BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 5929–5941 5929
DOI: 10.1021/bi100623g

Unexpected Tolerance of Glycosylation by UDP-GalNAc:Polypeptide α-*N*-Acetylgalactosaminyltransferase Revealed by Electron Capture Dissociation Mass Spectrometry: Carbohydrate as Potential Protective Groups[†]

Yayoi Yoshimura,[‡] Takahiko Matsushita,[‡] Naoki Fujitani,[‡] Yasuhiro Takegawa,[‡] Haruhiko Fujihira,[‡] Kentarou Naruchi,[‡] Xiao-Dong Gao,[‡] Naomi Manri,[§] Takeshi Sakamoto,[§] Kentaro Kato,[‡] Hiroshi Hinou,[‡] and Shin-Ichiro Nishimura*,[‡]

[‡]*Graduate School of Life Science and Research Center for Post-Genome Science and Technology, Hokkaido University, Sapporo 001-0021, Japan, and* [§]*Central Research Laboratory, Hitachi, Ltd., Kokubunji, Tokyo 185-8601, Japan*

Received April 23, 2010; Revised Manuscript Received June 9, 2010

ABSTRACT: UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferases (ppGalNAcTs, EC 2.4.1.41), a family of key enzymes that initiate posttranslational modification with *O*-glycans in mucin synthesis by introduction of α-GalNAc residues, are structurally composed of a catalytic domain and a lectin domain. It has been known that multiple Ser/Thr residues are assigned in common mucin glycoproteins as potential *O*-glycosylation sites and more than 20 distinct isoforms of this enzyme family contribute to produce densely *O*-glycosylated mucin glycoproteins. However, it seems that the functional role of the lectin domain of ppGalNAcTs remains unclear. We considered that electron capture dissociation mass spectrometry (ECD-MS), a promising method for highly selective fragmentation at peptide linkages of glycopeptides to generate unique *c* and *z* series of ions, should allow for precise structural characterization to uncover the mechanism in *O*-glycosylation of mucin peptides by ppGalNAcTs. In the present study, it was demonstrated that a system composed of an electrospray source, a linear RFQ ion trap that isolates precursor ions, the ECD device, and a TOF mass spectrometer is a nice tool to identify the preferential *O*-glycosylation sites without any decomposition of the carbohydrate moiety. It should be noted that electrons used for ECD are accelerated within a range from 1.75 to 9.75 eV depending on the structures of glycopeptides of interest. We revealed for the first time that additional installation of a α-GalNAc residue at potential glycosylation sites by ppGalNAcT2 proceeds smoothly in various unnatural glycopeptides having α-Man, α-Fuc, and *β*-Gal residues as well as α-GalNAc residues. The results may suggest that ppGalNAcT2 did not differentiate totally presubstituted sugar residues in terms of configuration of functional groups, D-, L-configuration, and even α-, *β*-stereochemistry at an anomeric carbon atom when relatively short synthetic peptides were employed for the acceptor substrates. Unexpected characteristics of ppGalNAcT2 motivated us to challenge site-directed installation of α-GalNAc residues at desired position(s) by protecting some hydroxyl groups of Thr/Ser residues with selectively removable sugars, notably a novel concept as "carbohydrate as protective groups", toward a goal of the systematic chemical and enzymatic synthesis of biologically important mucin glycopeptides.

Human cell surface mucins are one of the most important classes of glycoproteins to influence significantly specific cellular adhesion during differentiation, proliferation, or malignant alteration in embryogenesis, organogenesis, carcinogenesis, and cancer metastasis (*1*−*3*). Glycan chains of mucin glycoproteins, namely, *O*-glycans densely substituted at multiple Thr/Ser residues, are commonly constructed by a variety of glycosyltransferases at ER/Golgi membranes through highly complicated biosynthetic pathways (Figure 1) (*4*). It is likely that *O*-glycan structures of mucins such as MUC1, MUC2, MUC4, MUC5AC, and MUC16 in the normal human epithelial cells alter markedly during cancer progression and malignant alteration (*2*). Therefore, aberrantly *O*-glycosylated mucins by unusual glycoforms expressed on various cancer cell surfaces are considered to be promising diagnostic biomarkers (*5*). Extensive attention has also been paid toward discovery research of cancer-dominant mucin glycopeptides for the development of novel cancer vaccines and improved antibody immunotherapeutics (*6*, *7*).

UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferases (ppGalNAcTs,[1] EC 2.4.1.41), a family of key enzymes that initiate posttranslational modification with *O*-glycans in mucin synthesis, are structurally composed of a catalytic domain and a lectin domain (*4*). These enzymes utilize UDP-GalNAc as the sugar nucleotide donor substrate to modify Ser/Thr residues of proteins trafficking through the ER/Golgi membranes. The newly formed α-GalNAc-Ser/Thr moiety is termed the "Tn-antigen" and is generally further modified to construct complex oligosaccharide

[†]This work is supported partly by a program grant for the national project "Innovation COE program for future drug discovery and medical care" from the Ministry of Education, Culture, Science, and Technology, Japan.

*Corresponding author. E-mail: shin@glyco.sci.hokudai.ac.jp. Tel: +81-(0)11-706-9043. Fax: +81-(0)11-706-9042.

[1]Abbreviations: ppGalNAcT, UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferase; ECD-MS, electron capture dissociation mass spectrometry; MALDI-TOFMS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

Published on Web 06/11/2010

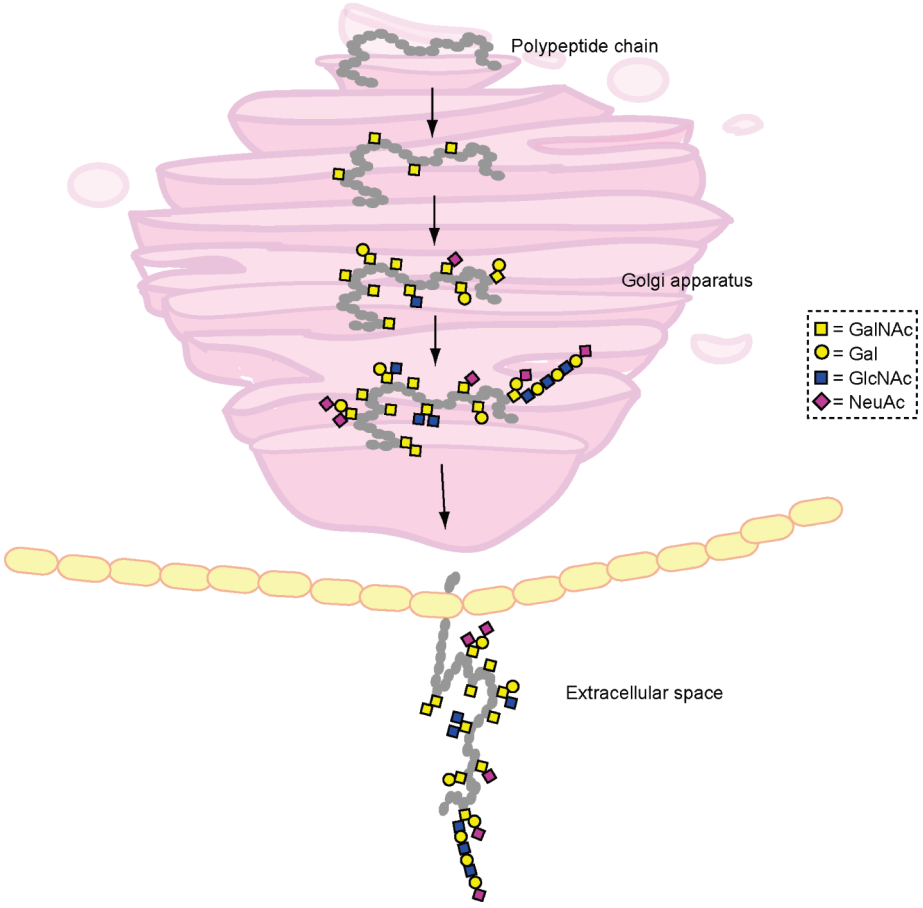


FIGURE 1: A schematic image of the biosynthetic path of mucin glycoproteins.

structures. It has been documented that multiple Ser/Thr residues are assigned in common mucin glycoproteins as potential *O*-glycosylation sites and more than 20 distinct isoforms of this enzyme family contribute to produce densely *O*-glycosylated mucin glycoproteins (*4*). Among isoforms of mammalian ppGalNAcTs, ppGalNAcT1, -T2, and -T4 are distributing in a wide range of tissues and supposed to act upon a large repertoire of protein scaffolds in an early stage during mucin domain assembly, while some isoforms such as ppGalNAcT7, -T10, and -T11 seem to function in the late or final stage (Figure 1) (*8*, *9*). Since the construction of such mucin domain assembly should be influenced significantly by preceding glycosylation status at neighboring Thr/Ser residue(s) (*10*), lectin domains of ppGalNAcTs appear to define the preferred glycosylation sites in partially glycosylated peptide substrates. Although it seems likely that there is a complex balance between redundant substrate preferences and hierarchical relationships among the members of ppGalNAcTs family, the functional role of their lectin domains remains to be unclear (*9*, *11*, *12*). Considering the existence of rare *O*-glycosylation at Ser/Thr residues by such as α-Man, α-Fuc, *β*-GlcNAc, *β*-Glc, and *β*-Xyr (*13*−*16*) as well as unknown lectin function of the ppGalNAcTs (*17*, *18*), we hypothesized that ppGalNAcTs should have some tolerance in the additional installation of GalNAc residues toward synthetic unnatural/unusual glycopeptides as acceptor substrates, in which one or some of the potential *O*-glycosylation sites are modified preferentially with above minor *O*-glycosides. If this really is the case, then it might become a breakthrough to expand chemical and enzymatic synthesis of mucin glycoproteins based on the site-directed glycosylation with various glycosyltransferases.

