Supporting Information for

Chemoenzymatic Synthesis of Glycosylphosphatidylinositol-Anchored Glycopeptides

Zhimeng Wu, Xueqing Guo, and Zhongwu Guo*

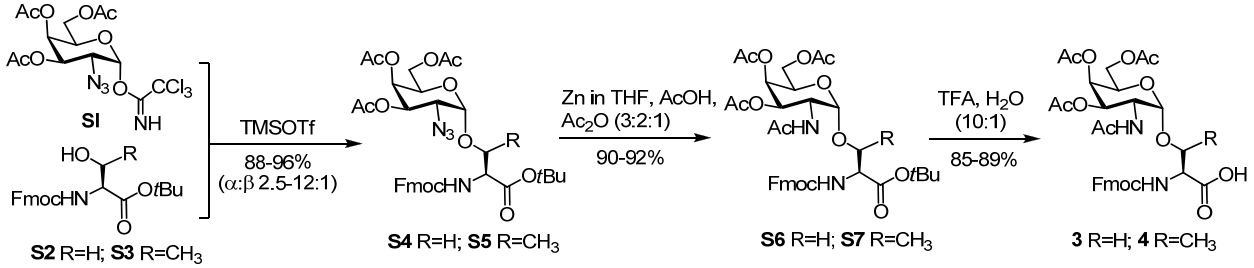
Department of Chemistry, Wayne State University, 5101 Cass Avenue, Detroit 48202, USA

Table of Contents:

1. General Experimental Methods

NMR spectra were recorded at 400 MHz. Proton chemical shifts are reported in ppm (δ) downfield from tetramethylsilane (TMS) or in reference to the solvent DHO (δ 4.79). Carbon chemical shifts are reported in ppm (δ) in reference to $CDCl_3$ (δ 77.16). Coupling constants (*J*) are reported in hertz (Hz). ESI MS spectra were obtained on a Waters' ZQ2000 single quadrupole mass spectrometer, and MALDI-TOF MS was obtained on a Bruker Ultraflex mass spectrometer. Thin layer chromatography (TLC) was performed on silica gel GF_{254} plates and detections were achieved by charring with phosphomolibdic acid/EtOH or 1% H_2SO_4/EtOH solutions. Molecular sieves were dried under high vacuum at 170-180 ^oC for 6 to 10 h immediately before use. Commercial reagents were used without further purification unless otherwise mentioned.

2. Synthesis of Glycosyl Amino Acids 3 and 4



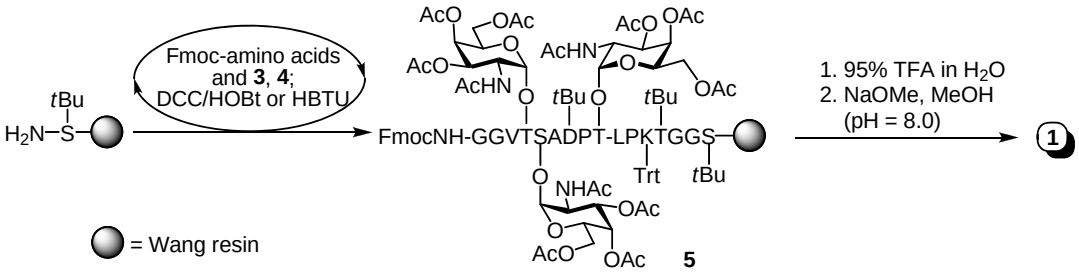
***N*-(Fluoren-9-ylmethoxycarbonyl)-*O*-(3,4,6-tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactopyranosyl-L-serine and threonine tert-butyl ester (S4 and S5):** Trichloroacetimidate **S1** (2.14g, 4.50 mmol), which was prepared according to a reported procedure,[1-3] and serine derivative **S2** (1.06 g, 2.78 mmol) or threonine derivative **S3**[4] (1.18 g, 3.00 mmol) were dissolved in a mixture of CH_2Cl_2 and Et_2O (1:1, 30 mL). After the mixture was cooled to -30°C, TMSOTf (90 μL, 0.45 mmol) was added under an Argon atmosphere. The mixture was stirred at -30°C for 30 min, and TLC showed that the reaction was completed. Hunig's base was then added to quench the reaction. The solution was diluted with CH_2Cl_2 (100 mL) and washed with 0.1 M aq. HCl solution. After dried over $MgSO_4$, the organic layer was concentrated, and the residue was purified by a silica gel column (EtOAc:Hexane 3:7) to give **S4** (1.2 g, 62%) and its β anomer (0.5 g, 25%) or **S5** (1.9 g, 89%) and its β anomer (0.15 g, 7%). The spectral data of **S4** and **S5** were identical to that reported in the literature.[5]

N-(Fluoren-9-ylmethoxycarbonyl)-*O*-(3,4,6-tri-*O*-acetyl-2-acetamido-2-deoxy-α-D-galactopyranosyl-L-serine and threonine tert-butyl ester (S6 and S7):[5] After compound **S4** (1.20 g, 1.72 mmol) or **S5** (1.90 g, 2.68 mmol) was dissolved in a mixture of THF, HOAc and acetic anhydride (3:2:1, 24 mL), activated zinc dust (2.00 g, 31.25 mmol) was added, and the mixture was stirred at rt until the starting material disappeared as shown by TLC. The zinc dust was then filtered off and washed with CH_2Cl_2 (3 x 10 mL). The solution was combined and washed with saturated aq. $NaHCO_3$ and dried over $MgSO_4$. After evaporation of the solvent under vacuum, the residue was purified by a silica gel column (CH_2Cl_2:MeOH=15:1) to give **S6** (1.10 g, 90%) or **S7** (1.79 g, 92%). **S6**: ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.75 (d, J = 8.0 Hz, 2 H), 7.60 (d, J = 6.4 Hz, 2 H), 7.39 (t, J = 7.6 Hz, 2 H), 7.31 (t, J = 7.6 Hz, 2 H), 5.76 (d, J = 8.4 Hz, 1 H), 5.70 (d, J = 8.8 Hz, 1 H), 5.36 (d, J = 3.2 Hz, 1 H), 5.10 (dd, J = 11.2, 2.4 Hz, 1 H), 4.83 (s, 1 H), 4.58 (m, 1 H), 4.40 (m, 2 H), 4.23 (t, J = 7.6 Hz, 1 H), 4.14-4.03 (m, 4 H), 3.95 (d, J = 9.6 Hz, 1 H), 3.84 (d, J = 8.4 Hz, 1 H), 2.15 (s, 3 H), 2.02 (s, 3 H), 1.98 (s, 3 H), 1.97 (s, 3 H), 1.48 (s, 9 H). **S7**: ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.77 (d, J = 8.4 Hz, 2 H), 7.63 (d, J = 6.4 Hz, 2 H), 7.40 (t, J = 7.6 Hz, 2 H), 7.33 (t, J = 7.6 Hz, 2 H), 6.00 (d, J = 10.0 Hz, 1 H), 5.60 (d, J = 10.0 Hz, 1 H), 5.38 (s,1 H), 5.09 (dd, J = 11.2, 3.2 Hz, 1 H), 4.88 (d, J = 3.2 Hz, 1 H), 4.64-4.58 (m, 1 H), 4.48-4.45 (m, 2 H), 4.28-4.19 (m, 4 H), 4.12-4.05 (m, 3 H), 2.16 (s, 3 H), 2.03 (s, 3 H), 2.00 (s, 3 H), 1.99 (s, 3 H), 1.46 (s, 9 H), 1.31 (d, J = 6.4 Hz, 3 H).

