# Supporting information

# Synthesis and Immunological Evaluation of the Unnatural $\beta$ -linked Mucin-1 Thomsen-Friedenreich Conjugate

Xuanjun Wu,<sup>§†</sup> Hunter McFall-Boegeman,<sup>†,‡</sup> Zahra Rashidijahanabad,<sup>†,‡</sup> Kunli Liu,<sup>†,‡</sup> Christian Pett,<sup>∥,#</sup> Jin Yu,<sup>∥</sup> Manuel Schorlemer,<sup>∥,#</sup> Sherif Ramadan,<sup>†,‡¶</sup> Sandra Behren,<sup>∥,#</sup> Ulrika Westerlind,<sup>∥,#</sup> and Xuefei Huang<sup>\*,†,‡,</sup>

<sup>§</sup>National Glycoengineering Research Center, Shandong University, Qingdao, Shandong 266237, China

<sup>†</sup>Department of Chemistry, <sup>‡</sup>Institute for Quantitative Health Science and Engineering, and <sup>v</sup>Department of Biomedical Engineering, Michigan State University, East Lansing, Michigan 48824, USA

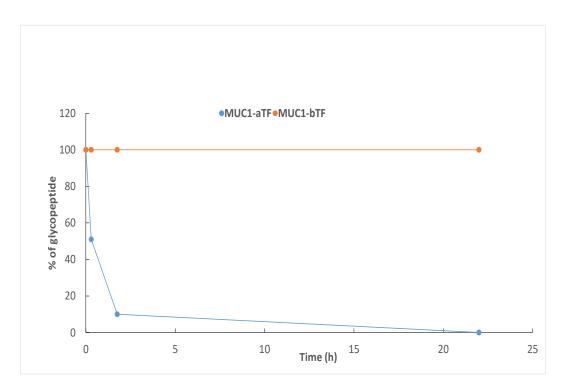
Leibniz-Institut für Analytische Wissenschaften–ISAS–e.V., 44227, Dortmund, Germany

<sup>#</sup>Department of Chemistry, Umeå University, 901 87 Umeå, Sweden

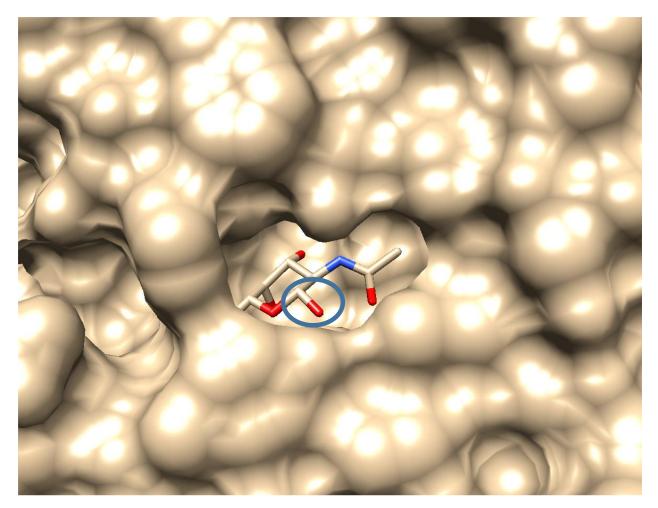
<sup>¶</sup>Chemistry Department, Faculty of Science, Benha University, Benha, Qaliobiya 13518, Egypt

## **Table of Contents**

Figure S1. Stability of MUC1 glycopeptides to glycosidase enzymes						
<b>Figure S2.</b> Crystal structure of the <i>Enterococcus faecalis</i> endo- $\alpha$ - <i>N</i> -acetylgalactosar (PDB:6M77) with $\beta$ -GalNAc						
Figure S3. MALDI-TOF MS spectra obtained for Qβ-MUC1-β-Tf conjugates						
Scheme S1. Synthesis of BSA-MUC1 conjugates for ELISA						
Figure S4. MALDI-TOF MS spectra of BSA-MUC1-β-Tf conjugates for ELISA	<b>S</b> 6					
Figure S5. Flow cytometry analysis of B16-MUC1 melanoma cells by IgG antibodies immune sera	in post- S7					
Figure S6. Glycopeptide microarray screening results of antisera induced by $Q\beta$ -MUC1-	β-Tf S10					
General Experimental Procedures and Methods for Synthesis	S11					
Synthesis of $\beta$ -Tf Glyco-AA Building Block <b>11</b>	S12					
Synthesis of MUC1-β-Tf glycopeptide 12	S15					
Endoglycosidase cleavage experiment	S16					
Synthesis of Qβ-MUC1-β-Tf	S16					
Synthesis of BSA-MUC1 conjugates	S16					
Immunization of MUC1.Tg mice	S16					
Evaluation of Antibody Titers by ELISA	S17					
Detection of Antibody Binding to Tumor Cells by FACS	S17					
Complement Dependent Cytotoxicity	S17					
Glycopeptide Microarray Analysis	S18					
Characterization Data and Spectra of Building Blocks and MUC1 Glycopeptides	S20					
References	S49					



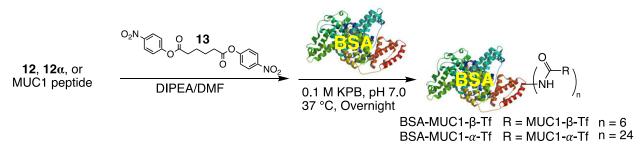
**Figure S1**. MUC1- $\beta$ -Tf has much enhanced stability toward the *Enterococcus faecalis* endo- $\alpha$ -*N*-acetylgalactosaminidase compared to MUC1- $\alpha$ -Tf over 24h reaction time.



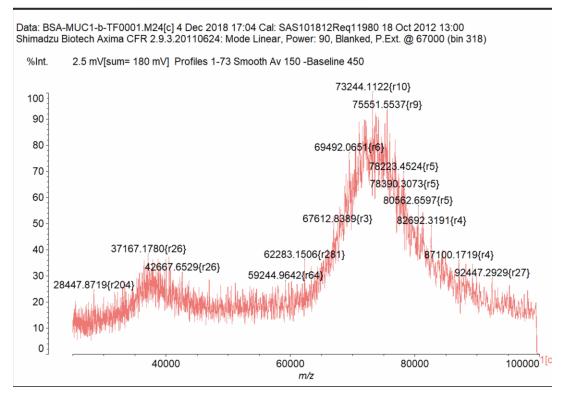
**Figure S2.** Crystal structure of the *Enterococcus faecalis* endo- $\alpha$ -*N*-acetylgalactosaminidase (PDB:6M77) with  $\beta$ -GalNAc showed that the 1-hydroxyl group (in blue oval) pointing toward the enzyme, presumably hindering the binding of MUC1- $\beta$ -Tf glycopeptide **12** with the enzyme. This is consistent with the enhanced stability of MUC1- $\beta$ -Tf glycopeptide **12** toward the enzyme compared to the corresponding MUC1- $\alpha$ -Tf glycopeptide **12a**. The enzyme is shown as spacing filling model in light brown color.  $\beta$ -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).

b-Tf-MG0001.L2[c] 22 Jun 2017 10:33 Cal: SAS101812Req11980 18 Oct 2012 13:00 zu Biotech Axima CFR 2.9.3.20110624: Mode Linear, Power: 90, Blanked, P.Ext. @ 120 mV Profile 297 Q	14000 (bin 144) β + 2 glycopep.
Q <b>β</b> + 1 glycopep. 13989.21	15459.86 Qβ + 3 glycopep. 16929.45
56 mV[sum= 16499 mV]. Profiles 1-297 Smooth Av 150 -Baseline 450	

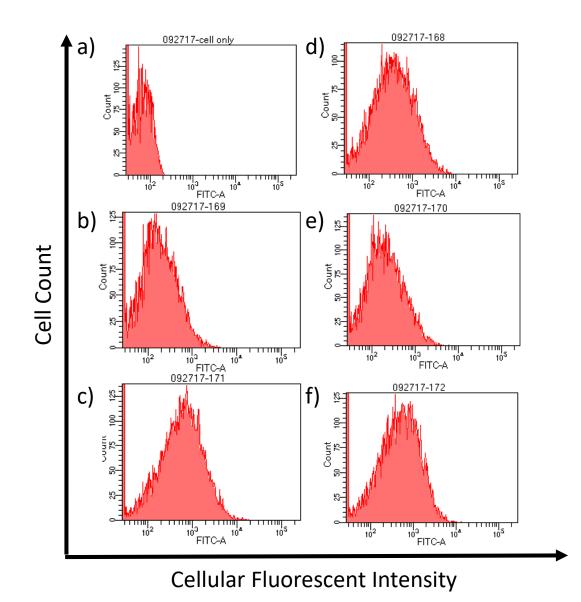
**Figure S3.** MALDI-TOF MS of Q $\beta$ -MUC1- $\beta$ -Tf conjugate. Based on the ratio of peak intensities, the average number of MUC1- $\beta$ -Tf per capsid was calculated to be 220.



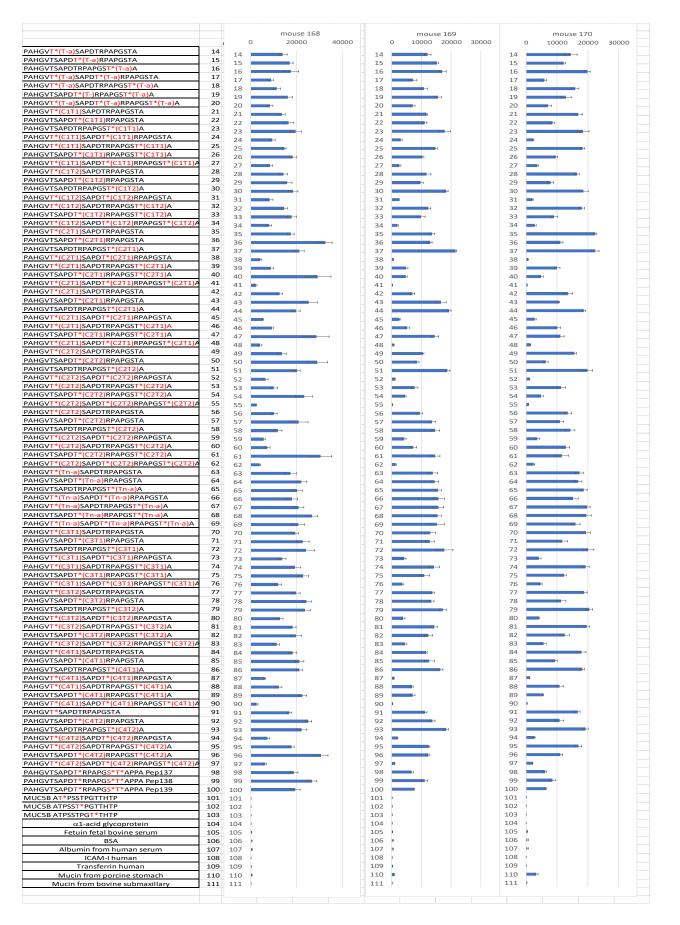
Scheme S1: Synthesis of BSA-MUC1 conjugates for ELISA. BSA-MUC1, BSA-MUC1- $\alpha$ -Tf were reported previously.<sup>1</sup>



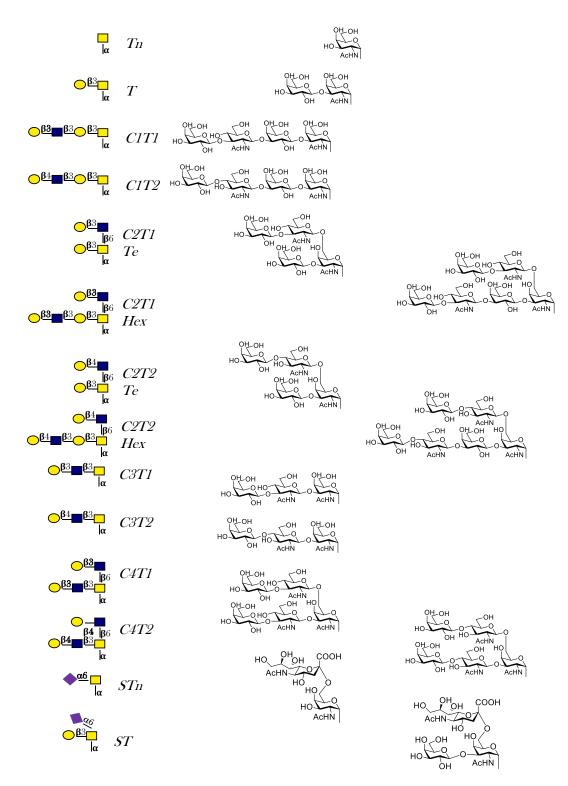
**Figure S4**. MALDI-TOF MS of BSA-MUC1- $\beta$ -Tf conjugates. Based on the molecular weight difference between the BSA- MUC1- $\beta$ -Tf conjugate and unmodified BSA, the number of MUC1- $\beta$ -Tf per BSA was calculated to be 6.



**Figure S5.** B16-MUC1 melanoma cells staining by IgG antibodies in post-immune sera elicited by Q $\beta$ -MUC1- $\beta$ -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 Q $\beta$ -MUC1- $\beta$ -Tf. The binding was tested with 1:20 dilution of the sera.



