BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 11117–11133 11117
DOI: 10.1021/bi901557a

Functional Neoglycopeptides: Synthesis and Characterization of a New Class of MUC1 Glycoprotein Models Having Core 2-Based *O*-Glycan and Complex-Type *N*-Glycan Chains†

Takahiko Matsushita,[‡] Reiko Sadamoto,[‡,△] Naoki Ohyabu,[‡,§] Hideki Nakata,[‡] Masataka Fumoto,[∥] Naoki Fujitani,[‡] Yasuhiro Takegawa,[‡] Takeshi Sakamoto,[⊥] Masaki Kurogochi,[‡] Hiroshi Hinou,[‡] Hiroki Shimizu,[#] Takaomi Ito,[‡,§] Kentarou Naruchi,[‡] Hiroko Togame,[§] Hiroshi Takemoto,[§] Hirosato Kondo,[∥] and Shin-Ichiro Nishimura*[,‡,#]

[‡]*Division of Advanced Chemical Biology, Graduate School of Life Science, Frontier Research Center for the Post-Genome Science and Technology and* [§]*Shionogi Innovation Center for Drug Discovery, Hokkaido University, N21, W11, Kita-ku, Sapporo 001-0021, Japan,* [∥]*Discovery Research Laboratories, Shionogi & Company, Ltd., 12-4, Sagisu 5-chome, Fukushima-ku, Osaka 541-0045, Japan,* [⊥]*Central Research Laboratory, Hitachi, Ltd., Kokubunji, Tokyo 185-8601, Japan, and* [#]*National Institute of Advanced Industrial Science and Technology (AIST), 2-17-2-1 Tsukisamu-higashi, Toyohira-ku, Sapporo 062-8517, Japan.* [△]*Present address: Ochanomizu University, Tokyo, Japan.*

Received September 5, 2009; Revised Manuscript Received October 22, 2009

ABSTRACT: An efficient protocol for the construction of MUC1-related glycopeptide analogues having complex *O*-glycan and *N*-glycan chains was established by integrating chemical and enzymatic approaches on the functional polymer platforms. We demonstrated the feasibility of sortase A-mediated ligation between two glycopeptide segments by tagging with signal peptides, LPKTGLR and GG, at each C- or N-terminal position. Structural analysis of the macromolecular *N*,*O*-glycopeptides was performed by means of ESI-TOFMS (MS/MS) equipped with an electron-captured dissociation device. Immunological assay using MUC1 glycopeptides synthesized in this study revealed that *N*-glycosylation near the antigenic *O*-glycosylated PDTR motif did not disturb the interaction between the anti-MUC1 monoclonal antibody and this crucial *O*-glycopeptide moiety. NMR study indicated that the N-terminal immunodominant region [Ala-Pro-Asp-Thr(*O*-glycan)-Arg] forms an inverse γ-turn-like structure, while the C-terminal region composed of *N*-glycopeptide and linker SrtA-peptide was proved to be an independently random structure. These results indicate that the bulky *O*- and *N*-glycan chains can function independently as disease-relevant epitopes and ligands for carbohydrate-binding proteins, when both are combined by an artificial intervening peptide having a possible effect of separating N- and C-terminal regions. The present strategy will greatly facilitate rapid synthesis of multiply functionalized complex neoglycopeptides as new types of convenient tools or models for the investigation of thhe structure−function relationship of various glycoproteins and development of novel class glycopeptide-based biopharmaceuticals, drug delivery systems, and biomedical materials.

Protein glycosylation is one of the most important processes in the posttranslational modifications, and carbohydrates covalently attached to proteins have been shown to play significant functional roles in protein folding, cellular differentiation, aging, cancers, and basic immunological systems (*1*−*6*). In general, *N*-glycans and *O*-glycans are known as two major groups and widely substituted at the potential glycosylation sites such as Asn of the Asn-Xyz-Ser(Thr) consensus sequence in common glycoproteins or Ser/Thr residues highly distributed in mucin-type glycoproteins. Clustering of multiple *N*-glycans or *O*-glycans on a single polypeptide appears to be a unique mechanism to provide proteins with some beneficial characteristics by glycosylation such as enhancement of solubility, stability against proteolytic enzymes, and the strength of affinity in the molecular recognition between glycans and a variety of carbohydrate-binding proteins involving enzymes, lectins, and antibodies, although required glycoforms (glycan structures) should define individual functions in terms of the specificity (*7*−*9*). Pioneering work by Lee et al. demonstrated the above "clustering effects" by the multiple (*N*-) glycan chains substituted on proteins by means of unique glycoprotein models, namely, neoglycoproteins (*7*, *10*). It is clear that neoglycoproteins and related compounds have become nice tools for demonstrating the importance of multivalency in *N*-glycan functions as ligands for carbohydrate−protein interaction and cell−cell adhesion and in its application toward the discovery research of carbohydrate-based therapeutic reagents (*7*−*9*, *11*−*17*).