EXPERIMENTAL PROCEDURES

Materials and Methods. All commercially available solvents and reagents were used without further purification. Tentagel S RAM resin (Hipep Laboratories, 0.25 mmol/g) was used for the synthesis of glycopeptides. The following protected amino acids were used: Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OPfp, Fmoc-Asp(Boc)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Pro-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Thr(*t*Bu)-OH, and Fmoc-Val-OH (Novabiochem). Glycosylated amino acid derivatives, Fmoc-Thr(Ac_3-α-L-Fuc)-OH, Fmoc-Thr-(Ac_4-*β*-D-Gal)-OH, and Fmoc-Thr(Ac_4-α-D-Man)-OH, were prepared according to the method reported previously (*19*, *20*). All solid-phase reactions for glycopeptide synthesis were performed manually in a polypropylene tube equipped with a filter. The reaction vessel for solid-phase synthesis was placed inside a cavity of a microwave instrument and was stirred with a vortex mixer. Single-mode microwave was irradiated at 2450 MHz using temperature control at 50 °C (*21*). Uridine-5′-diphospho-*N*-acetylgalactosamine, 2Na (UDP-GalNAc), and cytidine-5′-phospho-*N*-acetylneuraminic acid, 2Na (CMP-Neu5Ac), were purchased from Yamasa Co. α-Mannosidase (from jack bean) was obtained from Seikagaku Kogyo. α-Fucosidase (from bovine kidney) and *β*-galactosidase (from *Aspergillus niger*) were purchased from Sigma. Matrix-associated laser desorption ionization time-of-fright mass spectrometry (MALDI-TOFMS) was recorded and measured by Bruker Ultraflex machines using 2,5-dihydroxybenzoic acid (DHB) as matrix. HPLC analysis was performed on a Hitachi HPLC system equipped with an L-7100 intelligent pump and an L-7405 UV detector, using a reversed-phase (RP)

C18 column, Inertsil ODS-3, 4.6 × 250 mm^2 (GL Sciences Inc.). For purification of the products, Hitachi HPLC system equipped with an L-6250 intelligent pump and an L-7400 UV detector or using a reversed-phase (RP) C18 column, Inertsil ODS-3, 20 × 250 mm^2 (GL Sciences Inc.), was used. NMR spectra were recorded on a Bruker AVANCE DRX 600.

Preparation of Peptide Substrates Having Sugar Protective Groups. Synthesis of glycopeptide substrates bearing α-Man, α-Fuc, and β-Gal was performed by a standard method for solid-phase glycopeptide synthesis described in the previous publications (*22*−*26*). Tentagel S RAM resin (60 mg, 15 μmol) was swollen in 5 mL of DMF for 30 min and drained. All subsequent cycles required removal of an Fmoc group. To accomplish this, the swollen resin was covered with 20% piperidine in DMF under microwave irradiation for 5 min, drained, and washed five times with 5 mL of DMF. The coupling solutions were prepared by dissolving Fmoc-protected amino acid (45 μmol) in 1.5 mL of DMF, to which was added HBTU (1.0 equiv for amino acids), HOBt (1.0 equiv for amino acids), and DIEA (2.0 equiv for amino acids). The coupling solution was then added to the resin and agitated for 5−10 min under microwave irradiation, drained, and washed five times with 5 mL DMF. In the case of glycoamino acid, couplings were performed with the solution of glycoamino acid (18 μmol) in DMF with HBTU (1.8 equiv for sugar amino acids), HOBt (1.8 equiv for sugar amino acids), and DIEA (2.0 equiv for sugar amino acids), and the reaction mixtures were treated for 20 min under microwave irradiation (0−40 W) at 50 °C. After the coupling, unreacted amino groups were capped by acetylation with 13 mM HOBt in Ac_2O/DIEA/DMF solution (4.75:2.25:93.0, 1 mL) for 5 min, and the resin was washed with DMF. After removal of the last Fmoc groups, the resins were treated with "cleavage cocktail" solution [TFA/EDT/H_2O/triisopropylsilane (94:2.5:2.5:1)] for 2 h at room temperature, were then filtered off, and were washed with 0.1% TFA. The filtrates were concentrated, and the residue was poured into cold ether. The precipitates were centrifuged and washed with cold ether, and the crude glycopeptides were subsequently dissolved in MeOH (5 mL). To the solution was added 1 N NaOH to keep the solution at pH 13. After being stirred for 30 min, the reaction mixtures were neutralized by addition of AcOH. The solvents were evaporated *in vacuo*. The deprotected glycopeptides were dissolved in H_2O and purified by RP-HPLC under the following condition: a linear gradient of mobile phase 1−40% B (mobile phase A, water with 0.1% TFA; mobile phase B, acetonitrile with 0.1% TFA) in 60 min at a flow rate of 10.0 mL/min and ultraviolet (UV) detection at 220 nm. Totally, nine kinds of unnatural glycopeptides as substrates tested in this study were obtained in 20−40% yield (approximately 4−10 mg scale).

Expression and Purification of Human Recombinant ppGalNAcT2 (27). The human *ppGalNAc-T2* (hT2) gene lacking the sequence encoding the putative transmembrane region was amplified directly from a small intestine cDNA library (Takara Bio, Japan) by using 5′hT2-Xho and 3′hT2-Sal as primers. The 5′hT2-Xho primer (5′-CCCTCGAGAAAAGACATCATCATCATCATCATGATTACAAAGATGATGTGATAAAAAAAAGAAAGACCTTCAT-3′) encodes six histidine residues followed by a FLAG peptide sequence and residues 52−57 of hT2 protein. The 3hT2-Sal (5′-GTTCTTGTCGACCTACTGCTGCAGGTTGAGCG-3′) encodes the last five residues of hT2 protein followed by a stop codon. The PCR fragment encoding the stem region and the entire catalytic and lectin domains of hT2 protein was then digested with *Xho*I and *Sal*I and cloned into the corresponding sites of the plasmid pPICZαA (Invitrogen). This generated the expression plasmid pPICZαA-hT2, which expresses a soluble hT2 protein with N-terminal tag of hexahistidine followed by a FLAG peptide under the methanol-inducible *AOX1* promoter. Standard molecular biology techniques were used for this construction, and the correct sequence of the PCR-amplified product was verified by DNA sequencing. The expression plasmid (pPICZαA-hT2) was linearized with *Sac*I and used to transform the *Pichia pastoris* SMD1168 (Invitrogen) strain by the electroporation method. The transformants were grown on YPD plate (1% yeast extract, 2% peptone, 2% glucose, 2% agar) containing 1 M sorbitol and Zeocine (100 μg/mL). Colonies with the integration of expression cassette were determined by direct PCR analysis using 5′ and 3′ *AOX1* primers (Invitrogen). The transformation and selection were performed according to the manufacturer's instructions.

Production of human ppGalNAc-T2 was performed in an 8 L jar-fermentor with a 6 L working volume by using *P. pastoris* cells transformed with pPICZαA-hT2. A process controller system (EPC-2000; EYELA, Tokyo, Japan) was used to maintain dissolved oxygen level at 15% saturation by automatically adjusting both agitation and the air or oxygen supply. The jar-fermentor with 6 L of modified BMGY media [2% yeast extract, 6% peptone, 1.34% yeast nitrogen base, 1% glycerol, 400 μg/mL biotin, and 100 mM potassium phosphate (pH6.0)] was inoculated with 100 mL of overnight culture and grown about 16 h at 30 °C. After the carbon source (glycerol) was completely consumed, methanol was then fed into the jar-fermentor by a peristaltic pump at 10−15 mL/h to induce the expression of recombinant human ppGalNAc-T2, and the temperature was simultaneously reduced to 20 °C. After 2 days of methanol induction, the culture media were recovered by centrifugation. The supernatant was then concentrated and equilibrated with buffer A [50 mM $NaPO_4$ (pH 8.0), 300 mM NaCl] by using an ultrafiltration system (cutoff M_r = 10 kDa; MICROZA UF, Asahikasei, Tokyo, Japan). The sample was applied to a 5 mL HisTrap HP column (GE Healthcare). After the column was washed with buffer A containing 25 mM imidazole, elution was performed using a 5 column volume gradient of 25−500 mM imidazole in buffer A. Product fractions were collected and dialyzed against 50 mM Tris buffer (pH8.0) and then lyophilized for preservation.

Expression and Purification of Recombinant Human ST6GalNAc-I (28). Full-length human ST6GalNAc-I was amplified by PCR using human small intestine cDNA (Takara-bio, Tokyo, Japan) as template. Expression constructs of a soluble human ST6GalNAc-I were designed to encode the enzyme lacking the C-terminal 36 residues that include the cytoplasmic and transmembrane regions and tagged with six His sequence followed by a FLAG peptide (DYKDDDDK) at its C-terminus. pFastBac-HBMs is a baculovirus vector derived from pFastBac1 (Invitrogen, La Jolla, CA) and contains a honey bee mellitin secretion signal (HBMs) sequence (MKFLVNVALVFMVVYISYIYA). The C-terminal His6FLAG-tagged ST6GalNAc-I fragment was finally PCR-cloned into pFastBac-HBMs to generate pFastBac-His6FLAG-ST6GalNAcI. Expression of this plasmid resulted in secreting a soluble His6FLAG-tagged ST6GalNAc-I enzyme into the cell culture. Standard molecular biology techniques were used for all plasmid constructions. The correction sequence of all PCR-amplified products was verified by DNA sequencing. To produce bacmid DNA, *Escherichia coli* DH10Bac (maximum efficiency; Invitrogen), which harbors