DALGONTATIC TRANSPORT         0         2000         4000         90000         90000           DALGONTATIC TRANSPORT         15         1						
DARGY TI LOSAPPTERANSSIA         44         14         14         14           DARGY TI LOSAPPTERANSSIA         55         15         16         16           DARGY TI LOSAPPTERANSSIA         57         15         16         16           DARGY TI LOSAPPTERANSSIA         77         15         16         16           DARGY TI LOSAPPTERANSSIA         26         16         16         16           DARGY TI LOSAPPTERANSSIA         26         16         16         16           DARGY TI LOSAPPTERANSSIA         26         16         26         16           DARGY TI LOSAPPTERANSIA         26         16         26         16           DARGY TI LOSAPPTERANSIA         26         16         16         16           DARGY TI LOSAPPTERANSIA         26         16         16         16           DARGY TI LOSAPPTERANSIA         26         16         16         16         16         16         16         16						
DALESCY DAPPT T1 SIRANGESTA         15         15         15         16           DALESCY T1 SIRANGEST (1) A         36         16         16         16           DALESCY T1 SIRANGEST (1) A         36         16         16         16           DALESCY T1 SIRANGEST (2) A         36         16         16         16           DALESCY T1 SIRANGEST (2) A         36         16         16         16           DALESCY T1 SIRANGEST (2) A         36         16         16         16           DALESCY T1 SIRANGEST (2) A         36         16         16         16           DALESCY T2 SIRANGEST (2) A         36         16         <	PAHGVT*(T-a)SAPDTRPAPGSTA	14	14			
DAMESON TIL DEADED TIL DIRAPADESTA         17         17         17           DAMESON TIL DIRAPADESTIL DIRAPADE	PAHGVTSAPDT*(T-a)RPAPGSTA	15				
PARLSONT CT 3: BARPETER PROPERTY 1: ALA         18         10         10           PARLSONT CT 3: BARPETER PROPERTY 1: ALA         20         20         20           PARLSONT CT 1: BARPETER PROPERTY 1: ALA         20         20         20           PARLSONT CT 1: BARPETER PROPERTY 1: ALA         20         20         20           PARLSONT CT 1: BARPETER PROPERTY 1: ALA         20         20         20           PARLSONT CT 1: BARPETER CT 1: BARPETER PROPERTY 1: ALA         20         20         20           PARLSONT CT 1: BARPETER CT 1: BARPETER CT 1: ALA         20         20         20           PARLSONT CT 1: BARPETER CT 1: BARPETER CT 1: ALA         20         20         20           PARLSONT CT 1: BARPETER CT 1: BA			16	16		
DARASYLANDICTI INCARESITE ALA         39						
PARESVIT CT 1280001         C1 200         C1 200         C1 200           PARESVIT CT 1280001         C1 200         C1 200         C1 200           PARESVIT ADDITEARDEST (C1 1)A         C1 200         C1 200         C1 200           PARESVIT ADDITEARDEST (C1 1)A         C1 200         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C1 1280001         C1 218000         C1 200         C1 200           PARESVIT C2 11800000000000000000000000000000000000						
DARSON TARADIT CG113RARDESTA         22         22         23         24         25         <	PAHGVT*(T-a)SAPDT*(T-a)RPAPGST*(T-a)A				-	
DARGYTCANDTERAPECT (C11)A         23         23         23           DARGYTCANDTERAPECT (C11)A         26         25         24           DARGYTCANDTERAPECT (C11)A         26         26         26           DARGYTCANDTERAPECT (C11)A         26         26         26           DARGYTCANDTERAPECT (C11)AA         26         26         26           DARGYTCANDTERAPECT (C11)AA         26         26         26           DARGYTCANDTERAPECT (C11)AAASS1A         26         26         26           DARGYTCANDTERAPET (C11)AAASS1A         26         26         26           DARGYTCANDARDET (C11)AAAASS1A         26         26         26           DARGYTCANDARDET (C11)AAAASS1A         26         26         26           DARGYTCANDARDET (C11)AAAASS1A         26         26         26           DARGYTCANDTARDET (C11)AAAASS1A         26         26         26           DARGYTCANDTARDET (C11)AAAASS1A         26         26         26           DARGYTCANDTARDET (			21			
PARSYLICAT IGAT BARAPERA.         24         24         24         24           PARSYLICAT IGAT BARAPERA.         24         27         27         27           PARSYLICAT IGAT BARAPERA (CIT JUA AGAST ICATION AGAST ICATI						
DARASCHICCT LOLADAPOSTICATION         20						
PARLEGY PARTY INCLUING PARTY INCLUING         20						
PARLEYT-COTJEAPORTSA         29         20         20         20           PARLEYT-COTJEAPORTSA         29         20         20         20           PARLEYT-COTJEAPORTSA         20         20         20         20           PARLEYT-COTJEAPORTSA         30         31         31         31         31           PARLEYT-COTJEAPORTSA         30         31         31         31         31         31           PARLEYT-COTJEAPORTSA         30         31						
DALESCYTANDITICLI JAPACESTA         20         20         20         20           DALESCYTANDITICLI JAPACESTA         20         20         20         20           DALESCYTICLI JAPACESTACIA         20         20         20         20           DALESCYTICCI JAPACESTACICIA         20         20         20         20         20           DALESCYTICCI JAPACESTACICIA         20         20         20         20         20         20           DALESCYTICCI JAPACESTACICIA         20         20         20         20         20         20         20         20         20        20         2	PAHGVT*(C1T1)SAPDT*(C1T1)RPAPGST*(C1T) PAHGVT*(C1T2)SAPDTRPAPGSTA					
DALESCYLANDTHEPAROSTICCT2DA         30         30         30         30           DALESCYLANDTHEPAROSTICCT2DA         32         33         30         30           DALESCYLANDTECLIZAPOTENCICT2DA         30         30         30         30           DALESCYLANDTECLIZAPOTENCICT2DA         30         30         30         30           DALESCYLANDTECLIZAPOTENCICT2DA         30         30         30         30         30           DALESCYLANDTECLIZAPOTENCICT2DA         40         40         30         30         30           DALESCYLANDTECLIZAPOTENCICT2DA         40 <t< th=""><th>PAHGVTSAPDT*(C1T2)BPAPGSTA</th><th>-</th><th></th><th></th><th></th><th></th></t<>	PAHGVTSAPDT*(C1T2)BPAPGSTA	-				
PARGOTACUTICATIBRANESSIA         33         31         31         31           PARGOTACUTICATIBRANESSIA         33         31         31         31           PARGOTACUTICATIBRANESSIA         33         33         31         31           PARGOTACUTICATIBRANESSIA         33         33         31         31           PARGOTACUTICATIBRANESSIA         36         31         31         31           PARGOTACUTICATIBRANESSIA         36         31         31         31           PARGOTACUTICATIBRANESSIA         36         31         31         31           PARGOTACUTICATIBRANESSIA         43         43         31         31           PARGOTACUTICATIBRANESIA         43         43         31         31           PARGOTACUTICATIBRANESIA         32		30				
DALGOYTANDTICLINAPARAGESTICATIA       38       38       38       38         DALGOYTANDTICATIBRANGSTICATIA       48       44       44       44         DALGOYTANDTICATIBRANGSTICATIA       48       44       44       44         DALGOYTANDTICATIBRANGSTICATIA       48       46       44       44         DALGOYTANDTICATIBRANGSTICATIA       46       46       44       44         DALGOYTANDTICATIBRANGSTICATIA       46       46       46       46         DALGOYTANDTICATIBRANGSTICATIA       46       46       46       46						
DARAGY (C11)AAPDIPARASIX.C11       34       34       34         DARAGY (C11)AAPDIPARASIX.21       35       35         DARAGY (C11)AAPDIPARASIX.21       36       36         DARAGY (C11)AAPDIPARASIX.21       36       36         DARAGY (C11)AAPDIPARASIX.21       36       36         DARAGY (C11)AAPDIP (C11)RPARGSIX.21       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX.21       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX.21       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       47       47         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       47       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       46       46         DARAGY (C11)AAPDIP (C11)RPARGSIX (C11)A       47       46 <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
PARSOTYLCT1]SARDTRAPSSTA         36         35         35         36           PARSOTRADIC CT1]BRARSSTA         36         35         36         36           PARSOTRADIC CT1]BRARSSTA         36         35         36         36           PARSOTRADIC CT1]BRARSSTA         36         36         36         36           PARSOTRECTISATION CT121A         36         36         36         36           PARSOTRECTISATION CT21A         36         36         36         36         36           PARSOTRECTISATION CT21A         36         36         36         36         36         36         36         36         36         36         36         36         36         36         <					-	
PARLOY TCATLINGANESIA         36         36         36           PARLOY TCATLINGANESIA         38         36         37           PARLOY TCATLINGANESIA         38         38         38         38           PARLOY TCATLINGANESIA         38         38         38         38           PARLOY TCATLINGANESIA         48         48         38         38           PARLOY TCATLINGANESIA         50         50         50         50		35				
PARGYTICCT13BAPDTIFAACSTICCT10         38         38         38         38           PARGYTICCT13BAPDTIFAACSTICCT10         38         38         38         38           PARGYTICCT13BAPDTIFACT3IPAACSTICCT10         48         48         48         38           PARGYTICCT13BAPDTIFACT3IPAACSTICCT10         44         44         44         44           PARGYTICCT13BAPDTIFACT3IPAACSTICCT10         44         44         44         44           PARGYTICCT13BAPDTIFACT3IPAACSTICCT10A         44         44         44         44           PARGYTICCT13BAPDTICCT3IPAACSTICCT10A         44         44         44         44           PARGYTICCT3BAPDTIFACT3IPAACSTICCT3DA         44         44         44         44           PARGYTICCT3BAPDTIFACT3IPAACSTICCT3DA         45         45         44         44           PARGYTICCT3BAPDTIFACT3PIRAACSTICCT3DA         50         50         50         50         50           PARGYTICCT3BAPDTIFACT3PIRAACSTICCT3DA         50						
PARSOTTICCT 10 APPLIFICATOST (CCT 10)         39         30         30         30           PARSOTTICCT 10 SAPPLIFICATOSTA         40         40         40         40           PARSOTTICCT 10 SAPPLIFICATOSTA         44         40         40         40           PARSOTTICCT 10 SAPPLIFICATOSTICCT 10A         44         40         40         40           PARSOTTICCT 10 SAPPLIFICATOSTICCT 10A         44         40         40         40           PARSOTTICCT 10 SAPPLIFICATOSTICCT 10A         40         40         40         40           PARSOTTICCT 10 SAPPLIFICATOSTICCT 10A         50         50         40         40         40           PARSOTTICCT 10 SAPPLIFICATOSTICCT 10A         50<						
DARGYTAPD''(CT) RAPGST*(CT)]A         40         40         40           DARGYTAPD''(CT) RAPGST*(CT)]A         41         41         41           DARGYTAPD''(CT) RAPGSTA         44         45         44           DARGYTAPD''(CT) RAPGSTA         44         45         44           DARGYTAPD''(CT) RAPGSTA         44         45         44           DARGYTAPD''(CT) RAPGSTA         44         45         46           DARGYTAPD''(CT) RAPGSTA         44         45         46           DARGYTAPD''(CT) RAPGSTA         45         53         53           DARGYTAPD''(CT) RAPGSTA         53         54         54           DARGYTAPD''(CT) RAPGSTA         55         54         56           DARGYTAPD''(CT) RAPGSTA         55         54         56           DARGYTAPD''(CT) RAPGSTA         56         56         56           DARGYTAPD''(CT) RAPGSTA         56         56         56           DARGYTAPD''(CT) RAPGSTA         56         56         56						
PAREGYT*(C71)AAPD/T*C71/JRAPAGST*(C71)         44         41         42         43           PAREGYT*(C71)AAPD/T*C71/JRAPAGST*(C71)A         44         44         44         44           PAREGYT*(C71)AAPD/T*C71/JRAPAGST*(C71)A         45         45         46         46           PAREGYT*(C71)AAPD/T*C71/JRAPAGST*(C71)A         55         56         46         46           PAREGYT*(C71)AAPD/T*C71/JRAPAGST*(C71)A         58         58         56         56         56         56         56         56         56         56         56         56         56						
PAHGVT:(C21)BAPOTRPARSIA       42       42       43         PAHGVT:(C21)BAPOTRPARSIA       43       43         PAHGVT:(C21)BAPOTRPARSIA       44         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       54         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       54         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       54         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       54         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       54         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       54         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       55         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       56         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       56         PAHGVT:(C21)BAPOTRPARSIT:(C21)A       56         PAHGVT:(C21)BAPOTRSIT:(C21)A       56         PAHGVT:(C21)BAPOTST:(C21)A       56         PAHGVT:(C21)BAPOTST:(C21)A       56         PAHGVT:(C21)BAPOTST:(C21)A	PAHGVT*(C2T1)SAPDT*(C2T1)RPAPGST*(C2T	41				
DALGOVICAPDTEPAROSIT(CIT)IA         44         44         44           DALGOVICAPDIAGADIC(CIT)IRAPAGSIT(CIT)IA         45         45           DALGOVICAPDIC(CIT)IRAPAGSIT(CIT)IA         46         46           DALGOVICAPDIC(CIT)IRAPAGSIT(CIT)IA         47         46           DALGOVICAPDIC(CIT)IRAPAGSIT(CIT)IA         47         46           DALGOVICAPDICAPDIPAROSIT(CIT)IA         47         46           DALGOVICAPDIPAROSIT(CIT)IRAPAGSIT(CIT)IA         43         46           DALGOVICAPDIPAROSIT(CIT)IA         53         53           DALGOVICAPDIPAROSIT(CIT)IA         53         53           DALGOVICAPDIPAROSIT(CIT)IA         53         53           DALGOVICAPDRAFONICATIIA         53         53           DALGOVICAPDRAFONICATIIA         53         54           DALGOVICAPDRAFONICATIIA         57         57           DALGOVICAPDRAFONICATIIA         57         57           DALGOVICAPDRAFON			42	42		
DAHGOT*(C2T)ISAPDTRAPASTIC         45         45         46           DAHGOT*(C2T)ISAPDTRAPASTIC         46         46         46           DAHGOT*(C2T)ISAPDTRAPASTIC         47         47         47           DAHGOT*(C2T)ISAPDTRAPASTIC         47         47         47           DAHGOT*(C2T)ISAPDTRAPASTIC         48         47         47           DAHGOT*(C2T)ISAPDTRAPASTIC         48         47         47           DAHGOT*(C2T)ISAPDTRAPASTIC         48         47         47           DAHGOT*(C2T)ISAPDT(C2T)IAPARSTIC         48         51         51           DAHGOT*(C2T)ISAPDT(C2T)IAPARSTIC/C2T)IA         54         55         51           DAHGOT*(C2T)ISAPDT(C2T)IAPARSTIC/C2T)IA         55         55         56         56           DAHGOT*(C2T)ISAPDT(C2T)IAPARSTIC/C2T)IA         55         56         56         56         56           DAHGOT*(C2T)ISAPDT(C2T)IAPARSTIC/C2T)IA         56						
DALGOVIT (C27) ISAPDITEAPAGST (C27) A         46         46         47           DALGOVIT (C27) ISAPDITEAPAGST (C37) A         47         47         47           DALGOVIT (C27) ISAPDITEAPAGST (C37) A         47         47         47           DALGOVIT (C27) ISAPDITEAPAGST (C37) A         47         47         47           DALGOVIT (C27) ISAPDITEAPAGST (C37) A         51         51         51           DALGOVIT (C37) IRAPAGST (C37) A         51         51         51           DALGOVIT (C37) IRAPAGST (C37) A         53         55         51           DALGOVIT (C37) IRAPAGST (C37) A         53         55         56           DALGOVIT (C37) IRAPAGST (C37) A         59         55         56           DALGOVIT (C37) IRAPAGST (C37) A         59         56         56           DALGOVIT (C37) IRAPAGST (C37) A         57         57 </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
PAHGVTSAPDT*CC11JRAAPGST*CC2T         47         47         47         47           PAHGVTSAPDT*CC11JRAAPGST*CC2T         48         48         44           PAHGVTSAPDT*CC12JRAAPGST*CC1T         40         48           PAHGVTSAPDT*CC12JRAAPGSTA         50         40           PAHGVTSAPDT*CC12JRAAPGSTA         50         50           PAHGVTSAPDT*CC12JRAAPGSTA         52         53           PAHGVTSAPDT*CC12JRAAPGST*CC17         54         53           PAHGVTSAPDT*CC12JRAAPGST*CC17         54         55           PAHGVTSAPDT*CC12JRAPGSTA         59         55           PAHGVTSAPDT*CC12JRAPGSTA         59         55           PAHGVTSAPDT*CC12JRAPGST*CC12JA         60         60           PAHGVTSAPDT*CC12JRAPGST*CC12JA         61         61           PAHGVTSAPDT*CC12JRAPGST*CC12JA         62         6	PAHGVT*(C2T1)SAPDTRPAPGST*(C2T1)A					
DAHGVT*C2T]25APDTRPARGSTA         49         40           PAHGVTSAPDTRAPGST*(C2T)2A         51         51           PAHGVTSAPDTRAPGST*(C2T)2A         53         51           PAHGVTSAPDTRAPGST*(C2T)2A         53         51           PAHGVTSAPDT*(C2T)2PAPGST*(C2T)2A         53         55           PAHGVT*(C2T)2APDTCST*(C2T)2A         54         55           PAHGVT*(C2T)2APDTRAPGST*(C2T)2A         56         55           PAHGVT*(C2T)2APDTRAPGST*(C2T)2A         60         60           PAHGVT*(C2T)2APDTRAPGST*(C1T)2A         60         66           PAHGVT*(C1T)2APDTRAPGST*(C1T)A         66         66           PAHGVT*(C1T)2APDTRAPGST*(C1T)A         66         66           PAHGVT*(C1T)2APDTRAPGST*(C1T)A)         68         66           PAHGVT*(C1T)2APDTRAPGST*(C1T)A)         68         66           PAHGVT*(C1T)2APDT*(C1T)2APDTRAPGST*(C1T)A)         77         77	PAHGVTSAPDT*(C2T1)RPAPGST*(C2T1)A		47			
DAHGUTSAPDT*(C27)RPARGST         50         50         50           DAHGUTSAPDT*(C27)RPARGST*(C27)A         51         52         51           DAHGUTSAPDT*(C27)RPARGST*(C27)A         53         52         51           DAHGUTSAPDT*(C27)RPARGST*(C27)A         53         52         51           DAHGUTSAPDT*(C27)RPARGST*(C27)A         53         55         51           DAHGUTSAPDT*(C27)RPARGST*(C27)A         56         55         51           DAHGUTSAPDT*(C27)RPARGST*(C27)A         59         55         51           DAHGUTSAPDT*(C27)RPARGST*(C37)A         59         55         51           DAHGUTSAPDT*(C27)RPARGST*(C37)A         59         55         51           DAHGUT*(C37)RPARGST*(C37)A         59         56         51           DAHGUT*(C37)RPARGST*(C37)A         60         60         61           DAHGUT*(C37)RPARGST*(C37)A         60         60         61           DAHGUT*(C37)RPARGST*(C37)A         60         60         61           DAHGUT*(C37)RPARGST*(C37)A         61         62         61           DAHGUT*(C37)RPARGST*(C37)A         61         62         61           DAHGUT*(C37)RPARGST*(C37)A         61         62         61           DAHGUT*(C37)RPARG						
PAHGVTSAPDTRPAPGST*(C2T2)A       51       51       51         PAHGVTC2T2]AAPDTRPAPGST*(C2T2)A       52       53       53         PAHGVTCC12]AAPDTRPAPGST*(C2T2)A       53       53       53         PAHGVTCC12]AAPDTGTC2T2]PAPGSTA       56       56       56         PAHGVTCCT2]AAPDTGTC2T2]PAPGSTA       57       57       57         PAHGVTCCT2]AAPDTGTC2T2]PAPGSTA       57       57       57         PAHGVTCCT2]AAPDTGTC2T2]PAPGSTA       57       57       57         PAHGVTCCT2]AAPDTGTC2T2]PAPGSTA       57       57       57         PAHGVTCCT2]AAPDTGTC2T2]PAPGSTCCT2       56       56       56         PAHGVTCT2]SAPDTTGTAPGST       57       57       57         PAHGVTSAPDTCCT2]PAPGSTCCT2]A       60       60       60         PAHGVTSAPDTCTCT3]PAPGSTCCT3       60       60       61         PAHGVTSAPDTCTCT3]PAPGSTCCT3       60       62       62         PAHGVTSAPDTCTA]SAPDTTGAPGSTCA       60       66       66         PAHGVTSAPDTTRPAPGSTCA       60       67       64         PAHGVTSAPDTTRPAPGSTCA       70       70       71         PAHGVTSAPDTCCT3]PAPGSTA       72       72       71         PAHGVTSAPDTCCT3]PAPGSTA						
PAHEQU'T(CIT2)SAPDTRAPGST(CIT2)APAG						
PAHAGYTSAPDT*CCT2]PAPGST*CCT2] A       S4       S4       S4         PAHAGYT*CCT2]PAPGST*CCT2]       S5       S5       S5         PAHAGYT*CCT2]PAPAGSTA       S5       S5       S6         PAHAGYT*CCT2]PAPAGSTA       S5       S5       S6         PAHAGYT*CCT2]PAPAGSTA       S5       S7       S6         PAHAGYT*CCT2]PAPAGST*CCT2]A       G0       G0       S6         PAHAGYT*CCT2]PAPAGST*CCT2]A       G1       S6       S6         PAHAGYT*CCT2]PAPAGST*CCT2]A       G1       G1       S6         PAHAGYT*CCT2]PAPAGST*CCT2]A       G1       G1       G1         PAHAGYT*CATSPAPGST*CCT2]PAPAGST*CCT2]A       G2       G2       G2         PAHAGYT*CATSPAPGST*CT0*A]A       G6       G6       G4         PAHAGYT*CATSPAPGST*CT0*A]A       G6       G6       G4         PAHAGYT*CATSPAPGST*CT0*A]A       G6       G6       G6         PAHAGYT*CATSPAPGST*CT0*A]A       G6       G6       G6       G6         PAHAGYT*CATSPAPGST*CT0*A]A       G6       G6       G6       G6       G6         PAHAGYT*CATSPAPGST*CCT1]A       G7	PAHGVT*(C2T2)SAPDT*(C2T2)RPAPGSTA					
PAHAGVT*(C212)SAPDT*RAPGSTA       55       55         PAHAGVT*(C212)APARAPGSTA       55       57         PAHAGVTSAPDT*(C212)APARGSTA       57       57         PAHAGVTSAPDT*(C212)APARGSTA       50       57         PAHAGVTSAPDT*(C212)APARGST*(C212)A       60       59         PAHAGVTSAPDT*(C212)APARGST*(C212)A       60       59         PAHAGVTSAPDT*(C212)APARGST*(C212)A       60       61         PAHAGVTSAPDT*(C12)APARGST*(C212)A       60       61         PAHAGVTSAPDT*(C12)APARGST*(C212)A       60       63         PAHAGVTSAPDT*(C10)SAPCTSTAPAGST*(C212)A       66       64         PAHAGVTSAPDT*(T0)SAPARGSTA       66       65         PAHAGVTSAPDT*(T0)SAPARGSTA       66       66         PAHAGVTSAPDT*(T0)SAPARGST*(T0)A       66       66         PAHAGVTSAPDT*(T0)SAPARGST*(T0)A       66       66         PAHAGVTSAPDT*(T0)SAPARGSTA       70       70         PAHAGVTSAPDT*(T0)SAPARGSTA       72       71         PAHAGVTSAPDT*(T0)SAPARGSTA       72       72         PAHAGVTSAPDT*(T0)SAPARGSTA       72       72         PAHAGVTSAPDT*(C31)SAPARGSTA       72       72         PAHAGVTSAPDT*(C31)SAPARGSTA       72       72         PAHAGVTSAPD				<b>53</b>		
PARAGVT*(C212)SAPDT*(Z212)RAPGSTA         56         57           PARAGVTSAPDT*(Z212)RAPGST*(C212)A         58         58         58         58           PARAGVTSAPDT*(Z212)RAPGST*(C212)A         60         59         57           PARAGVTSAPDT*(Z212)RAPGST*(C212)A         60         59         57           PARAGVTSAPDT*(C212)RAPGST*(C212)A         60         60         60           PARAGVTSAPDT*(C212)RAPGST*(C212)A         60         60         60           PARAGVTSAPDT*(C212)RAPGST*(C212)A         60         60         60           PARAGVTSAPDT*(T0-a)RAPGST*(C212)A         66         66         66           PARAGVT*(T0-a)SAPDT*RAPGST*(T0-a)A         66         66         66           PARAGVT*(T0-a)SAPDT*RAPGST*(T0-a)A         66         66         66           PARAGVT*(T0-a)SAPDT*RAPGST*(T0-a)A         66         66         66           PARAGVT*(C311)SAPDT*(C311)RAPGST*(T0-a)A         66         66         66           PARAGVT*(C311)SAPDT*(C311)RAPGST*(C1-a)A         70         70         70           PARAGVT*(C311)SAPDT*(C311)RAPGST*(C317)A         72         75         76           PARAGVT*(C131)SAPDT*(C311)RAPGST*(C317)A         73         74         75           PARAGVT*(C131)SAPDT*(C311)RAPGST*(C317)A </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
PARAGVTSAPDT*(C212)         APAGVTSAPDT*(C212)         APAGVTSAPDT*(C12)         APAGVTSAPDT*(C21)         APAGVTSAPDT*						
PAHGVTSAPDTRPARGST(C212)A       58       58       58         PAHGVTS(C212)SAPDTRPARGST(C212)A       60       60       60         PAHGVTS(C212)SAPDTRPARGST(C212)A       60       60       60         PAHGVTS(C212)SAPDTRPARGST(C212)A       60       60       60         PAHGVTSAPDTRPARGSTA       62       60       60         PAHGVTSAPDTRPARGSTA       63       63       64         PAHGVTSAPDTRPARGST(Tn-a)A       66       65       66         PAHGVTSAPDTRPARGST(Tn-a)A       66       66       66         PAHGVTSAPDTRPARGST(Tn-a)A       67       67       66         PAHGVTSAPDTRPARGST(Tn-a)A       67       67       66         PAHGVTSAPDTRPARGST(Tn-a)A       67       67       67         PAHGVTSAPDTRPARGST(TC31)A       70       70       71         PAHGVTSAPDTRPARGST(C31)A       70       72       72       73         PAHGVTSAPDTRPARGST(C31)A       77       72       73       74       74         PAHGVTSAPDTRPARGST(C31)A       77       75       74       74       74       74       74       74       74       74       74       74       74       74       74       74       74       74						
PAHAGVT:(C212)APDTRAPGST:(C212)A       60       60         PAHGVT:(C212)APAPGST:(C212)A       61       61         PAHGVT:(C212)APAPGST:(C212)A       62       62         PAHGVT:(C12)APAPGSTA       64       63         PAHGVT:(D)SAPDT:(Tn-a)RPAPGSTA       66       66         PAHGVT:(Tn-a)SAPDTRAPGST:(Tn-a)A       66       66         PAHGVT:(Tn-a)RPAPGST:(Tn-a)A       67       68         PAHGVT:(Ta)SAPDT*(C31)A       77       71       71         PAHGVT:(C31)SAPDTRAPGST:(C31)A       77       72       72         PAHGVT:(C31)SAPDTRAPGST:(C31)A       77       75       74         PAHGVT:(C31)SAPDTRAPGST:(C31)A       77       75       76         PAHGVT:(C31)SAPDTRAPGST:(C31)A       77       76       76         PAHGVT:(C31)SAPDTRAPGST:(C31)A       77       76       76         PAHGVT:(C31)SAPDTRAPGST:(C31)A       78       78       78         PAHGVT:(C31)SAPDTRAPGST:(C31)A       78       78       78 <th>PAHGVTSAPDTRPAPGST*(C2T2)A</th> <th></th> <th>58</th> <th></th> <th></th> <th></th>	PAHGVTSAPDTRPAPGST*(C2T2)A		58			
PAHAGVTSAPDT*(212)RAPGST*(C2T2)A       61       61       61         PAHGVT*(C12)SAPDT*(C2T2)RAPGSTA       63       63       63         PAHGVT*(Tr-a)SAPDT*(Tr-a)RAPGSTA       64       64       64         PAHGVTSAPDT*(Tr-a)RAPGSTA       66       65       66         PAHGVTSAPDT*(Tr-a)RAPGST*(Tr-a)A       66       66       66         PAHGVTSAPDT*(Tr-a)RAPGST*(Tr-a)A       66       66       66         PAHGVT*(Tr-a)SAPDT*(Tr-a)RAPGST*(Tr-a)A       66       66       66         PAHGVT*(Tr-a)SAPDT*(Tr-a)RAPGST*(Tr-a)A       66       66       66         PAHGVT*(G13)SAPDT*(Tr-a)RAPGST*(Tr-a)A       70       70       70         PAHGVT*(G13)SAPDT*(G13)ARAPGST*(C3T)A       77       71       71         PAHGVT*(G13)SAPDT*(G13)ARAPGST*(C3T)A       77       72       72       73         PAHGVT*(G13)SAPDT*(G13)ARAPGST*(C3T)A       77       77       77       77         PAHGVT*(G13)SAPDT*(G13)ARAPGST*(C3T)A       77       77       77       77         PAHGVT*(G13)SAPDT*(G13)ARAPGST*(C3T)A       78       78       78       78       78       78       78       78       78       78       78       78       78       78       78       78       78       <						
PAHSUT*(C212)SAPDT*(C212)PAPGSTA       63       63         PAHSUT*(C112)SAPDT*(Tn-a)SAPDTRAPGSTA       64       64         PAHSUT*(Tn-a)SAPDTRAPGSTA       64       64         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       65       65         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       66       66         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       66       66         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       66       66         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       67       67         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       66       66         PAHSUT*(Tn-a)SAPDT*(Tn-a)A       67       67         PAHSUT*(C21)SAPDT*(C31)APAPGSTA       70       60       67         PAHSUT*(C31)SAPDT*(C31)A       72       72       71         PAHSUT*(C31)SAPDT*(C31)A       72       73       74       74         PAHSUT*(C31)SAPDT*(C31)A       75       75       75       75         PAHSUT*(C31)SAPDT*(C31)APAPGSTA       78       76       77       74         PAHSUT*(C31)SAPDT*(C31)APAPGST*(C31)A       78       76       76       77         PAHSUT*(C31)SAPDT*(C31)APAPGSTA       78       76       77       77       77         PAHSUT*(C31)SAPDT*(C31)AA       78       76       77       7						
PAHGVT*(Tr.a)SAPDT*(Tr.a)RAPGSTA       63       63       63         PAHGVTSAPDT*(Tr.a)RAPGST*(Tr.a)A       65       65       65         PAHGVT*(Tr.a)SAPDT*(Tr.a)RAPGSTA       66       66         PAHGVT*(Tr.a)SAPDT*(Tr.a)A       67       67         PAHGVT*(Tr.a)SAPDT*(Tr.a)A       67       67         PAHGVT*(Tr.a)SAPDT*(Tr.a)A       67       67         PAHGVT*(TT.a)SAPDT*(DS)APDT*(CS)APDT*(DS)APDT*(CS)APDT*(DS)APDT*(CS)APAPGST*(CS)A       73         PAHGVT*(CS)APDT*(DS)APAPGST*(CS)A       73       74       74         PAHGVT*(CS)APDT*(DS)APAPGST*(CS)A       73       74       74         PAHGVT*(CS)APAPGST*(CS)A       75       76       76         PAHGVT*(CS)APAPGST*(CS)A       77       76       77         PAHGVT*(CS)APAPGST*(CS)A       78       78       78         PAHGVT*(CS)A						
