Supporting information

Synthesis and Immunological Evaluation of Disaccharide Bearing MUC-1 Glycopeptide Conjugates with Virus-Like Particles

Xuanjun Wu,[†,‡] Craig McKay,[⊥] Christian Pett,[〃,#] Jin Yu,[〃] Manuel Schorlemer,[〃,#] Sherif Ramadan,[‡,¶] Shuyao Lang,[‡,§] Sandra Behren,[〃,#] Ulrika Westerlind,[〃,#] M. G. Finn,[⊥] and Xuefei Huang*[,‡,§,▽]

[†]National Glycoengineering Research Center, Shandong University, Qingdao, Shandong 266237, China

[‡]Department of Chemistry, [§]Institute for Quantitative Health Science and Engineering, and [▽]Department of Biomedical Engineering, Michigan State University, East Lansing, Michigan 48824, USA

[⊥]School of Chemistry & Biochemistry and School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

[〃]Leibniz-Institut für Analytische Wissenschaften–ISAS–e.V., 44227, Dortmund, Germany

[#]Department of Chemistry, Umeå University, 901 87 Umeå, Sweden

[¶]Chemistry Department, Faculty of Science, Benha University, Benha, Qaliobiya 13518, Egypt

Email: huangxu2@msu.edu

Table of Contents

Supporting Figures and Schemes

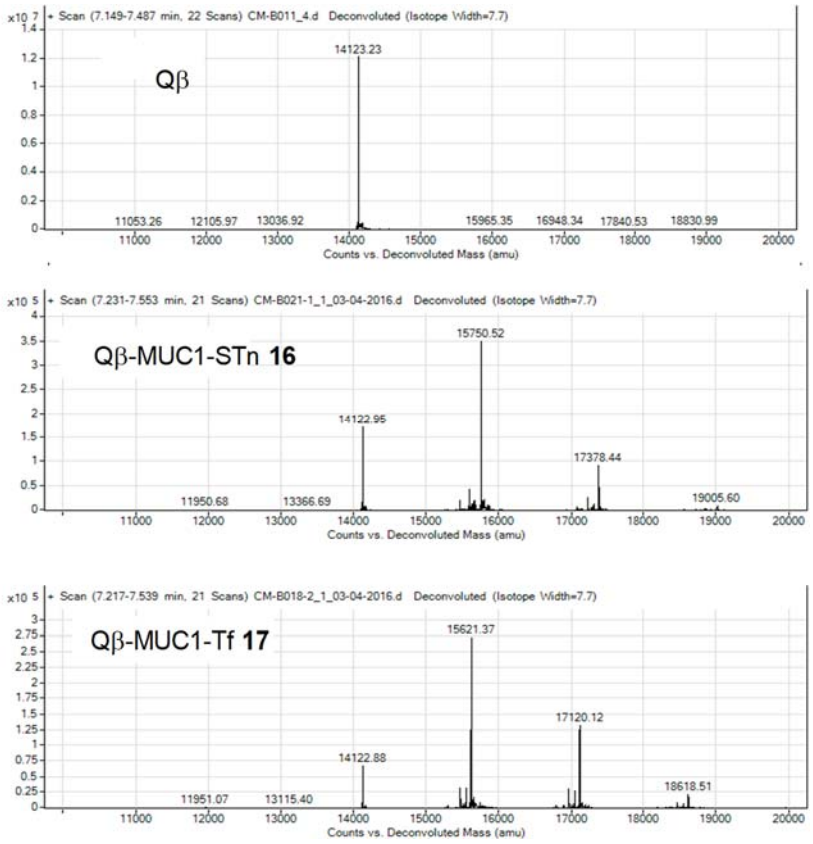


Figure S1. ESI-TOF HRMS spectra obtained for Qβ, Qβ-MUC1 conjugates **16** and **17**. Based on the intensities of the MS peaks corresponding to M^+ (M is the monomer of Qβ coat protein, MW 14,123), $(M+MUC1)^+$, $(M+2MUC1)^+$, and $(M+3MUC1)^+$, the average loading per particle was determined to be around 270 per particle.

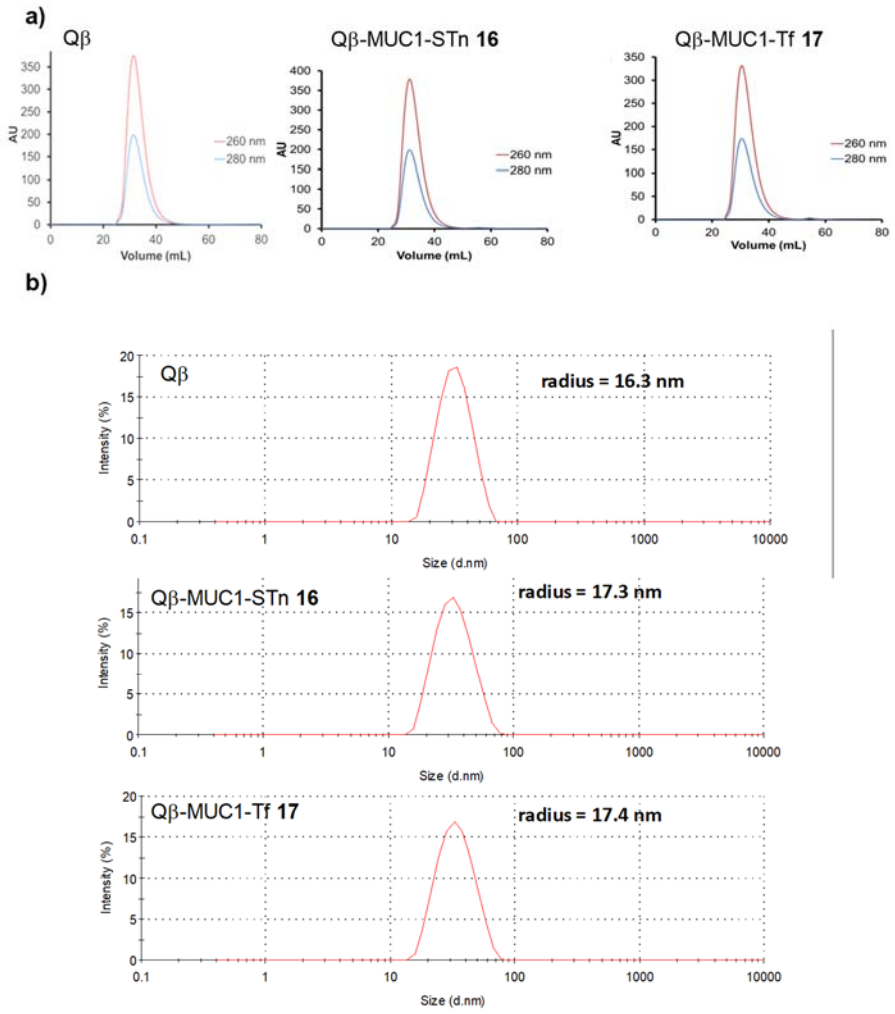
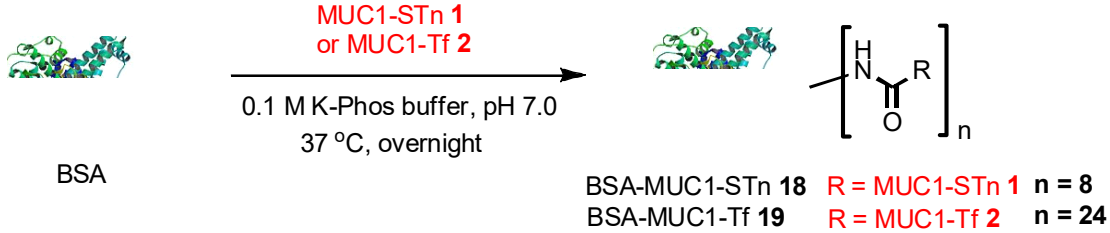


Figure S2. a) FPLC chromatograms and b) DLS analysis of Qβ, Qβ-MUC1 conjugates **16** and **17**.

Scheme S1. Synthesis of BSA-MUC1 conjugates **18** and **19** for ELISA.



a) BSA-MUC1-STn **18** **b)** BSA-MUC1-Tf **19**

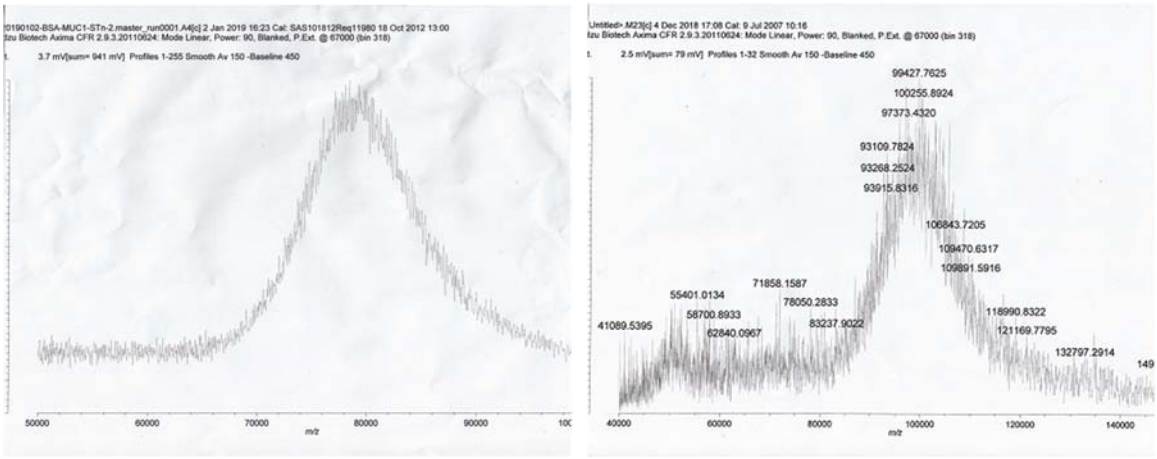


Figure S3. MALDI-TOF MS spectra for a) BSA-MUC1-STn **18** and b) BSA-MUC1-Tf **19**. Mass spectrometry analysis of the BSA-MUC1 conjugates **18** and **19** showed that the numbers of glycopeptides per BSA were 8 and 24 on average respectively based on the increase of molecular weights following glycopeptide conjugation.

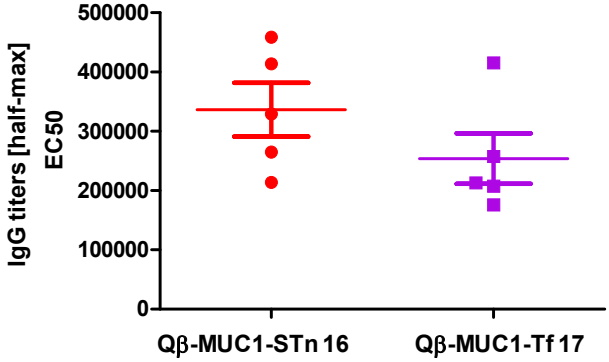


Figure S4. IgG (half max) titers (EC50) of anti-MUC1 antibodies from MUC1.Tg mice immunized with Qβ-MUC1 conjugates **16** and **17**. For determination of anti-MUC1 IgG titers, the ELISA measurements were performed against the corresponding BSA-MUC1 conjugates **18** and **19**. The EC50 antibody titers were calculated as half-max by a nonlinear regression, followed by a four-parameter logistic curve analysis. Each symbol represents one mouse (n = 5 mice for each group).

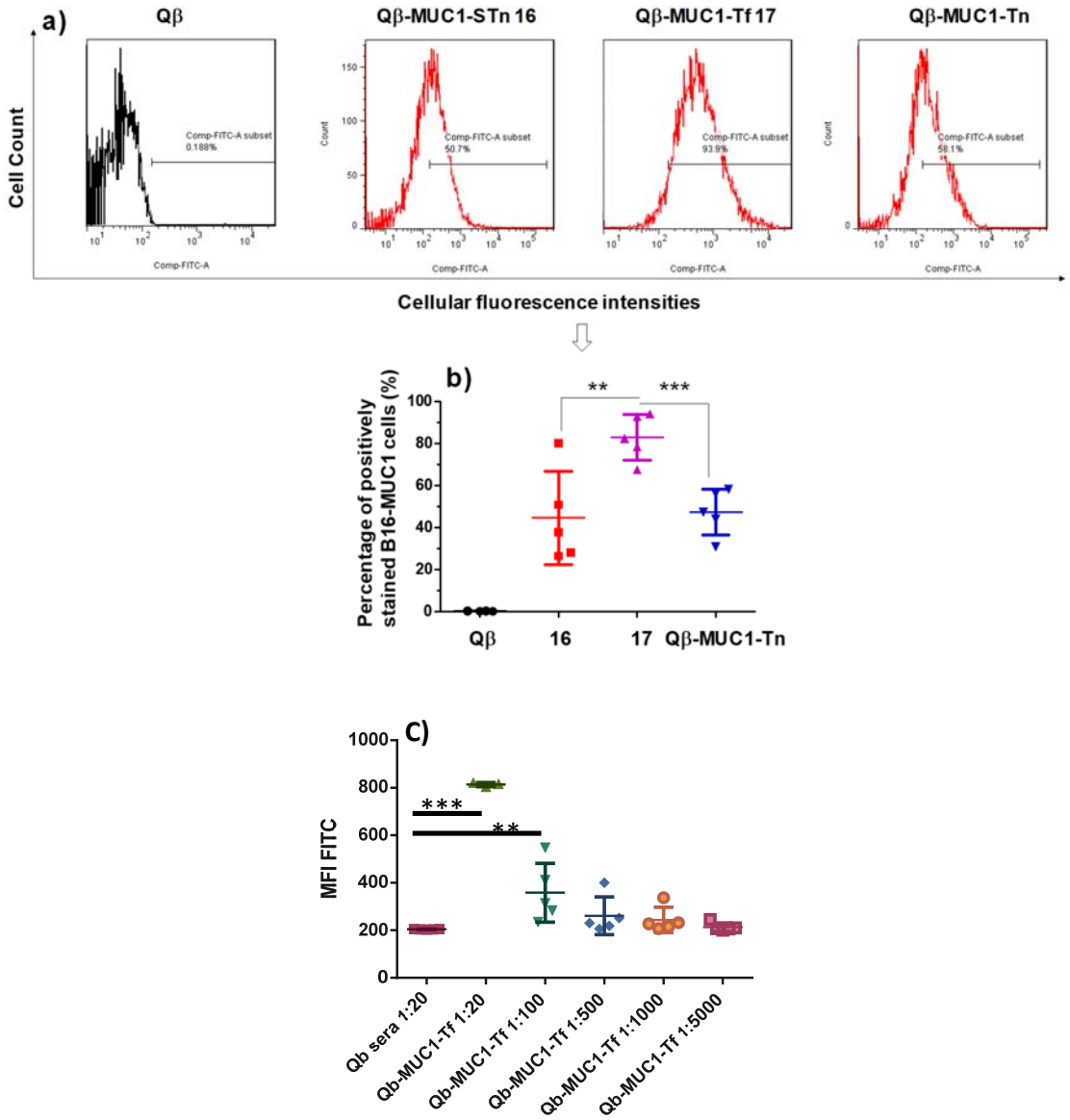


Figure S5. B16-MUC1 melanoma cells staining by IgG antibodies in postimmune sera elicited by various conjugates. a) For each conjugate, a representative FACS curve was shown. b) Analysis of the percentages of positively stained B16-MUC1 cells by postimmune sera elicited by various conjugates. The binding was tested with 1:20 dilution of the sera. Each symbol represents one mouse (n = 4–5 mice for each group). c) Dose dependent responses of postimmune sera from Qβ-MUC1-Tf immunized mice in binding with B16-MUC1 cells. Statistically significant differences were observed between the control sera at 1:20 dilution and Qβ-MUC-Tf sera up to 1:100 dilution. ** $p < 0.01$, *** $p < 0.001$. The p values were determined through a two tailed t-test using GraphPad Prism.

