Supporting information

Synthesis and Immunological Evaluation of the Unnatural β-linked Mucin-1 Thomsen-Friedenreich Conjugate

Xuanjun Wu,[§†] Hunter McFall-Boegeman,[†,‡] Zahra Rashidijahanabad,[†,‡] Kunli Liu,[†,‡] Christian Pett,[‖,#] Jin Yu,[‖] Manuel Schorlemer,[‖,#] Sherif Ramadan,[†,‡¶] Sandra Behren,[‖,#] Ulrika Westerlind,[‖,#] 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

[‖]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

Table of Contents

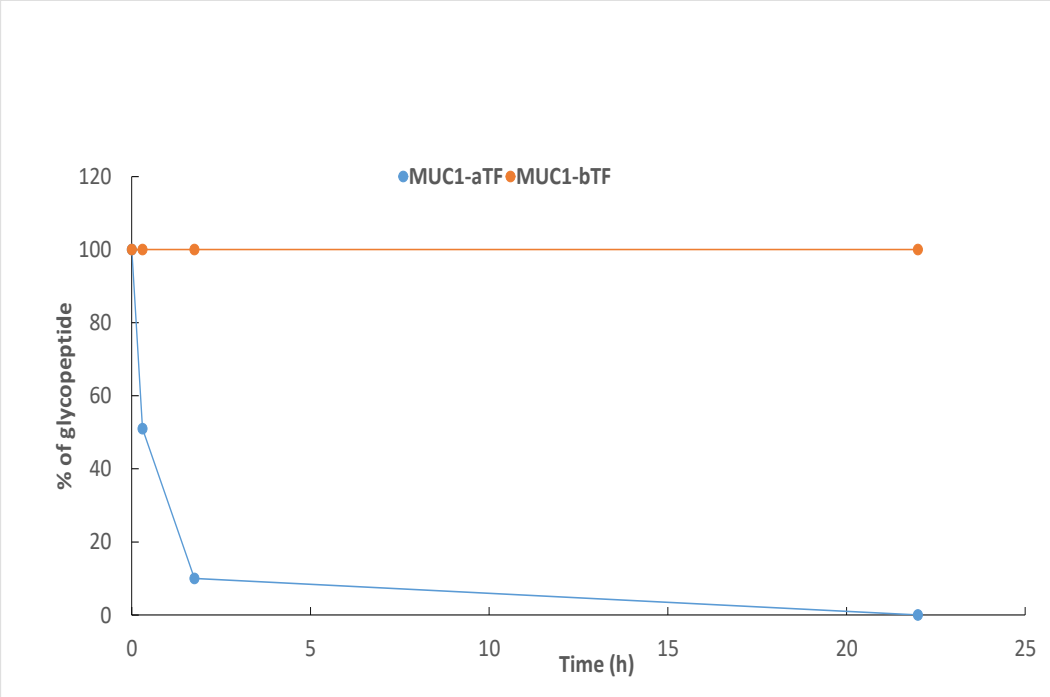


Figure S1. MUC1-β-Tf has much enhanced stability toward the *Enterococcus faecalis* endo-*α*-*N*-acetylgalactosaminidase compared to MUC1-α-Tf over 24h reaction time.

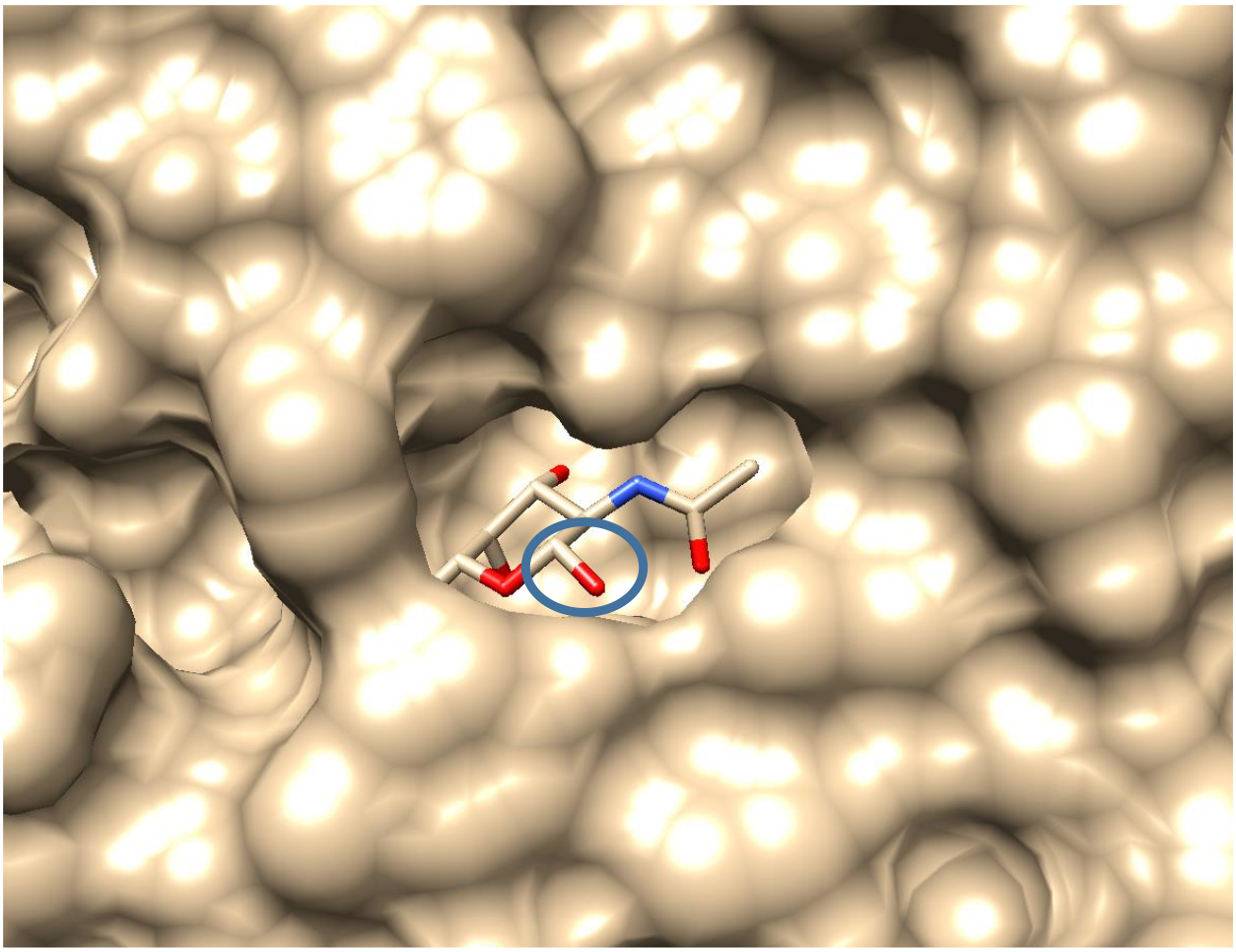


Figure S2. Crystal structure of the *Enterococcus faecalis* endo-*α*-*N*-acetylgalactosaminidase (PDB:6M77) with β-GalNAc showed that the 1-hydroxyl group (in blue oval) pointing toward the enzyme, presumably hindering the binding of MUC1-β-Tf glycopeptide **12** with the enzyme. This is consistent with the enhanced stability of MUC1-β-Tf glycopeptide **12** toward the enzyme compared to the corresponding MUC1-α-Tf glycopeptide **12α**. The enzyme is shown as spacing filling model in light brown color. β-GalNAc is shown as a stick structure (oxygen atoms are shown in red, nitrogen atom is shown as blue color, and carbon atoms are shown in gold color).

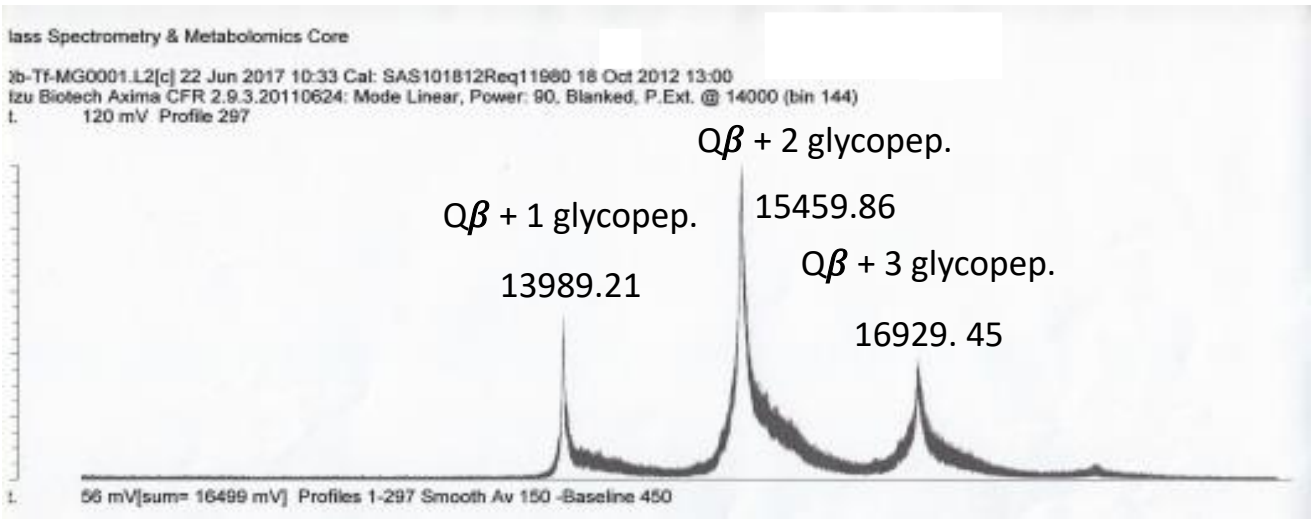
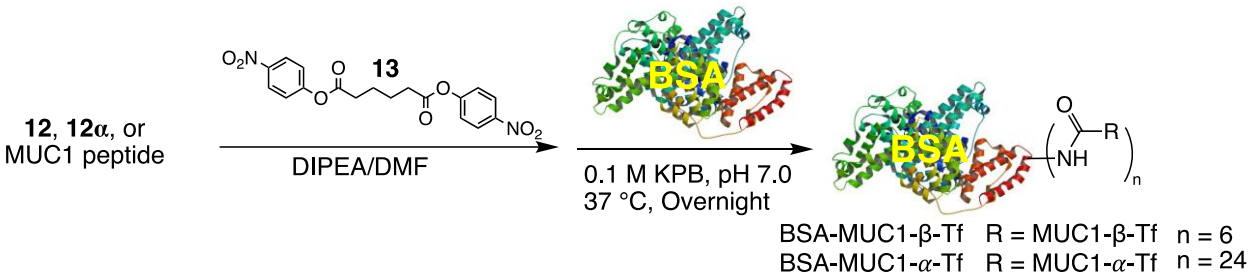


Figure S3. MALDI-TOF MS of Qβ-MUC1-β-Tf conjugate. Based on the ratio of peak intensities, the average number of MUC1-β-Tf per capsid was calculated to be 220.



Scheme S1: Synthesis of BSA-MUC1 conjugates for ELISA. BSA-MUC1, BSA-MUC1-α-Tf were reported previously.[1]

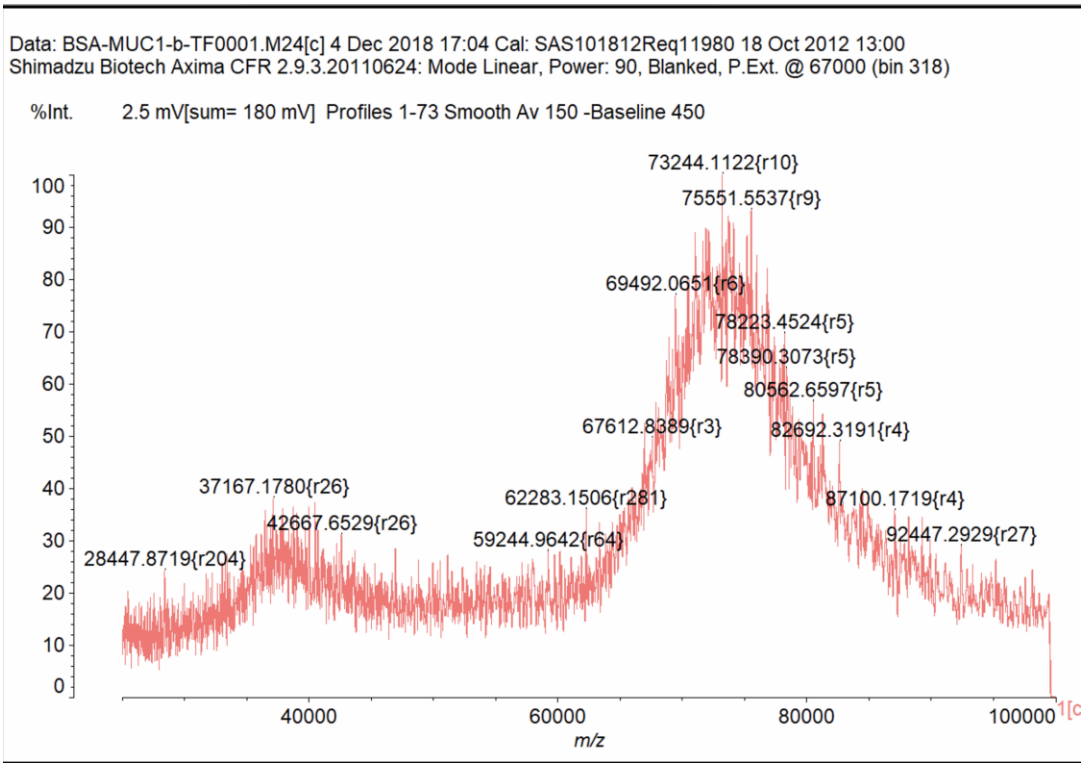


Figure S4. MALDI-TOF MS of BSA-MUC1-β-Tf conjugates. Based on the molecular weight difference between the BSA- MUC1-β-Tf conjugate and unmodified BSA, the number of MUC1-β-Tf per BSA was calculated to be 6.

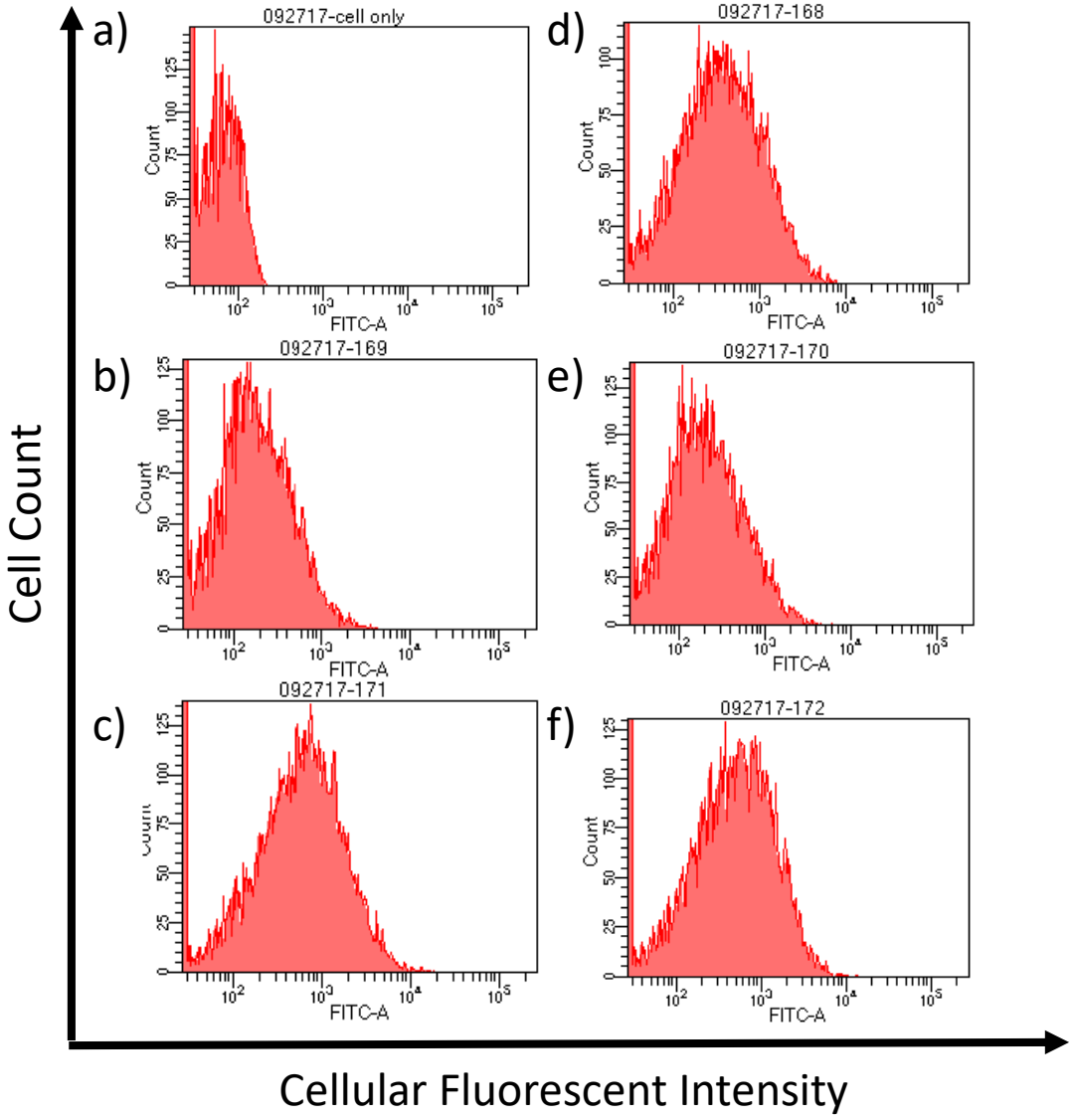
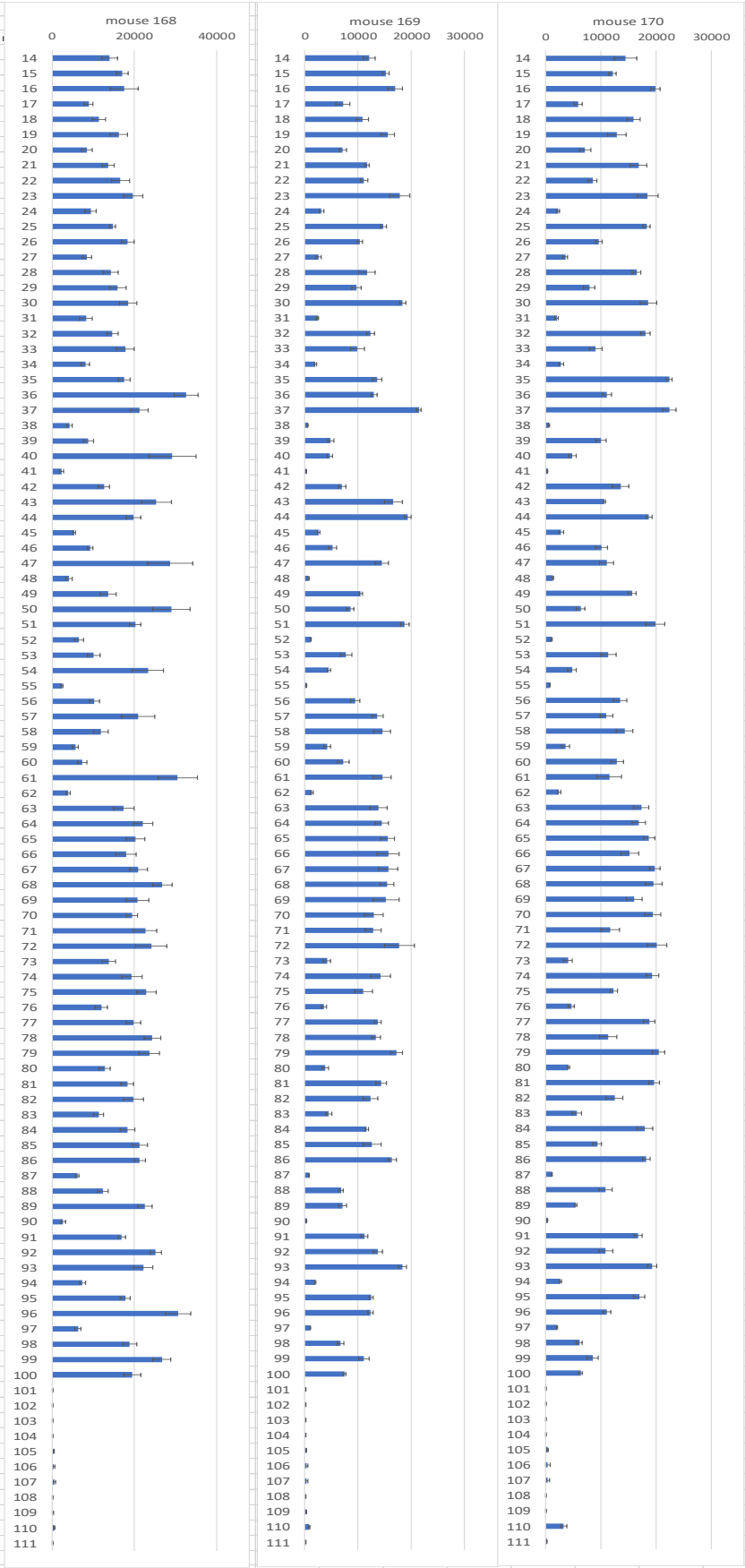
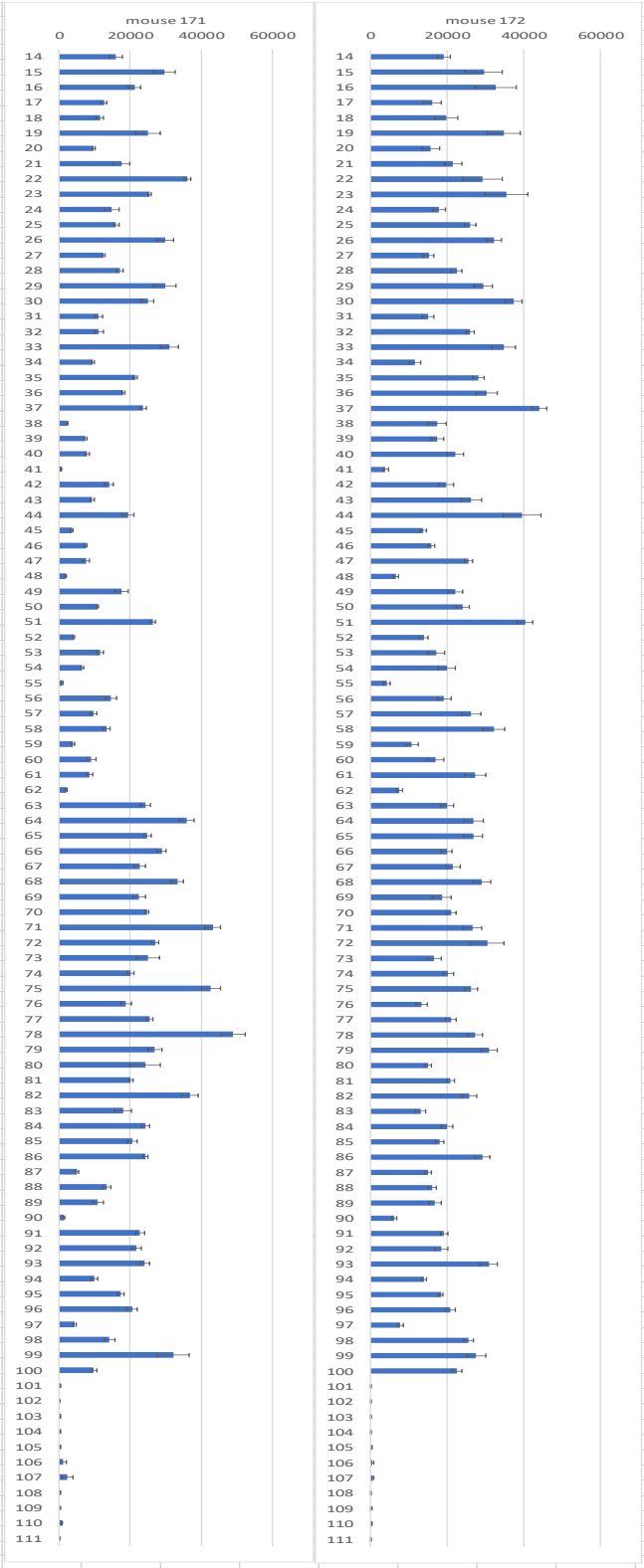


Figure S5. B16-MUC1 melanoma cells staining by IgG antibodies in post-immune sera elicited by Qβ-MUC1-β-Tf as measured by flow cytometry. a) B16-MUC1 unstained control; b-f) Sera staining of B16-MUC1 cells. Each curve represents serum from one mouse immunized with Qβ-MUC1-β-Tf. The binding was tested with 1:20 dilution of the sera.

