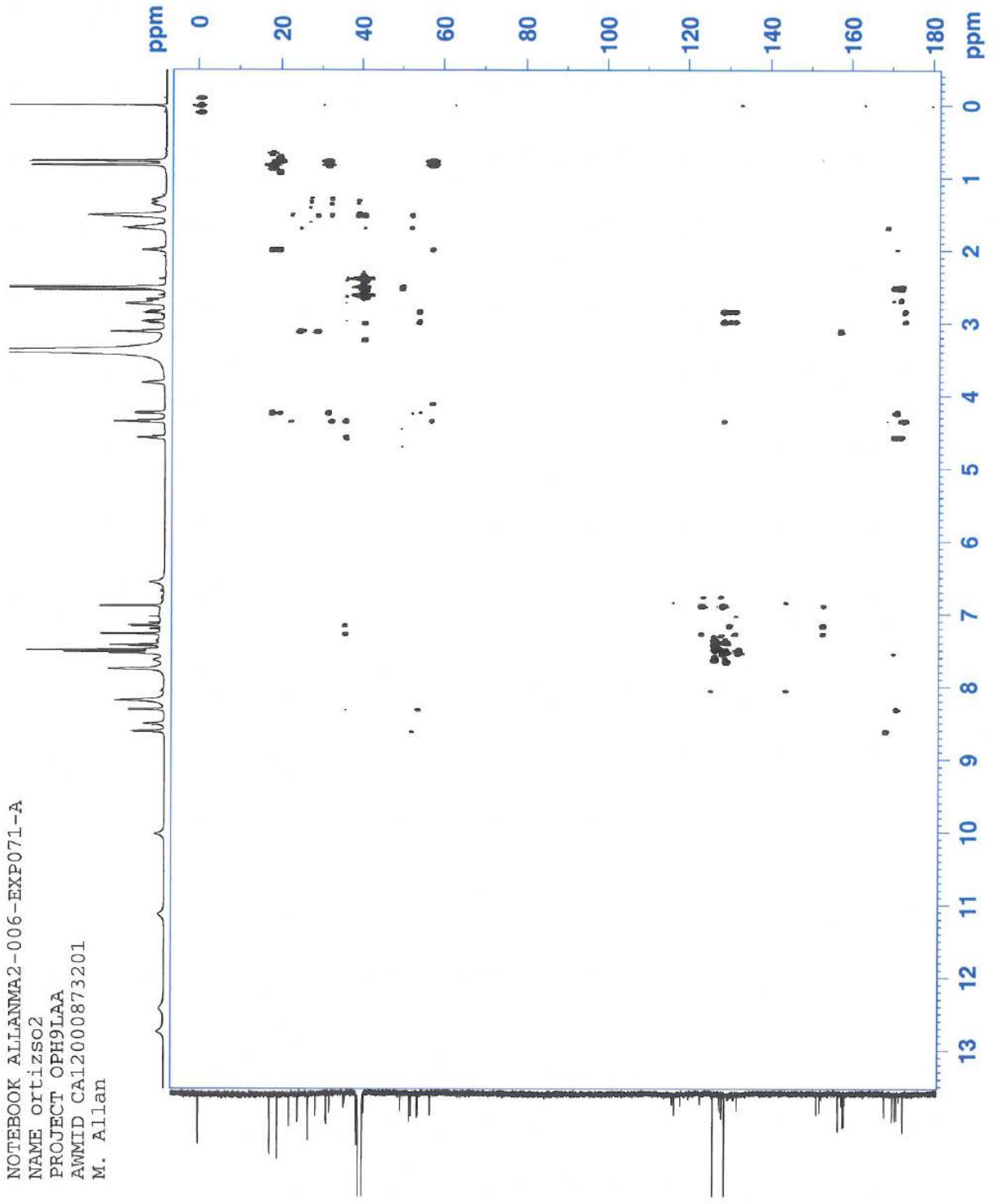

NOTEBOOK ALLANMA2-006-EXP071-A
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873201
M. Allan
173.18
172.24
171.62
171.22
170.65
168.68
158.80
158.59
158.38
158.18
157.19
153.07
152.66
151.91
132.39
129.39
128.73
128.33
126.60
123.49
118.68
117.16
116.70
57.33
54.23
52.58
52.22
49.96
40.64 DMSO
40.50 DMSO
40.38 DMSO
40.25 DMSO
40.11 DMSO
39.97 DMSO
39.83 DMSO
39.69 DMSO
39.55 DMSO
39.11
27.18
22.57
19.65
17.89
0.59
f1 (ppm)

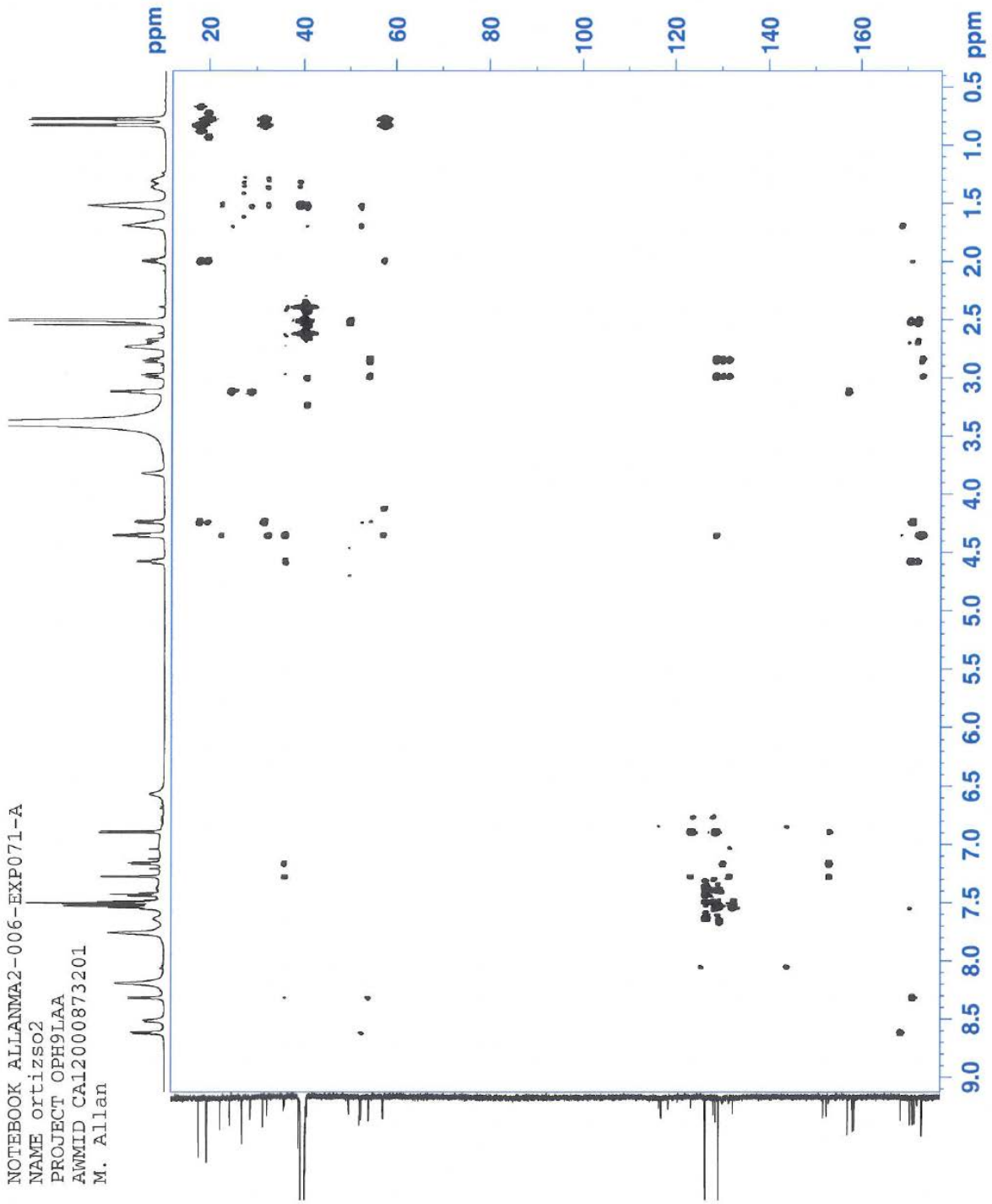


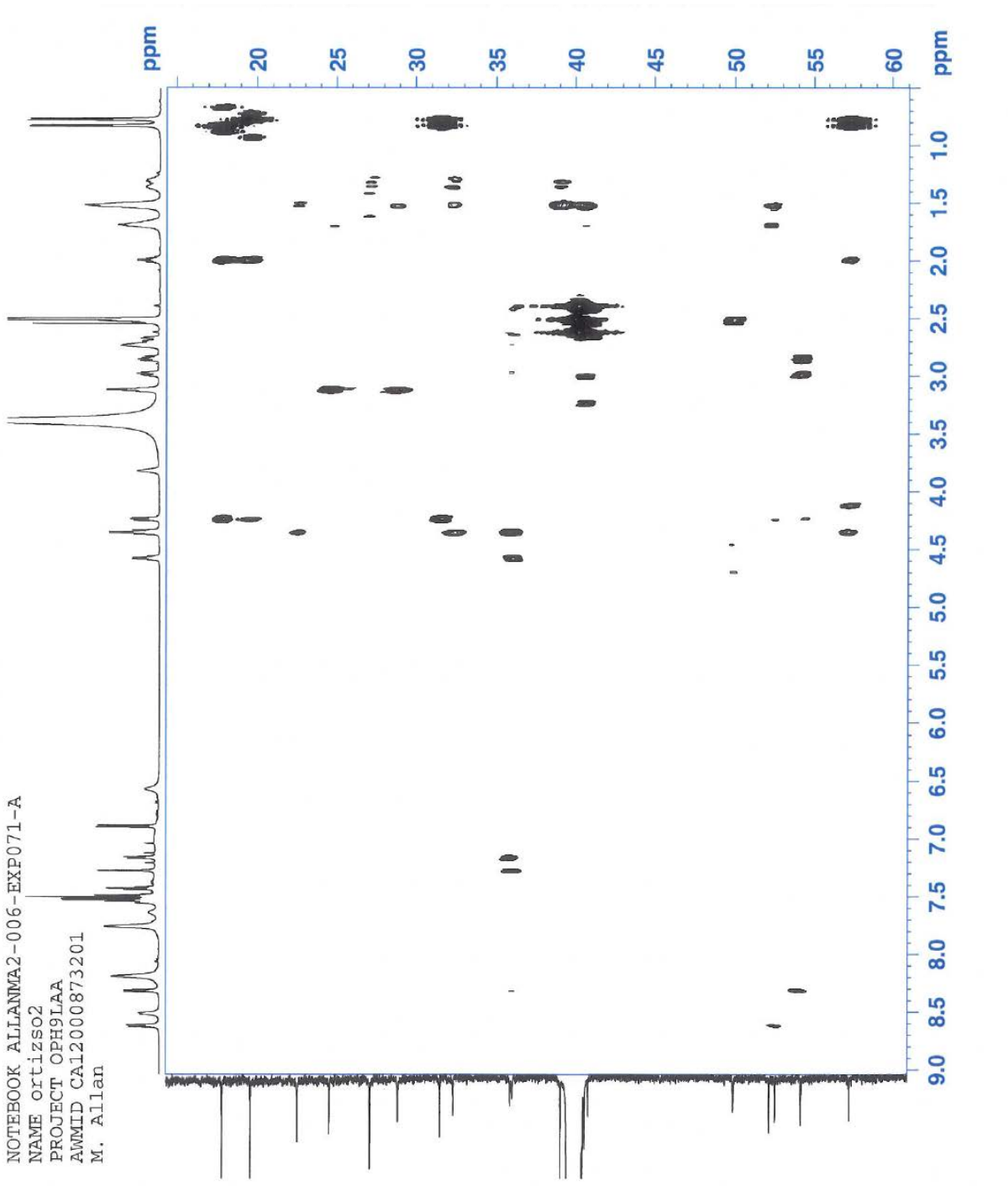
HSQC:

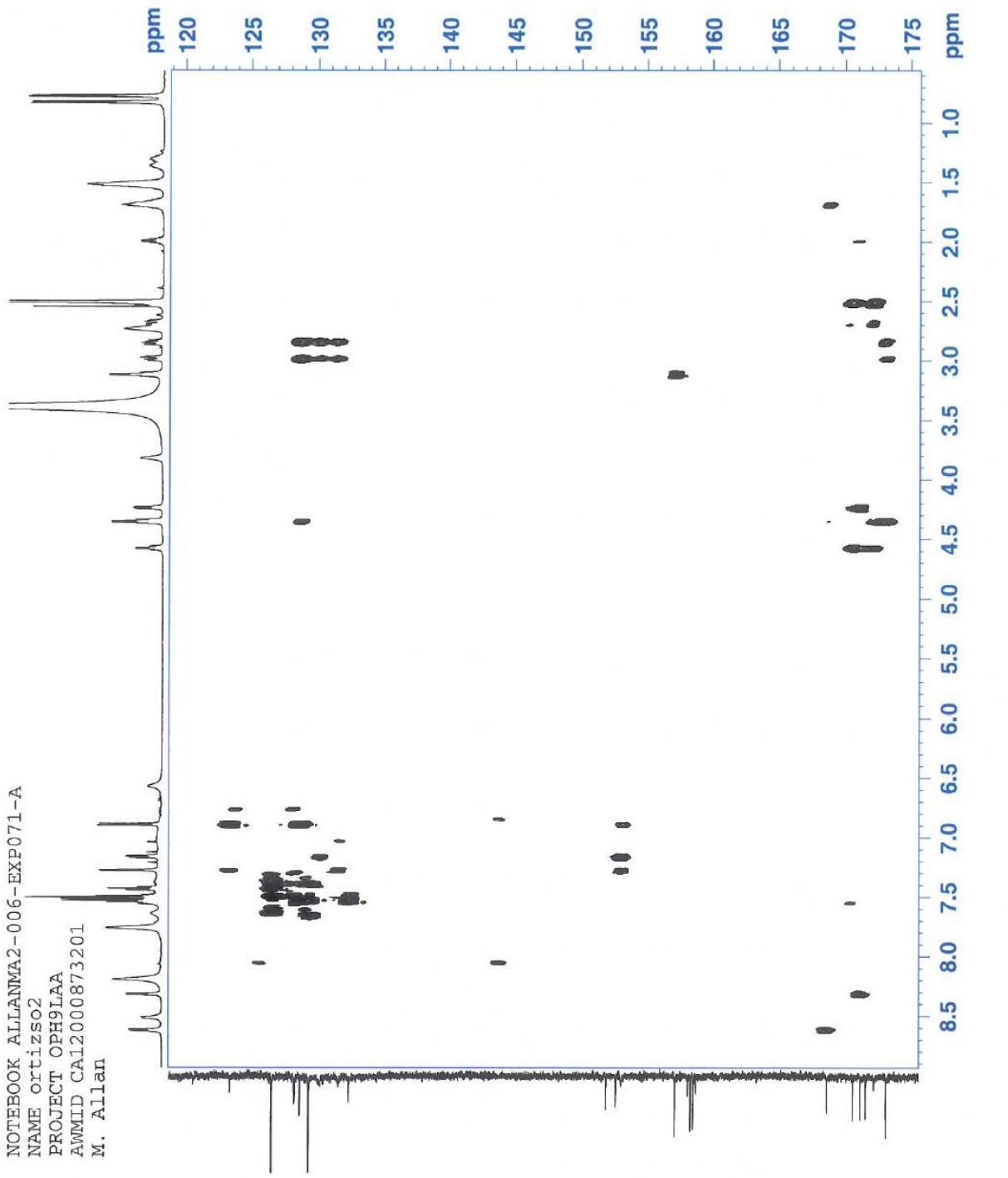

ppm
20
30
40
50
60
70
80
90
100
110
120
130
ppm
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
NOTEBOOK ALLANMA2-006-EXP071-A
NAME ortizso2
PROJECT OPH9LAB
AWMID CA120008773201
M. Allan

HMBC:

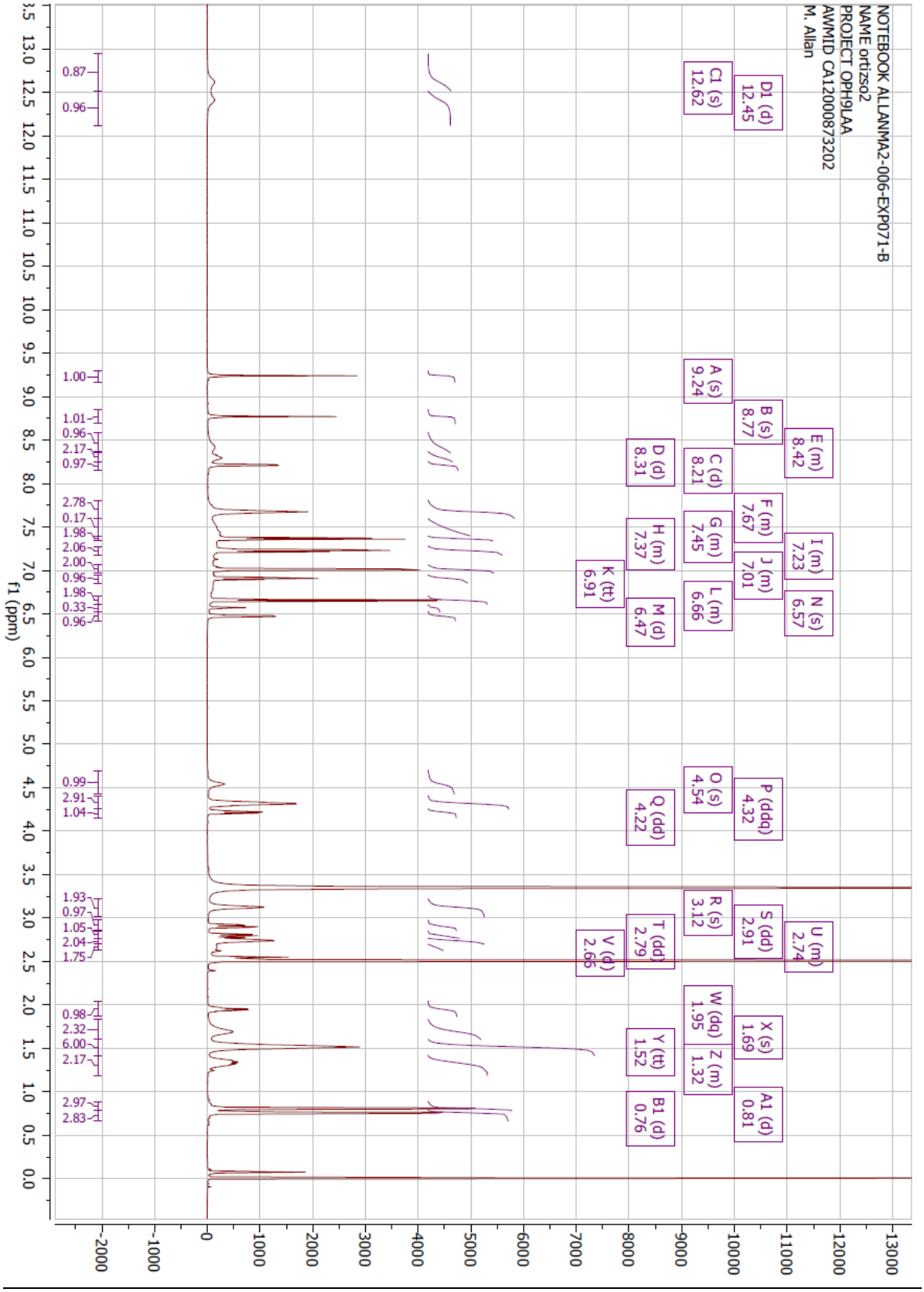

NOTEBOOK ALLANMA2-006-EXP071-A
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873201
M. Allan
ppm
0
20
40
60
80
100
120
140
160
180
ppm
0
1
2
3
4
5
6
7
8
9
10
11
12
13

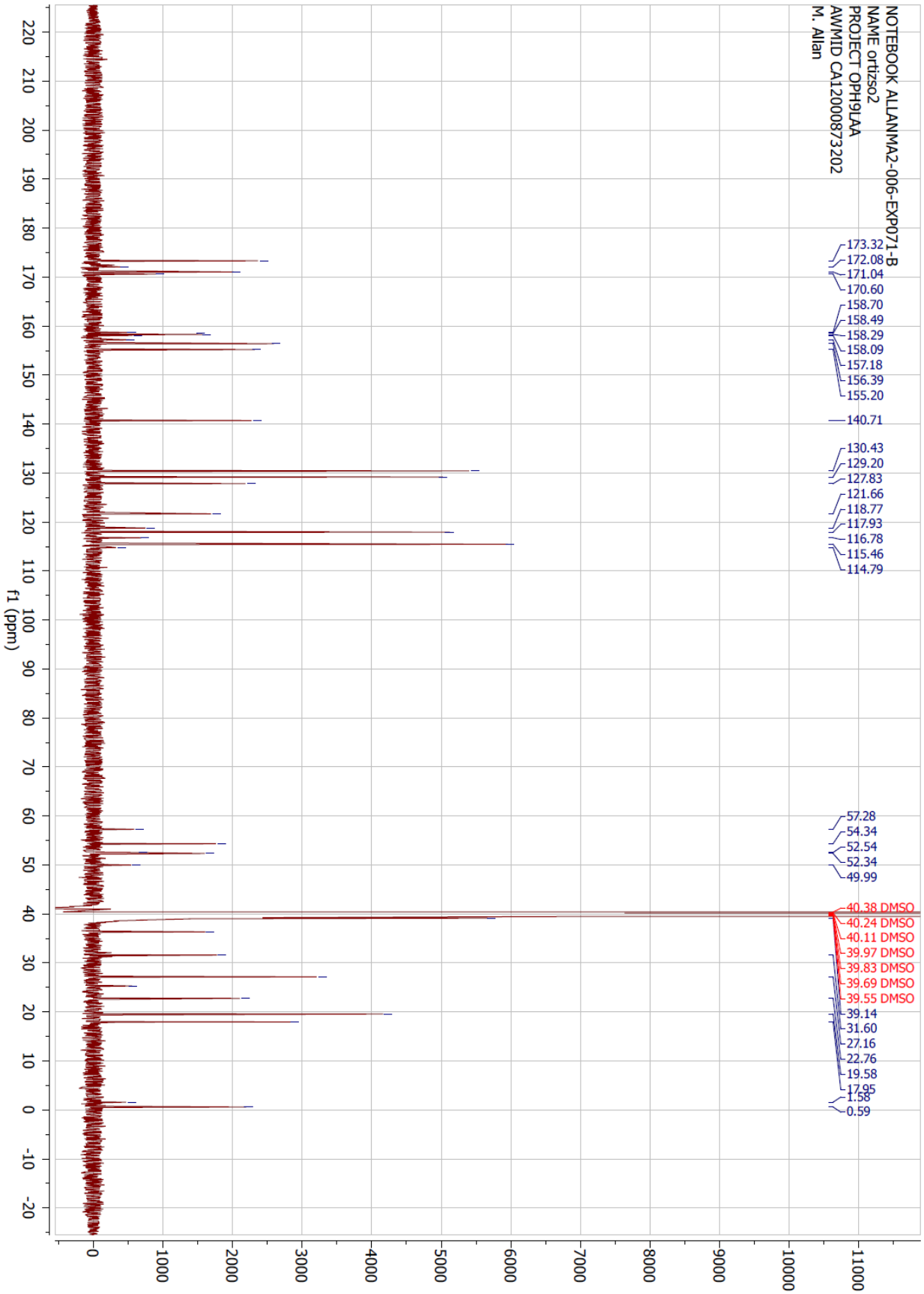

NOTEBOOK ALLANMA2-006-EXP071-A
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873201
M. Allan
ppm
20
40
60
80
100
120
140
160
ppm
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0

