A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Synthetic MUC1 glycopeptide bearing βGalNAc-Thr as Tn antigen isomer induces antibodies production against tumor cells

Authors: Vanessa Leiria Campo, Thalita Riul, Leandro Bortot, Maristela Martins-Teixeira, Marcelo Marchiori, Emanuela Iaccarino, Menotti Ruvo, Marcelo Dias-Baruffi, and Ivone Carvalho

To be cited as: *ChemBioChem* 10.1002/cbic.201600473

Link to VoR: http://dx.doi.org/10.1002/cbic.201600473

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

Synthetic MUC1 glycopeptide bearing βGalNAc-Thr as Tn antigen isomer induces antibodies production against tumor cells

Vanessa Leiria Campo,[*1a] Thalita B. Riul,[a] Leandro Oliveira Bortot,[a] Maristela B. Martins-Teixeira,[a] Marcelo Fiori Marchiori,[a] Emanuela Iaccarino,[b,c] Menotti Ruvo,[b] Marcelo Dias-Baruffi[a] and Ivone Carvalho[a]

a. Faculdade de Ciências Farmacêuticas de Ribeirão Preto, USP, Av. Café S/N, CEP 14040-903, Ribeirão Preto, SP, Brazil, E-mail: vlcampo@fcfrp.usp.br

b. Istituto di Biostrutture e Bioimmagini, CNR, via Mezzocannone, 16, 80134, Napoli, Italy

c. Second University of Naples, via Vivaldi, 43, 81100, Caserta, Italy

Abstract

This study presents the synthesis of the novel protected *O*-glycosyl amino acids βGalNAc-SerOBn (**1** and βGalNAc-ThrOBn (**2**) as mimetics of the natural Tn antigen (αGalNAc-Ser/Thr), along with th solid-phase assembly of the glycopeptides NHAcSer-Ala-Pro-Asp-Thr[αGalNAc]-Arg-Pro-Ala-Pro-Gly BSA (**3**–BSA) and NHAcSer-Ala-Pro-Asp-Thr[βGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA (**4**–BSA), bearing αGalNAc-Thr or βGalNAc-Thr, as mimetics of MUC1 tumor mucin glycoproteins. According to ELISA tests, mice immunizations with βGalNAc-glycopeptide **4**–BSA induced higher sera titers (1:320000) i compared to immunizations with αGalNAc-glycopeptide **3**–BSA (1:40000). Likewise, flow cytometry assays showed higher capacity of the obtained anti-glycopeptide **4**–BSA antibodies to recognize MCF-7 tumor cells. Cross recognition between immunopurified anti-βGalNAc antibodies and αGalNAc glycopeptide and vice versa was also verified. Lastly, molecular dynamics simulations and surface plasmon resonance (SPR) showed that βGalNAc-glycopeptide **4** can interact with a model antitumo monoclonal antibody (SM3). Taken together, these data highlight the improved immunogenicity of th unnatural glycopeptide **4**–BSA, bearing βGalNAc-Thr as Tn antigen isomer.

Keywords

MUC1-Mucins; Tumor; Tn antigen; Glycopeptides; Antibodies

[1] Corresponding author. Tel.: +55 16 33154285; fax: +55 16 33154879; e-mail: vlcampo@fcfrp.usp.br

1. Introduction

Mucins are highly *O*-glycosylated (up to 80% of carbohydrates by weight) glycoproteins, containing typical tandem repeat motifs rich in serine (Ser) and threonine (Thr), which represent potential sites for abundant *O*-glycosylation. Characteristically, the sugar residue directly linked to these amino acids is 2-acetamido-2-deoxy-α-D-galactopyranose (α-D-GalNAc).[1, 2] Mucins are present on cell surfaces of many epithelial tissues, for instance, respiratory mucosa, reproductive system and gastrointestinal tract, among others, where they exert essential roles of protection, hydration and lubrication. Additionally, apart fron these primary functions, mucins are also involved in complex biological processes, such as cel differentiation, signaling, adhesion and migration, fertilization, infections and inflammation.[1, 3, 4] I tumor-related mucins (e.g MUC1, MUC4 and MUC16), however, the glycosylation process i deregulated, resulting in abnormal, incomplete and truncated glycans, where the exposure o overexpression of *O*-α-D-GalNAc residue, also denominated Tn antigen when linked to the amino aci Ser/Thr (αGalNAc-Ser/Thr), is a common feature.[3, 5, 6]

Among different classes of tumor mucins, MUC1 is the most extensively studied, as it is present o nearly all epithelial tissues, being strongly overexpressed on epithelial tumor cells, where it i characteristically changed if compared with the glycan profile of MUC1 on normal cells, particularly, i the extracellular domain.[3, 7, 8] Thus, a variable number (25–125) of tandem repeat sequence (HGVTSAPDTRPAPGSTAPPA) can be observed, containing five potential *O*-glycosylation (*O* αGalNAc) sites (Scheme 1A).[7, 9] Due to its functional properties of adhesion, invasion and metastasis besides association to several types of cancer, such as breast, prostate, lung and pancreas, Tn antigen i known as tumor associated carbohydrate antigen (TACA) and is considered a clinically relevant tumo marker.[10] Thus, Tn antigen up-regulation acts as a signature associated with malignancy and, since it occurrence in normal tissues is limited, it represents a potential target for development of synthetic antitumor vaccines.[6, 11]

Accordingly, new synthetic methods for obtaining Tn antigen as well as its counterpart analogs have been constantly investigated.[10] To date, several Tn antigen analogs have been reported for use as vaccine constructs, for instance, *C*-glycosides,[12, 13] *S*-glycosides,[14] deoxyfluoro sugars[15, 16] and Tn

glycopeptides mimics containing non-natural amino acids.[17, 18] *C*-glycosides and *S*-glycosides are more stable toward enzymatic and chemical hydrolysis than natural Tn, as they have the interglycosidic oxygen replaced by methylene group or sulfur, respectively.[12, 14] Deoxyfluoro Tn analogs possessing fluorine atom in substitution of a sugar hydroxy group are supposed to enhance both the immunogenicity and the bioavailability of the antigen, considering the absence of fluorine in most organisms and its high bond strength to carbon.[15] The use of non-natural amino acids, such as α-methylserine or those containing long aliphatic side chains, has proved to be effective for obtaining non-natural Tn peptide-linked vaccine with improved immunogenicity.[17, 18] The application of Cu(I)-assisted 1,3-dipolar azide-alkyn cycloaddition (CuAAC) reactions between propargylated α-GalNAc sugar and azido-functionalize amino acids, or vice-versa, represents another strategy for synthesis of triazole-derived Tn neoglycopeptide vaccines with advantageous physicochemical properties, inherent to the triazol group.[19]

However, despite the vast repertoire of Tn antigen (αGalNAc-Ser/Thr) and related glycopeptide analogs described so far, there are only few studies concerning the synthesis of its correspondin stereoisomer (βGalNAc-Ser/Thr).[20] Furthermore, comparative immunological assays involvin glycopeptides bearing the distinct isomers have not been described yet. As judged by its structura similarity to Tn antigen, differing only on the anomeric configuration, βGalNAc-Ser/Thr may be frontline building block for development of antitumor vaccines, depending on its efficiency for inducin active immunity. In this regard, the crystal structure of the antitumor monoclonal antibody SM3 (SM mAb), able to recognize the fundamental epitope of the MUC1 peptide (SAPDTRPAP) and th corresponding glycopeptide [SAPDT(*O*-αGalNAc)RPAP], may represent a model for theoretica investigations of the βGalNAc-Ser/Thr core as a novel Tn mimic.[18, 21]

Therefore, this paper presents the synthesis of the novel protected *O*-glycosyl amino acids βGalNAc-SerOBn (**1**) and βGalNAc-ThrOBn (**2**) as mimetics of the natural Tn antigen (αGalNAc-Ser/Thr) (Scheme 1B), along with the solid-phase assembly of the MUC1 glycopeptides NHAcSer-Ala-Pro-Asp-Thr[αGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA (**3**) and NHAcSer-Ala-Pro-Asp-Thr[βGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA (**4**), bearing αGalNAc-Thr or βGalNAc-Thr (Scheme 1C). The immunological

assays comparing the antibodies production against tumor cells induced by glycopeptides **3**-BSA and **4**-BSA, as well as their theoretical and experimental capacities to interact with a model antitumor monoclonal antibody (SM3) by corresponding molecular dynamic simulations and surface plasmon resonance (SPR) assays, will also be reported herein.

(SCHEME 1)

2. Results and Discussion

2.1. Synthesis

The target glycopeptides NHAcSer-Ala-Pro-Asp-Thr[αGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA **3** and NHAcSer-Ala-Pro-Asp-Thr[βGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA **4** were obtained by solid-phase synthesis, which involved the previous preparation of the glycosyl amino acids αGalNAc-ThrOH (Tn) (**5**) and βGalNAc-ThrOH (**6**) as building blocks for obtaining corresponding tumor MUC1 fundamental scaffold and its counterpart analog.

2.1.1. Synthesis of glycosyl amino acids

Owing to the polyfunctionality of carbohydrates and amino acids, the synthesis of glycosyl amino acids demands the utilization of specific protecting groups, which must be suitably selected considering the lability of the glycosidic bond under acidic conditions and the possibility of β-elimination reaction under strong basic conditions.[22] Moreover, specific stereoselective methods are essential for synthesis of distinct *O*-glycosyl amino acids isomers (α and β),[23] requiring the use of glycosyl donors containing non-participating or participating groups at C-2 for synthesis of 1,2-*cis* (αGalNAc-Ser/Thr) or 1,2-*trans* glycosides (βGalNAc-Ser/Thr), respectively.[24, 25] One of the most frequently employed strategy for synthesis of glycosyl-amino acids relies on Koenigs-Knorr method,[26] which involves the use of glycosyl halides, such as chlorides, bromides and iodides, as glycosyl donors, along with typically heavy metal salts as catalysts.[27, 28] For instance, silver salts (Ag_2CO_3/$AgClO_4$, AgOTf),[29] trimethylsilyl triflate (TMSOTf),[30] dimethyl(methylthio)sulfonium triflate (DMTST)[31] and mercury salts ($HgBr_2$, $HgCl_2$,

$Hg(CN)_2$)[32] are some representative catalysts used in Koenigs-Knorr type glycosylation reactions,[26] since they are effective to assist leaving group displacement by complexation with anomeric halides.

In this context, for synthesis of the glycosyl amino acid αGalNAc-ThrOH (Tn) **5** we made use of our established glycosylation method using $HgBr_2$, by which glycosylation of FmocThr benzyl ester (**7**) with the azido-derived glycosyl donor αGalN_3Cl (**8**),[33] unable to exhibit neighboring group effect, afforded exclusively the α isomer αGalN_3-ThrOBn (**9**) (78%) (Scheme 1S).[34] Subsequently, the sugar azido group of **9** was reductively acetylated using zinc powder in THF-acetic anhydride-acetic acid, and its benzyl ester group was removed by standard hydrogenation reaction (10% Pd-C/H_2, 1 h),[34] giving the αGalNAc-ThrOH **5** (60%) to be utilized in the solid-phase synthesis of glycopeptide **3** (Scheme 1S). Regarding the synthesis of β-linked GalNAc amino acids, Danishefsky *et al.* reported the synthesis of the tumor associated Lewis^y (Le^y) pentasaccharides bearing both αGalNAc- and βGalNAc-Ser (anomeric mixture α/β 3:1) by careful glycosylation conditions utilizing an azido-functionalized glycosyl donor and TMSOTf as catalyst, followed by conversion to acetamides; however, similar glycosylation reaction with threonine acceptor afforded only the α derivative.[35] In fact, our described $HgBr_2$ catalyzed glycosylation reaction of FmocSer benzyl ester (**10**) with αGalN_3Cl **8** led to the predicted α glycoside (50%), along with very low yield of βGalN_3-SerOBn (4%), whereas none of the azido threonine-derived β-anomer could be obtained by glycosylation of threonine acceptor **7** with the donor **8**.[33] Thus, the fact that glycosylation reactions utilizing the azido-donor **8** were ineffective to get both βGalNAc-Ser/Thr, prompted us to apply an alternative route using the glycosyl donor αGalNAcCl (**11**), possessing an acetamide C-2 participating group. Firstly, the synthesis of the donor **11** was carried out by peracetylation (Ac_2O/ Py) of 2-amino-2-deoxy-D-galactose hydrochloride (**12**), followed by treatment with acetyl chloride and HCl (g) for 18 h being obtained in 85% yield (Scheme 2).[25] This synthetic procedure differs from those previously described conditions that utilize the titanium tetrachloride ($TiCl_4$) as catalyst, affording **11** in similar yield.[36] Subsequently, glycosylation reactions of FmocSer benzyl ester **10** or FmocThr benzyl ester **7** with the obtained αGalNAcCl **11,** in the presence of 1,2-dichloroethane as solvent and $HgBr_2$ as catalyst, afforded exclusively the novel protected βGalNAc-SerOBn **1** (55%) or βGalNAc-ThrOBn **2** (52%) after purification by column chromatography, highlighting the applicability of $HgBr_2$ as a catalyst also for this

particular glycosyl donor **11** (Scheme 2). Noteworthy, these moderate yields are comparable to other related β-glycosyl amino acids described in the literature.[37] The structures of the β isomers **1** and **2** were confirmed by ^{1}H NMR spectroscopy, being identified characteristic doublets at δ 4.60 ($J_{1,2}$ 8.3 Hz) and δ 4.65 ($J_{1,2}$ 8.3 Hz), confirming their β configuration. ESI-MS analysis also showed characteristic adducts for **1** ($[M+H]^+$ 747.2762) and **2** ($[M+H]^+$ 761.2916). The glycosyl amino acid βGalNAc-ThrOBn **2** was next submitted to hydrogenation reaction for removal of its benzyl ester group, affording βGalNAc-ThrOH **6** (88%) as prompt intermediate for solid-phase synthesis of glycopeptide **4** (Scheme 2) Regarding glycosyl amino acid **1**, though this work did not involve its use for synthesis of glycopeptide **3** and **4**, it can be a valuable building block for future synthesis of other MUC1 glycopeptide analogs with distinct glycosylation patterns.