PAHGVT:APDTRAPGST*(In-a)A       65       65       66         PAHGVT:(In-a)SAPDTRPAPGST*(In-a)A       67       67       67         PAHGVT:(In-a)SAPDTRPAPGST*(In-a)A       67       67       67         PAHGVT:APDT*(In-a)BRAPGST*(In-a)A       68       68       68         PAHGVT:APDT*(In-a)BRAPGST*(In-a)A       67       69       69         PAHGVT:APDT*(In-a)BRAPGST*(In-a)A       67       70       70         PAHGVT:APDT*(In-a)BRAPGST*(C3T1)A       72       71       70         PAHGVT:CATI)SAPDTRPAGST*(C3T1)A       72       71       71         PAHGVT:CATI)SAPDT*(C3T1)BRAPGST*(C3T1)A       74       73       74         PAHGVT:CATISAPDT*(C3T1)BRAPGST*(C3T1)A       74       74       76         PAHGVT:CATISAPDT*(C3T2)RAPGSTA       77       77       77       77         PAHGVT:CATISAPDT*(C3T2)RAPGSTA       78       78       78       78         PAHGVT:CATISAPDT*(C3T2)RPARGST       77       77       77       77       77         PAHGVT:CATISAPDTRAPGST       77       77       77       77       77       77       77       77       77       77       77       77       77       77       77       77       77       78       78<						
PAHGVT*(In-a)SAPDT*(In-a)RPAPGST*(In-a)A       66       66       67       66         PAHGVT*(In-a)SAPDTAPGST*(In-a)A       68       68       68       68         PAHGVT*(In-a)SAPDTSS*(In-a)A       68       68       69       69         PAHGVT*(In-a)SAPDTSS*(In-a)A       68       68       69       69         PAHGVT*(In-a)SAPDT*(	PAHGVTSAPDT*(Tn-a)RPAPGSTA			64 H		
PAHAGVT*(Tn-a)SAPDT*(Tn-a)RPAPGST*(Tn-a)A       67       67         PAHAGVT*APDT*(Tn-a)RPAPGST*(Tn-a)       69       69       67         PAHAGVT*(Tn-a)SAPDT*(Tn-a)RPAPGST*(Tn-a)       69       69       69         PAHAGVT*(Tn-a)SAPDT*(Tn-a)RPAPGST*(Tn-a)       69       69       69         PAHAGVT*(C3T1)SAPDT*(C3T1)RPAPGSTA       70       70       70         PAHAGVT*(C3T1)SAPDT*(C3T1)RPAPGSTA       73       73       74         PAHAGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T1)A       74       74       74         PAHAGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T1)A       74       74       74         PAHAGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T1)A       76       76       76         PAHAGVT*(C3T1)SAPDT*(C3T2)RPAPGST*(C3T2)A       80       76       77         PAHAGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       80       80       80         PAHAGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       80       81         PAHAGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       80       81         PAHAGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       82       81         PAHAGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       82       81         PAHAGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       82       81 </th <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
PAHGVTSAPD1*(1r-a)SAPDTS1*(1r-a)A       68       68       68         PAHGVT*(C3T1)SAPDTRPAPGST       70       70       70         PAHGVT*(C3T1)SAPDTRPAPGSTA       71       71       71         PAHGVTSAPDT*(C3T1)RPAPGSTA       71       71       71         PAHGVTSAPDTRPAPGST*(C3T1)A       72       72       73         PAHGVTSAPDTRPAPGST*(C3T1)A       72       73       74         PAHGVTSAPDTRPAPGST*(C3T1)A       74       74       74         PAHGVTSAPDTRPAPGST*(C3T1)A       75       75       76         PAHGVTSAPDT*(C3T1)RPAPGST*(C3T1)A       76       76       76         PAHGVTSAPDT*(C3T2)RPAPGST*(C3T2)A       78       78       78         PAHGVTSAPDT*(C3T2)RPAPGST*(C3T2)A       80       80       81       81         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       80       80       81       81       81         PAHGVTSAPDT*(C3T2)RPAPGST*(C3T2)A       81 <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
PAHGVT*(C311)SAPDTRPAPOSTA       70       70       70       70         PAHGVT*APDTRPAPOST*(C311)A       72       72       71       71         PAHGVT*(C311)SAPDTRPAPOST*(C311)A       72       72       73       73         PAHGVT*(C311)SAPDTRPAPOST*(C311)A       74       74       74       74         PAHGVT*(C311)SAPDTRPAPOST*(C311)A       75       75       76       76       76         PAHGVT*(C312)SAPDT*(C312)RPAPOSTA       77       77       77       77       77       77         PAHGVT*(C312)SAPDT*(C312)RPAPOSTA       78       78       77       78       78       78       78       78       78       77       77       77       77       77       77       77       77       77       77       78       78       78       78       78       78       78       78       78       78 <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>						
PAHQVTSAPDT*(C3T1)APPCSTA       71       71       71       71         PAHGVTSAPDT*(C3T1)APDT*(C3T1)APPCST       73       73       74       73         PAHGVT*(C3T1)SAPDT*(C3T1)APPCST*(C3T1)A       73       73       74       74         PAHGVT*(C3T1)SAPDT*(C3T1)APPCST*(C3T1)A       75       76       76       76         PAHGVT*(C3T1)SAPDT*(C3T1)APPCST*(C3T1)A       77       77       76       76         PAHGVT*(C3T1)SAPDT*(C3T1)APPCST*(C3T2)A       78       78       76       76         PAHGVT*(C3T2)SAPDTRPAPCST       77       77       76       77         PAHGVT*(C3T2)SAPDTRPAPCST       78       78       78       79         PAHGVT*(C3T2)SAPDTRPAPCST*(C3T2)A       81       81       81       81       81         PAHGVT*(C3T2)SAPDTRPAPCST*(C3T2)A       82       82       82       82       82       82       83       84						
PAHGVTSAPDTRPAPGST*(C3T1)A       72       72       72         PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGSTA       73       74       73         PAHGVT*(C3T1)SAPDTRPAPGST*(C3T1)A       74       73       74         PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T)A       75       75       75         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA       77       76       76         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA       77       76       76         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA       77       77       78         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA       78       77       78         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       80       80       80         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       81       77         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       83       81       81       77         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST       83       82       81       81       77         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST       83       81       81       77 <t< th=""><th>PAHGVT*(C3T1)SAPDTRPAPGSTA PAHGVTSAPDT*(C3T1)BPAPGSTA</th><th></th><th></th><th></th><th></th><th></th></t<>	PAHGVT*(C3T1)SAPDTRPAPGSTA PAHGVTSAPDT*(C3T1)BPAPGSTA					
PAHGV1*(C311)SAPDT*(C311)A       73       73       74         PAHGV1*(C311)SAPDT*(C311)A       74       74       74         PAHGV1*(C311)SAPDT*(C311)APAPGST*(C311)A       75       75       76         PAHGV1*(C311)SAPDT*(C311)RAPAGST*(C31)A       77       77       76         PAHGV1*(C312)SAPDT*(C312)RAPAGSTA       77       77       77         PAHGV1*(C312)SAPDT*(C312)RAPAGSTA       78       78       78         PAHGV1*(C312)SAPDT*(C312)RAPAGSTA       78       78       78         PAHGV1*(C312)SAPDT*(C312)RAPAGST*(C312)A       81       81       1         PAHGV1*(C312)SAPDT*(C312)RAPAGST*(C312)A       82       81       1         PAHGV1*(C312)SAPDT*(C312)RAPAGST*(C312)A       82       83       81       1         PAHGV1*(C312)SAPDT*(C312)RAPAGST*(C312)A       82       83       83       1         PAHGV1*(C411)SAPDT*(C411)RAPAGSTA       83       84       1       84       1         PAHGV1*(C411)SAPDT*(C411)RAPAGST*(C41)A       88       86       87       1       10         PAHGV1*(C411)SAPDT*(C411)RAPAGST*(C41)A       88       86       1       10       1       1         PAHGV1*SAPDT*(C41)RAPAGST*(C41)A       89       89       1       1 <t< th=""><th></th><th></th><th></th><th></th><th></th><th></th></t<>						
PAHGVTSAPDT*(C3T1)RPAPGST*(C3T)A       75       75       76         PAHGVT*(C3T1)SAPDT*(C3T1)RAPGST*(C3T)A       77       76       76         PAHGVT*(C3T2)SAPDT*RPAPGSTA       77       76       76         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA       78       78       78         PAHGVTSAPDT*(C3T2)RPAPGSTA       80       80       78       78         PAHGVT*(C3T2)SAPDT*(C3T2)A       81       81       81       81         PAHGVT*(C3T2)SAPDT*(C3T2)A       81       81       81       81         PAHGVT*(C3T2)SAPDT*(C3T2)A       81       81       81       81         PAHGVT*(C3T2)SAPDT*(C3T2)A       82       82       83       83       84         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C3T       83       83       83       84       84       84       84       84       84       84       84       84       84       85       85       85       85       85       85       85       85       85       85       86       86       86       86       86       86       86       86       86       86       86       86       86       86       86       86       86       86       87       87       87       8	PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGSTA					
PAHGVT*(C3T1)SAPDT*(C3T1)RPAPGST*(C3T       76       76       76         PAHGVT*(C3T2)SAPDT*RPAPGSTA       77       77       77         PAHGVT*(C3T2)SAPDT*(C3T2)A       79       71       78         PAHGVT*(C3T2)SAPDT*(C3T2)A       79       71       78         PAHGVT*(C3T2)SAPDT*(C3T2)A       80       80       79         PAHGVT*(C3T2)SAPDT*(C3T2)A       80       80       81         PAHGVT*(C3T2)SAPDT*(C3T2)A       81       81       81         PAHGVT*(C3T2)SAPDT*(C3T2)A       82       82       82         PAHGVT*(C3T2)SAPDT*(C3T2)ARPAPGST*(C3T2)A       83       83       84         PAHGVT*(C3T2)SAPDT*(C4T1)RPAPGST*(C3T1)A       84       84       84         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       86       86       87         PAHGVT*C4C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       88       88       87         PAHGVT*SAPDT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       88       87       87         PAHGVT*SAPDT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89       91       91         PAHGVT*SAPDT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       93       91       91         PAHGVT*SAPDT*(C4T2)SAPDT*(C4T2)SAPDT*(C4T2)A       93       91       91         PAHGVT*SAPDT*(C4T2)SAPDT*(C				<b>74</b>		
$\begin{array}{c c c c c c c c c c c c c c c c c c c $						
PAHGVTSAPDT*(C3T2)RPAPGSTA       78       78       78         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       80       79       79         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       81       81       81       81         PAHGVT*(C3T2)SAPDT*PAPGST*(C3T2)A       81       81       81       81       81         PAHGVT*(C3T2)SAPDT*PAPGST*(C3T2)A       81       81       81       81       81       81         PAHGVT*(C3T2)SAPDT*PAPGST*(C3T2)A       82       83       83       81 <t< th=""><th></th><th></th><th></th><th></th><th></th><th></th></t<>						
PAHGVTSAPDTRPAPGST*(C3T2)A       79       79         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGSTA       80       80         PAHGVT*(C3T2)SAPDT*(C3T2)A       81         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       82         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       82         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T2)A       84         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       84         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       85         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       86         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       87         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       88         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89         PAHGVT*SAPDT*(C4T2)RPAPGSTA       91         PAHGVT*SAPDT*(C4T2)RPAPGSTA       92         PAHGVT*SAPDTRPAPGST*(C4T2)A       93         PAHGVT*SAPDT*(C4T2)RPAPGSTA       92         PAHGVT*SAPDT*(C4T2)RPAPGST*(C4T2)A       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       94         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       94         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95         PAHGVT*(C4T2)SAPDT*(PAPGST*(C4T2)A       93<	PAHGVTSAPDT*(C3T2)RPAPGSTA	78				
PAHGVT*(C312)SAPDTRPAPGST*(C312)A       81       81         PAHGVT*(C312)RAPGST*(C312)A       82       82         PAHGVT*(C312)SAPDT*(C312)RPAPGST*(C3T       83       83         PAHGVT*(C311)SAPDTRPAPGSTA       84       84         PAHGVT*(C411)SAPDT*(C411)RPAPGSTA       86       86         PAHGVT*(C411)SAPDT*(C411)RPAPGSTA       86       86         PAHGVT*(C411)SAPDT*(C411)RPAPGSTA       87       87         PAHGVT*(C411)SAPDT*(C411)RPAPGST*(C411)A       88       88         PAHGVT*(C411)SAPDT*(C411)RPAPGST*(C411)A       89       89         PAHGVT*(C411)SAPDT*(C412)RPAPGST*(C411)A       89       89         PAHGVT*(C412)SAPDT*(C412)RPAPGST*(C411)A       89       89         PAHGVT*SAPDTRPAPGST*(C412)A       93       93         PAHGVT*SAPDTRPAPGST*(C412)RPAPGSTA       92       90         PAHGVT*SAPDTRPAPGST*(C412)RPAPGSTA       92       92         PAHGVT*(C412)SAPDT*(C412)RPAPGSTA       93       93         PAHGVT*(C412)SAPDT*(C412)RPAPGST*(C412)A       93       93         PAHGVT*(C412)SAPDT*(C412)RPAPGST*(C412)A       93       94         PAHGVT*(C412)SAPDT*(C412)RPAPGST*(C412)A       93       94         PAHGVT*(C412)SAPDT*(A12)RPAPGST*(C412)A       93       94						
PAHGVTSAPDT*(C3T2)RPAPGST*(C3T2)A       82       82         PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T       83       83         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       84       83         PAHGVT*(C4T1)SAPDTRPAPGST*(C4T1)A       86       86         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       87       87         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       88       86         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89       89         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89       89         PAHGVT*(C4T1)SAPDT*(C4T2)RPAPGST*(C4T1)A       89       89         PAHGVT*SAPDT*(C4T2)RPAPGST*(C4T1)A       90       90         PAHGVT*SAPDT*(C4T2)RPAPGST*(C4T2)A       92       92         PAHGVT*SAPDT*(C4T2)RPAPGST*(C4T2)A       93       93         PAHGVT*SAPDT*(C4T2)RAPGST*(C4T2)A       93       93         PAHGVT*SAPDT*(C4T2)RAPGST*(C4T2)A       94       94         PAHGVT*SAPDT*(C4T2)RAPGST*(C4T2)A       94       94         PAHGVT*C4T2)SAPDT*(PAPGST*(C4T2)A       94       94         PAHGVT*C4T2)SAPDT*(PAPGST*(C4T2)A       95       96         PAHGVT*SAPDT*(C4T2)SAPDT*(PAPGST*(C4T2)A       96       97         PAHGVT*SAPDT*(PAPGST*(PAPAPGST*(C4T2)A       94       94						
PAHGVT*(C312)SAPDT*(C312)RPAPGST*(C31       83       83       83       84         PAHGVT*(C411)SAPDT*(C411)RPAPGSTA       84       84       84       84         PAHGVTSAPDT*(C411)RPAPGSTA       85       85       84       84         PAHGVTSAPDT*(C411)RPAPGSTA       85       87       87       87         PAHGVT*(C411)SAPDT*(C411)RPAPGSTA       87       87       87       87         PAHGVT*(C411)SAPDT*(C411)RPAPGST*(C411)A       88       88       88       87         PAHGVT*(C411)SAPDT*(C411)RPAPGST*(C41       90       90       90       90         PAHGVT*(C412)RPAPGST*(C411)A       88       88       88       88       88         PAHGVT*(C412)SAPDT*PAPGST*(C412)RPAPGST*(C41       90						
PAHGVT*(C4T1)SAPDTRPAPGSTA       84       84       84         PAHGVTSAPDT*(C4T1)RPAPGSTA       85       85       85         PAHGVTSAPDT*(C4T1)SAPDT*(C4T1)RPAPGSTA       87       87         PAHGVTSAPDT*(C4T1)SAPDTRPAPGST*(C4T1)A       86       86         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       88       86         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89       89         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89       89         PAHGVT*SAPDT*(C4T2)RPAPGST*(C4T2)A       90       90         PAHGVT*SAPDT*(C4T2)RPAPGSTA       91       91         PAHGVT*SAPDT*(C4T2)RPAPGSTA       92       92         PAHGVT*SAPDT*(C4T2)RPAPGSTA       92       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA       92         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95         PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A       95         PAHGVTSAPDT*RPAPGST*(C4T2)A       96         PAHGVT*C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       96         PAHGVTSAPDT*RPAPGST*(C4T2)A       96         PAHGVTSAPDT*RPAPGST*(C4T2)A       97         PAHGVTSAPDT*RPAPGST*(C4T2)A       96         PAHGVTSAPDT*RPAPGST*1       90	PAHGVT*(C3T2)SAPDT*(C3T2)RPAPGST*(C3T					
PAHGVTSAPDTRPAPGST*(C4T1)A       86       86         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGSTA       87         PAHGVT*(C4T1)SAPDT*(C4T1)A       88         PAHGVT*(C4T1)SAPDT*(C4T1)A       88         PAHGVT*(C4T1)SAPDT*(C4T1)A       88         PAHGVT*(C4T1)SAPDT*(C4T1)A       89         PAHGVT*(C4T1)SAPDT*(C4T1)A       89         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T       90         PAHGVT*SAPDT*(C4T2)RPAPGSTA       91         PAHGVT*SAPDTRPAPGST*(C4T2)RPAPGSTA       92         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA       94         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       96         PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A       96         PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A       96         PAHGVTSAPDT*RPAPGS**(C4T2)A       96         PAHGVTSAPDT*RPAPGS**(C4T2)A       96         PAHGVTSAPDT*RPAPGS**(C4T2)A       96         PAHGVT*APAPGS**       97         PAHGVT*APD*P       91         PAHGVT*APD*P       91         PAHGVT*APD*       91         PAHGVT*APD**PAPGS***APPA       96         PAHGVTSAPDT*RPAPGS****APPA       96	PAHGVT*(C4T1)SAPDTRPAPGSTA					
PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*       87       87       87         PAHGVT*(C4T1)SAPDTRPAPGST*(C4T1)A       88       88       87         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       88       88       87         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       89       89       89         PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T1)A       90       90       90         PAHGVT*SAPDT*(C4T2)RPAPGSTA       91       91       91         PAHGVT*SAPDT*(C4T2)RPAPGSTA       92       92       92         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       93       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95         PAHGVT*C4DPT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95         PAHGVT*CAPDT*RPAPGS*T*(C4T2)A       95       95         PAHGVTSAPDT*RPAPGS*T*(C4T2)A       95       96         PAHGVTSAPDT*RPAPGS*T*(C4T2)A       96       97         PAHGVTSAPDT*RPAPGS*T*(C4T2)A       96       97         PAHGVTSAPDT*RPAPGS*T*APPA Pep137       98       98         PAHGVTSAPDT*RPAPGS*T*APPA Pep138       99       99         PAHGVTSAPDT*RPAPGS*T*APPA Pep138       99         PAHGVTSAPOT*RP						
PAHGVT*(C4T1)SAPDTRPAPGST*(C4T1)A       88       88         PAHGVT*CAT1)RAPAPGST*(C4T1)A       89       89         PAHGVT*C4T1)RAPAPGST*(C4T1)A       89       89         PAHGVT*C4T2)RAPDT*(C4T1)RPAPGST*(C4T       90       90         PAHGVT*SAPDTRPAPGSTA       91       91         PAHGVT*SAPDT*(C4T2)RPAPGSTA       93       93         PAHGVT*C4T2)SAPDT*(C4T2)RPAPGSTA       93       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA       94         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA       94         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST       95         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T7       97         PAHGVTSAPDT*RAPGS*T*APPA Pep137       98         PAHGVTSAPDT*RPAPGS*T*APPA Pep138       99         PAHGVTSAPDT*RPAPGS*T*APPA Pep138       99         PAHGVTSAPDT*RPAPGS*T*APPA Pep139       100         MUC5B ATPSSTPGTTHTP       102         MUC5B ATPSST*PGTTHTP       102         MUC5B ATPSST*PGTTHTP       103         MUC5B ATPSST*CHTTP       103         MUC5B ATPSST*PGT*HTP       103         MUC5B ATPSST*PGT*HTP       103         MUC5B ATPSSTPGT*HTP       103         MUC5B ATPSSTPGT*HT						
PAHGVTSAPDT*(C4T1)RPAPGST*(C4T1)A     89     89     89       PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T     90     90       PAHGVT*SAPDTRPAPGSTA     91     91       PAHGVT*SAPDTRPAPGSTA     91     91       PAHGVT*SAPDTRPAPGSTA     92     92       PAHGVT*C4T2)SAPDTRPAPGSTA     93       PAHGVT*C4T2)SAPDTRPAPGSTA     93       PAHGVT*C4T2)SAPDTRPAPGST*(C4T2)A     93       PAHGVT*C4T2)SAPDT*(C4T2)RPAPGSTA     94       PAHGVT*C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A     95       PAHGVT*C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A     96       PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A     96       PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A     97       PAHGVTSAPDT*(C4T2)SAPDT*(PAPAPST*(C4T2)A     96       PAHGVTSAPDT*(APAPAPST*0PA Pep137     98       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99       PAHGVTSAPDT*RPAPGST*THP     100       MUC5B ATPSSTPGTTHTP     102       MUC5B ATPSSTPGTTHTP     103       MUC5B ATPSSTPGTTHTP     106 <t< th=""><th></th><th></th><th></th><th></th><th></th><th></th></t<>						
PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T       90       90       90         PAHGVT*SAPDT*RAPGSTA       91       91       91         PAHGVT*SAPDT*(C4T2)RPAPGSTA       92       92       91         PAHGVT*(C4T2)SAPDT*(C4T2)RAPGSTA       92       92       91         PAHGVT*(C4T2)SAPDT*(C4T2)RAPGSTA       94       93       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95       95         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95       95         PAHGVT*CAT2)SAPDT*(C4T2)RPAPGST*(C4T2)A       96       96       97         PAHGVT*SAPDT*RPAPGS**(C4T2)A       96       96       97         PAHGVTSAPDT*RPAPGS**       91       97       97         PAHGVTSAPDT*RPAPGS**       91       99       99         PAHGVTSAPDT*RPAPGS**       91       90       90         PAHGVTSAPDT*RPAPGS**       91       101       101         MUCSB AT*PSSTPGTH*TP       102	PAHGVTSAPDT*(C4T1)RPAPGST*(C4T1)A					
PAHGVTSAPDT*(C4T2)RPAPGSTA       92       92         PAHGVTSAPDT*(C4T2)RAPGST*(C4T2)A       93       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA       94       93         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95         PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A       96       96         PAHGVT*CAT2)SAPDT*(C4T2)RPAPGST*(C4T2)A       95       95         PAHGVT*CAT2)SAPDT*(C4T2)RPAPGST*(C4T2)A       96       96         PAHGVTSAPDT*RAPGS*T*(C4T2)RPAPGST*(C4T1       97       97         PAHGVTSAPDT*RPAPGS*T*APPA Pep137       98       98         PAHGVTSAPDT*RPAPGS*T*APPA Pep138       99       99         PAHGVTSAPDT*RPAPGS*T*APPA Pep139       100       101         MUC58 ATPSSTPGTTHTP       102       102         MUC58 ATPSST*PGTTHTP       103       103         MUC58 ATPSST*PGTT*TP       103       103         MUC58 ATPSST*PGTT*TP       103       103         MUC58 ATPSST*PGT*T*TP       103       103         MUC58 ATPSST*PGT*T*TP       103       103         MUC58 ATPSST*PGT*T*TP       103       103         MUC58 ATPSST*QAT*********************************	PAHGVT*(C4T1)SAPDT*(C4T1)RPAPGST*(C4T		90			
$\begin{array}{c c c c c c c c c c c c c c c c c c c $						
PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA     94     94     94       PAHGVT*(C4T2)SAPDTRPAPGST*(C4T2)A     95     95       PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A     96     96       PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGST*(C4T2)A     96     96       PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A     96     97       PAHGVTSAPDT*RPAPGS*T*APPA Pep137     98     98       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     90     100       MUCSB AT*PSSTPGTTHTP     101     101       MUCSB AT*PSSTFGTTHTP     102     102       MUCSB ATPSST*GTTHTP     103     103       MUCSB ATPSST*PGTTHTP     103     103       MUCSB ATPSST*PGTTHTP     103     103       MUCSB ATPSST*PGTTHTP     103     104       Fetuin fetal bovine serum     105     105       BSA     106     106     106       Albumin from human serum     107     107     107       ICAM+human     108     108     108						
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	PAHGVT*(C4T2)SAPDT*(C4T2)RPAPGSTA	94				
PAHGVTSAPDT*(C4T2)RPAPGST*(C4T2)A     96     96       PAHGVTSAPDT*(C4T2)SRPAPGST*(C4T2)A     97     96       PAHGVTSAPDT*(PAPGST*(AT2)RPAPGST*(C4T2)A     98     98       PAHGVTSAPDT*(PAPACS*T*APPA Pep137     98     98       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     100     100       MUC58 ATPSSTPGTTHTP     101     101       MUC58 ATPSSTPGTTHTP     102     102       MUC58 ATPSSTPGTTHTP     102     102       MUC58 ATPSSTPGTTHTP     103     103       Galacid glycoptein     104     104       Fetuin fetal bovine serum     105     105       BSA     106     106     106       Albumin from human serum     107     107     107	PAHGVT*(C4T2)SAPDTRPAPGST*(C4T2)A					
PAHGVTSAPDT*RPAPGS*T*APPA Pep137     98     98     98       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep139     100       MUC58 AT*PSSTPGTHTP     101       MUC58 ATPSSTPGTTHTP     102       MUC58 ATPSSTPGTTHTP     103       MUC58 ATPSSTPGTTHTP     103       MUC58 ATPSSTPGT*THTP     104       MUC58 ATPSSTPGT*THTP     104       MUC58 ATPSSTPGT*THTP     105       MUC58 ATPSSTPGT*THTP     104       MUC58 ATPSSTPGT*THTP     105       MUC58 ATPSSTPGT*THTP     107       MUC58 ATPSSTPGT*THTP     107       MUC58 ATPSSTPGT*THTP     106       MUC58 ATPSSTPGT*THTP     107       MUC58 ATPSSTPGT*THTP     106       MUC58 ATPSSTPGT*THTP     106       MUC58 ATPSSTPGT*THTP     107       MUC58 ATPSSTPGT*THTP     106       MUC58 ATPSSTPGT*THTP     107       MUC58 ATPSSTPGT*THTP     107       MUC58 ATPSSTPGT*THTP						
PAHGVTSAPDT*RPAPGS*T*APPA Pep138     99     99     99       PAHGVTSAPDT*RPAPGS*T*APPA Pep139     100     100       MUCSB AT*PSSTPGTTHTP     101     101       MUCSB AT*PSSTPGTTHTP     102     102       MUCSB ATPSSTPGTTHTP     103     103       MUCSB ATPSSTPGTTHTP     103     104       Fetuin fetal bovine serum     105     105       BSA     106     106     106       Albumin from human serum     107     107       ICAM+human     108     108						
PAHGVTSAPD1*RPAPGS*1*APPA Pep139         100         100         100         100         100         100         100         100         100         100         100         100         100         100         101         102         103         103         103         103         103         103         103         103         103         103         104         104         104         104         104         104         104         105         105         105         105         107         107         107         107         107 <t< td=""><td>PAHGVTSAPDT*RPAPGS*T*APPA Pep138</td><th>99</th><td></td><td></td><td></td><td></td></t<>	PAHGVTSAPDT*RPAPGS*T*APPA Pep138	99				
MUCSB AT*PSSTPGTTHTP         101         101         101           MUCSB ATPSSTPGTTHTP         102         102         102           MUCSB ATPSSTPGT*THTP         103         103         103						
MUC5B ATPSSTPGT*THTP         103         103         103           α1-acid glycoprotein         104         104         104           Fetuin fetal bovine serum         105         105         105           BSA         106         106         106           Albumin from human serum         107         107         107           ICAM+human         108         108         108			101	101		
α1-acid glycoprotein         104         104           Fetuin fetal bovine serum         105         105         105           BSA         106         106         105         105           Albumin from human serum         107         107         107         107           ICAM-I human         108         108         108         108         108						
Fetuin fetal bovine serum         105         105         105           BSA         106         106         106         106         106         106         106         106         106         106         106         106         106         106         106         107         107         107         107         107         107         107         107         107         108						
BSA         106         H         106         H           Albumin from human serum         107         107         107         107         107         107         107         107         107         107         107         107         107         107         107         108	Fetuin fetal bovine serum	105				
ICAM-I human 108 108 108			106	H 106 H		
Transferrin human         109         109         109						
Mucin from porcine stomach 110 110 110 110 110	Mucin from porcine stomach	110				
Mucin from bovine submaxillary 111 111 111 111	Mucin from bovine submaxillary	111	111			