N-(Fluoren-9-ylmethoxycarbonyl)-*O*-(3,4,6-tri-*O*-acetyl-2-acetamido-2-deoxy-α-D-galactopyranosyl-L-serine and threonine (3 and 4):[5] After the solution of **S6** (1.10 g, 1.55 mmol) or **S7** (1.79 g, 2.47 mmol) in TFA and H_2O (10:1) was stirred at rt for 2h, the solvent was removed in vacuo. The

residue was coevaporated with toluene, and the crude product was purified by a silica gel column (Toluene:Ethanol 5:1) to give **3** (1.0 g, 98%) or **4** (1.4 g, 85%). **3:** ^{1}H NMR (DMSO-d_6, 400 MHz): δ 7.87 (d, J = 7.2 Hz, 2 H), 7.69 (d, J = 8.4 Hz, 2 H), 7.39 (t, J = 7.2 Hz, 2 H), 7.31 (t, J = 7.2 Hz, 2 H), 5.28 (d, J = 2.4 Hz, 1 H), 5.01 (dd, J = 3.2, 12 Hz, 1 H), 4.85 (d, J = 3.2 Hz, 1 H), 4.31 (d, J = 6.4 Hz, 2 H), 4.23-4.19 (m, 2 H), 4.18-4.15 (m, 2 H), 4.04-3.95 (m, 2 H), 3.83-3.74 (m, 2 H), 2.07 (s, 3 H), 1.92 (s, 3 H), 1.88 (s, 3 H), 1.79 (s, 3 H). **4:** ^{1}H NMR (DMSO-d_6, 400 MHz): δ 7.88 (d, J = 7.2 Hz, 2 H), 7.73-7.71 (m, 2 H), 7.63 (d, J = 9.6 Hz, 1 H), 7.58 (d, J = 10.4 Hz, 1 H), 7.42 (t, J = 7.2 Hz, 2 H), 7.31 (t, J = 7.2 Hz, 2 H), 5.28 (d, J = 2.4, Hz, 1 H), 5.00 (dd, J = 3.2, 12 Hz, 1 H), 4.78 (d, J = 3.2 Hz, 1 H), 4.48-4.37 (m, 2 H), 4.29-4.26 (m, 2 H), 4.22-4.11 (m, 3 H), 4.00 (d, J = 6.4 Hz, 2 H), 2.09 (s, 3 H), 1.99 (s, 3 H), 1.88 (s, 3 H), 1.81 (s, 3 H), 1.14 (d, J = 6.4 Hz, 3 H).

3. Solid-Phase Synthesis of Glycopeptide 1



The glycopeptide synthesis started from a commercial Wang resin loaded with a serine on the 0.1 mmol scale with the peptide chain elongated essentially on an automatic peptide synthesizer using amino acids Fmoc-Gly-OH, Fmoc-Thr(*t*Bu)-OH, Fmoc-Lys(Trt)-OH, Fmoc-Pro-OH, Fmoc-Leu-OH, Fmoc-Asp(*t*Bu)-OH, Fmoc-Ala-OH, Fmoc-Val-OH as building blocks. The synthetic protocols were: Fmoc deprotection was achieved with 20% piperidine in NMP and coupling reactions were achieved with 5 equiv. of amino acids using HBTU/DIPEA (5 equiv.) as the condensation reagents. Glycosyl amino acids **3** or **4** were introduced manually. Again, 20% piperidine was used to deprotect Fmoc. The coupling reactions were carried out in DMF with HBTU, HOBt and N-methylmorpholine as the condensation reagents and using 2.5 equiv. of **3** or **4**, while the reactions were kept at rt overnight. After the glycopeptide assembly on the resin was finished, the resin was treated with 95% aq. TFA containing 2.5% Et_3SiH at rt for 2h to release the glycopeptide, which was then characterized with MALDI–TOF MS (Figure 1) [calcd for $C_{117}H_{168}N_{20}O_{50}$ (m/z) 2653.1, observed 2654.7 $(M + H)^+$]. The resultant glycopeptide was dissolved in MeOH, to which was added 0.1 M solution of NaOMe in MeOH to adjust the pH to *ca*. 8.0. The reaction progress was monitored by MALDI-TOF. After the reaction was finished, it was neutralized by addition of acetic acid. The solvent was removed in

vaccuo, and the residue was dissolved in water and purified by preparative RP-HPLC (20% to 50% MeCN in H_2O, both containing 0.1% TFA) to give **1** as a white solid, which was characterized with MALDI-TOF MS (Figure 2) [calcd for $C_{99}H_{150}N_{20}O_{41}$ (m/z) 2275.1, observed 2276.7 $(M + H)^+$].

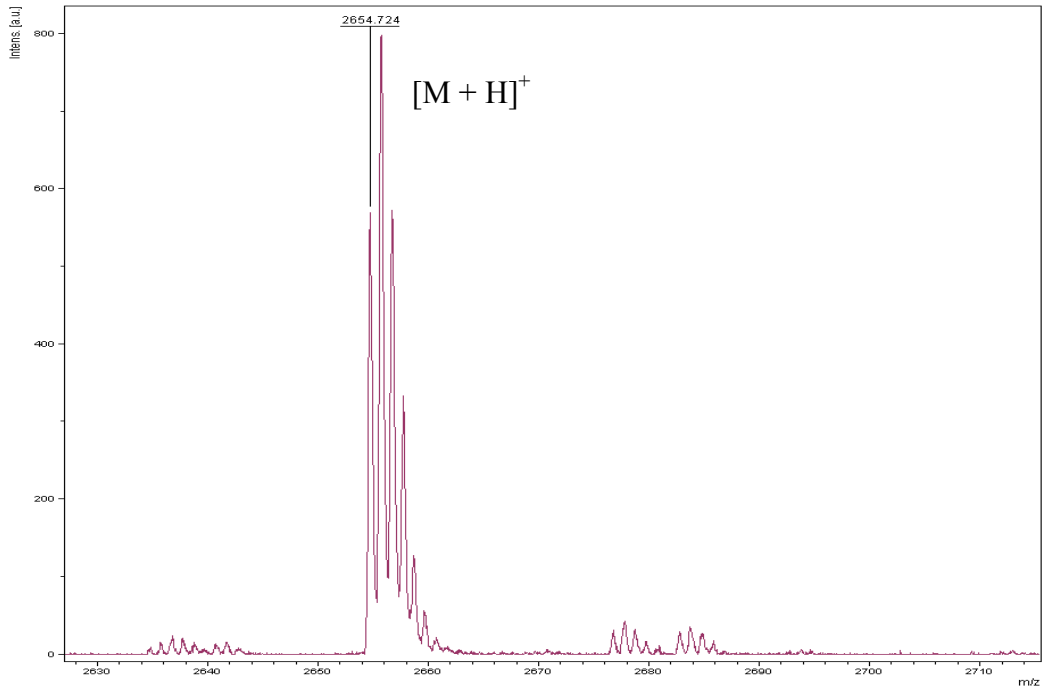


Figure 1. MALDI-TOF MS of the released glycopeptide before deacetylation

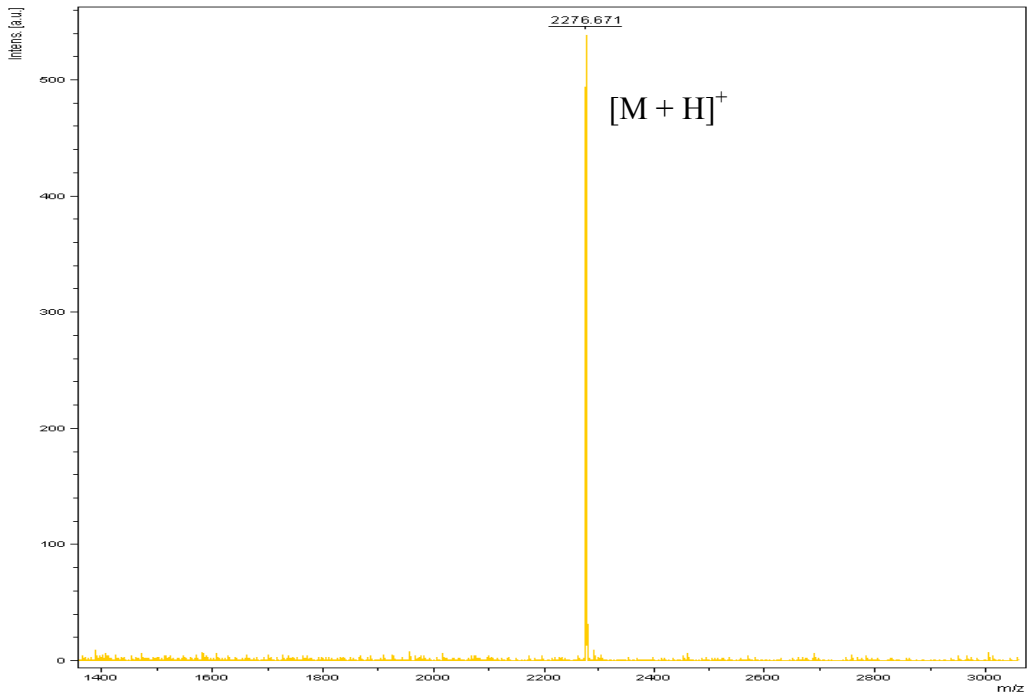


Figure 2. MALDI-TOF MS of glycopeptide **1**

4. Enzyme Reactions

Reactions of glycopeptide 1 with GPI analogues 6, 7 and 8. Glycopeptide **1** (0.5 mM), a GPI analogue **6**, **7** or **8** (2.5 mM) and SrtA (12.6 μM) were mixed in 0.3 M of Tris-HCl buffer (pH = *ca.* 7.5) containing 0.15 M of NaCl, 5 mM of $CaCl_2$, and 0.2 mM of mercaptoethanol. The reaction mixtures were incubated at 37 ^{o}C with the reaction progress monitored by RP HPLC and MALDI-TOF MS. After 24h of incubation, the reactions were quenched by addition of the same volume 0.1% TFA aq. solution, and the reaction mixtures were finally analyzed with RP HPLC (results shown in Figure 3) to determine reaction yields. The reaction products **9**, **10** and **11** were purified with HPLC and finally characterized with MALDI-TOF MS (see Figures 4, 5 and 6).