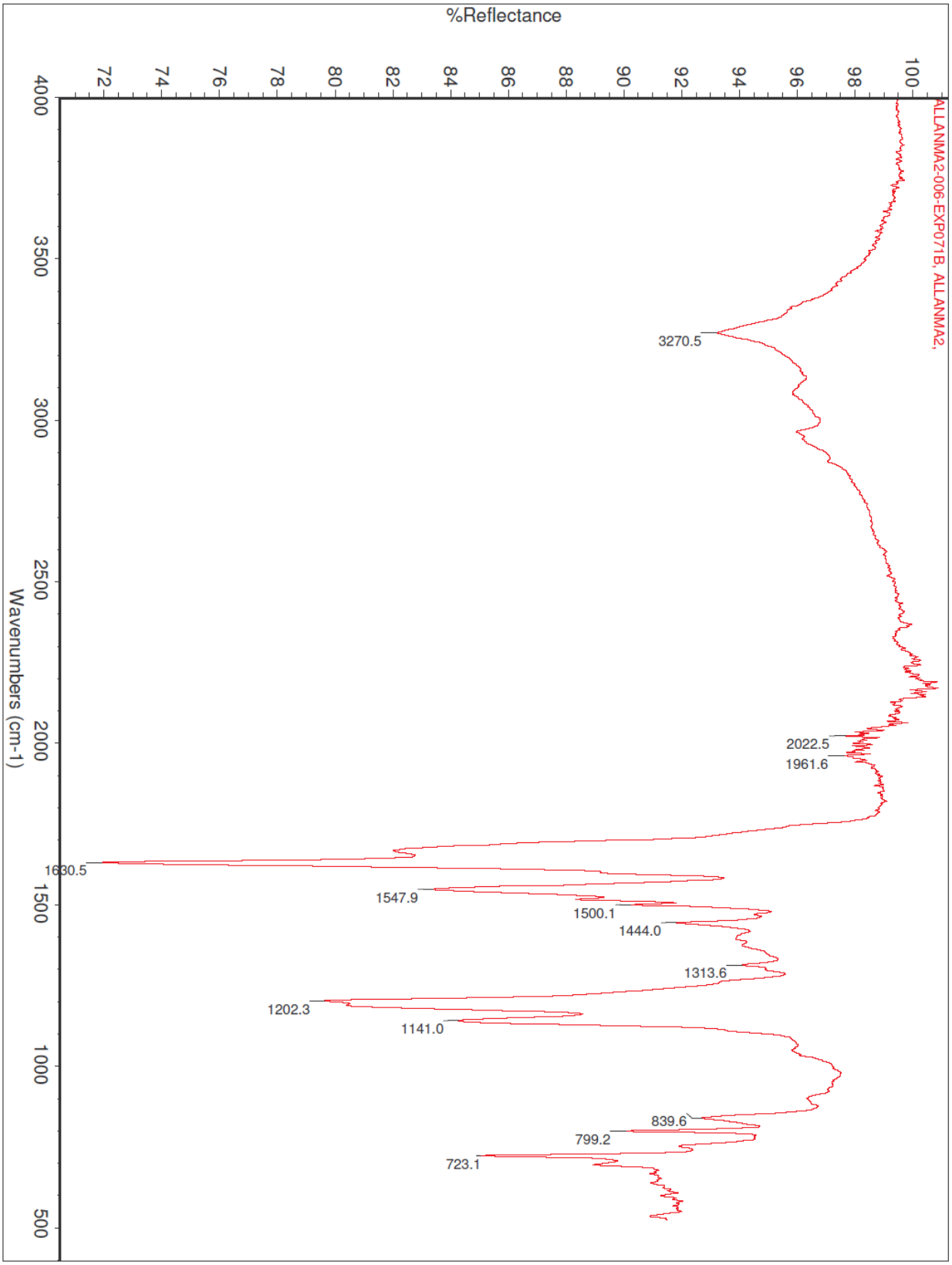

NOTEBOOK ALLANMA2-006-EXP071-A
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873201
M. Allan
ppm
20
25
30
35
40
45
50
55
60
ppm
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0


NOTEBOOK ALLANMA2-006-EXP071-A
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873201
M. Allan
ppm
120
125
130
135
140
145
150
155
160
165
170
175
ppm
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5

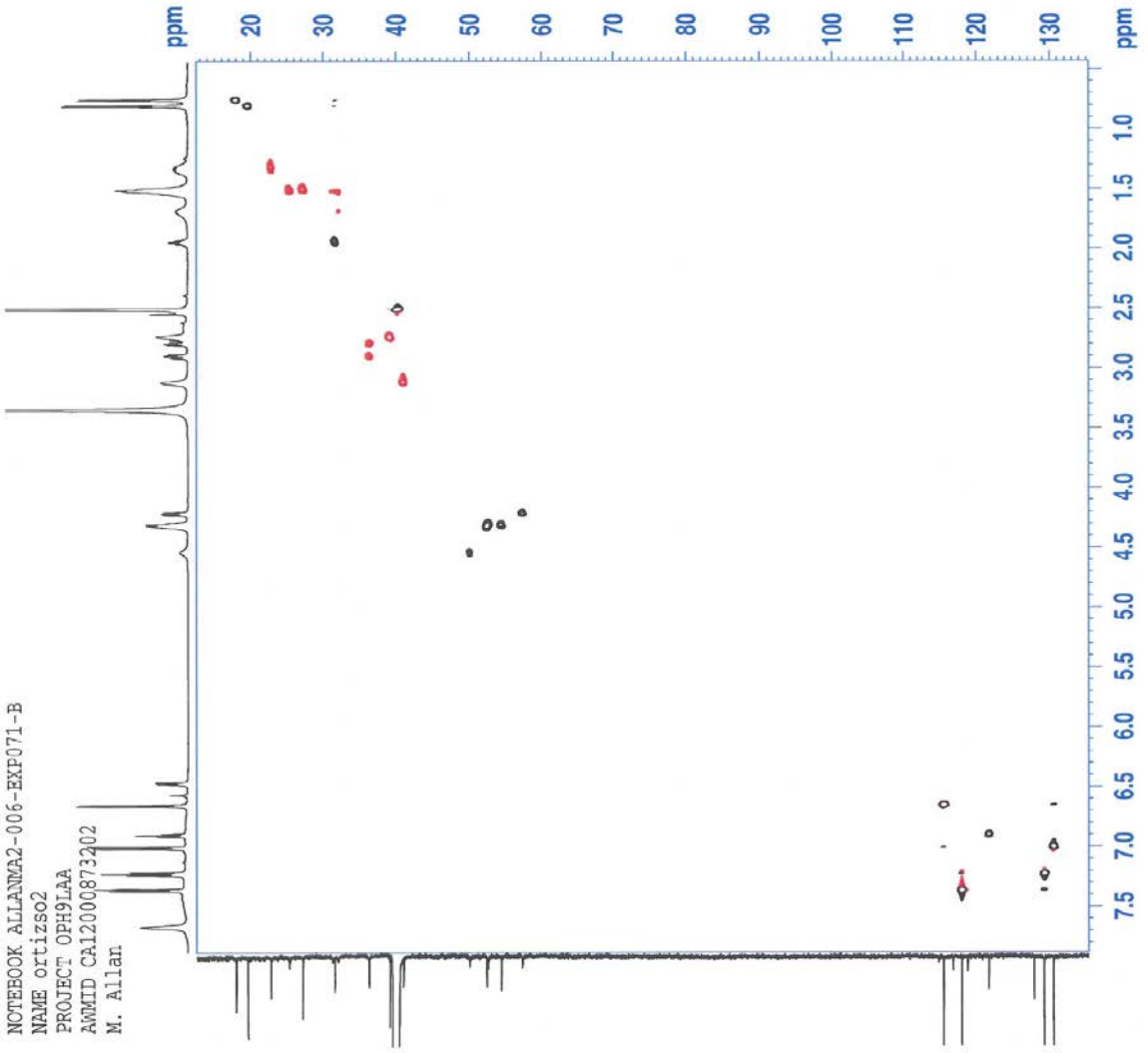
Lysine-mediated thymopentin conjugate (**2b**)



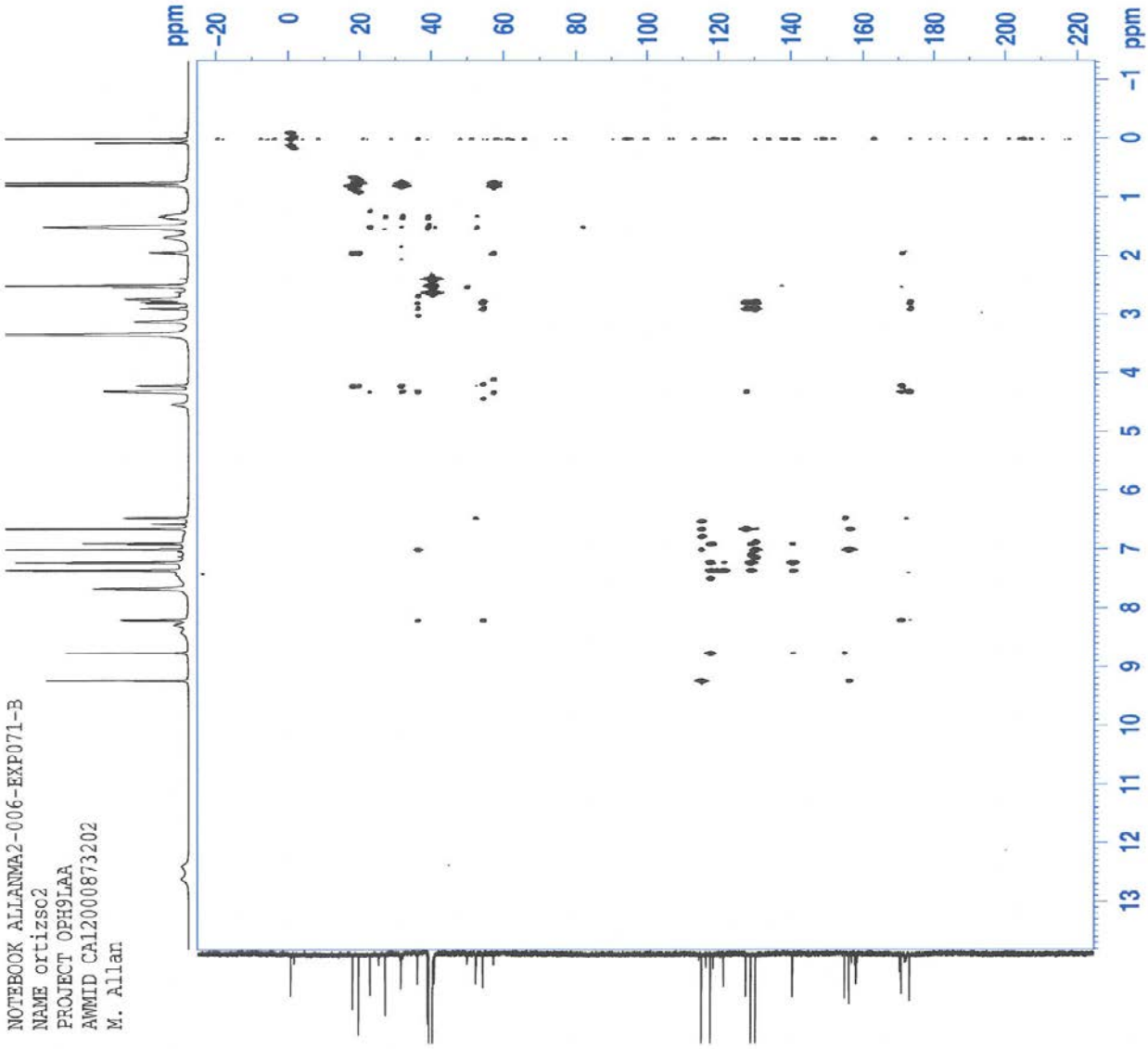

NOTEBOOK ALLANMA2-006-EXP071-B
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873202
M. Allan
173.32
172.08
171.04
170.60
158.70
158.49
158.29
158.09
157.18
156.39
155.20
140.71
130.43
129.20
127.83
121.66
118.77
117.93
116.78
115.46
114.79
57.28
54.34
52.54
52.34
49.99
40.38 DMSO
40.24 DMSO
40.11 DMSO
39.97 DMSO
39.83 DMSO
39.69 DMSO
39.55 DMSO
39.14
31.60
27.16
22.76
19.58
17.95
1.58
0.59
f1 (ppm)

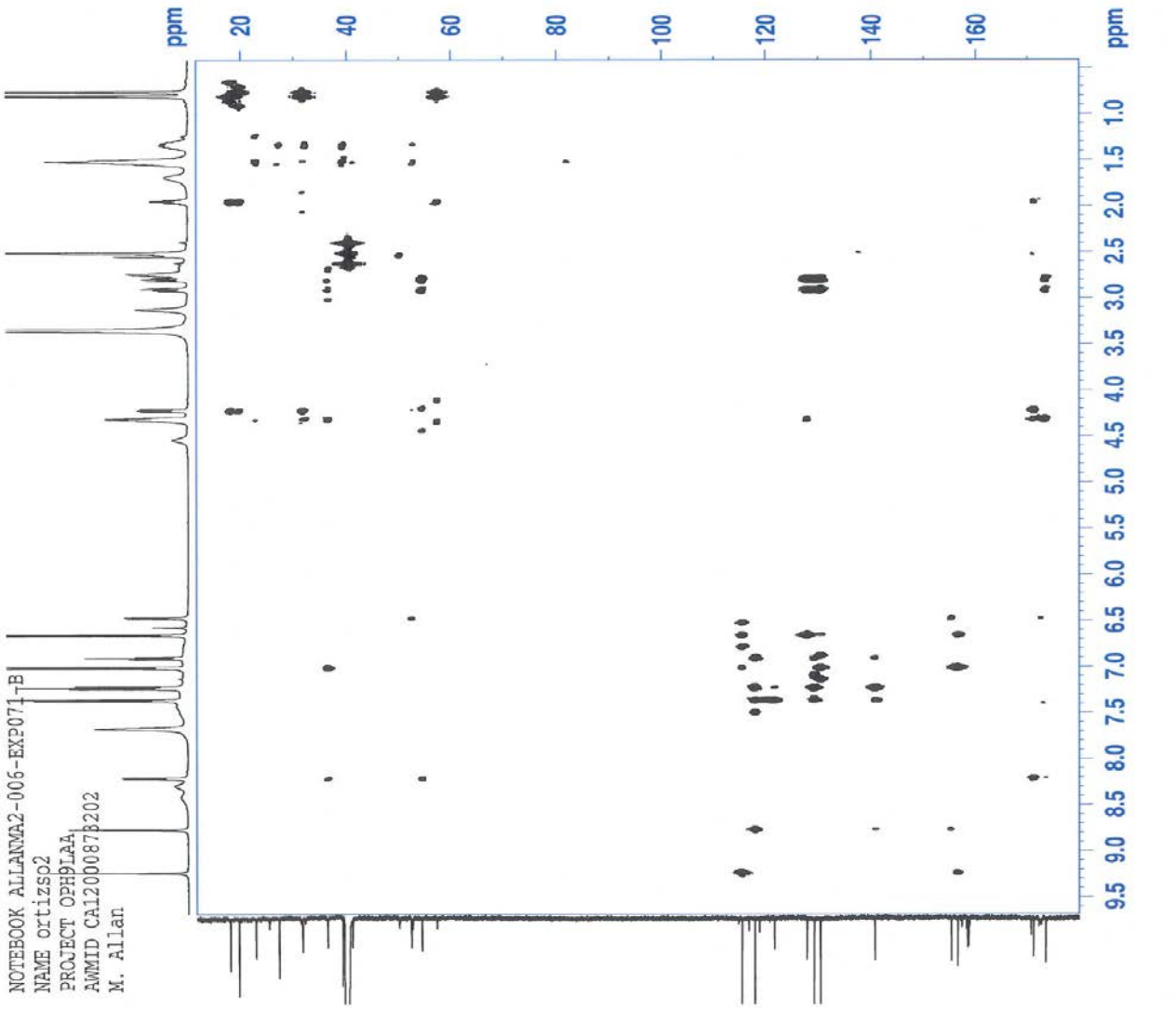


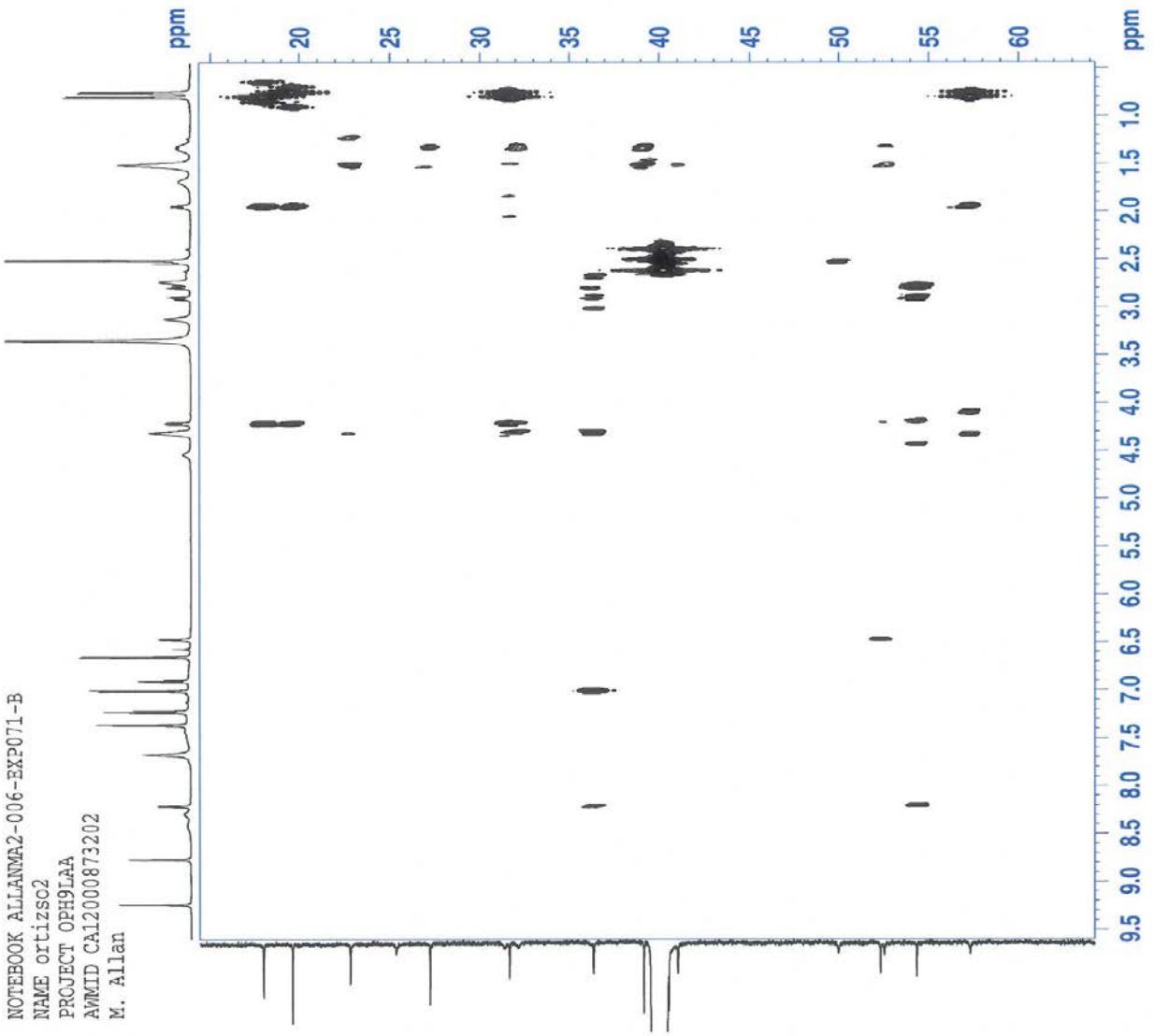
HSQC:

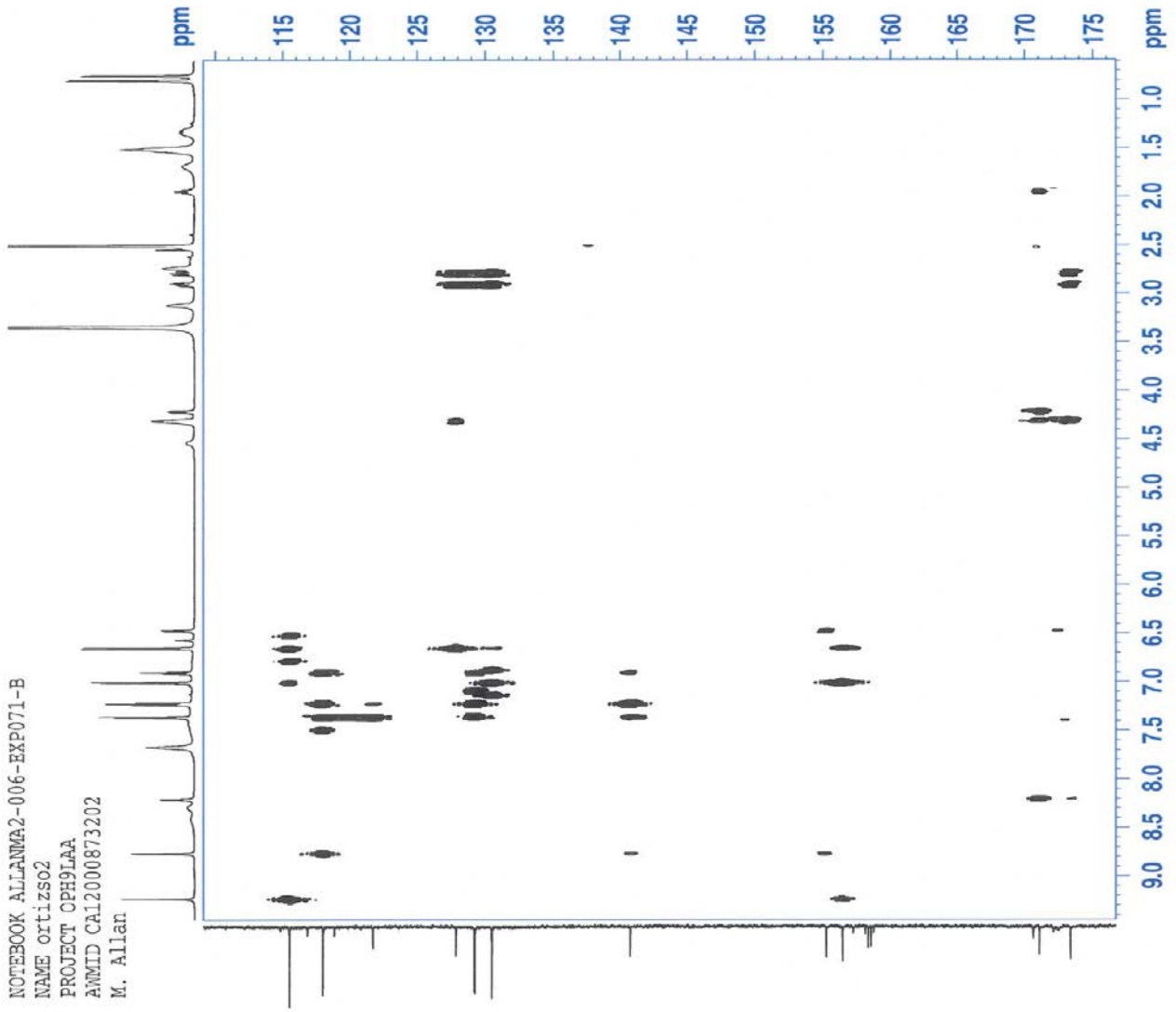

ppm
20
30
40
50
60
70
80
90
100
110
120
130
ppm
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
NOTEBOOK ALLANMA2-006-EXP071-B
NAME ortizso2
PROJECT OPH9LAA
AMMID CA12000873202
M. Allan

HMBC:



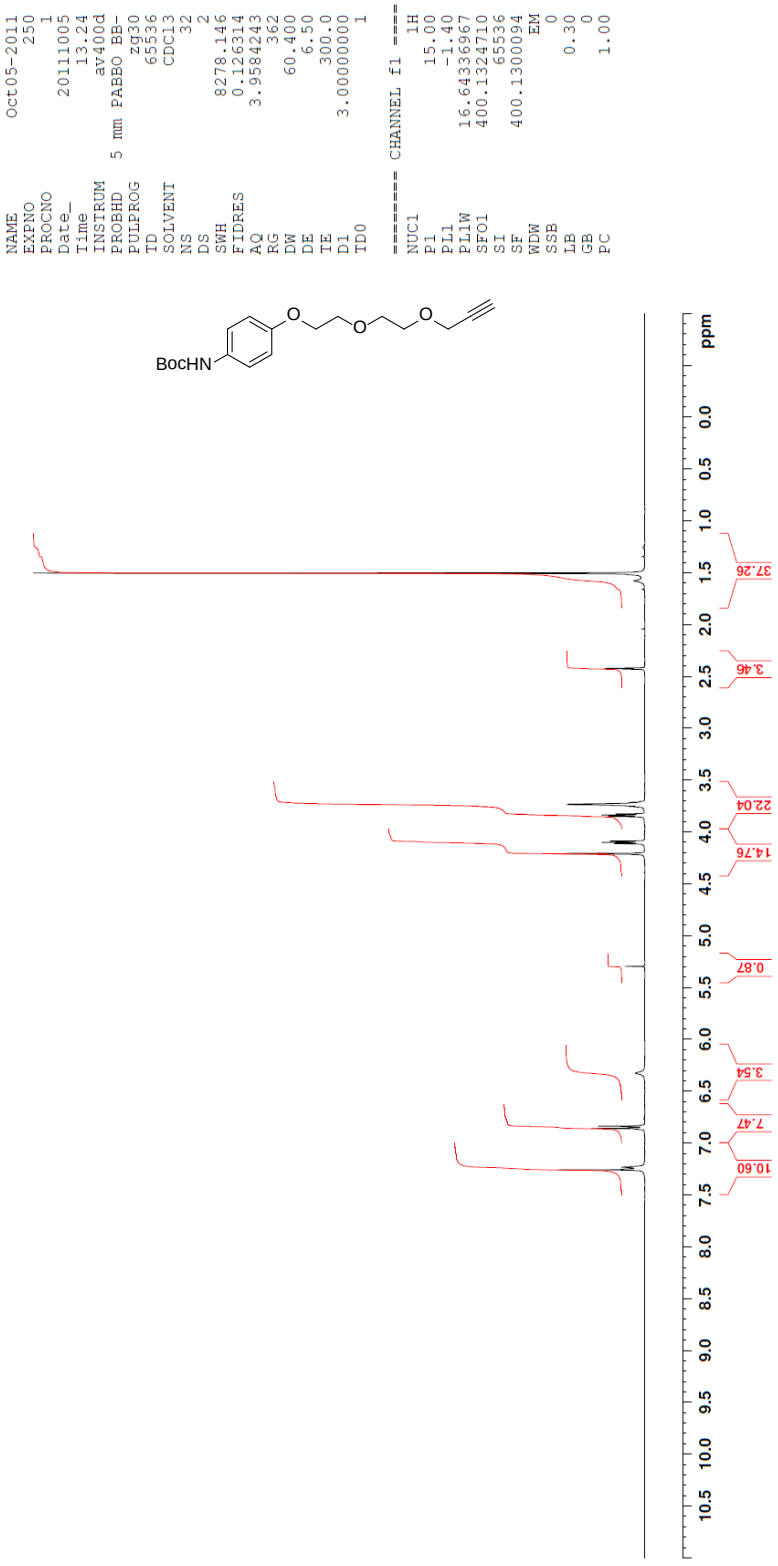

NOTEBOOK ALLANMA2-006-EXP071-B
NAME ortizso2
PROJECT OPHBLAA
AWMID CA120D087B202
M. Allan
ppm
20
40
60
80
100
120
140
160
ppm
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