In contrast to well-investigated *N*-glycan functions in glycoproteins, functional roles of most *O*-glycans remain to be unclear. Considerable interests, however, have been focused on the effects of protein *O*-glycosylation on the conformational impact to the core peptide structures because peptide conformational alteration induced by multiple *O*-glycosylation should often influence secondary structure and some functions of the original peptides/proteins (*18*−*21*). Live et al. revealed that the mode of α-*O*-glycosylation by the GalNAc residue within a clustered locus has a profound effect on the conformation of the core peptide by employing synthetic glycopeptides related to the N-terminal fragment Ser-Thr-Thr-Ala-Val of the cell surface glycoprotein

†This work is supported by a program grant "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: +81-11-706-9043. Fax: +81-11-706-9042.

Published on Web 10/23/2009

CD43 (*19*). We found that antifreeze glycoprotein (AFGP) forms a polyproline type II helix structure by attaching multivalent disaccharide (Galβ1,3GalNAcα1→) at Thr residues within the tandem repeating polypeptide (Ala-Ala-Thr)$_n$ to achieve specific antifreeze activity (*21*). It was demonstrated that hydrogen bonds between the GalNAc residue and threonyl peptide moiety are crucial for the formation of this active conformation. However, an AFGP analogue lost antifreeze activity by the replacement of GalNAc to Glc due to the disruption of the above key conformation. When the stereochemistry of *O*-glycoside linkage between GalNAc and Thr residues was converted from α- into β-configuration, this unnatural glycoprotein exhibited neither detectable AFGP activity nor the distinct secondary structure (*21*).

Mammalian glycoproteins often display both *O*- and *N*-glycans on a single polypeptide chain, though the functional roles of these glycans had little been discussed concurrently. New class synthetic glycopeptides bearing both *O*- and *N*-glycans (*22*), notably neoglycopeptides, would allow for insight into intramolecular influence between *O*- and *N*-glycans as well as structural and functional impacts of the individual glycosylation. It is noteworthy that recent progress in the construction of glycans and glycoproteins by chemical synthesis, native ligation-based synthesis, and site-directed glycosylation by cotranslational strategy(*23*−*31*) had greatly expanded the repertoire in the molecular design of synthetic glycoproteins as well as recombinant glycosyltransferase-based synthesis (*32*−*44*). However, there is no suited model compound feasible for investigating this subject at the molecular level due to the difficulty in rapid and large-scale synthesis of such highly complicated macromolecular glycopeptides. We considered that glycoprotein mimics having both desired *N*- and *O*-glycans are potential candidates for new class biopharmaceuticals when these glycan chains can function independently as a ligand of tissue specific lectins or a disease-specific epitope. Therefore, it is clear that advent of a facile and versatile protocol for the construction of such complex neoglycopeptides is now strongly required.