PAHGVT*(T-a)SAPDTRPAPGSTA	14
PAHGVTSAPDT*(T-a)RPAPGSTA	15
PAHGVTSAPDTRPAPGST*(T-a)A	16
PAHGVT*(T-a)SAPDT*(T-a)RPAPGSTA	17
PAHGVT*(T-a)SAPDTRPAPGST*(T-a)A	18
PAHGVTSAPDT*(T-)RPAPGST*(T-a)A	19
PAHGVT*(T-a)SAPDT*(T-a)RPAPGST*(T-a)A	20
PAHGVT*(C1T1)SAPDTRPAPGSTA	21
PAHGVTSAPDT*(C1T1)RPAPGSTA	22
PAHGVTSAPDTRPAPGST*(C1T1)A	23
PAHGVT*(C1T1)SAPDT*(C1T1)RPAPGSTA	24
PAHGVT*(C1T1)SAPDTRPAPGST*(C1T1)A	25
PAHGVTSAPDT*(C1T1)RPAPGST*(C1T1)A	26
PAHGVT*(C1T1)SAPDT*(C1T1)RPAPGST*(C1T1)A	27
PAHGVT*(C1T2)SAPDTRPAPGSTA	28
PAHGVTSAPDT*(C1T2)RPAPGSTA	29
PAHGVTSAPDTRPAPGST*(C1T2)A	30
PAHGVT*(C1T2)SAPDT*(C1T2)RPAPGSTA	31
PAHGVT*(C1T2)SAPDTRPAPGST*(C1T2)A	32
PAHGVTSAPDT*(C1T2)RPAPGST*(C1T2)A	33
PAHGVT*(C1T2)SAPDT*(C1T2)RPAPGST*(C1T2)A	34
PAHGVT*(C2T1)SAPDTRPAPGSTA	35
PAHGVTSAPDT*(C2T1)RPAPGSTA	36
PAHGVTSAPDTRPAPGST*(C2T1)A	37
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGSTA	38
PAHGVT*(C2T1)SAPDTRPAPGST*(C2T1)A	39
PAHGVTSAPDT*(C2T1)RPAPGST*(C2T1)A	40
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGST*(C2T1)A	41
PAHGVT*(C2T1)SAPDTRPAPGSTA	42
PAHGVTSAPDT*(C2T1)RPAPGSTA	43
PAHGVTSAPDTRPAPGST*(C2T1)A	44
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGSTA	45
PAHGVT*(C2T1)SAPDTRPAPGST*(C2T1)A	46
PAHGVTSAPDT*(C2T1)RPAPGST*(C2T1)A	47
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGST*(C2T1)A	48
PAHGVT*(C2T2)SAPDTRPAPGSTA	49
PAHGVTSAPDT*(C2T2)RPAPGSTA	50
PAHGVTSAPDTRPAPGST*(C2T2)A	51
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGSTA	52
PAHGVT*(C2T2)SAPDTRPAPGST*(C2T2)A	53
PAHGVTSAPDT*(C2T2)RPAPGST*(C2T2)A	54
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGST*(C2T2)A	55
PAHGVT*(C2T2)SAPDTRPAPGSTA	56
PAHGVTSAPDT*(C2T2)RPAPGSTA	57
PAHGVTSAPDTRPAPGST*(C2T2)A	58
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGSTA	59
PAHGVT*(C2T2)SAPDTRPAPGST*(C2T2)A	60
PAHGVTSAPDT*(C2T2)RPAPGST*(C2T2)A	61
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGST*(C2T2)A	62
PAHGVT*(Tn-a)SAPDTRPAPGSTA	63
PAHGVTSAPDT*(Tn-a)RPAPGSTA	64
PAHGVTSAPDTRPAPGST*(Tn-a)A	65
PAHGVT*(Tn-a)SAPDT*(Tn-a)RPAPGSTA	66
PAHGVT*(Tn-a)SAPDTRPAPGST*(Tn-a)A	67
PAHGVTSAPDT*(Tn-a)RPAPGST*(Tn-a)A	68
PAHGVT*(Tn-a)SAPDT*(Tn-a)RPAPGST*(Tn-a)A	69
PAHGVT*(C3T1)SAPDTRPAPGSTA	70
PAHGVTSAPDT*(C3T1)RPAPGSTA	71
PAHGVTSAPDTRPAPGST*(C3T1)A	72
PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGSTA	73
PAHGVT*(C3T1)SAPDTRPAPGST*(C3T1)A	74
PAHGVTSAPDT*(C3T1)RPAPGST*(C3T1)A	75
PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T1)A	76
PAHGVT*(C3T2)SAPDTRPAPGSTA	77
PAHGVTSAPDT*(C3T2)RPAPGSTA	78
PAHGVTSAPDTRPAPGST*(C3T2)A	79
PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA	80
PAHGVT*(C3T2)SAPDTRPAPGST*(C3T2)A	81
PAHGVTSAPDT*(C3T2)RPAPGST*(C3T2)A	82
PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A	83
PAHGVT*(C4T1)SAPDTRPAPGSTA	84
PAHGVTSAPDT*(C4T1)RPAPGSTA	85
PAHGVTSAPDTRPAPGST*(C4T1)A	86
PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA	87
PAHGVT*(C4T1)SAPDTRPAPGST*(C4T1)A	88
PAHGVTSAPDT*(C4T1)RPAPGST*(C4T1)A	89
PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A	90
PAHGVT*SAPDTRPAPGSTA	91
PAHGVTSAPDT*(C4T2)RPAPGSTA	92
PAHGVTSAPDTRPAPGST*(C4T2)A	93
PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA	94
PAHGVT*(C4T2)SAPDTRPAPGST*(C4T2)A	95
PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A	96
PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A	97
PAHGVTSAPDT*RPAPGS*T*APPA Pep137	98
PAHGVTSAPDT*RPAPGS*T*APPA Pep138	99
PAHGVTSAPDT*RPAPGS*T*APPA Pep139	100
MUC5B AT*PSSTPGTTHTP	101
MUC5B ATPSST*PGTTHTP	102
MUC5B ATPSSTPGT*THTP	103
α1-acid glycoprotein	104
Fetuin fetal bovine serum	105
BSA	106
Albumin from human serum	107
ICAM-I human	108
Transferrin human	109
Mucin from porcine stomach	110
Mucin from bovine submaxillary	111



PAHGVT*(T-a)SAPDTRPAPGSTA	14
PAHGVTSAPDT*(T-a)RPAPGSTA	15
PAHGVTSAPDTRPAPGST*(T-a)A	16
PAHGVT*(T-a)SAPDT*(T-a)RPAPGSTA	17
PAHGVT*(T-a)SAPDTRPAPGST*(T-a)A	18
PAHGVTSAPDT*(T-)RPAPGST*(T-a)A	19
PAHGVT*(T-a)SAPDT*(T-a)RPAPGST*(T-a)A	20
PAHGVT*(C1T1)SAPDTRPAPGSTA	21
PAHGVTSAPDT*(C1T1)RPAPGSTA	22
PAHGVTSAPDTRPAPGST*(C1T1)A	23
PAHGVT*(C1T1)SAPDT*(C1T1)RPAPGSTA	24
PAHGVT*(C1T1)SAPDTRPAPGST*(C1T1)A	25
PAHGVTSAPDT*(C1T1)RPAPGST*(C1T1)A	26
PAHGVT*(C1T1)SAPDT*(C1T1)RPAPGST*(C1T	27
PAHGVT*(C1T2)SAPDTRPAPGSTA	28
PAHGVTSAPDT*(C1T2)RPAPGSTA	29
PAHGVTSAPDTRPAPGST*(C1T2)A	30
PAHGVT*(C1T2)SAPDT*(C1T2)RPAPGSTA	31
PAHGVT*(C1T2)SAPDTRPAPGST*(C1T2)A	32
PAHGVTSAPDT*(C1T2)RPAPGST*(C1T2)A	33
PAHGVT*(C1T2)SAPDT*(C1T2)RPAPGST*(C1T	34
PAHGVT*(C2T1)SAPDTRPAPGSTA	35
PAHGVTSAPDT*(C2T1)RPAPGSTA	36
PAHGVTSAPDTRPAPGST*(C2T1)A	37
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGSTA	38
PAHGVT*(C2T1)SAPDTRPAPGST*(C2T1)A	39
PAHGVTSAPDT*(C2T1)RPAPGST*(C2T1)A	40
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGST*(C2T	41
PAHGVT*(C2T1)SAPDTRPAPGSTA	42
PAHGVTSAPDT*(C2T1)RPAPGSTA	43
PAHGVTSAPDTRPAPGST*(C2T1)A	44
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGSTA	45
PAHGVT*(C2T1)SAPDTRPAPGST*(C2T1)A	46
PAHGVTSAPDT*(C2T1)RPAPGST*(C2T1)A	47
PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGST*(C2T	48
PAHGVT*(C2T2)SAPDTRPAPGSTA	49
PAHGVTSAPDT*(C2T2)RPAPGSTA	50
PAHGVTSAPDTRPAPGST*(C2T2)A	51
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGSTA	52
PAHGVT*(C2T2)SAPDTRPAPGST*(C2T2)A	53
PAHGVTSAPDT*(C2T2)RPAPGST*(C2T2)A	54
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGST*(C2T	55
PAHGVT*(C2T2)SAPDTRPAPGSTA	56
PAHGVTSAPDT*(C2T2)RPAPGSTA	57
PAHGVTSAPDTRPAPGST*(C2T2)A	58
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGSTA	59
PAHGVT*(C2T2)SAPDTRPAPGST*(C2T2)A	60
PAHGVTSAPDT*(C2T2)RPAPGST*(C2T2)A	61
PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGST*(C2T	62
PAHGVT*(Tn-a)SAPDTRPAPGSTA	63
PAHGVTSAPDT*(Tn-a)RPAPGSTA	64
PAHGVTSAPDTRPAPGST*(Tn-a)A	65
PAHGVT*(Tn-a)SAPDT*(Tn-a)RPAPGSTA	66
PAHGVT*(Tn-a)SAPDTRPAPGST*(Tn-a)A	67
PAHGVTSAPDT*(Tn-a)RPAPGST*(Tn-a)A	68
PAHGVT*(Tn-a)SAPDT*(Tn-a)RPAPGST*(Tn-a)	69
PAHGVT*(C3T1)SAPDTRPAPGSTA	70
PAHGVTSAPDT*(C3T1)RPAPGSTA	71
PAHGVTSAPDTRPAPGST*(C3T1)A	72
PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGSTA	73
PAHGVT*(C3T1)SAPDTRPAPGST*(C3T1)A	74
PAHGVTSAPDT*(C3T1)RPAPGST*(C3T1)A	75
PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T	76
PAHGVT*(C3T2)SAPDTRPAPGSTA	77
PAHGVTSAPDT*(C3T2)RPAPGSTA	78
PAHGVTSAPDTRPAPGST*(C3T2)A	79
PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA	80
PAHGVT*(C3T2)SAPDTRPAPGST*(C3T2)A	81
PAHGVTSAPDT*(C3T2)RPAPGST*(C3T2)A	82
PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T	83
PAHGVT*(C4T1)SAPDTRPAPGSTA	84
PAHGVTSAPDT*(C4T1)RPAPGSTA	85
PAHGVTSAPDTRPAPGST*(C4T1)A	86
PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA	87
PAHGVT*(C4T1)SAPDTRPAPGST*(C4T1)A	88
PAHGVTSAPDT*(C4T1)RPAPGST*(C4T1)A	89
PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T	90
PAHGVT*SAPDTRPAPGSTA	91
PAHGVTSAPDT*(C4T2)RPAPGSTA	92
PAHGVTSAPDTRPAPGST*(C4T2)A	93
PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA	94
PAHGVT*(C4T2)SAPDTRPAPGST*(C4T2)A	95
PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A	96
PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T	97
PAHGVTSAPDT*RPAPGS*T*APPA Pep137	98
PAHGVTSAPDT*RPAPGS*T*APPA Pep138	99
PAHGVTSAPDT*RPAPGS*T*APPA Pep139	100
MUC5B AT*PSSTPGTTHTP	101
MUC5B ATPSST*PGTTHTP	102
MUC5B ATPSSTPGT*THTP	103
α1-acid glycoprotein	104
Fetuin fetal bovine serum	105
BSA	106
Albumin from human serum	107
ICAM-I human	108
Transferrin human	109
Mucin from porcine stomach	110
Mucin from bovine submaxillary	111



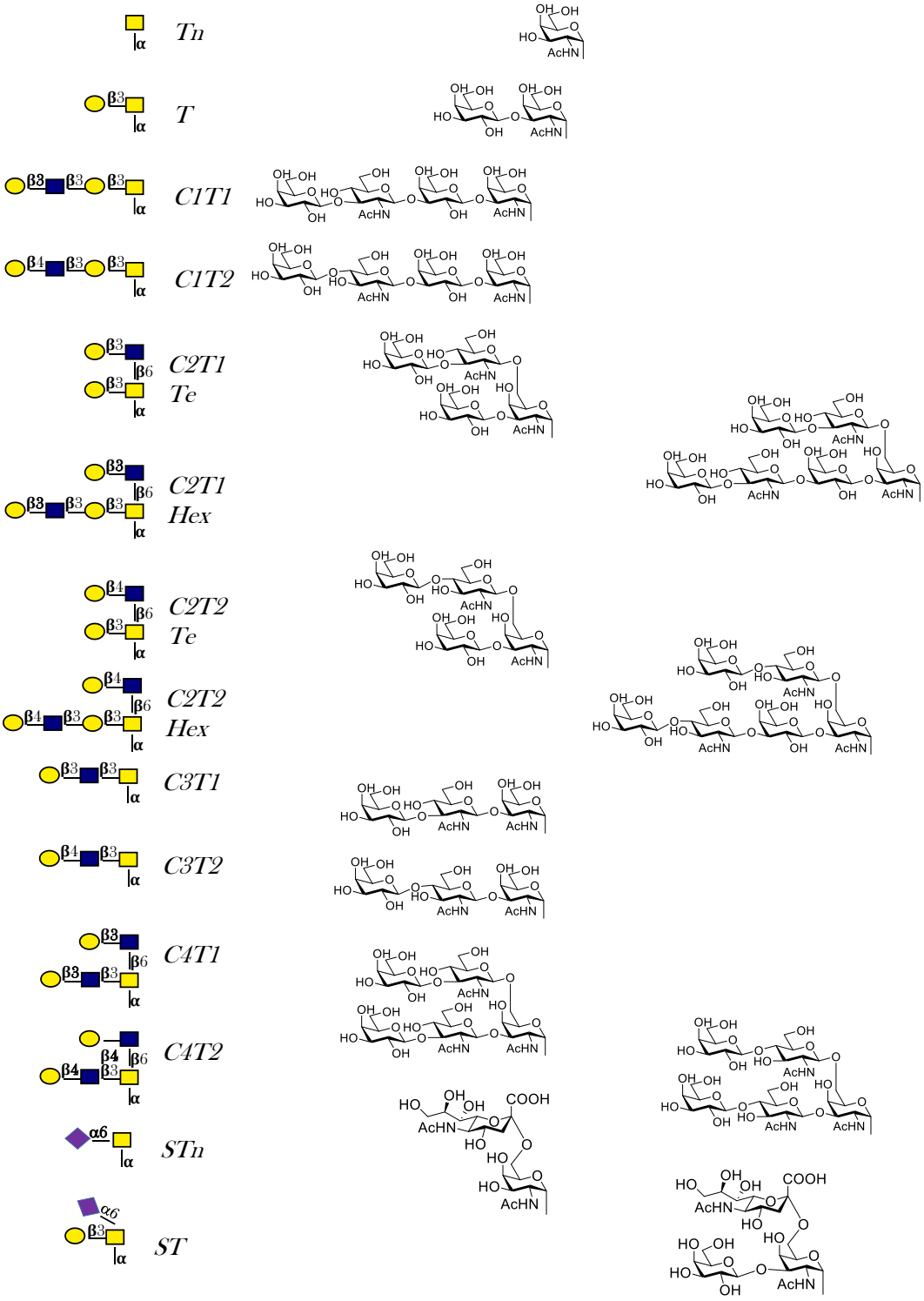


Figure S6. Glycopeptide microarray screening results of anti-sera induced by Qβ-MUC1-β-Tf . The results from five anti-sera (1/100 dilution) were shown.

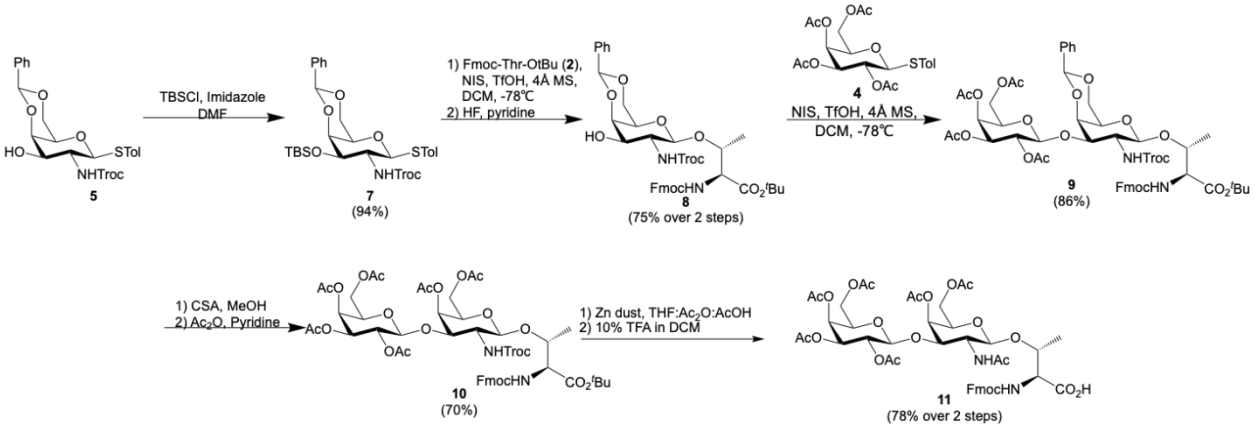
General Experimental Procedures and Methods for Synthesis:

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 flame 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 $(NH_4)_6MO_7O_{24}*4H_2O$ (24.0 g) in 6% H_2SO_4 (500 mL) or 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 Pharmacia0 instrument equipped with a Superose-6 column. 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 to the mixture. LCMS was performed on Waters Xevo G2-XS quadrupole/time-of-flight UPLC/MS/MS. The liquid chromatography was done on ACUITY UPLC® Peptide BEH C18 column, 130Å, 1.7 μm, 2.1 mm x 150 mm, using gradient eluent from 95% 0.1% formic acid in CH_3CN (0.3 mL min^{-1} flowrate) at a column temperature of 40 °C. The spectra were deconvoluted using MaxEnd148a. The average numbers of MUC1/subunit were analyzed by signal intensity in the mass spectra. 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 Reafent (Bradford Assay, Pierce) with BSA as the standard. *O*-Glycosidase, endo-α-*N*-acetylgalactosaminidase from *E. faecalis*, was purchased from New England Biolabs (p0733s).

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.

Synthesis of Fmoc protected β-Tf glyco-AA building block 11:



***p*-Tolyl-4,6-*O*-benzylidene-3-*O*-(*tert*-butyl-dimethylsilyl)-2-deoxy-1-thio-2-*N*-((2,2,2-trichloroethoxy)carbonyl)-ß-D-galactopyranoside (7).**

5 (0.450g, 0.820mmol, 1 equiv.),[2] imidazole (0.112g, 1.64mmol, 2 equiv.) and *tert*-butyldimethylsilyl chloride (0.161g, 1.07mmol, 1.3 equiv.) were dissolved in DMF (4mL). The reaction was stirred, under N_2, overnight. Upon completion of the reaction, the solution was poured into EtOAc and washed 3 times each with 1M HCl, sat. $NaHCO_3$, and sat. NaCl solutions. The organic layers were collected and dried over anhydr. Na_2SO_4 and the excess solvent was removed. The crude product was purified by silica gel flash column chromatography using EtOAc:hexanes:DCM as the eluent. **7** (0.513g) was obtained as a white solid in a 94% yield. ^{1}H NMR (500MHz, $CDCl_3$) δ 7.53 (d, *J* = 7.8Hz, 2H, Ar-H), 7.50-7.45 (m, 2H, Ar-H), 7.40-7.35 (m, 3H, Ar-H), 7.02 (d, *J* = 7.8Hz, 2H, Ar-H), 5.50 (s, 1H, benzylidene CH), 5.10-5.20 (m, 2H, N-H H-1), 4.75 (d, *J* = 12.0Hz, 1H, Troc CH_2), 4.66 (d, *J* = 12.0Hz, 1H, Troc CH_2), 4.38 (dd, *J* = 1.7Hz, 12.4Hz, 1H, H-6a), 4.32 (dd, *J* = 3.3Hz, 10.2Hz, 1H, H-5), 4.07 (d, *J* = 3.3Hz, 1H, H-4), 4.01 (dd, *J* = 1.7Hz, 12.4Hz, 1H, H-6b), 3.58-3.48 (m, 2H, H-2, H-3), 2.31 (s, 3H, STol CH_3), 0.85 (s, 9H, tBu CH_3), 0.06 (s, 3H, silyl CH_3), 0.05 (s, 3H, silyl CH_3). ^{13}C NMR (125MHz, $CDCl_3$) δ 153.5, 138.0, 137.9, 133.5, 129.7, 128.8, 128.2, 128.0, 126.3, 100.7, 95.3, 84.6, 77.3, 76.3, 74.6, 74.5, 70.9, 69.9, 69.5, 68.8, 53.3, 25.7, 21.2, 18.1, -4.5, -4.7. ESI-TOF ($C_{29}H_{38}Cl_3NNaO_6SSi$): calculated ([M+$Na^+$]): 684.1152, found 684.1151.