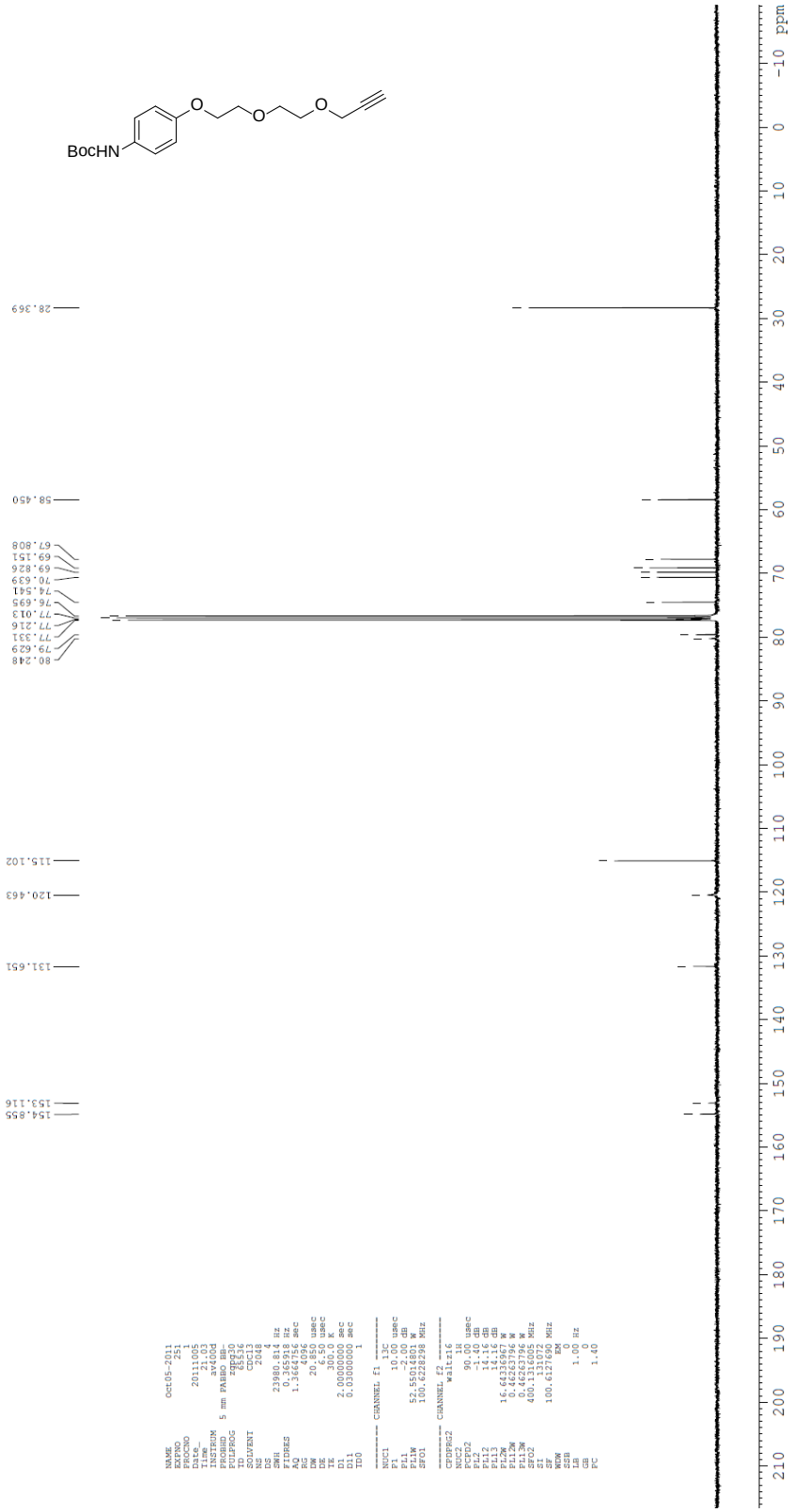

NOTEBOOK ALLANMA2-006-EXP071-B
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873202
M. Allan
ppm
20
25
30
35
40
45
50
55
60
ppm
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5

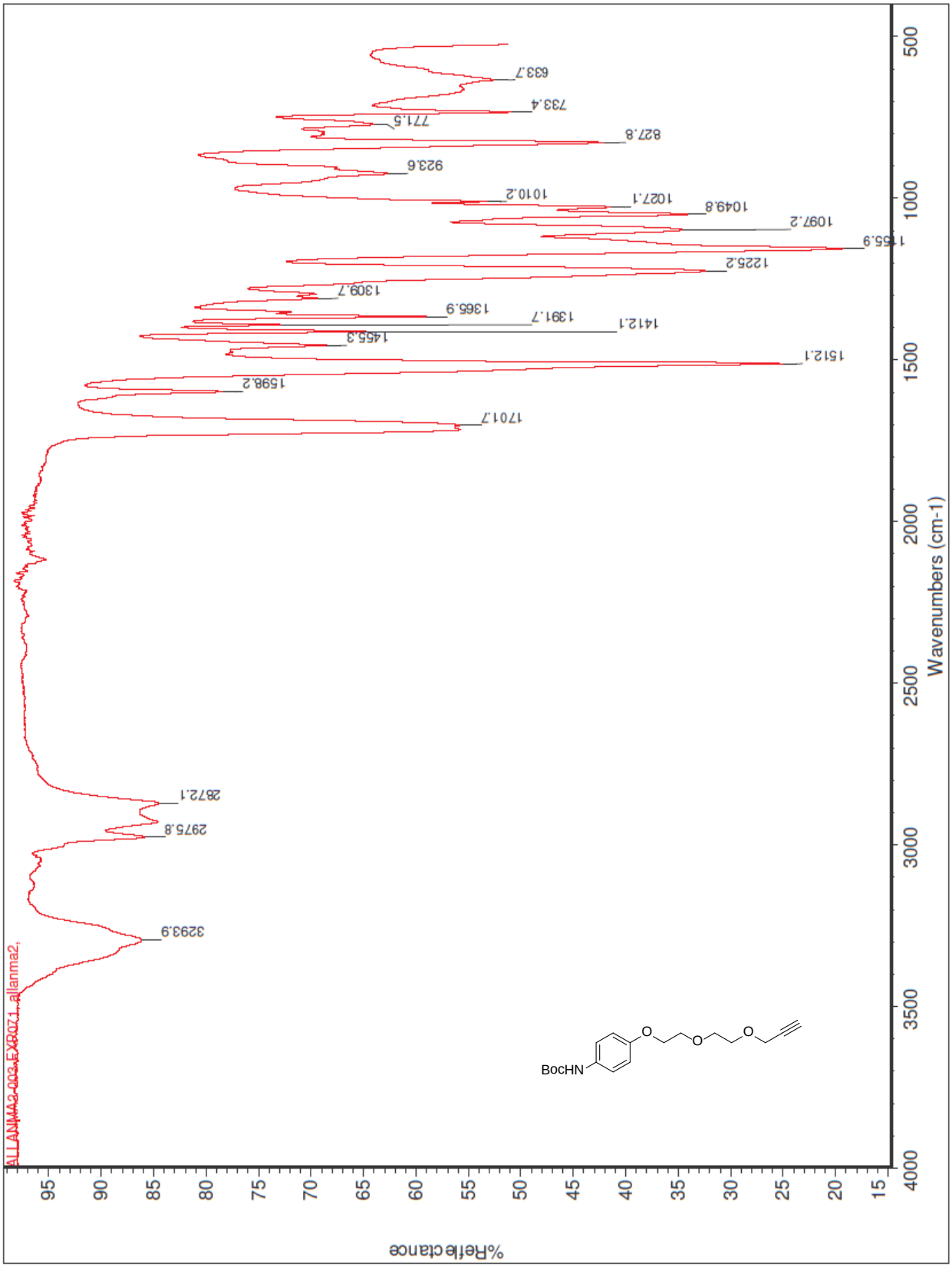

NOTEBOOK ALLANMA2-006-EXP071-B
NAME ortizso2
PROJECT OPH9LAA
AWMID CA12000873202
M. Allan
ppm
115
120
125
130
135
140
145
150
155
160
165
170
175
ppm
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0

2.7 NMR and IR spectra of linkers for modified CRM_{197}

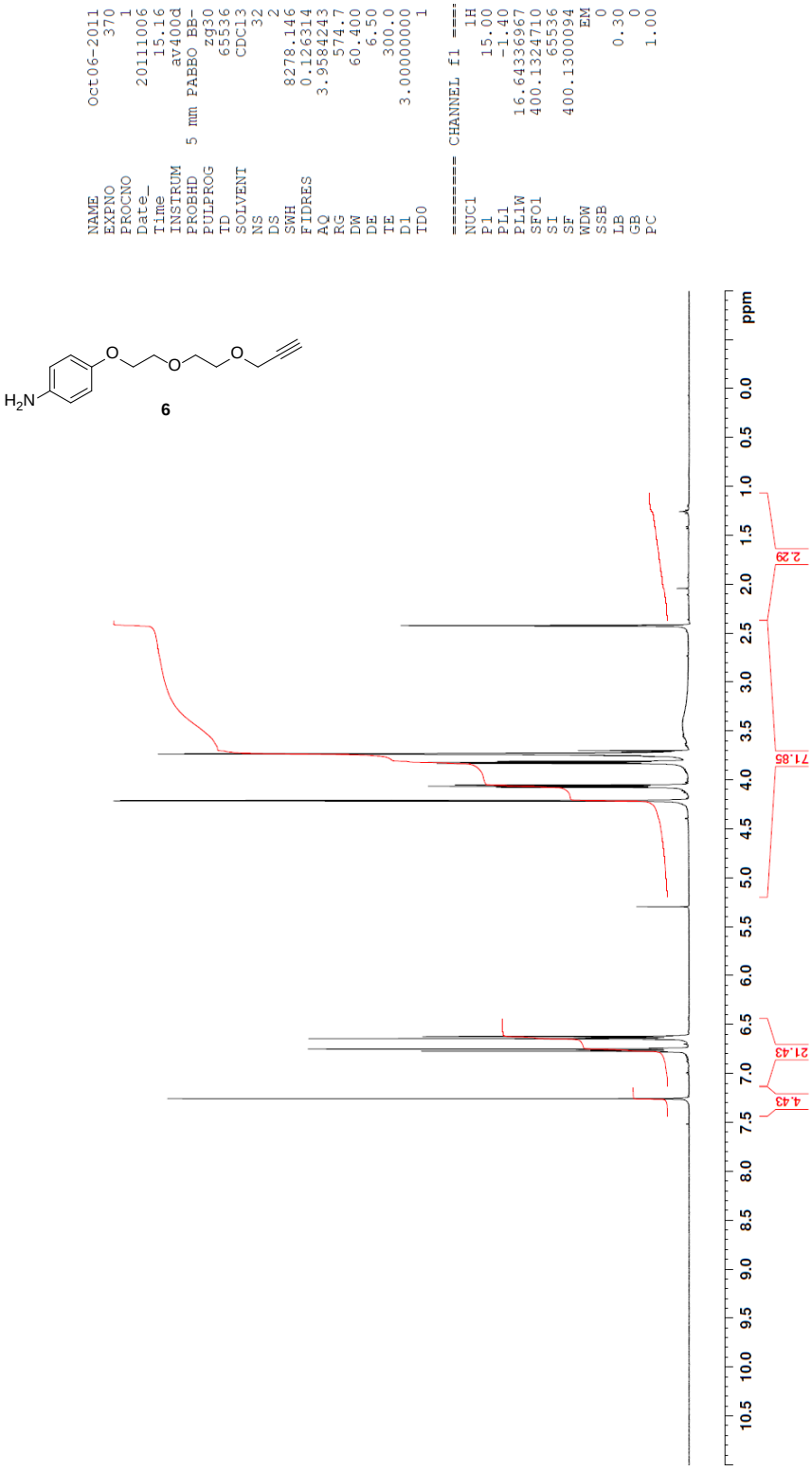
tert-butyl 4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamate (**5**)

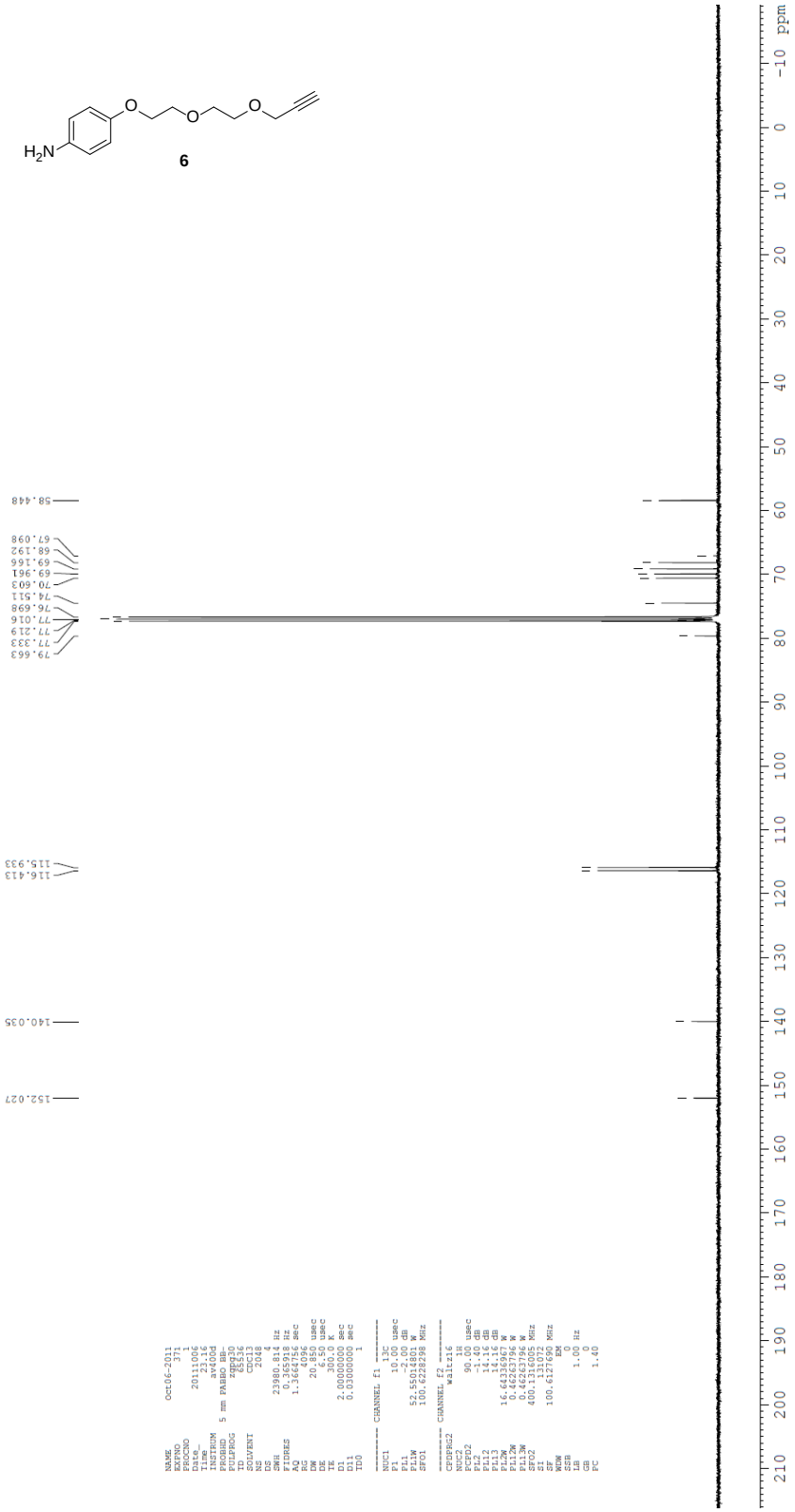


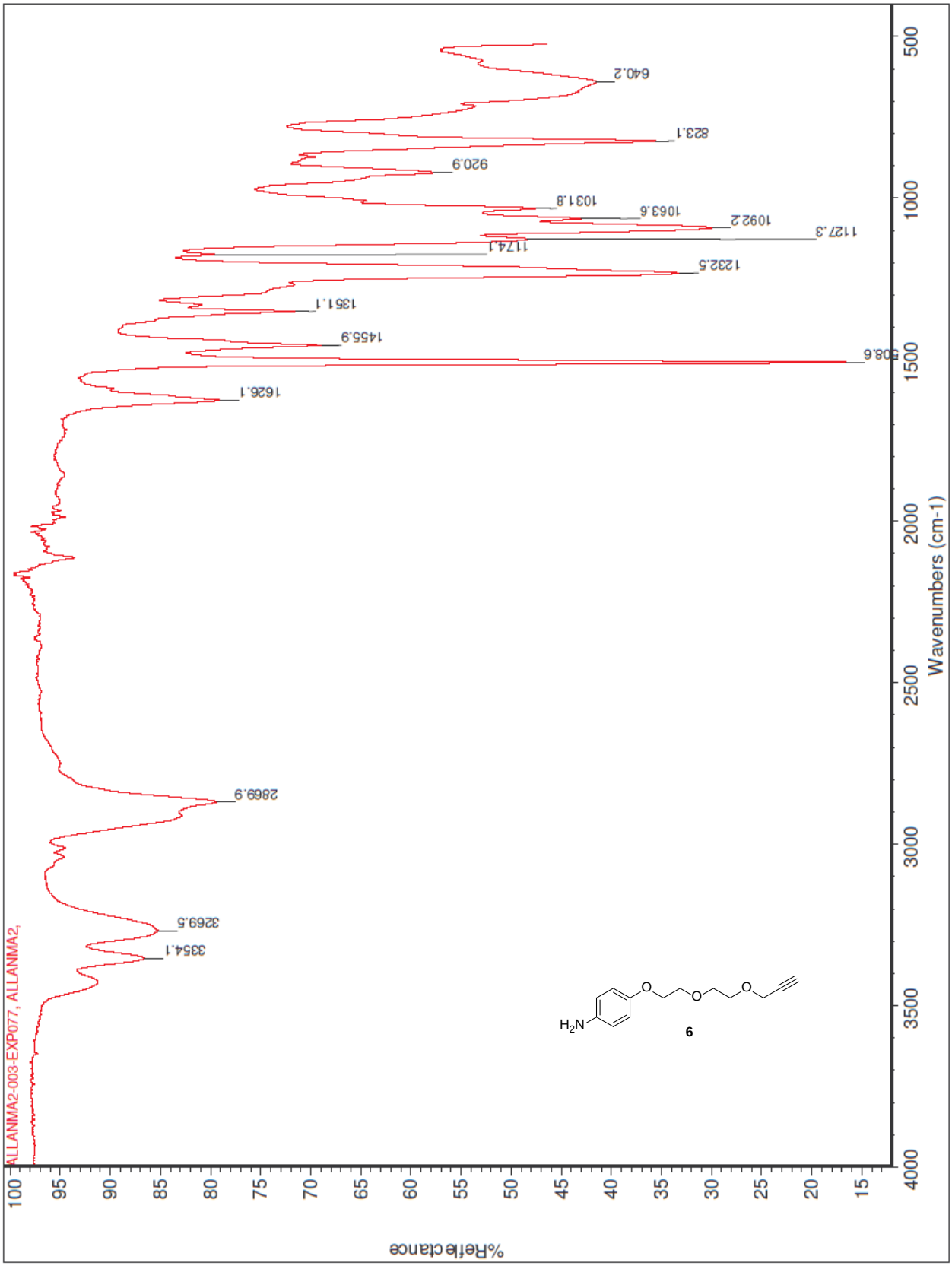

BocHN
O
O
O
28.369
58.450
67.808
69.151
69.826
70.639
74.541
76.695
77.013
77.216
77.331
79.629
80.248
115.102
120.463
131.651
153.116
154.855
ppm


%Reflectance
Wavenumbers (cm-1)
4000
3500
3000
2500
2000
1500
1000
500
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
3293.9
2975.8
2872.1
1701.7
1598.2
1512.1
1455.3
1412.1
1391.7
1365.9
1309.7
1225.2
1155.9
1097.2
1049.8
1027.1
1010.2
923.6
827.8
771.5
733.4
633.7
BocHN
O
O
O

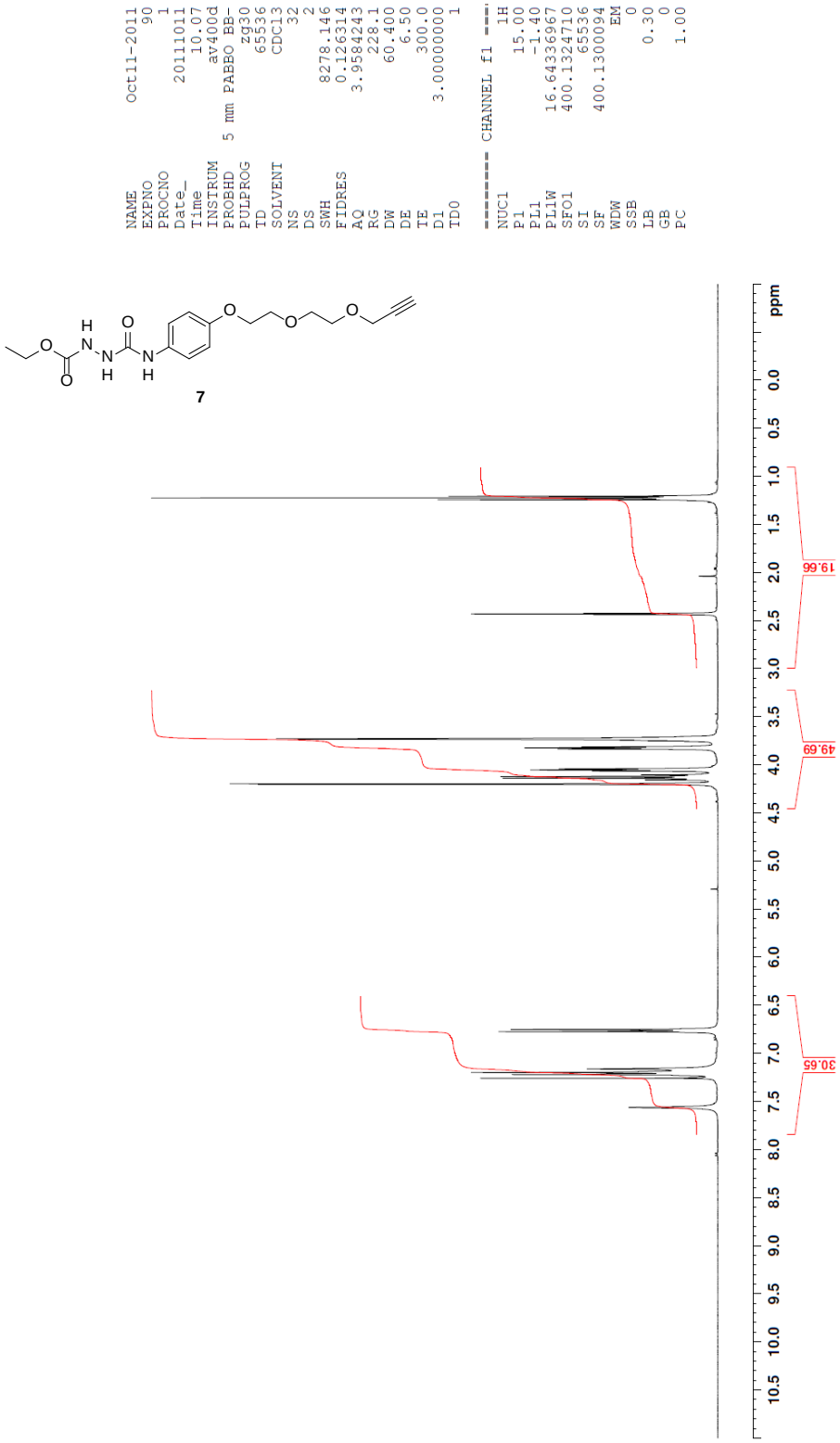
4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)aniline (**6**)