Aplysia californica polyhedrin virus bacmid, was transformed with pFastBac-HisFLAG-ST6GalNAcI and incubated on LB plates containing 50 μg/mL geneticin disulfate, 7 μg/mL gentamycin, 10 μg/mL tetracycline, 100 μg/mL X-gal, and 40 μg/mL IPTG for 48 h. White colonies containing recombinant bacmid ST6GalNAc-I were then isolated and cultured overnight. The bacmid DNA was then prepared from the cells and transfected into *Spodoptera frugiperda* (Sf9) cells using CellFectin (Invitrogen) and cultured for 72 h. After four successive amplifications in Sf9 cells, recombinant baculovirus stocks were obtained according to the manufacturer's description of BAC-TO-BAC baculovirus expression systems (Invitrogen). For large scale expression, Sf9 cells were grown (27 °C, with shaking at 114 rpm) in a 250 mL cell culture flask containing 1×10^8 cells/mL in 100 mL of Sf900II medium in the absence of serum. Cells were infected with 3:2000 of virus stock from the fourth amplification. The expression of ST6GalNAcI was monitored for 4 days after infection of the cells by Western blot with monoclonal anti-FLAG M2-peroxidase (HRP) antibody produced in mouse (Sigma-Aldrich, St. Louis, MO). Purification of recombinant soluble sialyltransferase (ST6GalNAcI) was performed at 4 °C as follows: insect cell culture medium was centrifuged at 7000*g* and filtered through a 0.45 μm filter for removal of cellular residues. Supernatant of ST6GalNAc-I (isoelectric point 10.3) was submitted to cation-exchange chromatography in a CM-Sepharose (GE Healthcare) column equilibrated with the buffer [50 mM MES−NaOH (pH 6.0) and 1% (v/v) glycerol]. The samples were eluted with a gradient of NaCl (from 0 to 1 M). Fractions of the eluted samples were assessed by Western blot with monoclonal anti-FLAG M2-peroxidase (HRP) antibody and SDS−PAGE on 8% polyacrylamide gels (Coomassie staining). After dialysis against buffer [25 mM Tris-HCl (pH 8.0) and 250 mM NaCl], the fractions containing recombinant human ST6GalNAc-I were submitted to an affinity chromatography in a Ni^{2+}-chelating Sepharose (GE Healthcare, Buckinghamshire, U.K.) column equilibrated with the buffer [250 mM Tris-HCl (pH 8.0), 1% (v/v) glycerol, and 10 mM imidazole]. The samples were then eluted with a gradient of imidazole (from 10 to 500 mM). Fractions of the eluted samples were assessed by Western blot with monoclonal anti-FLAG M2-peroxidase (HRP) antibody and SDS−PAGE on 8% polyacrylamide gels (Coomassie staining). The fractions containing ST6GalNAc-I were dialyzed against buffer [MES−NaOH (pH 6.0)] and concentrated by using Amicon Ultra-15 centrifugal filter units using a membrane with 30000 cutoff (Millipore, Bedford, MA). The purified and concentrated sample was assessed by SDS−PAGE and Coomassie staining and Western blot using monoclonal anti-FLAG M2-peroxidase (HRP) antibody.

*General Conditions of Modifications by ppGalNAcT2 and Glycosidases for the Preparation of Glycopeptides **1**−**18***. The enzymatic reactions contained the following components in a final volume of 40 μL: 10 mM $MnCl_2$, 100 mM Tris-HCl buffer (pH 7.5), 250 μM acceptor glycopeptide, 250 μM UDP-GalNAc, and 3 μg of ppGalNAcT2. The reaction mixtures were incubated at 37 °C for 6 h and were terminated by adding 160 μL of MeCN. The reactions were monitored and confirmed by MALDI-TOFMS. The enzyme reaction mixtures were evaporated with a SpeedVac, and the residue was applied to the next glycosidase treatment without any purification process. The residues after ppGalNAcT2 reaction were incubated for 15 h at 37 °C as 30 μL of reaction solution under the following conditions: (a) removal of α-Man residue, 150 mM sodium citrate buffer (pH 4.5), 50 mM $ZnCl_2$, and 20 milliunits of α-mannosidase; (b) removal of β-Gal residue, 150 mM sodium citrate buffer (pH 5.0) and 50 milliunits of β-galactosidase; and (c) removal of α-Fuc residue, 150 mM citric phosphate buffer (pH 5.0), and 50 milliunits of α-fucosidase, respectively. Finally, the reactions were terminated by adding 160 μL of MeCN. All reactions were monitored and confirmed by MALDI-TOFMS. The reaction mixtures were evaporated with a SpeedVac, and the samples were subjected to the purification by RP-HPLC. The condition of RP-HPLC is as follows: a linear gradient of mobile phase 2−14% B (mobile phase A, water with 0.1% TFA; mobile phase B, MeCN with 0.1% TFA) in 30 min at a flow rate of 1.0 mL/min and ultraviolet (UV) detection at 220 nm. Purified products were used for the structural characterization by ECD-MS experiments.

*Synthesis of MUC1 Glycopeptide **22***. The first enzymatic reaction with ppGalNAcT2 contained the following components in a final volume of 40 μL: 10 mM $MnCl_2$, 100 mM Tris-HCl buffer (pH 7.5), 125 μM MUC1 glycopeptide (AHGVTSAPD-TRPAPGS$\underline{T}_{17}$APP, Thr underlined is protected by α-Man residue), 125 μM UDP-GalNAc, and 3 μg of ppGalNAcT2. The reaction mixture was incubated at 37 °C for 12 h and was terminated by adding 160 μL of MeCN. The reaction mixture was evaporated by a SpeedVac, and crude product was directly employed for the next sialylation reaction. The following modification by ST6GalNAc1 was performed in a reaction mixture as final volume 25 μL: crude **19**, 10 mM $MnCl_2$, 100 mM Tris-HCl buffer (pH 6.5), 2.5 mM CMP-Neu5Ac, and 0.4 μg of ST6GalNAc1. The sialylation was carried out at 20 °C for 12 h and was stopped with 100 μL of MeCN. The reaction mixture was concentrated to dryness by a SpeedVac, and the crude compound **20** was dissolved in 20 μL of 150 mM sodium citrate buffer (pH 4.5) containing 50 mM $ZnCl_2$. The solution was incubated with 5 μL of α-mannosidase (20 milliunits) at 37 °C for 15 h. To the mixture was added 100 μL of MeCN to stop the reaction, and the mixture was concentrated by a SpeedVac. The crude product **21** was diluted with the following ppGalNAcT2 reaction mixture (total volume 20 μL): 10 mM $MnCl_2$, 100 mM Tris-HCl buffer (pH 7.5), and 3 μg of ppGalNAcT2. After this second ppGalNAcT2 reaction for 12 h at 37 °C, the mixture was added by 100 μL of MeCN to terminate the reaction. The reaction mixture was dried by a SpeedVac, and the crude product was diluted with 20 μL of Milli-Q water and injected to a RP-HPLC system under the condition described above. Purified product was employed for further structural elucidation by means of ECD-MS. All enzymatic reactions were monitored by MALDI-TOFMS.

ECD-MS Analysis. All samples tested were prepared as a 5 μM solution in 50:50 water/ACN with 0.1% formic acid and ionized in positive ion mode using a nanoelectron spray ionization (nanoESI) source at a flow rate of 700 nL/min with a syringe pump (Harvard Apparatus, Halliston, MA). Experiments on the structural characterization of glycopeptides were performed using a modified NanoFrontier LD system (Hitachi High-Technologies, Tokyo, Japan). The system is composed of an electrospray source, a linear RFQ ion trap that isolates precursor ions, the ECD device, and a TOF mass spectrometer. Electrons used for ECD were accelerated with 1.75−9.75 eV. Obtained ECD spectra were further analyzed by an in-house software tool covering peak picking, peak assignment, and sequence validation. The analytical condition and the systems of instruments were described in the previous papers (*29*, *30*).

Docking Simulation. The 3D structures were calculated by using a software package of the Schrödinger Suite (Schrödinger,

LLC, Portland, OR). We used the structure of ppGalNAc-T2 determined by X-ray crystal structure analyses (PDB entry: 2FFU (*31*)) as a template structure for the docking simulation

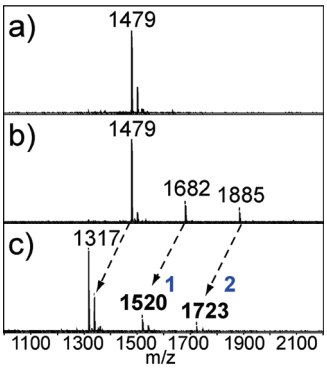
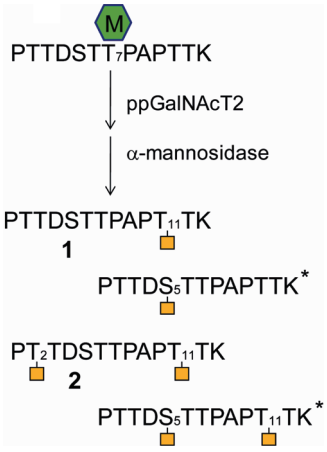


FIGURE 2: MALDI-TOFMS analysis during sequential enzymatic reactions by ppGalNAcT2 and α-mannosidase toward EA2 peptide having α-Man at Thr7. (a) MS of the starting EA2 glycopeptide, (b) MS after ppGalNAcT2-catalyzed glycosylation, and (c) MS after the treatment with α-mannosidase, respectively.

Scheme 1: Identified Structures of Glycopeptides **1** and **2** Produced by Sequential Enzymatic Modifications of Unnatural Glycopeptide Bearing α-Mannose Residue at Thr7[a]



[a]ECD-MS spectra of minor products represented by an asterisk were listed in the Supporting Information (Figure S-1).