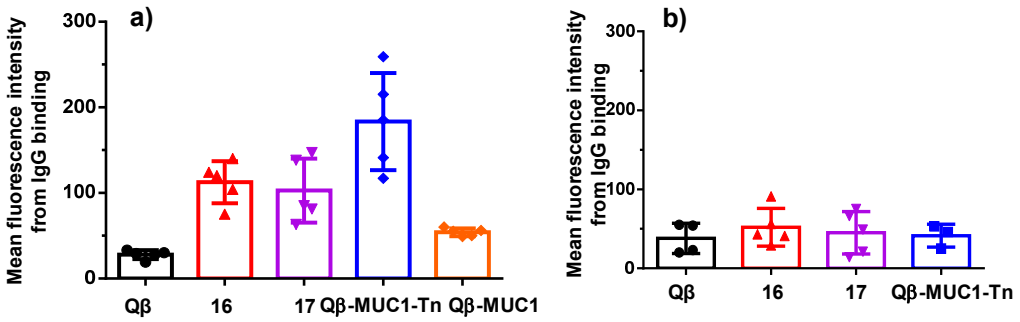


Figure S6. Flow cytometry analysis of the binding of anti-MUC1 IgG antibodies by synthetic conjugates to a) MCF-7 breast tumor cells and b) MCF-10A breast normal cells. Significant enhancements in cellular binding to MCF-7 cells were observed with both sera compared to those from control mice immunized with Qβ. Immunolabelling with monoclonal antibodies revealed surface expression of the MUC1-Tn epitope in MCF-7 cells,[1] which may have contributed to higher binding of MCF-7 by the Qβ-MUC1-Tn induced sera. All post-immune sera had low bindings with MCF-10A similar to those from the control mice. Binding to MCF-7 cells and MCF-10A cells was tested with 1:20 dilution of the corresponding serum. Each symbol represents one mouse (n = 3–5 mice for each group).

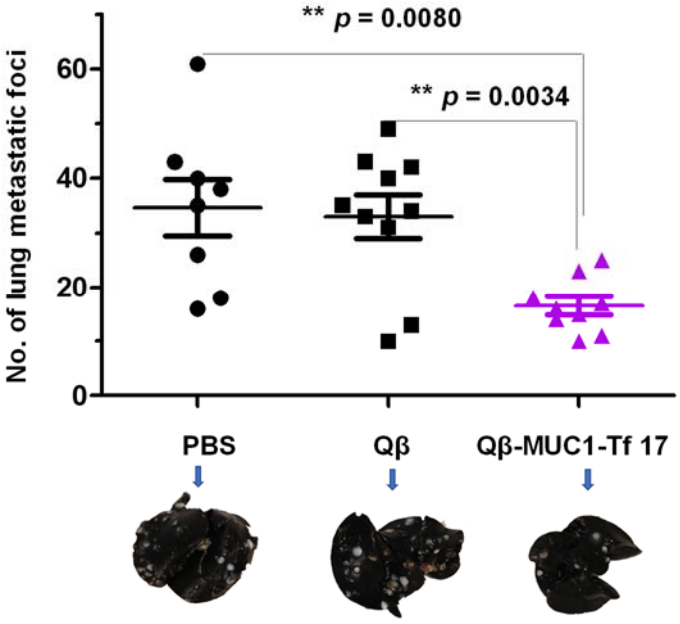
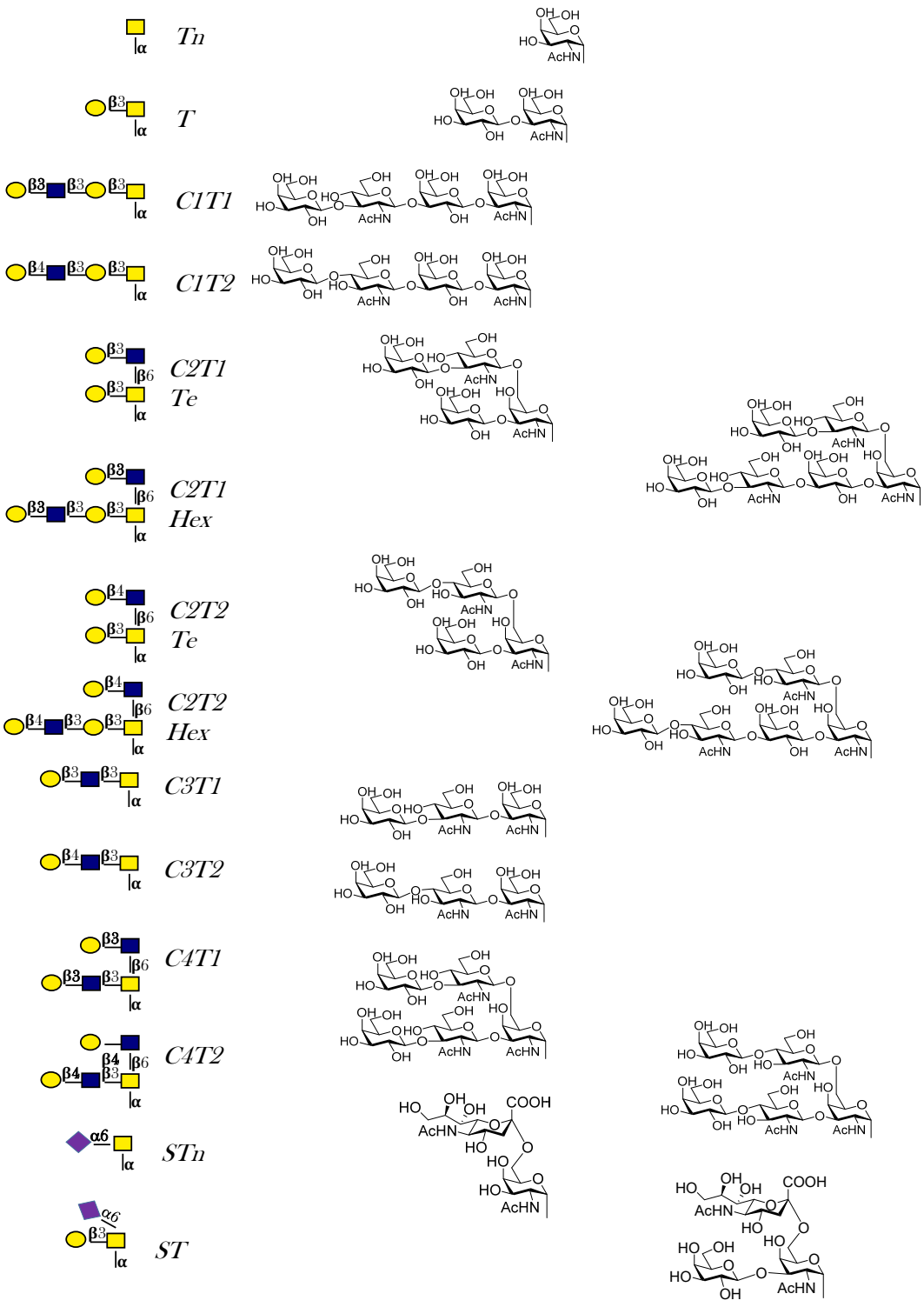


Figure S7. Representative pictures of the lung from mice receiving PBS, Qβ, and Qβ-MUC1-Tf **17** respectively. For clarity, **figure 3** (# of lung metastatic foci) is included in this figure.



Scheme S2. The abbreviations of glycan structures for glycopeptide microarray screening (see **Figures S7 and S8**).

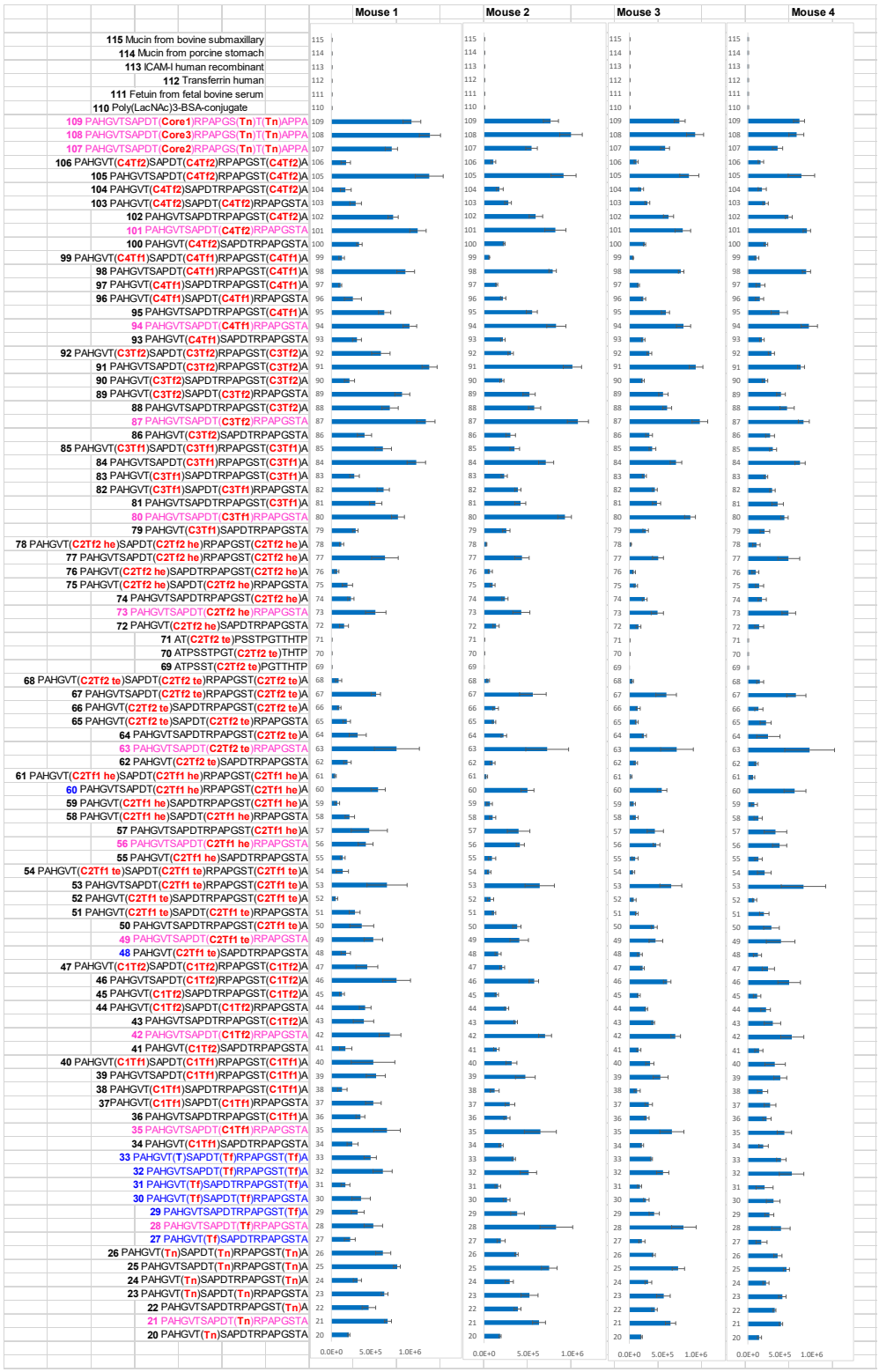


Figure S8. Glycopeptide microarray screening results of antisera induced by Qβ-MUC1-Tf **17**. The results from four anti-sera (1/25 dilution for mouse 1–3 and 1/100 for mouse 4) were shown.

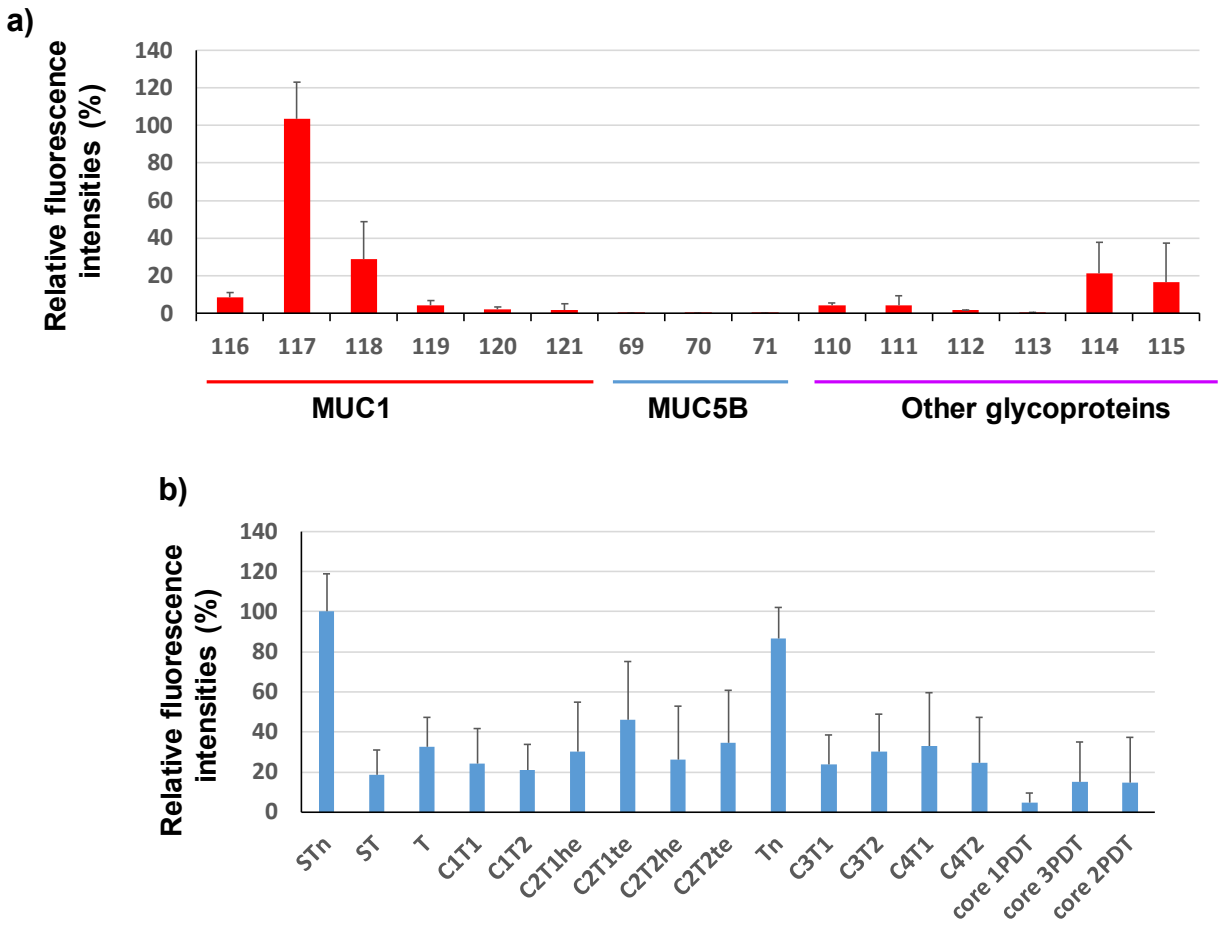


Figure S9. Representative results of MUC1 glycopeptide microarray screening of anti-sera from Qβ-MUC1-STn **16** immunized mice. a) Comparison of fluorescence intensities of microarray components containing MUC1 glycopeptides bearing STn antigen at various locations showed that glycosylation at PDT*R region led to strongest recognition by post-immune sera. T* denotes STn bearing threonine; Glycopeptide **116**: PAHGVT*SAPDTRPAPGSTAP; **117**: PAHGVTSAPDT*RPAPGSTAP; **118**: PAHGVTSAPDTRPAPGST*AP; **119**: PAHGVT*SAPDTRPAPGST*AP; **120**: PAHGVT*SAPDTRPAP; **121**: GSTAPPAHGVT*SAP. Glycopeptides **69–71** are various MUC5B glycopeptides. **110-115** are poly(LacNAc)-BSA, fetuin, transferrin, ICAM-1, porcine stomach mucin, and bovine submaxillary mucin respectively. b) Comparison of fluorescence intensities of microarray components containing MUC1 glycopeptides bearing various glycans at PAHGVTSAPDT*RPAPGSTA showed that while Tf gave the strongest recognition, other glycans can be recognized as well. Glycan structures: glycopeptide **117**: STn; **28**: Tf; **35**: C1Tf1; **42**: C1Tf2; **56**: C2Tf1he; **49**: C2Tf1te; **73**: C2Tf2he; **63**: C2Tf2te; **21**: Tn; **80**: C3Tf1; **87**: C3Tf2; **94**: C4Tf1; **101**: C4Tf2; **109**: core 1 PDT; **108**: core 3 PDT; **107**: core 2 PDT. The error bars represent standard deviation (SD) of eight replicates.