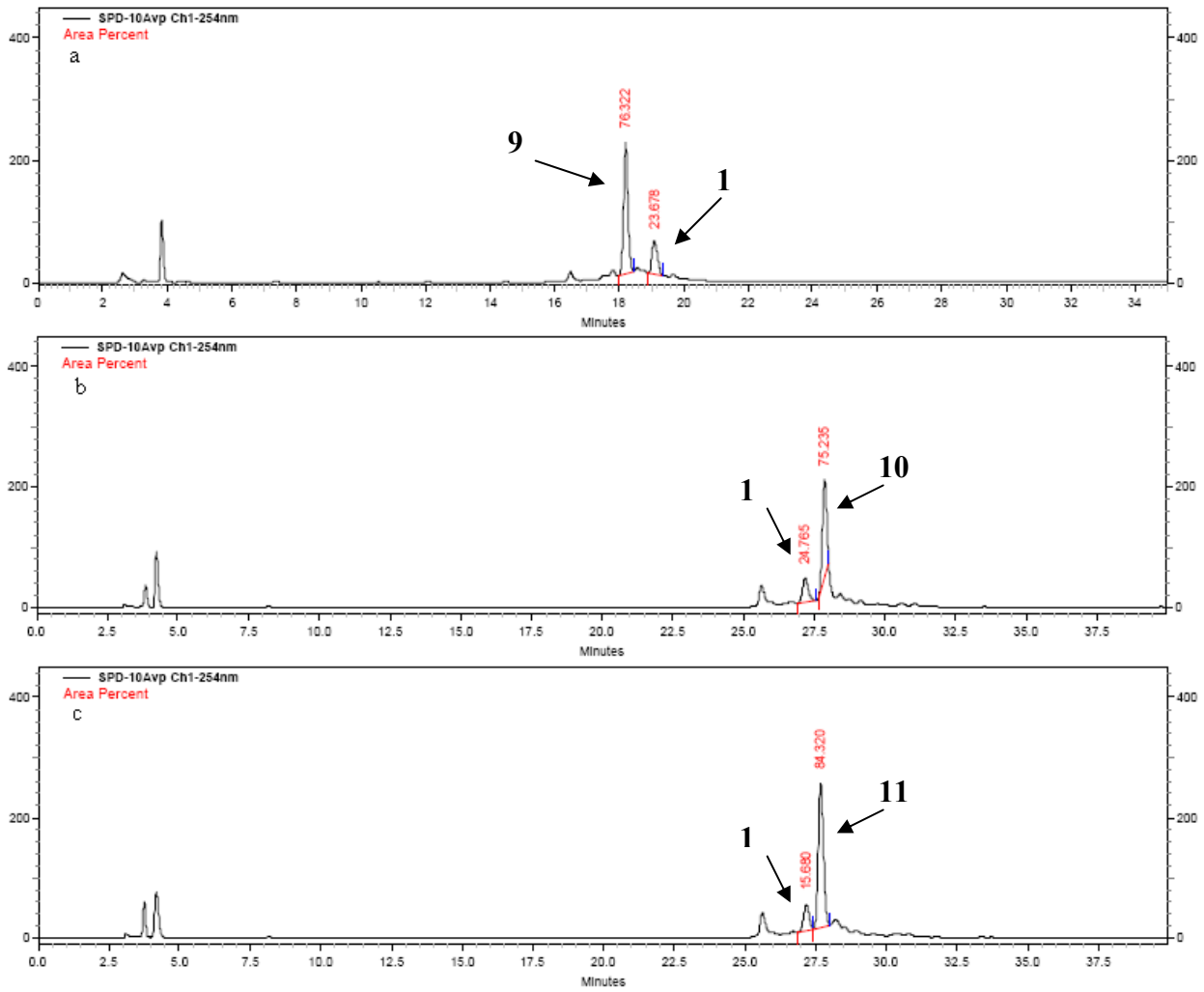


Figure 3: HPLC diagrams of the reactions between **1** and **6** (a), **7** (b) or **8** (c) in the presence of SrtA monitored with UV at 254 nm. HPLC conditions: eluent: (a) 20% to 50% MeCN in H_2O (both containing 0.1% TFA) in 30 min, (b and c) 20% MeCN for 5min then to 40% MeCN in H_2O (both containing 0.1% TFA) in 30 min; flow rate: 1 mL/min; column: C-18 (250 X 4.6 mm).

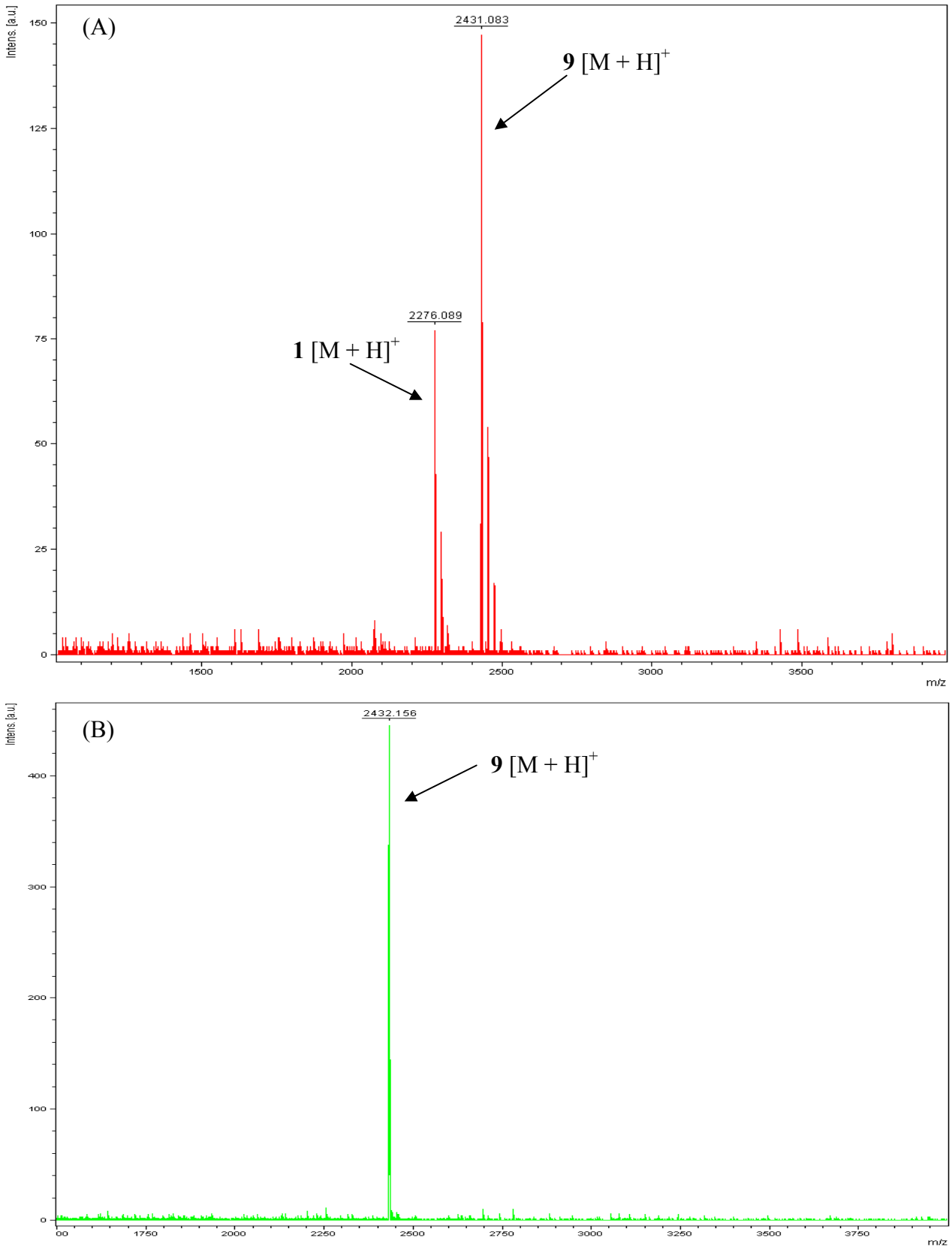


Figure 4. MALDI-TOF MS spectra of (A) the reaction mixture of **1** and **6** and (B) HPLC purified **9**: calcd for $C_{103}H_{160}N_{19}O_{46}P$ (m/z) 2431.4, observed 2432.2 $(M + H)^+$