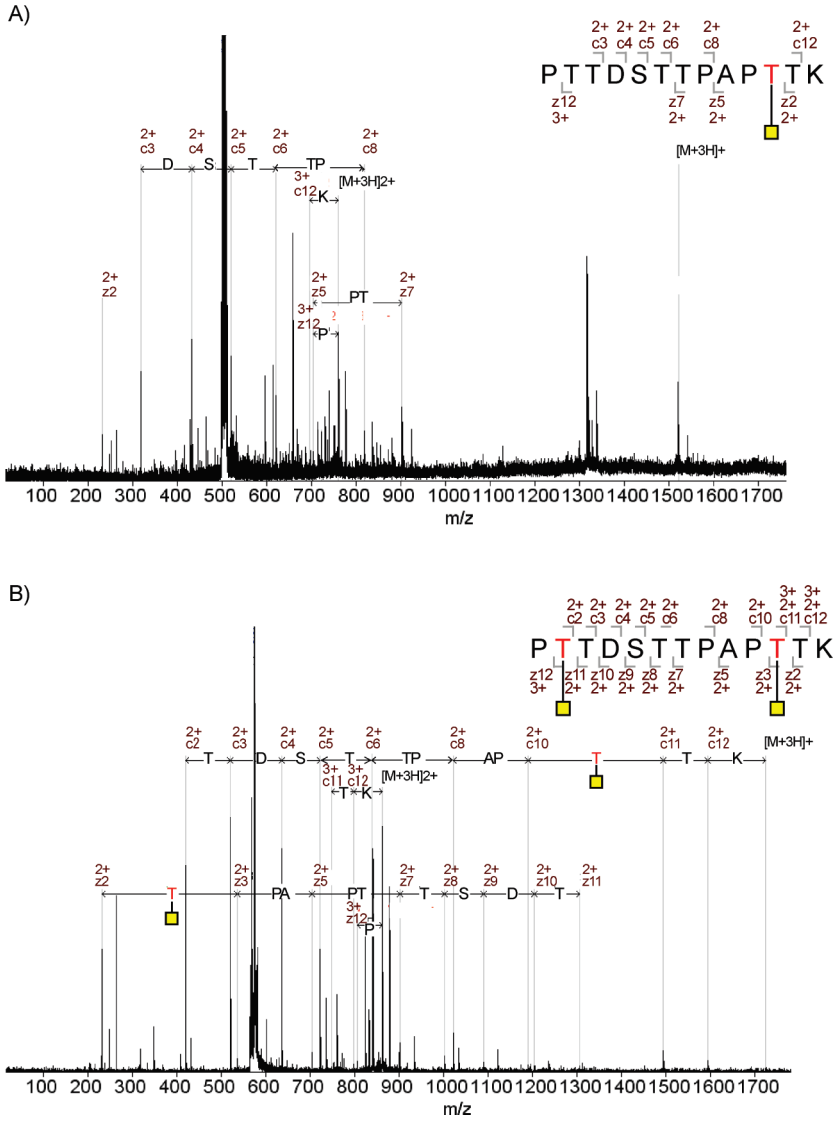


FIGURE 3: ECD-MS spectra of glycopeptides **1** (A) and **2** (B) and the corresponding fragmentation schemes derived from the precursor ions at m/z 507 for **1** ($[M + 3H]^{3+}$) and m/z 575 for **2** ($[M + 3H]^{3+}$). Electrons used for ECD were accelerated with 1.75 eV for compound **1** and 5.75 eV for compound **2**, respectively.

between the enzyme and the acceptor glycopeptides. The structure of 2FFU is a cocrystal structure including ppGalNAc-T2, UDP, Mn^{2+}, and the acceptor peptide EA2. We prepared the extended structure of glycopeptide protected by α-Man at Thr7 (EA2-7M) on the MAESTRO software included in the Schrödinger Suite, and the glycopeptide was energy minimized by the software MacroModel 9.6 in order to be applied as the initial structure for the docking calculations. The docking position was designated as the center of the site of EA2 in the template structure of 2FFU, and the docked glycopeptide was confined to the enclosing box whose volume was 40 $Å^3$. To soften the potential for nonpolar parts of the ligand, we set the cutoff value of 0.8 to the van der Waals radii of ligand atoms with partial atomic charge. All docking calculations were performed by the software Glide 5.0 included in the Schrödinger Suite without any distance and angle constraints. The force field OPLS2005 was used for the all docking simulation and energy minimization (*32*).

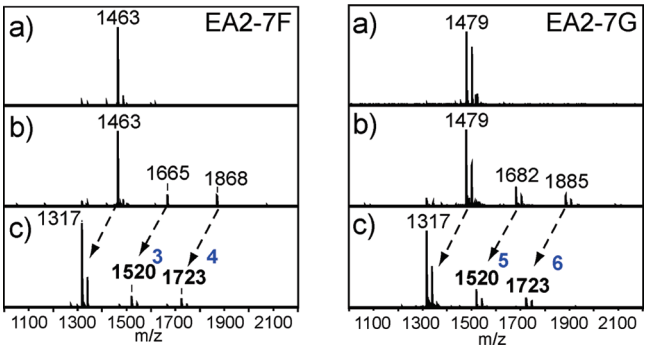


FIGURE 4: MALDI-TOFMS analysis during sequential enzymatic reactions by ppGalNAcT2 and α-fucosidase/β-galactosidase toward EA2 peptide having α-Fuc (EA2-7F)/β-Gal (EA2-7G) at Thr7. (a) MS of the starting EA2 glycopeptide, (b) MS after ppGalNAcT2-catalyzed glycosylation, and (c) MS after the treatment with α-fucosidase or β-galactosidase, respectively. ECD-MS spectra of compounds **3**−**6** were listed in the Supporting Information (Figures S-2−S-5).

RESULTS

Unexpected Glycosylation by ppGalNAcT2 toward Unnatural Glycopeptides. To test our hypothesis of the tolerance during additional installation of GalNAc residues by ppGalNAcTs, we decided to investigate preliminarily the acceptability of recombinant human ppGalNAcT2 in the glycosylation of unnatural synthetic glycopeptide, a derivative of EA2 peptide (PTTDSTTPAPTTK), in which Thr7 is modified by α-Man residue because this position is known as a predominant *O*-glycosylation site of ppGalNAcT2 (*31*). An artificial glycopeptide with unusual sugar residue at Thr7 was readily synthesized by solid-phase (glyco)peptide synthesis (SPPS) using Fmoc amino acid building block carrying per-*O*-acetylated mannose residue (*19*) under microwave irradiation (*21*). When this compound (250 μM) was subjected to the ppGalNAcT2-catalyzed glycosylation in the presence of 250 μM UDP-GalNAc, the signals at *m*/*z* 1682 and 1885 due to mono- and diglycosylated products were newly observed by MALDI-TOFMS, indicating that additional glycosylation by ppGalNAcT2 proceeded significantly in this unnatural glycopeptide detected at *m*/*z* 1479 (Figure 2a,b). As anticipated, α-Man residue at Thr7 could be removed by treating this intermediate with α-mannosidase to give two products, **1** and **2**, at *m*/*z* 1520 and 1723 in addition to the naked EA2 (*m*/*z* 1317) (Figure 2c). Structural characterization of compounds **1** and **2** was carried out by means of ESI-TOFMS (MS/MS) equipped

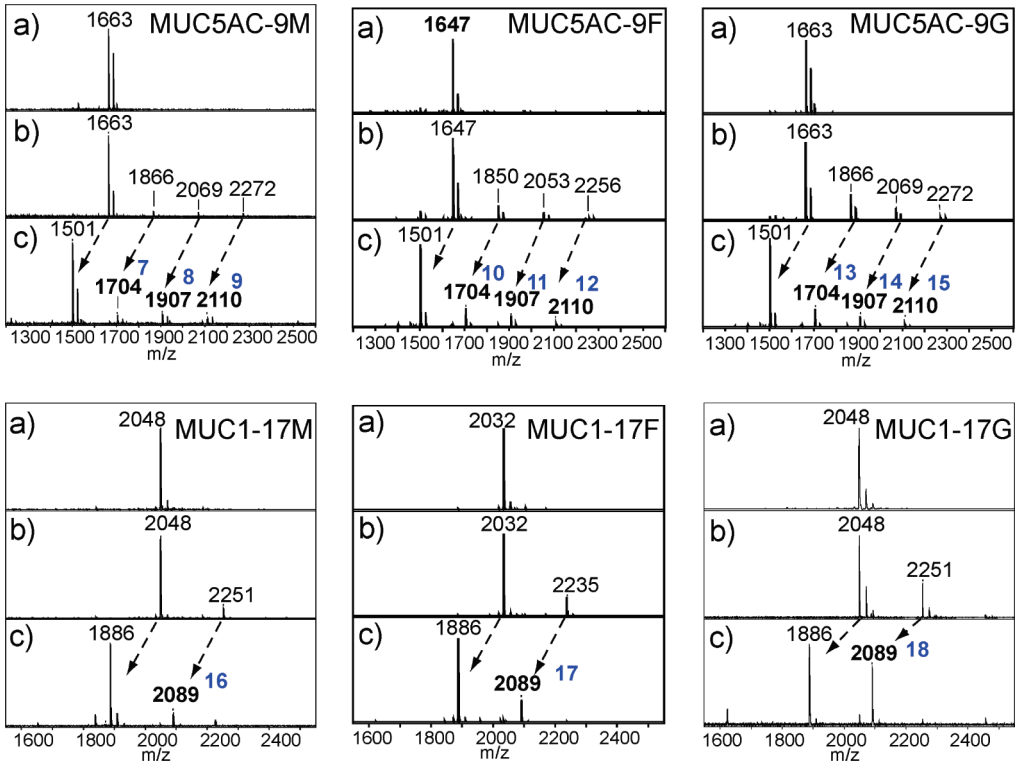


FIGURE 5: MALDI-TOFMS analysis during sequential enzymatic reactions by ppGalNAcT2 and α-mannosidase/α-fucosidase/β-galactosidase toward MUC5AC glycopeptides (MUC5AC-9M, -9F, and -9G) and MUC1 glycopeptides (MUC1-17M, -17F, and -17G). (a) MS of the starting glycopeptides, (b) MS after ppGalNAcT2-catalyzed glycosylation, and (c) MS after the treatment with α-mannosidase, α-fucosidase, or β-galactosidase, respectively.