***O*-(4,6-*O*-Benzylidene-2-2-deoxy-*N*-((2,2,2-trichloroethoxy)carbonyl)-ß-D-galactopyranosyl)-*N*-(fluoren-9-ylmethoxycarbonyl)-threonine *tert*-butyl ester (8).**

7 (0.759g, 1.14 mmol, 1 equiv.) and *N*-(fluoren-9-ylmethoxycarbonyl)-L-threonine *tert*-butyl ester **2** (Fmoc-Thr-O*t*Bu) (0.455g, 1.14mmol, 1 equiv.) were dissolved in dry DCM (5mL). Freshly activated 4Å molecular sieves (MS) were added and the reaction was stirred for 30 min under N_2. The reaction was cooled to -78°C. Sequentially *N*-iodosuccinimide (NIS) (0.515g, 2.29mmol, 2 equiv.) and trifluoromethanesulfonic acid (TfOH) (12.1µL, 0.137mmol, 0.12 equiv.) were added to the reaction. The reaction was allowed to slowly warm to room temperature. Upon reaching room temperature, the reaction was neutralized by addition of *N*,*N*-diisopropylethylamine

(DIPEA). The reaction was filtered through celite and excess solvent was removed. The mixture was run through a silica gel column, using EtOAc:hexanes:DCM as eluent, to obtain a mixture of the glycosylated amino acid and Fmoc-Thr-O*t*Bu. The recovered mixture was dissolved in pyridine (2mL) and cooled to 0°C before HF•Pyridine (1mL) was added. The solution was stirred and allowed to warm to room temperature. After 2h, the reaction was neutralized using a sat. $NaHCO_3$ solution and dissolved in EtOAc. The solution was washed 3 times each with sat. copper sulfate, 1M HCl, sat. $NaHCO_3$, and sat. NaCl solutions. The organic layers were collected, dried over anhydr. Na_2SO_4 and excess solvent was removed. The crude product was purified by silica gel flash column chromatography using EtOAc:hexanes:DCM as the eluent. **8** (0.704g) was obtained as a white solid in a 75% yield over 2 steps. ^{1}H NMR (500MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.6Hz, 2H, Ar-H), 7.66 (t, *J* = 7.3Hz, 2H, Ar-H), 7.48-7.45 (m, 2H, Ar-H), 7.41-7.34 (m, 5H, Ar-H), 7.32-7.27 (m, 2H, Ar-H), 5.86 (d, *J* = 9.4Hz, 1H, N-H), 5.55 (s, 1H, benzylidene CH), 5.32 (d, *J* = 7.8Hz, 1H, N-H), 4.72 (s, 2H, Troc CH_2), 4.68 (d, *J* = 7.7Hz, 1H, H-1), 4.55 (m, 1H, Thr CH), 4.42 (m, 1H, Fmoc CH_2), 4.35-4.22 (m, 4H, Fmoc CH, Fmoc CH_2, Thr C_αH, H-6a), 4.20 (bs, 1H, H-5), 4.10-4.00 (m, 2H, H-3, H-6b), 3.52 (bs, 1H, H-2), 3.46 (s, 1H, H-5), 2.71 (d, *J* = 10.5Hz, 1H, OH), 1.49 (s, 9H, tBu CH_3), 1.23 (d, *J* = 6.3Hz, 3H, Thr CH_3). ^{13}C NMR (125MHz, $CDCl_3$) δ 181.9, 157.0, 154.5, 144.0, 143.9, 141.2, 141.1, 137.4, 129.3, 128.2, 127.7, 127.1, 127.1, 126.5, 125.4, 125.4, 119.9, 119.8, 101.4, 96.9, 95.4, 82.4, 74.9, 74.5, 72.7, 69.4 69.0, 67.4, 66.5, 59.0, 55.9, 47.1, 28.0, 15.9. ESI-TOF ($C_{39}H_{43}Cl_3N_2NaO_{11}$): calculated ([M+$Na^+$]): 843.1830, found 843.1798.

_O_-(4,6-_O_-Benzylidene-2-deoxy-_O_-(2,3,4,6-tetra-_O_-acetyl-ß-D-galactopyranosyl)-2-_N_-((2,2,2-trichloroethoxy)carbonyl)-ß-D-galactopyranosyl)-_N_-(Fluoren-9-ylmethoxycarbonyl)-L-threonine *tert*-butyl ester (9).

8 (0.283g, 0.344mmol, 1 equiv.) and **4** (0.313g, 0.688mmol, 2 equiv.) were dissolved in dry DCM (3mL). Freshly activated 4Å MS were added and the reaction stirred for 30min under N_2. The reaction was cooled to -78°C. Sequentially, NIS (0.310g, 1.38mmol, 4 equiv.) and TfOH (6.20µL, 0.0413mmol, 0.12 equiv.) were added to the reaction. The reaction was allowed to slowly warm to room temperature. Upon reaching room temperature, the reaction was neutralized by addition of DIPEA, filtered through celite and excess solvent was removed The crude product was purified by silica gel flash column chromatography using EtOAc:hexanes:DCM as eluent. **9** (0.340g) was obtained as a white solid in an 86% yield. ^{1}H NMR (500MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.6Hz, 2H), 7.65 (t, *J* = 8.8Hz, 2H, Ar-H), 7.47 (m, 2H, Ar-H), 7.38 (t, *J* = 7.5Hz, 2H, Ar-H), 7.35-7.25 (m, 5H, Ar-H), 5.98 (d, *J* = 6.1Hz, 1H, N-H), 5.92 (d, *J* = 9.4Hz, 1H, N-H), 5.55 (s, 1H, benzylidene CH), 5.36 (d, *J* = 2.8Hz, 1H, H-4'), 5.14 (dd, *J* = 8.1Hz, 10.1Hz, 1H, H-2'), 5.00 (d, *J* = 7.8Hz, 1H, H-1), 4.95 (dd, *J* = 3.3Hz, 10.3Hz, 1H, H-3'), 4.83 (d, *J* = 12.1Hz, 1H, Troc CH_2), 4.72 (d, *J* = 8.0Hz, 1H, H-1'), 4.60 (d, *J* = 12.1Hz, 1H, Troc CH_2), 4.58-4.50 (m, 2H, H-4, Thr CH), 4.41-4.26 (m, 5H, Fmoc CH_2, Thr C_αH, H-2, H-6a/b), 4.23 (t, *J* = 7.6Hz, 1H, Fmoc CH), 4.18 (dd, *J* = 6.5Hz, 11.2Hz, 1H, H-6a'), 4.13 (dd, *J* = 6.5Hz, 11.2Hz, 1H, H-6b'), 4.05 (d, *J* = 12.1Hz, 1H, Fmoc CH_2), 3.89 (t, *J* = 6.3Hz, 1H, H-5'), 3.47 (m, 2H, H-3, H-5), 2.16 (s, 3H, Ac CH_3), 2.06 (s, 3H, Ac CH_3), 2.04 (s, 3H, Ac CH_3), 1.96 (s, 3H, Ac CH_3), 1.43 (s, 9H, tBu CH_3), 1.20 (d, *J* =

6.3Hz, 3H, Thr CH_3). ^{13}C NMR (125MHz, $CDCl_3$) δ 182.0 170.4, 170.2, 169.5, 169.5, 157.0, 154.1, 143.9, 143.9, 141.2, 141.2, 137.7, 129.0, 128.0, 127.6, 127.1, 126.5, 125.4, 125.3, 119.9, 119.9, 102.0, 100.9, 95.4, 95.2, 82.6, 75.9, 75.9, 74.3, 72.4, 70.8, 70.8, 69.0, 68.7, 67.3, 66.9, 66.3, 61.5, 59.0, 53.6, 47.1, 27.9, 20.8, 20.7, 20.6, 20.5, 15.8. ESI-TOF ($C_{53}H_{61}Cl_3N_2NaO_{20}$): calculated ([M+$Na^+$]): 1173.2781, found 1173.2784.

_O_-(4,6-Di-_O_-acetyl-2-deoxy-_O_-(2,3,4,6-_O_-tetra-_O_-acetyl-ß-D-galactopyranosyl)-2-_N_-((2,2,2-trichloroethoxy)carbonyl)-ß-D-galactopyranosyl)-_N_-(Fluoren-9-ylmethoxycarbonyl)-L-threonine _tert_-butyl ester (10).

9 (229mg, 0.199mmol, 1 equiv.) and *p*-toluenesulfonic acid (3.4mg, 0.0199mmol, 0.1 equiv.) were dissolved in MeOH (10mL). The reaction was stirred until complete as determined by TLC. The solvent removed by co-evaporation with toluene. The crude intermediate was then dissolved 3ml of pyridine. Excess Ac_2O was added and the reaction was stirred under N_2. The reaction was dissolved in EtOAc and washed 3 times each with sat. copper sulfate, 1M HCl, sat. $NaHCO_3$, and sat. NaCl solutions. The organic layers were collected, dried over anhydr. Na_2SO_4 and excess solvent was removed. The crude product was purified by silica gel flash column chromatography using EtOAc:hexanes:DCM as eluent. **10** (0.160g) was obtained as a white solid in a 70% yield. ^1H NMR (500MHz, $CDCl_3$) δ 7.71 (d, *J* = 7.6Hz, 2H, Ar-H), 7.61 (dd, *J* = 3.5Hz, 7.3Hz, 2H, Ar-H), 7.35 (t, *J* = 7.3Hz, 2H, Ar-H), 7.30-7.23 (m, 2H, Ar-H), 5.73-5.65 (m, 2H, N-H), 5.35 (d, *J* = 3.2Hz, 1H, H-4), 5.30 (d, *J* = 3.2Hz, 1H, H-4') 5.08 (dd, *J* = 8.2Hz, 10.3Hz, 1H, H-2'), 4.91 (dd, *J* = 3.3Hz, 10.3Hz, 1H, H-3'), 4.77-4.67 (m, 2H, H-1, Troc CH_2), 4.64 (d, *J* = 12.1Hz, 1H, Troc CH_2), 4.58 (d, *J* = 8.1Hz, 1H, H-1'), 4.45-4.35 (m, 2H, Thr CH, Fmoc CH_2), 4.30-4.17 (m, 4H, Fmoc CH, Fmoc CH_2, Thr $C_\alpha H$, H-3), 4.15-4.03 (m, 3H, H-6a/b', H-6a), 3.91 (dd, *J* = 7.1Hz, 11.6Hz, 1H, H-6b), 3.84 (t, *J* = 6.5Hz, 1H, H-5'), 3.77 (t, *J* = 5.6Hz, 1H, H-5), 3.42-3.33 (m, 1H, H-2), 1.04-1.98 (m, 15H, Ac CH_3), 1.92 (s, 3H, Ac CH_3), 1.43 (s, 9H, tBu CH_3), 1.14 (d, *J* = 6.3Hz, 3H, Thr CH_3). ESI-TOF ($C_{50}H_{61}Cl_3N_2NaO_{22}$): calculated ([M+$Na^+$]): 1169.2679, found 1169.2675.

_O_-(2-Acetamido-4,6-di-_O_-acetyl-2-deoxy-_O_-(2,3,4,6-tetra-_O_-acetyl-ß-D-galactopyranosyl)-ß-D-galactopyranosyl)-_N_-(Fluoren-9-ylmethoxycarbonyl)-L-threonine (11):

10 (160mg, 0.139mmol, 1 equiv.) and zinc dust (18.2mg, 0.755mmol, 2 equiv.) was dissolved in a 3:2:1 mixture of tetrahydrofuran (THF):Ac_2O:acetic acid (AcOH) (2mL). The reaction was stirred under N_2, until complete as measured by TLC After filtering through celite, the solvent was removed and the crude product was purified by silica gel flash column chromatography, using MeOH:DCM as eluent. The resulting solid was dissolved in a 10% solution of trifluoroacetic acid (TFA) in DCM. The reaction was stirred for an hour. The solvent was removed and the crude product was purified by silica gel flash column chromatography, using MeOH:DCM as the eluent. **11** (124mg) was obtained as a yellow glassy solid in 78% yield. ^1H NMR (500 MHz, Methanol-d_4) δ 7.79 (d, *J* = 7.5 Hz, 2H, Ar-H), 7.67 (dd, *J* = 7.6, 5.1 Hz, 2H, Ar-H), 7.39 (t, *J* = 7.4 Hz, 2H, Ar-H), 7.32 (tt, *J* = 7.5, 1.4 Hz, 2H, Ar-H), 5.38 (d, J = 3.1 Hz, 1H, H-4), 5.36 (d, *J* = 3.2 Hz, 1H,

H-4'), 5.09 (dd, *J* = 10.5, 3.5 Hz, 1H, H-3'), 5.01 (dd, *J* = 10.5, 7.7 Hz, 1H, H-2'), 4.90 (s, 2H, N-H), 4.76 (d, *J* = 7.8 Hz, 1H, H-1'), 4.52 (d, *J* = 8.2 Hz, 1H, H-1), 4.47 – 4.39 (m, 1H, Fmoc CH_2), 4.38 – 4.31 (m, 2H, Fmoc CH_2, Thr CH), 4.22 (t, *J* = 6.7 Hz, 1H, Fmoc CH), 4.16 – 3.94 (m, 8H, Thr $C_{\alpha}H$, H-2, H-5, H-5', H-6a/b, H-6a/b'), 3.94 – 3.87 (m, 1H, H-3), 2.13 (s, 3H, Ac CH_3), 2.05 (s, 3H, Ac CH_3), 2.05 (s, 3H, Ac CH_3), 2.03 (s, 3H, Ac CH_3), 2.03 (s, 3H, Ac CH_3), 1.96 (s, 3H, Ac CH_3), 1.94 (s, 3H, Ac CH_3), 1.18 (d, *J* = 6.3Hz, 3H, Thr CH_3). ^{13}C NMR (126 MHz, Methanol-d_4) δ 172.4, 171.1, 170.6, 170.6, 170.4, 170.0, 169.7, 157.3, 144.0, 143.7, 141.2, 141.1, 127.4, 126.8, 126.8, 124.8, 119.57, 101.1, 99.0, 76.2, 74.8, 70.9, 70.8, 70.4, 69.0, 68.8, 67.2 66.5, 62.0, 60.9, 54.2, 51.9, 47.0, 22.0 19.5, 19.5 19.4, 19.2, 19.1, 19.1, 15.9. ESI-Quadrupole ($C_{45}H_{53}N_2O_{21}$): calculated ([M-H$^+$]): 957.3141, found 957.3128.

Synthesis of MUC1-ß-Tf (12):

The MUC1-ß-Tf glycopeptide **12** was synthesized using *p*-nitrophenyl carbonate Wang resin. The *N*-Fmoc-1,4-diaminobutane linker was installed by dissolving *N*-Fmoc-1,4-diaminobutane*HCl (5 equiv.) and DIPEA (10 equiv.) in DMF before swelling the resin in the solution. The *N*-terminal Fmoc was deprotected using 20% piperidine in DMF. The amino acid coupling was carried out with Fmoc-amino acids (5 equiv.) using (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU)/hydroxybenzotriazole (HOBt) (4.9 equiv.) and DIPEA (10 equiv.). Coupling of **11** (2 equiv.) was performed using 1-[bis(dimethylamino)methylene-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxide hexafluorophosphate (HATU)/1-hydroxy-7-azabenzotriazole (HOAt) (1.9 equiv.) and DIPEA (4 equiv.). After assembly of the MUC1-ßTf glycopeptide, the *N*-terminal Fmoc group was removed and the resulting free amine was capped with a 1:1:8 solution of Ac_2O:DIPEA:DMF. The amino acid side chains were deprotected and the glycopeptide was cleaved from the resin using a solution of 18:1:1 TFA:triisopropylsilane (TIPS):H_2O. After 4hrs, the glycopeptide was precipitated by diethyl ether (EtO_2) and pelleted by centrifugation. The crude glycopeptide was purified by HPLC using a Shimadzu HPLC (LC-8A Liquid Chromatograph Pump, DGU-14A Degasser, and SPD-10A UV-Vis Detector), using a reverse phase SUPERCOSIL LC18, 25 cm × 10 mm 5 µm with an acetonitrile (ACN):H_2O (0.1% TFA) gradient. The gradient was 5% ACN for 5 min, 5-60% ACN in 5-45 min, 60-100% ACN in 45-50 min, 100% ACN for 50-55min, and 100-5% ACN in 55-60min. The flow rate was 5 ml/min. To remove the *O*-acetyl groups, the purified glycopeptide was treated with 5% (v/v) hydrazine in water overnight before being purified using the same HPLC gradient as above. For conjugation of MUC1-ß-Tf glycopeptide onto Qß, the purified glycopeptide was treated with adipate bis(4-nitrophenyl) ester (5 equiv.) in the presence of DIPEA (10 equiv.) in DMF and 1% H_2O for 1.5 hrs before being purified by HPLC to give a yield of 30-40%. The MUC1-β-Tf glycopeptide **12** was characterized by ESI-TOF MS and ^{1}H NMR before *p*-nitrophenyl ester functionalization. ^{1}H NMR (500MHz, deuterium oxide) δ 4.47 (t, *J* = 6.9Hz, 1H), 4.45-4.38 (m, 2H), 4.36 (d, *J* = 6.9Hz, 1H), 4.31 (d, *J* = 8.6Hz, 1H, anomeric proton 1), 4.28 (d, *J* = 3.3Hz, 1H), 4.24 (d, *J* = 8.1Hz, 1H, anomeric proton 2), 4.22-4.18 (m, 2H), 3.78 (t, *J* = 9.4Hz, 1H), 3.70 (d, *J* = 3.4Hz, 1H), 3.66-3.49 (m, 9H), 3.48-3.34 (m, 6H), 3.30 (dd, *J* = 7.8Hz, 9.7Hz, 1H), 3.06-2.96 (m, 4H), 2.79 (t, *J* = 7.4Hz, 2H), 2.65 (dd, *J* = 7.1Hz, 16.5Hz, 1H), 2.54 (dd, *J* =

7.1Hz, 16.5Hz, 1H), 2.12-2.01 (m, 3H), 1.85 (s, 3H, NHAc CH_3), 1.83 (s, 3H, NHAc CH_3), 1.82-1.75 (m, 4H), 1.75-1.60 (m, 5H), 1.60-1.33 (m, 9H), 1.16 (d, *J* = 2.3Hz, 3H, Ala CH_3), 1.15 (d, *J* = 2.3Hz, 3H, Ala CH_3), 1.09 (d, *J* = 10.4Hz, 2H), 0.93 (d, *J* = 6.4Hz, 3H, Thr CH_3); ESI-TOF ($C_{58}H_{97}N_{15}O_{24}$): calculated ([M+2H$^+$]/2): 694.8494, found 694.8515

Glycosidase cleavage experiments:

100 µg of glycopeptide, MUC1-α-Tf **12*α*** or MUC1-β-Tf **12**, 5 µl of 10X glycoBuffer, 10 µl endo-*α*-*N*-acetylgalactosaminidase from *Enterococcus faecalis* (400,000 units, New England Biolabs, P0733S) and H_2O was added to make a 50 µl total reaction volume. The reaction mixture was gently mixed and incubated at 37 °C. At indicated time points, the reaction was quenched by adding 50 ul MeOH. The reaction mixture was centrifuged to remove large proteins and reaction sample was analyzed by Xevo QTOF LC-MS/MS. The intensity of glycopeptide at each time point was compared to reaction sample at the beginning of the reaction. After only 20 minutes, a significant amount of the cleavage peptide product for the *α*-isomer **12*α*** was observed with nearly complete cleavage after 1.75 hours (**Figure S1**). In contrast, the β-anomer **12** showed no detectable cleavage even after 24 hours.

To further test the glycopeptide stability, MUC1-β-Tf glycopeptide **12** and MUC1-α-Tf glycopeptide **12*α*** with a β-*N*-acetylhexosaminidase (New England Biolabs, P0721S). 50 µg of glycopeptide, MUC1-α-Tf **12*α*** or MUC1-β-Tf **12**, 2.5 µl of 10X glycoBuffer, 10 µl β-*N*-acetylhexosaminidase (50 units) and H_2O were added to make a 25 µl total reaction volume. The reaction mixture was incubated at 37 °C for 24h and the reaction progress was determined by LC-MS/MS. Both glycopeptides were found stable at 37 °C for 24 hours. The unnatural β-Tf linkage to the peptide may have contributed to its resistance to hydrolysis by the enzyme. As a positive control for the enzyme activity, 50 µg of 4-nitrophenyl *N*-acetyl-β-D-glucosaminide was completely cleaved in 24 hours under the same reaction condition.

Synthesis of Qβ-MUC1 Conjugates.

For Qβ-MUC1-α-Tf: The conjugation was performed as previously reported.[1] For Qβ-MUC1-β-Tf: A solution of Qβ (1 mg, 0.07 µmol subunit, 0.49 µmol reactive amine) in 0.1 M KPB pH 7.0 (0.2 mL) was cooled in an ice bath. The solution was added to a frozen solution of dinitrophenyl linker functionalized **12** (28 µL from a 50 mM stock solution in DMSO, 1.4 µmol). The mixture was allowed to warm to room temperature while gently inverting to ensure mixing. Once the compounds 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 with 0.1 M KPB pH 7.0. The isolated fractions were subjected to centrifugal filtration (MWCO 100kDa) to concentrate the sample to 2-3 mg mL^{-1} (concentration should not be <5 mg mL^{-1} to prevent aggregation). Protein concentration was determined by the Bradford assay against BSA standards. The extent of particle modification was determined by MALDI-TOF MS and by electrophoretic

analysis. Percent protein recovery was ~65-75% with the particle integrity determined by FPLC and DLS analysis.