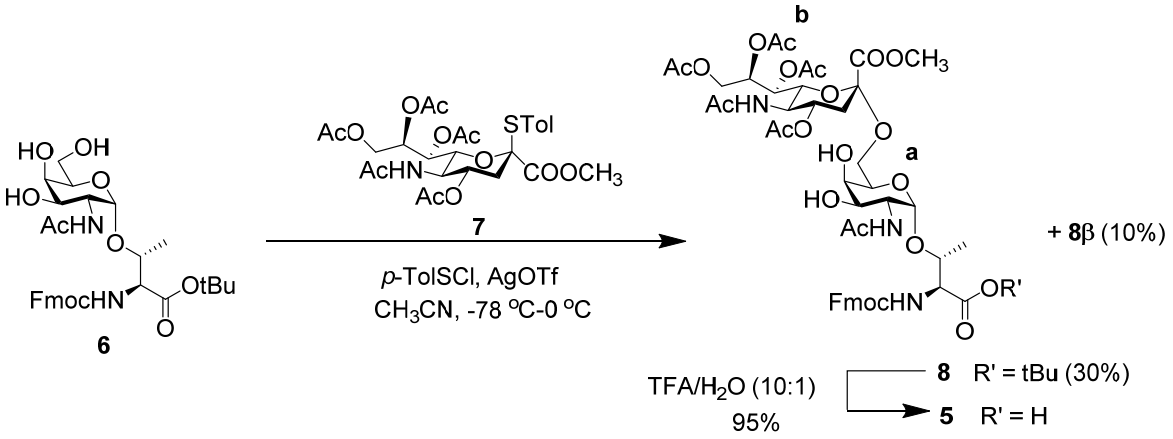
General experimental procedures and methods for synthesis

All chemicals were reagent grade and were used as received from the manufacturer, unless otherwise noted. Solvents were dried using a solvent purification system. Glycosylation reactions were performed with 4Å molecular sieves that were flamed dried under high vacuum. Reactions were visualized by UV light (254 nm) and by staining with either $Ce(NH_4)_2(NO_3)_6$ (0.5 g) and $(NH_4)_6Mo_7O_{24}\cdot 4H_2O$ (24.0 g) in 6% H_2SO_4 (500 mL), 5% H_2SO_4 in EtOH. Flash chromatography was performed on silica gel 601 (230–400 Mesh).

Centrifugal filter units of 10,000 and 100,000 molecular weight cut-off (MWCO) were purchased from EMD Millipore. Fast protein liquid chromatography (FPLC) was performed on a GE ÄKTA Explorer (Amersham Pharmacia) instrument equipped with a Superose-6 column. Microfluidic capillary gel electrophoresis was performed on a Bioanalyzer 2100 Protein 80 microfluidics chip (Agilent Technologies). For characterization of Qβ-MUC1 conjugates, liquid chromatography-mass spectrometry (LCMS) analysis was performed. The samples for LCMS were prepared as follows: 1:1 v/v of 40 μg mL^{-1} of Qβ-MUC1 stock solution and 100 mM DTT was mixed and incubated in a water bath at 37 °C for 30 min. One drop of 50% formic acid was added into the mixture. LCMS was performed on Waters Xevo G2-XS quadrupole/time-of-flight UPLC/MS/MS. The liquid chromatography was done on ACQUITY UPLC® Peptide BEH C18 column, 130Å, 1.7 μm, 2.1 mm × 150 mm, using gradient eluent from 95% 0.1% formic acid in water to 95% 0.1% formic acid in CH_3CN (0.3 mL min^{-1} flowrate) at column temperature 40 °C. The multiple charge mass spectra were transformed to single charge by using algorithm MaxEnd148a. The average numbers of MUC1/subunit were analyzed by signal intensity of mass spectrum. For characterization of BSA-MUC1 conjugates, MALDI-TOF MS analysis was performed. The samples for MALDI-TOF were prepared as follows: 1:1 v/v of 2 mg mL^{-1} of BSA-MUC1 conjugates and 100 mM DTT was mixed and incubated in a water bath at 37 °C for 30 min. After desalting using Cleanup C18 Pipette Tips (Agilent Technologies), the sample (2 μL) and matrix solution (2 μL, 10 mg mL^{-1} sinapic acid in 50/50/0.1 CH_3CN/H_2O/TFA) was mixed and spotted on a MALDI plate, air-dried (3 rounds) and then analyzed by MALDI-TOF mass spectrometry (Applied Biosystems Voyager DE STR). Protein concentration was measured using the Coomassie Plus Protein Reagent (Bradford Assay, Pierce) with bovine serum albumin (BSA) as the standard.

Mouse melanoma B16 line expressing human MUC1 (B16-MUC1) was kindly provided by Prof. Sandra J. Gendler (Mayo Clinic). B16-MUC1 cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS, 100 U mL^{-1} penicillin and 100 µg mL^{-1} streptomycin, 2 mM L-glutamine, 1 mM sodium pyruvate and 0.3 mg mL^{-1} G418 disulfate salt. MCF-7 cells were cultured in Eagle's Minimum Essential Medium (EMEM, ATCC® 30-2003™) with 10 % FBS, 100 U mL^{-1} penicillin and 100 µg mL^{-1} streptomycin, 2 mM L-glutamine, non-essential amino acids, sodium pyruvate and 10 µg mL^{-1} bovine insulin. MCF-10A cells were cultured in DMEM / Nutrient F12 (with 15 mM HEPES, L-glutamine and sodium bicarbonate) with 5% horse serum, 100 U mL^{-1} penicillin and 100 µg mL^{-1} streptomycin, 0.5 µg mL^{-1} hydrocortisone, 100 ng mL^{-1} cholera toxin, 10 µg mL^{-1} insulin and 20 ng mL^{-1} human recombinant epidermal growth factor (EGF).

Synthesis of Fmoc protected STn antigen building block 5



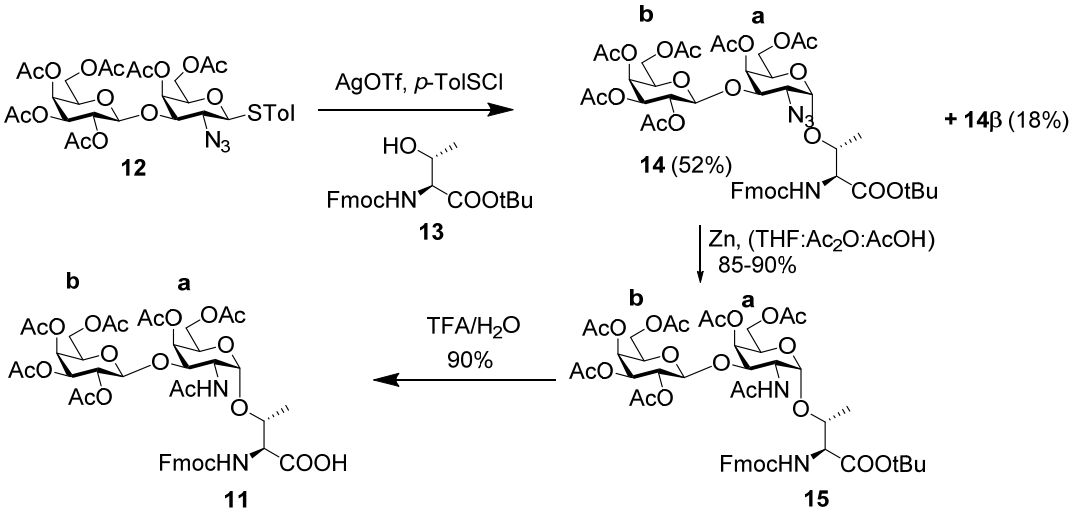
***O*-{2-Acetamido-2-deoxy-6-*O*-[methyl(5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosyl)onate]-α-D-galactopyranosyl}-*N*-[(9*H*-fluoren-9-ylmethoxy)carbonyl]-L-threonine *tert*-butyl ester (8):** Galactosamine threonine ester **6**[2-3] and sialyl donor **7**[4] were obtained according to literature. Sialyl donor **7** (1.2 g, 2.53 mmol) was dissolved in acetonitrile (25 mL) along with **6** (1.2 g, 2.00 mmol) and 4 Å molecular sieves (2.5 g). The reaction mixture was stirred at rt for 30 minutes then cooled to -78 °C. Silver triflate (1.44 g, 5.60 mmol) dissolved acetonitrile (5 mL) was added and the reaction was stirred for 10 min at -78 °C. After that, promoter, *p*-toluenesulfenyl chloride (*p*-TolSCl, 338 µL, 2.34 mmol) was added

directly into the solution. The reaction was allowed to slowly warm to 0 °C over 3 h. When the reaction completed as judged by TLC analysis, the mixture was diluted with dichloromethane (DCM) and filtered through Celite. The mixture was washed with sat. $NaHCO_3$ and concentrated. The crude product was purified by flash silica gel column (EtOAc/methanol 20:1) and then α/β anomers were separated by reverse phase HPLC (with gradient solvent CH_3CN and H_2O gradient 5–5% in 0–2 min, 5–50% in 2–5 min, 50–100% in 5–40 min, 100–100% in 40–45 min and 100–5% in 45–50 min; total flow rate: 5 mL min^{-1}) to obtain 30% of Fmoc-Neu5Ac-α-2,6-GalNAc-α-Thr **8** and 10% of its β anomer (Fmoc-Neu5Ac-β-2,6-GalNAc-α-Thr **8β**). For characterization of **8**, HRMS: $C_{51}H_{67}N_3O_{22}$ $[M+H]^+$ calcd: 1074.4294, obsd: 1074.4351. ^1H NMR (500 MHz, CD_3OD) δ 7.81 (d, *J* = 7.3 Hz, 2 H), 7.71 – 7.67 (m, 2 H), 7.43– 7.38 (m, 2 H), 7.35 – 7.30 (m, 2 H), 5.43 – 5.39 (m, 1 H, H8b), 5.34 (dd, *J* = 8.3, 2.0 Hz, 1 H, H7b), 4.86 – 4.80 (m, 1 H), 4.77 (d, *J* = 4.4 Hz, 1 H, H1a (*α*-Linkage)), 4.59 (dd, *J* = 10.8, 6.4 Hz, 1 H), 4.49 (dd, *J* = 11.0, 6.1 Hz, 1 H), 4.34 – 4.22 (m, 4 H), 4.16 – 4.08 (m, 4 H), 3.98 (d, *J* = 10.8 Hz, 1 H), 3.94 – 3.88 (m, 2 H), 3.85 (d, *J* = 2.9 Hz, 1 H), 3.83 (s, 3 H, COOMe), 3.64 – 3.56 (m, 2 H), 2.65 (dd, *J* = 12.7, 4.4 Hz, 1 H, H3b-*eq.*), 2.14 (s, 3 H, OAc), 2.08 (s, 3 H, OAc), 2.00 (s, 3 H, OAc), 1.98 (s, 3 H, OAc), 1.98 (s, 3 H, NHAc), 1.86 (t, *J* = 12.8 Hz, 1 H, H3b-*ax.*), 1.83 (s, 3 H, NHAc), 1.44 (s, 9 H, O*t*Bu), 1.22 (d, *J* = 6.8 Hz, 3 H, CH_3 (Thr)); ^{13}C NMR (126 MHz, CD_3OD) δ 172.28, 172.11, 170.96, 170.42, 170.34, 170.05, 169.64, 167.97 (C1b (CO_2Me)), 157.77, 143.97, 143.70, 141.30, 127.43, 127.40, 126.83, 126.78, 124.74, 124.63, 119.58, 119.52, 109.99, 99.38 (C1a, $^1J_{C1,H1}$= 173.0 Hz (*α*-Linkage)), 98.60, , 82.02, 75.29, 71.93, 69.75, 69.31, 68.56, 68.49, 67.31, 66.14, 63.50, 62.00, 59.43, 51.95, 49.44, 48.69, 37.35, 26.96, 21.83, 21.27, 19.81, 19.47, 19.29, 19.28, 18.43. For ^1H-^{13}C HMBC, $^3J_{C1,H3ax}$=6.4 Hz and $^3J_{H7,H8}$ = 8.3 Hz, indicate α-sialyl linkage formed.

***O*-{2-Acetamido-2-deoxy-6-*O*-[methyl(5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-non-2-ulopyranosyl)onate]-α-D-galactopyranosyl}-*N*-[(9*H*-fluoren-9-ylmethoxy)carbonyl]-L-threonine (5):** A solution of **8** (300 mg) in trifluoroacetic acid/H_2O (20 mL/2 mL) was stirred for 2 h at rt. Then the solvents were removed *in vacuo*, and the residue was co-evaporated with toluene (5 × 20 mL). The crude product was purified by flash chromatography on silica gel to give **5** as a colorless, amorphous solid in 90% yield. MS: $C_{47}H_{59}N_3O_{22}$ $[M-H]^-$ calcd: 1016.35, obsd: 1016.36. ^1H NMR (500 MHz, CD_3OD) δ 7.81 (d, *J* = 7.8 Hz, 2 H), 7.72 – 7.67 (m, 2 H), 7.43 – 7.38 (m, 2 H), 7.35 – 7.30 (m, 2 H), 5.44 – 5.39 (m, 1 H, H8b), 5.34 (d, *J* = 8.5

Hz, 1 H, H7b), 4.85 – 4.81 (m, 2 H, H1a, H4b), 4.58 (dd, *J* = 10.8, 6.4 Hz, 1 H), 4.47 (dd, *J* = 10.8, 6.4 Hz, 1 H), 4.37 – 4.23 (m, 4 H, CH (Thr), H9b), 4.22 – 4.07 (m, 4 H, H2a, H6b, H9b), 4.00 – 3.94 (m, 1 H, H5b), 3.93 – 3.87 (m, 2 H, H4a, H6a), 3.85 (d, *J* = 2.4 Hz, 1 H, H6a), 3.82 (s, 3 H, COOMe), 3.64 (dd, *J* = 10.8, 2.9 Hz, 1 H, H3a), 3.61 – 3.56 (m, 1 H, H5a), 2.65 (dd, *J* = 12.7, 4.4 Hz, 1 H, H3b-*eq.*), 2.13 (s, 3 H, OAc), 2.07 (s, 3 H, OAc), 2.00 (s, 3 H, OAc), 1.98 (s, 3 H, OAc), 1.98 (s, 3 H, NHAc), 1.86 (t, *J* = 12.8 Hz, 1 H, H3-*ax.*), 1.83 (s, 3 H, NHAc), 1.23 (d, *J* = 6.4 Hz, 3 H, CH_3 (Thr)); ^{13}C NMR (126 MHz, CD_3OD) δ 172.50, 172.11, 172.06, 171.00, 170.42, 170.36, 170.07, 167.99 (C1b (COOMe)), 157.76, 144.00, 143.69, 141.28, 127.43, 127.38, 126.82, 126.76, 124.75, 124.66, 119.56, 119.52, 99.58 (C1a, $^1J_{C1,H1}$= 173.0 Hz (*α*-Linkage)), 98.60, 75.89, 71.89, 69.63, 69.31, 68.59, 68.43, 68.36, 67.27, 66.20, 63.52, 61.96, 58.55, 51.94, 49.70, 48.69, 37.35, 21.63, 21.26, 19.80, 19.45, 19.28, 19.25, 18.01. For ^{1}H-^{13}C HMBC, $^3J_{C1,H3ax}$=6.4 Hz and $^3J_{H7,H8}$ = 8.5 Hz, indicate α-sialyl linkage formed.

Synthesis of Fmoc protected Tf antigen building block 11



***N*-(9*H*-Fluoren-9-yl)-methoxycarbonyl-*O*-(2-azido-4,6-di-*O*-acetyl-2-deoxy-3-*O*-[2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl]-α-D-galactopyranosyl)-L-threonine *tert*-butyl ester (14):** Disaccharide donor **12** was prepared according to a literature procedure.[5] A mixture of disaccharide donor **12** (0.18 g, 0.25 mmol) and threonine acceptor **13** (0.099 g, 0.25 mmol) was dissolved and co-evaporated with anhydrous toluene (2× 20 mL) and dried in high vacuum. Then the mixture was dissolved in anhydrous DCM (15 mL), molecular sieves (1 g, 4 Å) and 2,4,6-tri-tert-butylpyrimidine (TTBP, 0.062 g, 0.25 mmol) were added and the suspension was stirred for 1 h at rt under nitrogen. The mixture was cooled to -78 ^{O}C and a solution of silver triflate (0.192 mg, 0.75 mmol) in anhydrous Et_2O/DCM (4 mL/0.5 mL) was added. After 5 min, orange *p*-TolSCl (0.036 mL, 0.25 mmol) was directly added via a microsyringe to the reaction mixture to not be frozen on flask wall. The yellow color disappeared in a few seconds indicating consumption of the *p*-TolSCl promoter. After 5 min, TLC indicated the complete activation of donor. Then a solution of the acceptor in anhydrous DCM was added to the reaction via a syringe. The reaction was allowed to stir for 1 h at -78 ^{o}C then the temperature was raised to -10 ^{o}C with stirring continued for another 1 h. Finally, the reaction was quenched by DIPEA (3–4 drops), diluted with DCM and filtered over celite. The DCM solution was washed with water, dried over Na_2SO_4, concentrated and purified by silica gel flash chromatography (Hex/EtOAc, 3:1) to afford the products **14** in 52% yield and **14β** in 18% yield. For characterization of **14**, MS: $C_{47}H_{58}N_4O_{20}$ $[M+NH_4]^+$ calcd: 1016.40, obsd: 1016.35. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 – 7.74 (m, 2H), 7.67 – 7.60 (m, 2H),

7.45 – 7.38 (m, 2H), 7.35 – 7.29 (m, 1.2 Hz, 2H), 5.66 (d, *J* = 9.5 Hz, 1H, NH-Fmoc), 5.50 (d, *J* = 3.3 Hz, 1H, H3a), 5.38 (dd, *J* = 3.4, 1.1 Hz, 1H, H4b), 5.21 (dd, *J* = 10.5, 7.8 Hz, 1H, H2b), 5.06 (d, *J* = 3.9 Hz, 1H, H1a (*α*-Linkage)), 5.02 (dd, *J* = 10.5, 3.4 Hz, 1H, H3b), 4.74 (d, *J* = 7.8 Hz, 1H, H1b (*β*-Linkage)), 4.49 (dd, *J* = 10.3, 7.1 Hz, 1H), 4.44 – 3.39 (m, 1H, CH (Thr)), 4.35 – 4.25 (m, 3H), 4.21 – 4.05 (m, 6H, H5a, 2H6b), 4.00 – 3.91 (m, 2H, H4a, H5b), 3.61 (dd, *J* = 10.7, 3.8 Hz, 1H, H2a), 2.16 (s, 3H, OAc), 2.13 (s, 3H, OAc), 2.08 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.05 (s, 3H, OAc), 1.99 (s, 3H, OAc), 1.51 (s, 9H, O-*t*Bu), 1.35 (d, *J* = 6.5 Hz, 3H, CH_3(Thr)).; ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.44, 170.39, 170.24, 170.09, 169.60, 169.52, 169.14, 156.79, 143.86, 143.71, 141.28, 127.77, 127.09, 127.04, 125.19, 125.15, 120.04, 120.01, 101.58 (C1b), 99.34 (C1a), 82.93, 76.38, 74.66, 70.81, 70.79, 69.39, 68.73, 67.89, 67.35, 66.72, 62.92, 60.98, 60.40, 59.57, 59.20, 47.14, 29.71, 28.00, 20.75, 20.73, 20.71, 20.67, 20.58, 18.98, 14.21.