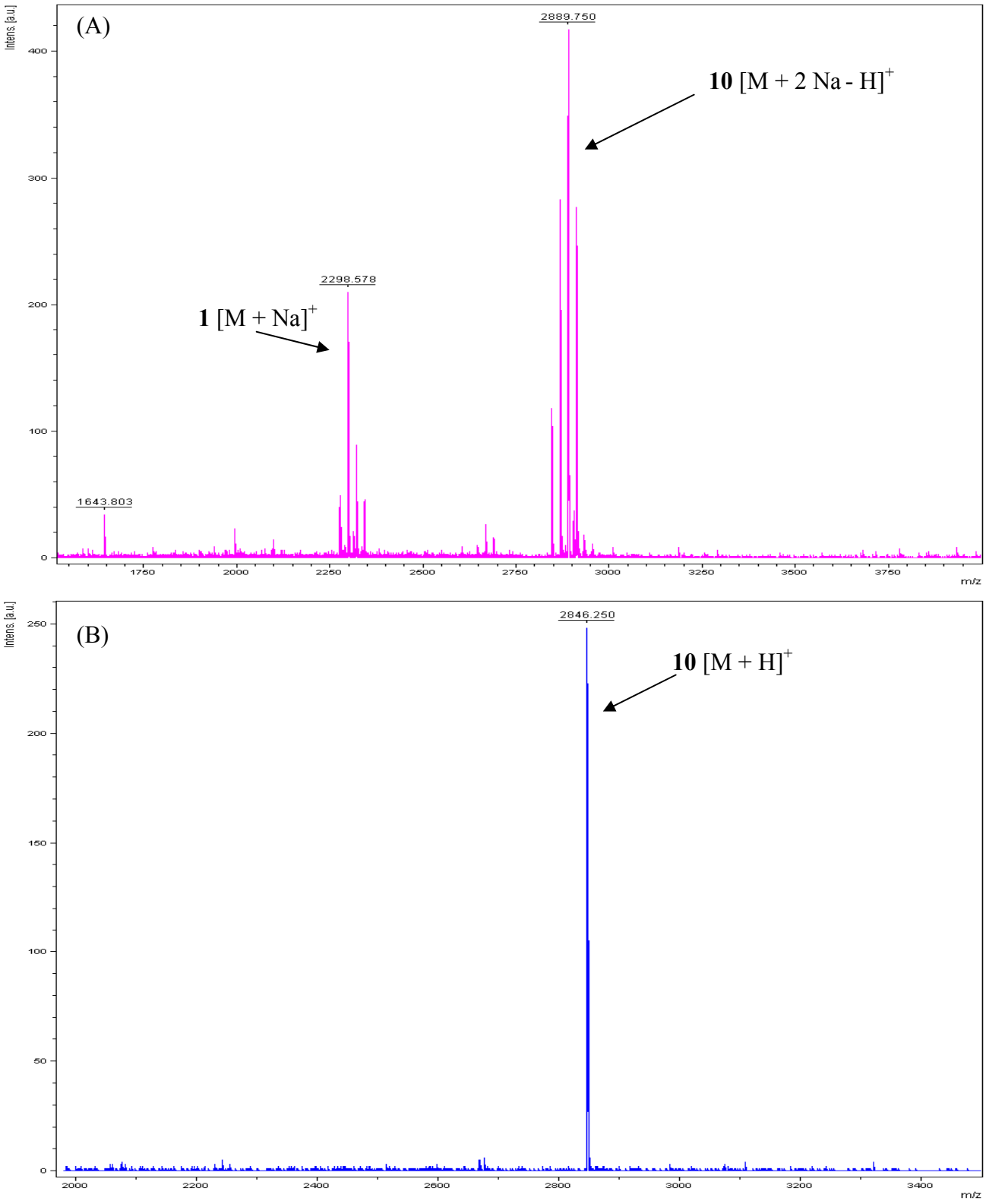


Figure 5. MALDI-TOF MS spectra of (A) the reaction mixture of **1** and **7** and (B) HPLC purified **10**: calcd for $C_{122}H_{186}N_{19}O_{56}P$ (m/z) 2845.8, observed 2846.3 $(M + H)^+$

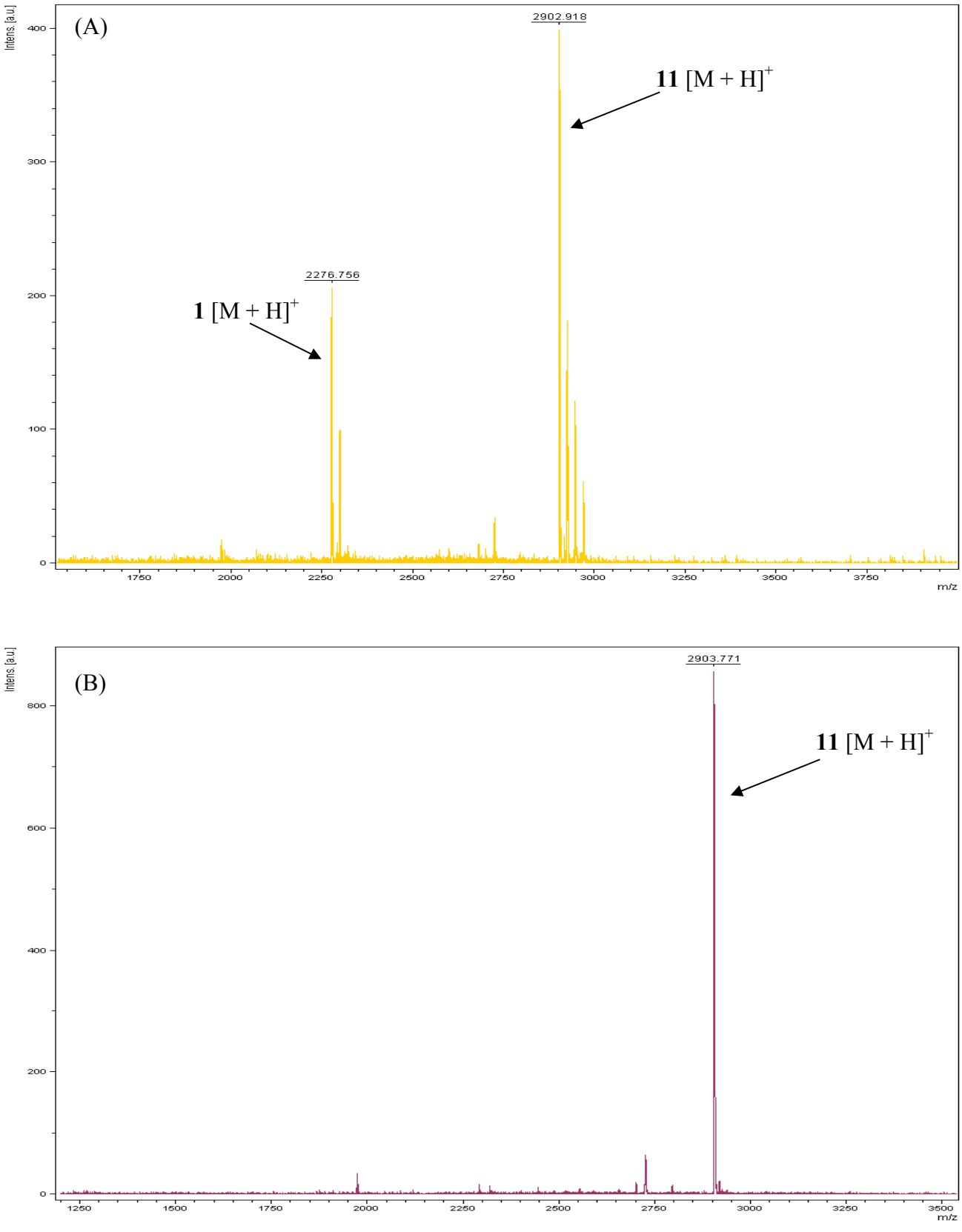


Figure 6. MALDI-TOF MS spectra of (A) the reaction mixture of **1** and **8** and (B) HPLC purified **11**: calcd for $C_{124}H_{189}N_{20}O_{57}P$ (m/z) 2902.9, observed 2903.8 $(M + H)^+$