(SCHEME 2)

2.1.2. Solid-phase synthesis of glycopeptides 3 and 4

Moving towards the MUC1 glycopeptides **3** and **4**, solid-phase glycopeptide synthesis (SPPS) was employed, utilizing commercial Wang resin pre-loaded with Fmoc-Gly (0.65 mmol/g resin) as the solid support and following the orthogonal Fmoc-based SPPS.[34, 38] Thus, after removal of *N*-Fmoc groups with 50% morpholine in DMF, the amino acids building blocks were coupled in the presence of the coupling reagents PyBOP/HOBt and *N*,*N*-diisopropylethylamine (DIPEA), in DMF. Optimal coupling times ranged from 6 h to 24 h for the incorporation of single amino acids, depending on preceding peptide sequence. For incorporation of the more sterically and electronically demanding glycosyl amino acids **5** or **6** the coupling times typically ranged from 24 h to 30 h.

The synthesis of glycopeptides **3** and **4** was initiated by cleavage of the Fmoc group from Fmoc-Gly-Wang resin, followed by nine rounds of coupling-deprotection with sequential incorporation of Fmoc-ProOH, Fmoc-AlaOH, Fmoc-ProOH, Fmoc-Arg(Pbf)OH, αGalNAc-ThrOH **5** or βGalNAc-ThrOH **6**, Fmoc-Asp(OtBu)OH, Fmoc-ProOH, Fmoc-AlaOH and Fmoc-SerOH, thus completing the preparation of the protected, resin-bound glycopeptides (**13**) and (**14**) (Scheme 3). The obtained peptide coupling efficiencies, as judged by measuring the released dibenzofulvene from the product, varied with the

position in the peptide chain: Pro1 (70%), Ala1 (80%), Pro2 (75%), Arg (72%), αGalNAc-ThrOH **5** (52%), βGalNAc-ThrOH **6** (50%), Asp (60%), Pro3 (70%), Ala2 (85%) and Ser (80%).[28, 39] Subsequently, the obtained glycopeptides **13** and **14** were *N*-acetylated (Ac_2O/ pyridine) for capping of terminal NH_2 and then cleaved from resin by treatment with aqueous TFA (80%), with concomitant removal of Pbf and OtBu protecting groups of the corresponding Arg and Asp amino acids. Lastly, removal of the sugar acetate protecting groups with catalytic NaOMe in MeOH gave crude glycopeptides **3** and **4**, which were then analysed and purified by reversed-phase HPLC. As confirmed by NMR ^{1}H and ESI-MS analysis, the isolated peaks corresponded to glycopeptides **3** and **4**, along with residual glycopeptides containing Arg(Pbf) due to incomplete removal of this protecting group. Thus, NMR ^{1}H spectra of glycopeptides **3** (18%) and **4** (15%) showed for both compounds methylene and methyl signals of amino acids between 3.0-1.12 ppm, 2 × NH*Ac*, $H_{\beta Asp}$ signals between 2.82-2.70 ppm, as well as characteristic signals for sugar H-2 of **3** (3.95 ppm, dd, $J_{1,2}$ 3.6 Hz, $J_{2,3}$ 10.8 Hz), and **4** (3.75 ppm, dd, $J_{1,}$ 8.5 Hz, $J_{2,3}$ 10.9 Hz). ESI-MS analysis also showed characteristic adducts for **3** (observed $[M + 3NH_4]$ 1266.5413) and **4** (observed $[M + 3NH_4]^+$ 1266.5303).

Envisioning further use in immunization assays, glycopeptides **3** and **4** were conjugated to BSA carrier protein. Thus, the coupling reactions of **3** and **4** to BSA were performed in aqueous Na_2HPO buffer at pH 9.0, in the presence of the reagents EDCI and NHS (Scheme 3).[7, 40] The conjugated glycopeptides were purified by ultrafiltration against deionized water using a membrane of 3 kDa. The loading rates of the glycopeptides **3**–BSA and **4**–BSA conjugates were determined by MALDI-TOF mass spectrometry, being found one or two molecules of **3** (67500.6) and **4** (68827.6), respectively, linked to one BSA protein.

(SCHEME 3)

2.2. Immunological assays

2.2.1. ELISA assays

In order to verify the capacity of glycopeptides **3**–BSA and **4**–BSA conjugates to induce a humoral immune response, each synthetic glycopeptide (10 μg) in combination with complete Freund's adjuvant (CFA) were injected into different groups of four Balb/c mice. After an interval of two weeks, the second and third immunizations were performed with incomplete Freund's adjuvant (IFA). Prior to the first immunization and two weeks after the third immunization, blood was drawn from the eyelid vein of eacl mouse, and the obtained sera were used for analysis of the induced antibodies by enzyme-linke immunosorbent assay (ELISA).[41, 42] For comparative purposes, another group of four Balb/c mice wa immunized solely with BSA under the same conditions described for glycopeptides **3**–BSA and **4**–BSA with subsequent sera collection and analysis by ELISA. In the ELISA tests, the microtiter plates wer coated separately with glycopeptides **3**–BSA and **4**–BSA, as well as only with BSA as control. The ser were titrated at increasing dilutions (1:2500 a 1:2560000) in a phosphate buffered saline (PBS) solution The antibodies concentrations were detected photometrically (450 nm) by using a secondary anti-mous antibody conjugated to horseradish peroxidase (HRP).[43] The titers of the ELISA tests were determined t be approximately 1:40000 and 1:320000 for mice immunized with glycopeptides **3**–BSA and **4**–BSA respectively, and using these corresponding conjugates as coats. However, when coated exclusively witl BSA, lower corresponding titers of 1:20000 and 1:40000 were verified for mice immunized witl glycopeptides **3**–BSA and **4**–BSA, as expected. For immunizations with BSA and its use as coat the tite of the ELISA test was close to 1:160000 (Figure 1). Thus, according to the obtained results, the employe mice immunization protocol proved to be effective to induce the production of antibodies against th antigenic structures of glycopeptides **3**–BSA and **4**–BSA related to MUC1 tumor mucins. Moreover, i order to estimate the relative efficient immune reaction of anti-glycopeptide **3**–BSA and anti-glycopeptide **4**–BSA antibodies in the polyclonal sera from immunized mice, 1 μL of each serum was incubated with PBS containing 3% BSA (m/v) for 1 h at room temperature, prior to additional ELISA assays (as previously described). This step allowed the depletion of non-specific anti-BSA antibodies, resulting in a decreased OD for all tested sera ($p<0.05$). However, this reduction was much less pronounced for

glycopeptide **4**–BSA antisera (Figure S9), suggesting that the polyclonal serum from mice immunized with glycopeptide **4**–BSA is richer in specific antibodies that recognize glycopeptide **4** epitopes. In fact, this result added to the higher sera titers induced by the glycopeptide **4**–BSA, containing the glycosyl-amino acid βGalNAc-ThrOH **6**, if compared to the glycopeptide **3**–BSA bearing the natural tumor antigen αGalNAc-ThrOH **5** (Tn), highlights the increased immunological potential of this unnatural β-configured building block, translating into an advantageous aspect for development of an antitumor vaccine, since it is not present in normal and tumor cells found in mammals.

(FIGURE 1)

2.2.2. Flow cytometry assays

Flow cytometry assays were subsequently performed to evaluate the binding of the glycopeptide **3**–BSA and **4**–BSA induced antibodies to cells of epithelial breast tumor (MCF-7).[39, 42] Previous to thes assays, anti-MCF-7 sera was obtained for use as reference by immunization of four Balb/c mice wit MCF-7 cells, following the conditions cited above. Thus, MCF-7 cells were incubated with glycopeptide **3**–BSA and **4**–BSA induced antisera at increasing dilutions (1:400 a 1:8000), along with MCF-7 antiser (1:1000). After washing, secondary antimouse-IgG antibodies from goat conjugated to a fluorescent labe (AlexaFluor 488) were added to the cells, which were then acquired and analyzed by flow cytometry. Th cells recognized by the antibodies from the mouse antisera showed fluorescence and were counte separately (right area-P3). According to Figure 2A, cells treated with preimmune sera as negative contro did not show positive staining, appearing in the left area (P2). The MCF-7 cells that were incubated wit the corresponding MCF-7 antisera (1:1000) as positive control were all labeled and appeared in the righ area (Figures 2A, 3). When MCF-7 cells were incubated with glycopeptides **3**–BSA and **4**–BSA antisera they were somehow recognized by these antibodies, being verified for glycopeptide **3**–BSA antisera binding values (percentage of positive staining) of 37% (1:400 and 1:800), 15% (1:1000), 12% (1:2000), 8% (1:4000) and 6% (1:8000) (Figures 2B and 3), whereas glycopeptide **4**–BSA antisera displayed higher binding values at increased dilutions, as follows, 54% (1:400), 50% (1:800), 18% (1:1000), 16% (1:2000), 15% (1:4000) and 10% (1:8000) (Figures 2C and 3).

Therefore, it is possible to assert that the obtained flow cytometry data using MCF-7 cells are in accordance with ELISA results, strictly based on the synthetic glycopeptides **3**–BSA and **4**–BSA. In fact, the observed binding capacities of glycopeptides **3**–BSA and **4**–BSA antisera towards MCF-7, mainly at 1:400 and 1:800 dilutions, are very significative, considering the high complexity of the natural MUC1 tumor mucins expressed by MCF-7 cells, whose essential scaffold is being mimicked by glycopeptides **3** and **4**. Moreover, the higher binding to MCF-7 cells presented by glycopeptide **4**–BSA antisera, if compared to the glycopeptide **3**–BSA antisera, reinforces the improved immunogenicity of the glycopeptide **4**–BSA, containing the unnatural glycosyl-amino acid βGalNAc-ThrOH **6**.

(FIGURE 2)

(FIGURE 3)

2.2.3. Immunopurification of antisera and binding to glycopeptides 3 and 4

In order to better assess the recognition between αGalNAc- and βGalNAc-glycopeptides and their corresponding anti-glycopeptides **3** and **4** antibodies, and to evaluate their cross recognition capacity, antisera were immunopurified and evaluated by ELISA assays. Their immunopurifications were carried out by initial immobilization of BSA-free glycopeptides **3** (αGalNAc) and **4** (βGalNAc) on the Sepharose 4B resin through their C-terminus after introduction of an hexamethylenediamine moiety, being their couplings to the resin monitored via HPLC analysis of glycopeptides on the supernatants. By this method, estimated incorporation values of 75% (about 1.3 mg) and 72% (about 1.3 mg) were verified for the αGalNAc- and βGalNAc-glycopeptides, respectively. Anti-glycopeptides **3** and **4** antibodies were then captured by the corresponding immobilized glycopeptides. After purification about 154 µg (0.22 mg/mL, 700 µL) antibodies were recovered from the anti-α serum and about 91 µg (0.13 mg/mL, 700 µL) from the anti-β serum.

Subsequently, recognition between αGalNAc- and βGalNAc-glycopeptides and corresponding anti-glycopeptides **3** and **4** antibodies, as well as the cross recognition between αGalNAc-glycopeptide and anti-βGalNAc-glycopeptide antibodies and vice versa, were assessed by ELISA. In a preliminary

experiment performed by coating BSA-conjugated glycopeptides **3** and **4** the expected strong binding (about 1.5 absorbance) was observed between immunopurified antibodies and the respective antigens (Figure 4). Interestingly, also a remarkable recognition was observed between anti-αGalNAc antibodies and βGalNAc-glycopeptide, and between anti-βGalNAc antibodies and αGalNAc-glycopeptide (on average 1.0 and 0.8 absorbance, respectively), even at 0.1 μg/mL antibodies. These findings suggest that antibodies raised with the βGalNAc-glycopeptide **4** can be recognized by the αGalNAc variant; therefore, such antibodies have the potential to also recognize the natural Tn tumor antigen. Importantly, very weal binding was recorded for the immunopurified anti-αGalNAc and anti-βGalNAc antibodies to immobilize BSA, thus indicating that they selectively recognize the antigens and not the carrier used fo immunization (Figure 4). In addition, binding of immunopurified anti-αGalNAc and anti-βGalNA antibodies to asialo-submaxillary mucin, which commonly expresses the alpha-anomer, was also assesse by ELISA assays, being verified that both anti-αGalNAc and anti-βGalNAc antibodies bound to asialo submaxillary mucin (See Supplementary Information, Figure S10).