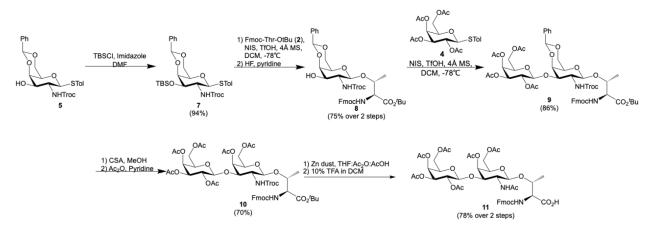
**Figure S6.** Glycopeptide microarray screening results of anti-sera induced by  $Q\beta$ -MUC1- $\beta$ -Tf . The results from five anti-sera (1/100 dilution) were shown.

#### **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<sub>4</sub>)<sub>2</sub>(NO<sub>3</sub>)<sub>6</sub> (0.5 g) and (NH<sub>4</sub>)<sub>6</sub>MO<sub>7</sub>O<sub>24</sub>\*4H<sub>2</sub>O (24.0 g) in 6% H<sub>2</sub>SO<sub>4</sub> (500 mL) or 5% H<sub>2</sub>SO<sub>4</sub> 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 QB-MUC1 conjugates, liquid chromatography-mass spectrometry (LCMS) analysis was performed. The samples for LCMS were prepared as follows: 1:1 v/v of 40  $\mu$ g mL<sup>-1</sup> 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<sub>3</sub>CN (0.3 mL min<sup>-1</sup> 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<sup>-1</sup> 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  $\mu$ L) and matrix solution (2  $\mu$ L, 10 mg mL<sup>-1</sup> sinapic acid in 50/50/0.1 CH<sub>3</sub>CN/H<sub>2</sub>O/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-Nacetylgalactosaminidase 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<sup>-1</sup> penicillin and 100  $\mu$ g mL<sup>-1</sup> streptomycin, 2 mM L-glutamine, 1 mM sodium pyruvate and 0.3 mg mL<sup>-1</sup> 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)-β-D-galactopyranoside (7).