Staphylococcus aureus sortase (SrtA)[1] is a bacterial *trans*-peptidase responsible for the covalent attachment of specific proteins to the peptidoglycan cross-bridge of the cell wall of Gram-positive bacteria (*45*, *46*). Proteins that become substrates of SrtA should have a sorting signal as the Lys-Pro-X-Thr-Gly (LPXTG) motif (X = Asp, Glu, Ala, Asn, Gln, or Lys) at the C-terminus. SrtA can cut an amide bond between T and G at LPXTG and form a new amide bond linking the carboxyl group of threonine to an amino group of the glycine oligomer (Gly$_n$, n = 1−5) at the N-terminus in a partner substrate molecule (*47*). It has also been demonstrated that SrtA can be applied for the production of a wide range of nonnatural polypeptides bearing different biological functions (*48*−*53*). Merit of this promising method is evident because the peptide-bond formation catalyzed by the *trans*-peptidases gives structurally defined and stereochemically pure linkages between two different peptide components. In addition, it was also suggested that at least 82 protein sequences of human genome involve SrtA motifs "LPXTG (X = D, E, A, N, Q, or K)" (*54*, *55*).[2] This indicates that the anchor motifs generated by SrtA-mediated ligation may not become any special antigenic structure to produce human antibody. We hypothesized that SrtA-mediated ligation employing various glycopeptide segments will expand the synthetic repertoire of much more complicated and larger glycopeptides having different complex carbohydrate moieties and appears to be applicable to rapid and large-scale production of various hybrid polypeptides that are multiply *N*-/*O*-glycosylated at fixed sites. It is clear that this class of glycoprotein mimics should greatly contribute to discovery research for potential candidates of novel biopharmaceuticals or biomedical materials (*13*−*17*). Our interest is now focused on the feasibility of *trans*-peptidase-mediated ligation between complex glycopeptide segments at the platform of "polymer-assisted enzymatic synthesis (*40*−*44*)".

EXPERIMENTAL PROCEDURES

Materials and Methods. All commercially available solvents and reagents were used without purification. Fmoc-TentaGel S RAM resin was purchased from Hipep Laboratories, and Fmoc-amino acid derivatives and Fmoc-Asn(Ac$_3$AcNH-β-Glc)-OH were purchased from Novabiochem. Fmoc-Thr(Ac$_7$core2)-OH was synthesized by the same manner in a previous report (*42*). All solid-phase reactions for glycopeptide synthesis were performed manually in a polypropylene tube equipped with a filter (LibraTube; Hipep Laboratories). Coupling reactions of Fmoc-amino acid derivatives and Fmoc removal reactions were conducted under microwave irradiation (0−40 w) at 50 °C, and acetyl capping reactions after coupling were carried out by conventional manner at ambient temperature (*56*). Microwave-assisted solid-phase reactions were performed mannually using a single mode microwave reactor (IDX Corp., GreenMotif I). The microwave reactor was a customized reaction cavity for installation of LibraTube and introduced an additional function of mechanically shaking for the reaction cavity in our laboratories. An aminooxy-functionalized polyacrylamide derivative was prepared in the same manner in the previous papers (*40*−*44*). Recombinant β1,4-galactosyltransferase (β1,4-GalT) was purchased from Toyobo, Ltd., and α2,3-*N*-sialyltransferase (α2,3-*N*-SialT) and α2,3-*O*-sialyltransferase (α2,3-*O*-SiaT) were purchased from Calbiochem. Endo-β-*N*-acetylglucosaminidase (endo-M)(*57*−*59*) was purchased from Tokyo Kasei Kogyo Co., Ltd. Uridine-5′-diphosphogalactose, 2Na (UDP-Gal), and cytidine-5′-monophospho-*N*-acetylneuraminic acid, 2Na (CMP-NANA), were purchased from Yamasa Co. Sialylglycopeptide (SGP) derived from hen egg yolk **4** was prepared as described Seko et al. (*60*, *61*). Purifications in enzymatic reactions for glycopeptide on a water-soluble polymer were performed by using centrifugal ultrafiltration (UF) units (Orbital Biosciences, LLC; 10K Apollo 20 mL, high-performance centrifugal concentrators). Matrix-associated laser desorption ionization time-of-fright mass spectrometry (MALDI-TOFMS) data were

[1]Abbreviations: SrtA, sortase A; β1,4-GalT, β1,4-galactosyltransferase; α2,3-*N*-SialT, α2,3-*N*-sialyltransferase; α2,3-*O*-SiaT, α2,3-*O*-sialyltransferase; endo-M, endo-β-*N*-acetylglucosaminidase; UDP-Gal, uridine-5′-diphosphogalactose; 2Na, CMP-NANA, cytidine-5′-monophospho-*N*-acetylneuraminic acid; 2Na, SGP, sialylglycopeptide; MALDI-TOFMS, matrix-associated laser desorption ionization time-of-fright mass spectrometry; ESI-MS, electrospray ionization mass spectrometry; CID, collision-induced dissociation; ECD, electron-captured dissociation; Fmoc, 9-fluorenylmethoxycarbonyl; HBTU, 2−2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, 1-hydroxy-1*H*-benzotriazole; DIPEA, *N*,*N*-diisopropylethylamine; Pbf, 2,2,5,7,8-pentamethyldihydrobenzofuran-5-sulfonyl; TFA, trifluoroacetic acid; Tris-HCl, tris-(hydroxymethyl)aminomethane hydrochloride.