Reaction of glycopeptide 1 with the de-Fmoc product of 9, which produced 12 and 13: After compound **9** (1 mg) was dissolved in methanol (0.5 mL), a catalytic amount of DBU (1 μL) was added, and the mixture was stirred at rt for 10 min to remove the N-terminal Fmoc group. After the reaction is finished as determined by MALDI-TOF MS, the de-Fmoc product was purified by RP HPLC and then reacted with glycopeptide **1** in the presence of SrtA. The enzymatic reaction of the de-Fmoc product of **9** with **1** was conducted under the same condition as described for the reactions between **1** and **6-8**. Thus, the de-Fmoc product of **9** was incubated with **1** and SrtA at 37 ^{o}C in the buffer (14.19 μL) mentioned above, and the reaction progress was monitored with HPLC. After 24h of incubation, the reaction was quenched by addition of the same volume 0.1% TFA, and the reaction mixture was analyzed by HPLC and MALDI-TOF MS with the results shown in Figure 7 and Figure 8. The reaction products **12** and **13** were finally purified by HPLC and characterized with MALDI-TOF MS (Figures 9 and 10).

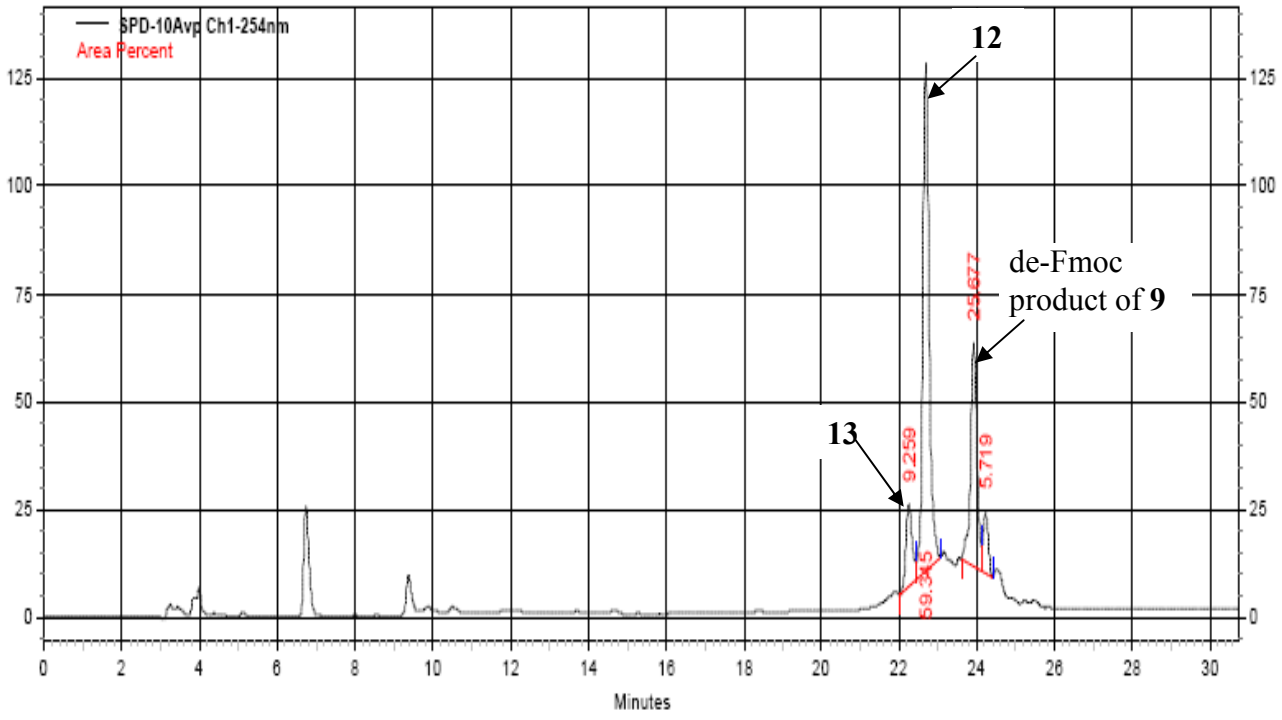


Figure 7: HPLC diagram of the reaction between **1** and the de-Fmoc product of **9** in the presence of SrtA monitored with UV 254 nm. HPLC conditions: eluent: 10% to 50% MeCN in H_2O (both containing 0.1% TFA) in 30 min; flow rate: 1 mL/min; column: C-18 (250X4.6 mm).

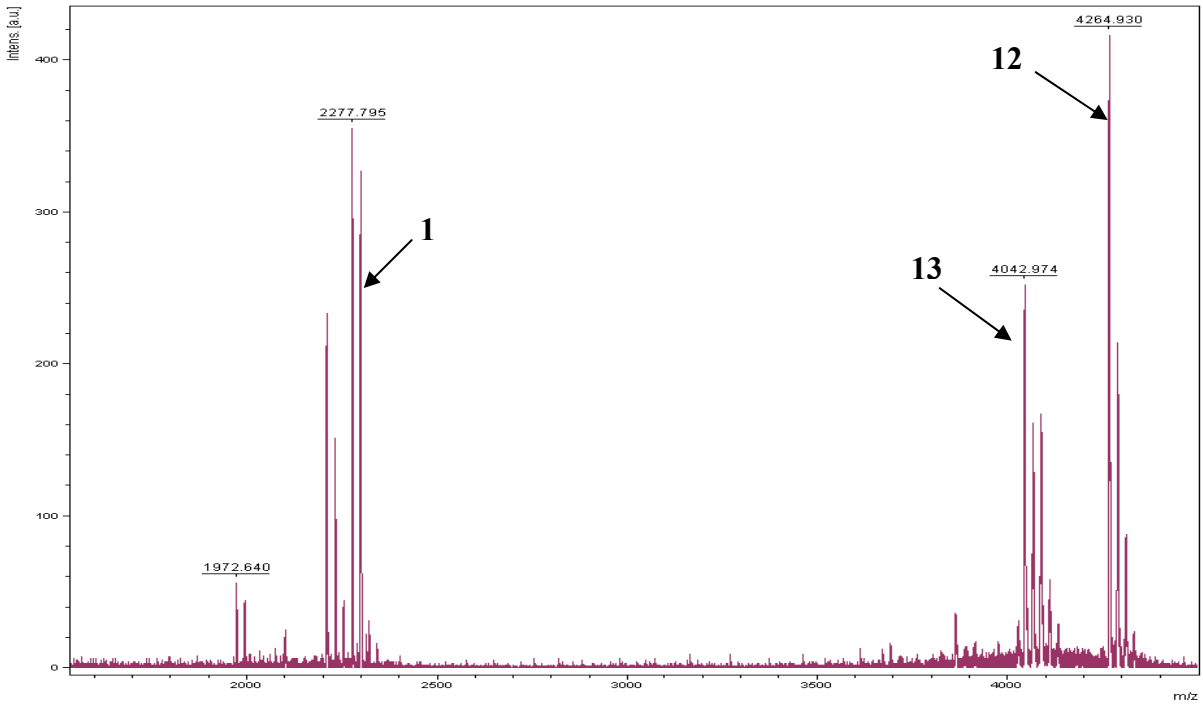


Figure 8. MALDI-TOF MS spectrum of the reaction mixture of **1** with the de-Fmoc product of **9**