5 (0.450g, 0.820mmol, 1 equiv.),<sup>2</sup> imidazole (0.112g, 1.64mmol, 2 equiv.) and tertbutyldimethylsilyl chloride (0.161g, 1.07mmol, 1.3 equiv.) were dissolved in DMF (4mL). The reaction was stirred, under N<sub>2</sub>, overnight. Upon completion of the reaction, the solution was poured into EtOAc and washed 3 times each with 1M HCl, sat. NaHCO<sub>3</sub>, and sat. NaCl solutions. The organic layers were collected and dried over anhydr. Na<sub>2</sub>SO<sub>4</sub> 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. <sup>1</sup>H NMR (500MHz, CDCl<sub>3</sub>) 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<sub>2</sub>), 4.66 (d, J = 12.0Hz, 1H, Troc CH<sub>2</sub>), 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<sub>3</sub>), 0.85 (s, 9H, tBu CH<sub>3</sub>), 0.06 (s, 3H, silyl CH<sub>3</sub>), 0.05 (s, 3H, silyl CH<sub>3</sub>). <sup>13</sup>C NMR (125MHz, CDCl<sub>3</sub>) 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<sub>29</sub>H<sub>38</sub>Cl<sub>3</sub>NNaO<sub>6</sub>SSi): calculated ([M+Na<sup>+</sup>]): 684.1152, found 684.1151.