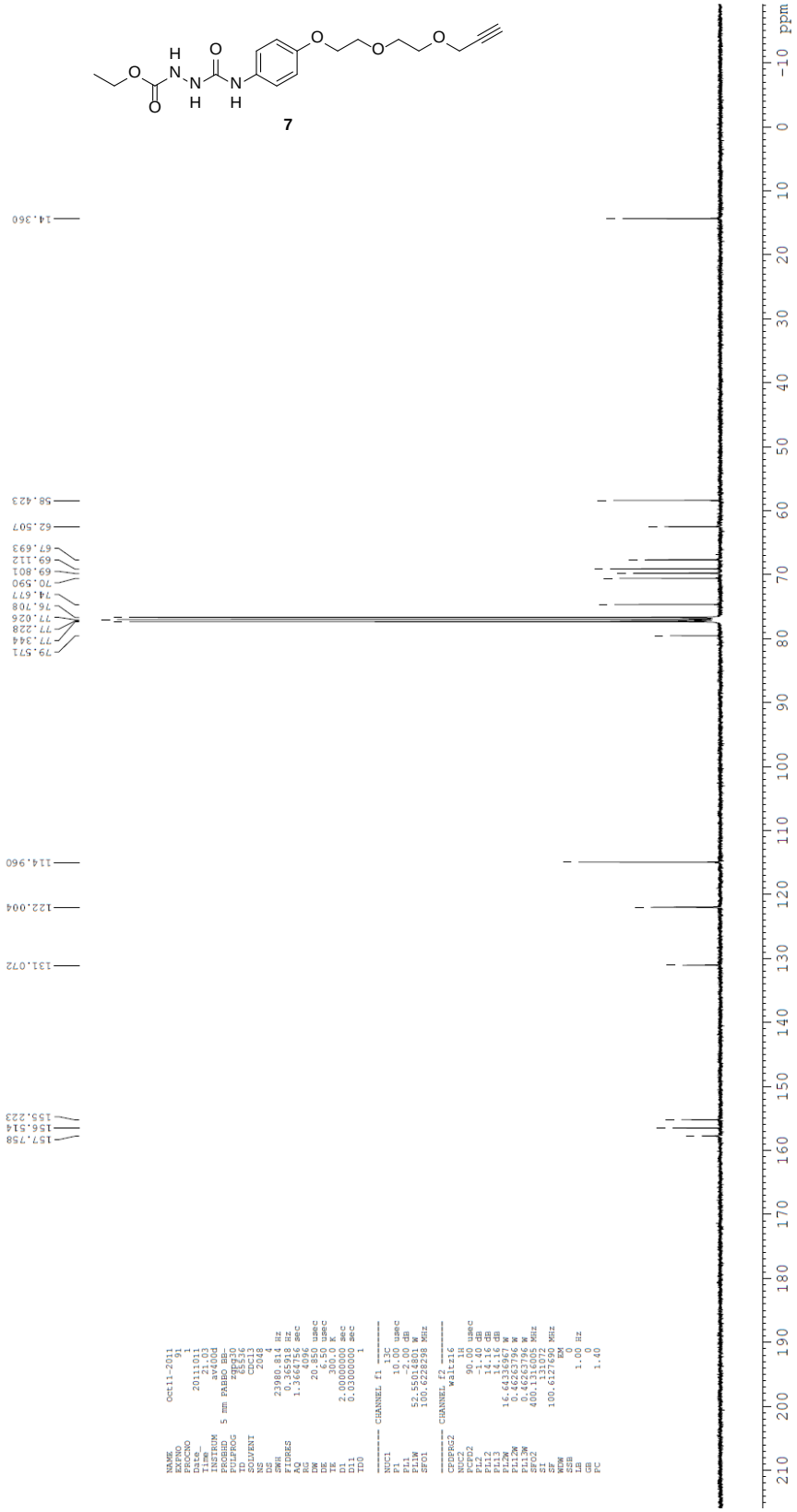


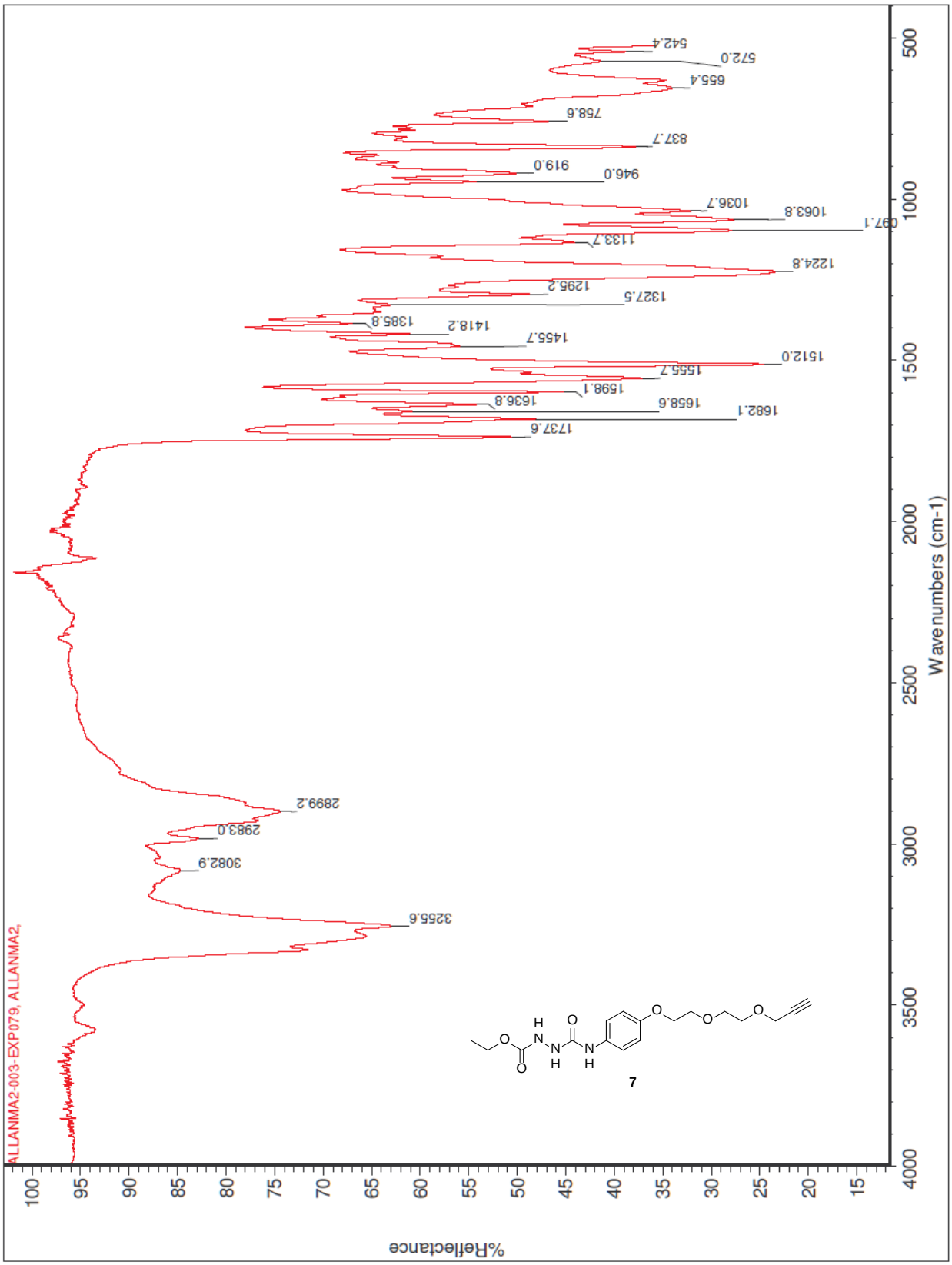

H2N
O
O
O
6
ppm
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
58.448
67.098
68.192
69.166
69.961
70.603
74.511
76.698
77.016
77.219
77.333
79.663
116.413
115.933
140.035
152.027


ALLANMA2-003-EXP077, ALLANMA2,
%Reflectance
Wavenumbers (cm-1)
640.2
823.1
920.9
1031.8
1063.6
1092.2
1127.3
1174.1
1232.5
1351.1
1455.9
1508.6
1626.1
2869.9
3269.5
3354.1
H_2N
O
O
O
6

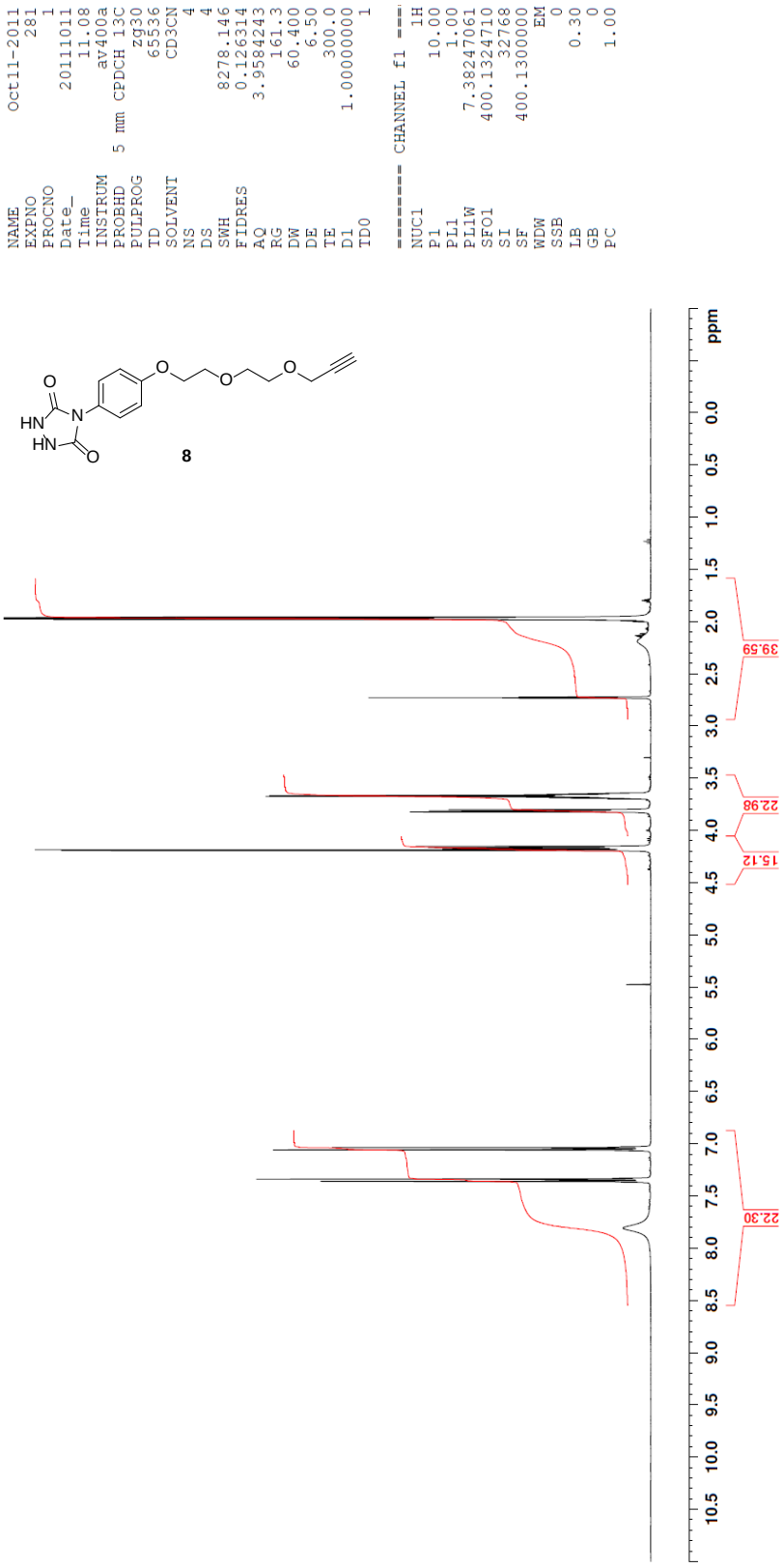
Ethyl 2-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenylcarbamoyl)hydrazinecarboxylate (**7**)

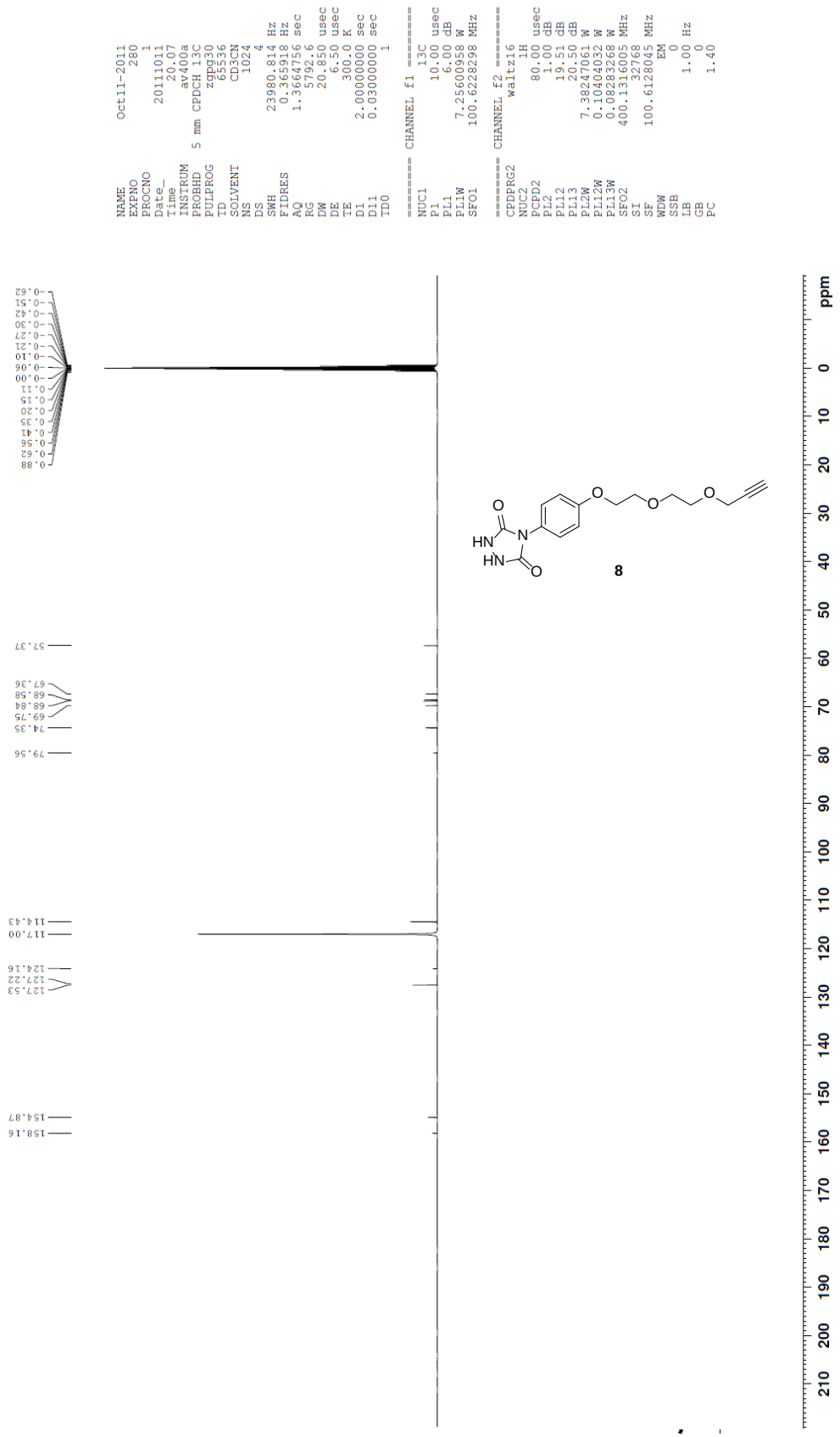


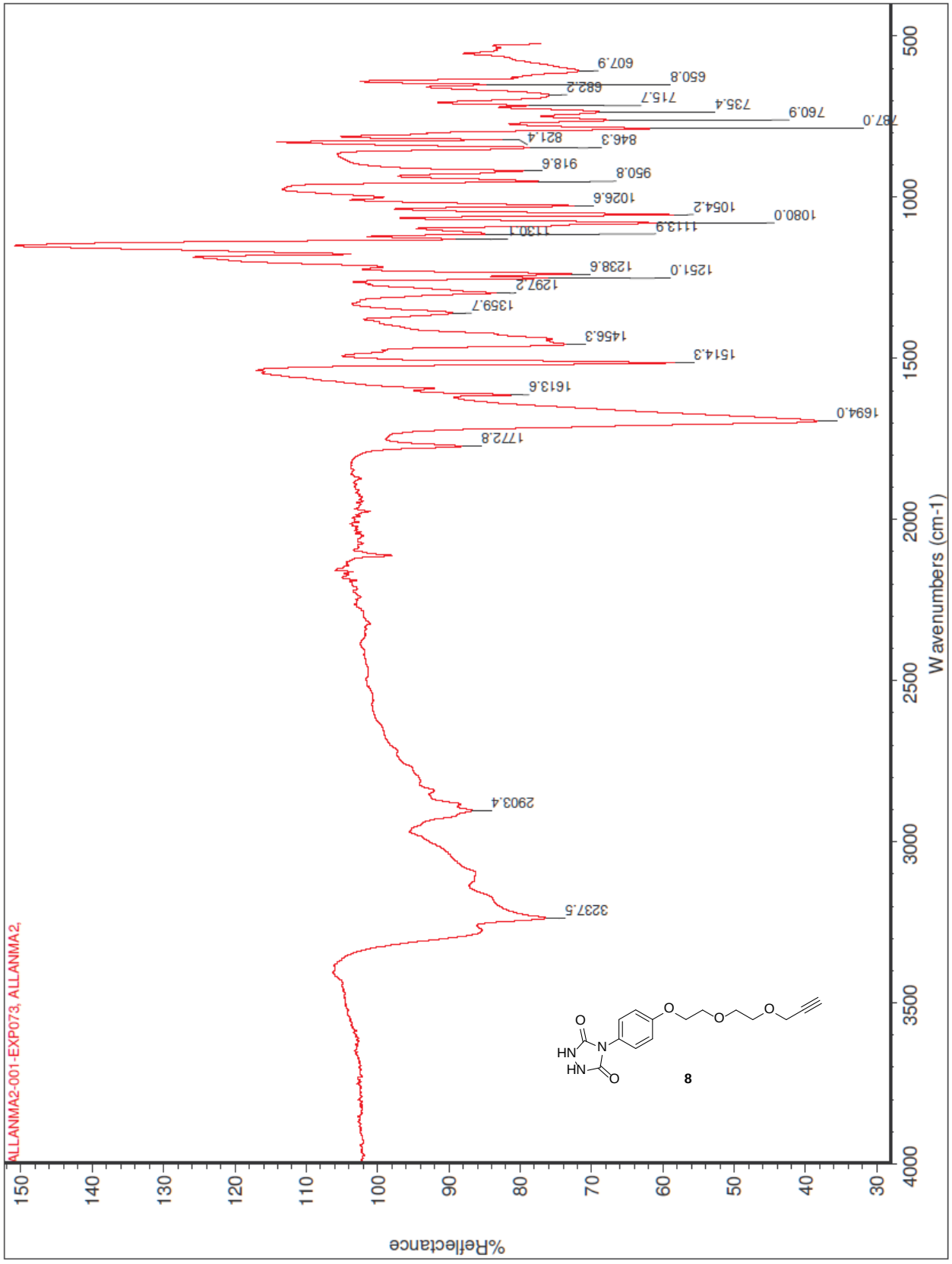

7
14.360
58.423
62.507
67.693
69.112
69.801
70.590
74.677
76.708
77.026
77.228
77.344
79.571
114.960
122.004
131.072
155.223
156.514
157.758
ppm


ALLANMA2-003-EXP079, ALLANMA2,
%Reflectance
Wavenumbers (cm-1)
542.4
572.0
655.4
758.6
837.7
919.0
946.0
1036.7
1063.8
1097.1
1133.7
1224.8
1295.2
1327.5
1385.8
1418.2
1455.7
1512.0
1555.7
1598.1
1636.8
1658.6
1682.1
1737.6
2899.2
2983.0
3082.9
3255.6
7

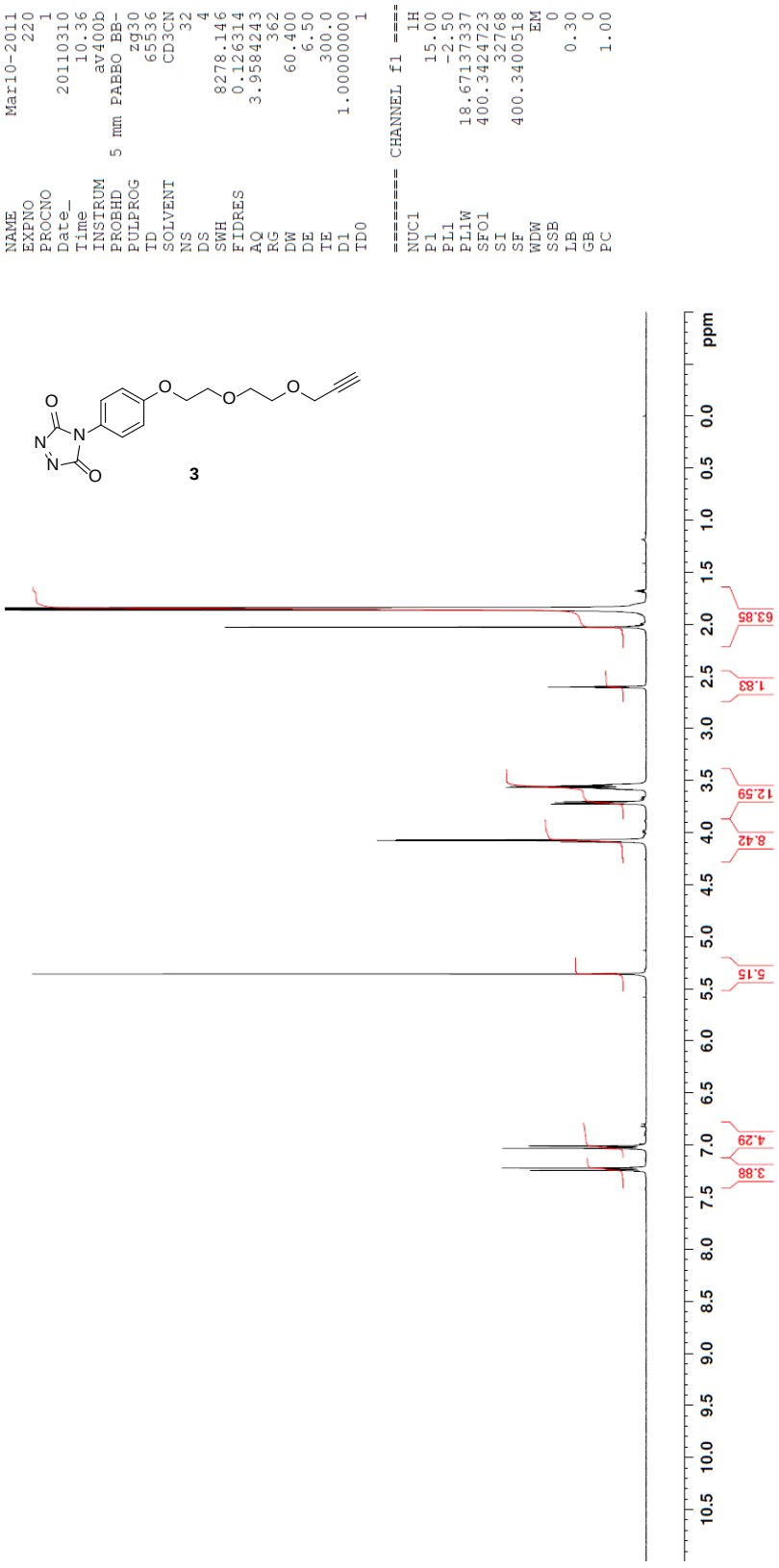
4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-1,2,4-triazolidine-3,5-dione (**8**)




NAME
EXPNO
PROCNO
Date_
Time
INSTRUM
PROBHD
PULPROG
TD
SOLVENT
NS
DS
SWH
FIDRES
AQ
RG
DW
DE
TE
D1
D11
TD0
Oct11-2011
280
1
20111011
20.07
av400a
5 mm CPDCH 13C
zgpg30
65536
CD3CN
1024
4
23980.814 Hz
0.365918 Hz
1.3664756 sec
5792.6
20.850 usec
6.50 usec
300.0 K
2.00000000 sec
0.03000000 sec
1
CHANNEL f1
13C
NUC1
P1
PL1
PL1W
SFO1
10.00 usec
6.00 dB
7.25600958 W
100.6228298 MHz
CHANNEL f2
waltz16
1H
CPDPRG2
NUC2
PCPD2
PL2
PL12
PL13
PL2W
PL12W
PL13W
SFO2
SI
SF
WDW
SSB
LB
GB
PC
80.00 usec
1.00 dB
19.51 dB
20.50 dB
7.38247061 W
0.10404032 W
0.08283268 W
400.1316005 MHz
32768
100.6128045 MHz
EM
0
1.00 Hz
0
1.40
0.88
0.62
0.56
0.41
0.35
0.20
0.15
0.11
0.00
0.06
0.10
0.21
0.27
0.30
0.42
0.51
0.62
57.37
67.36
68.58
68.84
69.75
74.35
79.56
114.43
117.00
124.16
127.22
127.53
154.87
158.16
HN
HN
O
O
N
O
O
O
8
ppm
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210