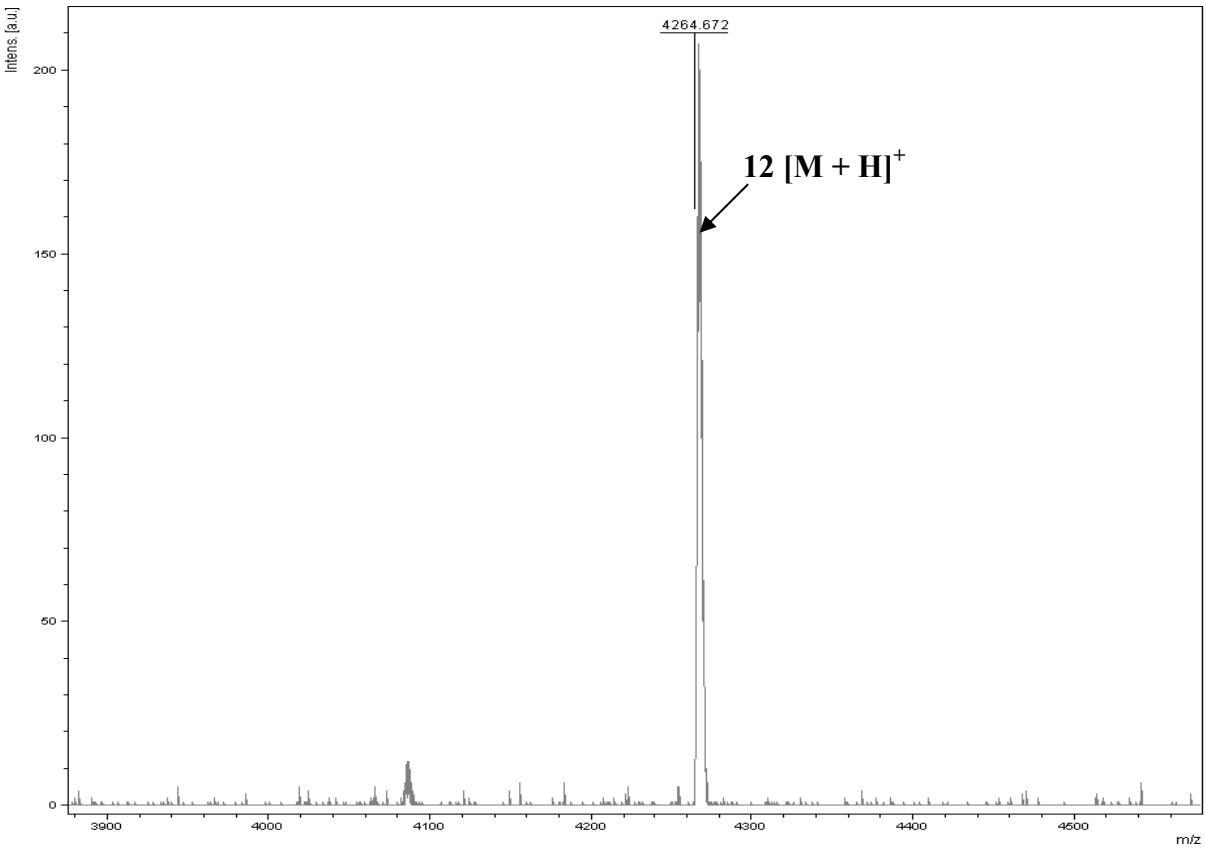


Figure 9. MALDI-TOF MS spectrum HPLC purified **12**: calcd for $C_{180}H_{287}N_{36}O_{80}P$ (m/z) 4263.9, observed 4264.7 $(M + H)^+$