Synthesis of BSA-MUC1 conjugates.

For synthesis of BSA-MUC1 conjugates, nitrophenyl linker functionalized MUC1 **12** (4 mg) or nitrophenyl linker functionalized **12a** was dissolved in 30 μL of DMSO, then added to a solution of BSA (10 mg mL^{-1}) in 0.1M KPB pH 7.0 (200 μL). The reaction was gently inverted several times and incubated at 37 °C for 16 h. The product was purified by centrifugal filtration (10kDa MWCO). Protein concentration was determined by the Bradford assay against BSA standards. The extent of modification was determined by MALDI-TOF MS.

Immunization of MUC1.Tg mice.

All animal experiments were approved by and performed in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC) of Michigan State University. The animal usage protocol number is PROTO201900423. MUC1.Tg mice were generated by breeding C57BL/6 wild-type female mice and MUC1.Tg male mice with a 10.6 kb genomic Sac II fragment of the human MUC1 gene and maintained in the University Laboratory Animal Resources facility of Michigan State University. MUC1.Tg female mice aged 6−10 weeks were used for studies.

In all studies, MUC1.Tg mice were subcutaneously injected under the scruff on day 0 with 0.2 mL of various Qβ-MUC1 vaccines in PBS containing MPLA (20 μL, 1 mg mL^{-1} in DMSO) for each mouse. Boosters were given subcutaneously at the same amounts of vaccines with MPLA under the scruff on days 14 and 28. All Qβ-MUC1 conjugates administered have the same amounts of MUC1 (8.6nmol). Serum samples were collected on days 0 (before immunization) and 35. The final bleeding was performed through cardiac bleed.

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 (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°C overnight. The coated plates were washed with PBS/0.5% Tween-20 (PBST) (4 × 200 μL) and blocked with 1 % BSA in PBS (100μL/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, 2 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.

Detection of antibody binding to tumor cells by FACS.

B16-MUC1 cells were 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, resuspended in FACS buffer and analyzed by LSR II (BD Biosciences). Data was processed by FlowJo software.

Complement dependent cytotoxicity:

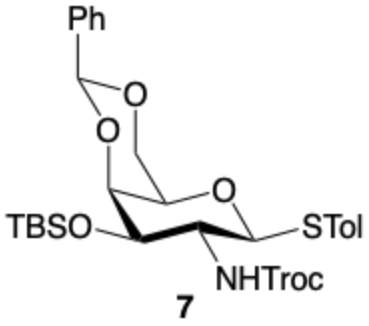
Complement dependent cytotoxicity of B16-MUC1 cells was determined by MTS assay. B16-MUC1 (7000 cells/well) were cultured in 96 well plate with DMEM (10% FBS, G-418 and 1%P.S.) for 12 h. The culture medium was carefully removed. A dilution of mouse sera (1/40) from different groups of immunized MUC1.Tg 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°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°C for 3 h. MTS (CellTiter 96® AQueous One Solution Cell Proliferation Assay; Promega, 20 µL) was added into each well and further incubated at 37°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

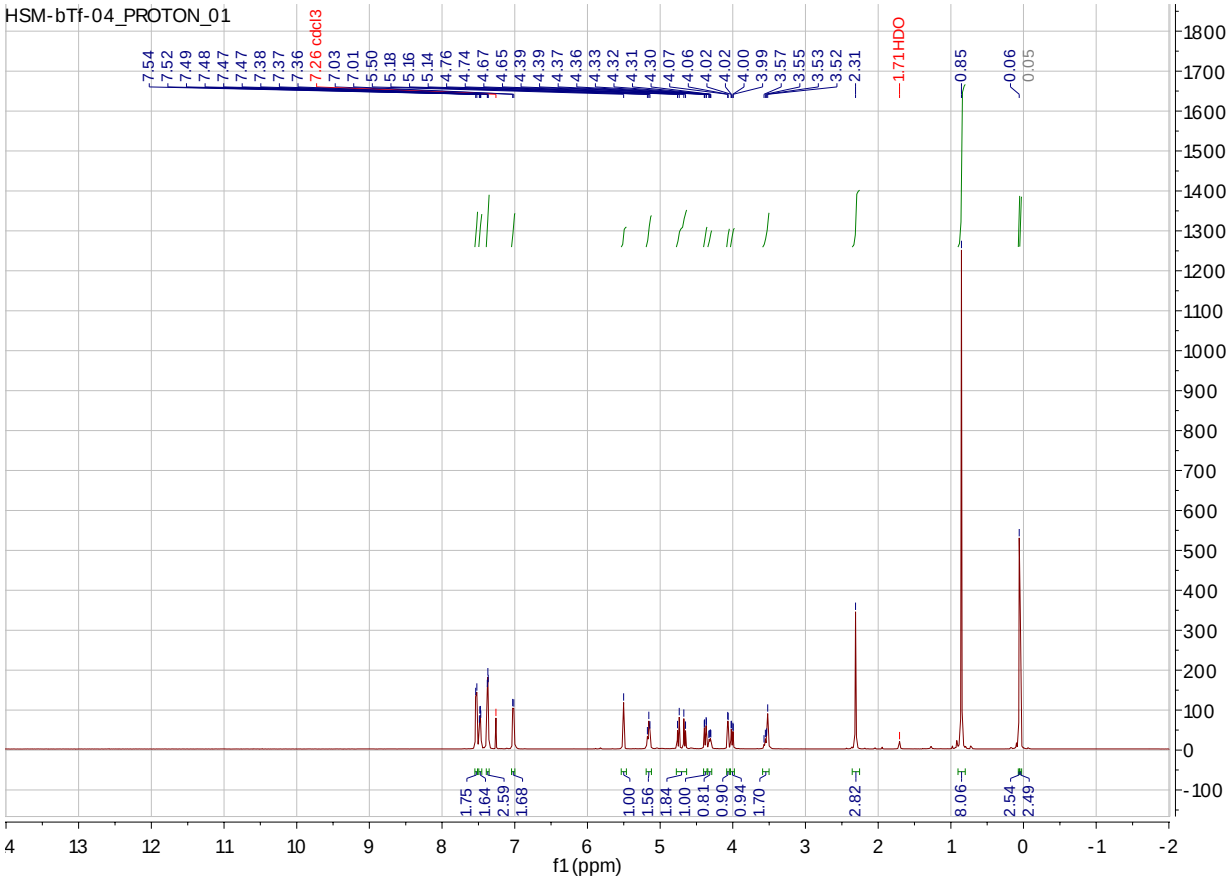
replicates with 450 μm pitch (spot to spot distance) in an array format containing 8 wells with 2 blocks of each 10 × 13 spots. During the spotting process the humidity was kept between 50–60%. The glycopeptides were immobilized on the microarray slides in a humidity chamber (85–95% humidity) by incubation 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 All Building Blocks and Glycopeptides:

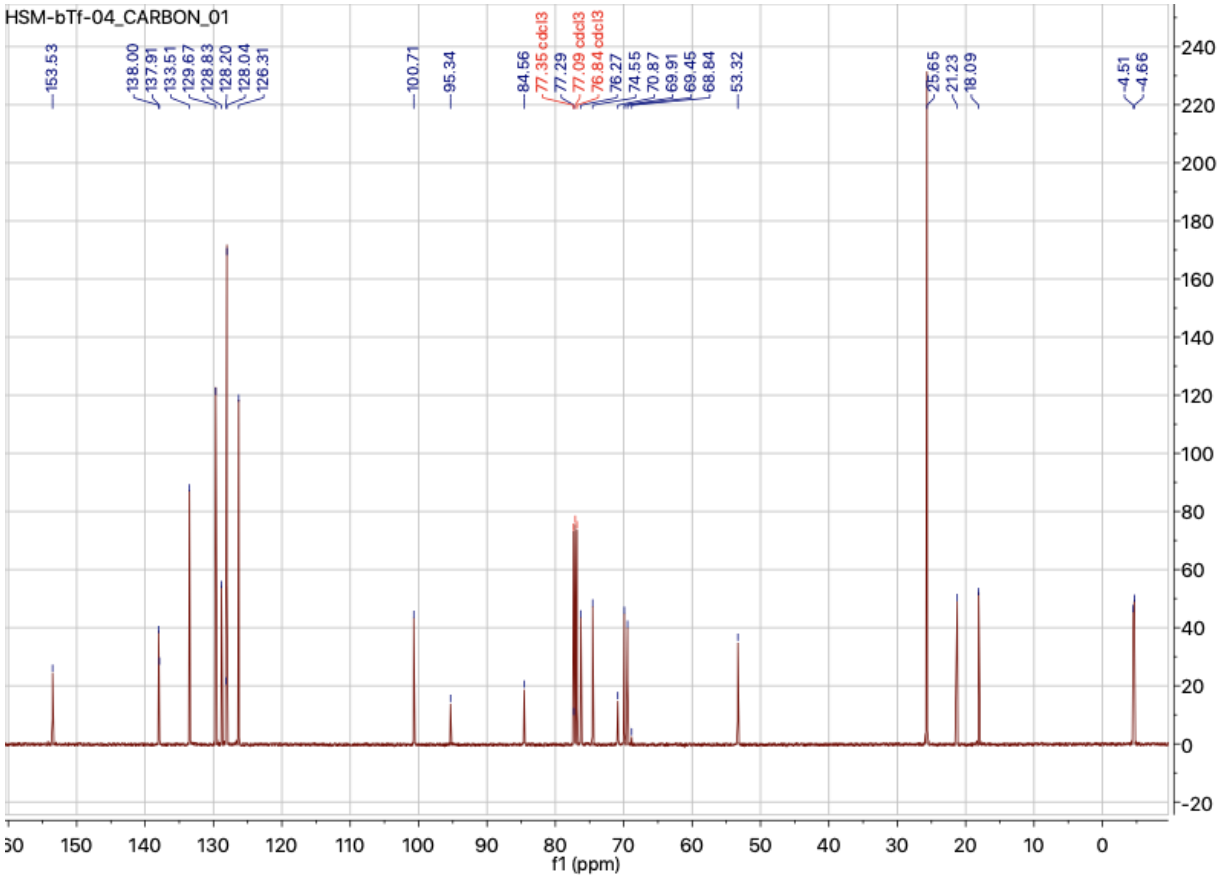


7

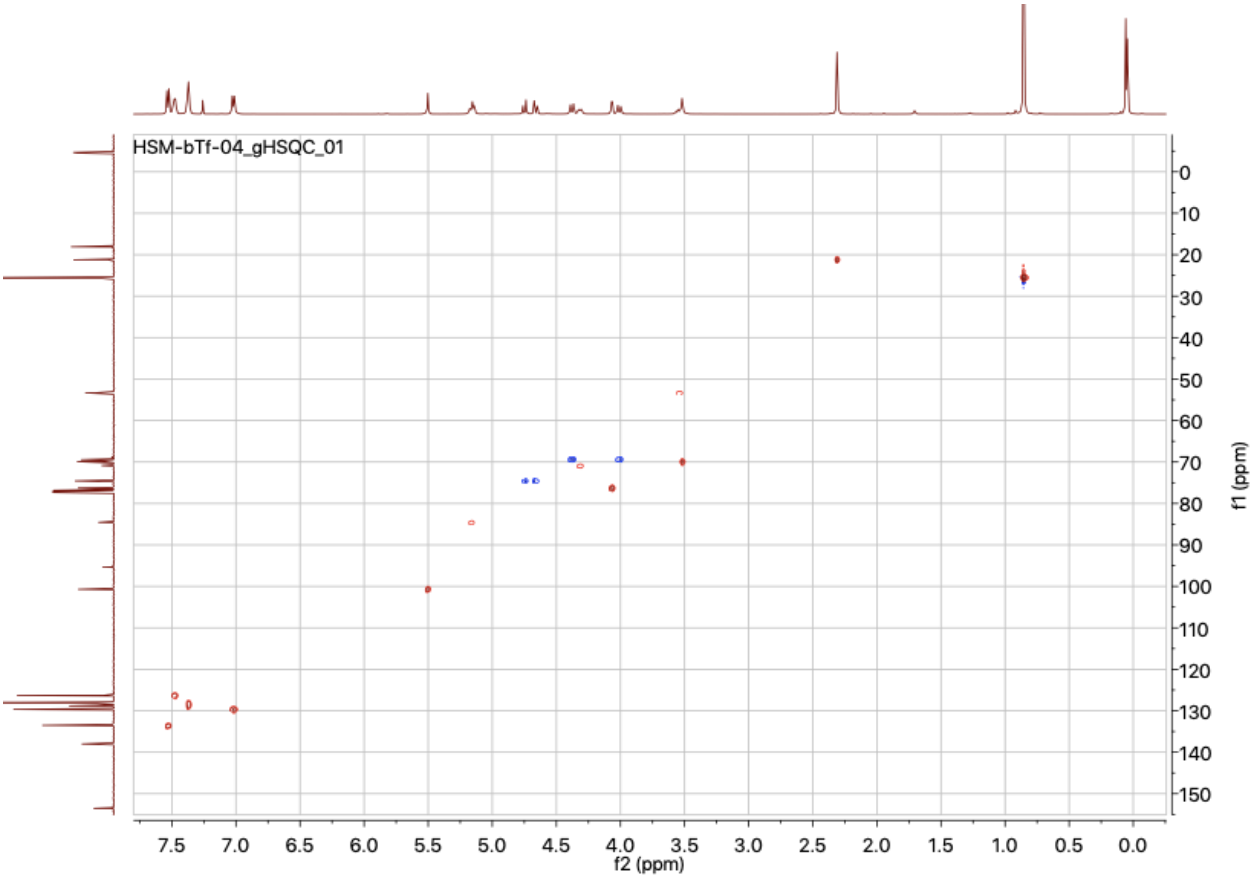
^{1}H NMR (500 MHz, $CDCl_3$) of **7**



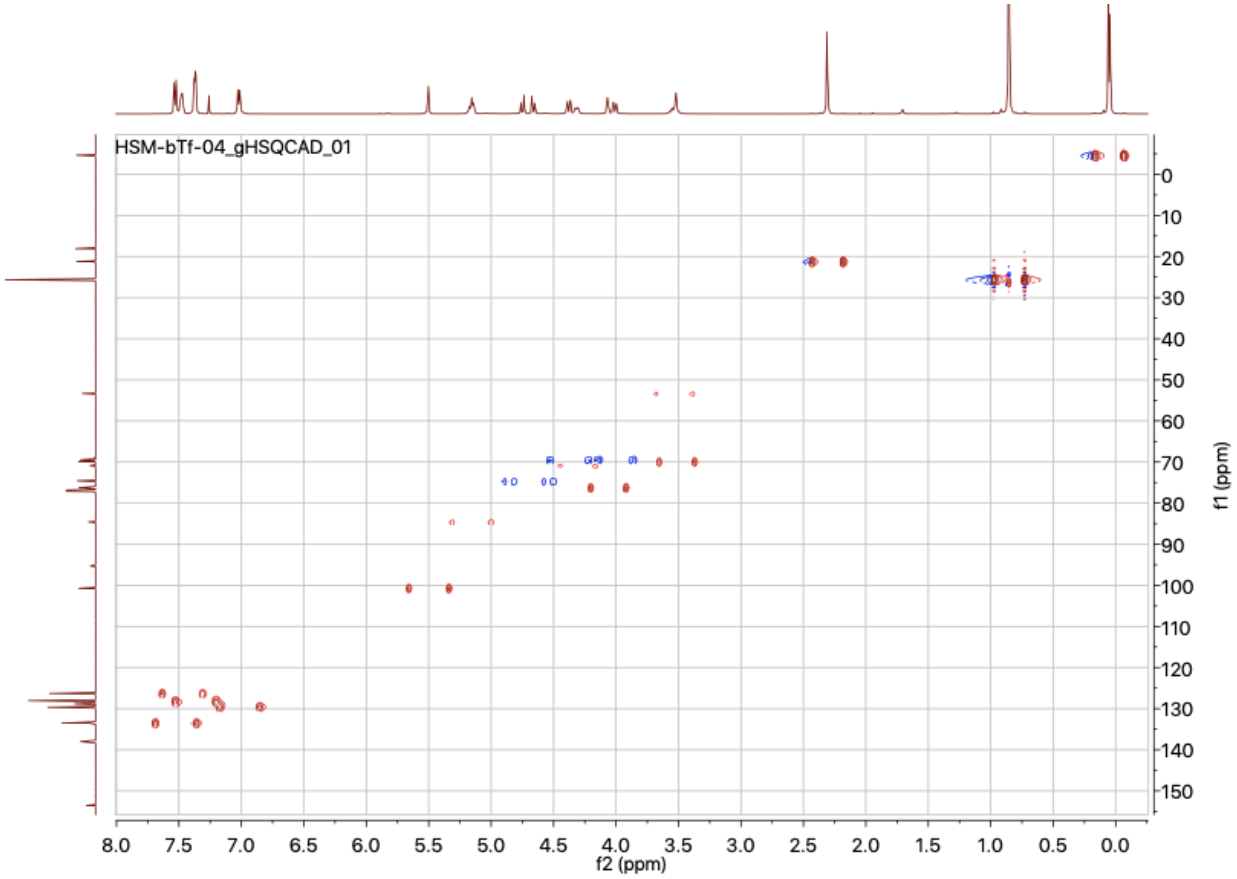
^{13}C NMR (126 MHz, $CDCl_3$) of **7**



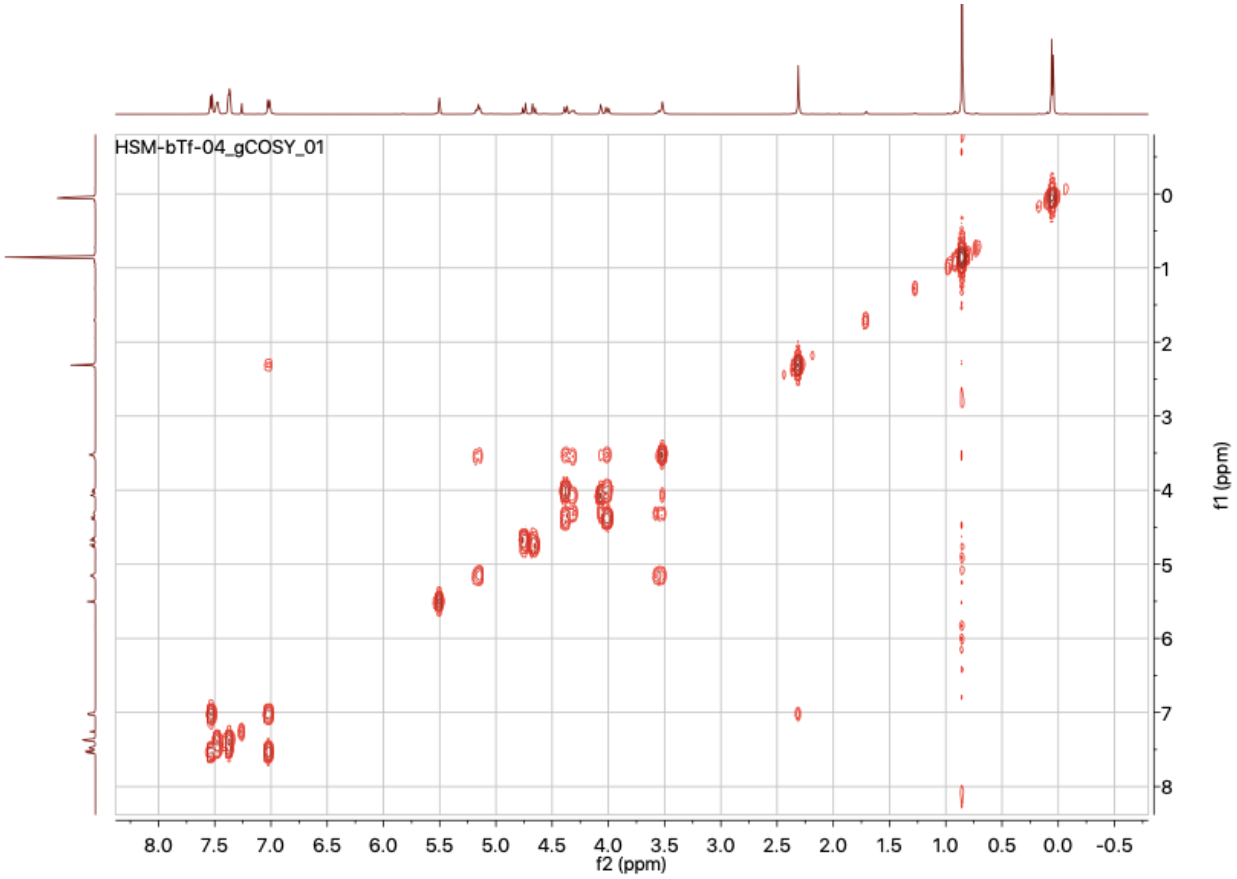
^{1}H-^{13}C HSQC (500 MHz, $CDCl_3$) of **7**