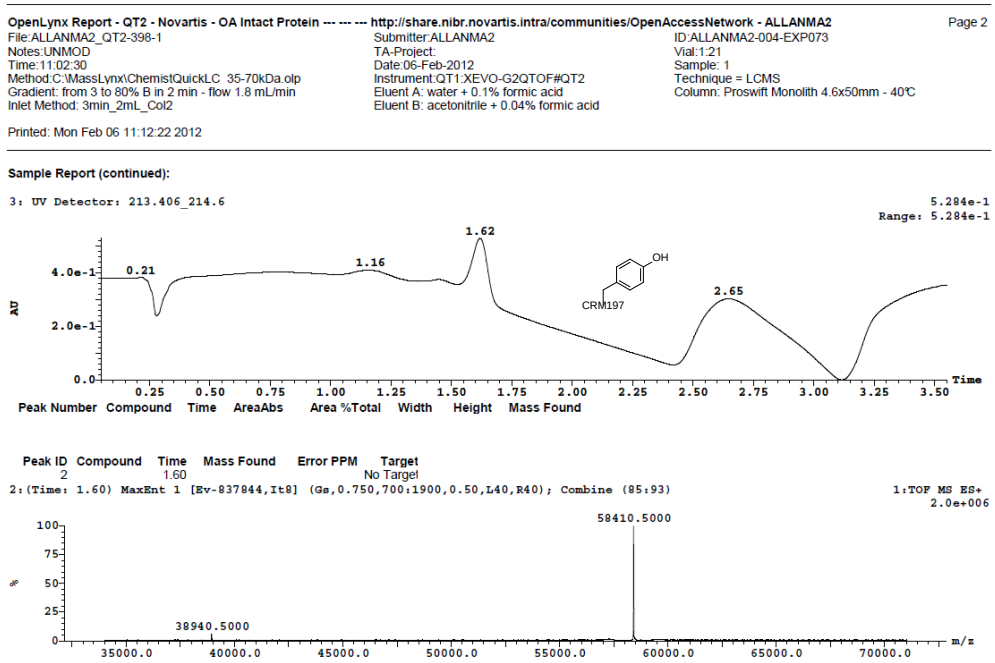

ALLANMA2-001-EXP073, ALLANMA2,
%Reflectance
Wavenumbers (cm-1)
607.9
650.8
682.2
715.7
735.4
760.9
787.0
821.4
846.3
918.6
950.8
1026.6
1054.2
1080.0
1113.9
1130.1
1238.6
1251.0
1297.2
1359.7
1456.3
1514.3
1613.6
1694.0
1772.8
2903.4
3237.5
O
HN
HN
N
O
O
O
O
8

4-(4-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)phenyl)-3H-1,2,4-triazole- 3,5(4H)-dione (**3**)

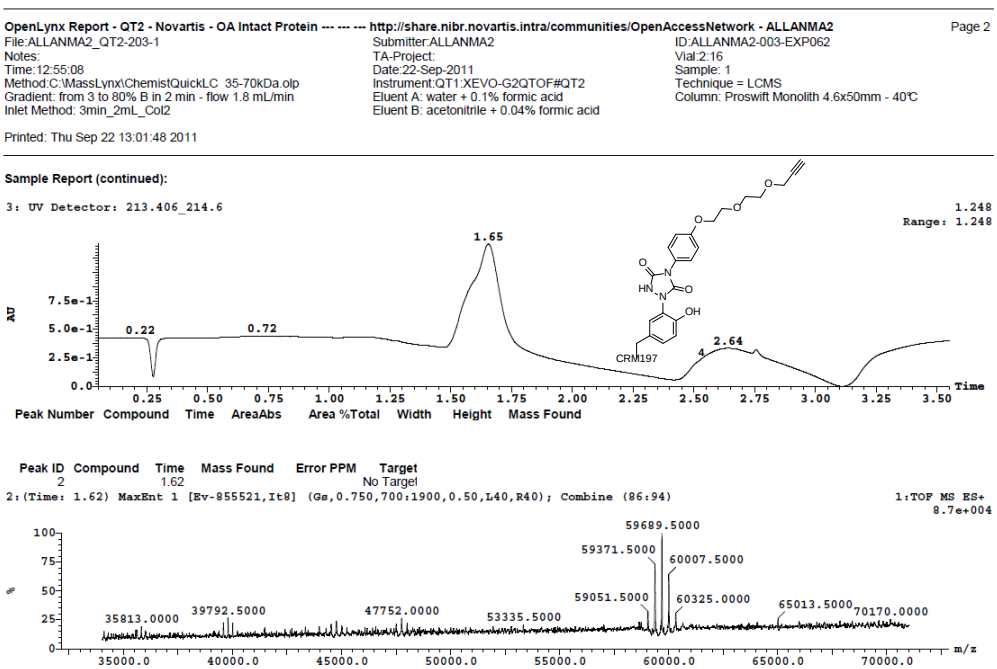


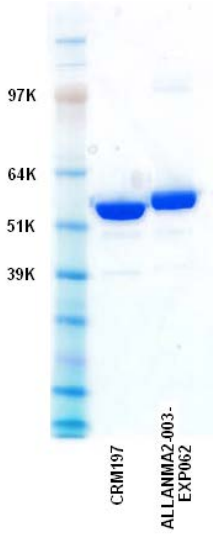
2.8 ESI MS spectra of modified proteins

CRM_{197}



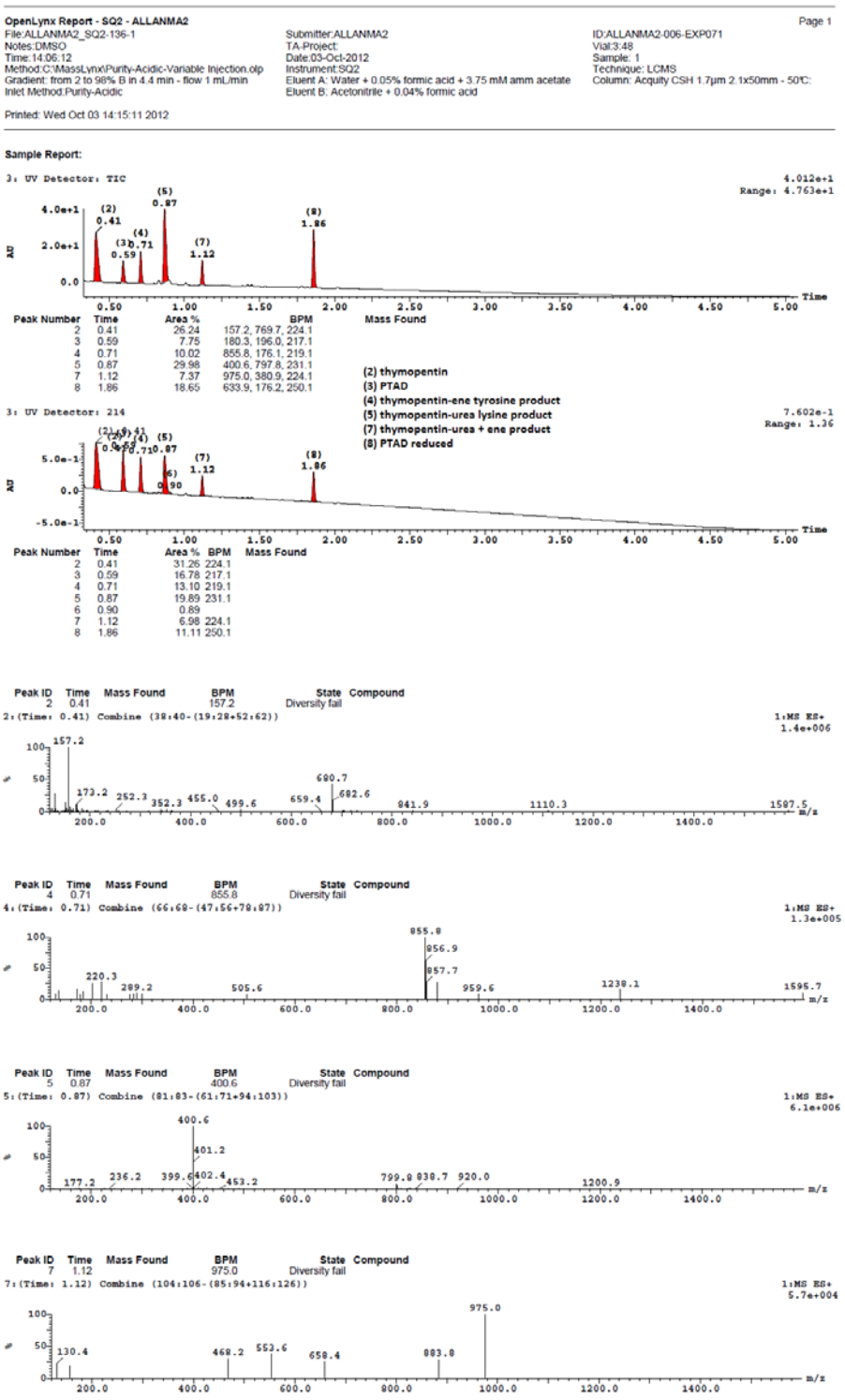
CRM_{197}-alkyne (**4**)