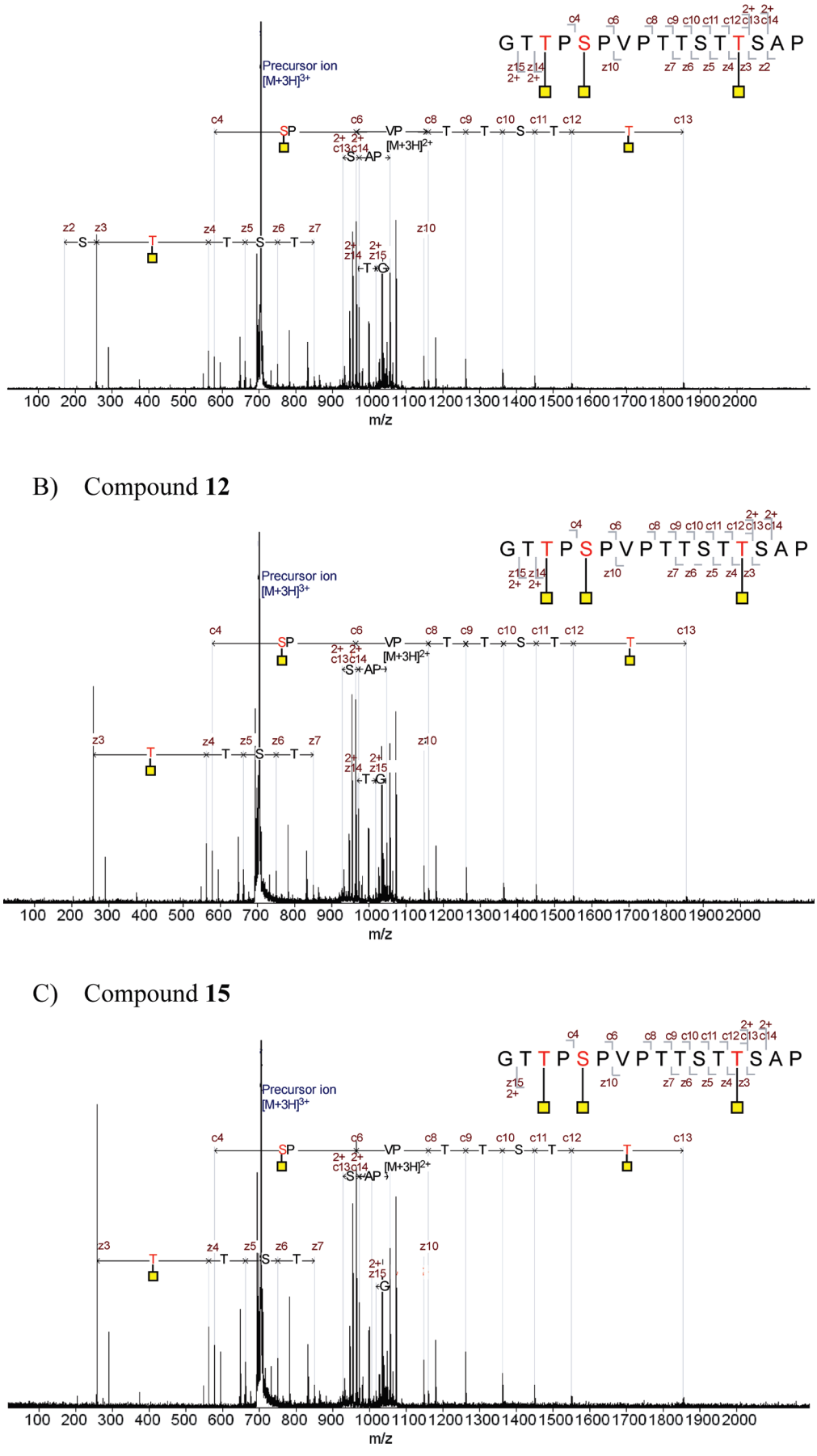
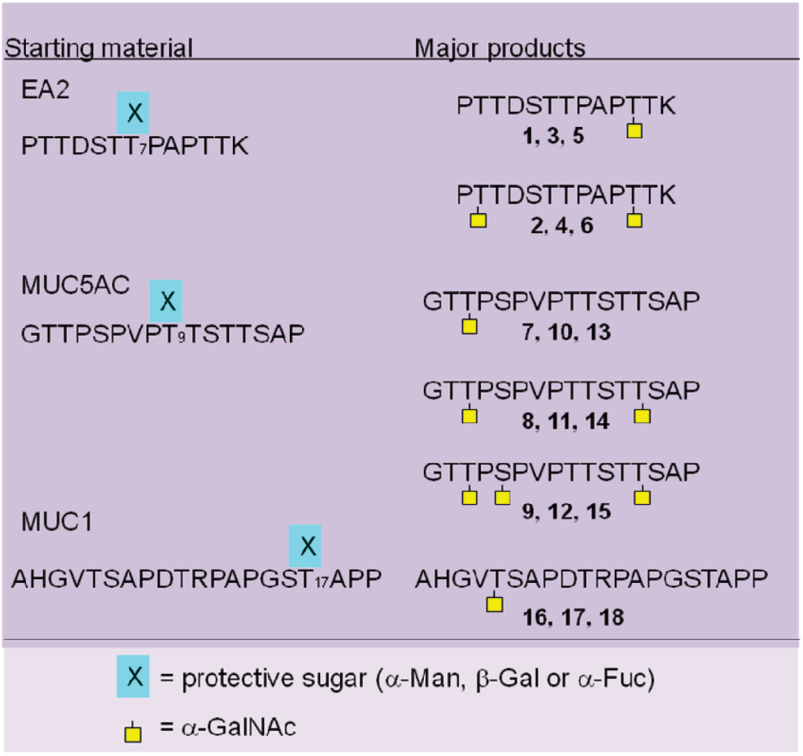
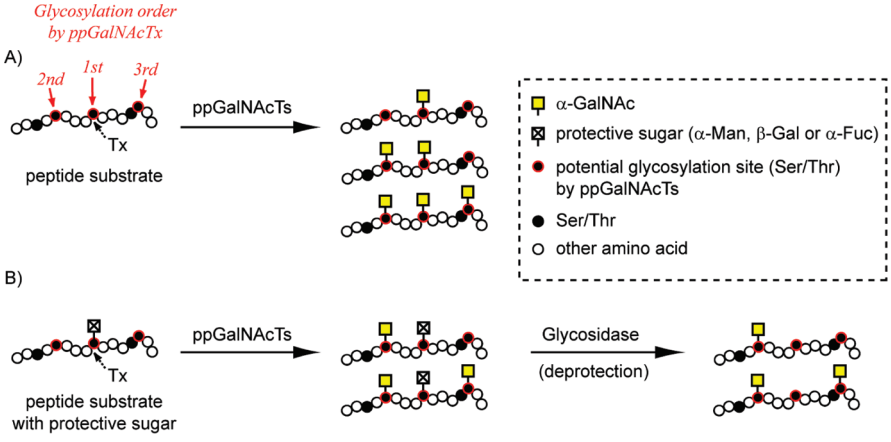


FIGURE 6: ECD-MS spectra of glycopeptides **9** (A), **12** (B), and **15** (C) and the corresponding fragmentation schemes derived from the precursor ions at m/z 704 ($[M + 3H]^{3+}$). Electrons used for ECD were accelerated with 1.75 eV for these three compounds.

Table 1: Summary of Products and Glycosylation Sites Uncovered through Accurate Structural Characterization by ECD-MS Analysis



Scheme 2: A Concept of "Carbohydrates as Protective Groups" for Site-Directed Installation of α-GalNAc Residues by ppGalNAcTs[a]



[a]Products when (A) naked peptide substrate or (B) peptide substrate with a protective sugar at the most predominant Thr/Ser residue is used as the acceptor substrate are indicated.

with an electron capture dissociation (ECD) device (*29*, *30*) because the above MALDI-TOFMS did not provide any information of the glycosylation site(s), although the numbers of newly installed α-GalNAc residues can be estimated from the molecular weight of the precursor ions. As shown in Figure 3, ECD-MS spectra of glycopeptides **1** and **2** allowed for accurate structural assignments by highly informative c and z series of fragment ions generated by preferential fragmentation at peptide bonds, in which α-GalNAc residues were safely attached to Thr or Ser residues during the ECD process while general fragmentation by conventional collision-induced dissociation (CID) occurred predominantly at glycoside linkages (*26*). Scheme 1 summarized products synthesized by means of the protection by mannose at Thr7.

Our interest was next focused on the effect of configurational and stereochemical features of carbohydrates substituted at Thr7 on the tolerance observed in the case for α-Man residue. We synthesized EA2 glycopeptides having α-Fuc and *β*-Gal at Thr7 instead of α-Man residue. They were also employed for the sequential enzymatic modifications and further structural elucidation in a similar manner to the modification study using EA2 having α-Man (EA2-7M). It was suggested that MALDI-TOFMS profiles appear to be quite similar to that observed in the case for EA2-7M (Figure 4). Interestingly, removal of α-Fuc or *β*-Gal at Thr7 from intermediates gave final products **3**/**4** and **5**/**6** exhibiting completely the same masses as compounds **1** (*m*/*z* 1520) and **2** (*m*/*z* 1723) (Figure 2c). Further ECD-MS

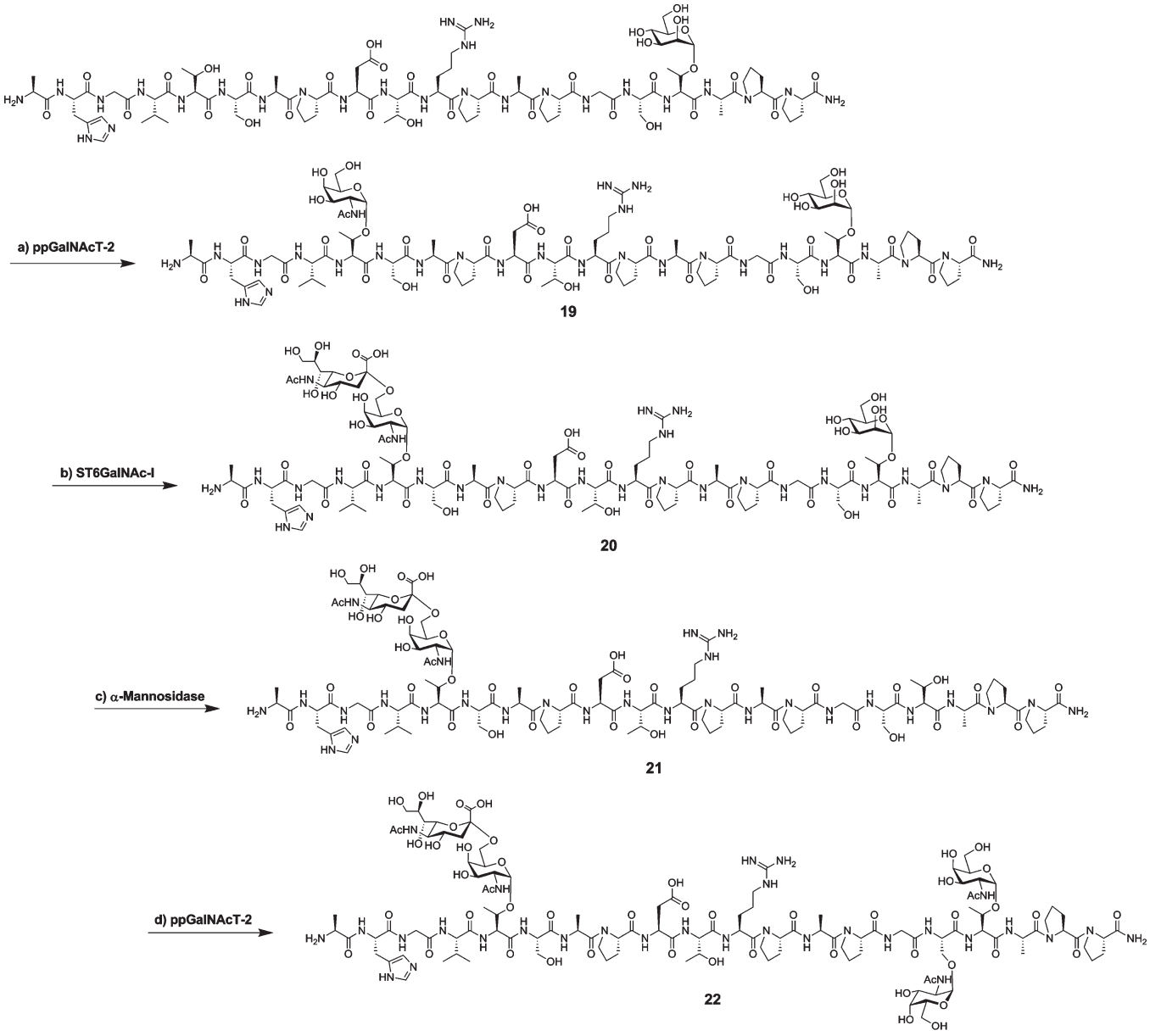


FIGURE 7: Synthetic route of MUC1 glycopeptide **22** through the protection/deprotection at Thr17 with α-Man as a protective group.

analysis (Supporting Information, Figures S-2−S-5) revealed that the composition of these products (**3**/**4** and **5**/**6**) is mostly the same as the result indicated in Scheme 1, and no detectable modification at other potential glycosylation sites was observed. This means that the structure of sugar residue introduced at Thr7 did not influence profiles in the additional glycosylation of EA2 peptide by ppGalNAcT2.