# *O*-(4,6-*O*-Benzylidene-2-2-deoxy-*N*-((2,2,2-trichloroethoxy)carbonyl)-β-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<sub>2</sub>. 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-OtBu. 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<sub>3</sub> solution and dissolved in EtOAc. The solution was washed 3 times each with sat. copper sulfate, 1M HCl, sat. NaHCO<sub>3</sub>, and sat. NaCl solutions. The organic layers were collected, dried over anhydr. Na<sub>2</sub>SO<sub>4</sub> 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. <sup>1</sup>H NMR (500MHz, CDCl<sub>3</sub>)  $\delta$  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<sub>2</sub>), 4.68 (d, J = 7.7Hz, 1H, H-1), 4.55 (m, 1H, Thr CH), 4.42 (m, 1H, Fmoc CH<sub>2</sub>), 4.35-4.22 (m, 4H, Fmoc CH, Fmoc CH<sub>2</sub>, Thr C<sub>α</sub>H, 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<sub>3</sub>), 1.23 (d, J = 6.3Hz, 3H, Thr CH<sub>3</sub>). <sup>13</sup>C NMR (125MHz, CDCl<sub>3</sub>)  $\delta$ 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<sup>+</sup>]): 843.1830, found 843.1798.

# *O*-(4,6-*O*-Benzylidene-2-deoxy-*O*-(2,3,4,6-tetra-*O*-acetyl-**B**-D-galactopyranosyl)-2-*N*-((2,2,2-trichloroethoxy)carbonyl)-**B**-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<sub>2</sub>. 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. <sup>1</sup>H NMR (500MHz, CDCl<sub>3</sub>)  $\delta$  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<sub>2</sub>), 4.72 (d, J == 8.0Hz, 1H, H-1'), 4.60 (d, J = 12.1Hz, 1H, Troc CH<sub>2</sub>), 4.58-4.50 (m, 2H, H-4, Thr CH), 4.41-4.26 (m, 5H, Fmoc CH<sub>2</sub>, Thr C<sub> $\alpha$ </sub>H, 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<sub>2</sub>), 3.89 (t, J = 6.3Hz, 1H, H-5'), 3.47 (m, 2H, H-3, H-5), 2.16 (s, 3H, Ac CH<sub>3</sub>), 2.06 (s, 3H, Ac CH<sub>3</sub>), 2.04 (s, 3H, Ac CH<sub>3</sub>), 1.96 (s, 3H, Ac CH<sub>3</sub>), 1.43 (s, 9H, tBu CH<sub>3</sub>), 1.20 (d, J =

6.3Hz, 3H, Thr CH<sub>3</sub>). <sup>13</sup>C NMR (125MHz, CDCl<sub>3</sub>)  $\delta$  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<sub>53</sub>H<sub>61</sub>Cl<sub>3</sub>N<sub>2</sub>NaO<sub>20</sub>): calculated ([M+Na<sup>+</sup>]): 1173.2781, found 1173.2784.

# *O*-(4,6-Di-*O*-acetyl-2-deoxy-*O*-(2,3,4,6-*O*-tetra-*O*-acetyl-B-D-galactopyranosyl)-2-*N*-((2,2,2-trichloroethoxy)carbonyl)-B-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<sub>2</sub>O was added and the reaction was stirred under N<sub>2</sub>. The reaction was dissolved in EtOAc and washed 3 times each with sat. copper sulfate, 1M HCl, sat. NaHCO<sub>3</sub>, and sat. NaCl solutions. The organic layers were collected, dried over anhydr. Na<sub>2</sub>SO<sub>4</sub> 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. <sup>1</sup>H NMR (500MHz, CDCl<sub>3</sub>) 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<sub>2</sub>), 4.64 (d, J = 12.1Hz, 1H, Troc CH<sub>2</sub>), 4.58 (d, J = 8.1Hz, 1H, H-1'), 4.45-4.35 (m, 2H, Thr CH, Fmoc CH<sub>2</sub>), 4.30-4.17 (m, 4H, Fmoc CH, Fmoc CH<sub>2</sub>, Thr C<sub> $\alpha$ </sub>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<sub>3</sub>), 1.92 (s, 3H, Ac CH<sub>3</sub>), 1.43 (s, 9H, tBu CH<sub>3</sub>), 1.14 (d, J = 6.3Hz, 3H, Thr CH<sub>3</sub>). ESI-TOF (C<sub>50</sub>H<sub>61</sub>Cl<sub>3</sub>N<sub>2</sub>NaO<sub>22</sub>): calculated ([M+Na<sup>+</sup>]): 1169.2679, found 1169.2675.

# *O*-(2-Acetamido-4,6-di-*O*-acetyl-2-deoxy-*O*-(2,3,4,6-tetra-*O*-acetyl-B-D-galactopyranosyl)-B-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<sub>2</sub>O:acetic acid (AcOH) (2mL). The reaction was stirred under N<sub>2</sub>, 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 by silica gel flash column chromatography, using MeOH:DCM as the eluent. 11 (124mg) was obtained as a yellow glassy solid in 78% yield. <sup>1</sup>H NMR (500 MHz, Methanol-*d*<sub>4</sub>)  $\delta$  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<sub>2</sub>), 4.38 – 4.31 (m, 2H, Fmoc CH<sub>2</sub>, Thr CH), 4.22 (t, J = 6.7 Hz, 1H, Fmoc CH), 4.16 – 3.94 (m, 8H, Thr C<sub>a</sub>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<sub>3</sub>), 2.05 (s, 3H, Ac CH<sub>3</sub>), 2.03 (s, 3H, Ac CH<sub>3</sub>), 2.03 (s, 3H, Ac CH<sub>3</sub>), 2.05 (s, 3H, Ac CH<sub>3</sub>), 1.96 (s, 3H, Ac CH<sub>3</sub>), 1.94 (s, 3H, Ac CH<sub>3</sub>), 1.18 (d, J = 6.3Hz, 3H, Thr CH<sub>3</sub>). <sup>13</sup>C NMR (126 MHz, Methanol-*d*<sub>4</sub>) 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<sub>45</sub>H<sub>53</sub>N<sub>2</sub>O<sub>21</sub>): calculated ([M-H<sup>+</sup>]): 957.3141, found 957.3128.

#### Synthesis of MUC1-B-Tf (12):

The MUC1- $\beta$ -Tf glycopeptide 12 was synthesized using *p*-nitrophenyl carbonate Wang resin. The N-Fmoc-1,4-diaminobutane linker was installed by dissolving N-Fmoc-1,4diaminobutane\*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-BTf glycopeptide, the N-terminal Fmoc group was removed and the resulting free amine was capped with a 1:1:8 solution of Ac<sub>2</sub>O: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<sub>2</sub>O. After 4hrs, the glycopeptide was precipitated by diethyl ether (EtO<sub>2</sub>) 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  $\times$  10 mm 5  $\mu$ m with an acetonitrile (ACN):H<sub>2</sub>O (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-B-Tf glycopeptide onto QB, 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<sub>2</sub>O for 1.5 hrs before being purified by HPLC to give a yield of 30-40%. The MUC1- $\beta$ -Tf glycopeptide 12 was characterized by ESI-TOF MS and <sup>1</sup>H NMR before pnitrophenyl ester functionalization. <sup>1</sup>H NMR (500MHz, deuterium oxide)  $\delta$  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<sub>3</sub>), 1.83 (s, 3H, NHAc CH<sub>3</sub>), 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<sub>3</sub>), 1.15 (d, J = 2.3Hz, 3H, Ala CH<sub>3</sub>), 1.09 (d, J = 10.4Hz, 2H), 0.93 (d, J = 6.4Hz, 3H, Thr CH<sub>3</sub>); ESI-TOF (C<sub>58</sub>H<sub>97</sub>N<sub>15</sub>O<sub>24</sub>): calculated ([M+2H<sup>+</sup>]/2): 694.8494, found 694.8515

#### **Glycosidase cleavage experiments:**

100 µg of glycopeptide, MUC1- $\alpha$ -Tf **12\alpha** or MUC1- $\beta$ -Tf **12**, 5 µl of 10X glycoBuffer, 10 µl endo- $\alpha$ -*N*-acetylgalactosaminidase from *Enterococcus faecalis* (400,000 units, New England Biolabs, P0733S) and H<sub>2</sub>O 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  $\alpha$ -isomer **12\alpha** was observed with nearly complete cleavage after 1.75 hours (**Figure S1**). In contrast, the  $\beta$ -anomer **12** showed no detectable cleavage even after 24 hours.

To further test the glycopeptide stability, MUC1- $\beta$ -Tf glycopeptide **12** and MUC1- $\alpha$ -Tf glycopeptide **12** $\alpha$  with a  $\beta$ -*N*-acetylhexosaminidase (New England Biolabs, P0721S). 50 µg of glycopeptide, MUC1- $\alpha$ -Tf **12** $\alpha$  or MUC1- $\beta$ -Tf **12**, 2.5 µl of 10X glycoBuffer, 10 µl  $\beta$ -*N*-acetylhexosaminidase (50 units) and H<sub>2</sub>O 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  $\beta$ -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- $\beta$ -D-glucosaminide was completely cleaved in 24 hours under the same reaction condition.

#### Synthesis of $Q\beta$ -MUC1 Conjugates.

For Q $\beta$ -MUC1- $\alpha$ -Tf: The conjugation was performed as previously reported.<sup>1</sup> For Q $\beta$ -MUC1- $\beta$ -Tf: A solution of Q $\beta$  (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<sup>-1</sup> (concentration should not be <5 mg mL<sup>-1</sup> to prevent aggregation). Protein concentration 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  $\mu$ L of DMSO, then added to a solution of BSA (10 mg mL<sup>-1</sup>) in 0.1M KPB pH 7.0 (200  $\mu$ L). 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 Q $\beta$ -MUC1 vaccines in PBS containing MPLA (20  $\mu$ L, 1 mg mL<sup>-1</sup> 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 Q $\beta$ -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  $\mu$ g mL<sup>-1</sup> of the corresponding BSA-MUC1 conjugates (100  $\mu$ L/well) in NaHCO<sub>3</sub>/Na<sub>2</sub>CO<sub>3</sub> buffer (0.05 M, pH 9.6) containing 0.02 % NaN<sub>3</sub> by incubation at 4°C overnight. The coated plates were washed with PBS/0.5% Tween-20 (PBST) (4 × 200  $\mu$ L) and blocked with 1 % BSA in PBS (100 $\mu$ L/well) at rt for 1 h. The plates were washed again with PBST (4 × 200  $\mu$ L) and incubated with serial dilutions of mice sera in 0.1 % BSA/PBS (100  $\mu$ L/well, 2 wells for each dilution). The plates were incubated for 2 h at 37 °C and then washed with PBST (4 × 200  $\mu$ L). 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  $\mu$ L) 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  $\mu$ L). A solution of enzymatic substrate 3,3',5,5'-tetramethylbenzidine(TMB, 200  $\mu$ L) 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  $\mu$ L of H<sub>2</sub>O<sub>2</sub>). Color was allowed to develop for 15 min and then quenched by adding 50  $\mu$ L of 0.5 M H<sub>2</sub>SO<sub>4</sub>. 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<sub>2</sub> in cell growth medium. The number of cells was determined using a hemocytometer. Suspensions of  $3.0 \times 10^5$  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<sub>3</sub>) 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<sup>-1</sup>) 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:**

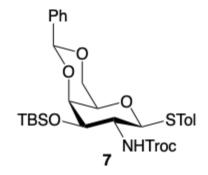
Complement dependent cytotoxicity of B16-MUC1 cells was determined by MTS assay. B16-MUC1 (7000 cells/well) were cultured in 96 well plate with DMEM (10% FBS, G-418 and 1% P.S.) for 12 h. The culture medium was carefully removed. A dilution of mouse sera (1/40) from different groups of immunized MUC1.Tg mice in 50  $\mu$ L 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  $\mu$ L 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  $\mu$ L) 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:**

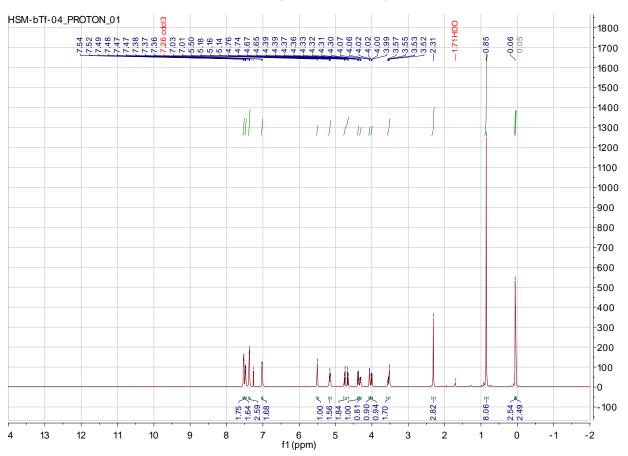
Glycopeptides and glycoproteins were dissolved in sodium phosphate buffer (150 mM, pH 8.5) at a final concentration of 50 mM and transferred into a 384-well microtiter plate. All arrays were printed on NHS-activated hydrogel microarray glass slides (Schott, Nexterion, slide H). The spotter settings were adjusted to generate substrate peptide spots of 100 pL  $\pm$  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  $10 \times 13$  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  $26 \times 20$ 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.