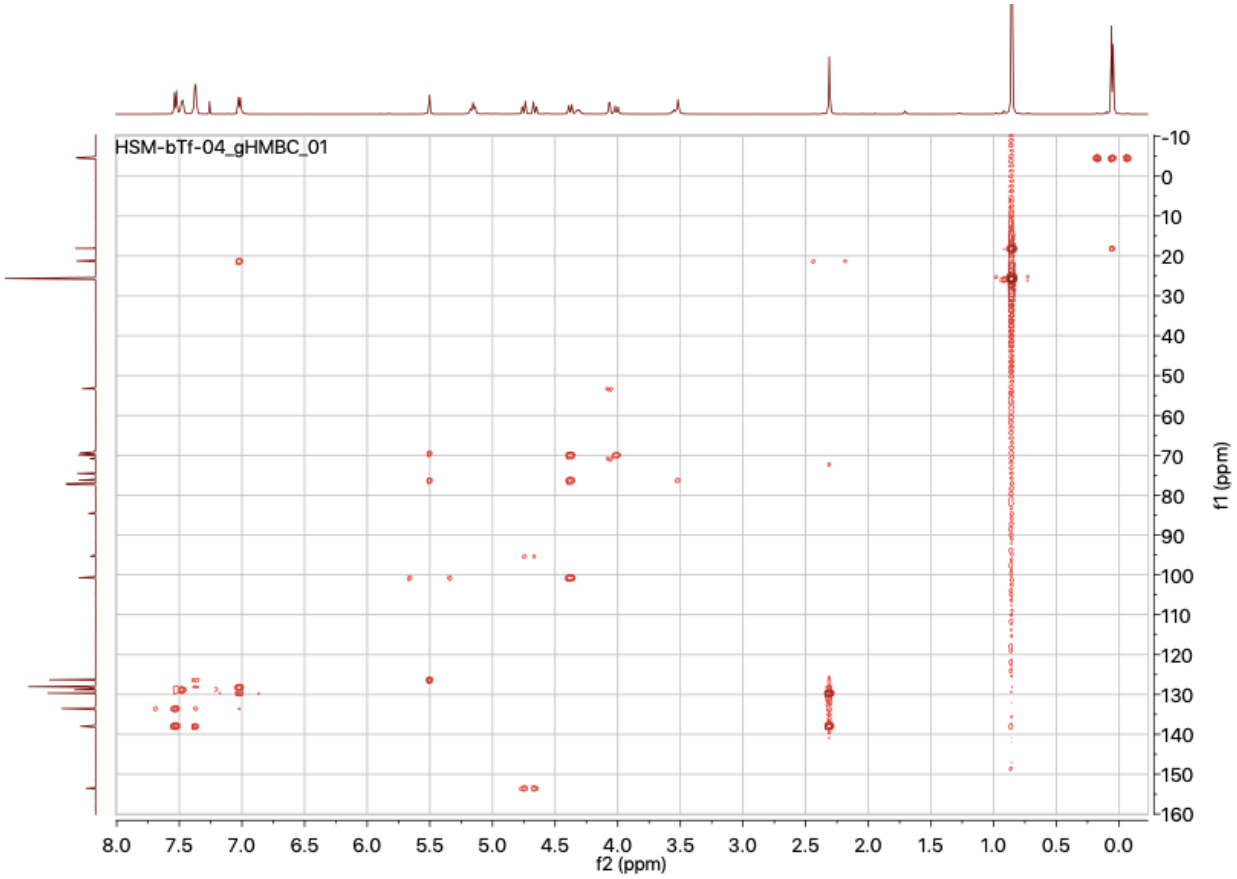
^{1}H-^{13}C HSQCAD (500 MHz, $CDCl_3$) of **7**

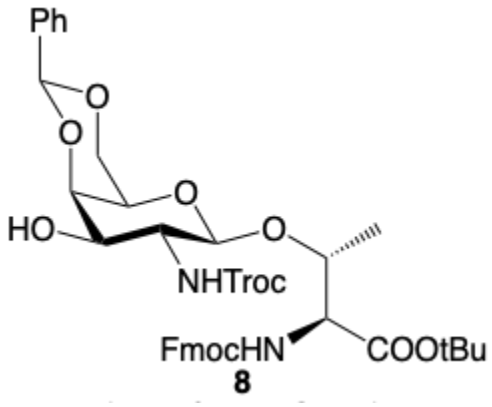


^{1}H-^{1}H COSY (500 MHz, $CDCl_3$) of **7**



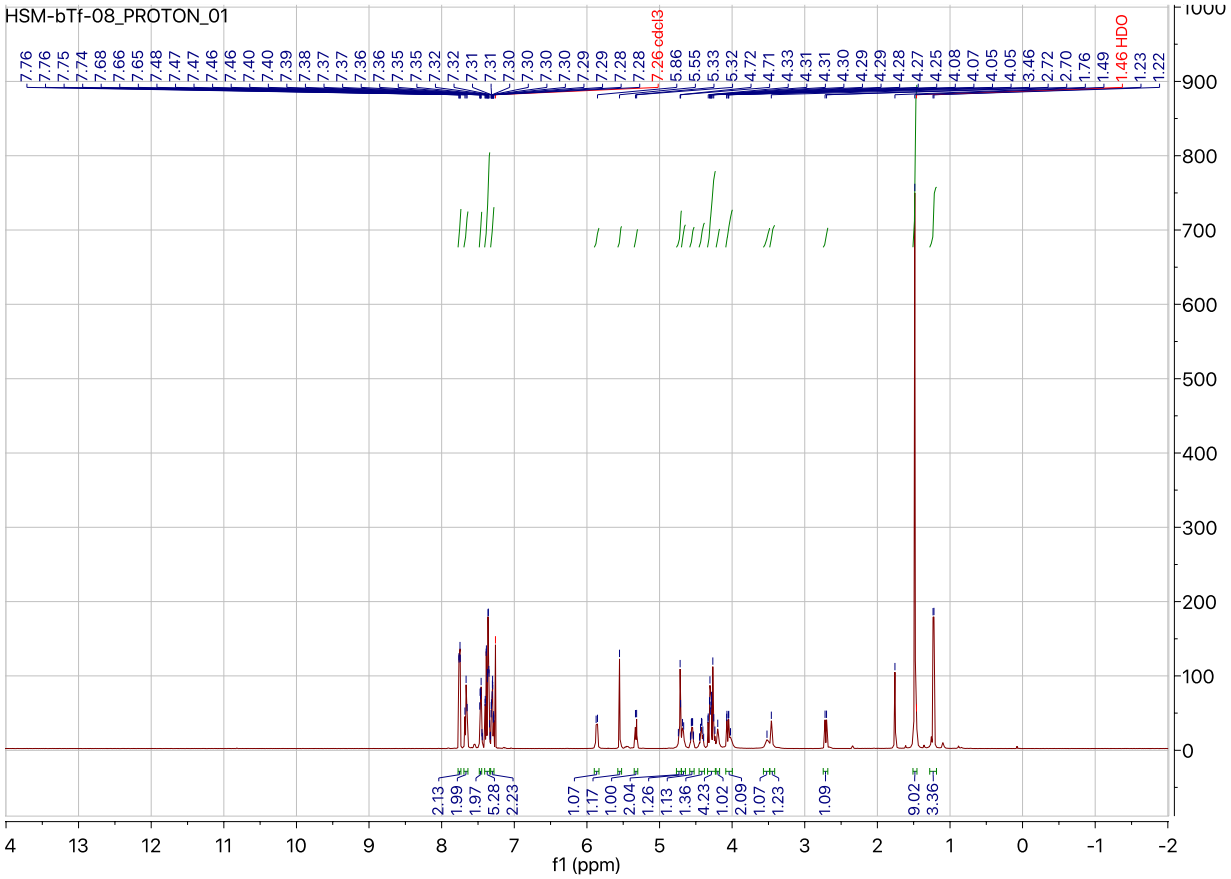
^{1}H-^{13}C HMBC (500 MHz, $CDCl_3$) of **7**



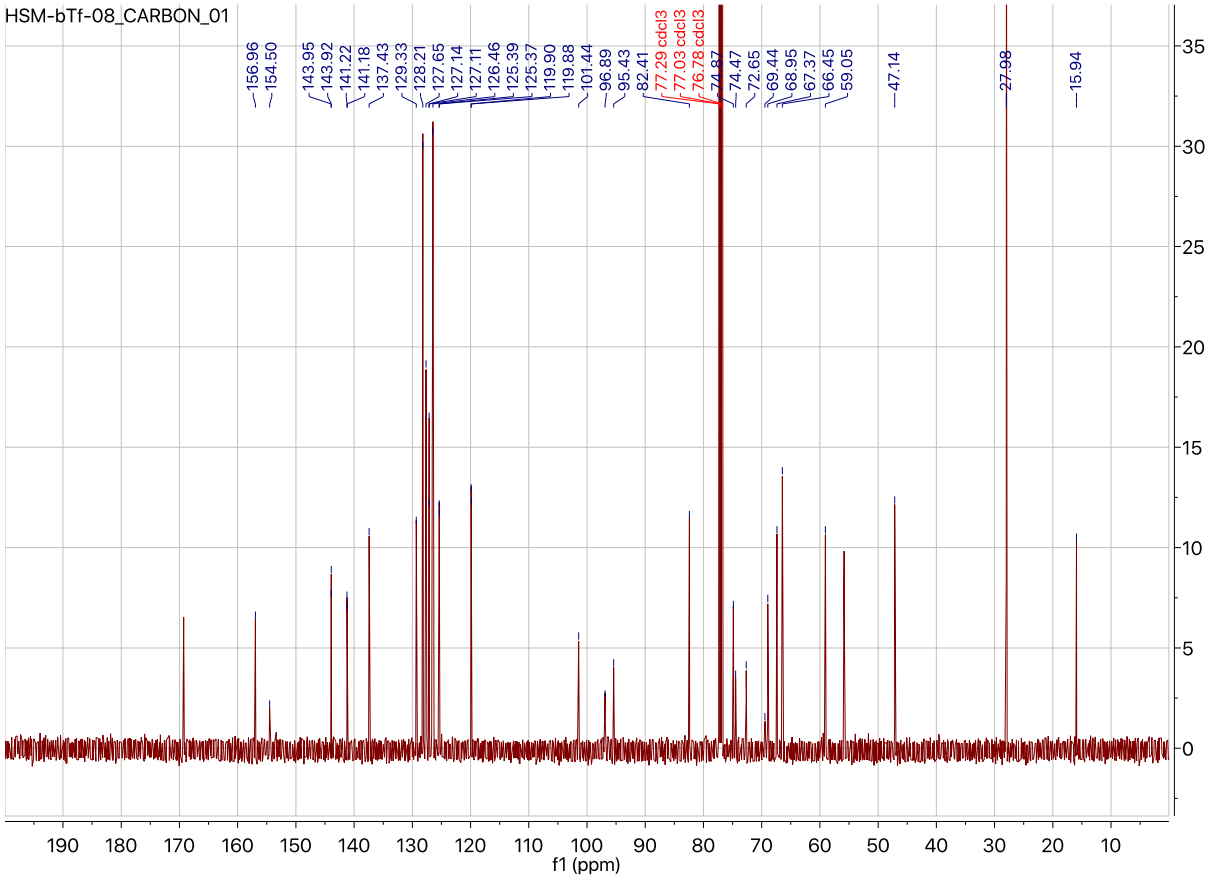


8

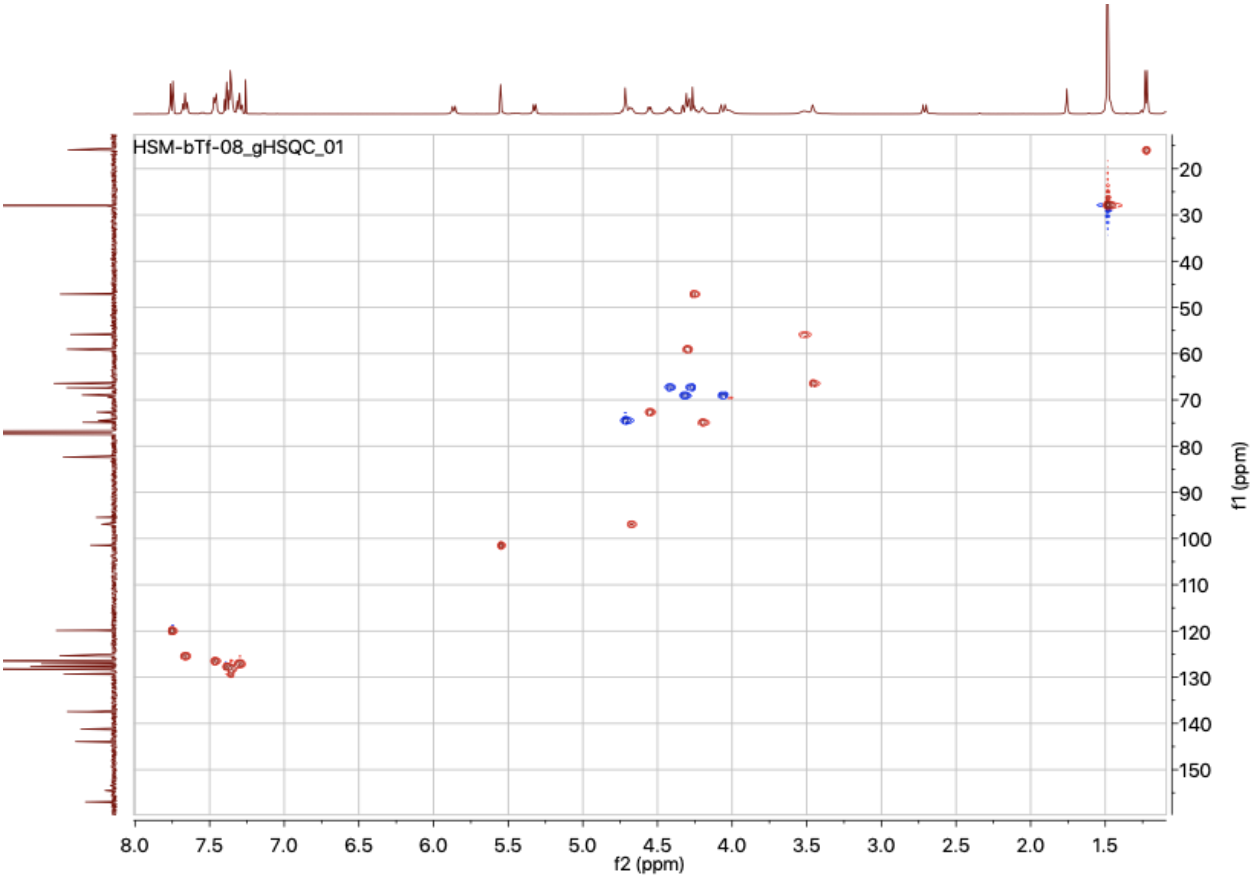
^{1}H NMR (500 MHz, $CDCl_3$) of **8**



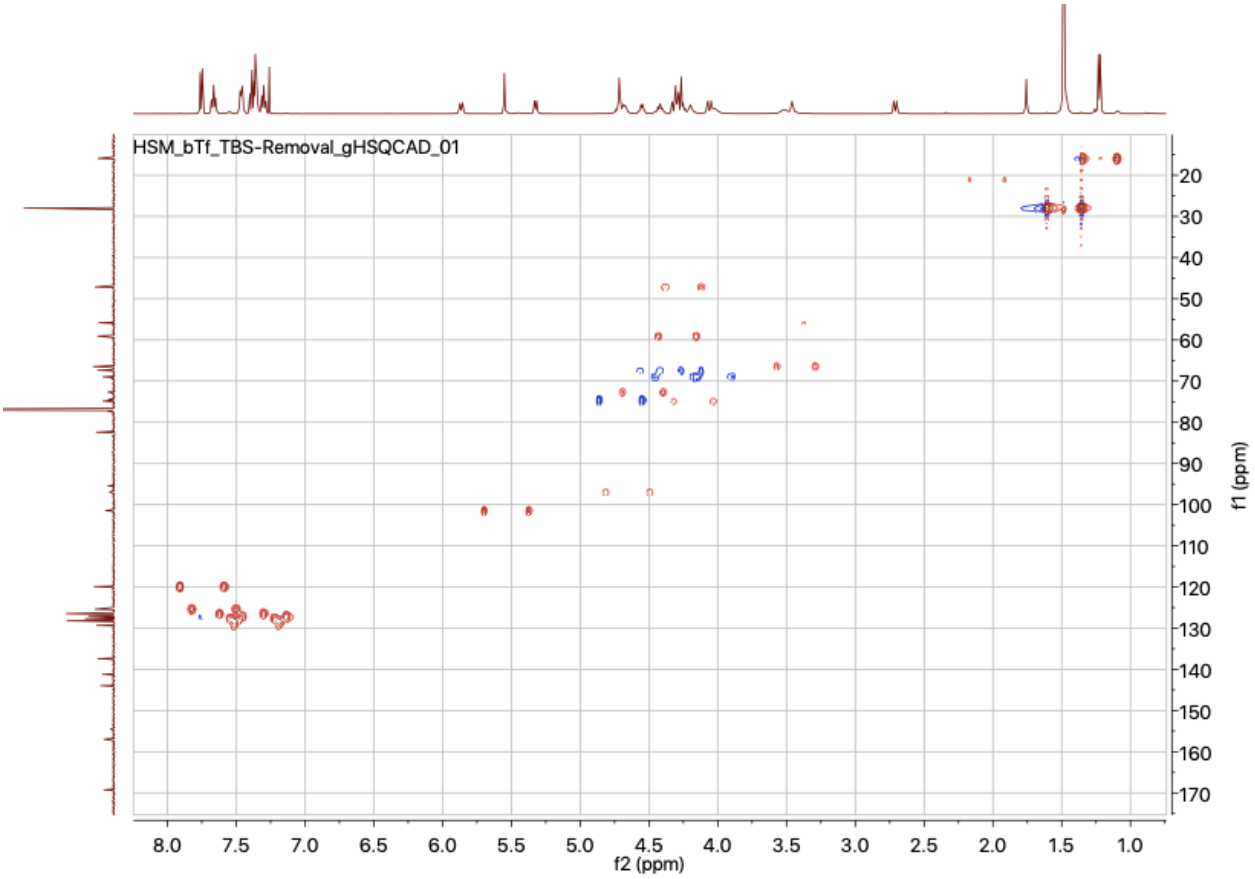
^{13}C NMR (126 MHz, $CDCl_3$) of 8



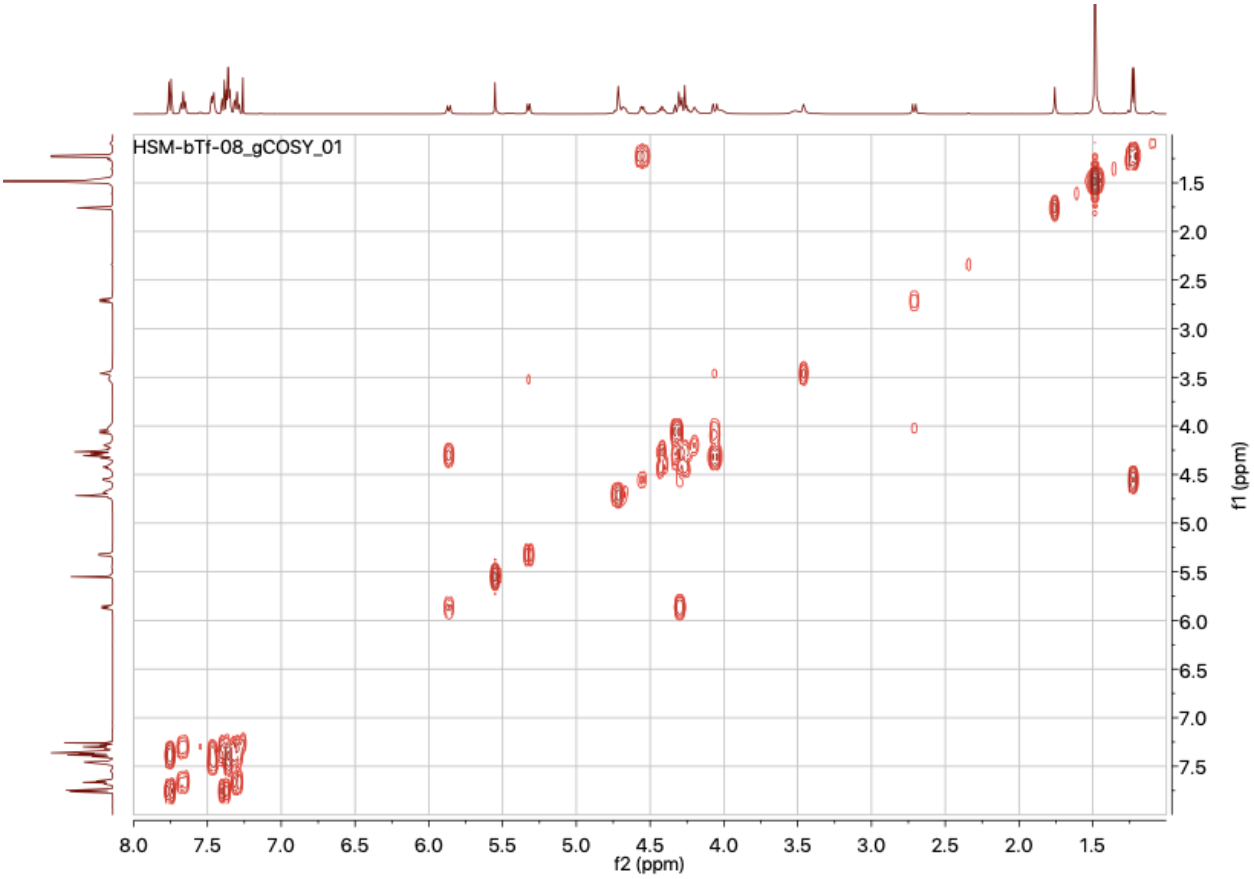
^{1}H-^{13}C HSQC (500 MHz, $CDCl_3$) of 8



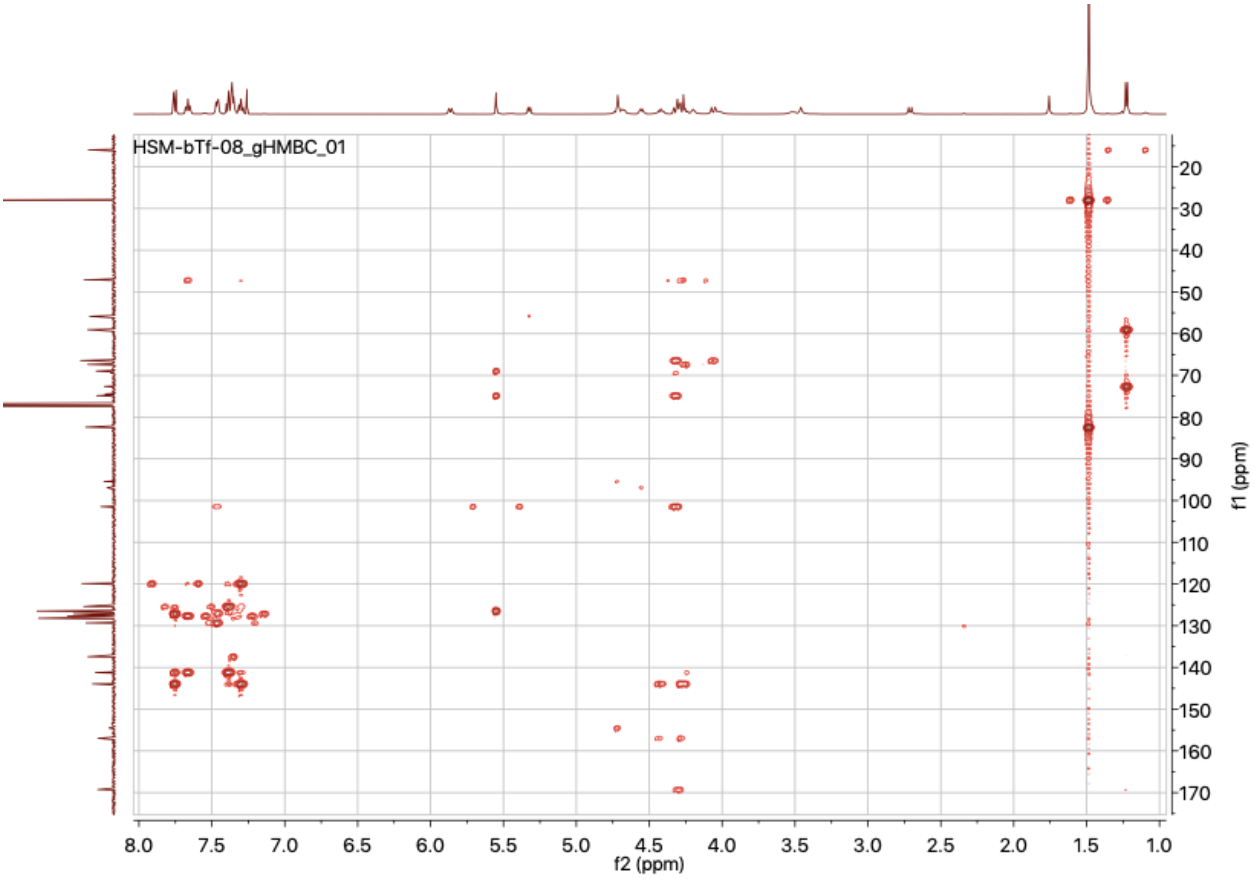
^{1}H-^{13}C HSQCAD (500 MHz, $CDCl_3$) of **8**