(FIGURE 4)

2.3. Molecular dynamics simulations and surface plasmon resonance (SPR) assays with a mode antitumor monoclonal antibody

Considering the better immunogenicity observed for the glycopeptide **4**–BSA, bearing the glycosyl amino acid βGalNAc-ThrOH **6**, we asked whether or not it could interact with a model anti-MUC1 mAb specifically SM3, since it is the only anti-MUC1 mAb for which there is a crystallographic structure.

Accordingly, molecular dynamics simulations were carried out using the crystallographic comple consisted of the SM3 mAb complexed with a nine-residue MUC-1 derived peptide (SAPDTRPAP) (Figure 5A),[21] from which two distinct SM3-peptide complexes differing only by the *O*-glycosylation type at the Thr residue (α or β GalNAc-Thr) were built, representing the glycopeptides **3** and **4** (Figure 5B-C).

(FIGURE 5)

According to molecular dynamics simulations, applying Chi1 dihedral angle of +60°, the interactions between the SM3 mAb and the glycopeptides **3** (α) and **4** (β) does not show major deviations from the crystallographic structure.[21] The carbohydrate moieties of both glycopeptides interact with SM3 via hydrogen bonds with Tyr-32L and Asn-53L (Figure 6A-B), suggesting that SM3 can also accommodate the β-configured glycopeptide. This finding prompted us to further investigate the SM3-αGalNAc **3** and SM3-βGalNAc **4** glycopeptides interactions experimentally.

(FIGURE 6)

Thus, SPR binding experiments of αGalNAc- and βGalNAc-glycopeptides were performed with immobilized SM3 mAb. Firstly, αGalNAc-glycopeptide **3** was injected at increasing concentration starting from 50 nM. As shown in Figure 7A, classical binding and dissociation curves were observed but no saturation was observed up to 2 μM. A rough estimation of the KD obtained by fitting the dissociation and association curves between 50 and 400 nM gave a value of about 52 nM. In relation to βGalNAc glycopeptide **4**, no binding was detected for concentrations below 500 nM (Figure 7B). Curves for the concentrations between 500 nM and 2 μM were very noisy and the maximum RU signal was about 25% that achieved with the αGalNAc-glycopeptide **3** at the same concentrations. A very rough estimation of the KD obtained by fitting dissociation and association curves at 500 nM, 1 μM and 2 μM gave a value of about 20 μM, which is quite similar to the reported binding affinity (KD 24 μM) between SM3 and the unglycosylated peptide.[21] In fact, this βGalNAc- KD value is higher than that estimated with the αGalNAc-glycopeptide variant, but it is still a representative binding to SM3 mAb.

(FIGURE 7)

3. Conclusions

In summary, we have synthesized the novel protected *O*-glycosyl amino acids βGalNAc-SerOBn **1** and βGalNAc-ThrOBn **2** as mimetics of the natural Tn antigen (αGalNAc-Ser/Thr), along with the solid-phase assembly of the glycopeptides NHAcSer-Ala-Pro-Asp-Thr[αGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA (**3**–BSA) and NHAcSer-Ala-Pro-Asp-Thr[βGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA (**4**–BSA), bearing αGalNAc-Thr or βGalNAc-Thr, as mimetics of MUC1 tumor mucin glycoproteins. In this regard, $HgBr_2$-

catalysed glycosylations, broadened applied to the glycosyl donor αGalNAcCl **11**, afforded exclusively the β-configured glycosyl amino acids **1** and **2**.

According to ELISA tests, mice immunizations with βGalNAc-glycopeptide **4**–BSA induced higher sera titers (1:320000) if compared to immunizations with αGalNAc-glycopeptide **3**–BSA (1:40000). Likewise, flow cytometry assays showed higher capacity of the obtained anti-glycopeptide **4**–BSA antibodies to recognize MCF-7 tumor cells, as judged by the corresponding binding values of 37% and 54% verified for glycopeptides **3**–BSA and **4**–BSA antisera at 1:400 dilution. Cross recognition betwee immunopurified anti-βGalNAc antibodies and αGalNAc-glycopeptide and vice versa was also verified Lastly, molecular dynamics simulations and surface plasmon resonance (SPR) showed that βGalNAc glycopeptide **4** can interact with a model antitumor monoclonal antibody (SM3). Overall, experimenta and theoretical data highlight the improved immunogenicity of the unnatural glycopeptide **4**–BSA bearing βGalNAc-Thr as Tn antigen isomer, representing a frontline glycopeptide for development o antitumor vaccines.

4. Experimental

4.1. General

All chemicals were purchased as reagent grade and used without further purification. Reaction were monitored by thin layer chromatography (TLC) on 0.25 nm precoated silica gel plates (Whatman AL SIL G/UV, aluminium backing) with the indicated eluents. Compounds were visualized under UV light (254 nm) and/or dipping in ethanol–sulfuric acid (95:5, v/v), followed by heating the plate for a few minutes. Column chromatography was performed on silica gel 60 (Fluorochem, 35–70 mesh). Nuclear magnetic resonance spectra were recorded on Bruker Advance DRX 300 (300 MHz), DPX 400 (400 MHz) or DPX 500 (500 MHz) spectrometers. Chemical shifts (δ) are given in parts per million downfield from tetramethylsilane. Assignments were made with the aid of HMQC and COSY experiments. Accurate mass electrospray ionization mass spectra (ESI-HRMS) were obtained using positive or negative ionization modes on a Bruker Daltonics MicroOTOF II ESI-TOF mass spectrometer. MALDI-TOF analyses were acquired using positive ionization mode on an UltrafleXtreme (Bruker, Billerica)

mass spectrometer. Absorbance of 96-well ELISA plates was measured at 450 nm using a spectrophotometer SpectraMax - Molecular Device. FACS analysis were performed using a FACSCanto II flow cytometer equipped with FACS Diva software (BD).

4.2. Synthesis

4.2.1. 2-Acetamide-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl chloride 11

2-Amino-2-deoxy-D-galactose hydrochloride **12** (200 mg, 0.928 mmol) was treated with Ac_2O (1.2 mL) and pyridine (2.0), and stirred at room temperature for 15 h. The obtained mixture was then extracted with 10% HCl and saturated $NaHCO_3$ and concentrated *in vacuo*, affording the peracetylated product 2-acetamide-1,3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl as a pale viscous oil (350 mg, 0.89 mmol 96%). The obtained compound (300 mg, 0.77 mmol) was treated with acetyl chloride (5 mL), previously saturated with HCl gas, and then stirred at 40°C for 18 h. The obtained amber mixture was diluted in dichloromethane, extracted with saturated $NaHCO_3$, dried over $MgSO_4$ and concentrated *in vacuo*. After purification by silica gel column chromatography (EtOAc/Hex 7:3, v/v), the chloride **11** was obtained as an yellowish solid in 85% yield (238.6 mg, 0.65 mmol). $\delta_{\mathbf{H}}$ ($CDCl_3$, 300 MHz) 6.26 (1H, d, $J_{1,2}$ 3.6 Hz H-1), 5.65 (1H, d, *J* 9.3 Hz, N*H*CO), 5.47 (1H, d, *J* 1.8 Hz, H-4), 5.29 (1H, dd, $J_{3,4}$ 3.3 Hz, $J_{2,3}$ 11.1 Hz H-3), 4.80 (1H, dt, $J_{1,2}$ 3.6 Hz, $J_{2,3}$ 11.2 Hz, H-2), 4.49 (1H, t, *J* 6.2 Hz, H-5), 4.19 (1H, dd, $J_{5,6a}$ 6.2 Hz $J_{6a,6b}$ 11.4 Hz, H-6a), 4.09 (1H, dd, $J_{5,6b}$ 6.7 Hz, $J_{6a,6b}$ 11.4 Hz, H-6b), 2.17, 2.06, 2.03, 2.0 (12H,4 s CH_3CO).

4.2.2. General procedure for glycosylation reaction

A mixture of 2-acetamide-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl chloride **11** (2 equiv.) and *N*-(fluoren-9-ylmethoxycarbonyl)-serine/threonine benzyl ester **7/10** (1 equiv.), (prepared from commercially available amino acid Fmoc-serine/threonine by treatment with caesium carbonate and benzyl bromide in DMF), in 1,2-dichloroethane (2.0 mL) was refluxed for 20 to 22 h with mercuric bromide (2.25 equiv.) and followed by TLC (EtOAc/hexane 7:3, v/v). The resulting amber mixture was

concentrated *in vacuo* and the residue was purified by a silica gel column chromatography (EtOAc/hexane 7:3, v/v).

_N_-(Fluoren-9-ylmethoxycarbonyl)-(2-acetamido-2-deoxy-3,4,6-tri-_O_-acetyl-β-D-galactopyranosyl)-L-serine benzyl ester 1

Following general glycosylation procedure, the reaction mixture of **11** (40 mg, 0.11 mmol) and **10** (22.5 mg, 0.054 mmol), in 1,2-dichloroethane (3.0 mL) was refluxed for 20 h with mercuric bromide (4(mg, 0.11 mmol), affording the β anomer **1** as an amorphous solid (22.4 mg, 0.029 mmol, 55%) afte purification by silica gel column chromatography (EtOAc/hexane 7:3, v/v). $[a]_D^{25}$ −23.2 (*c* 0.2 $CHCl_3$). δ_H ($CDCl_3$, 500 MHz) 7.70 (2H, d, *J* 7.0 Hz, C*H* Fmoc arom.), 7.59-7.54 (2H, m, C*H* Fmoc arom.), 7.35 7.24 (9H, m, C*H* Bn arom., C*H* Fmoc arom.), 5.75 (1H, d, *J* 8.0 Hz, N*H* Ser), 5.27 (1H, d, $J_{3,4}$ 2.9 Hz, H 4), 5.16-5.11 (3H, m, CH_2 Bn, H-3), 4.60 (1H, d, $J_{1,2}$ 8.3 Hz, H-1), 4.49-4.44 (1H, m, C*H* Ser), 4.42-4.3 (2H, m, CH_2 Fmoc), 4.21 (1H, dd, *J* 3.4 Hz, *J* 10.6 Hz, CH_2a Ser), 4.18-4.13 (1H, m, C*H* Fmoc), 4.07 4.01 (2H, m, H-6a, H-6b), 3.98-3.93 (1H, m, CH_2b Ser), 3.84-3.81 (1H, m, H-2), 3.76 (1H, t, *J* 6,5 Hz, H 5), 2.09, 2.07, 1.95 (12H, 3s, COCH_3). δ_C ($CDCl_3$, 125 MHz) 129.0, 128.6, 128.2, 127.5, 125.5, 120.4 (*C*H Fmoc arom., *C*H Bn arom.), 101.4 (C-1), 70.9, 69.8, 69.0 (C-5, C-3, CH_2 Ser), 67.7, 67.2, 66.7 (*C*H_2Bn, *C*H_2 Fmoc, C-4), 61.3 (C-6), 54.5 (CH Ser), 51.7 (C-2), 47.2 (*C*H Fmoc), 23.1 (NHCO*C*H_3) 20.8 (OCO*C*H_3). ESI-HRMS: calculated for $C_{39}H_{43}N_2O_{13}$ $[M+H]^+$ 747.2760, found 747.2762.

_N_-(Fluoren-9-ylmethoxycarbonyl)-(2-acetamido-2-deoxy-3,4,6-tri-_O_-acetyl-β-D-galactopyranosyl)-L-threonine benzyl ester 2

Following general glycosylation procedure, the reaction mixture of **11** (111 mg, 0.30 mmol) and ' (64.6 mg, 0.15 mmol), in 1,2-dichloroethane (4.0 mL) was refluxed for 22 h with mercuric bromide (109.8 mg, 0.30 mmol), affording the β anomer **2** as an amorphous solid (60.1 mg, 0.078 mmol, 52%) after purification by silica gel column chromatography (EtOAc/hexane 7:3, v/v). $[a]_D^{25}$ −19.5 (*c* 0.2 $CHCl_3$). δ_H ($CDCl_3$, 500 MHz) 7.69 (2H, d, *J* 7.5 Hz, C*H* Fmoc arom.), 7.57 (2H, d, *J* 7.3 Hz, C*H* Fmoc arom.), 7.37-7.20 (9H, m, C*H* Bn arom., C*H* Fmoc arom.), 5.80 (1H, d, *J* 9.1 Hz, N*H* Thr), 5.57 (1H, d, *J* 8.1

Hz, N*H*Ac), 5.23 (1H, d, *J* 3.1 Hz, H-4), 5.16 (1H, m, H-3), 5.13 (2H, AB, J_{AB} 12.5 Hz, CH_2 Bn), 4.57 (1H, d, $J_{1,2}$ 8.3 Hz, H-1), 4.42-4.25 (4H, m, βC*H* Thr, αC*H* Thr, CH_2 Fmoc), 4.17 (1H, t, *J* 7.0 Hz, C*H* Fmoc), 4.0 (1H, dd, $J_{5,6}$ 7.5 Hz, $J_{6a,6b}$ 11.4 Hz, H-6a), 3.95 (1H, dd, $J_{5,6}$ 6.5 Hz, $J_{6a,6b}$ 11.2 Hz, H-6b), 3.80 (1H, dd, $J_{1,2}$ 8.4 Hz, $J_{2,3}$10.6 Hz, H-2), 3.67 (1H, t, *J* 6.5 Hz, H-5), 2.06, 1.96, 1.93, 1.88 (12H, 4s, COCH_3), 1.13 (3H, d, *J* 6.2 Hz, CH_3 Thr). δ_C ($CDCl_3$, 125 MHz) 128.7, 128.3, 127.8, 127.8, 127.2, 125.2, 120.0 (*C*H Fmoc arom., *C*H Bn arom.), 99.0 (C-1), 74.8 (β*C*H Thr), 70.6 (C-5), 69.0 (C-3), 67.4 (*C*H_2Bn), 66.9 (*C*H_2 Fmoc), 66.5 (C-4), 61.1 (C-6), 58.7 (α*C*H Thr), 51.8 (C-2), 46.9 (*C*H Fmoc), 22.7 (NCO*C*H_3), 20.7, 20.6, 20.5 (OCO*C*H_3), 17.3 (*C*H_3 Thr). ESI-HRMS: calculated for $C_{40}H_{45}N_2O_{13}$ $[M+H]^+$ 761.2916, found 761.2916.