Effects of Peptide Sequences on the Tolerance in Controlled Glycosylation by ppGalNAcT2. Considering the importance of peptide sequence in the specific interaction with ppGalNAcTs (*9*, *11*), we should assess this feature in the tolerance revealed in the glycosylation for EA2 glycopeptides by employing some other suited mucin peptides. We selected two typical mucin tandem repeating peptides, MUC5AC (GTTPSPVPTTSTTSAP) (*31*) and MUC1 (AHGVTSAPDTRPAPSTAPP) (*33*), as scaffold peptides in which Thr residues underlined are known to be a predominant *O*-glycosylation site of ppGalNAcT2, respectively. We synthesized MUC5AC and MUC1 peptides bearing α-Man, α-Fuc, and *β*-Gal residues at designated Thr residues (abbreviated as MUC5AC-9M, -9F, -9G and MUC1-17M, -17F, -17G). When these six new glycopeptides were employed for glycosylation site mapping analysis according to the same procedure described above, ppGalNAcT2 shows significant activity of additional installation of α-GalNAc residue(s) in all of the glycopeptides in the presence of an equivalent molar ratio of UDP-GalNAc (Figure 5). As anticipated, presubstituted sugars, α-Man, α-Fuc, and *β*-Gal residues, in intermediates of MUC5AC and MUC1 glycopeptides were also trimmed efficiently by treating with own glycosidase. It seems that MUC5AC-9M, -9F, and -9G gave nine kinds of compounds (**7**−**15**) having mono-, di-, and trivalent α-GalNAc residues and MUC1-17M, -17F, and -17G afforded only three products (**16**−**18**) carrying a single α-GalNAc residue. ECD-MS spectra of all precursor ions generated from these glycopeptides allowed for precise identification of the glycosylation sites in compounds **7**−**18** (Figure 6 and Supporting Information Figures S-6−S-14). Surprisingly, it was demonstrated that the precursor ions detected at m/z 704 ($[M + 3H]^{3+}$) derived from compounds **9**, **12**, and **15** gave completely the same fragmentation scheme in individual ECD-MS spectra (Figure 6). This result clearly

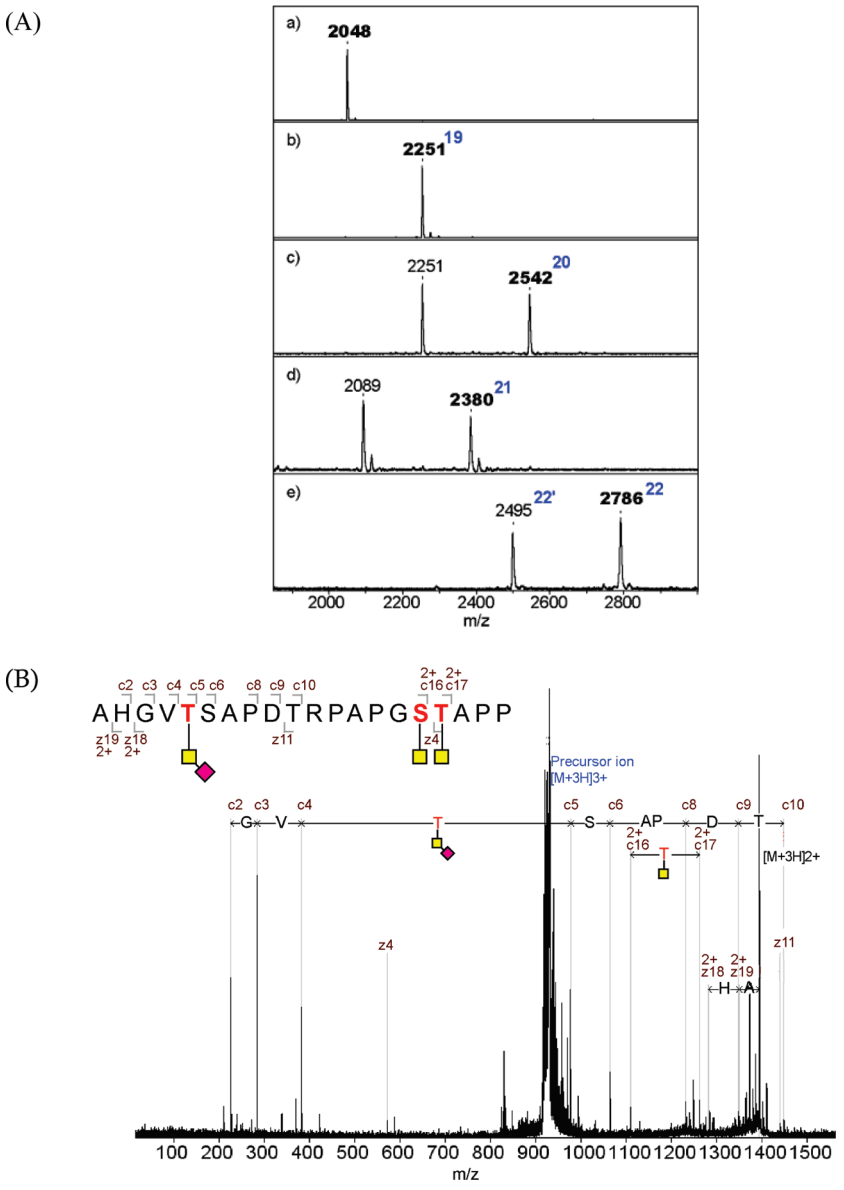


FIGURE 8: Monitoring the synthetic process by a concept "carbohydrate as a protective group" and structural characterization of the products. (A) Synthetic process monitored by MALDI-TOFMS. (B) MS/MS analysis of glycopeptide **22** by ECD-MS at 9.75 eV and the corresponding fragmentation scheme derived from the precursor ion at m/z 929 ($[M + 3H]^{3+}$; the ion peak at m/z 2786 in MALDI-TOFMS). Glycosylated sites and glycan structures were determined by the specific fragment ion peaks (c4-c5, c10, c17-c18^{2+}, and z4), demonstrating that sialyl Tn antigen was introduced only at the desired position (Thr5). The ion peak at m/z 2495, a byproduct, was identified to be compound **22′** carrying three GalNAc residues at T5, S16, and T17 derived from compound **19** due to the low reactivity of ST6GalNAc-I employed (Supporting Information Figures S-15 and S-16).

indicates that installation of α-GalNAc in MUC5AC-M, -F, and -G by ppGalNAcT2 seems to converge finally the single compound with three α-GalNAc at Thr3, Ser5, and Thr13 residues while the glycosylation initiated mainly at Thr3 or partially at Ser5 as suggested in the ECD-MS spectra of other products such as compound **7**, **8**, **10**, **11**, **13**, and **14** (Supporting Information, Figures S-6−S-11).

On the contrary, products **16**, **17**, and **18** were proved to be MUC1 tandem repeat peptide having single α-GalNAc residue at Thr5, suggesting that additional installation of α-GalNAc residue by ppGalNAcT2 proceeds specifically at this position when Thr17 is occupied preferentially by above unusual sugar residues (Supporting Information Figure S-12−S-14). Although ppGalNAcT2 did not differentiate these three sugar structures in terms of conformation, configuration, and stereochemistry at glycoside linkage, the interaction of ppGalNAcT2 with MUC1 peptide appears to be much more sequence specific than other cases. This tendency might be greatly affected by the concentration of UDP-GalNAc as donor substrate as well as that of acceptor substrates, since multiple glycosylation seems to occur in various mucin-type peptides (*9*). Table 1 summarized products assigned from the results of MALDI-TOFMS and their glycosylation sites revealed by precise structural characterization on the basis of ECD-MS experiments. As a result, it was concluded that the temporarily assigned compound numbers **1**−**18** were proved to converge six structures (**1** = **3** = **5**, **7** = **10** = **13**, **8** = **11** = **14**, **9** = **12** = **15**, and **16** = **17** = **18**) as major products during enzymatic glycosylation using three different protective groups (see also Figure 3, Figure 6, and Supporting Information Figures S-1−S-14).