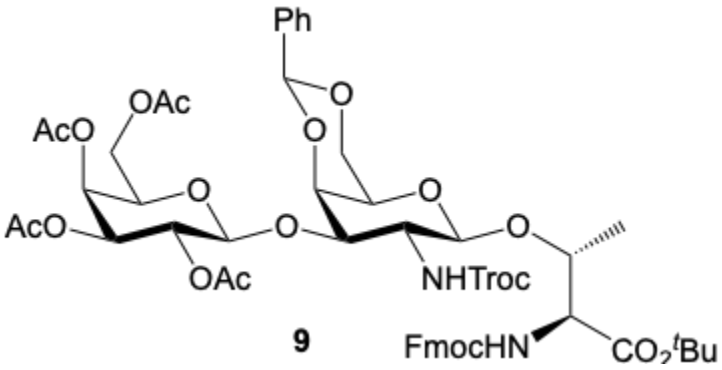


^{1}H-^{1}H COSY (500 MHz, $CDCl_3$) of **8**



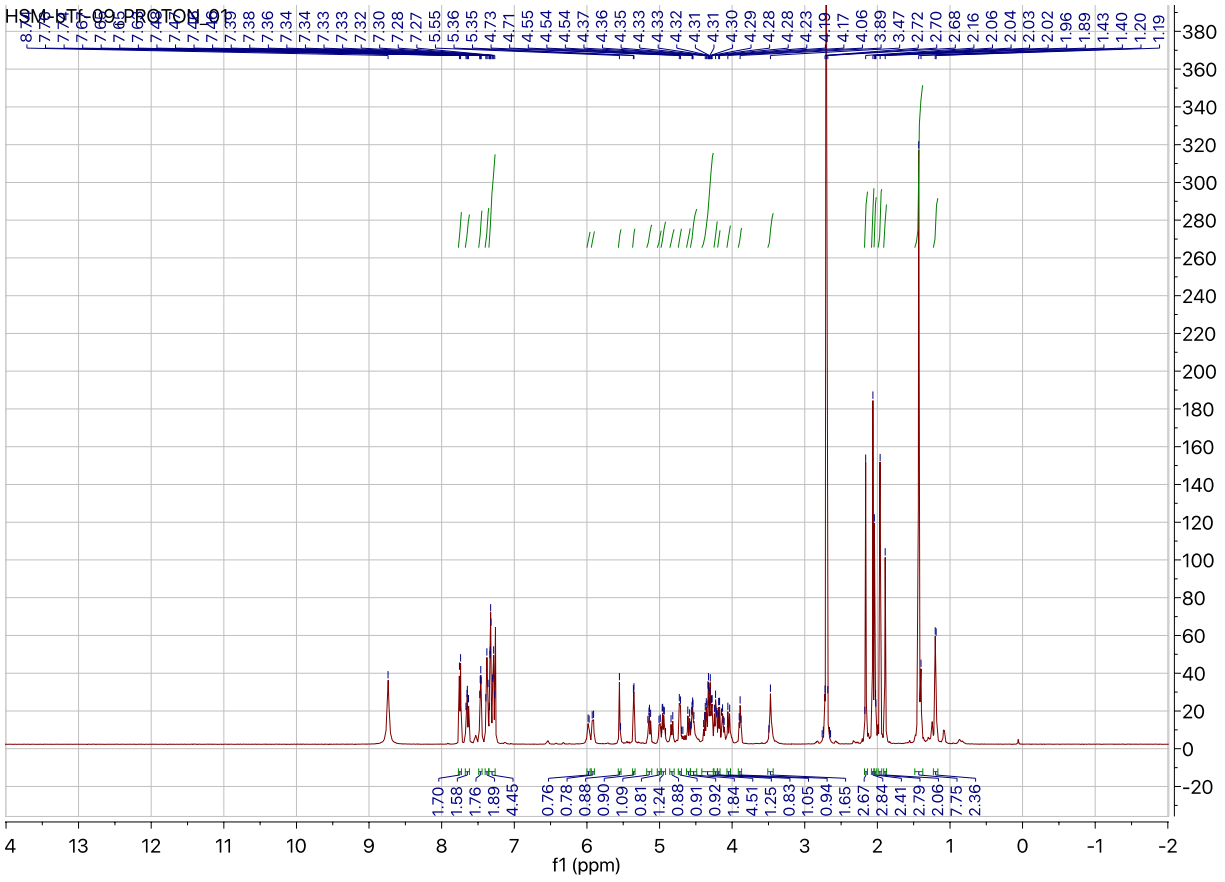
^{1}H-^{13}C HMBC (500 MHz, $CDCl_3$) of **8**



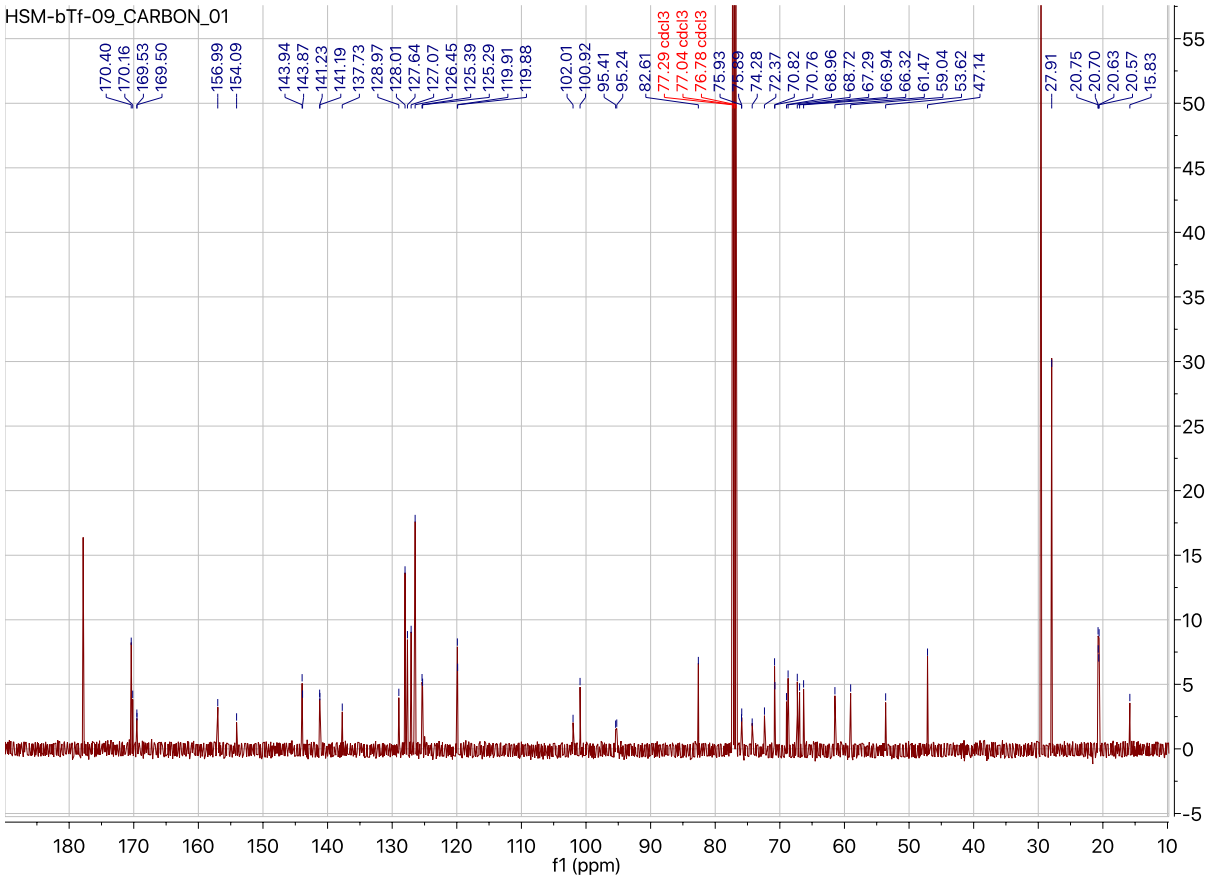


9

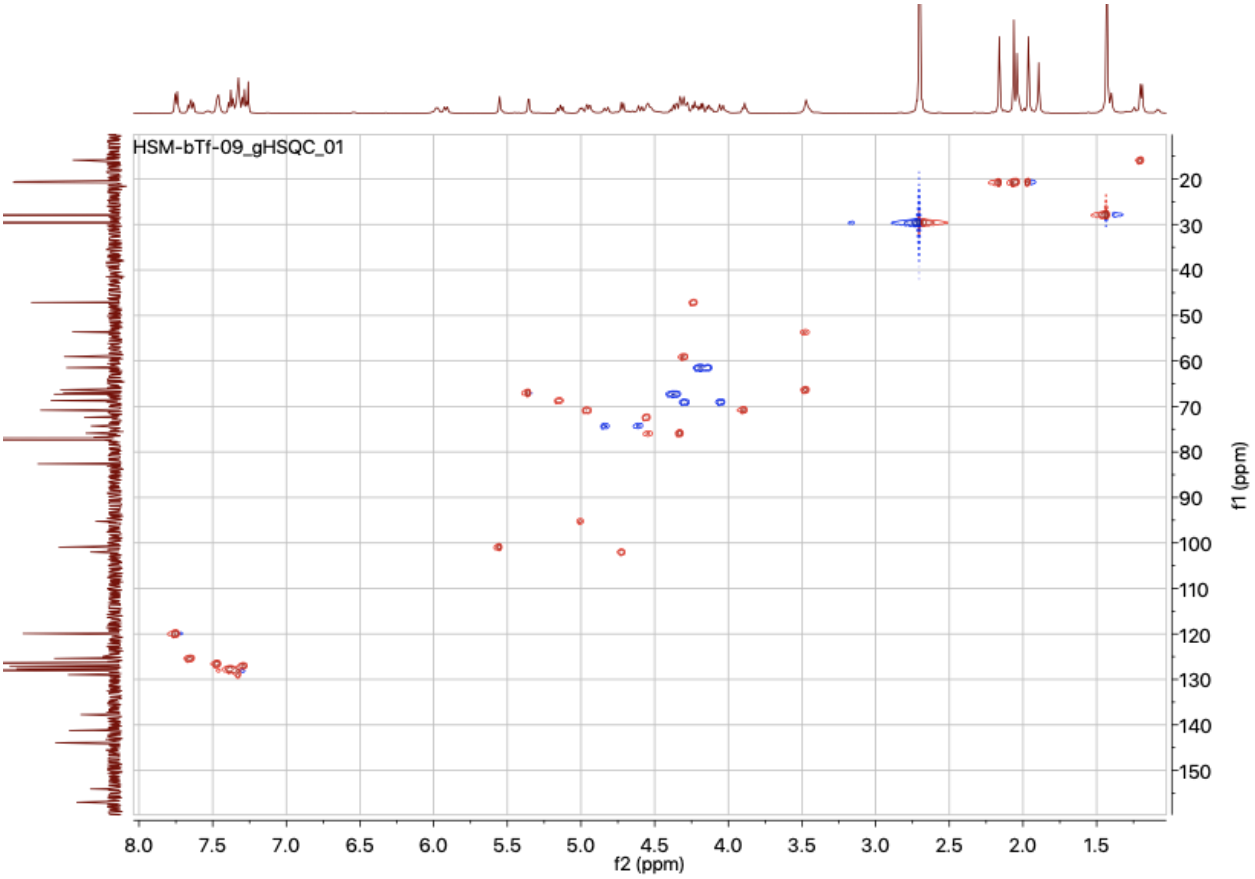
^{1}H NMR (500 MHz, $CDCl_3$) of **9**



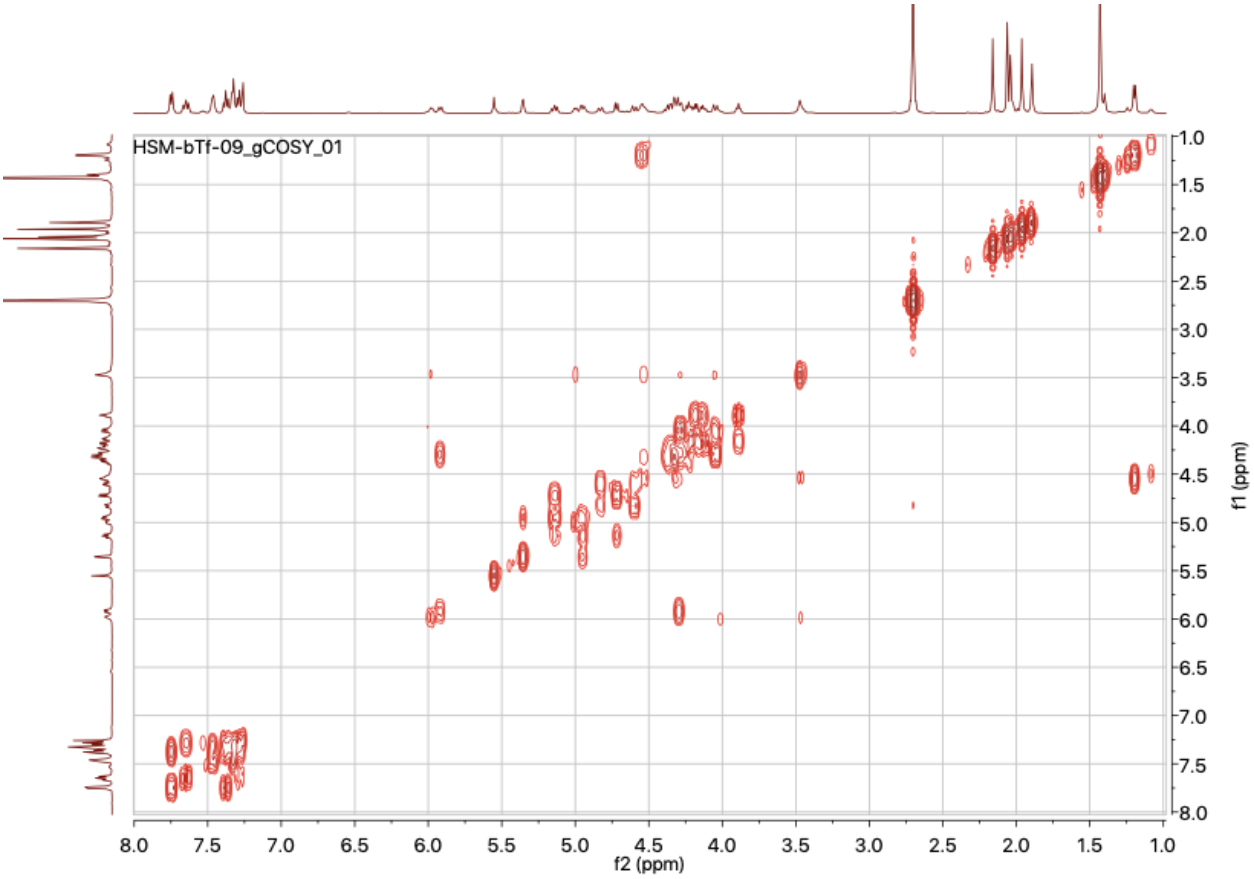
^{13}C NMR (126 MHz, $CDCl_3$) of **9**



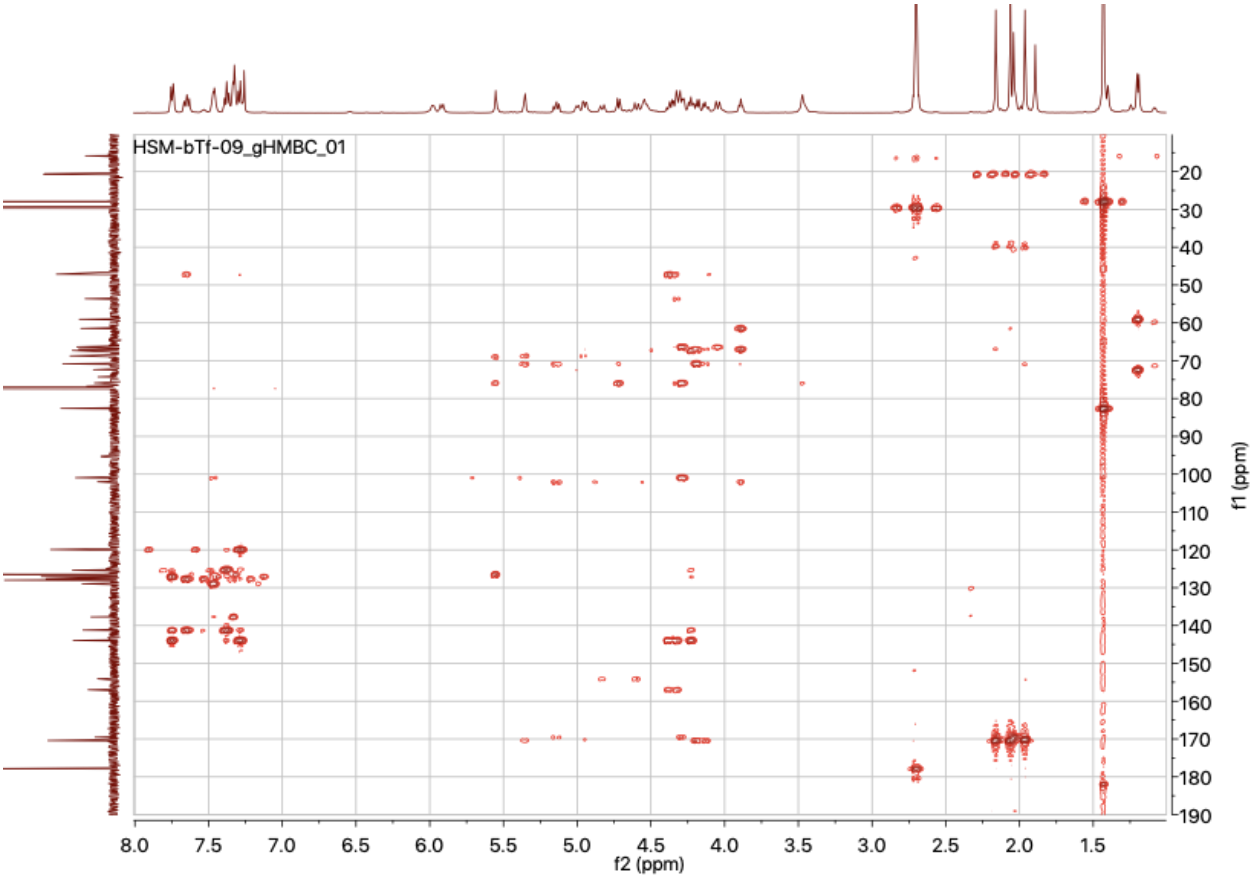
^{1}H-^{13}C HSQC (500 MHz, $CDCl_3$) of **9**

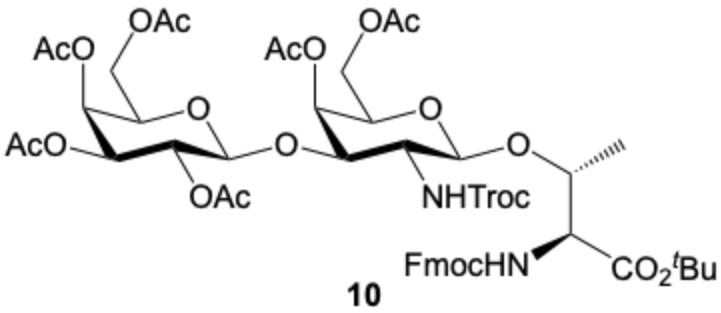


^{1}H-^{1}H COSY (500 MHz, $CDCl_3$) of **9**



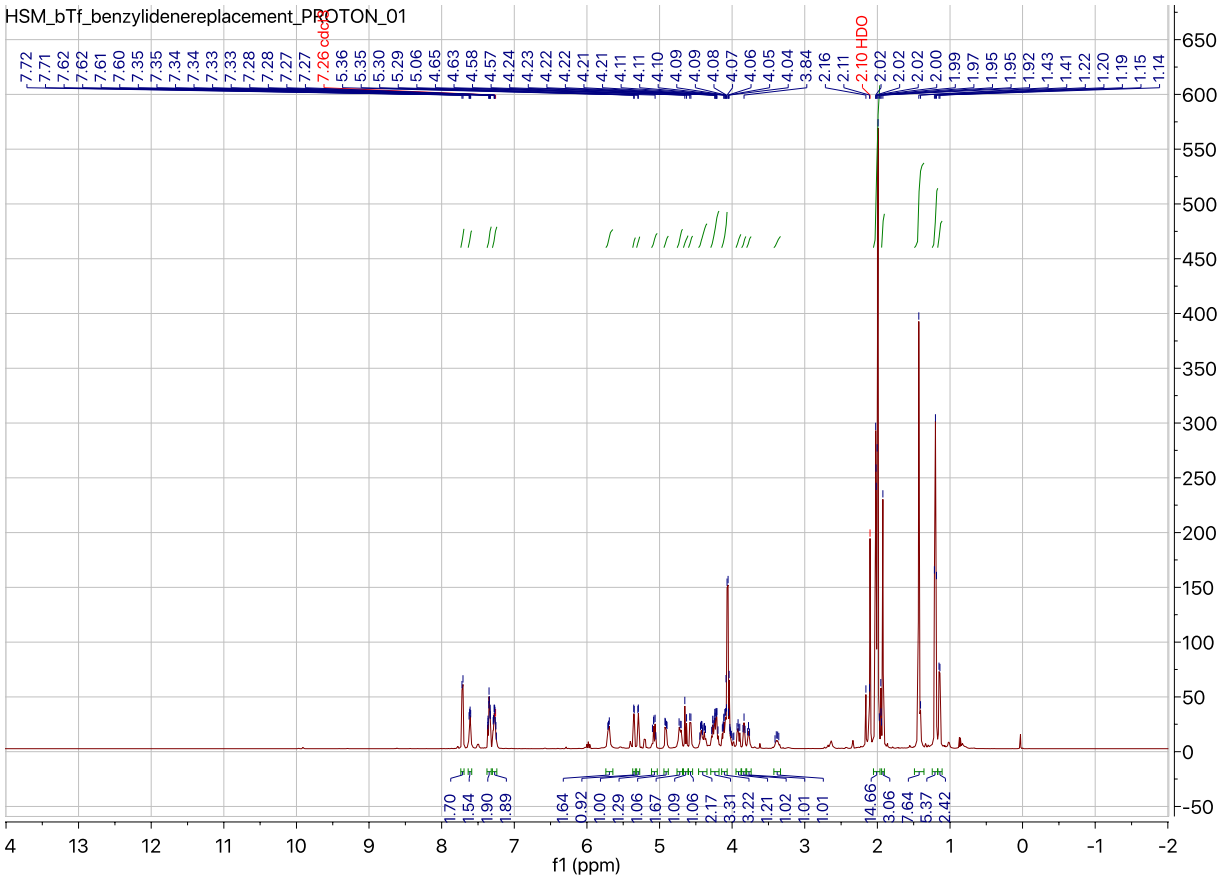
^{1}H-^{13}C HMBC (500 MHz, $CDCl_3$) of **9**