_N_-(Fluoren-9-ylmethoxycarbonyl)-(2-acetamido-2-deoxy-3,4,6-tri-_O_-acetyl-β-D-galactopyranosyl)-L-threonine 6

A solution of *N*-(fluoren-9-ylmethoxycarbonyl)-*O*-(2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-β-D galactopyranosyl)-L threonine benzyl ester **2** (60.1 mg, 0.078 mmol) in MeOH (2.5 mL) was treated with glacial AcOH (0.25 mL) and 10% Pd/C (30 mg), being stirred under H_2 (~1.5 atm) for 1 h. The reaction mixture was then filtered through Celite, concentrated *in vacuo* and purified by column chromatography (DCM-MeOH 9:1 v/v), affording the product **6** as an amorphous solid (48 mg, 0.078 mmol, 88%). δ_H (CD_3OD, 400 MHz) 7.69 (2H, d, *J* 7.5, C*H* Fmoc arom.), 7.58 (2H, t, *J* 7.1, C*H* Fmoc arom.), 7.30-7.19 (4H, m, C*H* Fmoc arom.), 5.22 (1H, d, *J* 2.5 Hz, H-4), 4.99 (1H, dd, $J_{3,4}$ 3.1, $J_{2,3}$ 11.2, H-3), 4.52 (1H, d, 8.3 Hz, H-1), 4.30-4.21 (3H, m, βC*H* Thr, αC*H* Thr, H-2), 4.13 (1H, t, *J* 7.0 Hz, C*H* Fmoc), 4.09-3.90 (4H CH_2 Fmoc, H-6a, H-6b, H-5), 1.97, 1.91, 1.87, 1.84 (12H, 4s, COCH_3), 1.09 (3H, d, *J* 6.0 Hz, CH_3 Thr). δ_C ($CDCl_3$, 100 MHz) 126.3-119.1 (*C*H Fmoc arom., *C*H Bn arom.), 99.5 (C-1), 75.7 (β*C*H Thr), 70.7 (C-3) 69.8(C-5), 66.8 (*C*H_2 Fmoc), 66.6 (C-4), 66.2 (C-2), 60.9 (C-6), 58.7 (α*C*H Thr), 46.7 (*C*H Fmoc), 21.3 19.4 (OCO*C*H_3), 16.4 (*C*H_3 Thr). ESI-HRMS: calculated for $C_{33}H_{38}N_2O_{13}Na^+$ $[M+Na]^+$ 693.2266, found 693.2269.

4.2.3. General method for glycopeptide synthesis

The glycopeptides **3** and **4** were assembled manually using a fritted glass reaction vessel with nitrogen purging for effective mixing. Pre-loaded FmocGly-Wang resin (typically 100 mg) was swollen

in CH_2Cl_2 for 2 h and then washed with DMF (3×). Deprotection of the *N*-α-Fmoc group was carried out using 50% morpholine–DMF followed by filtration and washing with DMF (3x). Coupling reactions were performed with 2.0 mol equiv of Fmoc-amino acid and coupling agents (PyBOP and HOBt) dissolved in DMF, and the base DIPEA. Coupling times were variable, ranging from 6 h to 24 h for the amino acids and between 24 h and 30 h for the glycosyl-amino acids **5** and **6**. The reaction mixtures were filtered after each coupling and the resin washed three times with DMF, CH_2Cl_2 and MeOH. After drying *in vacuo*, small aliquots of resin (1 mg) were treated with 20% morpholine in DMF for removal of *N*Fmoc group and consequent generation of dibenzofulvene product which absorb UV strongly (290 nm), offering potential for monitoring of coupling reactions by spectrophotometer. After removal of the *N*Fmoc group of the last amino acid, the glycopeptides were *N*-acetylated with Py and Ac_2O and then cleaved from resin in the presence of aqueous 80% TFA. After cleavage was complete, the solutions were filtered to remove the cleaved resin, concentrated and then treated with 1M sodium methoxide solution in MeOH for removal of the sugar acetate protecting groups. The purification of the obtained glycopeptides **3** and **4** was carried out on Shimadzu Shim-PaK HPLC equipment using a C18 reversed-phase column Shim PacK CLC-ODS (M) column (250 x 4.6 mm), under gradient conditions (A: 0.1% TFA/ H_2O, B: MeCN 0-70% B, within 30 min.), at a flow rate of 1.0 mL $min.^{-1}$, with UV detection (224 nm). Under these conditions the retention times for glycopeptides **3** and **4** were 25.8 min and 25.3 min, respectively.

NHAcSer-Ala-Pro-Asp-Thr[αGalNAc]-Arg-Pro-Ala-Pro-Gly-OH 3

Using the standard peptide synthesis procedure outlined, followed by purification by reversed-phase HPLC, glycopeptide **3** was obtained as an amorphous solid (15.0 mg, 0.012 mmol, 18 %). tR (224 nm) = 25.8 min; $\delta_{\mathbf{H}}$ (D_2O, 400 MHz) 4.52-4.48 (2H, m, $H_{\alpha Thr,}$ $H_{\alpha Ser}$), 4.41-4.38 (2H, m, $H_{\alpha Pro}$, $H_{\alpha Arg),}$ 4.30 (2H m, $2H_{\alpha Ala),}$ 4.20 (1H, m, $H_{\beta Thr}$), 3.95 (1H, dd, $J_{1,2}$ 3.6 Hz, $J_{2,3}$ 10.8 Hz, H-2), 3.87 (1H, t, *J* 6.2 Hz, H-5), 3.83-3.61 (11H, m, 2H-6, H-3, H-4, $3H_{\delta Pro,}$ $CH_{2Gly,}$ $2H_{\beta Ser}$), 3.54-3.50 (3H, m, $H_{\delta Pro}$), 3.07 (1H, t, *J* 7.3 Hz, $H_{\delta Arg}$), 2.82 (1H, dd, *J* 6.8 Hz, *J* 17.0 Hz, $H_{\beta Asp}$), 2.70 (1H, dd, *J* 6.7 Hz, *J* 17.0 Hz, $H_{\beta Asp}$), 2.18-2.10 (3H, m, $H_{\beta Pro}$),1.95, 1.93 (6H, 2 s, 2 x NHCOCH_3), 1.84-1.69 (6H, m, $H_{\beta Pro}$, $H_{\gamma Pro),}$ 1.35 (4H, d, *J* 5.5 Hz, $2H_{\beta Arg}$, $2H_{\gamma Arg}$), 1.25 (6H, d, *J* 7.3 Hz, CH_3 Ala), 1.12 (3H, d, *J* 6.2 Hz, CH_3 Thr). ESI-HRMS: calcd. for

$C_{50}H_{80}N_{14}O_{21}$ $3NH_4^+$ $[M + 3NH_4]^+$: 1266.6637, found: 1266.5413. Residual glycopeptide containing Arg(Pbf) - ESI-HRMS: calcd. for $C_{63}H_{96}N_{14}O_{24}S$ $[M + Na]^+$: 1487.6335, found: 1487.6091.

NHAcSer-Ala-Pro-Asp-Thr[βGalNAc]-Arg-Pro-Ala-Pro-Gly-OH 4

Using the standard peptide synthesis procedure outlined, followed by purification by reversed-phase HPLC, glycopeptide **4** was obtained as an amorphous solid (13.0 mg, 0.010 mmol, 15 %). tR (224 nm) = 25.3 min; δ_H (D_2O, 400 MHz) 4.47-4.39 (2H, m, $H_{\alpha Thr,}$ $H_{\alpha Ser}$), 4.35-4.33 (2H, m, $H_{\alpha Pro}$, $H_{\alpha Arg),}$ 4.31-4.28 (2H, m, $2H_{\alpha Ala),}$ 4.13-4.10 (1H, m, $H_{\beta Thr}$), , 3.88-3.77 (4H, H-5, $3H_{\delta Pro}$), 3.75 (dd, $J_{1,2}$ 8.5 Hz, $J_{2,3}$ 10.9 Hz H-2), 3.70-3.56 (8H, 2H-6, H-3, H-4, , $CH_{2Gly,}$ $2H_{\beta Ser}$), 3.53-3.49 (3H, m, $H_{\delta Pro}$), 3.09 (1H, t, *J* 7.3 Hz $H_{\delta Arg}$), 2.81 (1H, dd, *J* 6.9 Hz, *J* 16.8 Hz, $H_{\beta Asp}$), 2.73 (1H, dd, *J* 7.3 Hz, *J* 17.1 Hz, $H_{\beta Asp}$), 2.18-2.10 (3H m, $H_{\beta Pro}$), 1.96, 1.93 (6H, 2 s, 2 x NHCO*CH*$_3$), 1.83-1.69 (6H, m, $H_{\beta Pro}$, $H_{\gamma Pro}$), 1.35 (4H, d, *J* 5.5 Hz $2H_{\beta Arg}$, $2H_{\gamma Arg}$), 1.25 (6H, m, CH_3 Ala), 1.00 (3H, d, *J* 6.3 Hz, CH_3 Thr). ESI-HRMS: calcd. fo $C_{50}H_{80}N_{14}O_{21}$ $3NH_4^+$ $[M + 3NH_4]^+$: 1266.6637, found: 1266.5303. Residual glycopeptide containing Arg(Pbf) - ESI-HRMS: calcd. for $C_{63}H_{96}N_{14}O_{24}S$ $[M + Na]^+$: 1487.6335, found: 1488.6071.

4.2.4. General procedure for coupling of glycopeptides 3 and 4 to BSA

Glycopeptides **3** and **4** (2.5 mg, 2 μmol, 45 eq.) were dissolved in water (200 μL) and treated with the coupling reagents EDCI (380 μg, 2 μmol) and HOBt (270 μg, 2 μmol). The reaction mixtures were allowed to stand at room temperature with occasional gentle mixing and after 50 min they were added to a solution of BSA (1.5 mg) in Na_2HPO_4 buffer (20 mM, pH 9.0, 200 μL), previously precooled to 4°C. The solutions were then reacted for 48 h at room temperature and subsequently purified by ultrafiltration (AMICON ULTRA-0.5-MWCO 3 kDa). After lyophilization the obtained conjugates (approximately 7.0 mg) were analyzed by MALDI-TOF MS, being verified for glycopeptides **3-** and **4-**BSA corresponding peaks of mass 67500.6 and 68827.6, consistent with 1 and 2 epitopes per BSA.

4.3. Immunological assays

4.3.1. Mice Immunization

A solution of 10 μg of BSA (control) or glycopeptides **3** and **4**–BSA conjugates in PBS, in combination with complete Freund's adjuvant (CFA), was injected into 4 different Balb/c mice. After an interval of two weeks, the second and third immunizations were performed utilizing an emulsion with incomplete Freund's adjuvant (IFA). The immunization protocol for MCF-7 cells was similar to that used for glycopeptides **3** and **4**–BSA, with utilization of 1 x 10^6 of lysed MCF-7cells, obtained by extraction with cell lysis buffer. Preimmune sera were collected from all mice previously to immunization.

4.3.2. Enzyme-Linked Immunosorbent Assay (ELISA)

96-Well ELISA plates were coated with glycopeptides **3** or **4**–BSA in phosphate buffer solution (PBS), or only with BSA (control), and incubated overnight at 4 °C. The plates were then washed with PBS buffer and blocked with 3% gelatin in 0.05% Tween PBS buffer for 2 h at 37 °C. After washing with 0.05% Tween PBS buffer, the plates were incubated with different dilutions of preimmune and immune sera from all mice (ranging from 1:2500 to 1:2560000) in 0,05% Tween PBS for 2 h at 37 °C. Subsequently, the wells were washed and incubated with horseradish peroxidase conjugated to rabbit anti-mouse IgG Ab for 1 h at 37 °C. The plates were then washed and the substrate solution of 3,3′,5,5′ Tetramethylbenzidine (TMB) and H_2O_2 were added. After 10 min the solution of 2N sulfuric acid was added and the absorbance of each well was measured at 450 nm using a spectrophotometer (SpectraMax Molecular Device).