A Concept "Carbohydrates as Protective Groups". Unexpected tolerance of ppGalNAcT2-mediated glycosylation toward unusual glycopeptides strongly encouraged us to apply this

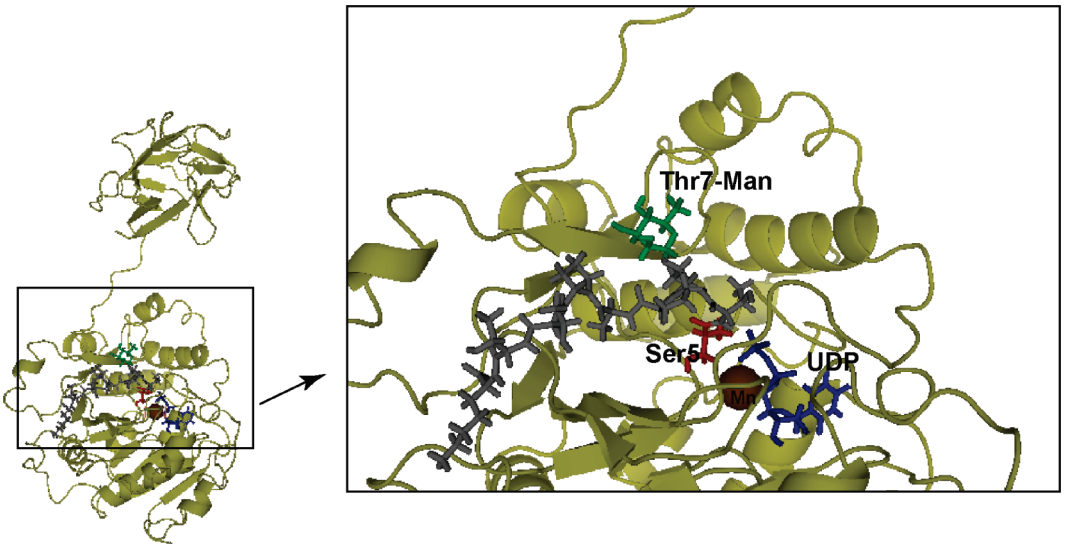


FIGURE 9: Proposed binding profile between ppGalNAcT2 and EA2 with α-Man protection at Thr7 generated by docking simulation with the 3D structure (PDB entry: 2ffu) (*31*).

unusual but unique characteristic for the development of highly selective enzymatic synthesis of complex mucin glycopeptides. Scheme 2 illustrates a new concept of "carbohydrate as protective groups" and a general synthetic protocol based on this new theory. Since each ppGalNAcTs must have its own preferential glycosylation site(s) as represented tentatively as **Tx**, random *O*-glycosylation by ppGalNAcTs in the presence of a large excess of UDP-GalNAc as donor substrates gives a mixture of the products, in which all **Tx** residues must be predominantly glycosylated (Scheme 2A). If this **Tx** residue has already been *protected* with some sugar residues such as α-Man, α-Fuc, or *β*-Gal which can be *deprotected* selectively by designated glycosidase(s) after ppGalNAcTs reaction, we considered that this procedure allows for site-directed installation of α-GalNAc residue(s) at desired positions and further sugar extension except at the **Tx** residue independent from the preferential pathways of individual ppGalNAcTs used (Scheme 2B).

Feasibility of this concept was demonstrated preliminarily by the synthesis of a cancer-related antigenic MUC1 glycopeptide **22** having a sialyl Tn antigen (Neu5Acα2,6GalNAcα) at Thr5 and Tn antigen (GalNAcα) at Ser16 and Thr17 residues by combined use of two glycosyltransferases, ppGalNAcT2 and ST6GalNAc-I (*28*), and a glycosidase (α-mannosidase) starting from MUC1 peptide in which Thr17 is preferentially modified with α-Man residue (one of the substrates abbreviated as MUC1-17M). Merit of this approach is evident because a designed protocol composed of four simple-step reactions using (a) ppGalNAcT2, (b) ST6GalNAc-I, (c) α-mannosidase, and (d) ppGalNAcT2 (Figure 7) greatly facilitated the synthetic pathway of compound **22** from a simple starting material in high efficiency as monitored by MALDI-TOFMS spectra (Figure 8A). Structural characterization by ECD-MS spectra revealed that the precursor ion at *m*/*z* 2786 detected in MALDI-TOFMS corresponds to compound **22**, and sialyl Tn and Tn antigens were installed correctly at desired Thr/Ser residues (Figure 8B), though the performance of recombinant ST6GalNAc-I should be improved in order to achieve quantitative sugar extension. It is clear that sequential installation of GalNAc and Neu5Ac residues by the above two glycosyltransferases occurs predominantly at Thr17 without the protection at this position and produces a highly complicated mixture of glycopeptides, in which position and number of the Thr/Ser residues modified by sialyl Tn and/or Tn antigens cannot be controlled (*33*).

DISCUSSION

Our finding clearly suggests that the lectin domain of ppGalNAcT2 did not contribute to the substrate recognition through its potential carbohydrate-binding ability with relatively small synthetic peptides such as tandem repeating unit of mucin glycoproteins. It is likely that the binding mode in the interaction between the synthetic (glyco)peptides and its own catalytic cavity of ppGalNAcT2 (*34*) defines a preferential pathway of *O*-glycosylation. In fact, the result of docking simulation of a glycosylated EA2 (EA2-7M) with three-dimensional structure of the catalytic sites in ppGalNAcT2 (PDB entry: 2ffu) (*31*) indicated that α-Man substituted at Thr7 residue of EA2 might not affect the binding mode in this pocket and further glycosylation at neighboring Thr/Ser residues (Figure 9). It should also be noted that conformational impacts of an initial or specific *O*-glycosylation step during mucin domain assembly influence significantly the dynamic mechanism in the molecular recognition by ppGalNAcTs (*31*, *34*, *35*) as reported in cases for some synthetic *O*-glycopeptides/glycoproteins such as antifreeze glycoproteins and disease-related mucin glycopeptides (*36*−*38*).

Although the functional roles of lectin domains of a variety of ppGalNAcTs are still unclear, the tolerance in the substrate specificity of this class enzymes uncovered in the present study should encourage us to expand the novel concept, carbohydrate as protective groups for Thr/Ser residues, for the construction of highly potential mucin glycopeptide library. However, it must be considered that improvement of recombinant glycosyltransferases has been strongly required in terms of the practical performance such as versatility and efficiency suited for large-scale synthesis in biomedical and drug discovery research. For this purpose, it seems that use of engineered glycosyltransferases might allow for the enhancement of protein stability and recycling through site-specific immobilization on polymer supports (*39*).

Advantage of ECD-MS analysis (*40*−*42*) was demonstrated by performing accurate structural characterization of highly complicated *O*-glycopeptides. It seems likely that our approach on the basis of the ECD device that can accelerate electron in a

range of 1.75−9.75 eV should greatly facilitate the specific fragmentations at peptide linkages without significant degradation at the carbohydrate moiety (*29*, *30*). Judging from the specific fragment ion peaks (c4-c5) in the ECD-MS analysis of compound **22**, in which electrons were accelerated at 9.75 eV (Figure 8B), it is clear that no detectable fragmentation at the glycoside linkages in the Neu5Acα(2→6)GalNAcα1→Thr moiety occurred during this process. Although MALDI-TOF/TOF MS may become an alternative choice of convenient tools to provide highly sensitive and informative fragmentation patterns from mucin glycopeptides (*43*), complexity of b and y series of fragment ions generated from multiply *O*-glycosylated peptides often makes identification of the glycosylation sites difficult. Therefore, this feature of ECD technology would greatly contribute to the structural and functional elucidation of much more complicated glycopeptides and other glycoconjugates.

CONCLUSION

A novel strategy for chemical and enzymatic synthesis of biologically important mucin glycopeptides based on the concept of "carbohydrate as protective groups" was inspired from an unexpected mechanism in the ppGalNAcT2-mediated *O*-glycosylation. Versatility of this promising strategy is evident because a combinational protection by using other considerable protective groups such as α-Fuc, *β*-Gal, or even *β*-GalNAc, *β*-GlcNAc, *β*-Xyl, and *β*-Glc as well as α-Man residue would expand reasonably the repertoire of much more complicated mucin glycopeptides. Deprotection of carbohydrates as protective groups at Thr/Ser residues can be performed in a site-specific manner by employing designated glycosidases that usually show high selectivity in sugar structure and stereochemistry at the anomeric position with broad tolerance in the nonsugar moiety (aglycons). It should be noted that this characteristic of sialidase was introduced by Venot et al. in the chemoenzymatic synthesis of oligosaccharides of CD65/VIM-2 epitope (*44*). It is our belief that the present approach should be greatly beneficial for rapid and precise synthesis of mucin glycopeptide library (*22*−*26*) in a robotic manner and the development of new cancer biomarkers and mucin-based therapeutic reagents such as cancer vaccine and monoclonal antibody drugs (*45*−*48*).