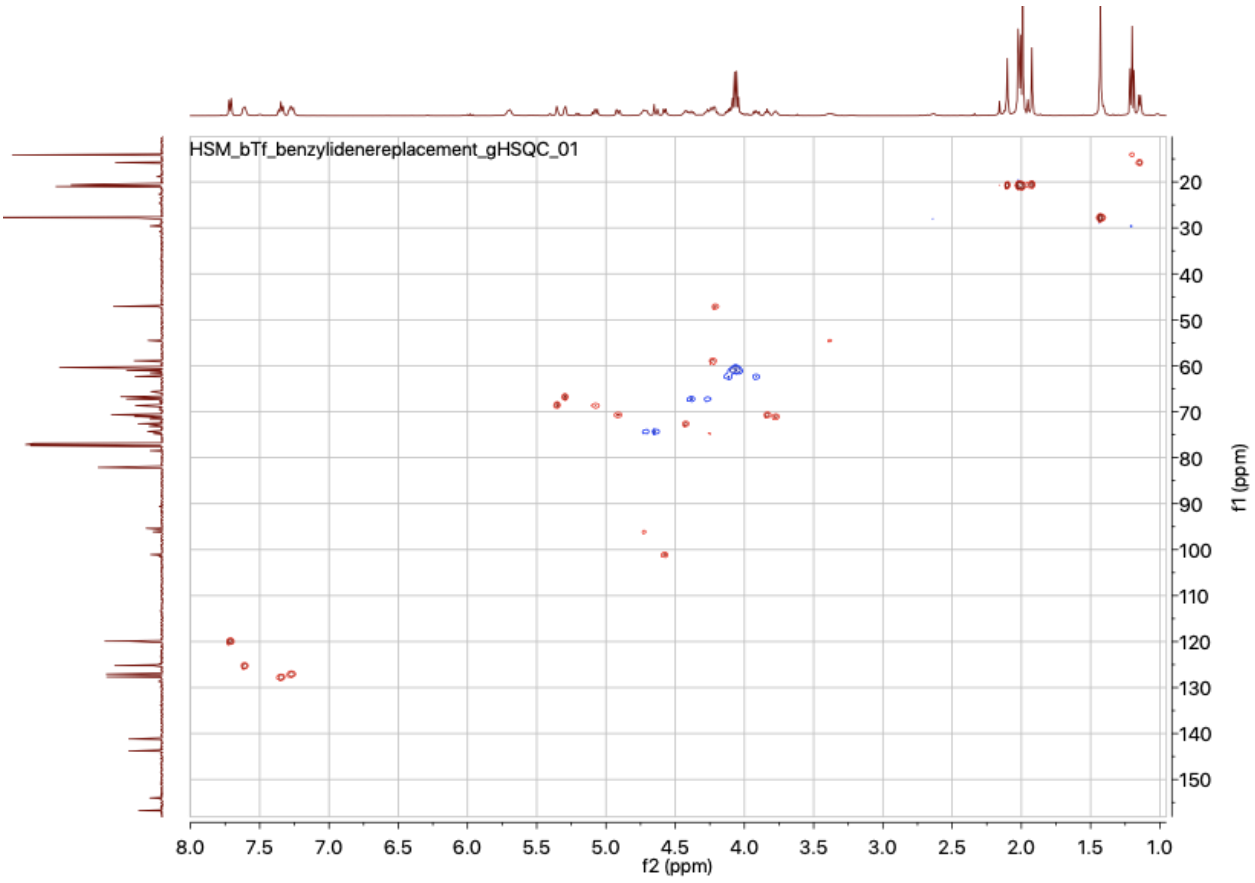


10

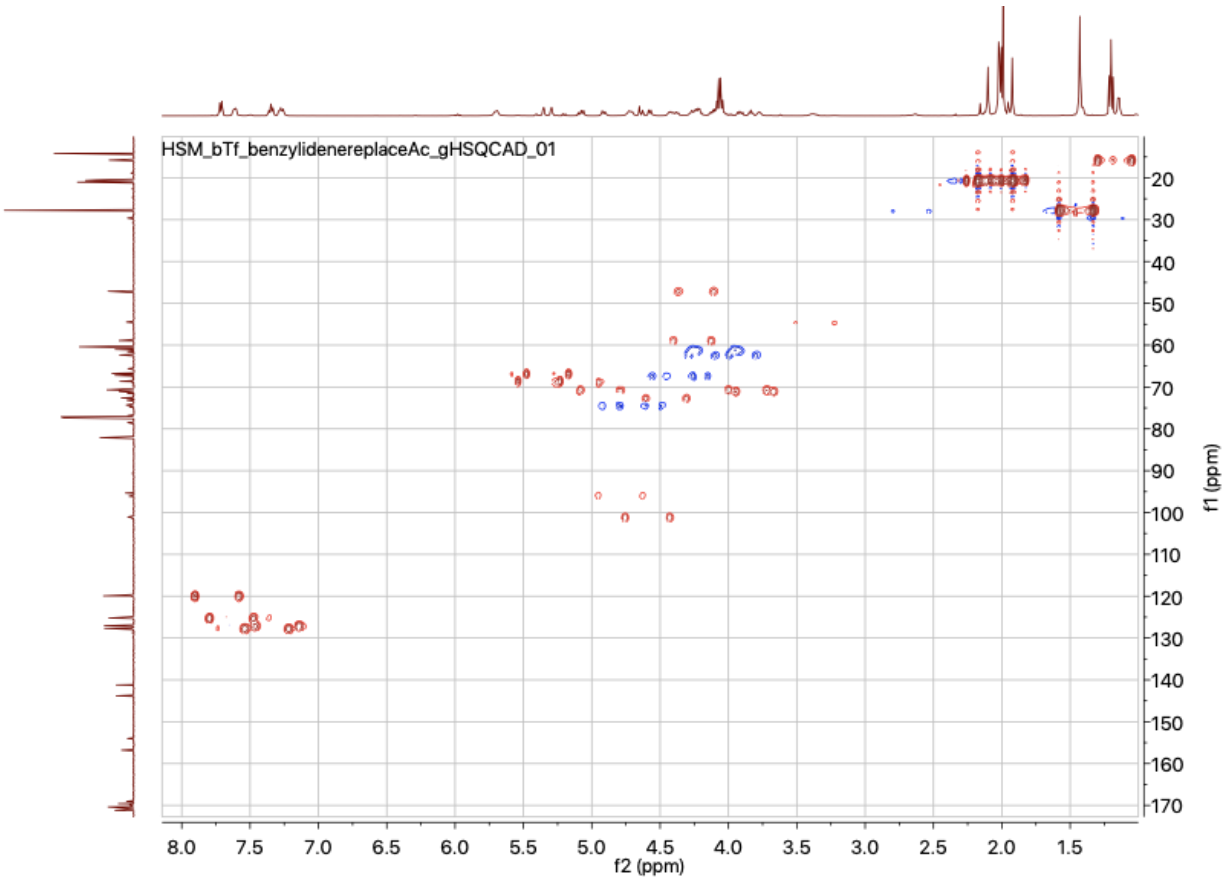
^{1}H NMR (500 MHz, $CDCl_3$) of **10**



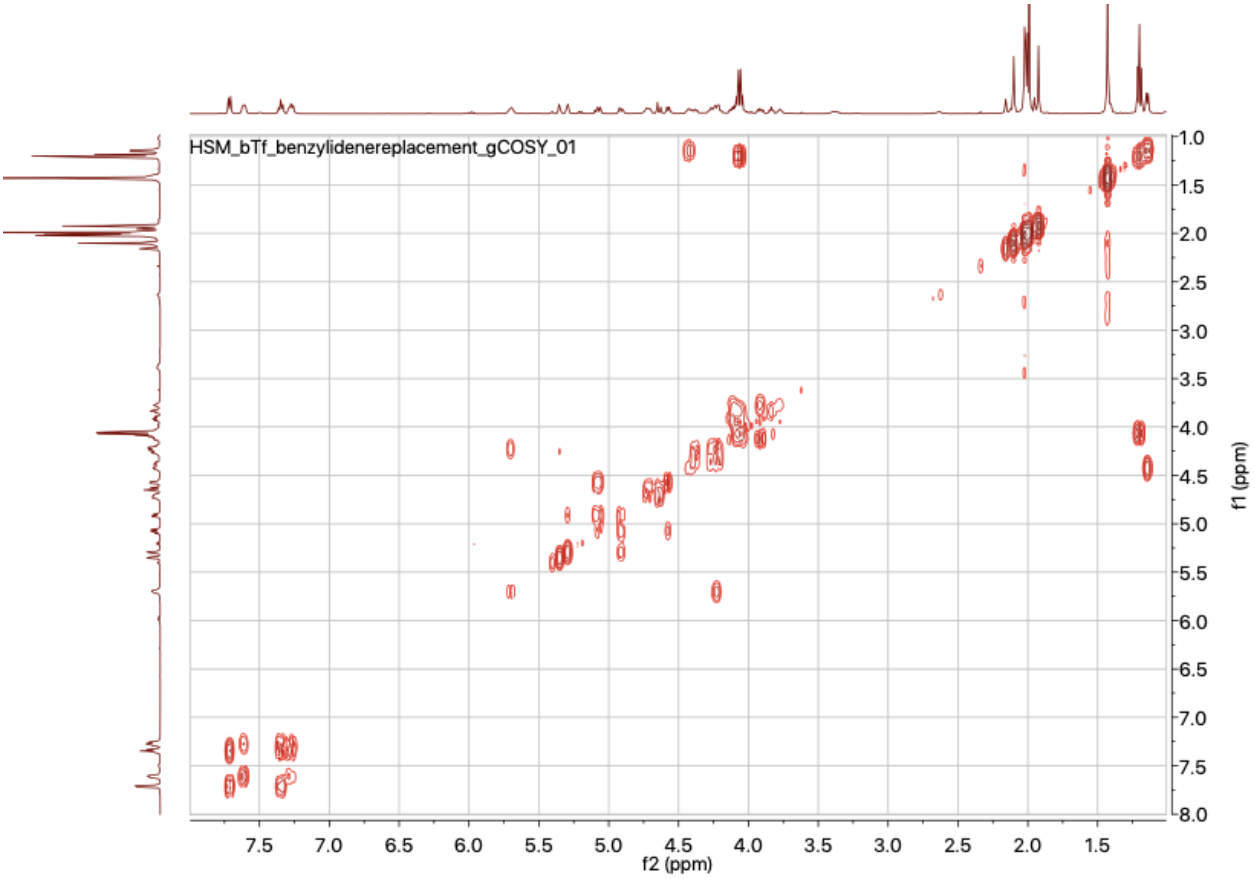
^{1}H-^{13}C HSQC (500 MHZ, $CDCl_3$) of **10**



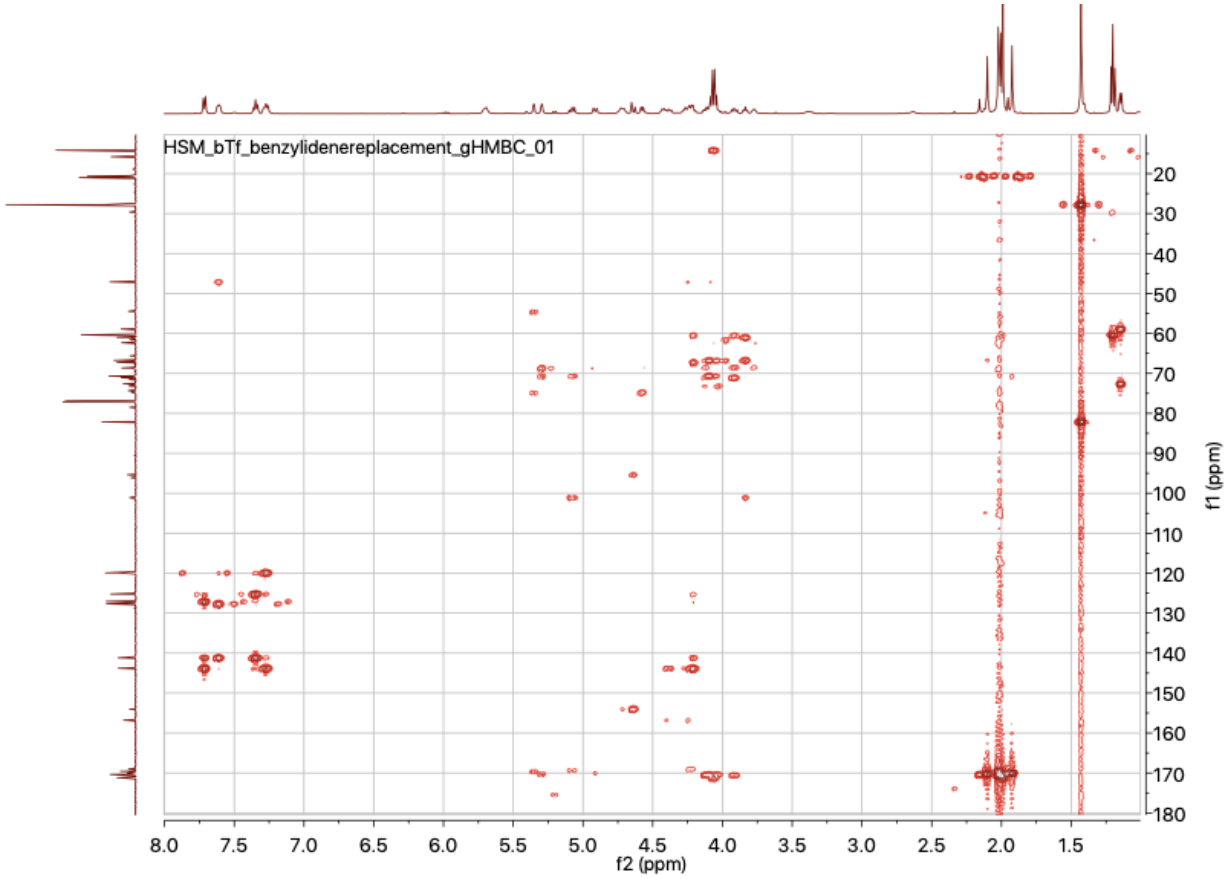
^{1}H-^{13}C HSQCAD (500 MHz, $CDCl_3$) of **10**

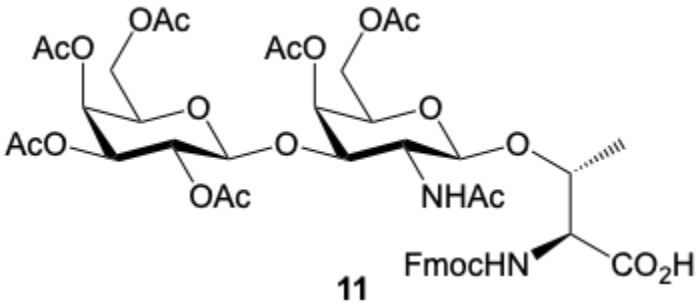


^{1}H-^{1}H COSY (500 MHz, $CDCl_3$) of **10**



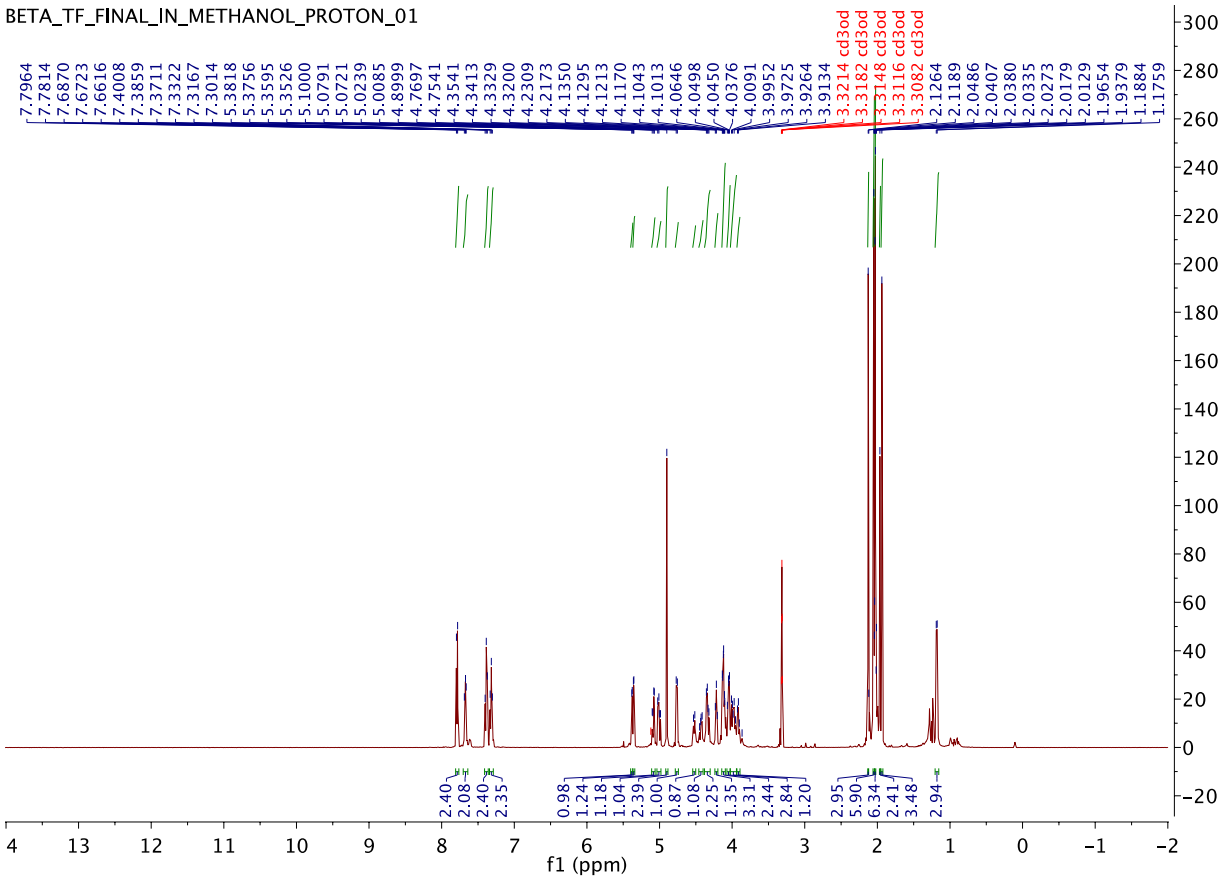
^{1}H-^{13}C HMBC (500 MHz, $CDCl_3$) of **10**