***N*-(9*H*-Fluoren-9-yl)-methoxycarbonyl-*O*-(2-acetamido-4,6-di-*O*-acetyl-2-deoxy-3-*O*-[2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl]-α-D-galactopyranosyl)-L-threonine *tert*-butyl ester (15):** A solution of **14** was dissolved in a mixture of THF:Ac_2O:AcOH (3:2:1). To that solution, Zn dust (10 equiv.) was added. The reaction mixture was stirred for 3 hours at rt. Then the Zn was filtered out and washed with MeOH. The solvents were removed in vacuo, and the residue was co-evaporated with toluene (5 × 20 mL). The crude product was purified by flash chromatography on silica gel (EtOAc/Hex, 2:1) to give **15** as a colorless, amorphous solid in 85-90%. HRMS: $C_{49}H_{62}N_2O_{21}$ $[M+H]^+$ calcd: 1015.3923, obsd: 1015.3970. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (d, *J* = 7.5 Hz, 2 H), 7.61 (d, *J* = 7.6 Hz, 2 H), 7.44 – 7.39 (m, 2 H), 7.37 – 7.29 (m, 2 H), 5.87 (d, *J* = 9.7 Hz, 1 H), 5.44 – 5.34 (m, 2 H), 5.13 – 5.05 (m, 1 H), 4.97 – 4.90 (m, 1 H), 4.84 (d, *J* = 4 Hz,1 H, H1a (*α*-Linkage)), 4.59 – 4.47 (m, 4 H, H1b (*β*-Linkage)), 4.25 (q, *J* = 10.5, 8.4 Hz, 2 H), 4.20 – 4.09 (m, 6 H), 3.99 (dd, *J* = 11.2, 7.3 Hz, 1 H), 3.87 (t, *J* = 6.8 Hz, 1 H), 3.79 (dd, *J* = 11.2, 3.0 Hz, 1 H), 2.16 (s, 3H, OAc), 2.13 (s, 3H, OAc), 2.07 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.04 (s, 3H, OAc), 2.02 (s, 3H, OAc), 1.97 (s, 3H, NHAc),, 1.46 (s, 9 H, O-*t*Bu), 1.29 (d, *J* = 6 Hz, 3 H, CH_3 (Thr)); ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.17, 170.46, 170.40, 170.19, 170.14, 169.84, 169.51, 156.37, 143.59, 141.34, 127.86, 127.10, 124.86, 120.11, 120.09, 100.91 (C1b), 100.07 (C1a), 83.29, 76.70, 73.34, 70.71, 70.65, 69.07, 68.61, 67.99, 66.98, 66.70, 63.01, 60.93, 60.40, 58.99, 48.36, 47.23, 28.09, 27.98, 23.42, 21.06, 20.75, 20.73, 20.71, 20.67, 20.60, 20.58, 18.53, 14.20, 14.13.

***N*-(9*H*-Fluoren-9-yl)-methoxycarbonyl-*O*-(2-acetamido-4,6-di-*O*-acetyl-2-deoxy-3-*O*-[2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl]-α-D-galactopyranosyl)-L-threonine (11):** A solution of **15** (300 mg) in trifluoroacetic acid/H_2O (20 mL/2 mL) was stirred for 3 h at rt. Then the solvents were removed *in vacuo*, and the residue was co-evaporated with toluene (5 × 20 mL). The crude product was purified by flash chromatography on silica gel (DCM:MeOH) to give **11** as a colorless, amorphous solid in 90% yield. HRMS: $C_{45}H_{54}N_2O_{21}$ $[M+H]^+$ calcd: 959.3297, obsd: 959.3304. ^{1}H NMR (500MHz, CD_3OD): δ 7.83 (d, *J* = 7.8 Hz, 2 H), 7.71 (d, *J* = 7.3 Hz, 2 H), 7.39 – 7.44 (m, 2 H), 7.32 – 7.36 (m, 2 H), 5.34 – 5.38 (m, 2 H, H3b), 4.96 – 4.99 (m, 2 H, H4a, H2b), 4.86 (d, *J* = 3.9 Hz, 1 H, H1a (*α*-Linkage)), 4.60 (dd, *J* = 10.8, 6.4 Hz, 2 H, H1b (*β*-Linkage), H5a), 4.52 – 4.56 (m, 1 H), 4.41 – 4.374.39 (m, 1 H, CH (Thr)), 4.28 – 4.36 (m, 2 H, H2a, H6a), 4.23 – 4.26 (m, 1 H), 4.08 – 4.22 (m, 5 H, H5b, H6b), 3.94 – 4.00 (m, 2 H, H4b), 3.87 (dd, *J* = 11.0, 3.2 Hz, 1 H, H3a), 2.13 (s, 3 H, OAc), 2.11 (s, 3 H, OAc), 2.05 (s, 3 H, OAc), 2.04 (s, 3 H, OAc), 2.01 (s, 3 H, OAc), 1.97 (s, 3 H, OAc), 1.93 (s, 3 H, NHAc), 1.24 (d, *J* = 6.8 Hz, 3 H, CH_3 (Thr)); ^{13}C NMR (126 MHz, CD_3OD) δ 171.85, 171.77, 170.90, 170.63, 170.62, 170.56, 170.05, 169.60, 157.70, 143.88, 143.76, 141.29, 141.27, 127.47, 127.46, 126.80, 124.73, 124.58, 119.65, 119.61, 101.07 (C1b, $^1J_{C1,H1}$= 162.0 Hz (*β*-Linkage)), 99.54 (C1a, $^1J_{C1,H1}$= 173.0 Hz (*α*-Linkage)), 75.94, 73.62, 70.75, 70.26, 69.76, 68.68, 67.53, 67.04, 66.23, 62.86, 60.81, 58.40, 48.52, 21.82, 19.43, 19.31, 19.27, 19.25, 19.06, 19.05, 17.77.

Synthesis of MUC1 glycopeptides 1 and 2

The MUC1 glycopeptides **1** and **2** were synthesized from *p*-nitrophenyl carbonate Wang resin pre-loaded with Fmoc-1,4-diaminobutane **9** using Fmoc chemistry. The *N*-terminal protecting group, Fmoc, was de-protected by 20% piperidine in DMF. The amino acid coupling was carried out with Fmoc amino acids (5 eq.) using (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU)/hydroxybenzotriazole (HOBt) (4.9 eq.) and DIPEA (10 eq). For coupling of the glycosyl amino acid **5** or **11**, Fmoc protected STn antigen building block **5** (2 eq.) or Fmoc protected Tf antigen block Fmoc-Tf-Thr **11** (2 eq.) was respectively introduced into the peptide chain mediated by 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU)/1-hydroxy-7-azabenzotriazole (HOAt) (1.9 eq.) and DIPEA (4 eq.). After assembly of glycopeptides, the *N*-terminal Fmoc group was removed and the resulting free amine was capped with acetic anhydride. The glycopeptides were cleaved from resins by TFA/TIPS/H_2O = 95/2.5/2.5 for 2 h. The excess TFA was evaporated. The glycopeptides were precipitated by diethyl ether and centrifuged to pellet the solid. To remove the acetyl protecting groups of the Tf, the crude glycopeptide was treated with 5% (v/v) hydrazine in water for 6 h. The crude reaction was neutralized to pH 7. To remove the methyl ester and *O*-acetate of the STn, the crude glycopeptide was treated with 1% 1 M NaOH (aq.) in methanol/H_2O (1:1) for 6 h. The deprotected glycopeptides were purified on a Shimadzu HPLC (LC-8A Liquid Chromatograph Pump, DGU-14A, Degasser and SPD-10A UV-Vis Detector), using reverse phase column SUPERCOSIL LC18, 25 cm × 10 mm 5 μm with gradient solvent CH_3CN and H_2O (0.1% TFA) gradient 0–5% in 2 min, 5–30% in 2–30 min (flow rate: 10 mL min^{-1}). For conjugation of the MUC1 glycopeptides onto Qβ, the purified glycopeptides were treated with adipate bis(4-nitrophenyl) ester **10** (5 eq.) in the presence of DIPEA (10 eq.) in DMF with 1% water for 1.5 h and then purified by HPLC to obtain MUC1 glycopeptides **1** and **2** with 20–30% yield. The products were characterized by ESI HRMS and ^{1}H NMR. MUC1-STn **1** (before installation of *p*-nitrophenyl ester linker), ESI ($C_{63}H_{104}N_{16}O_{27}$): calculated ([M+2H$^+$]/2): 759.3701, found: 759.3705; ^{1}H NMR (500 MHz, deuterium oxide) δ 4.49 – 4.44 (m, 2H), 4.42 – 4.35 (m, 4H), 4.29 – 4.25 (m, 2H), 4.20 – 2.13 (m, 4H), 3.94 – 3.88 (m, 3H), 3.81 – 3.62 (m, 12H), 3.60 – 3.39 (m, 12H), 3.10 – 3.01 (m, 5H), 2.85 – 2.77 (m, 4H), 2.69 (dd, *J* = 12.7, 4.4 Hz, 1 H, H3b-*eq* of sialic acid), 2.66 (d, *J* = 7.0 Hz, 1H), 2.56 – 2.51 (m, 2H), 2.18 – 2.07 (m, 5H), 1.89 (s, 3H, OAc), 1.86 (s, 3H, NHAc), 1.86 (s, 3H, NHAc), 1.78 – 1.66 (m, 6H), 1.56 – 1.45 (m, 7H), 1.44 – 1.38 (m,

3H), 1.24 – 1.19 (m, 5H, CH_3 (Thr)), 1.17 (dd, *J* = 6.9, 3.7 Hz, 1H), 1.12 (d, *J* = 6.3 Hz, 3H, CH_3 (Ala)). For MUC1-STn **1**, ESI ($C_{75}H_{115}N_{17}O_{32}$): calculated ([M+2$Na^+$]/2): 905.8839, found: 905.8835. For characterization of MUC1-Tf **2** (before installation of *p*-nitrophenyl ester linker): ESI ($C_{58}H_{97}N_{15}O_{24}$): calculated ([M+2H^+]/2): 694.8488, found: 694.8507. MUC1-Tf **2**, ESI ($C_{70}H_{108}N_{16}O_{29}$): calculated ([M+$H^+$+$Na^+$]/2): 830.3717, found: 830.3723; ^{1}H NMR (500 MHz, deuterium oxide) δ 8.28 (d, *J* = 5.5 Hz, 1H), 8.22 (d, *J* = 6.8 Hz, 1H), 8.19 – 8.16 (m, 2H), 7.80 – 7.73 (m, 2H, *p*-nitrophenyl), 7.23 – 7.20 (m, 2H, *p*-nitrophenyl), 4.49 – 4.44 (m, 2H), 4.39 – 4.34 (m, 3H), 4.30 (d, *J* = 7.8 Hz, 1H, H1b), 4.27 – 4.22 (m, 3H), 4.22 – 4.17 (m, 2H), 4.12 – 4.07 (m, 2H), 4.06 – 4.02 (m, 2H), 3.91 – 3.83 (m, 3H), 3.73 (d, *J* = 3.4 Hz, 1H), 3.69 – 3.53 (m, 10H), 3.52 – 3.30 (m, 10H), 3.09 – 2.93 (m, 8H), 2.77 – 2.71 (m, 2H), 2.62 – 2.53 (m, 4H), 2.15 – 2.02 (m, 7H), 1.89 (s, 3H, OAc), 1.87 – 1.78 (m, 7H, NHAc), 1.75 – 1.65 (m, 5H), 1.58 – 1.51 (m, 6H), 1.35 – 1.31 (m, 3H), 1.21 – 1.19 (m, 3H, CH_3 (Thr)), 1.11 – 1.06 (m, 3H, CH_3 (Ala)).

Synthesis of Qβ-MUC1 conjugates 16 and 17

For synthesis of Qβ-MUC1-STn **16:** A solution of Qβ (2.1 mg, 0.148 μmol subunit, 0.59 μmol reactive amine) in 0.1 M K-Phos buffer pH 7.0 (0.363 mL) was cooled on an ice batch, and added to a frozen solution of MUC1-STn **1** (57 μL from a 50 mM stock solution in DMSO, 2.83 μmol) kept in an Eppendorf tube. For synthesis of Qβ-MUC1-Tf **17:** A solution of Qβ (2.3 mg, 0.159 μmol subunit, 0.636 μmol reactive amine) in 0.1 M K-Phos buffer pH 7.0 (0.391 mL) was cooled on an ice batch, and was added to a frozen solution of MUC1-Tf **2** (61 μL from a 50 mM stock solution in DMSO, 3.0 μmol) kept in an Eppendorf tube. The mixture was allowed to warm to rt and was gently inverted several times to ensure mixing of the reactants. Once the DMSO had completely dissolved, the reaction was incubated at 37 °C for 16 h. The reaction mixture was purified by PD-10 size exclusion chromatography eluting 0.1 M K-Phos buffer (pH 7.0). The isolated fractions were subjected to centrifugal filtration (MWCO 100kDa) to concentrate the sample to 2-3 mg/mL (the sample should not be concentrated to more than 5 mg/mL to prevent aggregate formation). The extent of particle modification was determined by ESI-TOF LC-MS and by electrophoretic analysis. The total protein concentration was determined by Bradford assay against BSA standards. Percent protein recovery was approx. 65–75 %. The particle integrity was determined by FPLC and DLS analysis.

Synthesis of BSA-MUC1 conjugates 18 and 19

For synthesis of BSA-MUC1 conjugates, 4 mg of MUC1-STn **1** and MUC1-Tf **2** were dissolved in 30 μl of DMSO respectively and then added to a solution of BSA (10 mg ml^{-1}) in pH = 7.0, 0.1 M K-Phos buffer (200 μL) in an Eppendorf tube. The reaction was inverted gently several times and incubated at 37 ºC overnight. The product was purified by an Amicon Ultra filter (10 kDa MW cut-off) against 0.1 M K-Phos. The total protein content was quantified by Bradford assay against BSA standards. The extent of modification was determined by MALDI-TOF MS.

Evaluation of antibody titers by ELISA

The Nunc MaxiSorp® flat-bottom 96-well microtiter plates were coated with 10 μg mL^{-1} of the corresponding BSA-MUC1 conjugates **18** and **19** (100 μL/well) in $NaHCO_3/Na_2CO_3$ buffer (0.05 M, pH 9.6) containing 0.02 % NaN_3 by incubation at 4 ºC overnight. The coated plates were

washed with PBS/0.5% Tween-20 (PBST) (4 × 200 μL) and blocked with 1 % BSA in PBS (100 μL/well) at rt for 1 h. The plates were washed again with PBST (4 × 200 μL) and incubated with serial dilutions of mice sera in 0.1 % BSA/PBS (100 μL/well, 3 wells for each dilution). The plates were incubated for 2 h at 37 °C and then washed with PBST (4 × 200 μL). A 1:2000 dilution of HRP-conjugated goat anti-mouse IgG, IgG1, IgG2b, IgG2c, IgG3 or IgM (Jackson ImmunoResearch Laboratory) in 0.1% BSA/PBS (100 μL) was added to the wells respectively to determine the titers of antibodies generated. The plates were incubated for 1 h at 37 °C and then washed with PBST (4 × 200 μL). A solution of enzymatic substrate 3,3',5,5'-tetramethylbenzidine (TMB, 200 μL) was added to the plates (for one plate: 5 mg of TMB was dissolved in 2 mL of DMSO plus 18 mL of citric acid buffer containing 20 μL of H_2O_2). Color was allowed to develop for 15 min and then quenched by adding 50 μL of 0.5 M H_2SO_4. The readout was measured at 450 nm using a microplate reader. The titer was determined by regression analysis with log10 dilution plotted with optical density and reported as the highest fold of dilution giving the optical absorbance value of 0.1 over those of the pre-immune control sera (OD = 0.2). All samples were performed in three replicates. For **Figure S4**, the antibody titers were calculated as half-max by a nonlinear regression, followed by a four-parameter logistic curve analysis.