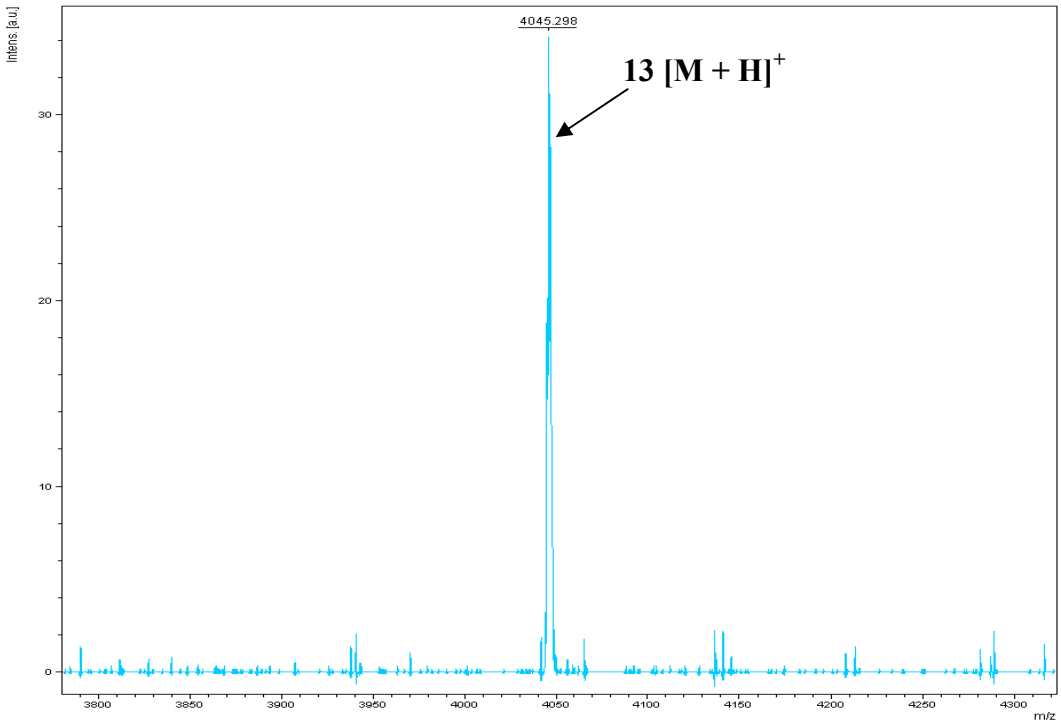
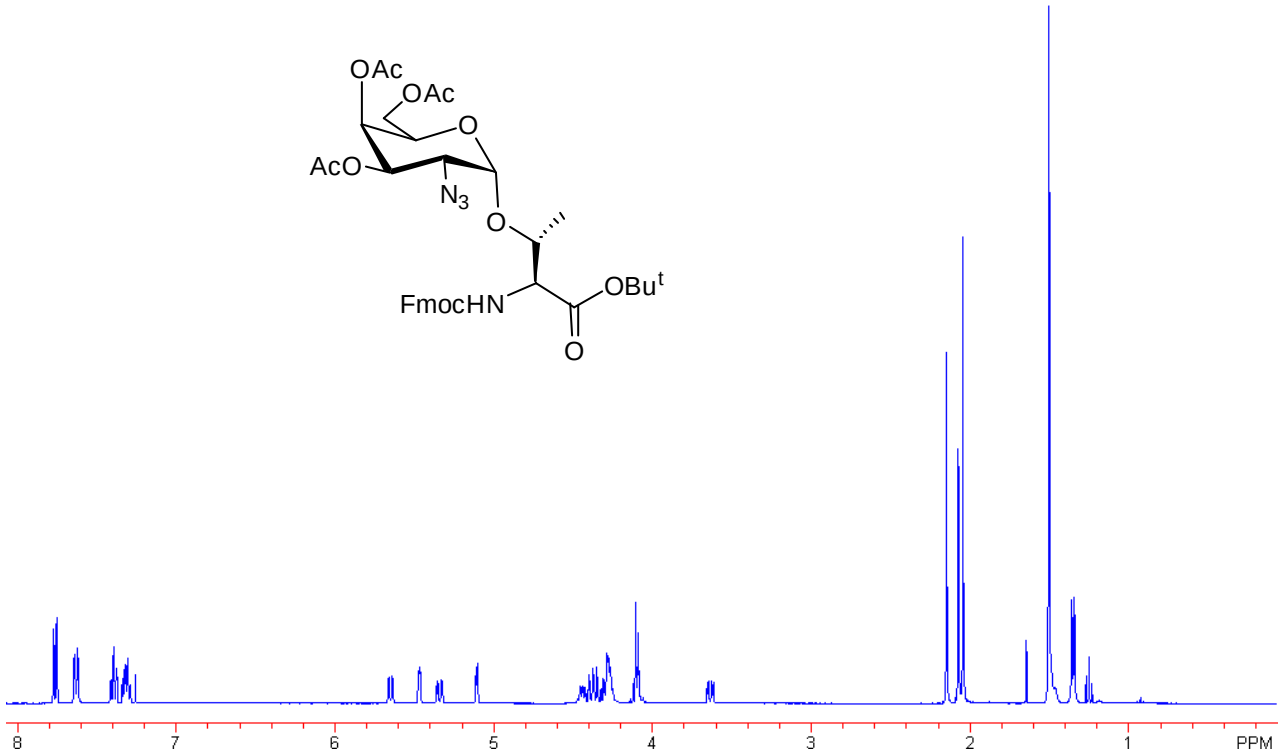
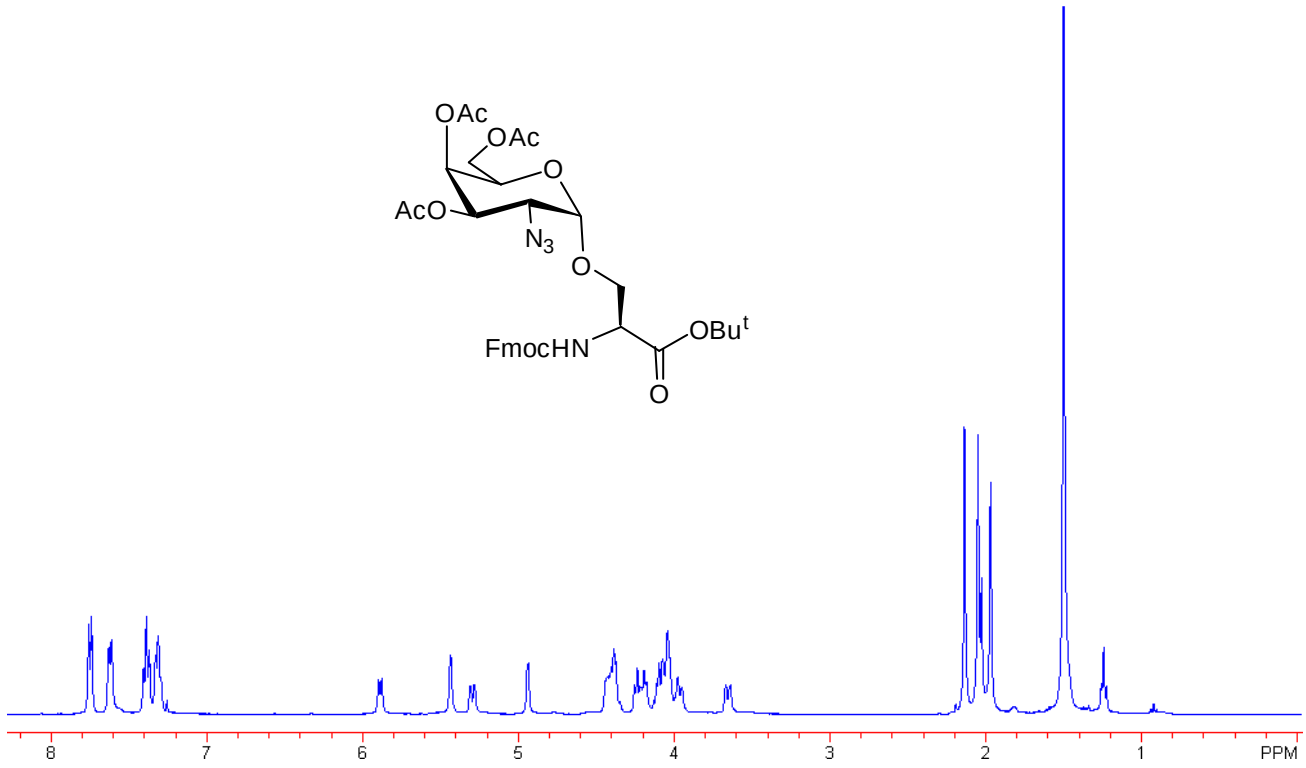
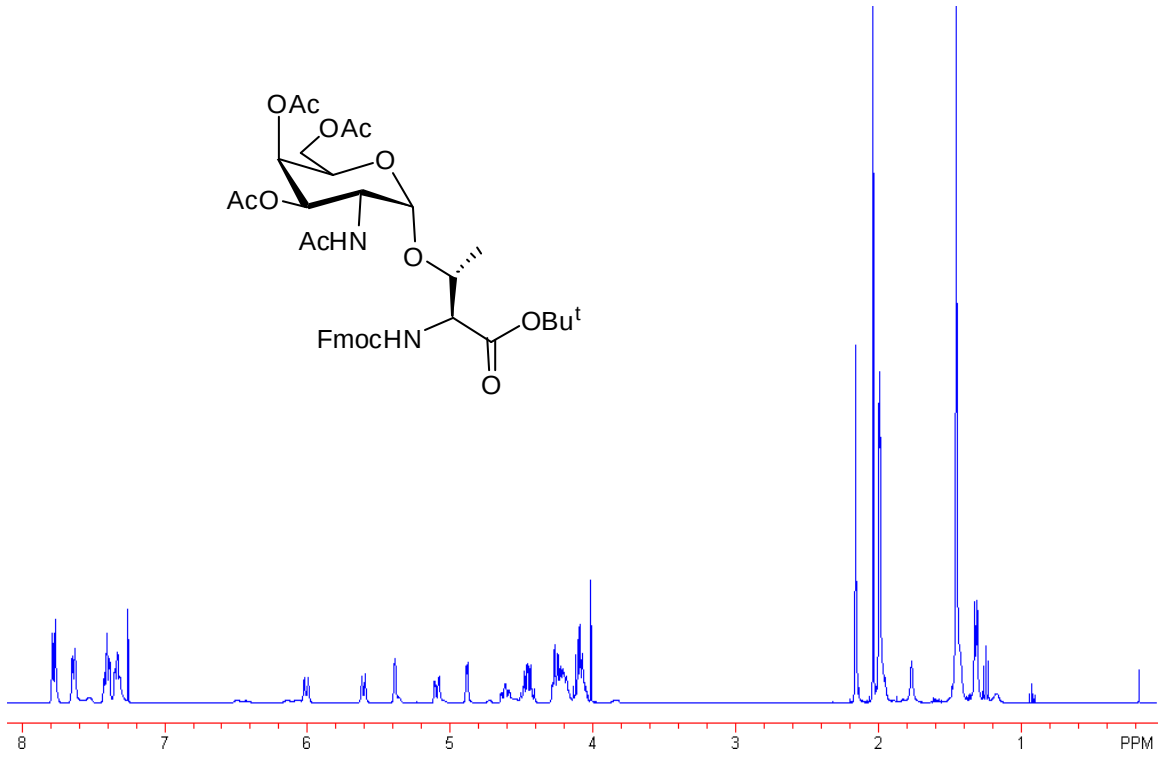


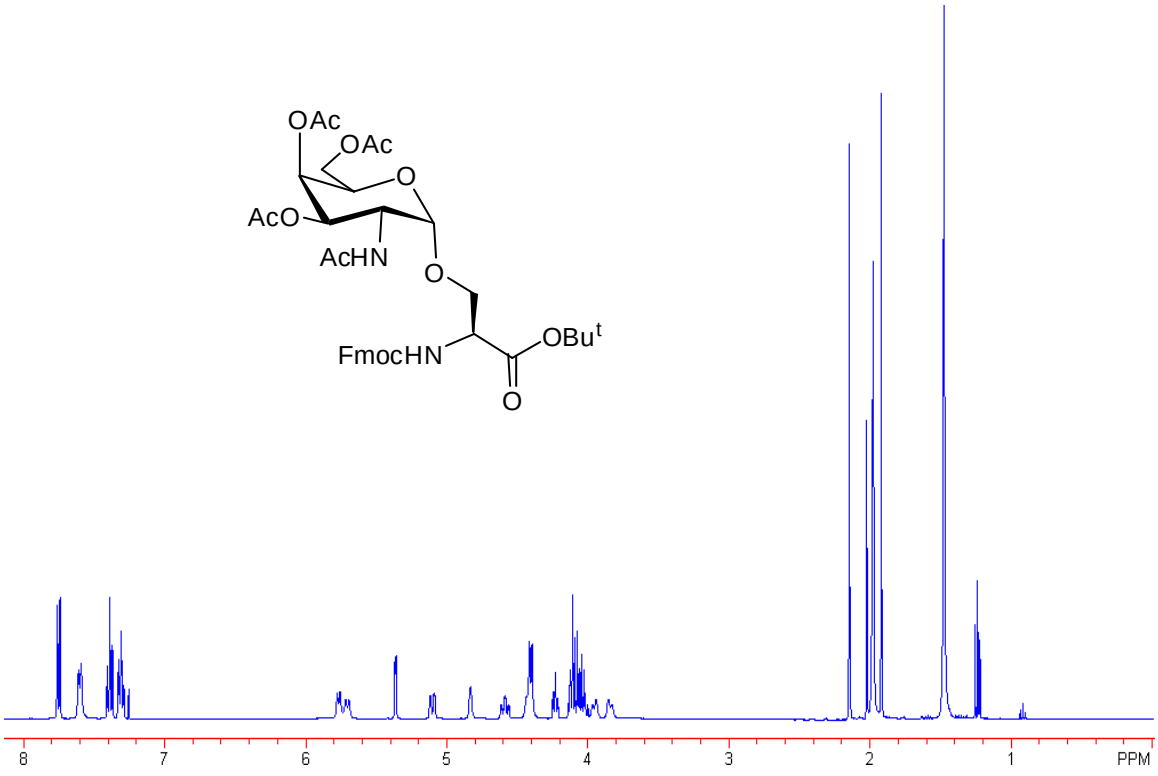
Figure 10. MALDI-TOF MS spectrum HPLC purified **13**: $C_{165}H_{277}N_{36}O_{78}P$ (m/z) 4044.1, observed 4045.3 $(M + H)^+$