[2]The sortase motif (LPxTG) found in the current human genome sequence was summarized in Table S-11 (Supporting Information). Protein sequences of human genome were obtained by the BLAST search program from the ncbi ftp site, and motifs were detected by genome regular expression (/LP[N|E|A|N|Q|K]TG/).

recorded by a Bruker REFLEX III or AUTOFLEX II using 2,5-dihydroxybenzoic acid as a matrix. Reversed-phase HPLC (RP-HPLC) analyses and purifications were performed on a Hitachi HPLC system equipped with an L-2130 intelligent pump and an L-2420 UV detector or equipped with an L-7100 intelligent pump and an L-7405 UV detector, using a reversed-phase C18 column, Inertsil ODS-3 (GL Sciences Inc.) or Mightysil RP-18 (Kanto Chemical Co., Inc.), at 25 °C. The chromatographies were monitored by UV absorption at 220 nm. Electrospray ionization mass spectra (ESI-MS; JEOL JMS-700TZ) and amino acid analyses (Procise491 cLC; Applied Biosystems) were performed in the Center of Instrumental Analysis at Hokkaido University.

Production and Purification of Sortase A. *S. aureus* sortase (SrtA) was produced in *Escherichia coli* (*E. coli*) transfected with the plasmid carrying the SrtA gene. The enzyme was prepared according to the procedure of Mao et al. (*49*) with some modifications. The transformed cells were inoculated in 2 × YT medium containing 50 mg/L ampicillin at 37 °C; then IPTG (final concentration 1 mM) was added. The culture was cultivated at 20 °C for 20 h. The cells were harvested and frozen. After thawing, the cells were suspended in 25 mM Tris-HCl (pH 7.5) containing 0.5 M NaCl, 20 mM imidazole, and 10% (w/v) glycerol and lysed by sonication. The cell debris was removed by centrifugation at 10000*g* for 10 min. The supernartant was applied to affinity chromatography using Ni-Sepharose 6 Fast Flow (Amersham Biosciences, Uppsala, Sweden). The column was washed with sonication buffer. The protein was eluted with 25 mM Tris-HCl (pH 7.5) containing 0.5 M NaCl, 0.5 M imidazole, and 10% (w/v) glycerol. The purified SrtA was concentrated by Cetriprep YM-10 (Millipore Co., Bedford, MA), Ultrafree-MC (Millipore Co., Bedford, MA), until protein concentration was 16 mg/mL.

Synthesis. Glycopeptides used in this study listed in Figure 1 were synthesized according to the general strategy illustrated in Figure 2 and Schemes 1−5.