4.3.3. Flow cytometry analysis

Freshly harvested MCF-7 cells were washed with cold PBS and incubated with F_cR blocking solution for 1 hour in ice bath. Then, cells were washed thrice (300 x *g*, 5 min.), and separated in flow cytometry tubes (1x10^6 cells per tube) for subsequent incubation with different dilutions of preimmune and immune sera from mice immunized with glycopeptide **3**- and **4**-BSA (ranging from 1:400 to 1:8000) in cold PBS and ice bath. Control cells were incubated with sera from those mice immunized with MCF-7

extract. After 1 h, cells were washed three times and goat antimouse-IgG antibodies labeled with fluorescent dye (AlexaFluor 488) were added to the cells. After staining, cells were analysed by flow cytometry with a FACSCanto flow cytometer equipped with FACSDiva software (BD Biosciences).

4.3.4. Immunopurification of antisera and binding to glycopeptides 3 and 4

Immobilization of peptides on Sepharose 4B resin

αGalNAc-glycopeptide **3** (1.7 mg, 1.4 μmol) and βGalNAc-glycopeptide **4** (1.8 mg, 1.5 μmol) were coupled to an NHS-activated Sepharose 4B Fast flow resin (NHS: N-hydroxy-succinimmide) to perform immunopurification of anti-glycopeptides **3** and **4** antibodies. The two glycopeptides had no free amine for direct coupling, but they contained one C-terminal carboxyl group. We therefore pre-treated the resin in order to introduce an amino group and then conjugate glycopeptides **3** and **4** through a canonical EDC/NHS chemistry. The resin had a nominal substitution of 16 to 23 μmol NHS/mL drained medium. Two 200 μL resin aliquots (about 3 – 8 μmol NHS groups) were washed twice with cold 1 mM HCl and then with cold water. Solvents were removed by centrifugation. Resins were then treated with hexamethylenediamine by suspending them in 500 μL of 1.0 M reagent in 50 mM phosphate buffer pH 8.0 and left under stirring overnight at 4 °C. Resins were then washed extensively with phosphate buffer. Glycopeptides **3** and **4** were preactivated by dissolution in 600 μL of 0.05M NHS/0.2M EDC in water. After 30 minutes solutions were transferred onto the two separate resins. Couplings of the NHS preactivated glycopeptides to the amine groups on the resin were allowed to proceed overnight and were monitored by RP-HPLC by taking resin samples (50 μL) at t=0 and after the overnight coupling. For the HPLC analysis, resin samples were centrifuged at 13000g and the supernatants analyzed on a C18 monolithic ONYX column 50 x 2 mm ID operated at 0.6 mL/min, applying a gradient of CH_3CN, 0.1% TFA from 1% to 70% in 15 minutes. The resin was finally washed twice with NaAc 0.1 M pH 5.0 and PBS pH 7.5.

Antibodies Immunopurification

Resins were equilibrated in PBS pH 7.5. Anti-αGalNAc-glycopeptide **3** and anti-βGalNAc-glycopeptide **4** sera (500 μL) were incubated with the corresponding resin-immobilized glycopeptides for 16 h under gentle shaking at room temperature. Supernatants were then recovered and resins were washed

5 times with PBS pH 7.5. Resins were then treated with 350 µL 10 mM NaOH for 10 minutes under gentle shaking. Supernatants were removed by centrifugation and immediately neutralized with 2M TRIS pH 7.0. The presence of antibodies was assessed by SDS-PAGE analysis of the bound fraction. Protein concentration was assessed by the Bradford method.

ELISA assays

ELISA was used to monitor the binding of immunopurified antibodies to the corresponding immunogens (αGalNAc-glycopeptide **3** and βGalNAc-glycopeptide **4**) and cross recognition of anti αGalNAc-glycopeptide antibodies to βGalNAc-glycopeptide and viceversa. The glycopeptides **3** and 4 were used as BSA-conjugates to improve plate adsorption. Initial binding assays were performed by coating BSA-glycopeptides **3** and **4** on microtiter plates at 10 µg/mL, 1 µg/mL and 0.1 µg/mL dissolved in PBS (100 µL, overnight, triplicate wells). To exclude the immunopurified antibodies that recognize BSA used as carrier for immunization, wells were coated with 10 µg/mL BSA and probed with the two different anti-glycopeptide **3** and anti-glycopeptide **4** antibodies. After extensive washing with PBS, wells were blocked by incubation with 300 µL of 3% BSA w/v for 3h at room temperature. Immunopurified anti-αGalNAc and βGalNAc-glycopeptide antibodies were then dispensed in the wells both at 1 µg/ml and incubated for 1h at 37 °C (100 µL). After washing, wells were treated with goat anti-mouse-HRP 1:1000 (1 µg/mL) for 1h at 37 °C (100 µL). After washing, wells were treated with o-phenylenediamine dihydrochloride 0.4 mg/mL in 0.1 M citric acid and 0.2 M Na_2HPO_4 buffer (pH 4.8) and adding 0.4 µL/mL of 30% H_2O_2 as substrate (100 µL). The chromogenic reaction was stopped adding 50 µL 2.5 M H_2SO_4 and absorbance at 490 nm was read using an EnSpire multiwell reader (Perkin Elmer). The obtained data were averaged, blanks subtracted and plotted using GraphPad vers. 4.

4.4. Surface plasmon resonance assays (SPR)

SPR assays were performed using the anti-MUC1 SM3 monoclonal antibody (ABCAM, concentration 1 mg/mL) and the two synthetic glycopeptides **3** and **4**. All binding assays were performed on a Biacore 3000 instrument (GE Healthcare). A CM5 sensor chip was functionalized with the anti-MUC1 SM3 monoclonal antibody running the Wizard method and applying a standard amine coupling chemistry.[44] Firstly, the sensor chip surface was activated with EDC/NHS. Then SM3 at 10 µg/mL was

injected (70 μL) in NaAc 10 mM pH 4.0. When ligand immobilization was complete, 1M ethanolamine, pH 8.5 was passed over the chip. After a washing pulse with NaOH 10 mM and baseline equilibration, we estimated an immobilization level of about 7000 Response Units (RUs). The reference channel was deactivated by treatment with 1 M ethanolamine after the NHS/EDC activation. Signals from this channel were subtracted from those of the sample channel. All binding assays were conducted at 25 °C, at a constant flow rate of 20 μL/min using HBS-EP as running buffer. The samples of glycopeptides **3** and **4** were prepared in HBS-EP buffer and a total volume of 60μL of each analyte solution was injected onto the surface at increasing concentrations providing a dose-response signal. Overall binding events were monitored for 550 sec; the association step for 180 sec while the dissociation step for 370 sec and after each injection the chip surface was regenerated with NaOH using a range of concentrations comprised between 2 μM and 0.05 μM. Fitting of binding data was performed by using BIAevaluation software (version 4.1), adopting the 1:1 Langmuir binding model to extrapolate kinetic and thermodynamic parameters. ka stands for association constant, kd stands for dissociation constant and KD stands for apparent affinity constant. The final dissociation constant value (KD) reported for each analyte was the average between the three KD values extrapolated from every channel.

4.5. Molecular dynamics simulations

All simulations were performed with the GROMACS 5.0.4 simulation package[45] using the AMBER99SB-ILDN forcefield[46] combined with the Glycam06 parameters for carbohydrates and glycoproteins[47] and TIP3P water model using PBC in a triclinic box in which the ionic strength was set to 150 mM with Na^+ and Cl^-. The time-step value was 2 fs with water and non-water bonds constrained by the LINCS and SETTLE algorithms, respectively. All intermolecular interactions were treated explicitly inside 1 nm cutoffs and the long-range electrostatic interactions were treated by PME.

The initial states of systems were according to the crystallographic structure of the SM3 antibody complexing the MUC1 9-mer peptide (PDB ID 1SM3).[21] The glycosylated peptides were built with the GLYCAM-Web server (http://glycam.org/). The dihedral angle χ1 of the glycosylated Thr residue was +60° as in the crystallographic structure[21] and other published reports. [20, 48]

The fully solvated, electrically neutral systems at the desired ionic strength were energy minimized

using the steepest descent algorithm. During the equilibration phase, temperature was set to 309 K by the v-rescale thermostat and pressure to 1 bar by the Berendsen barostat. Temperature and pressure were kept at the same values during production runs by the Nosé-Hoover thermostat and Parrinello-Rahman barostat. Total simulation time was 1 ns for the equilibration phase and 100 ns for the production phase for each system.

Images were generated with PyMOL. The hardware used for preparative and equilibration steps of the molecular dynamics simulations was composed by virtual machines hosted at the University of São Paulo cloud computing infrastructure. A total of 4 virtual machines, each with 8 Intel Xeon cores at 2.40 GHz and 32 GB RAM, were used. The production step were performed with the DAVinCI supercomputer located at Rice University (Texas-USA), through a scientific collaboration agreement with the University of São Paulo (São Paulo-Brazil). For each simulation, 36 processing cores at 2.83 GHz and 6 nVIDIA Tesla M2050 GPUs were used.

Supporting Information Available: The Supporting Information encloses the ^{1}H NMR, ESI-MS, MALDI-TOF analyses and chromatograms of the main compounds, as well as complementary figures related to biological assays.

Acknowledgements

We acknowledge the financial support and fellowships from FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo – Brazil, Process N° 2012/19390-0) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). We are grateful to Dra. Lílian Cataldi Rodrigues for technical assistance in some biological assays, to Vinícius Palaretti for NMR analyses, to José Carlos Tomaz for ESI-MS analyses, to José César Rosa for MALDI-TOF analyses, to Prof. Antonio Caliri for discussions about the molecular dynamics simulations, and to the Data Analysis and Visualization CyberInfrastructure funded by NSF (grant OCI-0959097) for supporting the molecular dynamics studies.