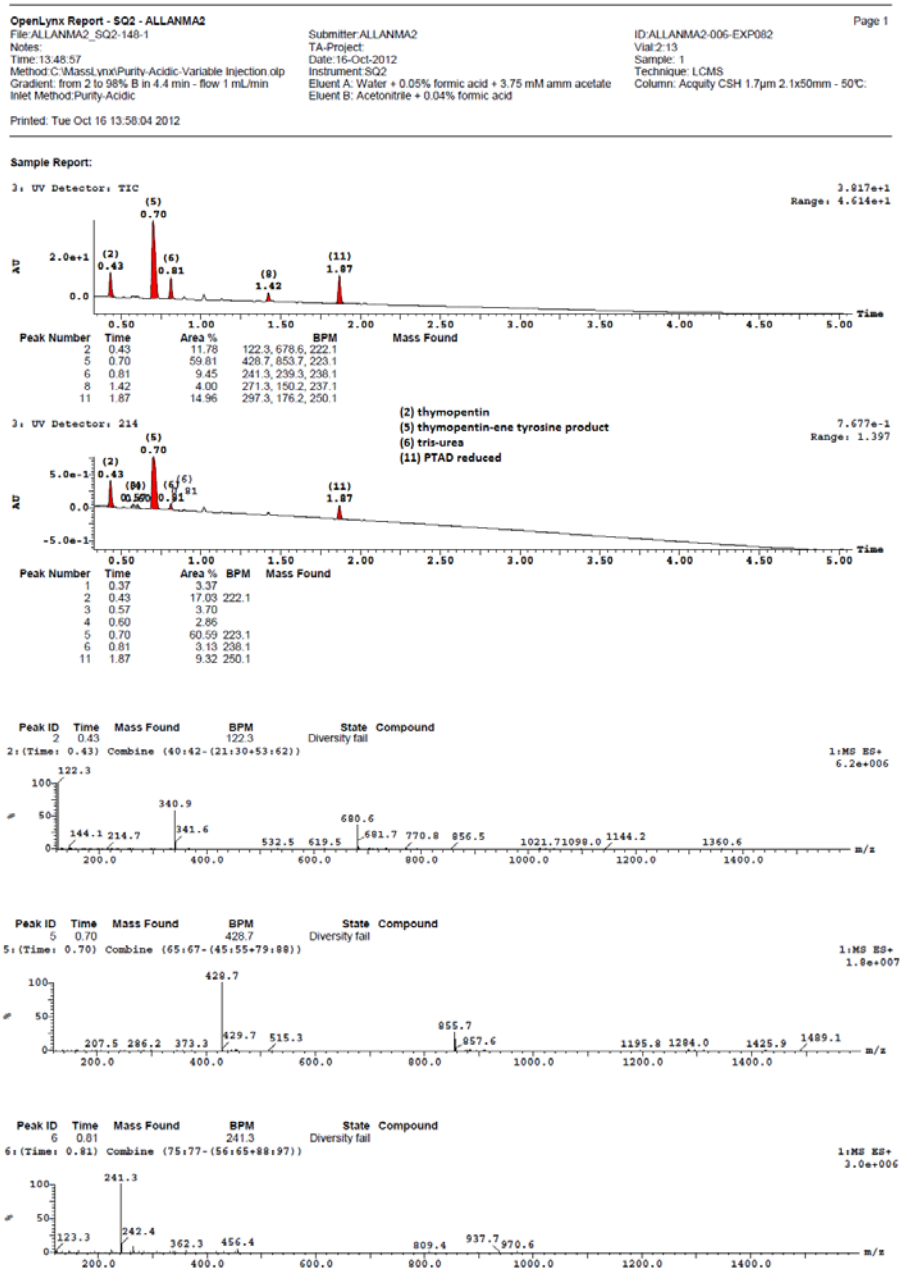


Condition optimization on thymopentin and proteins

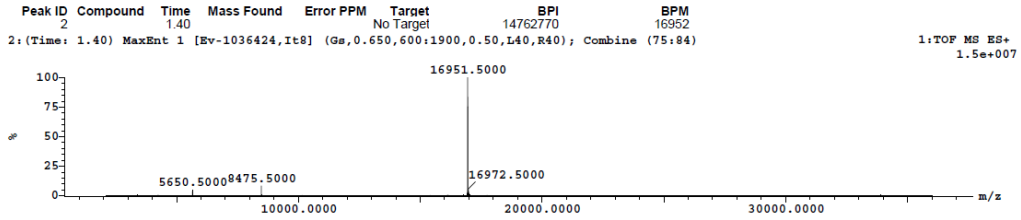
Thymopentin in PBS



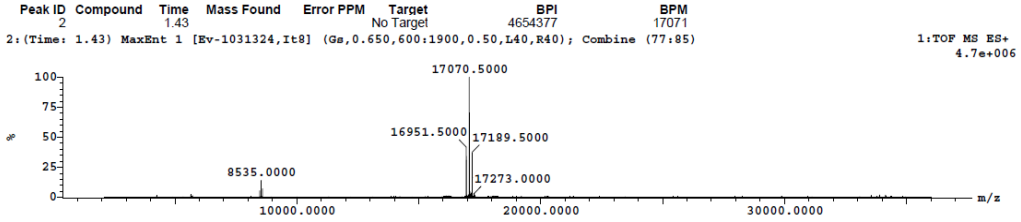
Thymopentin in Tris



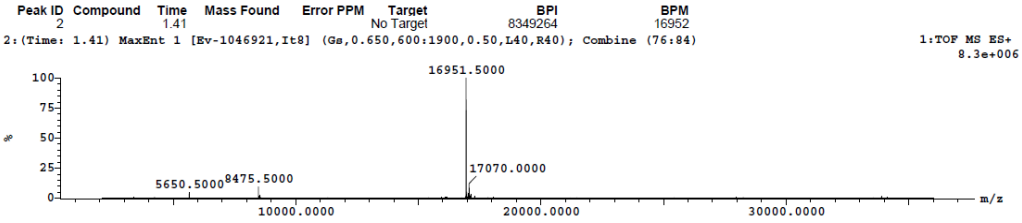
Myoglobin



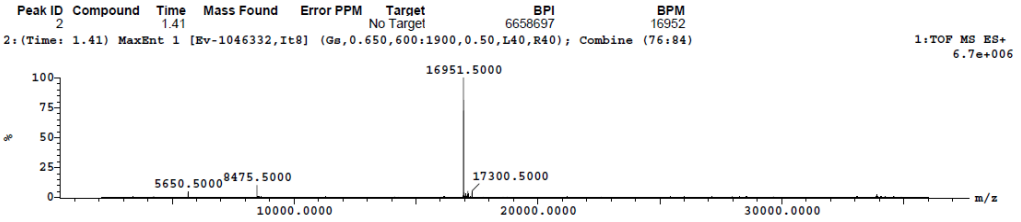
Myoglobin in PBS



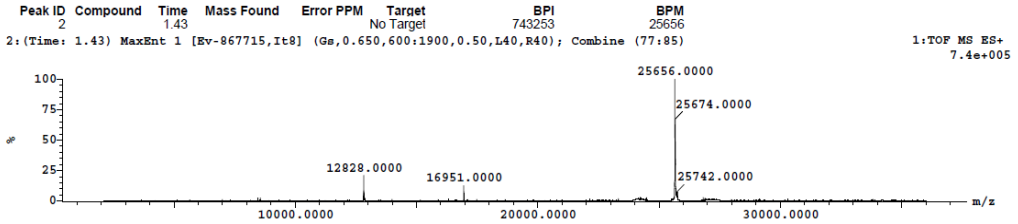
Myoglobin in Tris 0.2M



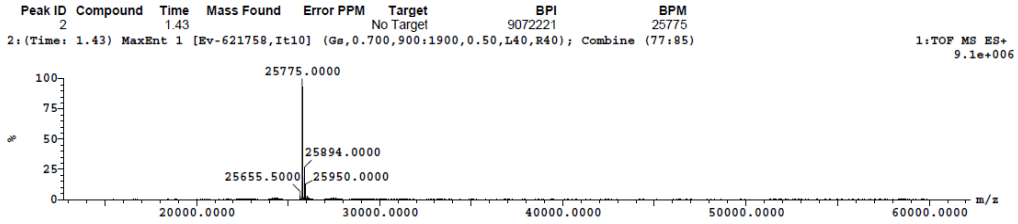
Myoglobin in Tris 1M



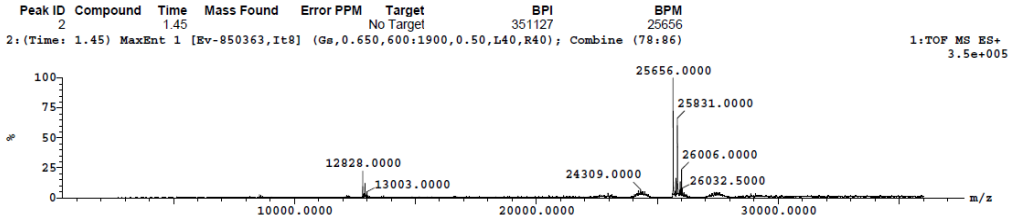
α-chymotrypsinogen A



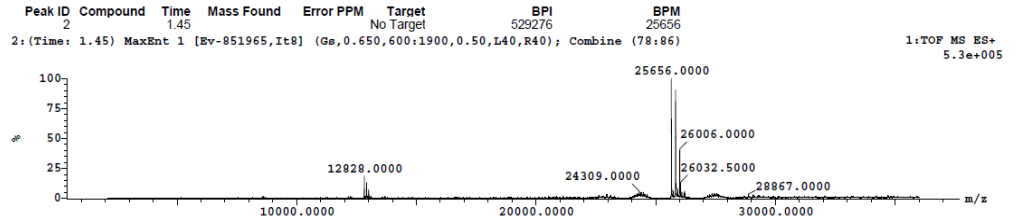
α-chymotrypsinogen A in PBS



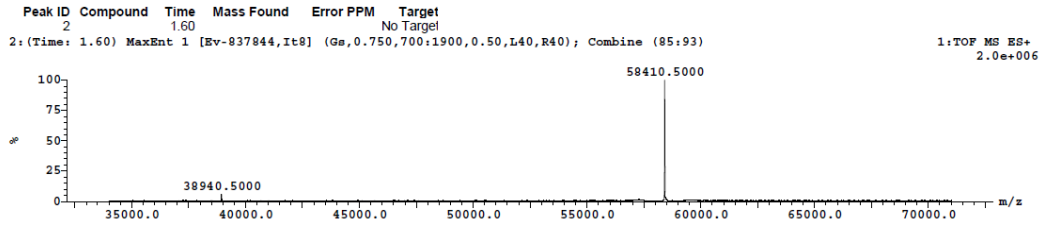
α-chymotrypsinogen A in Tris 0.2M



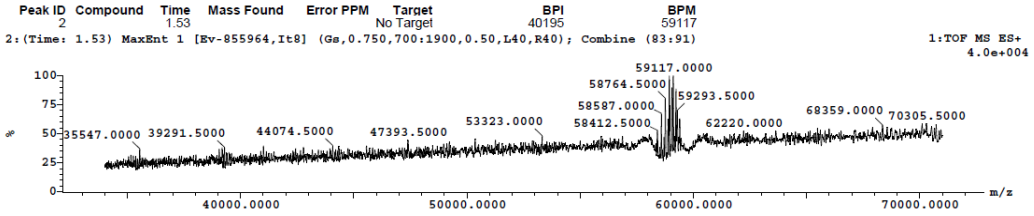
α-chymotrypsinogen A in Tris 1M



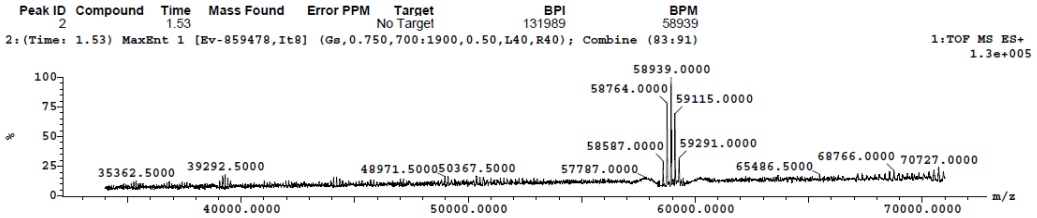
CRM_{197}



CRM_{197} in PBS

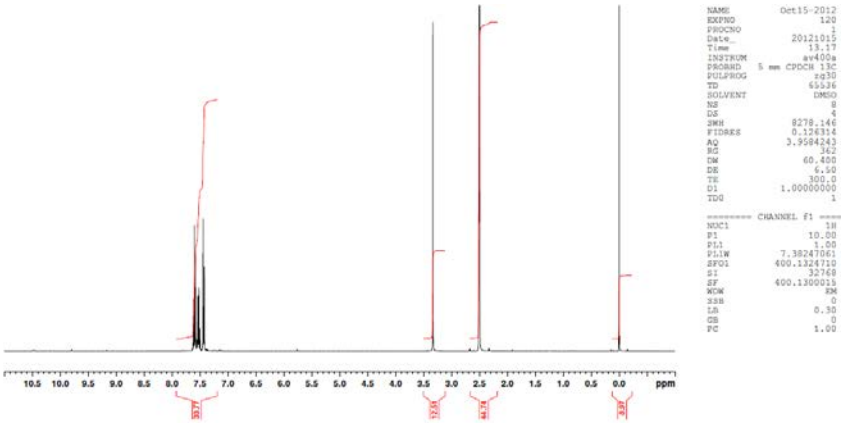


CRM_{197} in Tris 0.2M

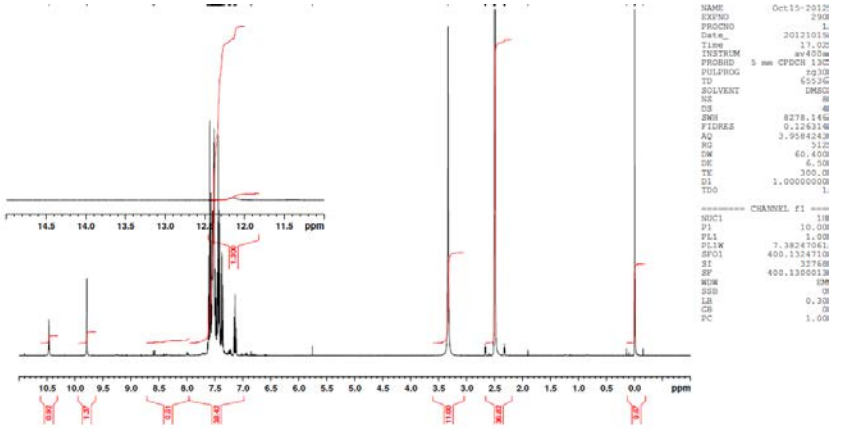


2.9 NMR study of PTAD stability

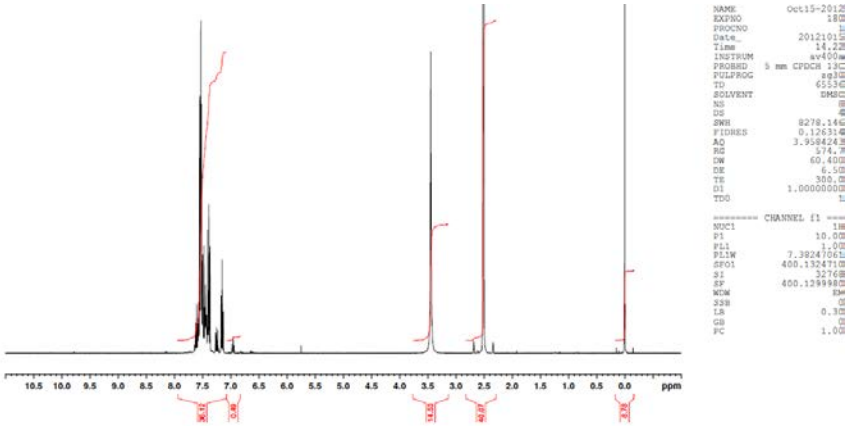
In DMSO-d6, 5min



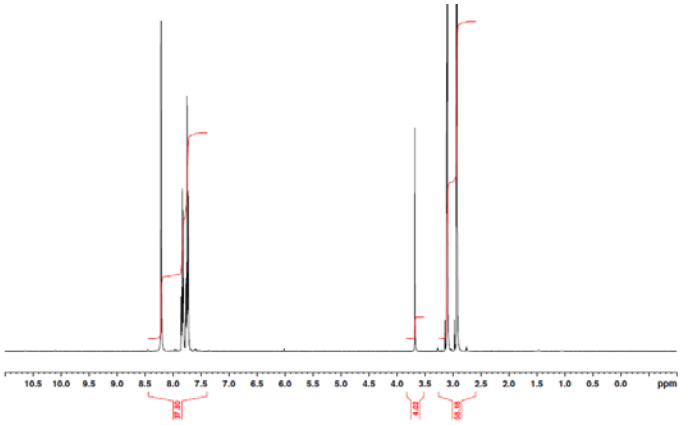
DMSO-d6, 3h



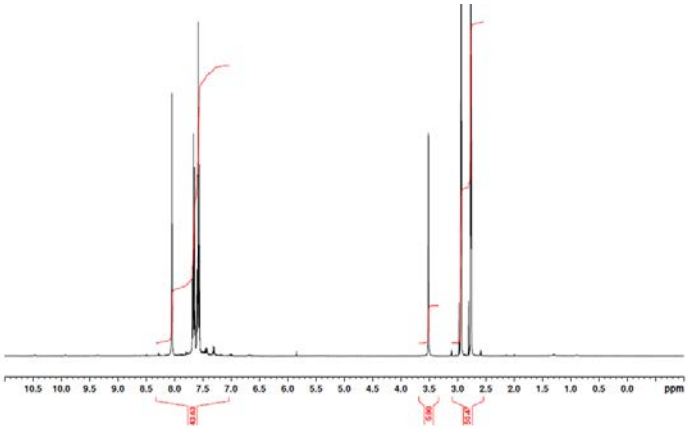
DMSO-d6 + one drop D_2O, 5min



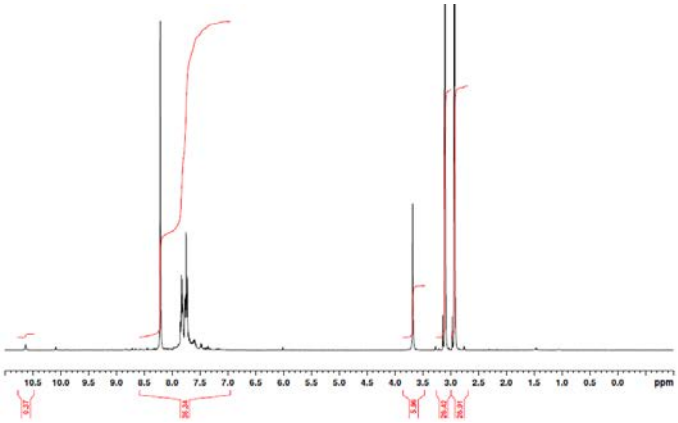
In DMF-d7, 5min



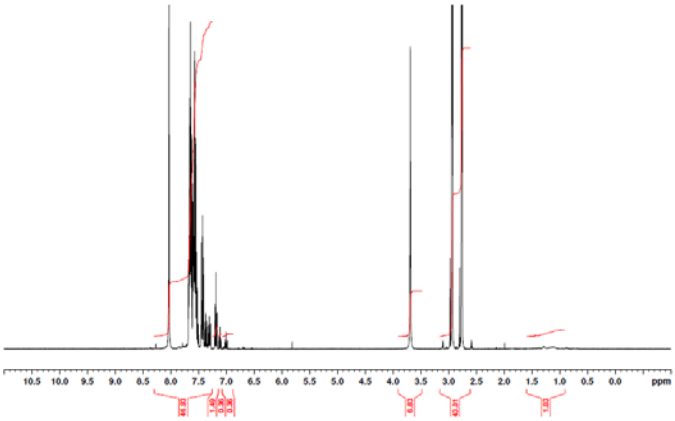
DMF-d7, 3h



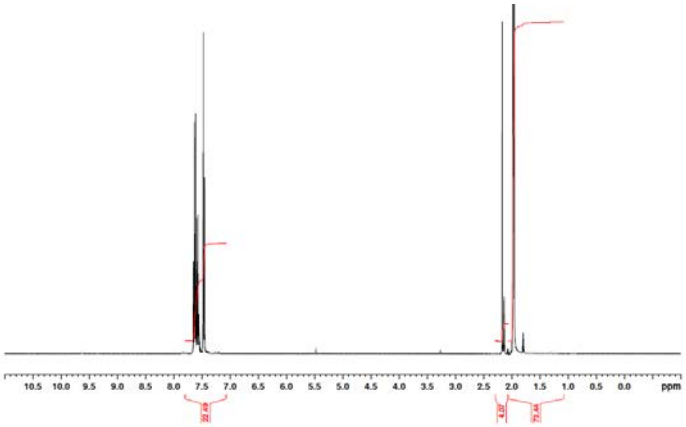
DMF-d7, 16h



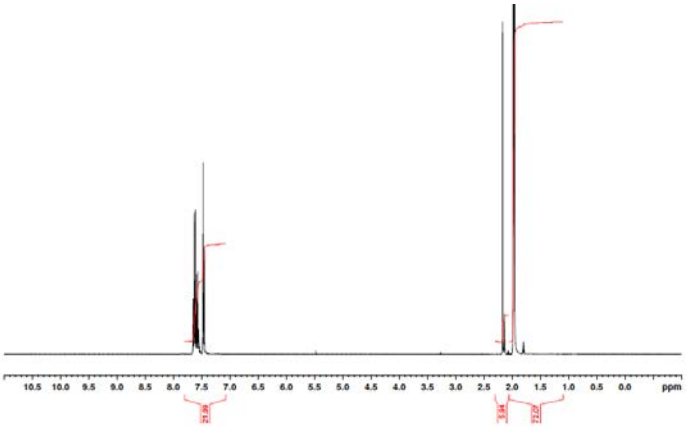
DMF-d7 + one drop D_2O, 5min



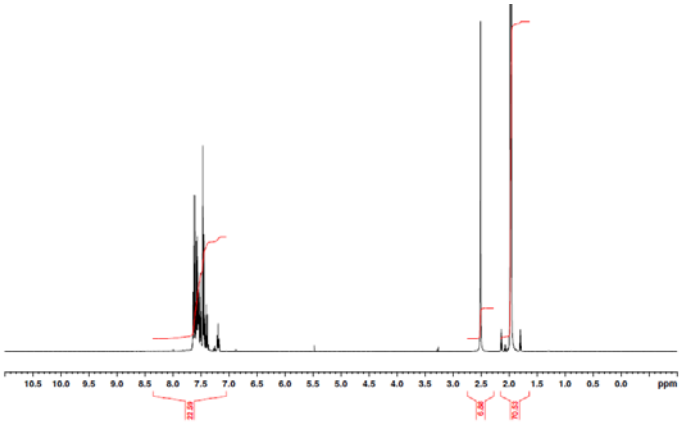
In CD_3CN, 5min



In CD_3CN, 16h



In CD_3CN + one drop D_2O, 5min



2.10 MS/MS analysis of proteolytic digests (Myoglobin conjugate):

```
  1 MGLSDGEWQQ VLNVWGKVEA DIAGHGQEVL IRLFTGHPET LEKFDKFKHL
 51 KTEAEMKASE DLKKHGTVVL TALGGILKKK GHHEAELKPL AQSHATKHKI
101 PIKYLEFISD AIIHVLHSKH PGDFGADAQG AMTKALELFR NDIAAKYKEL
151 GFQG
```

Red indicated the sequence was found in the search, and green indicated the modified residues. In addition, W14 on the N terminus and Y103 towards the C terminus were modified as well, but Lysine modification is much more prevalent compared with Y/W modification.

3.1 Synthesis of hexasaccharide 5

General procedure for glycosylation with trichloroacedimidate donors (Schemes 5 and 6)

To a stirred solution of acceptor (1 mmol) and donor (1.2 mmol) in anhydrous CH_2Cl_2 (15 mL) containing activated 4 Å MS (0.75 g), TMSOTf (0.2–0.4 mmol) was added at 0°C. The mixture was stirred for 30 min when TLC (2:1 cyclohexane-EtOAc) showed the reaction was complete. Then the mixture was neutralized with triethylamine, filtered through a celite pad, and the filtrate was concentrated. Chromatography of the residue (cyclohexane-EtOAc) gave the desired product.

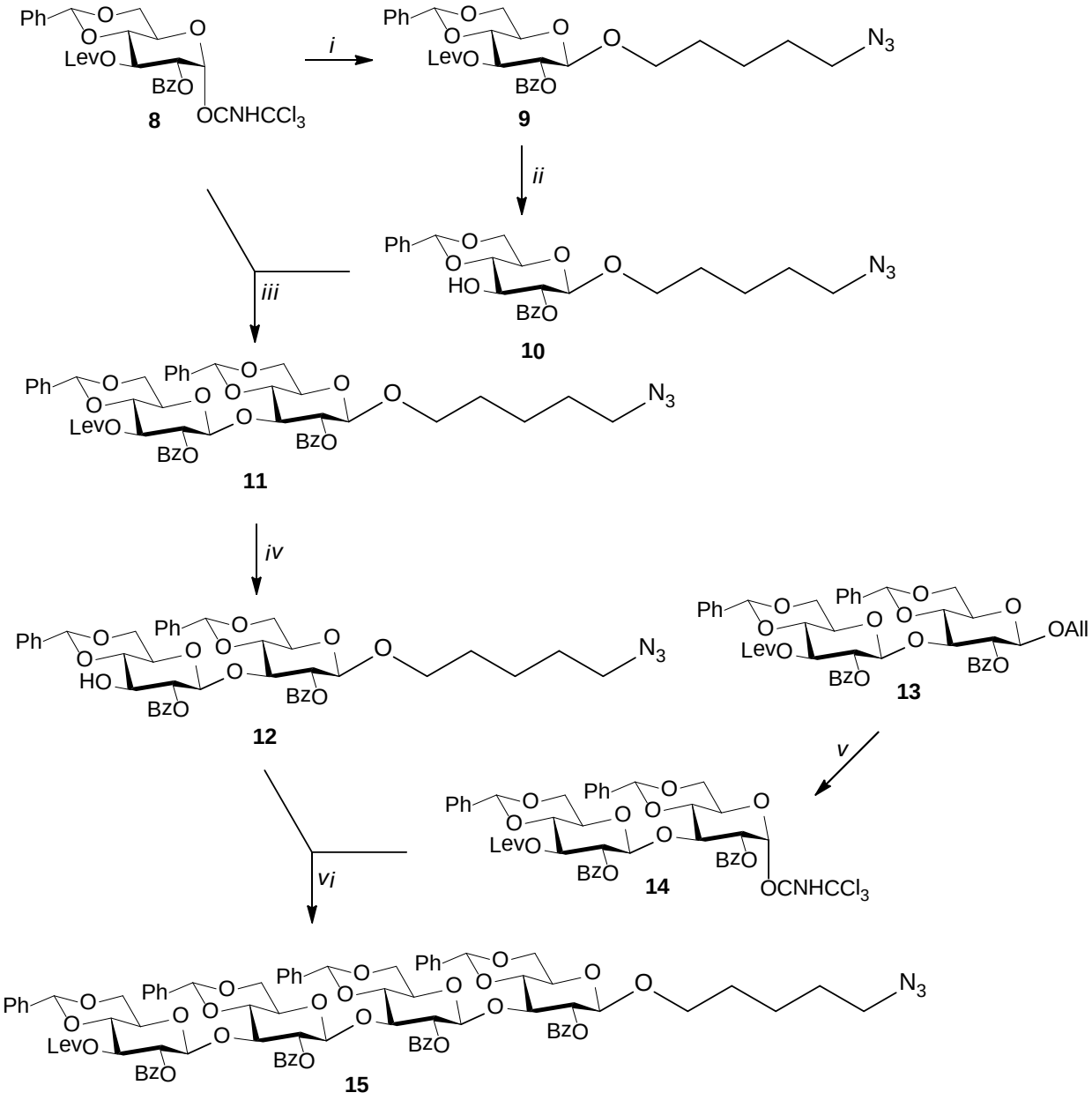
General procedure for delevulinoylation (Schemes 5 and 6)

To a solution of the 3-*O*-Lev oligosaccharide (1 mmol) in CH_2Cl_2 (25 mL) ethylenediamine (0.26 mL, 4 mmol) and AcOH (0.29 mL, 5 mmol) were added at 0°C. A white solid was formed, and the suspension was stirred for 5-6 h at 50°C, when the deprotection was complete (TLC, cyclohexane-EtOAc 2:1). The mixture was concentrated and chromatography of the residue (cyclohexane-EtOAc) yielded the delevulinoylated product.

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-levulinoyl-β-D-glucopyranoside 9. The general procedure for glycosylation was employed for 5-azidopentanol and **8** (53%). $[\alpha]_D^{25}$ = -21.7° (c 0.97, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.59 (s, 1 H, NH), 8.07–7.35 (m, 10 H, *Ph*CO), 5.54 (s, 1 H, PhC*H*), 5.51 (t, 1 H, J = 9.8 Hz, H-3), 5.28 (dd, 1 H, $J_{1,2}$ = 8.4 Hz, H-2), 4.69 (d, 1 H, $J_{1,2}$ = 8.4 Hz, H-1), 4.40 (dd, 1 H, , $J_{5,6a}$ = 4.8, $J_{6a,6b}$ = 10.2 Hz, H-6a), 3.97–3.87 (m, 1 H, H-1a′), 3.86 (t, 1 H, J = 10.4 Hz, H-6b), 3.78 (t, 1 H, J = 9.8 Hz, H-4), 3.62–3.56 (m, 1 H, H-5), 3.51–3.45 (m, 1 H, H-1b′), 3.01–2.84 (m, 2 H, H-5′), 2.59–2.47 (m, 4 H, CH_2CH_2), 1.98 (s, 3 H, CH_3), 1.57–1.46 (m, 2 H, H-2′), 1.47–1.36 (m, 2 H, H-3′), 1.32–1.36 (m, 2 H, H-4′). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 205.60 (CO), 171.71, 165.01 (COO), 136.70–126.02 (Ar), 101.57

(C-1), 101.25 (PhCH), 78.36 (C-4), 72.35 (C-2), 71.55 (C-3), 69.89 (C-1′), 68.43 (C-6), 66.60 (C-5), 50.91 (C-5′), 37.81 (CH_2CO), 29.36 (CH_3), 28.74 (C-2′), 28.20 (C-4′), 27.88 (CH_2COO), 22.89 (C-3′). ESI HR-MS ($C_{30}H_{35}N_3O_9$): *m/z* = found ($[M+Na]^+$ 604.2291; calc 604.2271).

Scheme 5. Reactions leading to tetrasaccharide **15**



i. 5-azidopentanol, 20%TMSOTf, CH_2Cl_2, 53%; *ii*. $H_2NCH_2CH_2NH_2{\cdot}AcOH$, CH_2Cl_2, 50°C, 69%; *iii*. 20% TMSOTf, CH_2Cl_2, 73%; *iv*. $H_2NCH_2CH_2NH_2{\cdot}AcOH$, CH_2Cl_2, 50°C, 98%; *v*. 1,5-Cyclooctadiene-bis(methyldiphenylphosphine)-Iridium-hexafluorophosphate catalyst, THF; I_2, H_2O; CCl_3CN, DBU, CH_2Cl_2, 88% (over two steps); *vi*. 20% TMSOTf, CH_2Cl_2, 83%.

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranoside 10. The 3-OLev group was removed in compound **9**[3] according to the general procedure (69%). $[\alpha]_D^{25}$ = -35.7° (c 0.92, $CHCl_3$). ^{1}H NMR ($CDCl_3$, 400 MHz) δ: 8.10–7.29 (m, 10 H, *Ph*CO), 5.58 (s, 1 H, PhC*H*), 5.19 (d, 1 H, $J_{1,2}$ = 8.6 Hz, H-2), 4.66 (d, 1 H, $J_{1,2}$ = 7.8 Hz, H-1), 4.39 (dd, 1 H , $J_{5,6a}$ = 5.0, $J_{6a,6b}$ = 10.5 Hz, H-6a), 4.06 (t, 1 H, J = 8.8 Hz, H-3), 3.94–3.88 (m, 1 H, H-1a′), 3.84 (t, 1 H, J = 10.4 Hz, H-6b), 3.68 (t, 1 H, J = 9.6 Hz, H-4), 3.56–3.47 (m, 1 H, H-5), 3.51–3.45 (m, 1 H, H-1b′), 3.03–2.91 (m, 2 H, H-5′), 2.73 (br. s, 1 H, OH-3), 1.59–1.35 (m, 4 H, H-2′,3′), 1.36–1.32 (m, 2 H, H-4′). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 165.11 (COO), 133.85–126.25 (Ar), 101.92 (PhCH), 101.57 (C-1), 80.91 (C-4), 74.74 (C-2), 72.34 (C-3), 69.91 (C-1′), 68.61 (C-6), 66.22 (C-5), 51.06 (C-5′), 28.91 (C-2′), 28.34 (C-4′), 23.06 (C-3′). ESI HR-MS ($C_{30}H_{35}N_3O_9$): *m/z* = found ([*M*+Na]$^+$ 604.2291; calc 604.2271). ESI HR-MS ($C_{25}H_{29}N_3O_7$): *m/z* = found ([*M*+H]$^+$ 484.2045; calc 404.2084); found ([*M*+Na]$^+$ 506.1887; calc 506.1903).

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-levulinoyl-β-D-glucopyranosyl-(1→3)-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranoside 11. The general procedure for glycosylation was applied to compound **10**. Yield: 73%. White crystals from EtOAc: m.p. 156–157°C. $[\alpha]_D^{25}$ = -45.1° (c 0.41; $CHCl_3$). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.72–7.23 (m, 20 H, 2 × *Ph*CO, 2 × *Ph*CH), 5.57 (s, 1 H, PhC*H*), 5.35 (s, 1 H, PhC*H*), 5.31 (t, 1 H, J = 9.3 Hz, H-3^B), 5.29–5.16 (m, 2 H, J = 8.2 Hz, H-2^A,2^B), 4.92 (d, 1 H, $J_{1,2}$ = 7.3 Hz, H-1^B), 4.75 (d, 1 H, $J_{1,2}$ = 7.5 Hz, H-1^A), 4.36 (dd, 1 H, $J_{5,6a}$ = 4.9, $J_{6a,6b}$ = 10.5 Hz, H-6a^B), 4.21–4.14 (m, 2 H, H-6a^A,3^B), 3.86–3.75 (m, 4 H, H-4^{A,B},6b^B,1a′), 3.70 (t, 1 H, J = 10.0 Hz, H-6b^A), 3.56–3.42 (m, 2 H, H-5^{A,B}), 3.36–3.31 (m, 1 H, H-1b′), 2.91–2.76 (m, 2 H, H-5′), 2.45–2.31 (m, 4 H, CH_2CH_2), 1.94 (s, 3 H, CH_3), 1.42–1.23 (m, 4 H, H-2′,3′), 1.17–1.04 (m, 2 H, H-4′). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 205.85 (CO), 171.68, 164.79, 164.45, 163.53 (COO), 137.93–125.92 (Ar), 101.43 (C-1^A), 101.25 (PhCH), 101.14 (PhCH), 100.70 (C-1^B), 79.17 (C-4^B), 78.55 (C-3^A), 77.32 (C-4^A), 73.16, 72.67 (C-2^{A,B}), 71.84 (C-3^B), 69.57 (C-1′), 68.50 (C-6^B), 68.37 (C-6^A), 66.32, 66.07 (C-5^{A,B}), 50.81 (C-

5′), 37.66 (*C*H_2CO), 29.54 (CH_3), 29.31 (C-2′), 29.12 (C-3′), 27.78 (*C*H_2COO), 22.80 (C-4′). ESI HR-MS ($C_{50}H_{53}N_3O_{15}$): *m/z* = found ([*M*+Na]$^+$ 958.3354; calc 958.3374).

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranoside 12. After applying the general procedure for delevulinoylation to disaccharide **11**, compound **12** was obtained in 98% yield. White crystals from EtOAc: m.p. 179–180°C. $[\alpha]_D^{25}$ = -3.6° (c 0.70; $CHCl_3$). ^{1}H NMR (400 MHz, $CDCl_3$) *δ*: 7.84–7.23 (m, 20 H, 2 × *Ph*CO, 2 × *Ph*CH), 5.56 (s, 1 H, PhC*H*), 5.37 (s, 1 H, PhC*H*), 5.25 (t, 1 H, *J* = 8.5 Hz, H-2^A), 5.12 (t, 1 H, *J* = 7.9 Hz, H-2^B), 4.93 (d, 1 H, $J_{1,2}$ = 7.1 Hz, H-1^B), 4.56 (d, 1 H, $J_{1,2}$ = 7.0 Hz, H-1^A), 4.37 (dd, 1 H, $J_{5,6a}$ = 4.9, $J_{6a,6b}$ = 10.5 Hz, H-6a^B), 4.22–4.18 (m, 2 H, H-6a^A,3^B), 3.87–3.76 (m, 4 H, H-3^B,4^B,6b^B,1a′), 3.70 (t, 1 H, $J_{6a,6b}$ = 10.0 Hz, 6b^A), 3.65 (t, 1 H, *J* = 10.0 Hz, H-4^A), 3.57–3.51 (m, 2 H, H-5^B), 3.41–3.31 (m, 2 H, H-5^A,1b′), 2.92–2.83 (m, 2 H, H-5′), 1.42–1.27 (m, 4 H, H-2′,3′), 1.16–1.08 (m, 2 H, H-4′). ^{13}C NMR ($CDCl_3$, 100 MHz) *δ*: 165.52, 164.52 (COO), 137.11–126.07 (Ar), 101.65 (Ph*C*H), 101.49 (C-1^A), 101.39 (Ph*C*H), 100.45 (C-1^B), 80.48 (C-4^A), 79.33 (C-4^B), 78.05 (C-3^A), 75.19 (C-2^A), 73.50 (C-2^B), 72.54 (C-3^B), 69.59 (C-1′), 68.70 (C-6^B), 68.60 (C-6^A), 66.44 (C-5^B), 66.04 (C-5^A), 50.99 (C-5′), 28.80 (C-2′), 28.26 (C-3′), 22.97 (C-4′). ESI HR-MS ($C_{45}H_{47}N_3O_{13}$): *m/z* = found ([*M*+Na]$^+$ 860.2998; calc 860.3007).

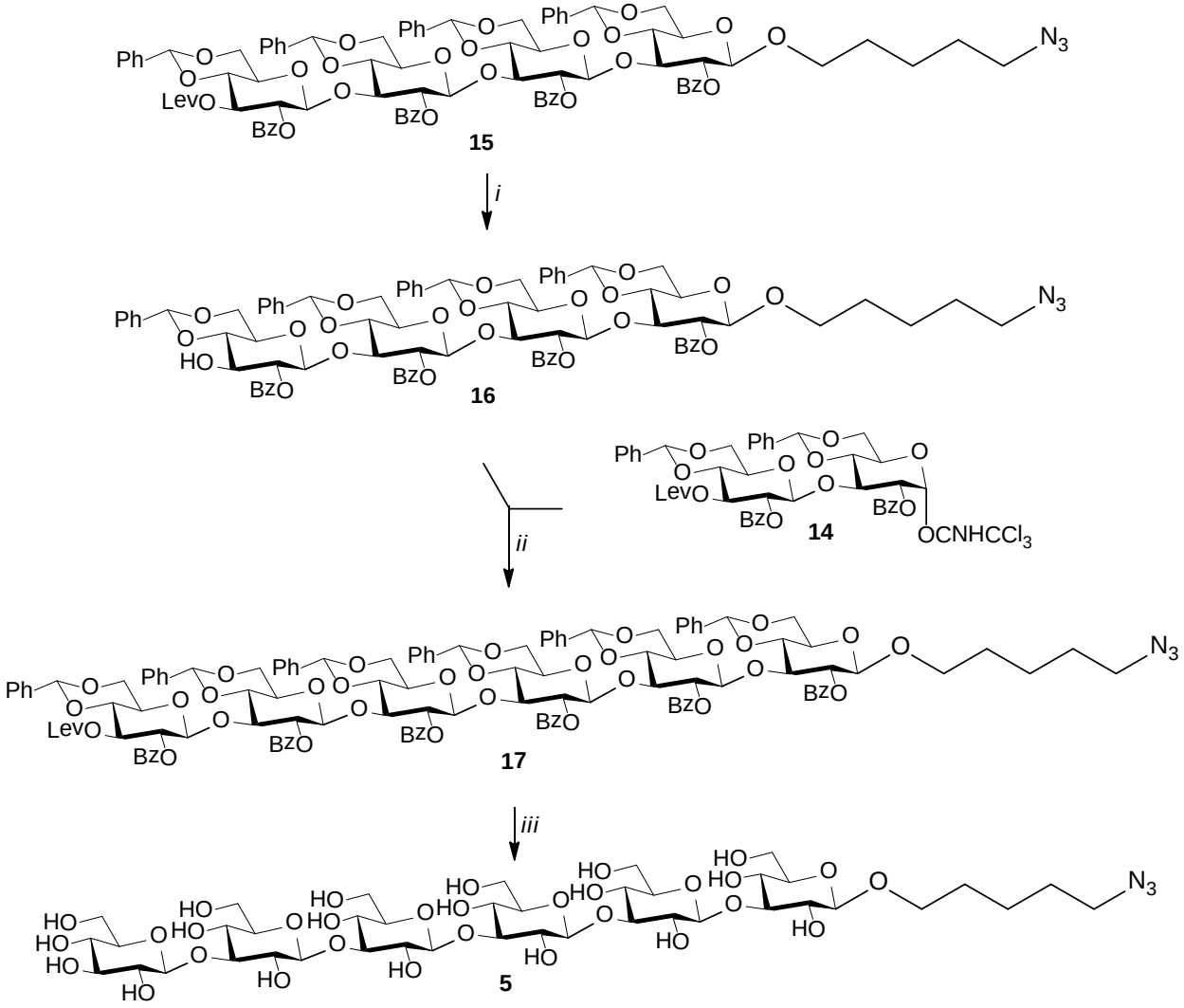
2-*O*-Benzoyl-3-*O*-levulinoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl trichloroacetimidate 14. A mixture of the 1-*O*-allyl compound **13**[3] (865 mg, 1 mmol) and 1,5-cyclooctadiene-bis(methyldiphenylphosphine)-Iridium(I)-hexafluorophosphate catalyst (2 mg) in dry THF (7 mL) was carefully degassed. The catalyst was activated under hydrogen atmosphere for 2 min, until the catalyst turned from red to pale yellow, and the reaction mixture was stirred at room temperature for 3 h. When NMR analysis of a small portion showed complete isomerization, water was added in order to get a 4:1 THF:H_2O mixture, followed by iodine (0.5 g, 2 mmol). After 15 min (TLC, 1:1 cyclohexane-EtOAc) the solution was diluted with water and extracted with CH_2Cl_2 (3 x 10 mL).