L-Glycyl-L-glycyl-L-prolyl-L-valyl-L-histidyl-(N⁴-2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-β-D-glucopyranosyl)-L-asparaginyl-L-valyl-L-threonyl-L-seryl-L-alanylamide (***12***). For the synthesis of *N*-linked glycopeptide **12**, TentaGel S RAM resin (0.26 mmol/g, 385 mg, 100 μmol), Fmoc-amino acids (300 μmol, 3.0 equiv), and Fmoc-Asn(Ac_3AcNH-β-Glc)-OH (150 μmol, 1.5 equiv) were used. Fmoc-amino acids were described below: Fmoc-Ala-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Pro-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Thr(*t*Bu)-OH, and Fmoc-Val-OH. TentaGel S RAM resin was placed in a 20 mL LibraTube and allowed to swell in DMF for a period of 30 min. The resin was filtered, and the Fmoc-removal reaction was conducted under microwave irradiation. Piperidine (20%)/DMF (4 mL) was added, and the mixture was shaken under microwave irradiation for 3 min. Following filtration and washing with DMF (10 mL, five times), the coupling reaction was performed under microwave irradiation. The corresponding Fmoc-amino acid (300 μmol, 3.0 equiv) dissolved in HBTU (300 μmol, 3.0 equiv), HOBt (300 μmol, 3.0 equiv), and DIPEA (600 μmol, 6 equiv) in DMF (3 mL) was added to the resin, and the mixture was shaken under microwave irradiation for 10 min. For introduction of the Fmoc-glycosylated amino acid, Fmoc-Asn(Ac_3AcNH-β-Glc)-OH (150 μmol, 1.5 equiv) dissolved in HBTU (150 μmol, 1.5 equiv), HOBt (150 μmol, 1.5 equiv), and DIPEA (300 μmol, 3 equiv) in DMF (1.5 mL) was treated to the resin under microwave irradiation for 15 min. Following filtration and washing with DMF (10 mL, five times), the unreacted amino groups on the resin were acetylated with a solution of Ac_2O (4.75%, v/v), DIPEA (2.25%, v/v), and HOBt (13 mM) in NMP (8 mL). After shaking for 5 min at ambient temperature, the resin was filtered and washed with DMF (10 mL, five times). Fmoc-removal, coupling, and capping procedures as described above were carried out repeatedly. After completion of the synthesis, the glycopeptidyl-resin **11** was treated with 90% TFA aqueous (4 mL) at room temperature for 2 h, and the resin was filtered. The resin was washed twice with the same cocktail, and the filtrates were combined and concentrated by streaming of nitrogen gas. The glycopeptide was precipitated by addition of cold *tert*-butyl methyl ether in an ice bath to give a colorless solid. The solid was washed twice with cold *tert*-butyl methyl ether, dried by streaming of nitrogen gas, dissolved in 50% acetonitrile aqueous (10 mL), and lyophilized. The crude material was purified by a preparative RP-HPLC (t_R = 16.3 min) to give **12** (78.7 mg, 62.2 μmol) in 62% yield. The RP-HPLC purification condition was described below: column, Inertsil ODS-3 (250 × 20 mm); eluent A, water containing 0.1% TFA; eluent B, acetonitrile containing 0.1% TFA. Eluent (A/B = 90/10) was employed; then the ratio of eluent B was increased lineally from 10% to 50% over 45 min with a flow rate of 7.0 mL/min. The purified **12** was analyzed by MALDI-TOFMS. MALDI-TOFMS: $C_{53}H_{83}N_{15}O_{21}$ $[M + H]^+$ calcd (*m*/*z*) 1266.60, found (*m*/*z*) 1226.59.

L-Glycyl-L-glycyl-L-prolyl-L-valyl-L-histidyl-{N⁴-2-acetamido-2-deoxy-β-D-glucopyranosyl}-L-asparaginyl-L-valyl-L-threonyl-L-seryl-L-alanylamide (***2***). Compound **12** (60.5 mg, 47.8 μmol) in methanol solution (12 mL) was adjusted and was kept to pH 12.4 (as indicated by pH meter) with 1 N sodium hydroxide for 4 h at ambient temperature. The reaction was monitored by RP-HPLC to confirm the completion of the reaction (Supporting Information Figure S-1). After being stirred for 5 h at ambient temperature, the mixture was neutralized by addition of 1 N acetic acid and evaporated *in vacuo*. The crude material was purified by a preparative RP-HPLC (t_R = 24.7 min) to give **2** (51.6 mg, 44.6 μmol) in 93% yield. RP-HPLC purification condition was described below: column, Inertsil ODS-3 (250 × 20 mm); eluent A, water containing 0.1% TFA; eluent B, acetonitrile containing 0.1% TFA. Eluent (A/B = 90/10) was employed from 0 to 5 min; then the ratio of eluent B was increased lineally from 5% to 25% over 60 min with a flow rate of 7.0 mL/min. The purified **2** was analyzed by RP-HPLC and ESI-HRMS. ESI-HRMS: $C_{47}H_{78}N_{15}O_{18}$ $[M + H]^+$ calcd (*m*/*z*) 1140.5644, found (*m*/*z*) 1140.5645. The 600 MHz ^{1}H NMR spectrum of **2** in H_2O/D_2O (90/10) at 300 K is indicated in Supporting Information (Figure S-10). The 600 MHz ^{1}H and 150 MHz ^{13}C NMR chemical shifts of **2** in H_2O/D_2O (90/10) at 300 K are listed in Supporting Information [Table S-1 (peptide moiety) and Table S-2 (carbohydrate moiety)]. Amino acid ratios (numbers in parentheses are theoretical values): Ala(1) 1.0, Asx(1) 1.0, Gly(2) 2.0, His(1) 1.0, Pro(1) 1.0, Ser(1) 0.9, Thr(1) 0.9, Val(2) 1.9.