Detection of antibody binding to tumor cells by FACS

B16-MUC1 cells, MCF-7 cells or MCF-10A cells were respectively cultured at 37 °C under 5% CO_2 in cell growth medium. The number of cells was determined using a hemocytometer. Suspensions of 3.0×10^5 cells were added to each of the 1.5 mL microcentrifuge tubes, then centrifuged at 1,600 rpm for 5 min to remove the supernatant. The cell pellets were washed with FACS buffer (1% FBS in PBS with 0.1 % NaN_3), and incubated with 1:20 dilution of mouse sera in FACS buffer (100 μL) for 30 min on ice. The incubated cells were washed twice with FACS buffer, followed by incubation with FITC conjugated goat anti-mouse IgG (minimal x-reactivity) antibody (Biolegend, 2 μL, 0.5 mg mL^{-1}) for 30 min on ice. The cells were washed twice, re-suspended in FACS buffer and analyzed by LSR II (BD Biosciences). Data was processed by FlowJo software.

Complement dependent cytotoxicity

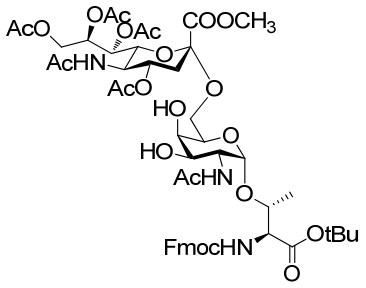
Complement dependent cytotoxicity of B16-MUC1 cells was determined by MTS assay. B16-MUC1 (7000 cells/well) were cultured in 96 well plate with DMEM (10% FBS, G-418 and 1% P.S.) for 12 h. The culture medium was carefully removed. A dilution of mouse sera (1/40) from different groups of immunized MUC1.Tg mice in 50 μL of DMEM (1% FBS, G-418 and 1% P.S.) were respectively added to the plate and incubated for 1 h at 37 ^{0}C. Then baby rabbit complement (CL3441, Cedarlane) at a dilution (1/5) in 50 μL of DMEM (10% FBS, G-418 and 1% P.S.) were added and incubated at 37 $^{\circ}$C for 3 h. MTS (CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay; Promega, 20 μL) was added into each well and further incubated at 37 ^{0}C for 3 h. The optical absorption of the MTS assay was measured at 490 nm. Complement alone treated cells were used as a positive control (maximum OD), and 5% Triton X-100 treated cells were used as a negative control (minimum OD). All data were performed in three replicates. Cytotoxicity was calculated as follows: Cytotoxicity (%) = (OD positive control – OD experimental) / (OD positive control – OD negative control) × 100.

Glycopeptide microarray analysis

Glycopeptides and glycoproteins were dissolved in sodium phosphate buffer (150 mM, pH 8.5) at a final concentration of 50 mM and transferred into a 384-well microtiter plate. All arrays were printed on NHS-activated hydrogel microarray glass slides (Schott, Nexterion, slide H). The spotter settings were adjusted to generate substrate peptide spots of 100 pL ± 3 pL using a piezo non-contact microarray spotter (M2-Automation). Each glycopeptide was printed in 8 spot replicates with 450 μm pitch (spot to spot distance) in an array format containing 8 wells with 2 blocks of each 10 × 13 spots. During the spotting process the humidity was kept between 50–60%. The glycopeptides were immobilized on the microarray slides in a humidity chamber (85–95% humidity) by incubation overnight. The unreacted NHS groups were capped by treatment with 25 mM ethanolamine in sodium borate buffer (100 mM, pH 8.5). The antisera were diluted at different concentrations in PBS/0.05% Tween-20 and incubated for 1 h. After washing with PBS/0.05% Tween-20, the slides were incubated with an Alexa Fluor 488-conjugated goat anti-mouse IgG antibody (dilution 1:2000 in PBS/0.05% Tween-20). Then the slides were washed and dried by centrifugation. For readout, the slides were scanned at 10 μm resolution by a fluorescence scanner

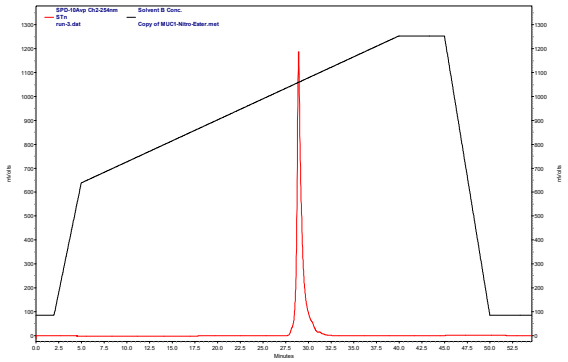
(Typhoon Trio+, Amersham) using a 520 nm emission filter (520 BP 40), a blue (488) excited mode laser and a photomultiplier tube (PMT) at 600 according to standard settings. The obtained image was analyzed by ImageQuant TL array image analysis software. An array grid of 26 × 20 was fitted around the spot area. Background was automatically removed according to the ImageQuant "Spot Edge Average" method, which provides a good localized background intensity and is relative tolerant of noise in the image. The obtained data was then imported into excel and the mean and standard deviation were calculated from 8 replicate spots per glycopeptide.

Characterization data and spectra of building blocks and MUC1 glycopeptides

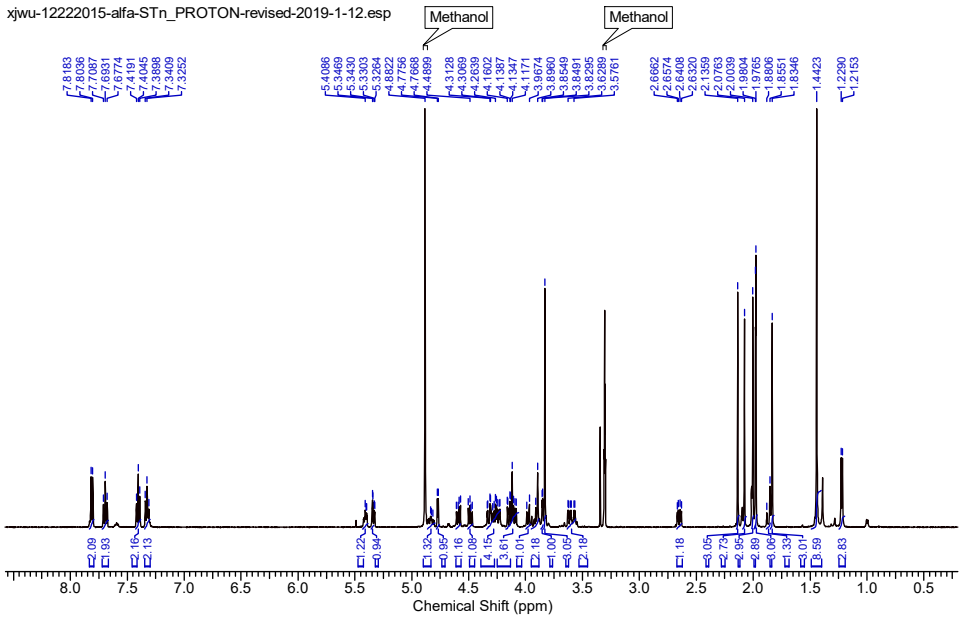


8

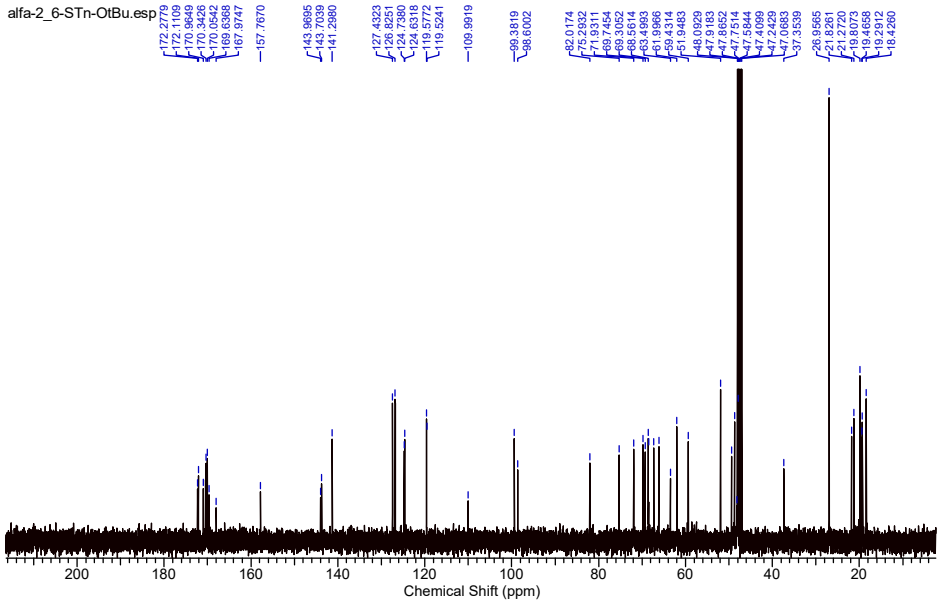
HPLC chromatogram of **8**



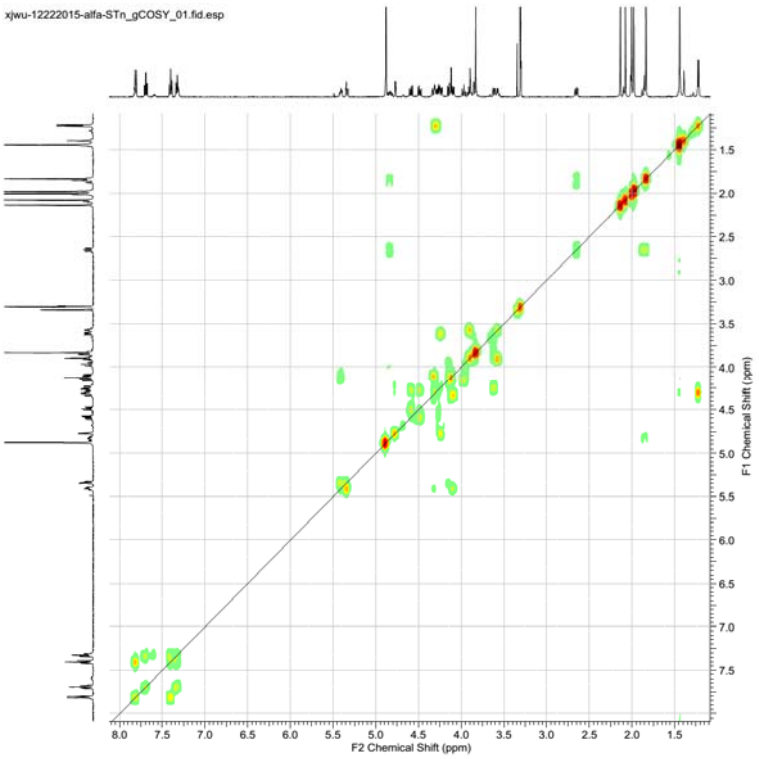
^{1}H NMR (500 MHz, CD_3OD) of **8**



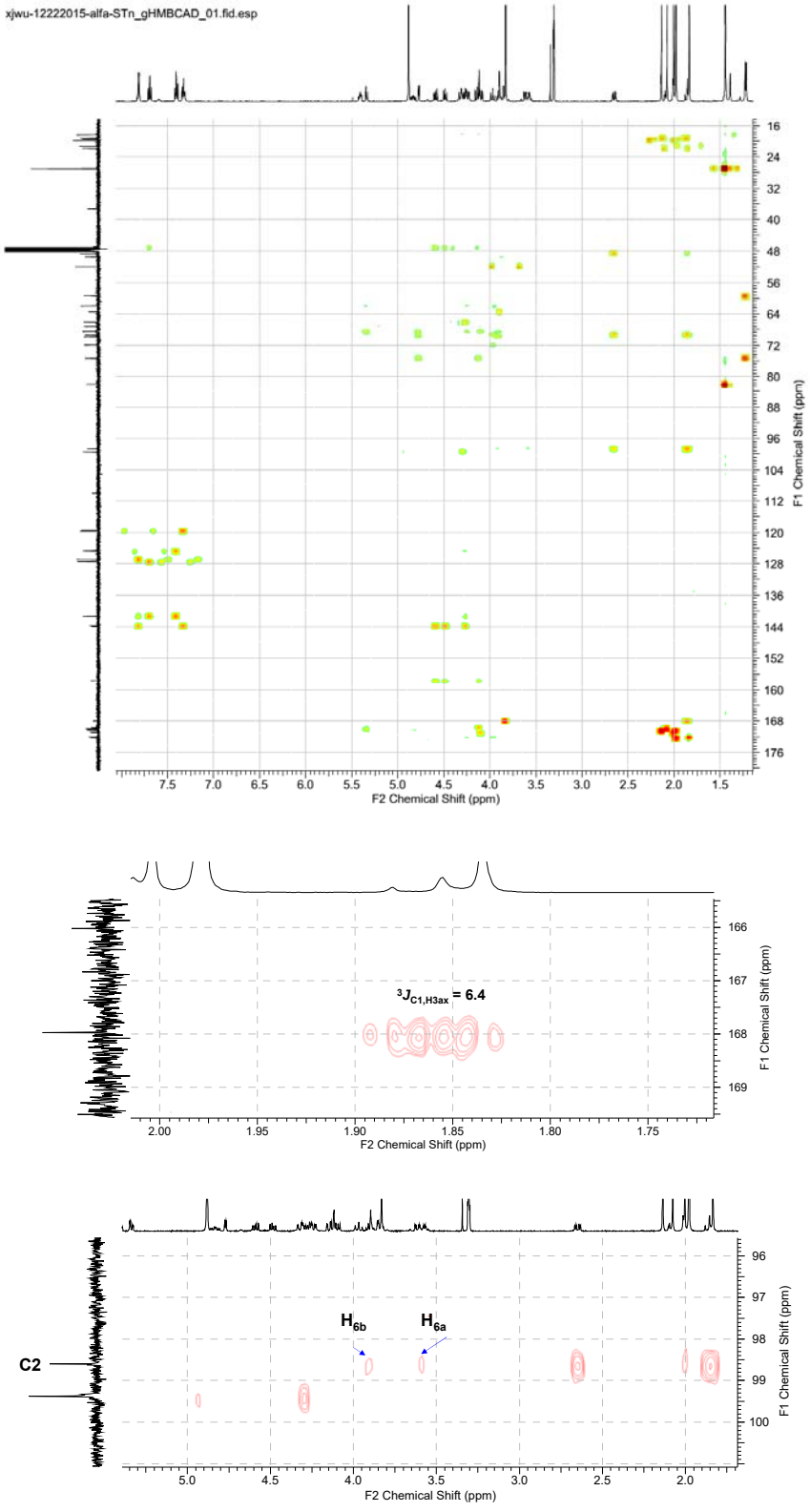
^{13}C NMR (126 MHz, CD_3OD) of **8**

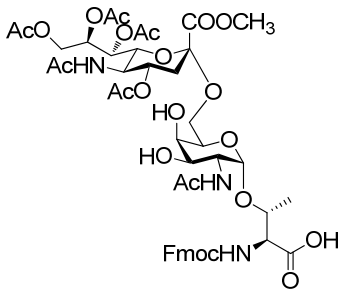


^{1}H-^{1}H COSY (500 MHz, CD_3OD) of **8**



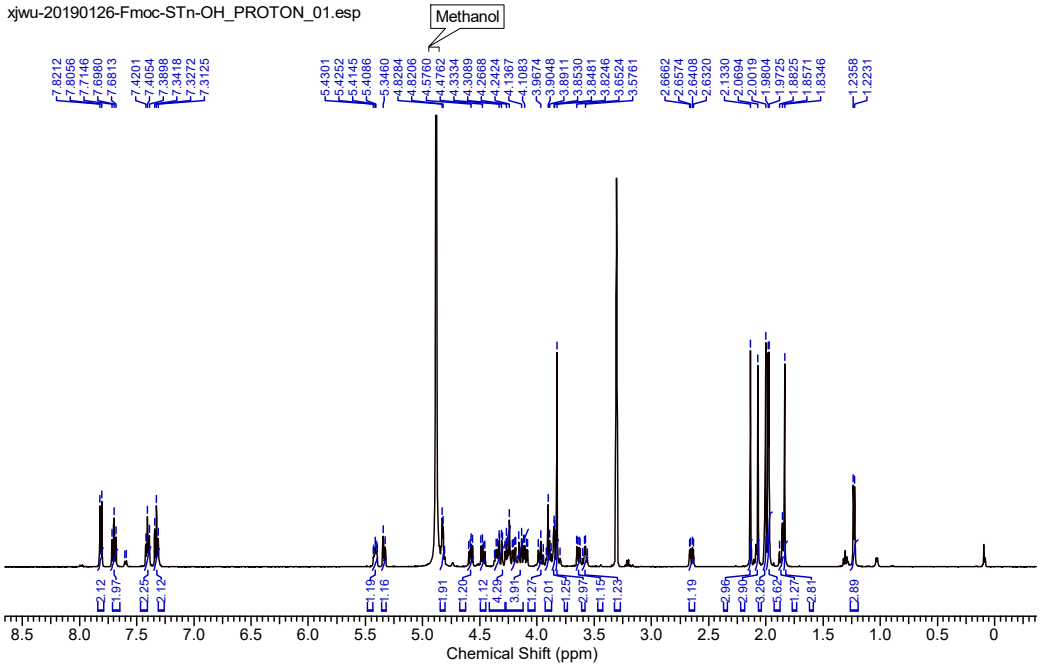
^{1}H-^{13}C HMBC (500 MHz, CD_3OD) of **8**