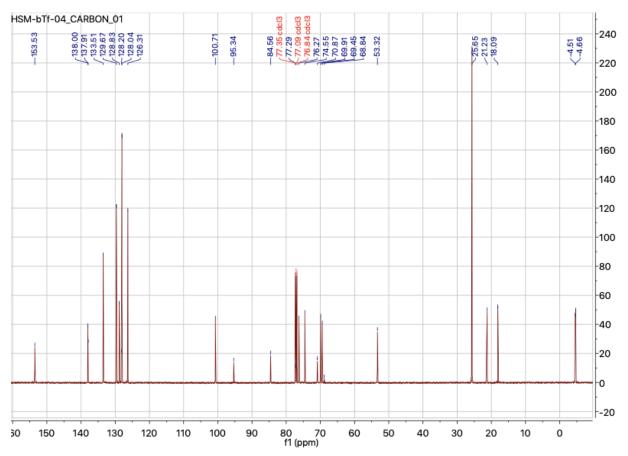
# Characterization Data and Spectra of All Building Blocks and Glycopeptides:



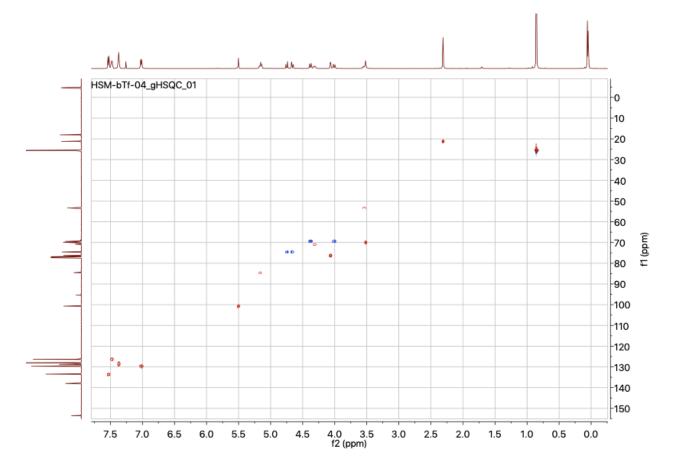
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) of **7** 



# $^{13}\text{C}$ NMR (126 MHz, CDCl\_3) of 7



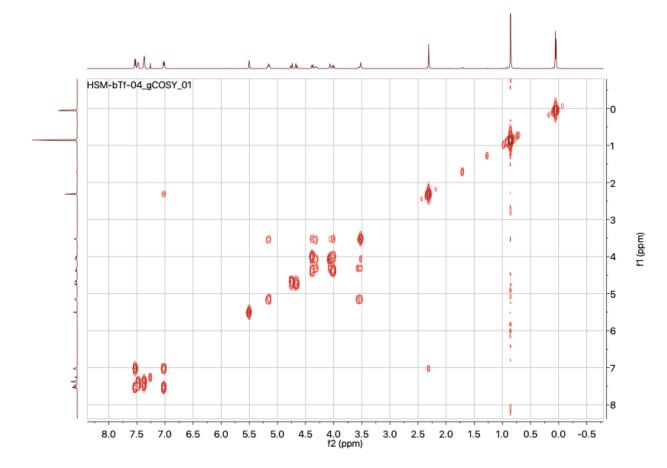
<sup>1</sup>H-<sup>13</sup>C HSQC (500 MHz, CDCl<sub>3</sub>) of **7** 

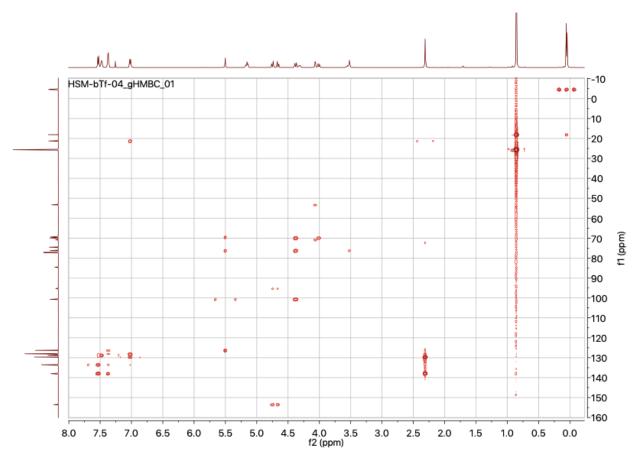


MI d h the la HSM-bTf-04\_gHSQCAD\_01 -780 0 ŀo -10 = -20 8 8 -30 -40 -50 ٠ . -60 f1 (ppm) -70 6 8 9 0 0 0 40 00 . -80 0 -90 -100 0 -110 120 • • • 8 -130 0 -140 -150 4.0 3.5 f2 (ppm) 7.5 8.0 7.0 6.0 5.5 4.5 3.0 2.5 2.0 1.0 0.5 0.0 6.5 5.0 1.5

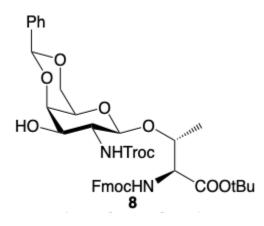
# $^{1}\text{H-}^{13}\text{C}$ HSQCAD (500 MHz, CDCl<sub>3</sub>) of **7**

 $^{1}\text{H}\text{-}^{1}\text{H}$  COSY (500 MHz, CDCl<sub>3</sub>) of **7** 

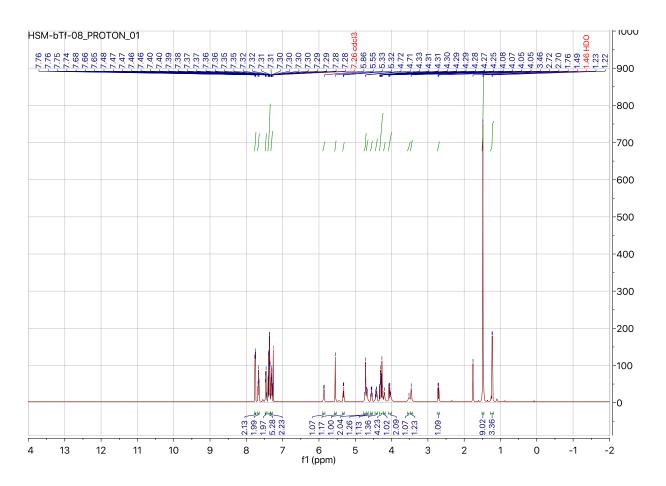




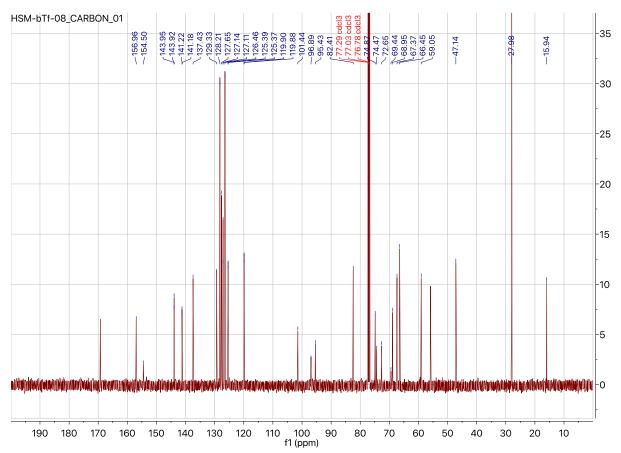
# $^1\text{H-}^{13}\text{C}$ HMBC (500 MHz, CDCl<sub>3</sub>) of 7



<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) of  $\bf{8}$ 



# $^{13}\text{C}$ NMR (126 MHz, CDCl\_3) of 8



\_linh HSM-bTf-08\_gHSQC\_01 0 -20 -30 -40 0 -50 -60 0 ° ° -70 ò 0 ۰ -80 f1 (ppm) -90 8 -100 -110 120 • ° 🌼 -130 -140 -150 7.5 4.5 f2 (ppm) 8.0 7.0 6.5 6.0 5.5 5.0 4.0 2.5 2.0 1.5 3.5 3.0

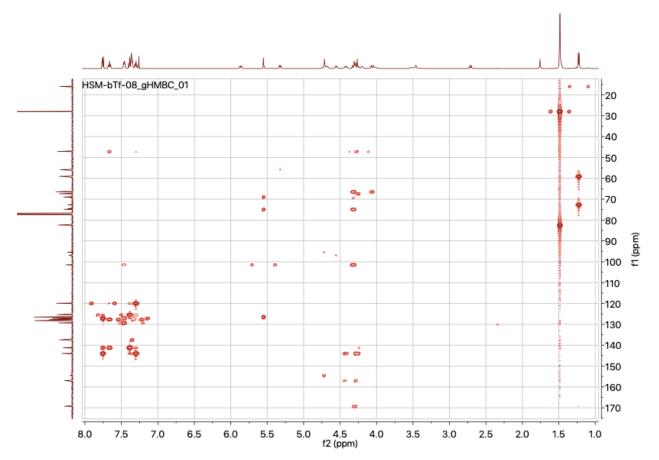
# <sup>1</sup>H-<sup>13</sup>C HSQC (500 MHz, CDCl<sub>3</sub>) of 8

Inth HSM\_bTf\_TBS-Removal\_gH\$QCAD\_01 . 0 -20 --30 -40 n e -50 . 4 0 0 -60 • السيدار ∘*₽ •∂* n -70 0 0 0 -80 90 (mdg) 100 F 0 0 . -110 -120 0 0 •. • & • ~ -130 H -140 150 -160 -170 4.5 f2 (ppm) 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.0 3.5 3.0 2.5 2.0 1.5 1.0

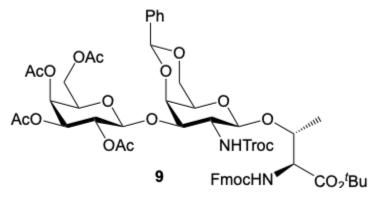
## <sup>1</sup>H-<sup>13</sup>C HSQCAD (500 MHz, CDCl<sub>3</sub>) of **8**

\_linh HSM-bTf-08\_gCOSY\_01 0 -1.5 -2.0 0 -2.5 0 -3.0 -3.5 • • • • • • • -4.0 0 June Û 4.5 (mdd) 0 -5.0 5 ۵ -5.5 0 0 6.0 6.5 -7.0 00.000 -7.5 00 7.5 7.0 4.5 f2 (ppm) 8.0 6.5 6.0 5.5 5.0 4.0 3.5 3.0 2.5 2.0 1.5

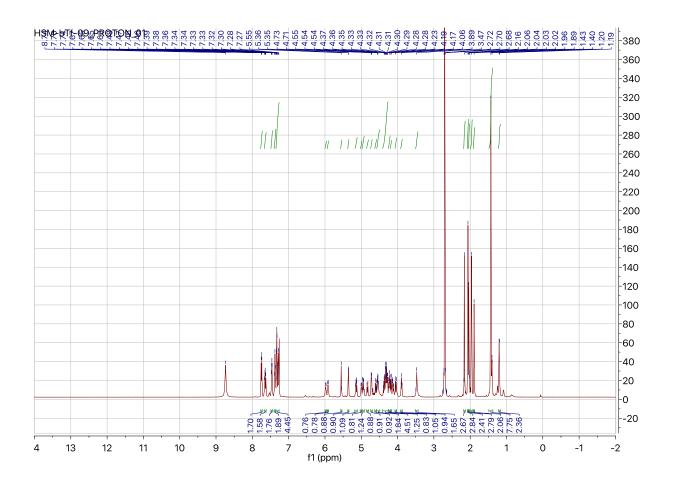
## <sup>1</sup>H-<sup>1</sup>H COSY (500 MHz, CDCl<sub>3</sub>) of **8**



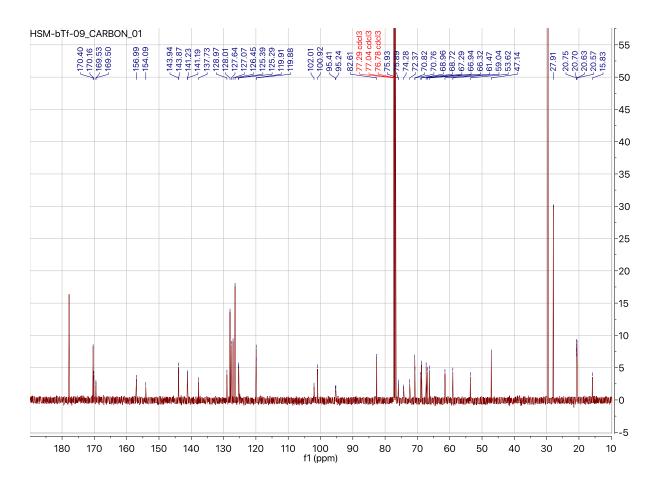
# $^1\text{H-}^{13}\text{C}$ HMBC (500 MHz, CDCl<sub>3</sub>) of $\boldsymbol{8}$



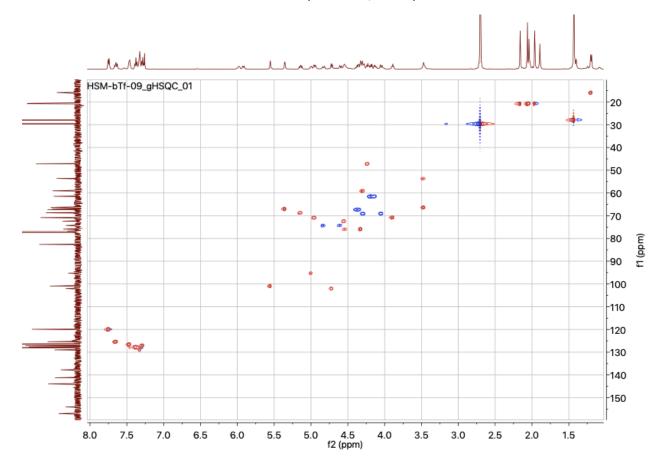
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) of  $\mathbf{9}$ 



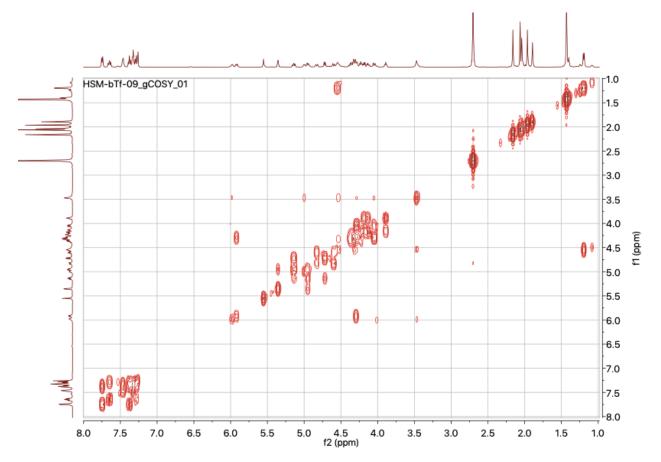
### $^{13}$ C NMR (126 MHz, CDCl<sub>3</sub>) of **9**