L-Glycyl-L-glycyl-L-prolyl-L-valyl-L-histidyl-{N⁴-(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→6)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl-(1→6)-[(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→6)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glcopyranosyl-(1→2)-α-D-mannopyranosyl-(1→3)]-β-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl}-L-asparaginyl-L-valyl-L-threonyl-L-seryl-L-alanylamide (***3***). To a

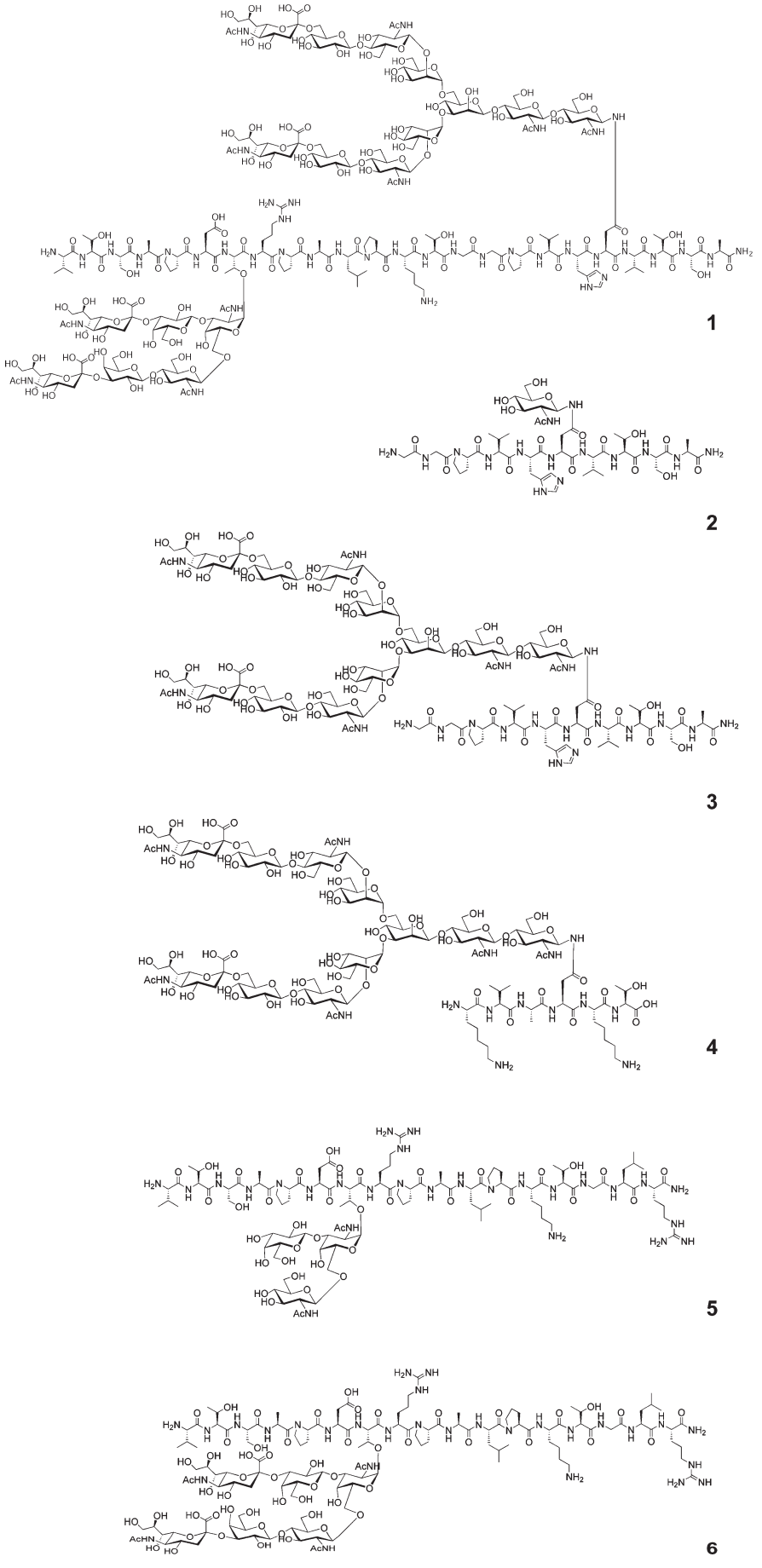


FIGURE 1: Chemical structure of MUC1-related glycopeptide **1** and some key compounds (**2**−**6**) used in this study.