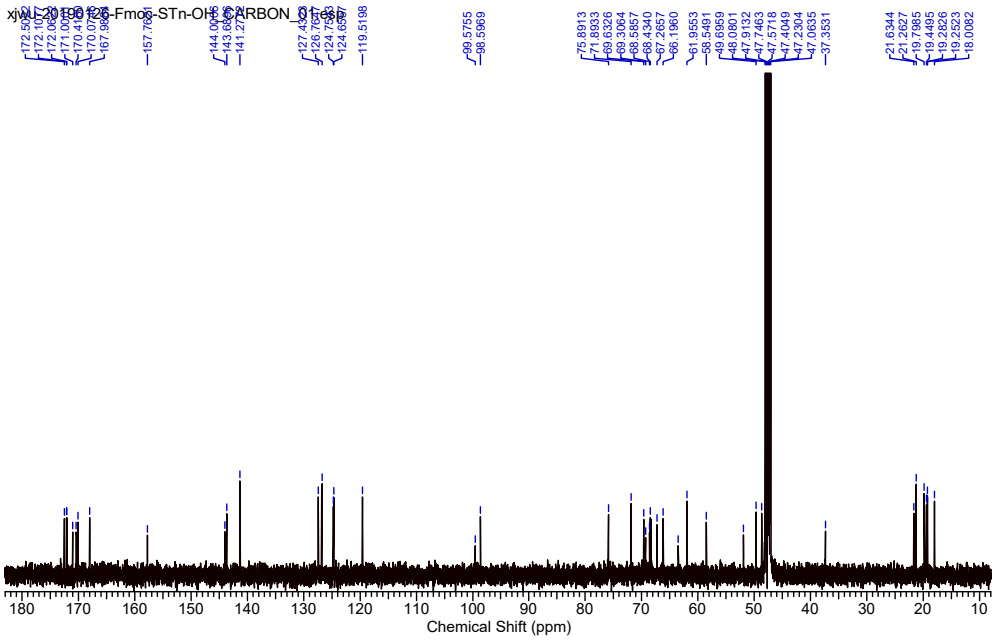


5

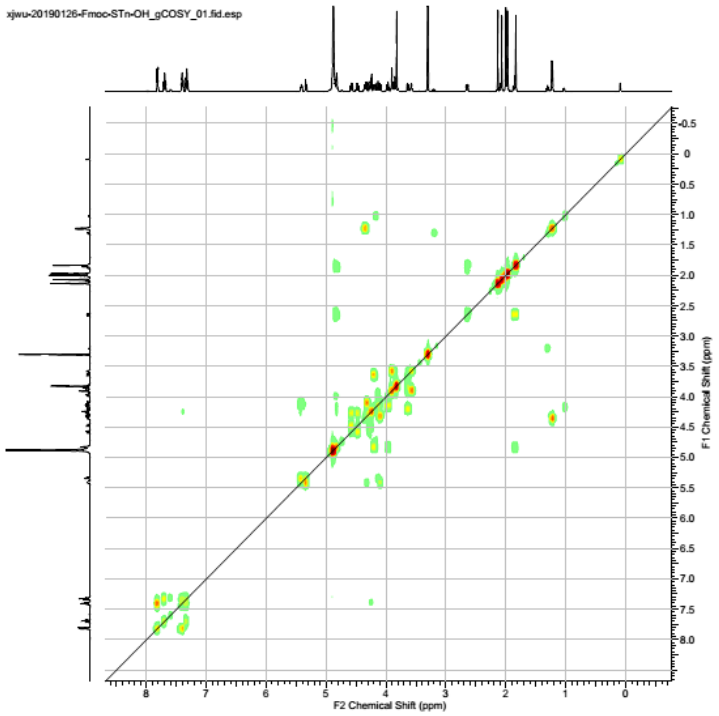
^{1}H NMR (500 MHz, CD_3OD) of **5**



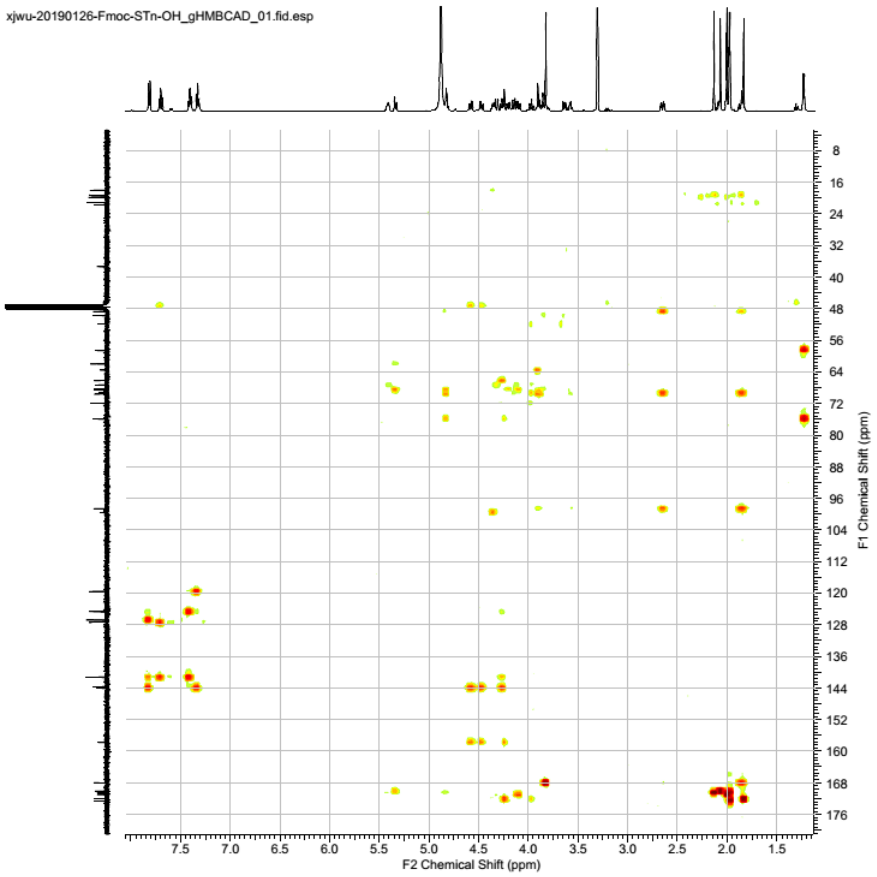
^{13}C NMR (126 MHz, CD_3OD) of **5**

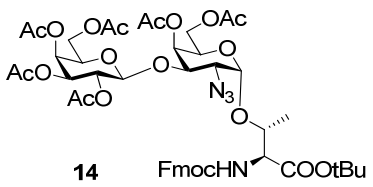


^{1}H-^{1}H COSY (500 MHz, CD_3OD) of **5**



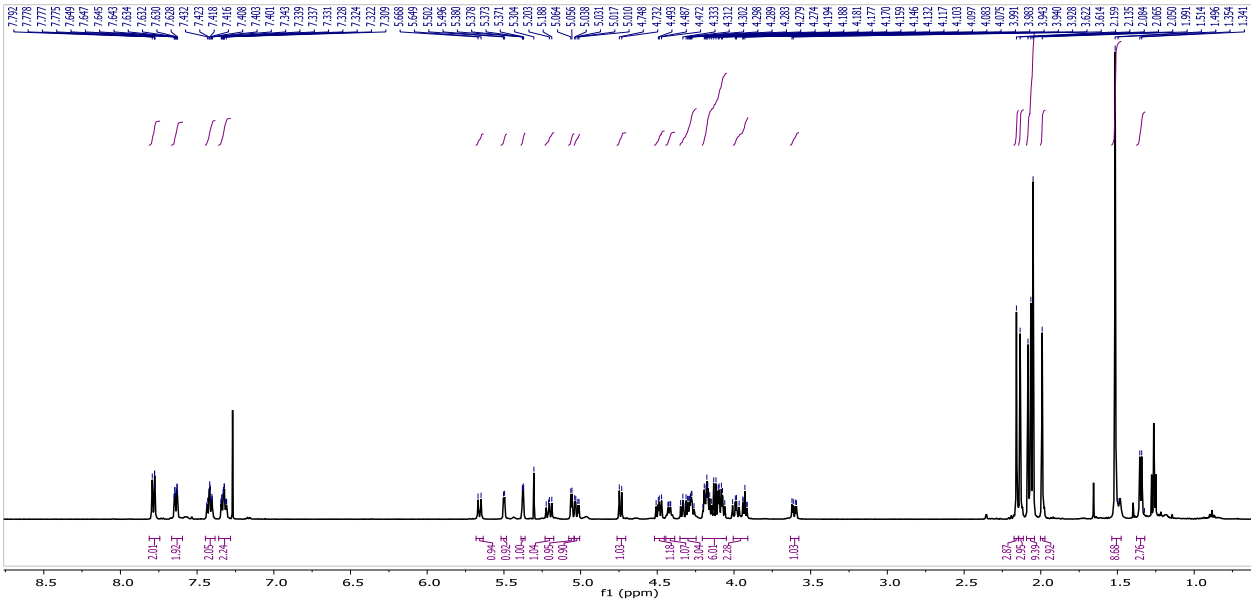
^{1}H-^{13}C HMBC (500 MHz, CD_3OD) of **5**



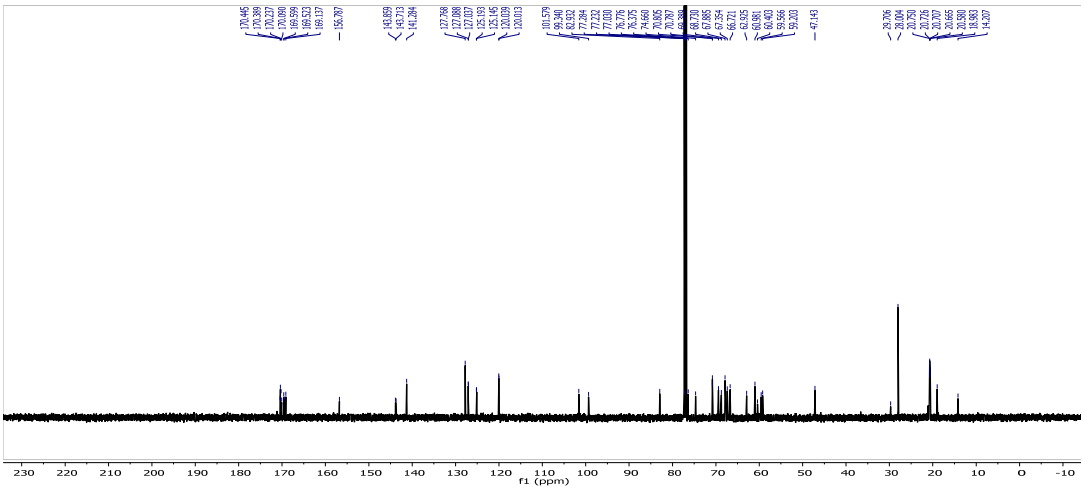


14

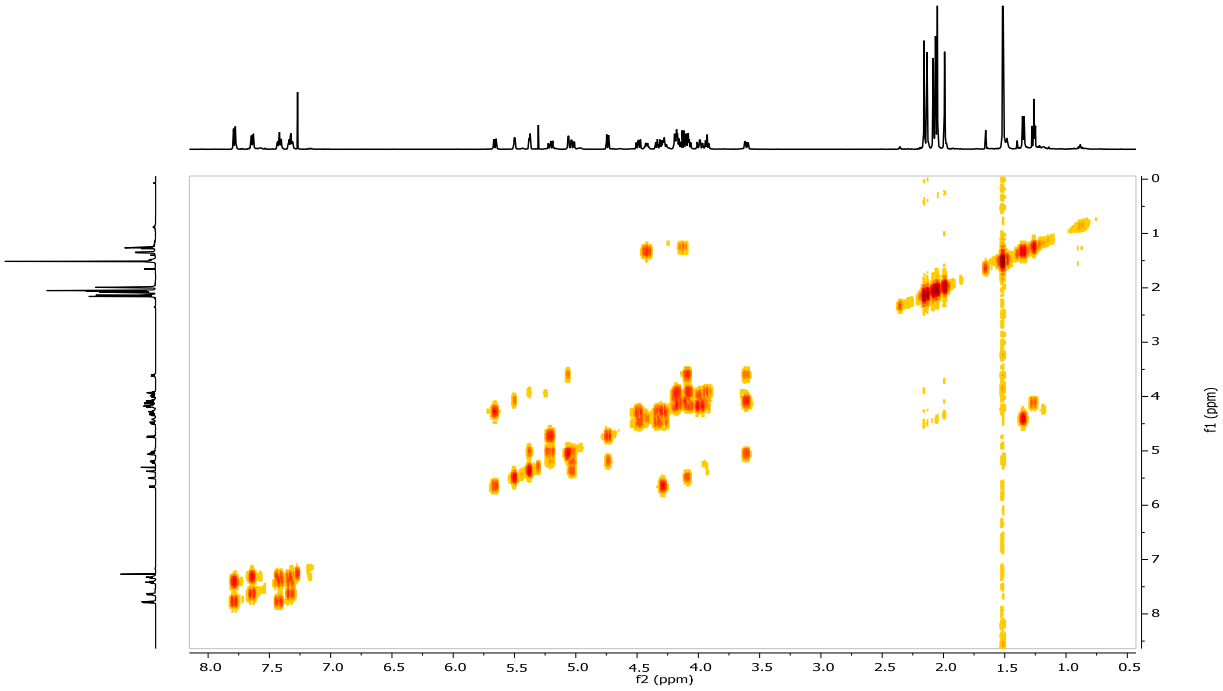
^{1}H NMR (500 MHz, $CDCl_3$) of **14**



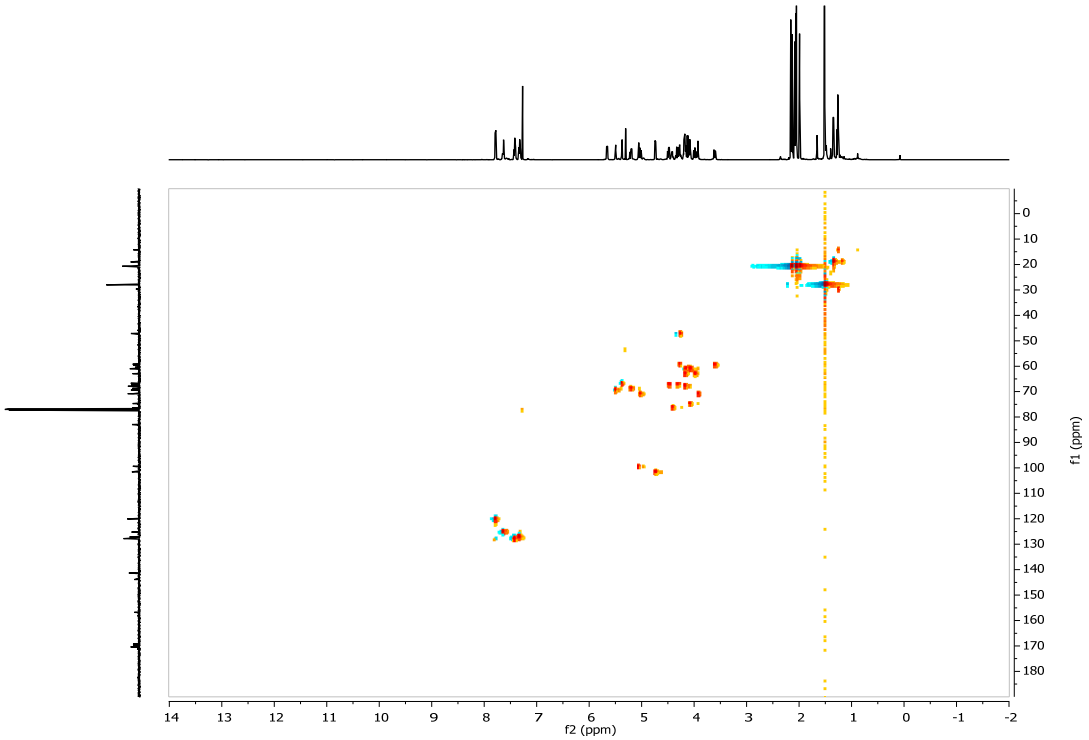
^{13}C NMR (126 MHz, $CDCl_3$) of **14**