References

[1] M. Andrianifahanana, N. Moniaux, S. K. Batra, *Biochimica Et Biophysica Acta-Reviews on Cancer* **2006**, *1765*, 189-222.
[2] C. A. Buscaglia, V. A. Campo, A. C. C. Frasch, J. M. Di Noia, *Nature Reviews Microbiology* **2006**, *4*, 229-236.
[3] H. Cai, Z.-Y. Sun, Z.-H. Huang, L. Shi, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Chemistry-a European Journal* **2013**, *19*, 1962-1970.
[4] K. Agarwal, R. Kaul, M. Garg, A. Shajahan, S. K. Jha, S.-G. Sampathkumar, *Journal of the American Chemical Society* **2013**, *135*, 14189-14197.
[5] E. Osinaga, *Iubmb Life* **2007**, *59*, 269-273; C. Brocke, H. Kunz, *Bioorganic & Medicinal Chemistry* **2002**, *10*, 3085-3112.
[6] D. J. Gill, T. Keit Min, J. Chia, S. C. Wang, C. Steentoft, H. Clausen, E. A. Bard-Chapeau, F. A. Bard, *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, E3152-E3161.
[7] U. Westerlind, A. Hobel, N. Gaidzik, E. Schmitt, H. Kunz, *Angewandte Chemie-International Edition* **2008**, *47*, 7551 7556.
[8] G. Rivalland, B. Loveland, P. Mitchell, *Expert Opinion on Biological Therapy* **2015**, *15*, 1773-1787.
[9] S. Dziadek, D. Kowalczyk, H. Kunz, *Angewandte Chemie-International Edition* **2005**, *44*, 7624-7630.
[10] D. Y. Feng, A. S. Shaikh, F. S. Wang, *Acs Chemical Biology* **2016**, *11*, 850-863.
[11] A. Holeman, P. H. Seeberger, *Current Opinion in Biotechnology* **2004**, *15*, 615-622.
[12] D. Urban, T. Skrydstrup, J. M. Beau, *Chemical Communications* **1998**, 955-956.
[13] F. Peri, L. Cipolla, M. Rescigno, B. La Ferla, F. Nicotra, *Bioconjugate Chemistry* **2001**, *12*, 325-328.
[14] D. A. Thayer, H. N. Yu, M. C. Galan, C. H. Wong, *Angewandte Chemie-International Edition* **2005**, *44*, 4596-4599.
[15] A. Hoffmann-Roeder, M. Johannesy, *Chemical Communications* **2011**, *47*, 9903-9905.
[16] S. Wagner, C. Mersch, A. Hoffmann-Roeder, *Chemistry-a European Journal* **2010**, *16*, 7319-7330; M. Johannes, T. Oberbillig, A. Hoffmann-Roeder, *Organic & Biomolecular Chemistry* **2011**, *9*, 5541-5546.
[17] S. J. Keding, A. Endo, S. J. Danishefsky, *Tetrahedron* **2003**, *59*, 7023-7031.
[18] F. Corzana, J. H. Busto, F. Marcelo, M. Garcia de Luis, J. L. Asensio, S. Martin-Santamaria, Y. Saenz, C. Torres, J. Jimenez-Barbero, A. Avenoza, J. M. Peregrina, *Chemical Communications* **2011**, *47*, 5319-5321.
[19] Q. Wan, J. Chen, G. Chen, S. J. Danishefsky, *Journal of Organic Chemistry* **2006**, *71*, 8244-8249; N. Miller, G. M. Williams, M. A. Brimble, *Organic Letters* **2009**, *11*, 2409-2412.
[20] F. Corzana, J. H. Busto, G. Jimenez-Oses, M. G. de Luis, J. L. Asensio, J. Jimenez-Barbero, J. M. Peregrina, A. Avenoza, *Journal of the American Chemical Society* **2007**, *129*, 9458-9467.
[21] P. Dokurno, P. A. Bates, H. A. Band, L. M. D. Stewart, J. M. Lally, J. M. Burchell, J. Taylor-Papadimitriou, D. Snary M. J. E. Sternberg, P. S. Freemont, *Journal of Molecular Biology* **1998**, *284*, 713-728.
[22] C.-H. Wong, *Carbohydrate-based drug discovery*, Wiley-VCH Verlag GmbH & Co. KgaA, Weinheim, **2003**; G. Arsequell, G. Valencia, *Tetrahedron-Asymmetry* **1997**, *8*, 2839-2876; Wakabaya.K, W. Pigman, *Carbohydrate Research* **1974**, *35*, 3-14.
[23] W. P. Stockl, H. Weidmann, *Journal of Carbohydrate Chemistry* **1989**, *8*, 169-198.
[24] H. Paulsen, C. Kolar, W. Stenzel, *Chemische Berichte-Recueil* **1978**, *111*, 2358-2369.
[25] R. L. B. J. N. Whistler, *Methods in carbohydrate chemistry, Vol. 6*, Academic, New York, **1972**.
[26] W. Koenigs, E. Knorr, *Berichte Der Deutschen Chemischen Gesellschaft* **1901**, *34*, 957-981.
[27] A. F. G. Bongat, A. V. Demchenko, *Carbohydrate Research* **2007**, *342*, 374-406; R. M. van Well, K. P. R. Kartha, R. A. Field, *Journal of Carbohydrate Chemistry* **2005**, *24*, 463-474; T. S. Kaerkkaeinen, K. P. R. Kartha, D. MacMillan, R. A. Field, *Carbohydrate Research* **2008**, *343*, 1830-1834; J. C. Manimala, Z. T. Li, A. Jain, S. VedBrat, J. C. Gildersleeve, *Chembiochem* **2005**, *6*, 2229-2241.
[28] S. A. Mitchell, M. R. Pratt, V. J. Hruby, R. Polt, *Journal of Organic Chemistry* **2001**, *66*, 2327-2342.
[29] E. Meinjohanns, M. Meldal, K. Bock, *Tetrahedron Letters* **1995**, *36*, 9205-9208.
[30] U. K. Saha, R. R. Schmidt, *Journal of the Chemical Society-Perkin Transactions 1* **1997**, 1855-1860.
[31] M. Schultz, H. Kunz, *Tetrahedron-Asymmetry* **1993**, *4*, 1205-1220.
[32] I. Carvalho, S. L. Scheuerl, K. P. R. Kartha, R. A. Field, *Carbohydrate Research* **2003**, *338*, 1039-1043.
[33] M. B. Martins-Teixeira, V. L. Campo, M. Biondo, R. Sesti-Costa, Z. A. Carneiro, J. S. Silva, I. Carvalho, *Bioorganic & Medicinal Chemistry* **2013**, *21*, 1978-1987.
[34] V. L. Campo, T. B. Riul, I. Carvalho, M.-D. Baruffi, *Chembiochem* **2014**, *15*, 1495-1507.
[35] P. W. Glunz, S. Hintermann, L. J. Williams, J. B. Schwarz, S. D. Kuduk, V. Kudryashov, K. O. Lloyd, S. J. Danishefsky, *Journal of the American Chemical Society* **2000**, *122*, 7273-7279.
[36] Z. Tarasiejska, R. W. Jeanloz, *Journal of the American Chemical Society* **1958**, *80*, 6325-6327; D. Stokmaier, O. Khorev, B. Cutting, R. Born, D. Ricklin, T. O. G. Ernst, F. Boeni, K. Schwingruber, M. Gentner, M. Wittwer, M. Spreafico, A. Vedani, S. Rabbani, O. Schwardt, B. Ernst, *Bioorganic & Medicinal Chemistry* **2009**, *17*, 7254-7264.
[37] V. L. Campo, I. Carvalho, S. Allman, B. G. Davis, R. A. Field, *Organic & Biomolecular Chemistry* **2007**, *5*, 2645-2657.
[38] W. C. Chan, P. D. White, *Fmoc solid phase peptide synthesis: a practical approach*, Oxford University Press Inc., New York, **2000**.

[39] A. Hoffmann-Roeder, A. Kaiser, S. Wagner, N. Gaidzik, D. Kowalczyk, U. Westerlind, B. Gerlitzki, E. Schmitt, H. Kunz, *Angewandte Chemie-International Edition* **2010**, *49*, 8498-8503.
[40] H. Kunz, *Journal of Peptide Science* **2003**, *9*, 563-573.
[41] J. S. Plested, P. A. Coull, M. A. J. Gidney, in *Haemophilus influenzae Protocols, Vol. 71* (Ed.: M. A. H. Herbert, Derek W. Moxon, E. Richard), Humana Press, **2003**, pp. 243-261.
[42] H. Cai, Z.-Y. Sun, M.-S. Chen, Y.-F. Zhao, H. Kunz, Y.-M. Li, *Angewandte Chemie-International Edition* **2014**, *53*, 1699-1703.
[43] A. Kurosaka, A. Yano, N. Itoh, Y. Kuroda, T. Nakagawa, T. Kawasaki, *Journal of Biological Chemistry* **1991**, *266*, 4168-4172.
[44] A. Sandomenico, V. Severino, A. Chambery, A. Foca, G. Foca, C. Farina, M. Ruvo, *Molecular Biotechnology* **2013**, *54*, 983-995.
[45] D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *Journal of Computational Chemistry* **2005**, *26*, 1701-1718.
[46] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. shaw, *Proteins-Structure Function and Bioinformatics* **2010**, *78*, 1950-1958.
[47] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. Gonzalez-Outeirino, C. R. Daniels, B. L. Foley, R. J. Woods, *Journal of Computational Chemistry* **2008**, *29*, 622-655.
[48] D. Madariaga, N. Martinez-Saez, V. J. Somovilla, H. Coelho, J. Valero-Gonzalez, J. Castro-Lopez, J. L. Asensio, J. Jimenez-Barbero, J. H. Busto, A. Avenoza, F. Marcelo, R. Hurtado-Guerrero, F. Corzana, J. M. Peregrina, *Acs Chemical Biology* **2015**, *10*, 747-756; D. M. Coltart, A. K. Royyuru, L. J. Williams, P. W. Glunz, D. Sames, S. D. Kuduk, J. B. Schwarz, X. T. Chen, S. J. Danishefsky, D. H. Live, *Journal of the American Chemical Society* **2002**, *124*, 9833-9844; U. Karsten, N. Serttas, H. Paulsen, A. Danielczyk, S. Goletz, *Glycobiology* **2004**, *14*, 681-692; L. Kinarsky, G. Suryanarayanan, O. Prakash, H. Paulsen, H. Clausen, F. G. Hanisch, M. A. Hollingsworth, S. Sherman, *Glycobiology* **2003**, *13*, 929-939; J. Schuman, A. P. Campbell, R. R. Koganty, B. M. Longenecker, *Journal of Peptid Research* **2003**, *61*, 91-108; S. Dziadek, C. Griesinger, H. Kunz, U. M. Reinscheid, *Chemistry-a European Journal* **2006**, *12*, 4981-4993; D. Madariaga, N. Martinez-Saez, V. J. Somovilla, L. Garcia-Garcia, M. A. Berbis, J. Valero-Gonzalez, S. Martin-Santamaria, R. Hurtado-Guerrero, J. L. Asensio, J. Jimenez-Barbero, A. Avenoza, J. H. Busto, F Corzana, J. M. Peregrina, *Chemistry-a European Journal* **2014**, *20*, 12616-12627; T. Matsushita, N. Ohyabu, N. Fujitani, K. Naruchi, H. Shimizu, H. Hinou, S. I. Nishimura, *Biochemistry* **2013**, *52*, 402-414.

FIGURES AND SCHEMES

Scheme 1. (A) Schematic representation of MUC1 tandem repeat sequence containing Tn antigen (αGalNAc-Ser/Thr). **(B)** Chemical structures of the protected *O*-glycosyl amino acids βGalNAc-SerOBn **1** and βGalNAc-ThrOBn **2** as Tn antigen analogs. **(C)** Chemical structures of the target MUC1 glycopeptides **3**–BSA and **4**–BSA.

Scheme 2. Synthesis of protected *O*-glycosyl amino acids βGalNAc-SerOBn **1**, βGalNAc-ThrOBn **2** and βGalNAc-ThrOH **6**. Reagents and conditions: (i) Ac_2O/ pyridine; (ii). Acetyl chloride, HCl(g), 40°C, 18 h; (iii) $HgBr_2$, 1,2-DCE, 90°C, 20 h; (iv) H_2/Pd-C 10%, MeOH, 0.5 h.

Scheme 3. Solid phase synthesis of glycopeptides **3**–BSA and **4**–BSA. Reagents and conditions: i. 50% morpholine in DMF; ii. Fmoc-ProOH, PyBOP, HOBt, DIPEA; iii. Fmoc-AlaOH, PyBOP, HOBt, DIPEA iv. Fmoc-Arg(Pbf)OH, PyBOP, HOBt, DIPEA; v. αGalNAc-ThrOH **5** or βGalNAc-ThrOH **6**, PyBOP HOBt, DIPEA; vi. Fmoc-Asp(OtBu)OH, PyBOP, HOBt, DIPEA; vii. Fmoc-SerOH, PyBOP, HOBt DIPEA; viii. Ac_2O/ pyridine; ix. TFA 80%; x. NaOMe, MeOH; xi. BSA, EDCI, NHS, Na_2HPO_4.

Figure 1. Sera titration from mice immunized with glycopeptides **A) 3**–BSA and **B) 4**–BSA by ELISA. OD_{450nm}: optical density at 450 nm. ■: Immunizations with glycopeptides **3**–BSA and **4**–BSA and coating with corresponding glycopeptides **3**–BSA and **4**–BSA.▼: Immunizations with glycopeptides **3**–BSA and **4**–BSA and coating with BSA. ●: Immunizations with BSA and coating with BSA. ▲: Preimmune (PI) sera and coating with corresponding glycopeptides **3**–BSA and **4**–BSA.

Figure 2. Flow cytometry analysis of the binding (percentages of positive staining) of glycopeptides **3**-BSA and **4**–BSA antisera to MCF-7 cells. **A)** MCF-7 cells incubated with preimmune (PI) and anti-MCF-7 sera (1:1000). **B)** MCF-7 cells incubated with glycopeptide **3**–BSA antisera at increasing dilutions of 1:400 to 1:8000. **C)** MCF-7 cells incubated with glycopeptide **4**–BSA antisera at increasing dilutions of 1:400 to 1:8000.

Figure 3. Percentages of positive staining, obtained by flow cytometry, of glycopeptides **3**– and **4**–BSA antisera to MCF-7 cells. MCF-7 cells incubated with anti-MCF-7 sera (black bar) and preimmune (PI) sera (white bars) at 1:1000 dilution, and with glycopeptide **3**–BSA antisera (light grey bars) and glycopeptide **4**–BSA antisera (dark grey bars) at increasing dilutions of 1:400 to 1:8000.

Figure 4. Binding between anti-αGalNAc and anti-βGalNAc immunopurified antibodies to coated BSA-conjugated glycopeptides **3** and **4**, and BSA, at 10, 1 and 0.1 µg/mL.

10.1002/cbic.201600473

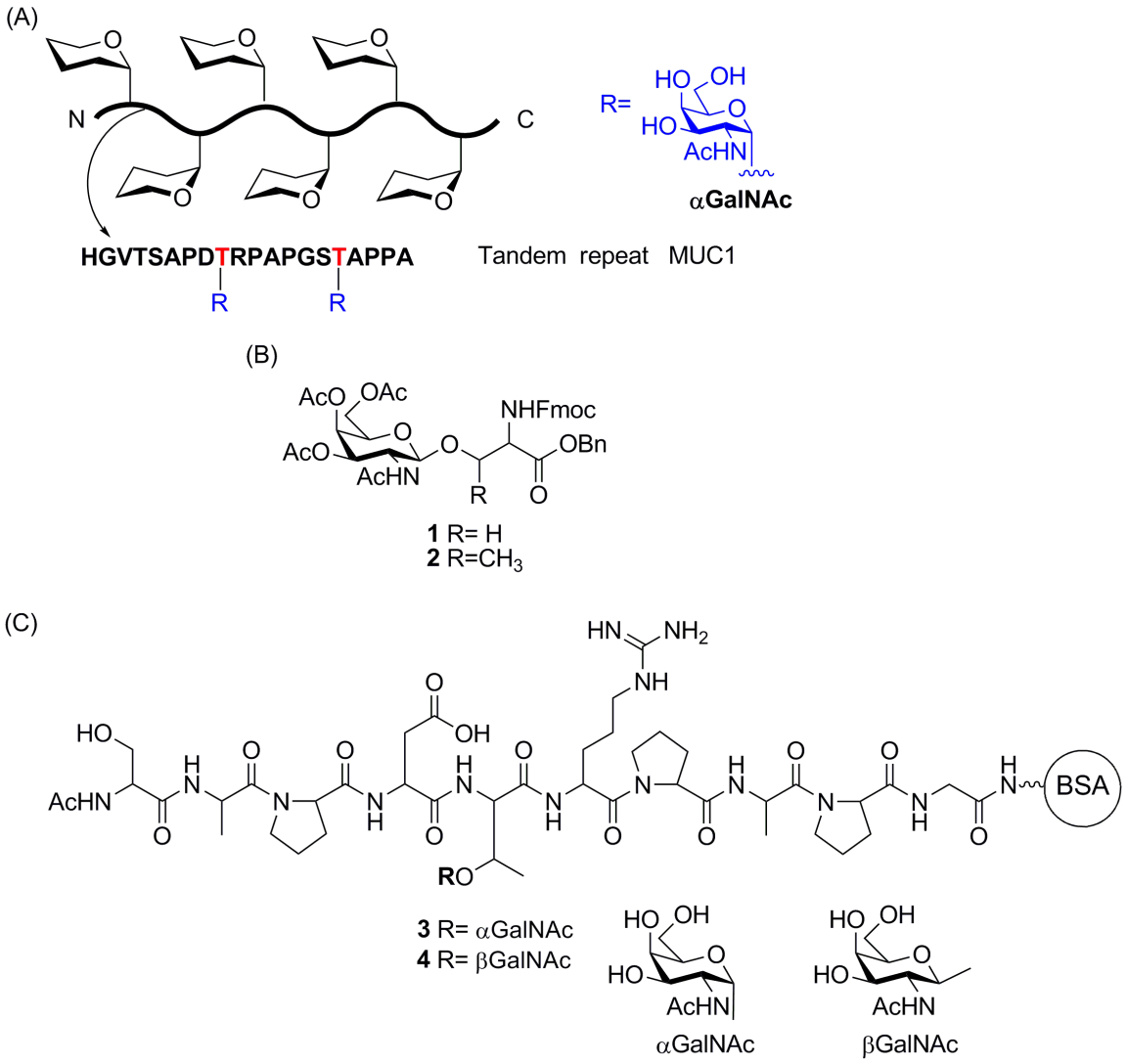
Figure 5. (**A**) Overall view of the simulated system. The SM3 monoclonal antibody chains A and B are shown as green and cyan cartoons, respectively. The MUC1-derived glycopeptide is shown as spheres. Yellow indicates a carbon atom, blue indicates nitrogen and red an oxygen. (**B**,**C**) Detailed view showing the difference in the carbohydrate ring orientation between the α and βGalNAc glycopeptides **3** and **4** shown as sticks.