SUPPORTING INFORMATION AVAILABLE

ECD-MS/MS spectra (Figures S-1−S-16). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hattrup, C. L., and Gendler, S. J. (2008) Structure and function of the cell surface (tethered) mucins. *Annu. Rev. Physiol. 70*, 431–457.
2. Hollingsworth, M. A., and Swanson, B. J. (2004) Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer 4*, 45–60.
3. Brockhausen, I. (1999) Pathways of *O*-glycan biosynthesis in cancer cells. *Biochim. Biophys. Acta 1473*, 67–95.
4. Ten Hagen, K. G., Fritz, T. A., and Tabak, L. A. (2003) All in the family: the UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases. *Glycobiology 13*, 1–16.
5. Ludwig, J. A., and Weinstein, J. N. (2005) Biomarkers in cancer staging, prognosis and treatment selection. *Nat. Rev. Cancer 5*, 845–856.
6. Dube, D. H., and Bertozzi, C. R. (2005) Glycans in cancer and inflammation—potential for therapeutics and diagnostics. *Nat. Rev. Drug Discovery 4*, 477–488.
7. Taylor-Papadimitriou, J. (1999) MUC1 and cancer. *Biochim. Biophys. Acta 1455*, 301–313.
8. Hang, H. C., and Bertozzi, C. R. (2005) The chemistry and biology of mucin-type *O*-linked glycosylation. *Bioorg. Med. Chem. 13*, 5021–5034.
9. Pratt, M. R., Hang, H. C., Ten Hagen, K. G., Rarick, J., Gerken, T. A., Tabak, L. A., and Bertozzi, C. R. (2004) Deconvoluting the functions of polypeptide *N*-α-acetylgalactosaminyltransferase family members by glycopeptide substrate profiling. *Chem. Biol. 11*, 1009–1016.
10. Gerken, T. A., Zhang, J., Levine, J., and Elhammer, A. (2002) Mucin core *O*-glycosylation is modulated by neighboring residue glycosylation status. *J. Biol. Chem. 277*, 49850–49862.
11. Hassan, H., Reis, C. A., Bennett, E. P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (2000) The lectin domain of UDP-*N*-acetyl-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. *J. Biol. Chem. 275*, 38197–38205.
12. Wandall, H. H., Irazoqui, F., Tarp, M. A., Bennett, E. P., Mandel, U., Takeuchi, H., Kato, K., Irimura, T., Suryanarayanan, G., Hollingsworth, M. A., and Clausen, H. (2007) The lectin domains of polypeptide GalNAc-transferases exhibit carbohydrate-binding specificity for GalNAc: lectin binding to GalNAc-glycopeptide substrates is required for high density GalNAc-*O*-glycosylation. *Glycobiology 17*, 374–387.
13. Taylor, M. E., and Drickamer, K. (2003) Introduction to Glycobiology, Oxford University Press, Tokyo, Japan.
14. Endo, T. (1999) *O*-Mannosyl glycosylation in mammals. *Biochim. Biophys. Acta 1473*, 237–246.
15. Harris, R. J., and Spellman, M. W. (1993) *O*-Linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology 3*, 219–224.
16. Wells, L., and Hart, G. W. (2003) *O*-GlcNAc turns twenty: functional implications for post-translational modification of nuclear and cytosolic proteins with a sugar. *FEBS Lett. 546*, 154–158.
17. Hazes, B. (1996) The $(QXW)_3$ domain: a flexible lectin scaffold. *Protein Sci. 5*, 1490–1501.
18. Imberty, A., Piller, V., Piller, F., and Breton, C. (1997) Fold recognition and molecular modeling of a lectin-like domain in UDP-GalNAc: polypeptide *N*-acetylgalactosaminyltransferases. *Protein Eng. 12*, 1353–1356.
19. Elofsson, M., Roy, S., Salvador, L. A., and Kihlberg, J. (1996) Building blocks for glycopeptides synthesis: preparation of α-O-fucosylated Fmoc serine and threonine in one step from L-fucose tetraacetate. *Tetrahedron Lett. 37*, 7645–7648.
20. Salvador, L. A., Elofsson, M., and Kihlberg, J. (1995) Preparation of building blocks for glycopeptide synthesis by glycosylation of Fmoc amino acids having unprotected carboxyl groups. *Tetrahedron 51*, 5643–5656.
21. Matsushita, T., Hinou, H., Kurogochi, M., Shimizu, H., and Nishimura, S.-I. (2005) Rapid microwave-assisted solid-phase glycopeptide synthesis. *Org. Lett. 7*, 877–880.
22. Fumoto, M., Hinou, H., Ohta, T., Ito, T., Yamada, K., Takimoto, A., Kondo, H., Shimizu, H., Inazu, T., Nakahara, Y., and Nishimura, S.-I. (2005) Combinatorial synthesis of MUC1 glycopeptides: polymer blotting facilitates chemical and enzymatic synthesis of highly complicated mucin glycopeptides. *J. Am. Chem. Soc. 127*, 11804–11818.
23. Matsushita, T., Hinou, H., Fumoto, M., Kurogochi, M., Fujitani, N., Shimizu, H., and Nishimura, S.-I. (2006) Construction of highly glycosylated mucin-type glycopeptides based on microwave-assisted solid-phase syntheses and enzymatic modifications. *J. Org. Chem. 71*, 3051–3063.
24. Naruchi, K., Hamamoto, T., Kurogochi, M., Hinou, H., Shimizu, H., Matsushita, T., Fujitani, N., Kondo, H., and Nishimura, S.-I. (2006) Construction and structural characterization of versatile lactosaminoglycan-related compound library for the synthesis of complex glycopeptides and glycosphingolipids. *J. Org. Chem. 71*, 9609–9621.
25. Ohyabu, N., Hinou, H., Matsushita, T., Izumi, R., Shimizu, H., Kawamoto, K., Numata, Y., Togame, H., Takemoto, H., Kondo, H., and Nishimura, S.-I. (2009) An essential epitope of anti MUC1 monoclonal antibody KL-6 revealed by focused glycopeptide library. *J. Am. Chem. Soc. 131*, 17102–17109.
26. Matsushita, T., Sadamoto, R., Ohyabu, N., Nakata, H., Fumoto, M., Fujitani, N., Takegawa, Y., Sakamoto, T., Kurogochi, M., Hinou, H., Shimizu, H., Ito, T., Naruchi, K., Togame, H., Takemoto, H., Kondo, H., and Nishimura, S.-I. (2009) Functional neoglycopeptides: synthesis and characterization of new class MUC1 glycoprotein models having core 2-based *O*-glycan and complex-type *N*-glycan chains. *Biochemistry 48*, 11117–11133.
27. White, T., Bennett, E. P., Takio, K., Sørensen, T., Bonding, N., and Clausen, H. (1995) Purification and cDNA cloning of a human UDP-*N*-acetyl-α-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase. *J. Biol. Chem. 270*, 24156–24165.
28. Ikehara, Y., Kojima, N., Kurosawa, N., Kudo, T., Kono, M., Nishihara, S., Issiki, S., Morozumi, K., Itzkowitz, S., Tsuda, T.,

Nishimura, S.-I., Tsuji, S., and Narimatsu, H. (1999) Cloning and expression of a human gene encoding an *N*-acetylgalactosamine-α2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology 9*, 1213–1224.
29. Baba, T., Hashimoto, Y., Hasegawa, H., Hirabayashi, A., and Waki, I. (2004) Electron capture dissociation in a radio frequency ion trap. *Anal. Chem. 76*, 4263–4266.
30. Satake, H., Hasegawa, H., Hirabayashi, A., Hashimoto, Y., Baba, T., and Masuda, K. (2007) Fast multiple electron capture dissociation in a linear radio frequency quadrupole ion trap. *Anal. Chem. 79*, 8755–8761.
31. Fritz, T. A., Raman, J., and Tabak, T. A. (2006) Dynamic association between the catalytic and lectin domains of human UDP-GalNAc: polypeptide α-*N*-acetylgalactosaminyltransferase-2. *J. Biol. Chem. 281*, 8613–8619.
32. Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. J. (2001) Evaluation and reparamerization of the OPLS-AA force field for proteins *via* comparison with accurate quantum chemical calculation on peptides. *J. Phys. Chem. B 105*, 6474–6487.
33. Sørenson, A. L., Reis, C. A., Tarp, M. A., Mandel, U., Ramachandran, K., Sankaranarayanan, V., Schwientek, T., Graham, R., Taylor-Papadimitriou, J., Hollingsworth, M. A., Burchell, J., and Clausen, H. (2006) Chemoenzymatically synthesized multimeric Tn/STn MUC1 glycopeptides elicit cancer-specific anti-MUC1 antibody responses and override tolerance. *Glycobiology 16*, 96–107.
34. Raman, J., Fritz, T. A., Gerken, T. A., Jamison, O., Live, D., Liu, M., and Tabak, L. A. (2008) The catalytic and lectin domains of UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferase function in concert to direct glycosylation site selection. *J. Biol. Chem. 283*, 22942–22951.
35. Fritz, T. A., Hurley, J. H., Trinh, L.-B., Shiloach, J., and Tabak, L. A. (2004) The beginnings of mucin biosynthesis: the crystal structure of UDP-GalNAc:polypeptide α-*N*-acetylgalactosaminyltransferase-T1. *Proc. Natl. Acad. Sci. U.S.A. 101*, 15307–15312.
36. Coltart, D. M., Royyuru, A. K., Williams, L. J., Glunz, P. W., Sames, D., Kuduk, S. D., Schwarz, J, B., Chen, X.-T., Danishefsky, S. J., and Live, D. H. (2002) Principles of mucin architecture: structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. *J. Am. Chem. Soc. 124*, 9833–9844.
37. Kinarsky, L., Suryanarayanan, G., Prakash, O., Paulsen, H., Clausen, H., Hanisch, F.-G., Hollingsworth, M. A., and Sherman, S. (2003) Conformational studies on the MUC1 tandem repeat glycopeptides: implication for the enzymatic *O*-glycosylation of the mucin protein core. *Glycobiology 13*, 929–939.
38. Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., and Nishimura, S.-I. (2004) Antifreeze glycoproteins: elucidation of the structural motifs that are essential for antifreeze activity. *Angew. Chem., Int. Ed. 43*, 856–862.
39. Ito, T., Sadamoto, R., Naruchi, K., Togame, H., Takemoto, H., Kondo, H., and Nishimura, S.-I. (2010) Highly oriented recombinant glycosyltransferases: site-specific immobilization of unstable membrane proteins by using *Staphylococcus aureus* sortase A. *Biochemistry 49*, 2604–2614.
40. Zubarev, R. A., Kelleher, N. L., and McLafferty, F. W. (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J. Am. Chem. Soc. 120*, 3265–3266.
41. Zubarev, R. A. (2003) Reactions of polypeptide ions with electrons in the gas phase. *Mass Spectrom. Rev. 22*, 57–77.
42. Zubarev, R. A. (2004) Electron-capture dissociation tandem mass spectrometry. *Curr. Opin. Biotechnol. 15*, 12–16.
43. Kurogochi, M., Matsushita, T., and Nishimura, S.-I. (2004) Post-translational modifications on proteins: facile and efficient procedure for the identification of *O*-glycosylation sites by MALDI-LIFT-TOF/TOF mass spectrometry. *Angew. Chem., Int. Ed. 43*, 4071–4075.
44. Kashem, M. A., Wlasichuk, K. B., Gregson, J. M., and Venot, A. P. (1993) Chemoenzymatic synthesis of sialylated and fucosylated oligosaccharides having an *N*-acetyllactosaminyl core. *Carbohydr. Res. 250*, 129–144.
45. Boltje, T. J., Buskas, T., and Boons, G.-J. (2009) Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nat. Chem. 1*, 611–622.
46. Tarp, M. A., and Clausen, H. (2008) Mucin-type *O*-glycosylation and its potential use in drug and vaccine development. *Biochim. Biophys. Acta 1780*, 546–563.
47. Ingale, S., Wolfert, M. A., Gaekwad, J., Buskas, T., and Boons, G.-J. (2007) Robust immune responses elicited by a fully synthetic three-component vaccine. *Nat. Chem. Biol. 3*, 663–667.
48. Wandal, H. H., Blixt, O., Tarp, M. A., Pedersen, J. W., Bennett, E. P., Mandel, U., Ragupathi, G., Livingston, P. O., Hollingsworth, M. A., Taylor-Papadimitriou, J., Burchell, J., and Clausen, H. (2010) Cancer biomarkers defined by autoantibody signatures to aberrant *O*-glycopeptide epitopes. *Cancer Res. 70*, 1306–1313.