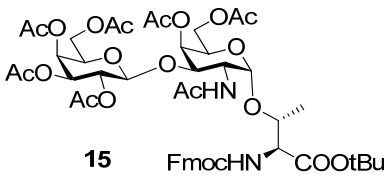


gCOSY (500 MHz, $CDCl_3$) of **14**



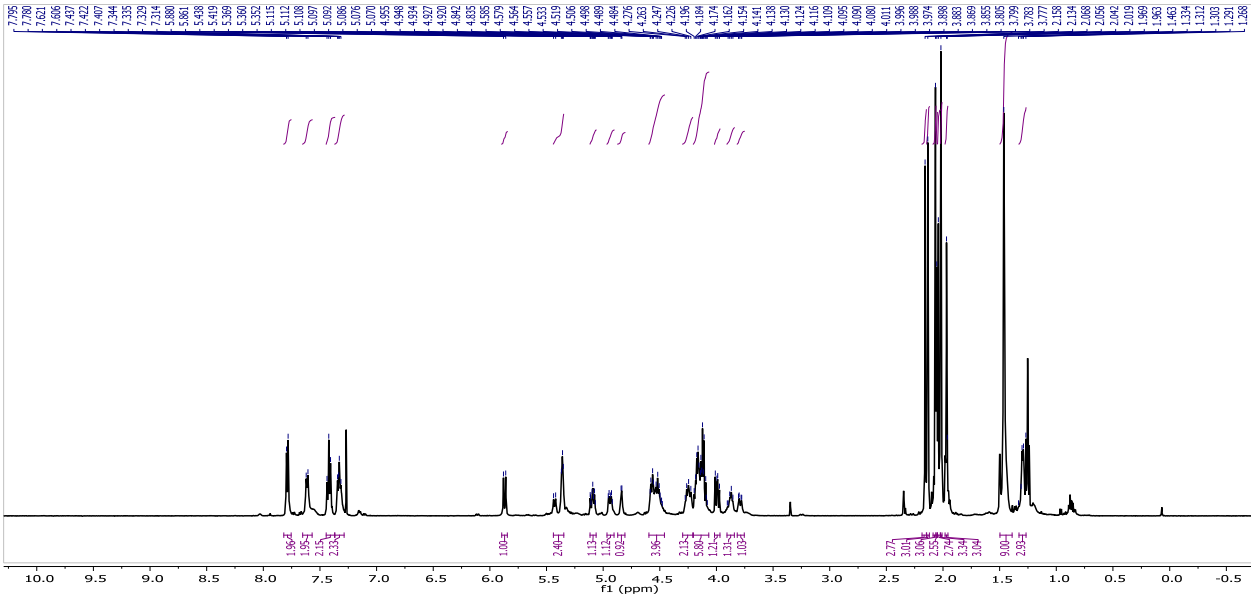
^{1}H-^{13}C HSQC (500 MHz, $CDCl_3$) of **14**



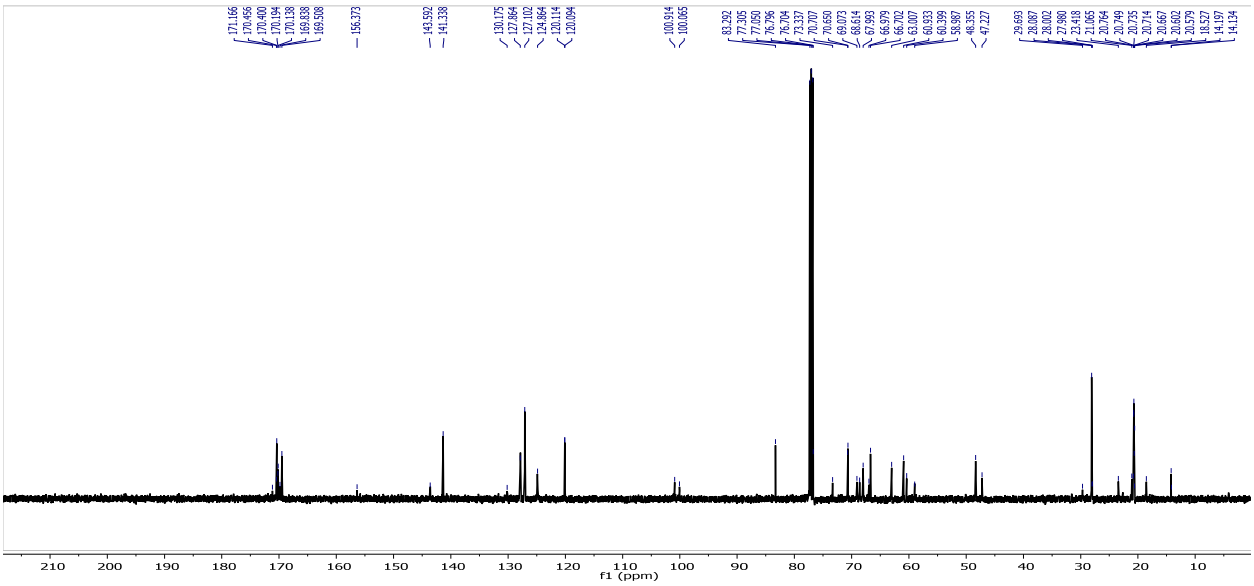


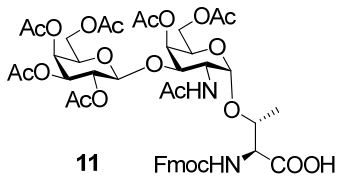
15

^{1}H NMR (500 MHz, $CDCl_3$) of **15**



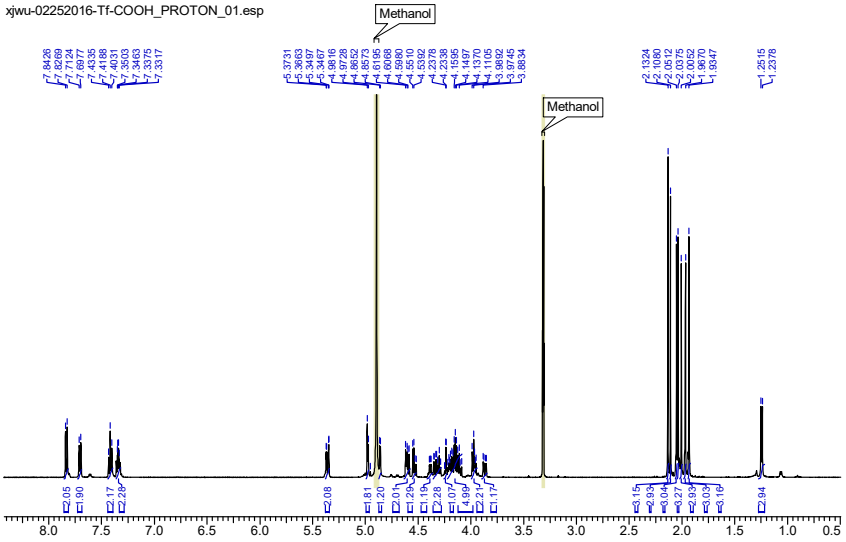
^{13}C-NMR (126 MHz, $CDCl_3$) of **15**



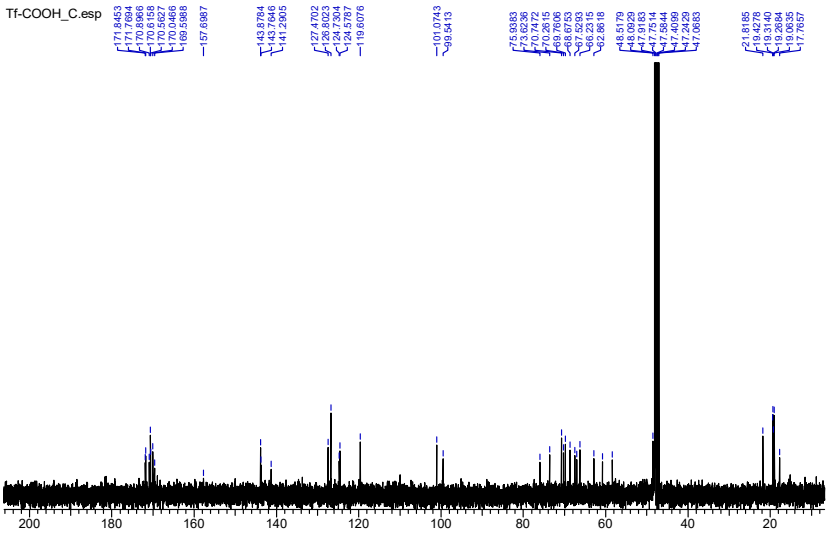


11

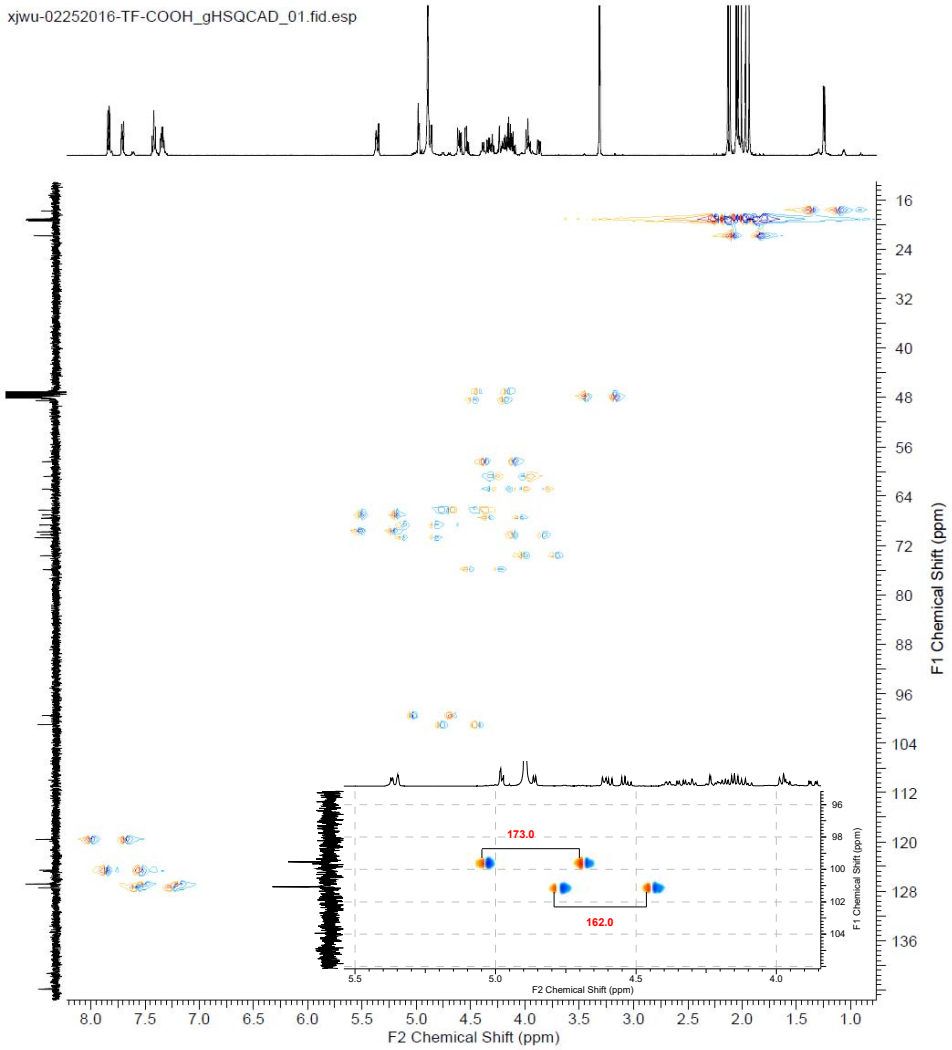
^{1}H NMR (500 MHz, CD_3OD) of **11**

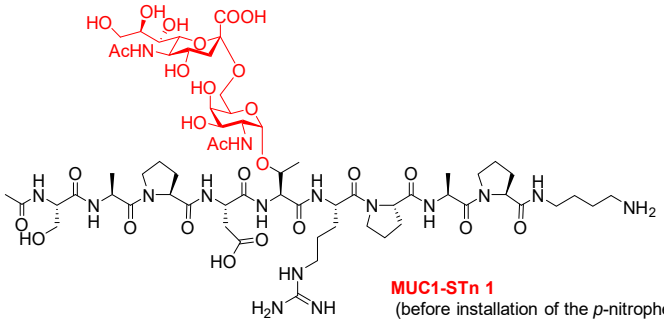


^{13}C NMR (126 MHz, CD_3OD) of **11**



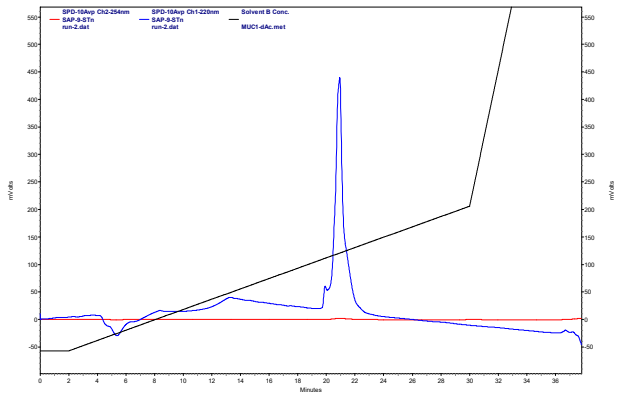
^{1}H-^{1}H COSY (500 MHz, CD_3OD) of **11**



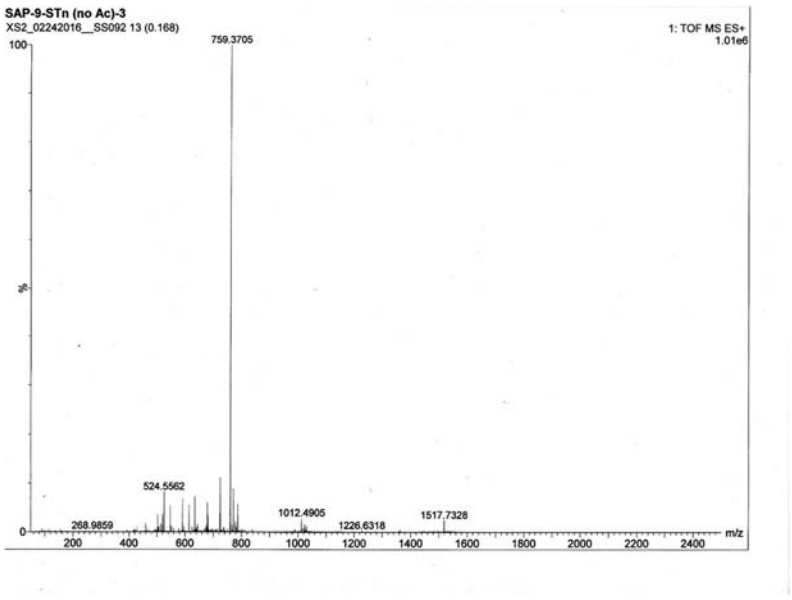


MUC1-STn 1
(before installation of the *p*-nitrophenyl ester linker)