<sup>1</sup>H-<sup>13</sup>C HSQC (500 MHz, CDCl<sub>3</sub>) of **9** 

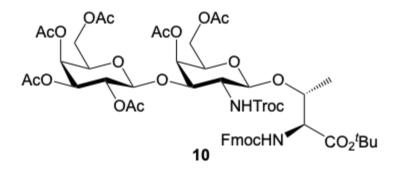


<sup>1</sup>H-<sup>1</sup>H COSY (500 MHz, CDCl<sub>3</sub>) of **9** 

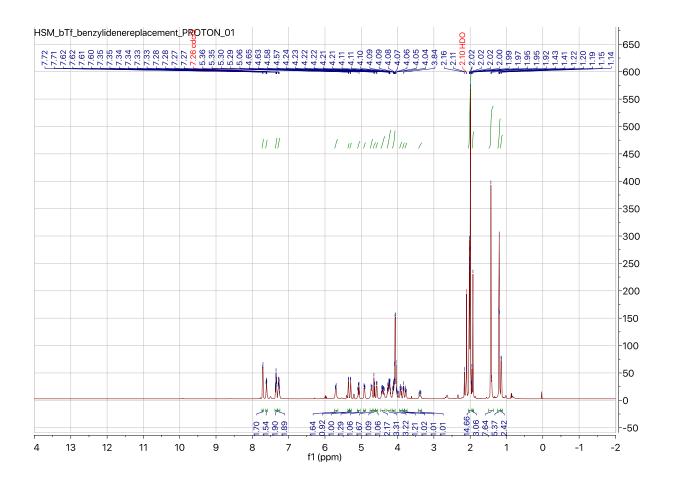


\_mM HSM-bTf-09\_gHMBC\_01 . . . o -20 00000 o 🍓 o -30 -40 68 . **6**00 \*\* . -50 -60 φ • 3 -70 **•** ° 0 ø •0 -80 -90 -90 100 g -110 <sup>g</sup> 0 • -110 -120 : ింద చిళ్లింం -130 -140 **ക**് -150 e a --160 ÷0 -170 ģ -180 <sup>L</sup>190 7.5 4.5 f2 (ppm) 8.0 7.0 6.5 6.0 5.5 5.0 3.0 2.5 2.0 1.5 4.0 3.5

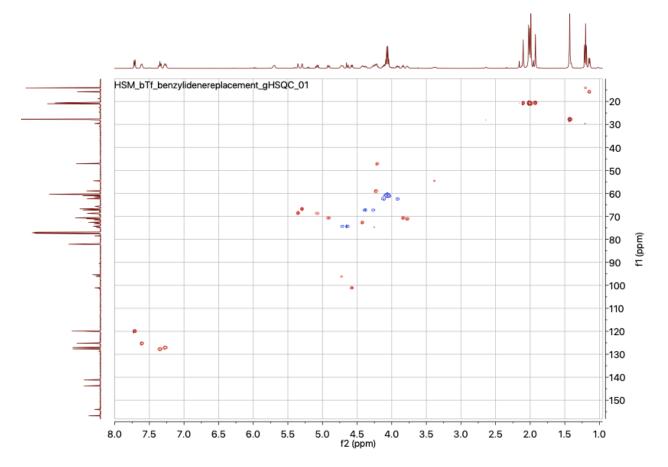
# <sup>1</sup>H-<sup>13</sup>C HMBC (500 MHz, CDCl<sub>3</sub>) of **9**



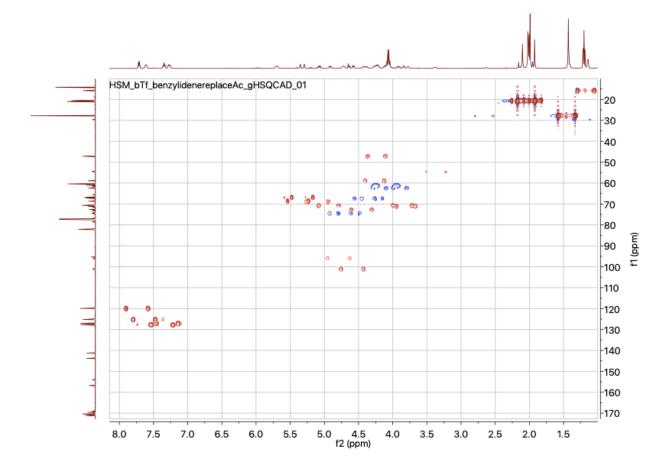
<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) of  $\mathbf{10}$ 



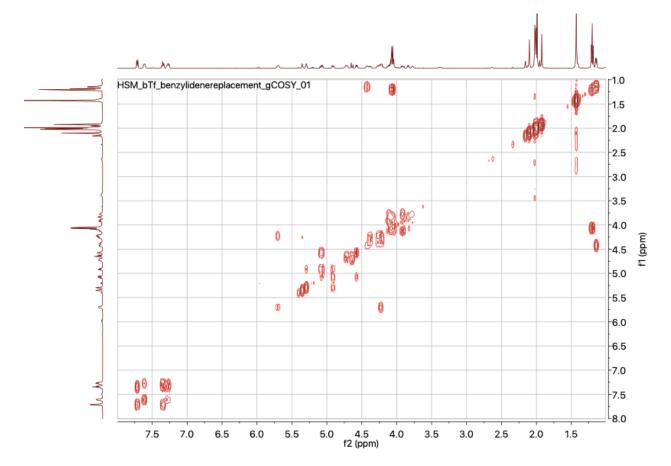
 $^{1}\text{H}\text{-}^{13}\text{C}$  HSQC (500 MHZ, CDCl<sub>3</sub>) of **10** 



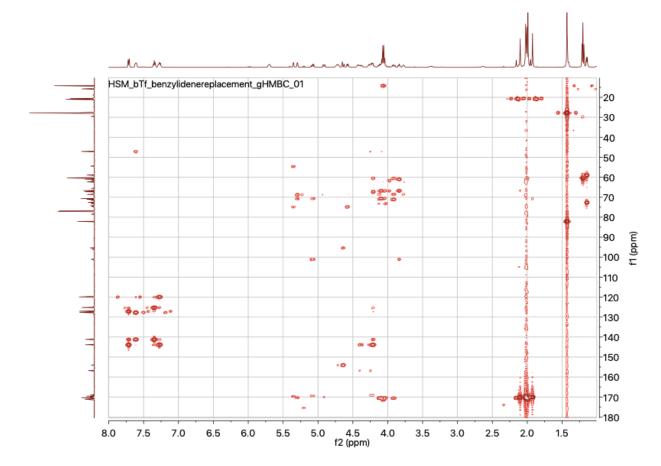
## $^{1}\text{H}\text{-}^{13}\text{C}$ HSQCAD (500 MHz, CDCl<sub>3</sub>) of **10**

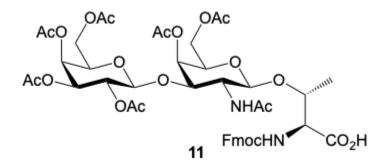


## <sup>1</sup>H-<sup>1</sup>H COSY (500 MHz, CDCl<sub>3</sub>) of **10**

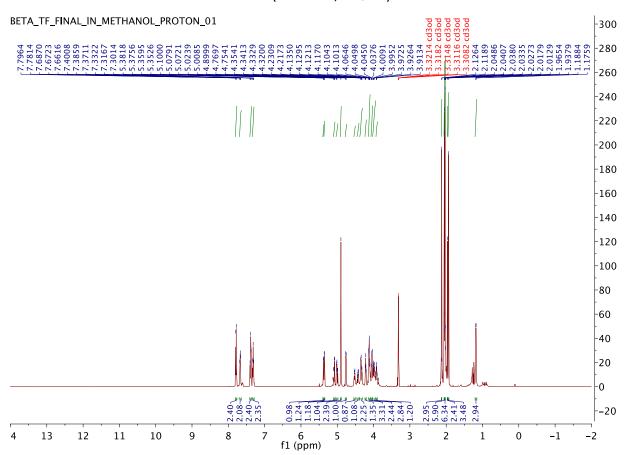


# <sup>1</sup>H-<sup>13</sup>C HMBC (500 MHz, CDCl<sub>3</sub>) of **10**

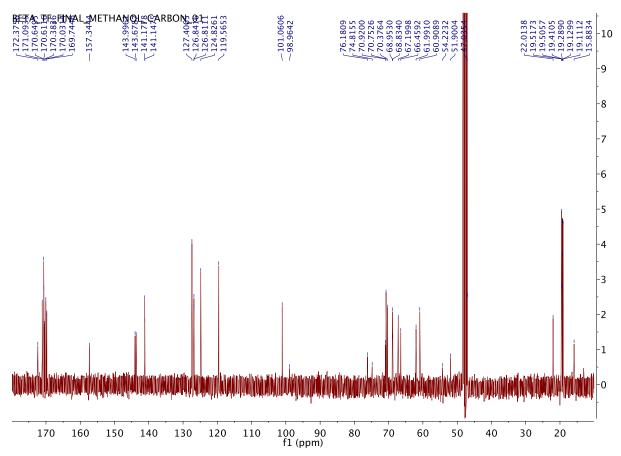


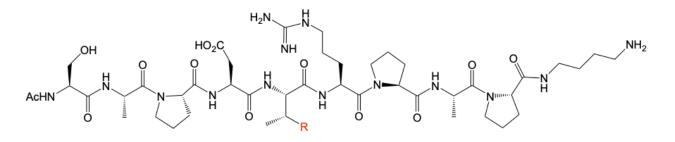


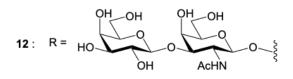
 $^1\text{H}$  NMR (500 MHz, CD<sub>3</sub>OD) of 11



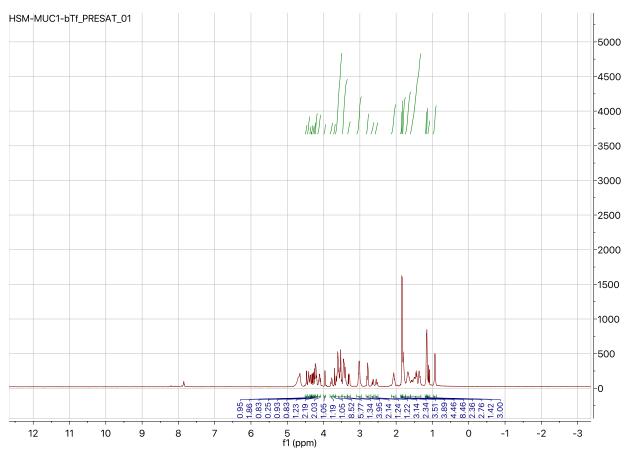


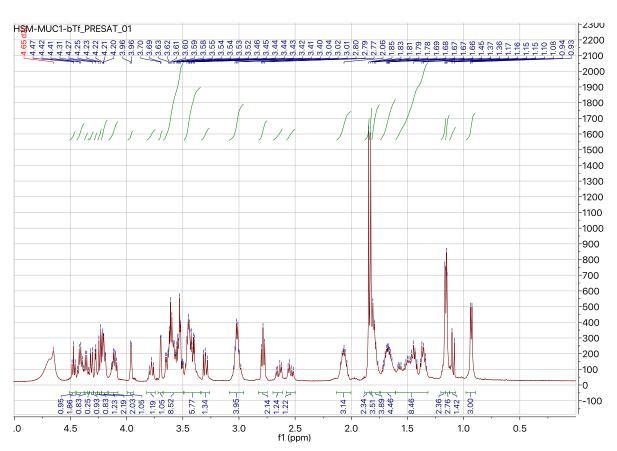




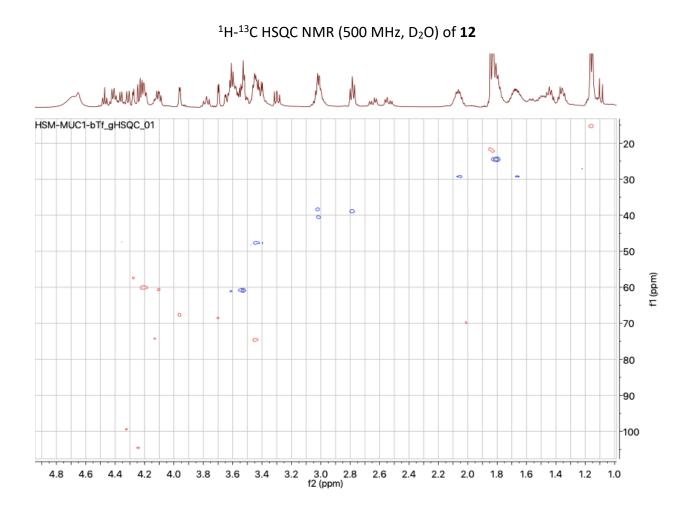


<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O) of **12** 

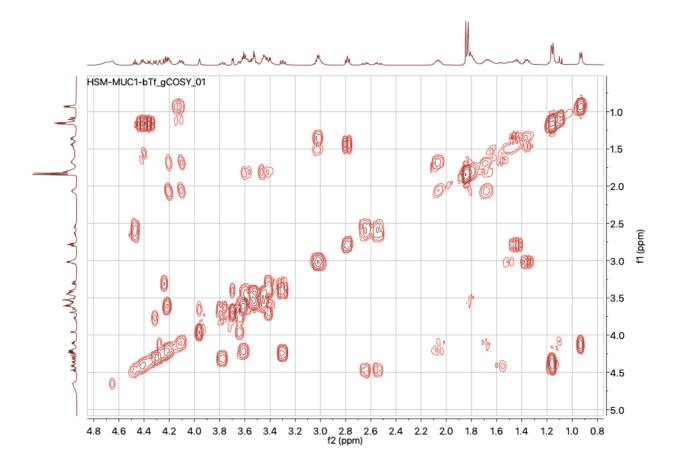


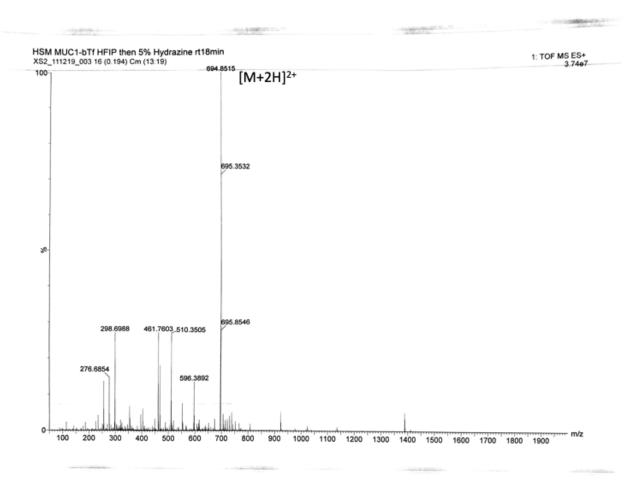


Zoomed in <sup>1</sup>H NMR (500 MHz,  $D_2O$ ) of **12** 



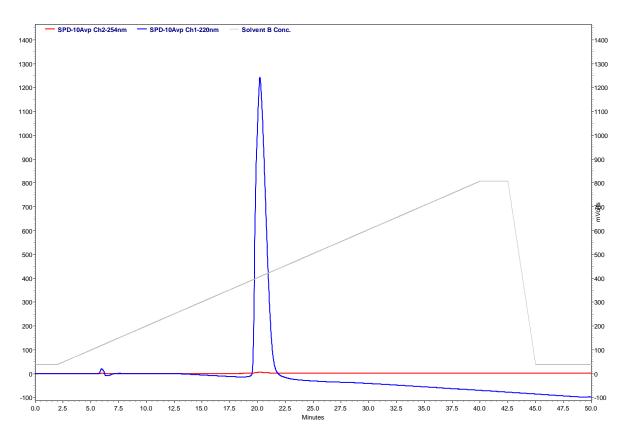
<sup>1</sup>H-<sup>1</sup>H COSY (500 MHz, D<sub>2</sub>O) of **12** 





## ESI-TOF MS of **12**

HPLC of purified 12



#### **References:**

- 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.