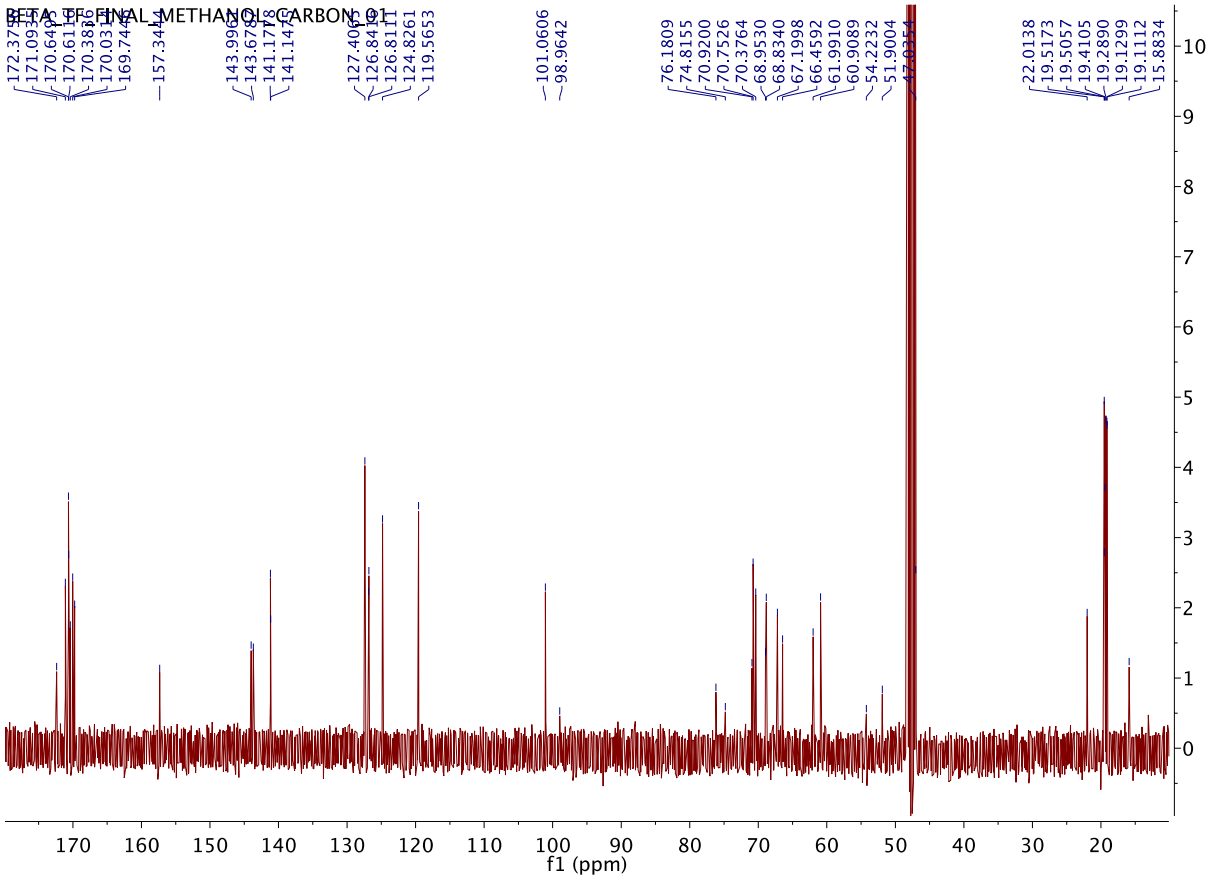


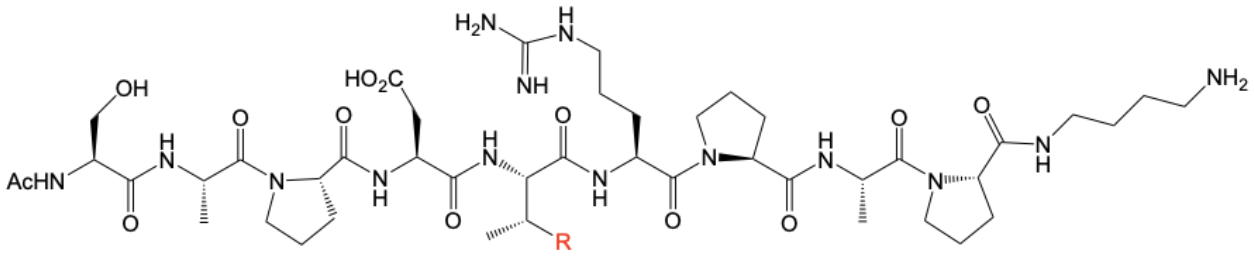
11

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

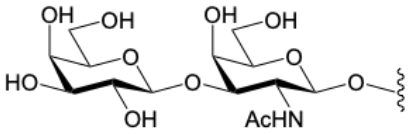


^{13}C NMR (126 MHz, CD_3OD) of **11**

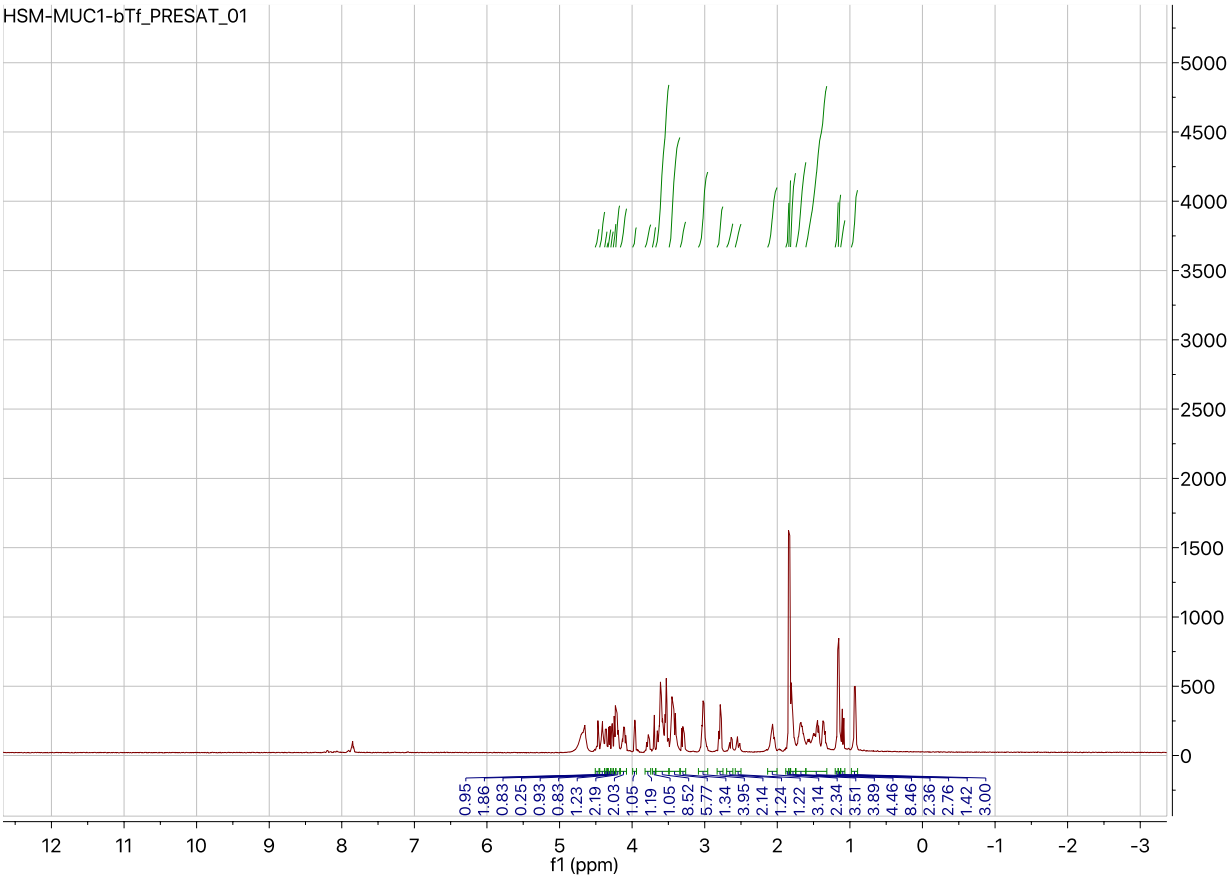




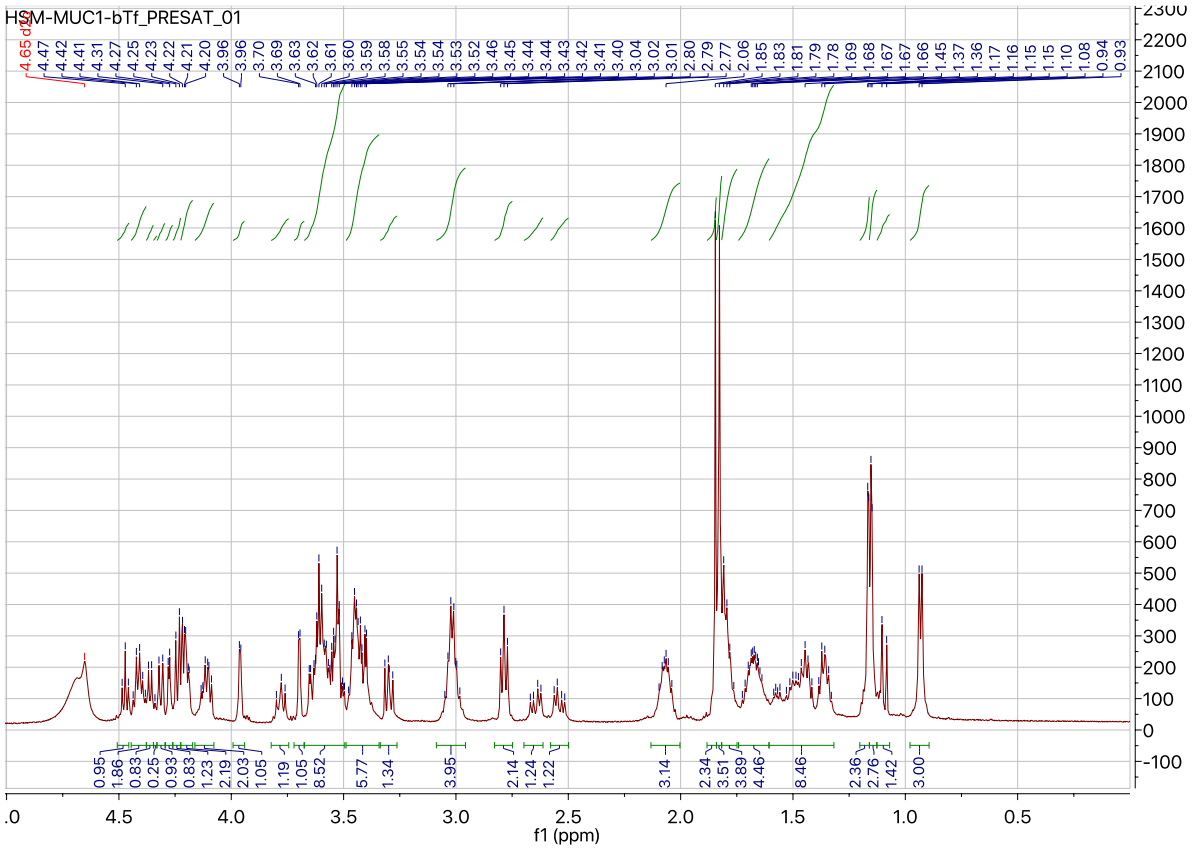
12 : R =



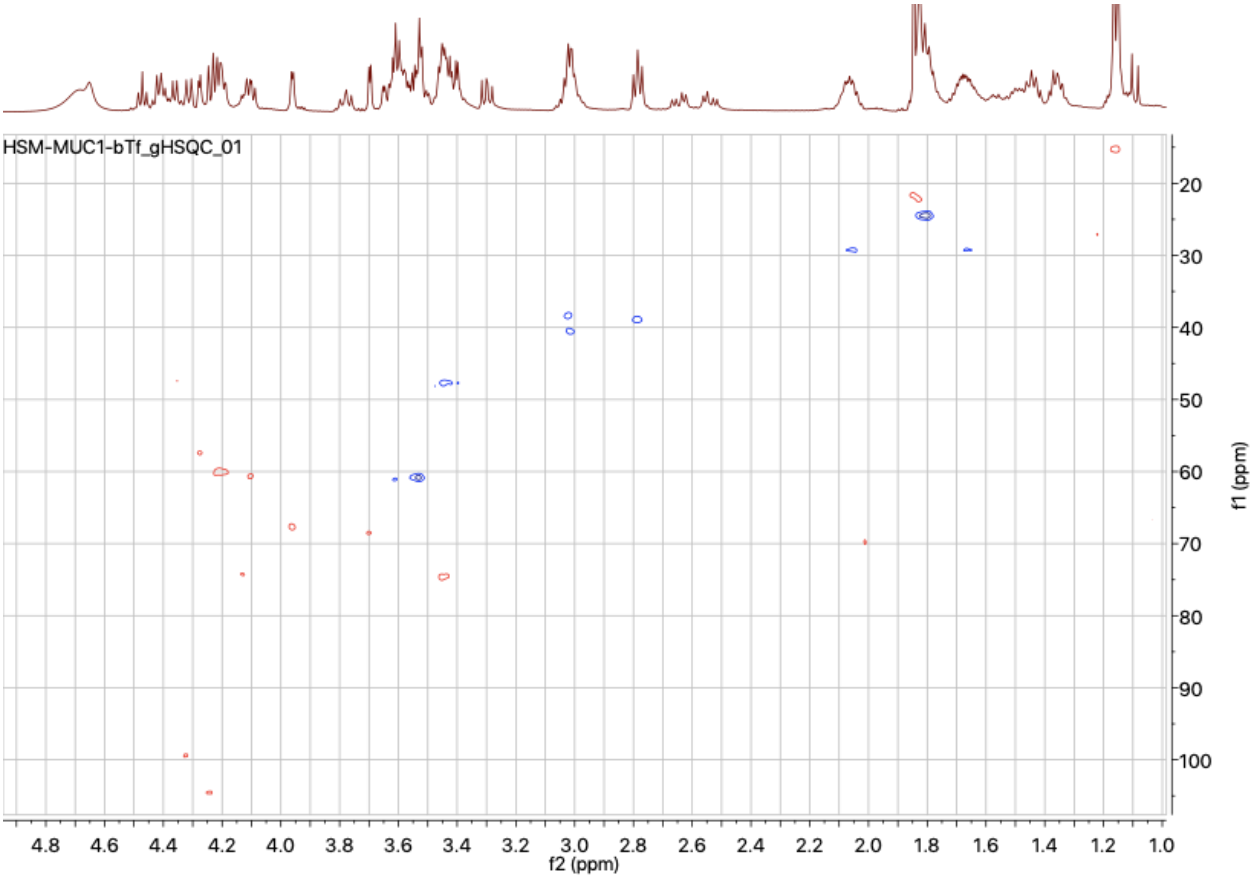
^{1}H NMR (500 MHz, D_2O) of **12**



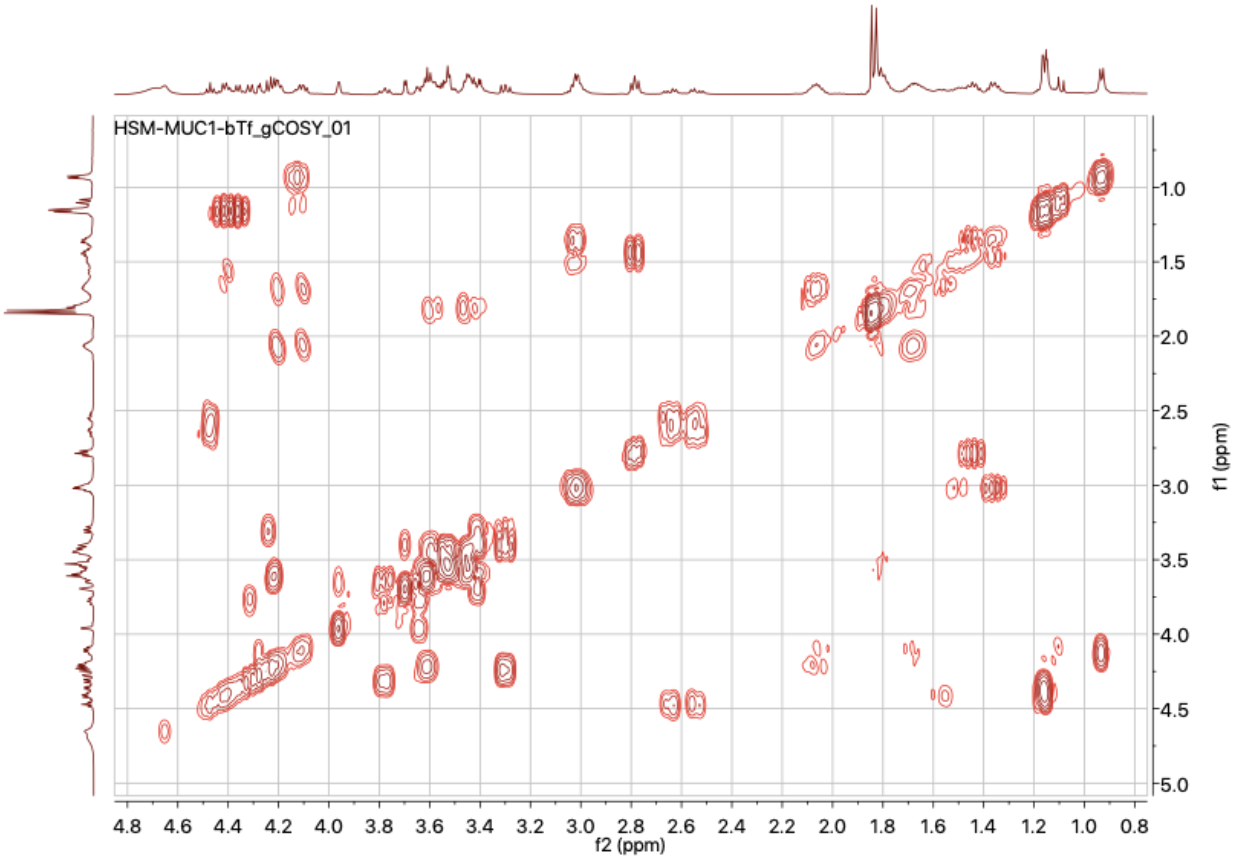
Zoomed in ^{1}H NMR (500 MHz, D_2O) of **12**



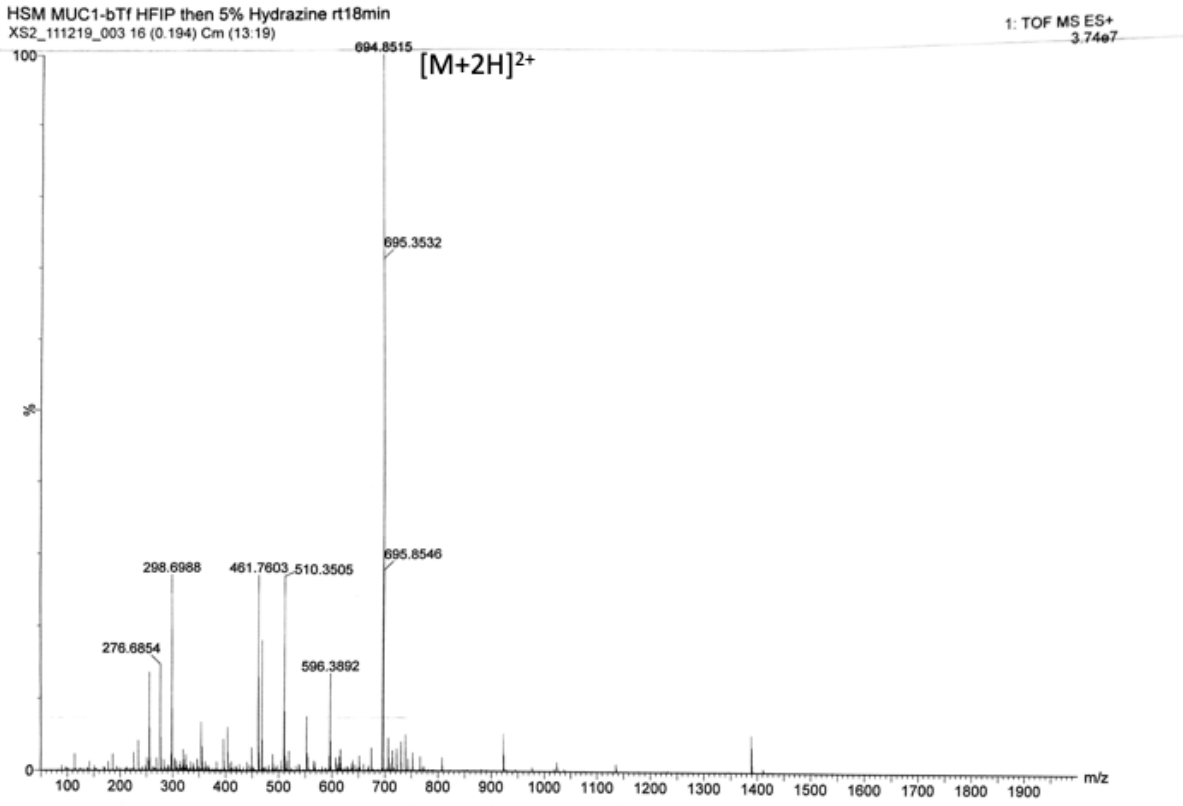
^{1}H-^{13}C HSQC NMR (500 MHz, D_2O) of **12**



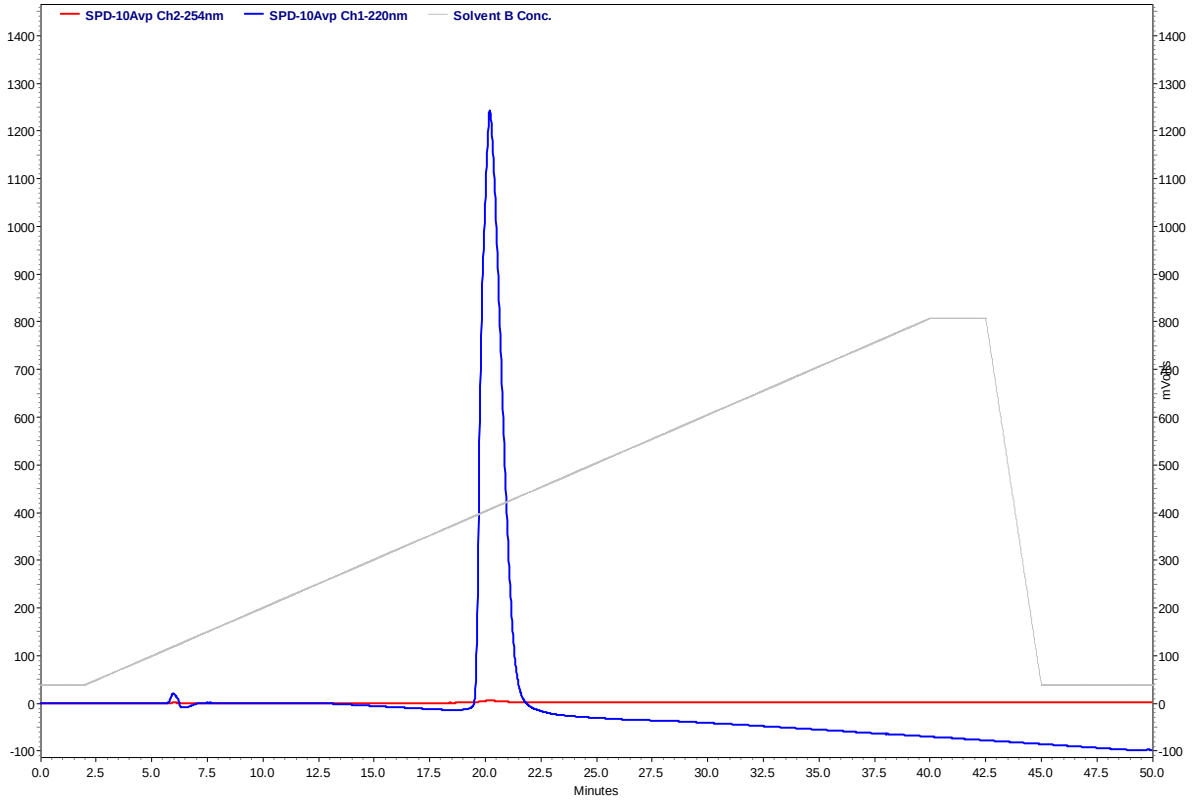
^{1}H-^{1}H COSY (500 MHz, D_2O) of **12**



ESI-TOF MS of **12**



HPLC of purified **12**



References:

(1) Wu, X.; McKay, C.; Pett, C.; Yu, J.; Schorlemer, M.; Ramadan, S.; Lang, S.; Behren, S.; Westerlind, U.; Finn, M. G.; Huang, X. Synthesis and Immunological Evaluation of Disaccharide Bearing MUC-1 Glycopeptide Conjugates with Virus-like Particles. *ACS Chem. Biol.* **2019**, *14*, 2176-2184.

(2) Wang, Z.; Zhou, L.; El-boubbou, K.; Ye, X.-S.; Huang, X. Multi-Component One-Pot Synthesis of the Tumor-Associated Carbohydrate Antigen Globo-H Based on Preactivation of Thioglycosyl Donors. *J. Org. Chem.* **2007**, *72*, 6409-6420.