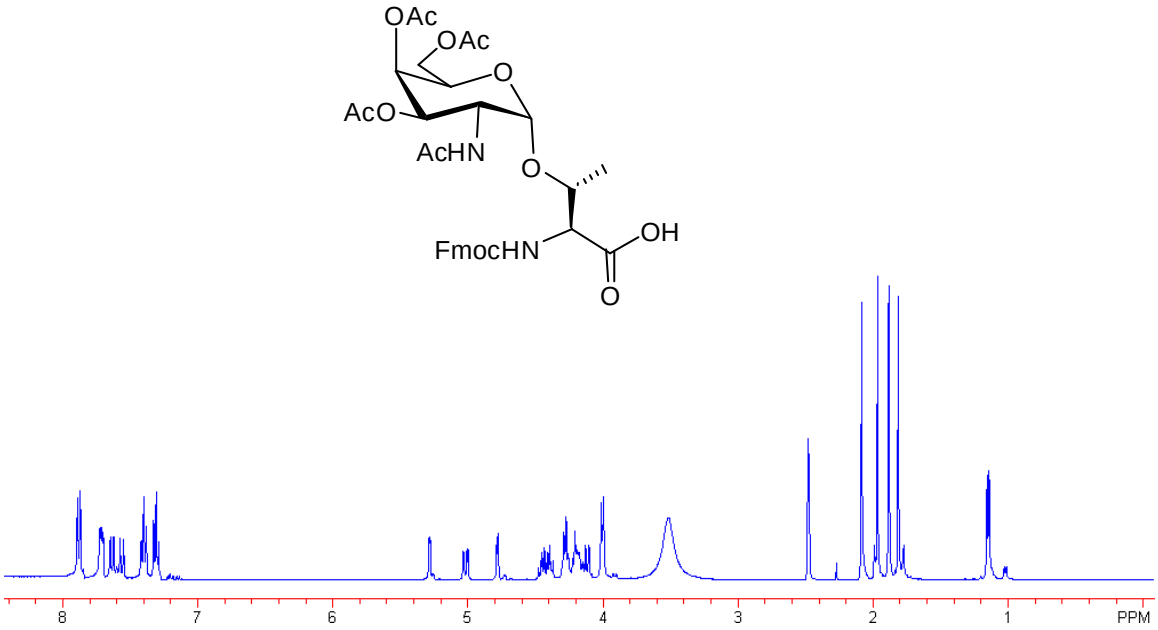
5. ^{1}H NMR Spectra of Synthetic Intermediates S1-S7, 3 and 4 (all are known compounds[1-4])

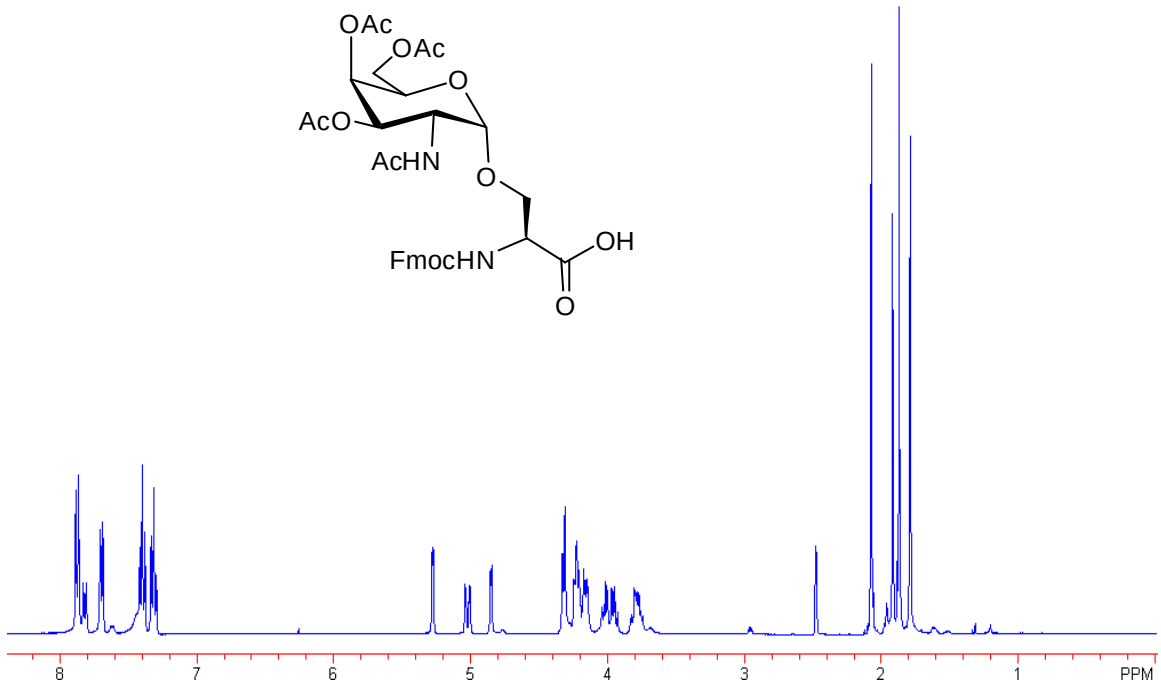



OAc
OAc
O
AcO
N3
O
FmocHN
OBut
O
8
7
6
5
4
3
2
1
PPM


OAc
OAc
O
AcO
AcHN
O
FmocHN
OBut
O
8
7
6
5
4
3
2
1
PPM


OAc
OAc
O
AcO
AcHN
O
FmocHN
OBu^t
O
8
7
6
5
4
3
2
1
PPM


OAc
OAc
O
AcO
AcHN
O
FmocHN
OH
O
8
7
6
5
4
3
2
1
PPM



6. Reference:

1. Broddefalk, J.; Nilsson, U.; Kihlberg, J., An Improved Synthesis of 3,4,6-Tri-O-Acetyl-2-Azido-2-Deoxy-Alpha-D-Galactopyranosyl Bromide - a Key Component for Synthesis of Glycopeptides and Glycolipids. *J. Carbohydr. Chem.* **1994,** *13*, 129-132.
2. Liu, M.; Young, V. G.; Lohani, S.; Live, D.; Barany, G., Syntheses of T-N building blocks N-alpha-(9-fluorenylmethoxycarbonyl)-O-(3,4,6-tri-O-acetyl-2-azido-2-deoxy-alpha-D-galactopyranosyl)-L-serine/-L-threonine pentafluorophenyl esters: comparison of protocols and elucidation of side reactions. *Carbohydr. Res.* **2005,** *340*, 1273-1285.
3. Guazzelli, L.; Catelani, G.; D'Andrea, F.; Giannarelli, A., Stereoselective entry into the D-GalNAc series starting from the D-Gal one: a new access to N-acetyl-D-galactosamine and derivatives thereof. *Carbohydr. Res.* **2009,** *344*, 298-303.
4. Schultz, M.; Kunz, H., Synthetic O-Glycopeptides as Model Substrates for Glycosyltransferases. *Tetrahedron Asym.* **1993,** 4, (6), 1205-1220.
5. Paulsen, H.; Adermann, K., Synthesis of O-Glycopeptides of the N-Terminus of Interleukin-2. *Lieb. Ann. Chem.* **1989**, 751-769.