The foregoing 1-OH disaccharide was dissolved in CH_2Cl_2 (15 mL) and CCl_3CN (1.5 mL) to which DBU (40 µl) was added. After 30 min the reaction was complete (1:1 cyclohexane-EtOAc). The mixture was concentrated and after purification with cyclohexane-EtOAc 850 mg of product **15** (88%, over two steps) were obtained. $[\alpha]_D^{25}$ = +48.3˚ (c 2.2; $CHCl_3$). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.51 (s, 1 H, NH), 7.73–7.21 (m, 20 H, 2 × *Ph*CO, 2 × *Ph*CH), 6.55 (d, 1 H, $J_{1,2}$ = 4.0 Hz, H-1^A), 5.56 (s, 1 H, PhC*H*), 5.27–5.24 (m, 2 H, PhC*H*, H-3^B), 5.10–5.04 (m, 2 H, H-2^{A,B}), 5.05 (d, 1 H, $J_{1,2}$ = 7.3 Hz, H-1^B), 4.48 (t, 1 H, J = 9.5 Hz, H-3^A), 4.35 (t, 1 H, $J_{6a,6b}$ = 10.5 Hz, H-6a$^{A/B}$), 4.33 (t, 1 H, $J_{6a,6b}$ = 10.3 Hz, H-6a$^{A/B}$), 4.12 (dd, 1 H, $J_{5,6b}$ = 4.9, $J_{6a,6b}$ = 10.5 Hz, H-6b^B), 3.90–3.57 (m, 4 H, H-4^{A,B},5^A,6b$^{A/B}$), 3.56–3.51 (m, 1 H, H-5^B), 2.55–2.31 (m, 4 H, CH_2CH_2), 1.93 (s, 3 H, CH_3). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 205.79 (CO), 171.70, 165.20, 164.88 (COO), 137.26–125.96 (Ar), 101.28 (Ph*C*H, C-1^B), 101.19 (Ph*C*H), 93.37 (C-1^A), 78.00 (C-4^A), 77.89 (C-4^B), 76.68 (C-3^A), 72.75 (C-2$^{A/B}$), 72.08 (C-2$^{A/B}$), 71.83 (C-3^B), 68.50 (C-6 $^{A/B}$), 68.49 (C-6 $^{A/B}$), 66.24 (C-5^B), 65.07 (C-5^A), 37.75 (*C*H$_2$CO), 29.62 (CH_3), 27.83 (*C*H$_2$COO). ESI HR-MS ($C_{47}H_{44}Cl_3NO_{15}$): *m/z* = found ([*M*+Na]$^+$ 990.1671; calc 990.1674).

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-3-*O*-levulinoyl-β-D-glucopyranosyl-(1→3)-bis[2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)]-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranoside 15. Donor **14** and acceptor **12** were coupled according to the general procedure for glycosylation. Yield: 83%. White crystals from EtOAc: m.p. 109–110°C. $[\alpha]_D^{25}$ = +17.2˚ (c 0.70; $CHCl_3$). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.86–7.20 (m, 40 H, 2 × *Ph*CO, 2 × *Ph*CH), 5.50 (s, 1 H, PhC*H*), 5.38 (m, 2 H, PhC*H*, H-3^D), 5.26 (t, 1 H, J = 8.7 Hz, H-2^D), 5.16 (t, 1 H, J = 7.9 Hz, H-2^B), 5.04 (d, 1 H, $J_{1,2}$ = 5.9 Hz, H-1^C), 4.98 (d, 1 H, $J_{1,2}$ = 8.6 Hz, H-1^D), 4.97 (d, 1 H, $J_{1,2}$ = 5.9 Hz, H-1^B), 4.86 (t, 1 H, J = 7.6 Hz, H-2^A), 4.83–4.80 (m, 2 H, PhC*H*, H-2^C), 4.71 (s, 1 H, PhC*H*), 4.43 (d, 1 H, $J_{1,2}$ = 7.8 Hz, H-1^A), 4.32 (dd, 1 H, $J_{5,6a}$ = 5.6, $J_{6a,6b}$ = 10.4 Hz), 4.21 (dd, 1 H, $J_{5,6a}$ = 5.0, $J_{6a,6b}$ = 10.4 Hz), 4.16–4.11 (m, 2 H), 4.10–4.04 (m, 2 H), 3.96–3.90 (m, 2 H), 3.82–3.68 (m, 4 H), 3.62–3.40 (m, 7 H), 3.35–3.30 (m, 1 H, H-1b′), 3.27 (t, 1 H, J = 9.4 Hz, H-4^A), 2.98–2.83 (m, 2 H, H-5′), 2.53–2.37 (m, 4 H, CH_2CH_2), 1.97 (s, 3 H, CH_3), 1.49–1.10 (m, 6 H, H-2′,3′,4′). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 205.83 (CO), 171.69, 165.03, 164.55, 164.45 (COO), 137.24–126.02 (Ar), 101.78 (Ph*C*H), 101.46 (C-1^A), 101.26

(Ph*C*H), 100.90 (Ph*C*H), 100.82 (Ph*C*H), 99.32 (C-1^{D}), 98.29 (C-1^{C}), 97.26 (C-1^{B}), 78.80 (C-4^{A}), 78.44, 78.24, 74.81, 74.68, 74.09, 73.27, 72.62, 71.98, 69.54 (C-1′), 68.61, 66.45, 66.14, 65.64, 51.02 (C-5′), 37.92 (*C*H_2CO), 29.52 (CH_3), 28.84 (C-2′), 28.30 (C-3′), 27.97 (*C*H_2COO), 23.00 (C-4′). ESI HR-MS ($C_{90}H_{89}N_3O_{27}$): *m/z* = found ([*M*+Na]$^+$ 1666.5562; calc 1666.5581).

Scheme 6. Reactions leading to hexasaccharide **5**



i. $H_2NCH_2CH_2NH_2$·AcOH, CH_2Cl_2, 50°C, 92%; *iii*. 40% TMSOTf, CH_2Cl_2, 92%; *iii*. 9:1 AcOH-H_2O, 50°C; NaOMe, MeOH 93% (over two steps).

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)-bis[2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)]-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-

glucopyranoside 16. Delevulinoylation of **15** afforded **16** in 92% yield. $[\alpha]_D^{25}$ = +13.6˚ (c 0.45; $CHCl_3$).

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.97–7.20 (m, 40 H, 2 × *Ph*CO, 2 × *Ph*CH), 5.56 (s, 1 H, PhC*H*), 5.40 (s, 1 H, PhC*H*), 5.20 (t, 1 H, *J* = 8.1 Hz, H-2^D), 5.16 (t, 1 H, *J* = 4.8 Hz, H-2^B), 5.06 (d, 1 H, $J_{1,2}$ = 7.4 Hz, H-1^D), 5.01–4.96 (m, 3 H, PhC*H*, H-1^B,2^A), 4.85 (s, 1 H, PhC*H*), 4.83–4.81 (m, 2 H, H-1^C,2^C), 4.48 (d, 1 H, $J_{1,2}$ = 7.8 Hz, H-1^A), 4.35 (dd, 1 H, $J_{5,6a}$ = 5.6, $J_{6a,6b}$ = 10.4 Hz), 4.24–4.06 (m, 5 H), 4.02–3.95 (m, 3 H), 3.91 (t, 1 H, *J* = 9.3, H-3^C), 3.84–3.29 (m, 1 H, H-1a′), 3.76 (t, 1 H, *J* = 9.3, H-3^D), 3.70 (t, 1 H, *J* = 9.3), 3.64–3.40 (m, 10 H), 2.98–2.85 (m, 2 H, H-5′), 1.49–1.25 (m, 4 H, H-2′,3′), 1.22–1.19 (m, 2 H, H-4′). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 165.78, 164.55, 164.58 (COO), 137.24–125.72 (Ar), 101.76 (Ph*C*H), 101.71 (Ph*C*H), 101.50 (C-1^A), 101.26 (Ph*C*H), 100.68 (Ph*C*H), 98.77 (C-1^C), 98.69 (C-1^D), 97.08 (C-1^B), 78.80 (C-4^A), 80.77, 78.98, 78.11, 77.57, 77.50, 77.18, 76.77, 75.54, 75.75, 73.98, 72.54, 73.35, 69.54 (C-1′), 68.67, 66.49, 66.05, 65.64, 65.48, 51.02 (C-5′), 28.85 (C-2′), 28.29 (C-3′), 23.00 (C-4′). ESI HR-MS ($C_{85}H_{83}N_3O_{25}$): *m/z* = found ([*M*+Na]$^+$ 1568.5248; calc 1568.5213).

5-Azidopentyl 2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)-tetrakis[2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranosyl-(1→3)]-2-*O*-benzoyl-4,6-*O*-benzylidene-β-D-glucopyranoside 17. Donor **14** and acceptor **16** were coupled according to the general procedure for glycosylation. Yield: 92%. $[\alpha]_D^{25}$ = +2.2˚ (c 0.67; $CHCl_3$).

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.85–7.19 (m, 60 H, 2 × *Ph*CO, 2 × *Ph*CH), 5.53 (s, 1 H, PhC*H*), 5.41 (t, 1 H, *J* = 8.6 Hz, H-3^F), 5.40 (s, 1 H, PhC*H*), 5.29 (t, 1 H, *J* = 7.9 Hz, H-2$^{B/F}$), 5.20 (t, 1 H, *J* = 5.4 Hz, H-2$^{B/F}$), 5.09 (d, 1 H, $J_{1,2}$ = 7.1 Hz, H-1$^{B/F}$), 5.00 (d, 1 H, $J_{1,2}$ = 5.7 Hz, H-1 $^{B/F}$), 4.96 (s, 1 H, PhC*H*), 4.85 (s, 1 H, PhC*H*), 4.92 (t, 1 H, *J* = 8.2 Hz, H-2^A), 4.88–4.77 (m, 9 H, 3 × PhC*H*, 3 × H-1$^{B/F}$, 3 × H-2 $^{B/F}$), 4.46 (d, 1 H, $J_{1,2}$ = 7.8 Hz, H-1^A), 4.33 (dd, 1 H, $J_{5,6a}$ = 4.8, $J_{6a,6b}$ = 10.7 Hz), 4.24 (dd, 1 H, $J_{5,6a}$ = 4.8, $J_{6a,6b}$ = 10.4 Hz), 4.20–4.05 (m, 6 H), 4.01–3.90 (m, 4 H), 3.82–3.67 (m, 3 H), 3.60–3.27 (m, 15 H), 2.98–2.82 (m, 2 H, H-5′), 2.54–2.31 (m, 4 H, CH_2CH_2), 1.97 (s, 3 H, CH_3), 1.48–1.28 (m, 4 H, H-2′,3′), 1.22–1.08 (m, 2 H, H-4′). ^{13}C NMR ($CDCl_3$, 100 MHz) δ: 205.83 (CO), 171.70, 165.07, 164.69, 164.64, 164.56, 164.48 (COO), 137.33–126.07 (Ar), 101.89 (Ph*C*H), 101.45 (C-1^A), 101.23 (Ph*C*H), 101.01 (Ph*C*H), 100.77 (Ph*C*H), 99.14 (C-1$^{B/F}$), 98.55 (C-1$^{B/F}$), 97.33 (C-1$^{B/F}$), 94.24 (C-1$^{B/F}$), 96.85 (C-1$^{B/F}$), 78.93, 78.36, 78.29, 77.79,

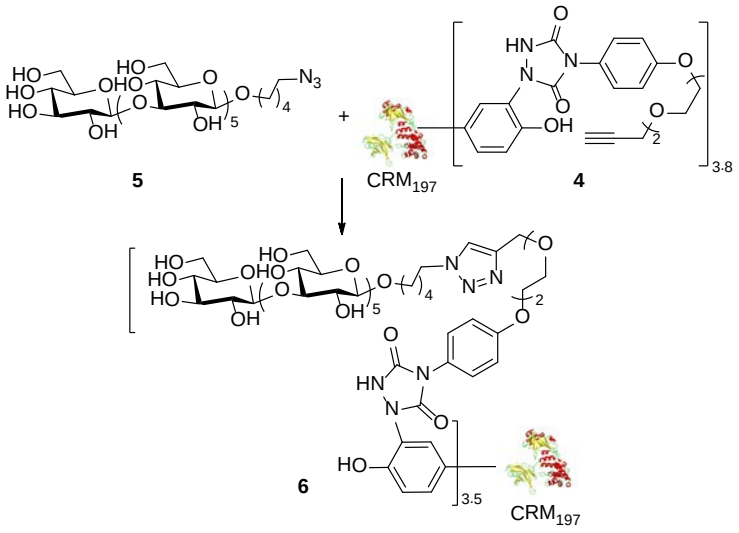
77.55, 75.21, 74.58, 74.30, 74.19, 74.04, 73.22, 73.06, 72.50, 72.05, 69.57 (C-1′), 68.59, 66.44, 66.16, 65.64, 65.51, 51.02 (C-5′), 37.93 (*C*H_2CO), 29.53 (CH_3), 28.84 (C-2′), 28.29 (C-3′), 28.01 (*C*H_2COO), 23.01 (C-4′). ESI HR-MS ($C_{130}H_{125}N_3O_{39}$): *m/z* = found ([*M*+Na]$^+$ 2374.7764; calc 2374.7788).

5-Azidopentyl β-D-glucopyranosyl-(1→3)-tetrakis[β-D-glucopyranosyl-(1→3)]-β-D-glucopyranoside 5. The protected hexasaccharide **17** (100 mg, 0.043 mmol) was dissolved in 9:1 AcOH-H_2O (2.5 mL). After stirring for 6 h at 50° the solvent was evaporated, and the material was dissolved in MeOH (2.5 ml) to which a solution of 1M NaOMe was added until pH was strongly alkaline (TLC, 9:1 CH_2Cl_2-MeOH). The mixture was stirred overnight, and then concentrated. The residue was purified on a Sephadex G-10 column, eluting with H_2O. Fractions containing the pure desired compound (NMR) were combined and freeze-dried to yield 47 mg of product **5** (93%). $[\alpha]_D^{25}$ = -1.4° (c 0.50; $CHCl_3$). ^{1}H NMR (D_2O, 400 MHz) δ: 4.74 (d, 5 H, $J_{1,2}$ = 7.1 Hz, H-1$^{B\text{-}F}$), 4.46 (d, 1 H, $J_{1,2}$ = 7.9 Hz, H-1^A), 3.92–3.89 (m, 7 H), 3.79–3.65 (m, 11 H), 3.56–3.31 (m, 22 H), 1.66–1.60 (m, 4 H, H-2′,3′), 1.46–1.43 (m, 2 H, H-4′). ^{13}C NMR (D_2O, 100 MHz) δ: 103.48 (C-1^F), 103.19 (C-1$^{B\text{-}E}$), 102.60 (C-1^A), 85.20, 84.83, 84.66, 76.65, 76.28, 76.20, 74.40, 73.96, 73.59, 71.03 (C-1′), 70.22, 68.82, 68.73, 63.17 (C-6^A), 61.32 (C-6$^{B\text{-}F}$), 51.73 (C-5′), 28.92 (C-2′), 28.37 (C-3′), 23.06 (C-4′). FT-IR: 2098.29 cm^{-1} (N_3). ESI HR-MS ($C_{41}H_{71}N_3O_{31}$): *m/z* = found ([*M*+Na]$^+$ 1102.4145; calc 1102.4150).

3.2 General procedure for click chemistry with modCRM$_{197}$ 4 (Scheme 4)

In a typical experiment, to a solution of modCRM$_{197}$ **4** (300 µg, 0.005 µmol) in 100 mM sodium phosphate pH 7.0 (70 µL) and azide **5** (0.10 µmol), a premixed solution of 5 mM $CuSO_4 \cdot 5H_2O$ (5 µL) and 25 mM THPTA (5 µL) was added under nitrogen atmosphere, followed by 5 mM aminoguanidine hydrochloride (5 µL) and 10 mM sodium ascorbate (5 µL). The mixture was stirred at ambient temperature for 1.5 h, at which time the glycoprotein was washed on a 30 kDa Amicon centrifugal filter with 10 mM EDTA/10 mM sodium phosphate pH 7.0 (2 × 100 µL) and 10 mM sodium phosphate pH 7.0 (8 × 100 µL), and subsequently reconstituted with 10 mM sodium phosphate (pH 7.0). Yield (recovered glycoprotein): 85%.

Scheme 7. Conjugation of **4** and **5**



The glycoconjugate loading was determined by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS; UltraFlex III MALDI-TOF/TOF instrument, Bruker Daltonics) in linear mode and with positive ion detection. The samples for analysis were prepared by mixing 2.5 µL of product and 2.5 µL of Super DHB matrix; 2.5 µL of each mixture was deposited on a samples plate, dried at room temperature for 10 min, and subjected to the spectrometer. A loading of 3.5 sugar moieties for determined for **6**. Conjugation efficiency (carbohydrate moieties used for reaction/conjugated) was 17.5%.

For SDS page analysis, the samples (5 µg) were electrophoresed on a 7% TrisAcetate gel or 4-12% Bis-Tris gel (NuPage, Invitrogen) and stained with Coomassie blue.

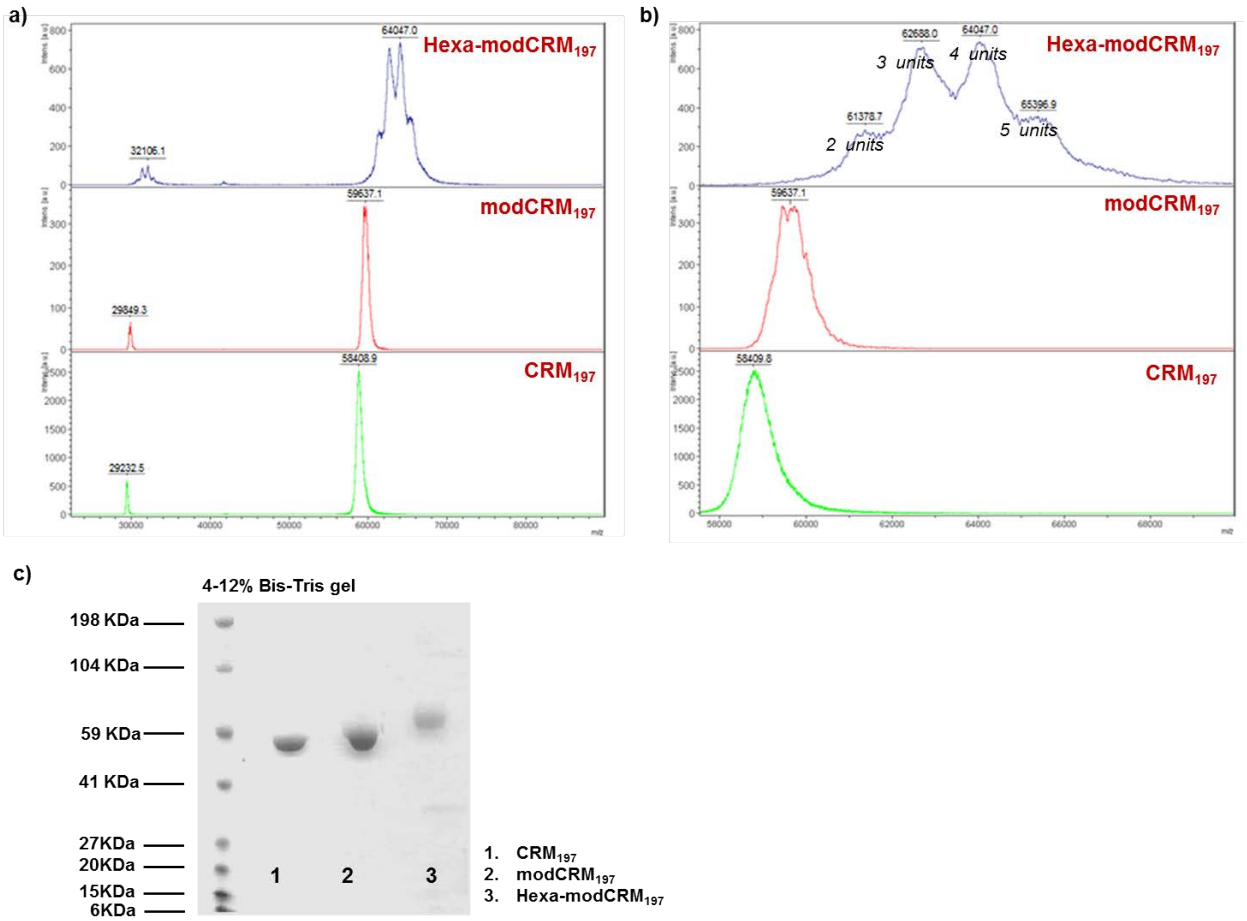
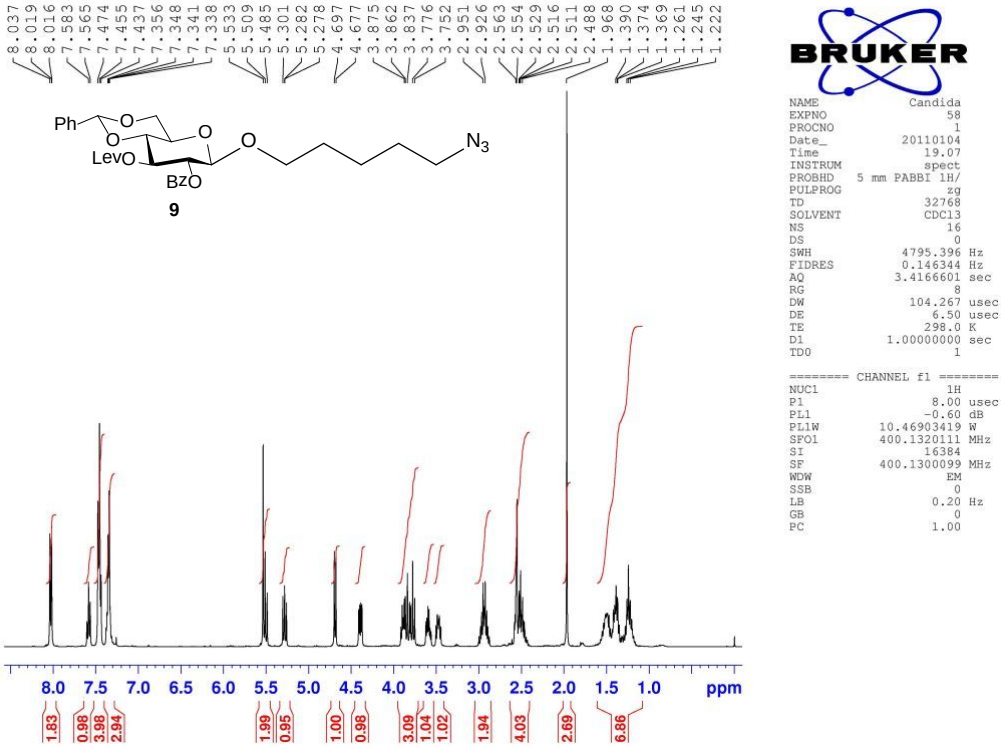


Figure 1. a) MALDI TOF spectra after conjugation of compound **5** to modCRM$_{197}$ **4**; b) zoom of MALDI TOF spectra; c) SDS page on 4-12% Bis-Tris gel stained with Coomassie blue.