Figure 6. (**A**) Detailed view of the interaction between the carbohydrate moiety of the αGalNAc-glycopeptide **3**. Its carbohydrate moiety interacts with Tyr-32L and Asn-53L via hydrogen bonds. (**B** Detailed view of the interaction between the carbohydrate moiety of the βGalNAc-glycopeptide **4**. It carbohydrate moiety also interacts with Tyr-32L and Asn-53L via hydrogen bonds.

Figure 7. Binding sensorgrams relative to the interactions of glycopeptide **3** (α) (**A**) and glycopeptide 4 (β) (**B**) with immobilized SM3 MAb. Glycopeptides were injected at increasing concentrations betwee 50 nM and 2 μM.

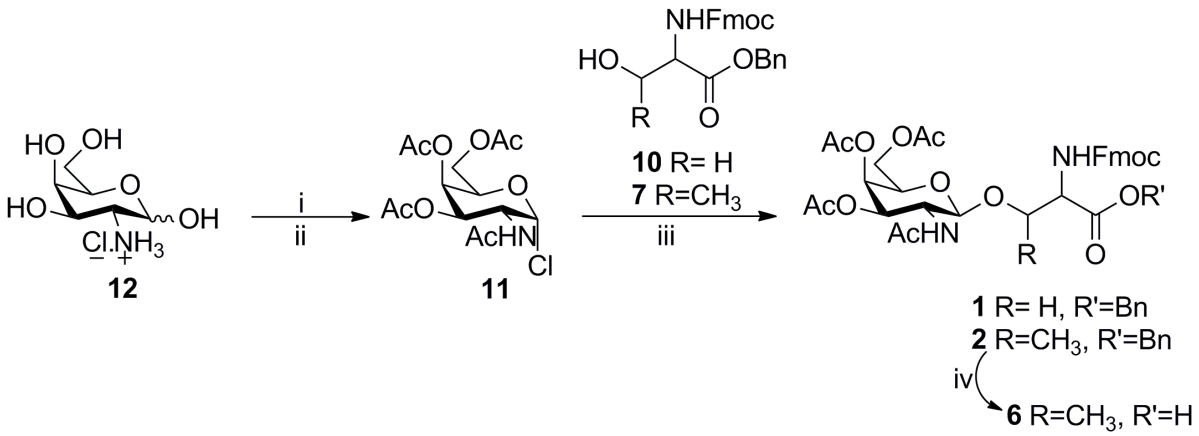
10.1002/cbic.201600473

SCHEME 1



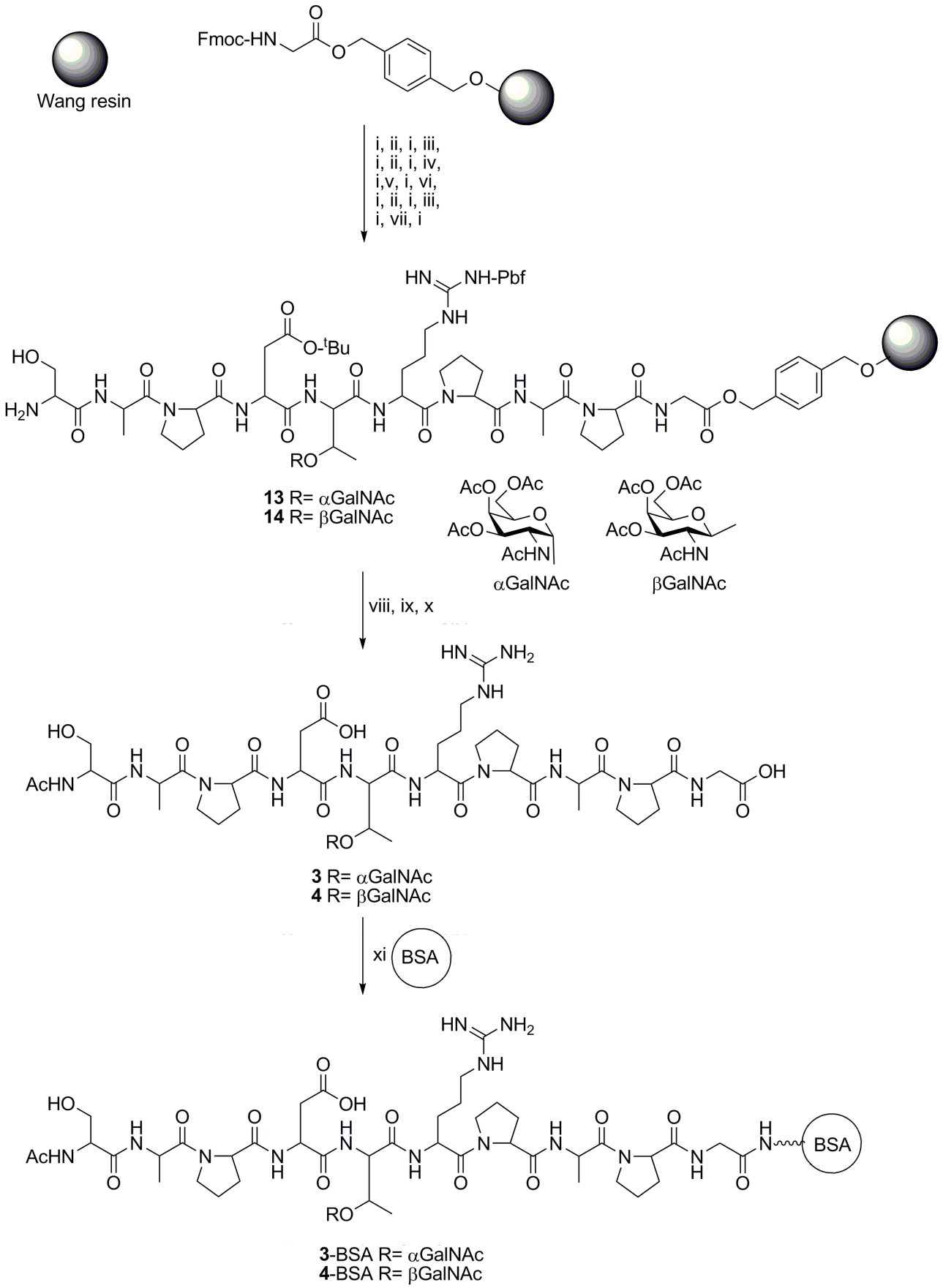
10.1002/cbic.201600473

SCHEME 2



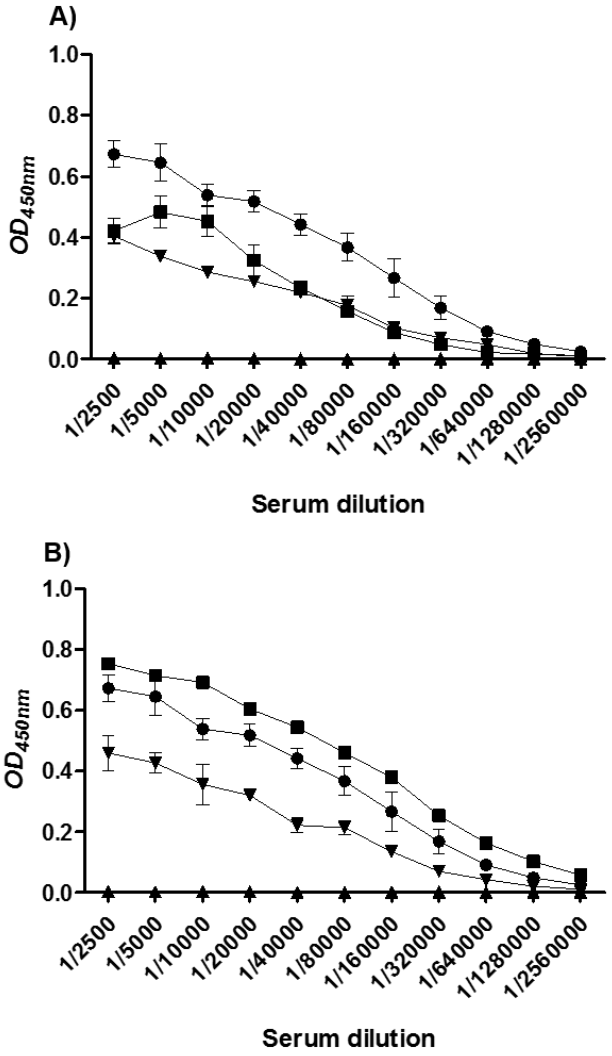
10.1002/cbic.201600473

SCHEME 3



10.1002/cbic.201600473

FIGURE 1



10.1002/cbic.201600473

FIGURE 2

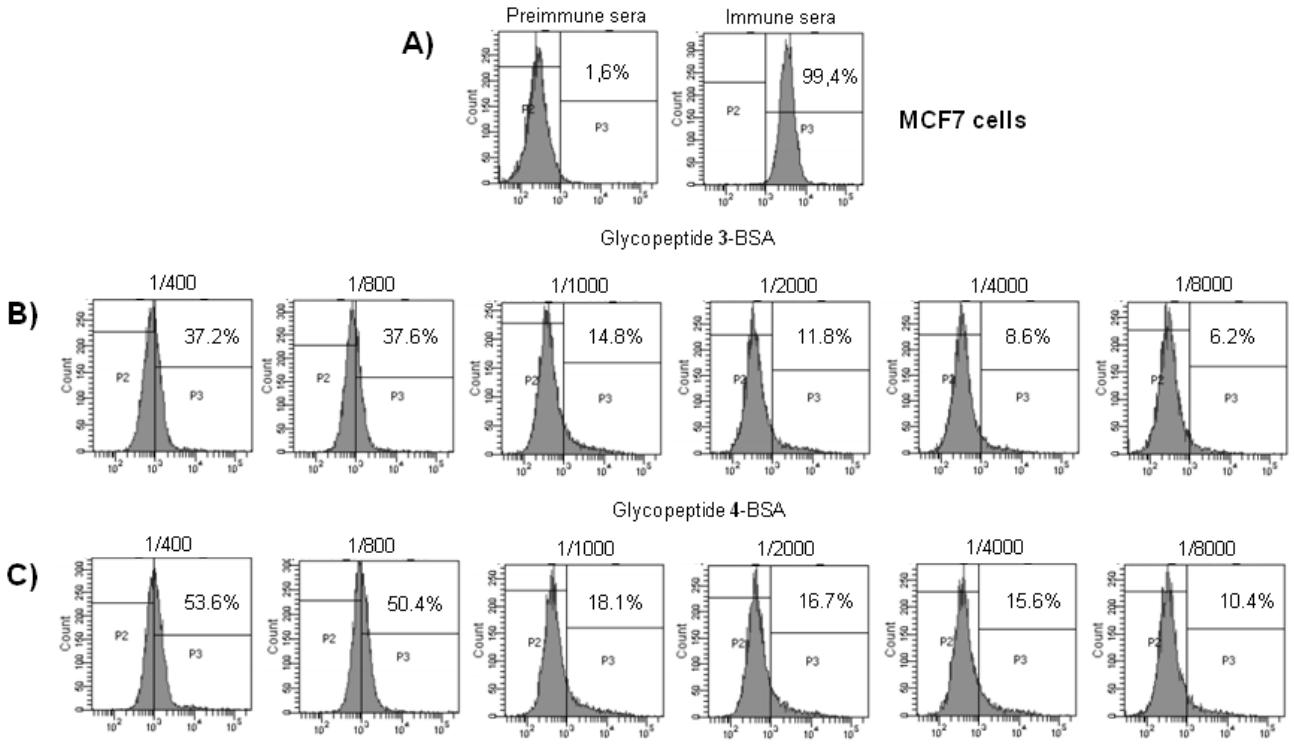
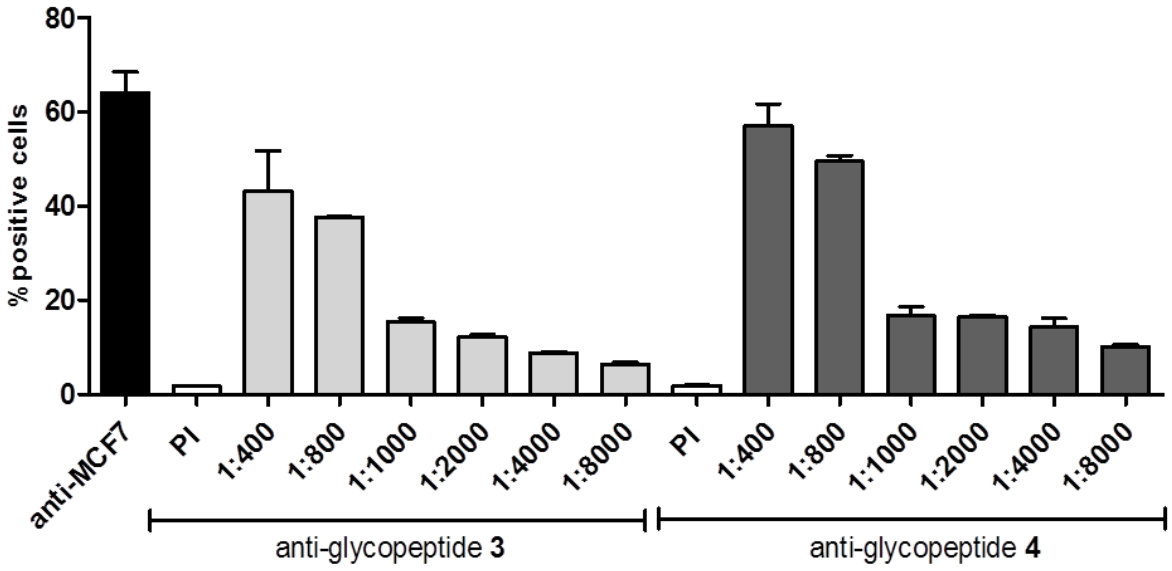