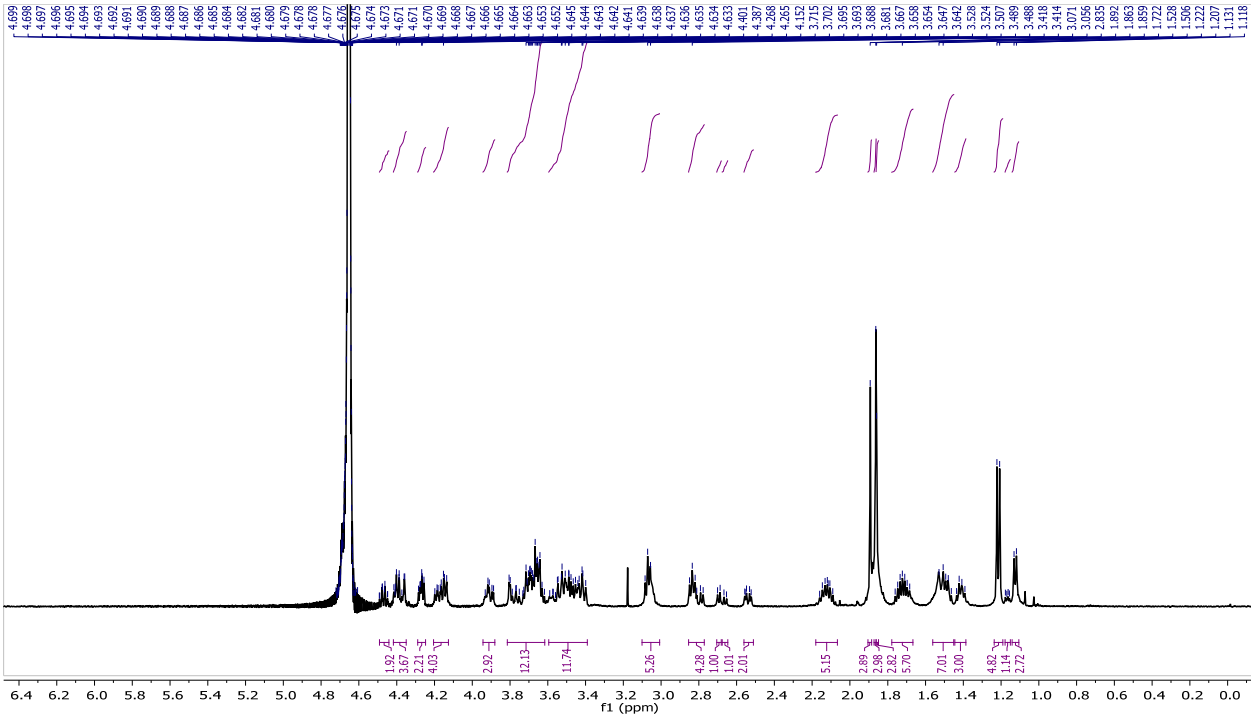
HPLC chromatogram of MUC1-STn **1** (before installation of the p-nitrophenyl ester linker)

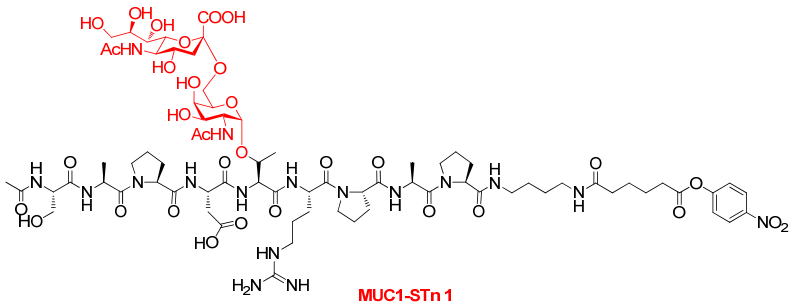


HRMS of MUC1-STn **1** (before installation of the p-nitrophenyl ester linker)

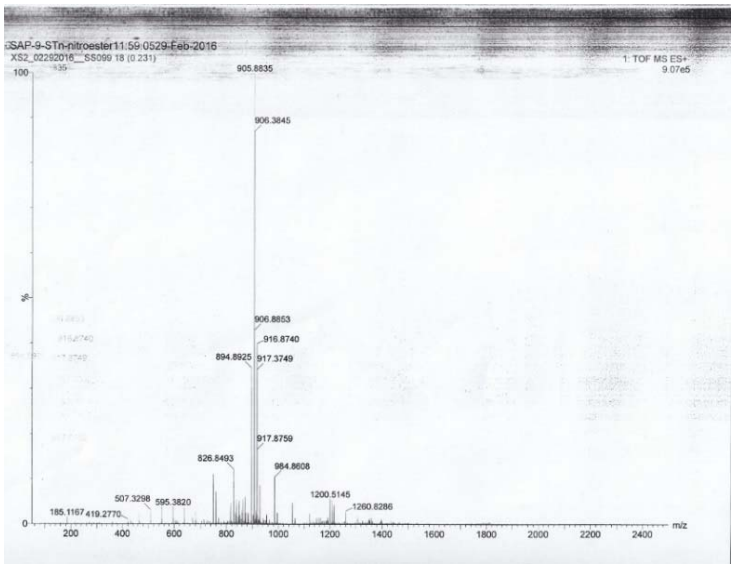


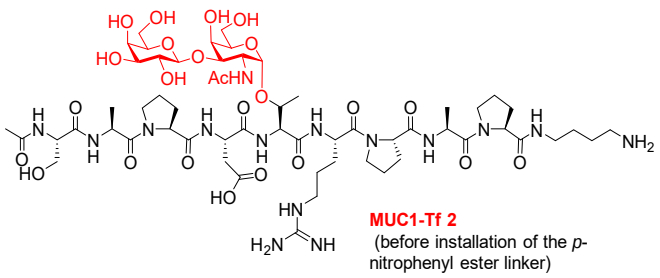
^{1}H NMR of MUC1-STn **1** (before installation of the *p*-nitrophenyl ester linker)





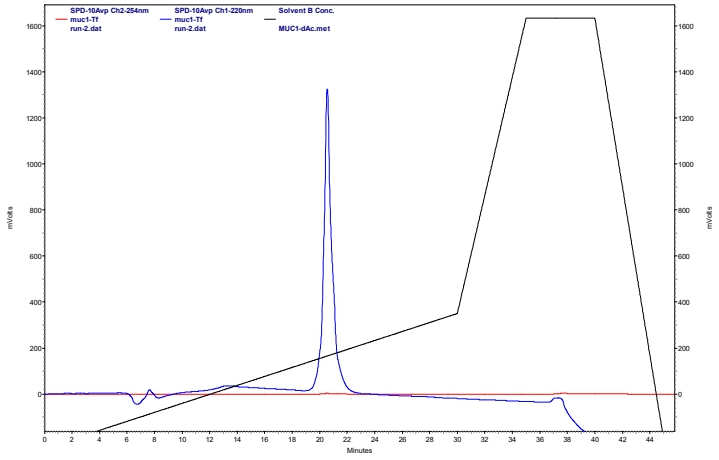
HRMS of MUC1-STn **1**



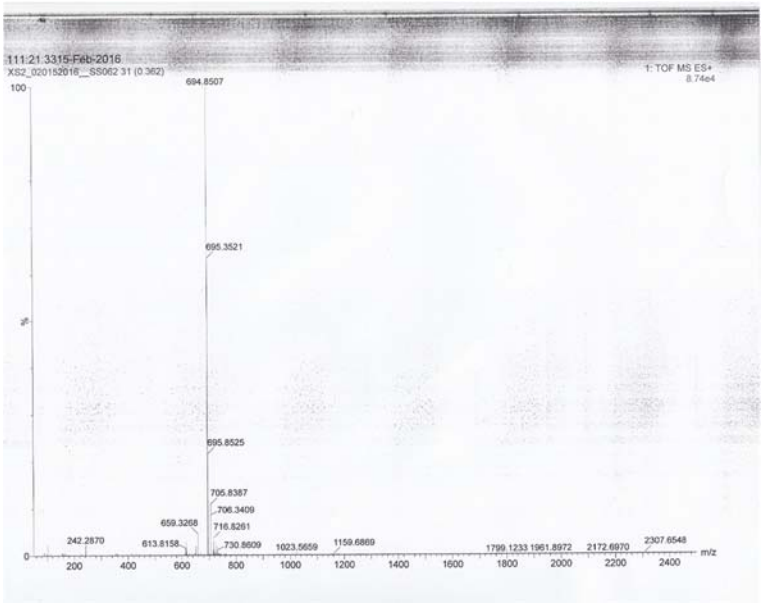


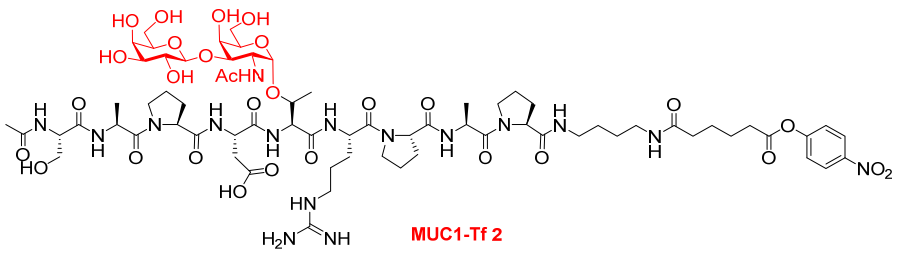
MUC1-Tf 2
(before installation of the *p*-nitrophenyl ester linker)

HPLC chromatogram of MUC1-Tf **2** (before installation of the *p*-nitrophenyl ester linker)

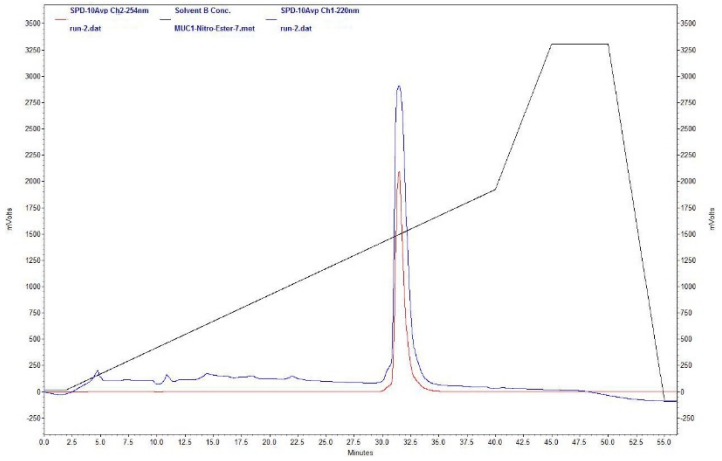


HRMS chromatogram of MUC1-Tf **2** (before installation of the p-nitrophenyl ester linker)

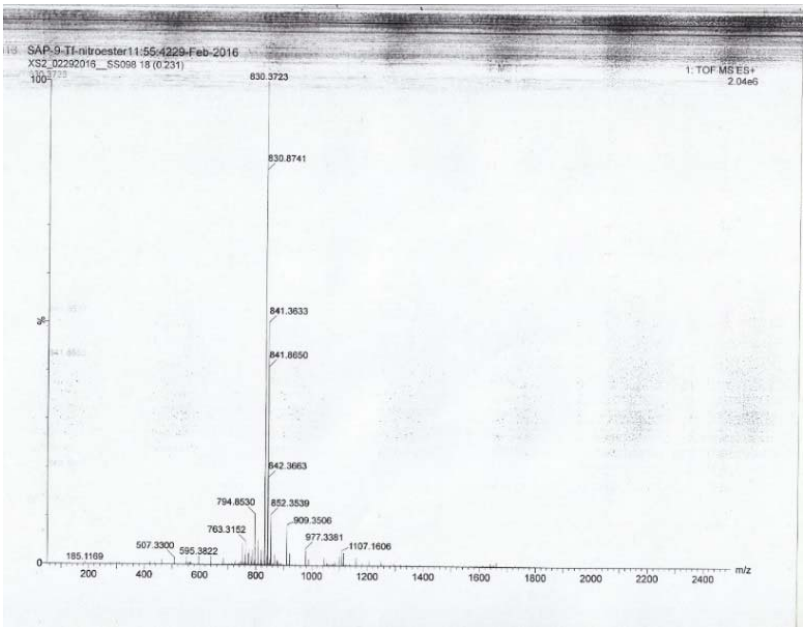




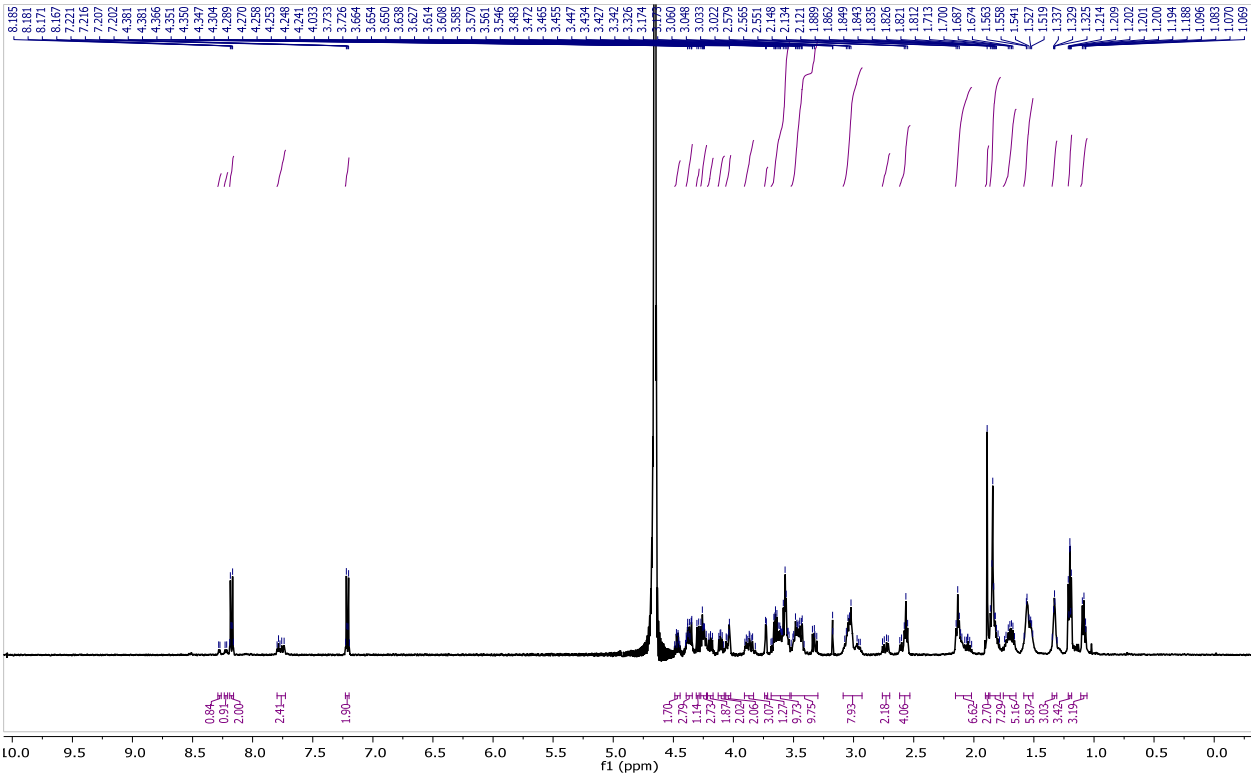
HPLC chromatogram of MUC1-Tf **2**



HRMS chromatogram of MUC1-Tf **2**



^{1}H NMR of MUC1-Tf **2**



References:

(1) Lavrsen, K., Madsen, C., Rasch, M., Woetmann, A., Ødum, N., Mandel, U., Clausen, H., Pedersen, A., and Wandall, H. (2013) Aberrantly glycosylated MUC1 is expressed on the surface of breast cancer cells and a target for antibody-dependent cell-mediated cytotoxicity. *Glycoconjugate J. 30*, 227-236.

(2) Sungsuwan, S., Yin, Z., and Huang, X. (2015) Lipopeptide-coated iron oxide nanoparticles as potential glycoconjugate-based synthetic anticancer vaccines. *ACS Appl. Mater. Interfaces 7*, 17535-17544.

(3) Liebe, B., and Kunz, H. (1997) Solid‐phase synthesis of a tumor‐associated sialyl‐T_N antigen glycopeptide with a partial sequence of the "tandem repeat" of the MUC‐1 mucin. *Angew. Chem. Int. Ed. 36*, 618-621.

(4) Yin, Z., Dulaney, S., McKay, C., Baniel, C., Kaczanowska, K., Ramadan, S., Finn, M. G., and Huang, X. (2016) Chemical synthesis of GM2 glycans, bioconjugation with bacteriophage Qβ and the induction of anti-cancer antibodies. *ChemBiochem 17*, 174-180.

(5) Nilsson, J., Brinkmalm, G., Ramadan, S., Gilborne, L., Noborn, F., Blennow, K., Wallin, A., Svensson, J., Abo-Riya, M. A., and Huang, X. (2019) Synthetic standard aided quantification and structural characterization of amyloid-beta glycopeptides enriched from cerebrospinal fluid of Alzheimer's disease patients. *Sci. Rep. 9*, 5522.