3.3 NMR spectra of carbohydrate compounds

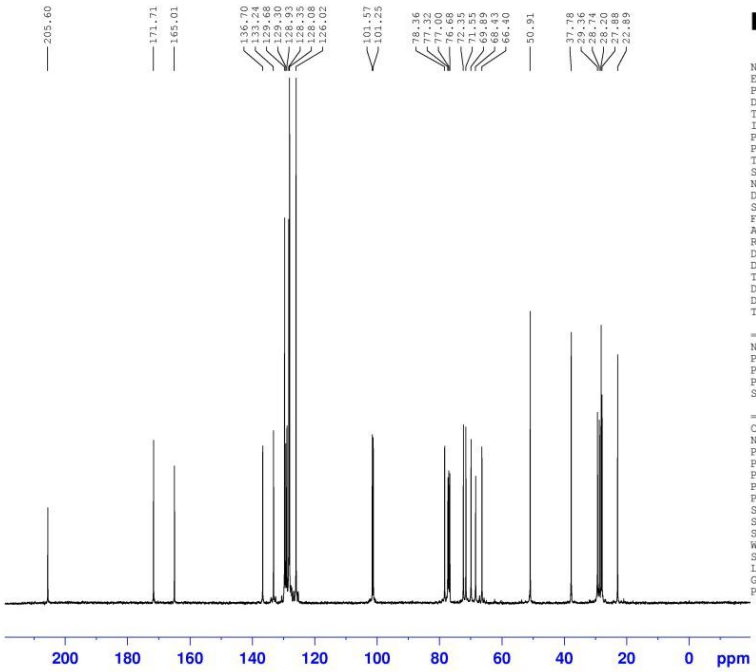
^{1}H NMR ($CDCl_3$, 400 MHz) of compound 9



BRUKER

NAME	Candida
EXPNO	58
PROCNO	1
Date_	20110104
Time	19.07
INSTRUM	spect
PROBHD	5 mm PABBI 1H/
PULPROG	zg
TD	32768
SOLVENT	CDCl3
NS	16
DS	0
SWH	4795.396 Hz
FIDRES	0.146344 Hz
AQ	3.4166601 sec
RG	8
DW	104.267 usec
DE	6.50 usec
TE	298.0 K
D1	1.00000000 sec
TD0	1
======== CHANNEL f1 ========	
NUC1	1H
P1	8.00 usec
PL1	-0.60 dB
PL1W	10.46903419 W
SFO1	400.1320111 MHz
SI	16384
SF	400.1300099 MHz
WDW	EM
SSB	0
LB	0.20 Hz
GB	0
PC	1.00

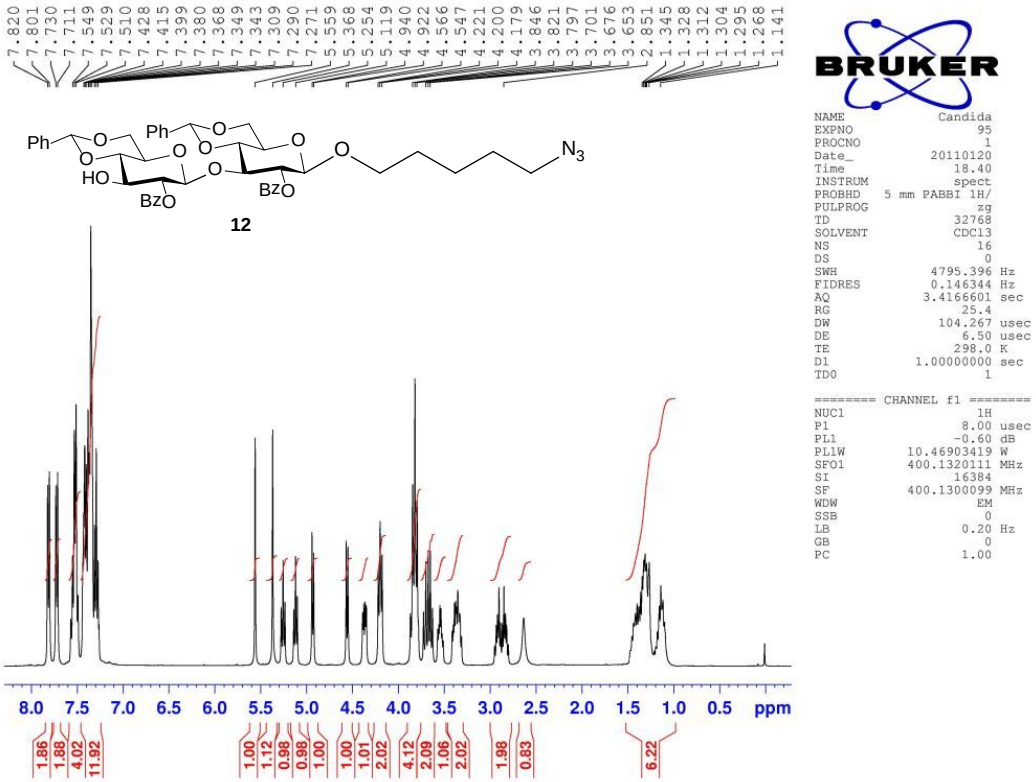
^{13}C NMR ($CDCl_3$, 100 MHz) of compound 9



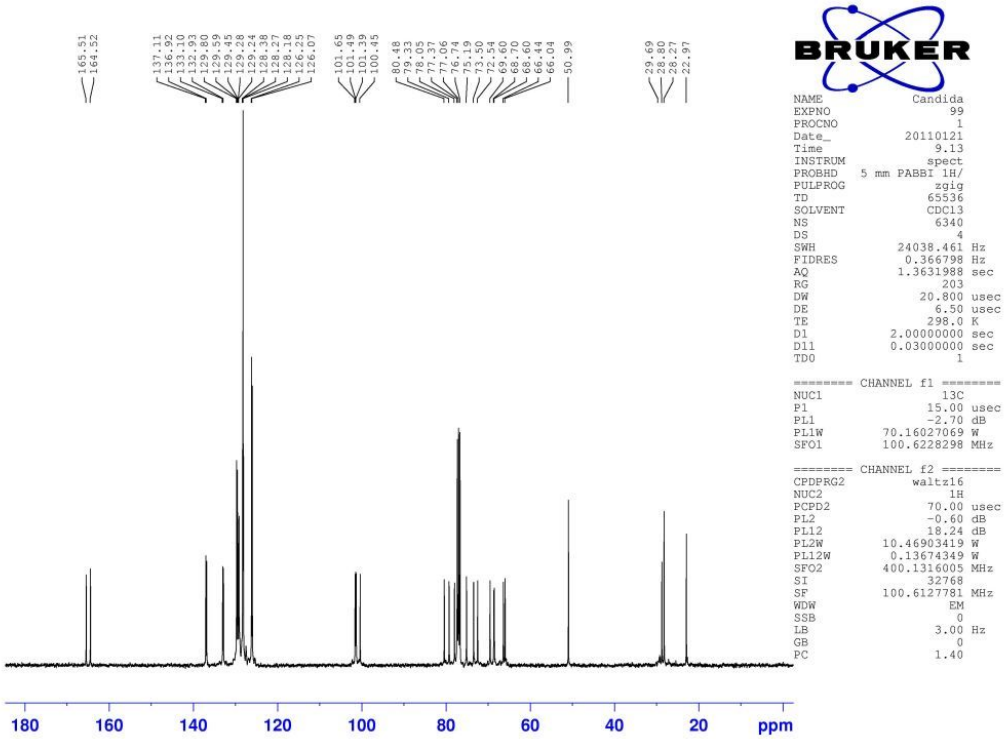
BRUKER

NAME	Candida
EXPNO	61
PROCNO	1
Date_	20110105
Time	8.24
INSTRUM	spect
PROBHD	5 mm PABBI 1H/
PULPROG	zgig
TD	65536
SOLVENT	CDCl3
NS	6317
DS	4
SWH	24038.461 Hz
FIDRES	0.366798 Hz
AQ	1.3631988 sec
RG	203
DW	20.800 usec
DE	6.50 usec
TE	298.0 K
D1	2.00000000 sec
D11	0.03000000 sec
TD0	1
======== CHANNEL f1 ========	
NUC1	13C
P1	15.00 usec
PL1	-2.70 dB
PL1W	70.16027069 W
SFO1	100.6228298 MHz
======== CHANNEL f2 ========	
CPDPRG2	waltz16
NUC2	1H
PCPD2	70.00 usec
PL2	-0.60 dB
PL12	18.24 dB
PL2W	10.46903419 W
PL12W	0.13674349 W
SFO2	400.1316005 MHz
SI	32768
SF	100.6127845 MHz
WDW	EM
SSB	0
LB	3.00 Hz
GB	0
PC	1.40

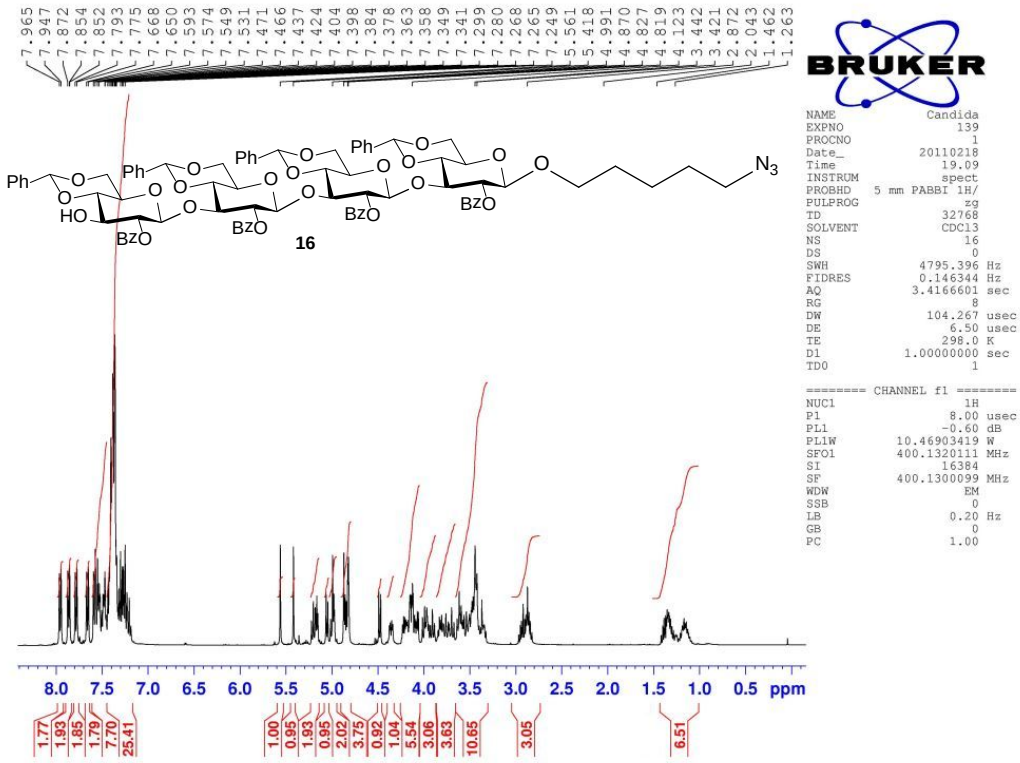
^{1}H NMR ($CDCl_3$, 400 MHz) of compound 12



^{13}C NMR ($CDCl_3$, 100 MHz) of compound 12

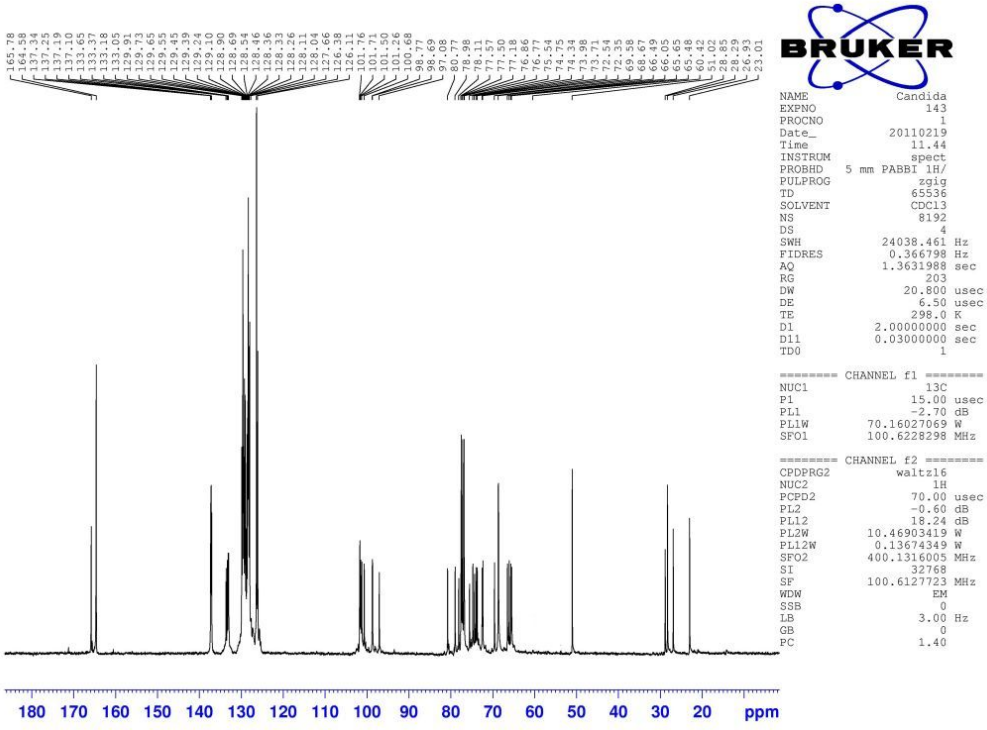


^{1}H NMR ($CDCl_3$, 400 MHz) of compound 16



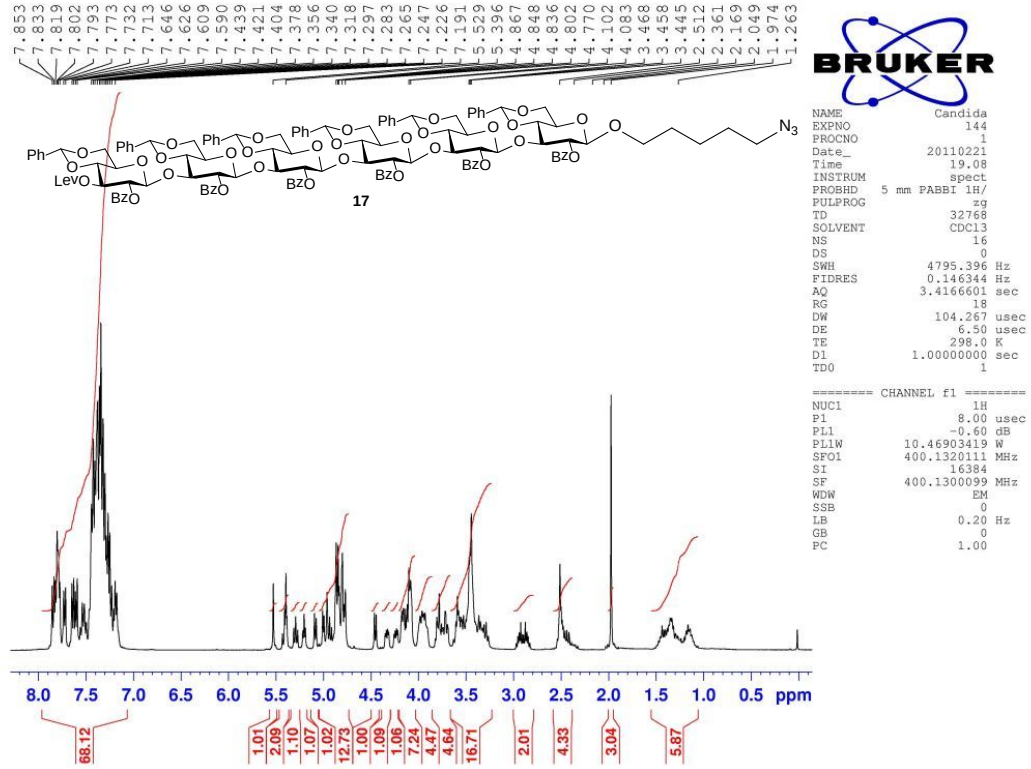

```
NAME                 Candida
EXPNO                    139
PROCNO                     1
Date_             20110218
Time                  19.09
INSTRUM                spect
PROBHD   5 mm PABBI 1H/
PULPROG                  zg
TD                    32768
SOLVENT               CDCl3
NS                       16
DS                        0
SWH                4795.396 Hz
FIDRES             0.146344 Hz
AQ                 3.4166601 sec
RG                        8
DW                  104.267 usec
DE                    6.50 usec
TE                  298.0 K
D1                 1.00000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                      1H
P1                     8.00 usec
PL1                   -0.60 dB
PL1W              10.46903419 W
SFO1              400.1320111 MHz
SI                      16384
SF                400.1300099 MHz
WDW                       EM
SSB                        0
LB                      0.20 Hz
GB                         0
PC                      1.00
```

^{13}C NMR ($CDCl_3$, 100 MHz) of compound 16

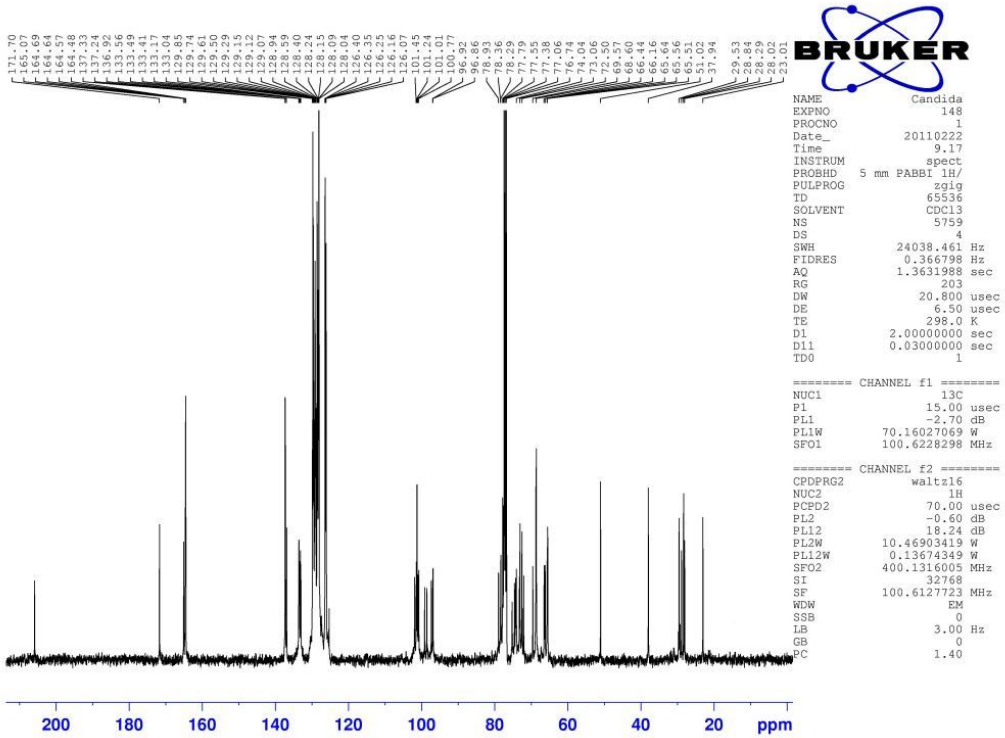



```
NAME                 Candida
EXPNO                    143
PROCNO                     1
Date_             20110219
Time                  11.44
INSTRUM                spect
PROBHD   5 mm PABBI 1H/
PULPROG                zgig
TD                    65536
SOLVENT               CDCl3
NS                     8192
DS                        4
SWH               24038.461 Hz
FIDRES             0.366798 Hz
AQ                 1.3631988 sec
RG                      203
DW                  20.800 usec
DE                    6.50 usec
TE                  298.0 K
D1                 2.00000000 sec
D11                0.03000000 sec
TD0                        1

======== CHANNEL f1 ========
NUC1                     13C
P1                    15.00 usec
PL1                   -2.70 dB
PL1W              70.16027069 W
SFO1              100.6228298 MHz

======== CHANNEL f2 ========
CPDPRG2              waltz16
NUC2                      1H
PCPD2                 70.00 usec
PL2                   -0.60 dB
PL12                  18.24 dB
PL2W              10.46903419 W
PL12W              0.13674349 W
SFO2              400.1316005 MHz
SI                      32768
SF                100.6127723 MHz
WDW                       EM
SSB                        0
LB                      3.00 Hz
GB                         0
PC                      1.40
```

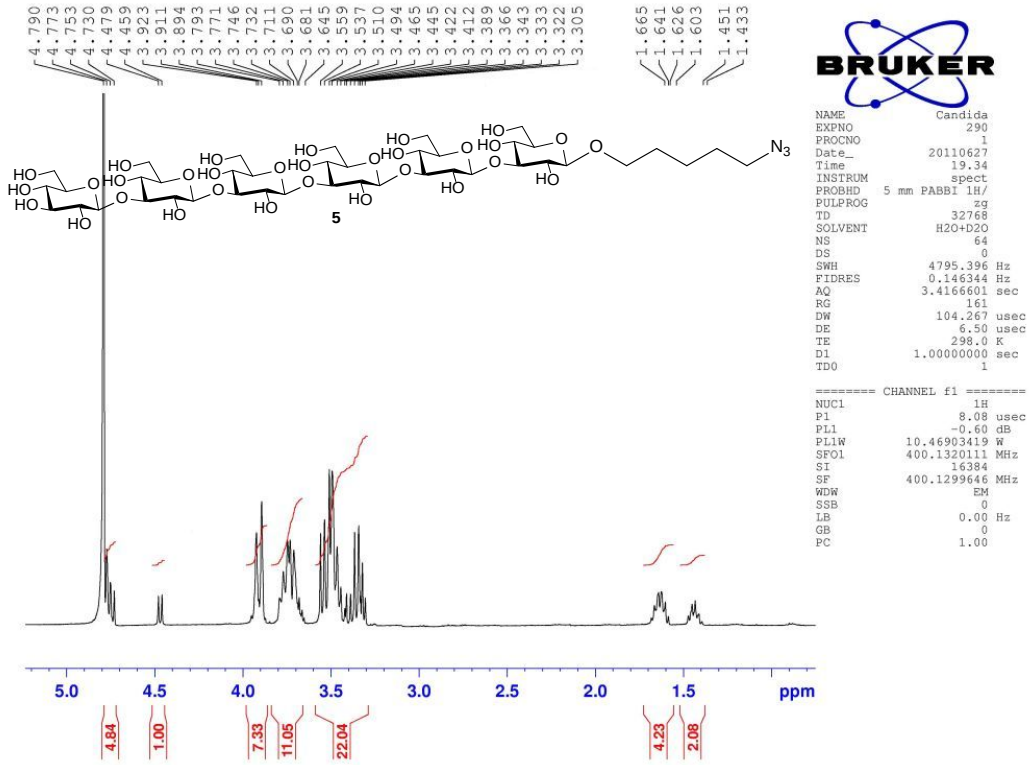
^{1}H NMR ($CDCl_3$, 400 MHz) of compound 17



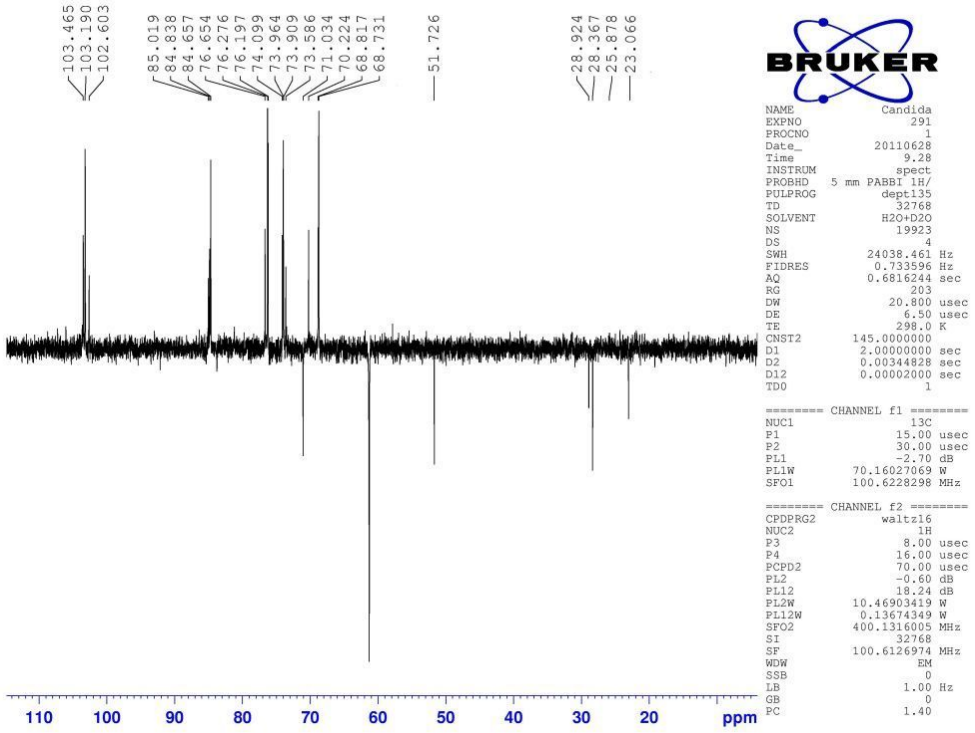
^{13}C NMR ($CDCl_3$, 100 MHz) of compound 17



^{1}H NMR (D_2O, 400 MHz) of compound 5



Dept 135 (D_2O, 100 MHz) of compound 5



[1] Piron, Flavia; Oprea, Cornelia; Cismaş, Crina; Terec, Anamaria; Roncali, Jean; Grosu, Ion; Synthesis of Podands with Cyanurate or Isocyanurate Cores and Terminal Triple Bonds, *Synthesis* **2010**, *10*, 1639-1644.

[2] Ghorbani-Choghamarani, Arash; Chenani, Zahra and Mallakpour, Shadpour ; Supported Nitric Acid on Silica Gel and Polyvinyl Pyrrolidone (PVP) as an Efficient Oxidizing Agent for the Oxidation of Urazoles and Bis-urazoles, *Synthetic Communications* **2009**, *39:23*, 4264 – 4270.

[3]Adamo, Roberto; Tontini, Marta; Brogioni, Giulia; Romano, Maria Rosaria; Costantini, Gabriele ; Danieli, Elisa; Proietti, Daniela; Berti, Francesco; Costantino, Paolo *J. Carbohydr. Chem.* **2011,** *30,* 249–280.