FIGURE 3



10.1002/cbic.201600473

FIGURE 4

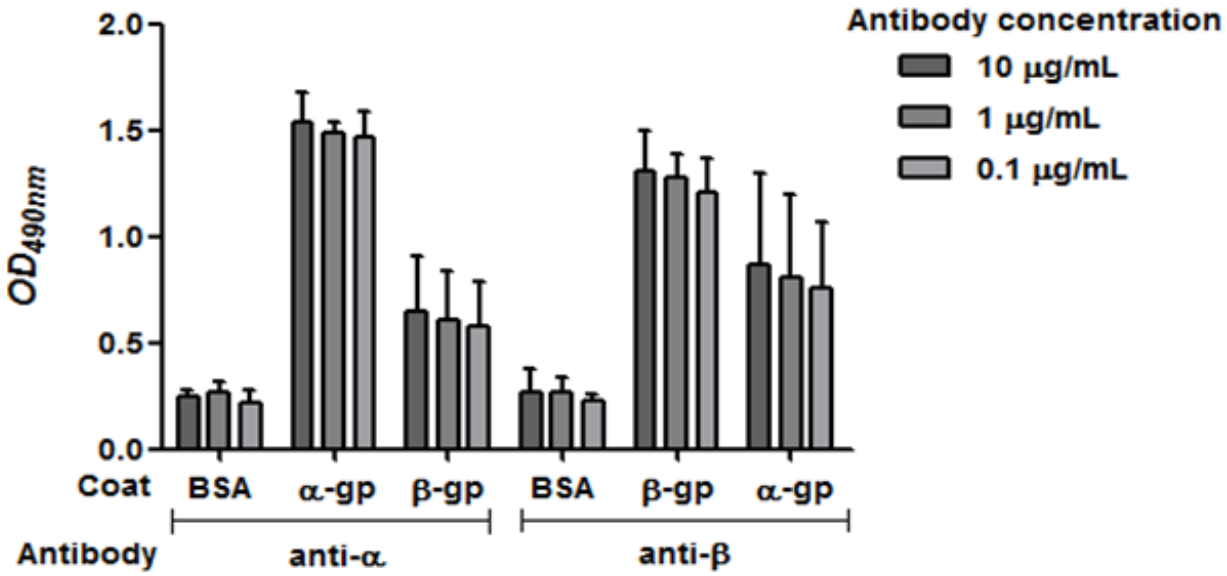
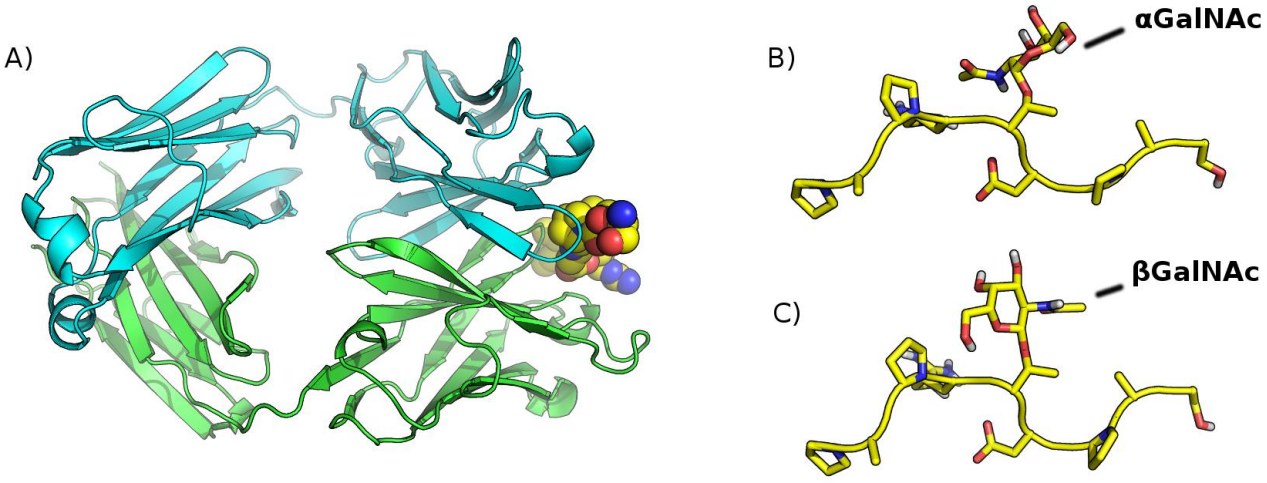


FIGURE 5



10.1002/cbic.201600473

FIGURE 6

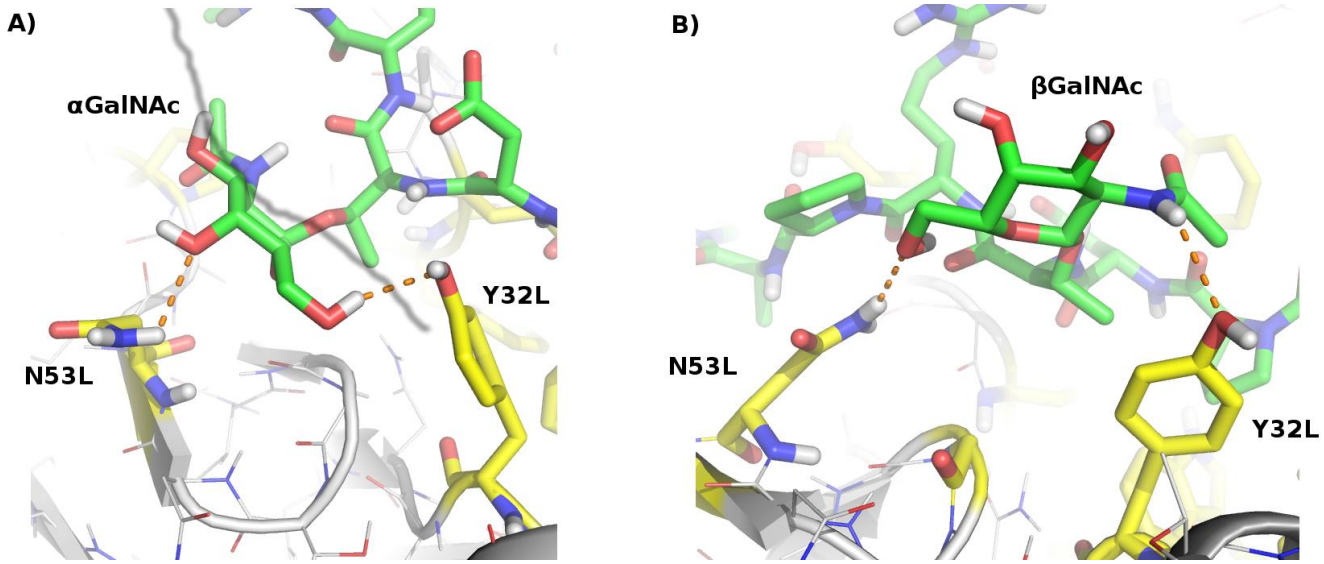
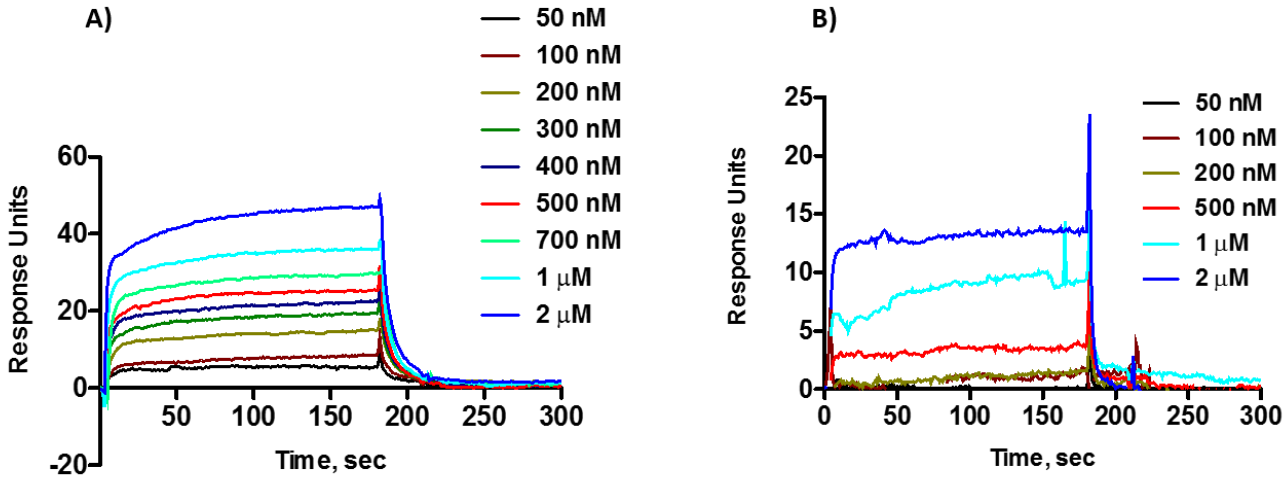
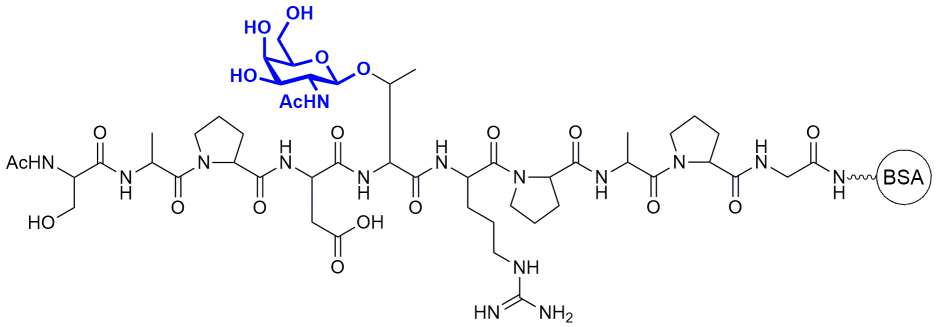


FIGURE 7



10.1002/cbic.201600473

Talking Graph – Table of Contents (TOC)



The glycopeptides NHAcSer-Ala-Pro-Asp-Thr[αGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA and NHAcSer-Ala-Pro-Asp-Thr[βGalNAc]-Arg-Pro-Ala-Pro-Gly-BSA, bearing the corresponding αGalNAc-Thr (Tn) and βGalNAc-Thr as Tn antigen isomer, were synthesized as mimetics of MUC1 tumor mucin glycoproteins. Experimental and theoretical data highlighted the improved immunogenicity of the unnatural βGalNAc-glycopeptide-BSA, thus representing a frontline glycopeptide for development of antitumor vaccines.