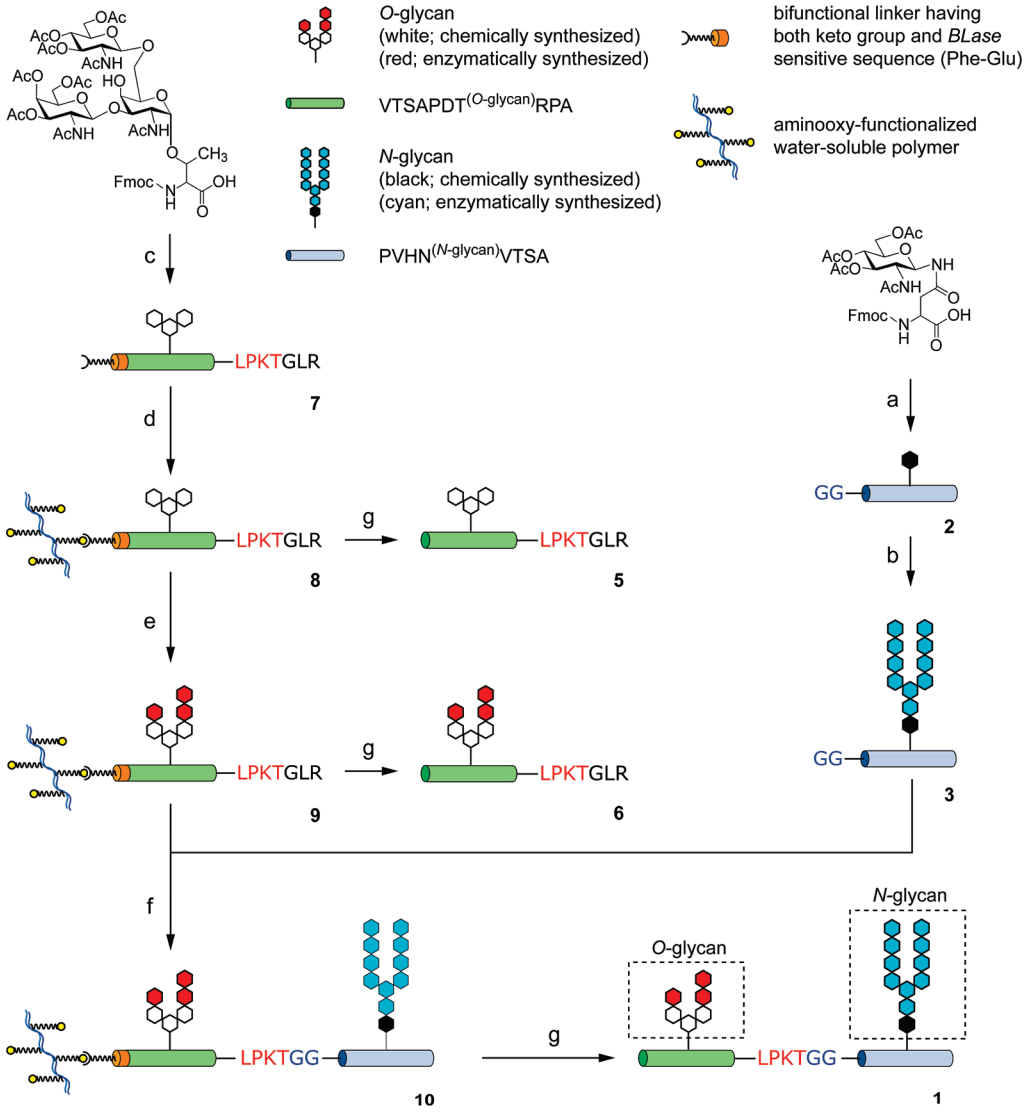
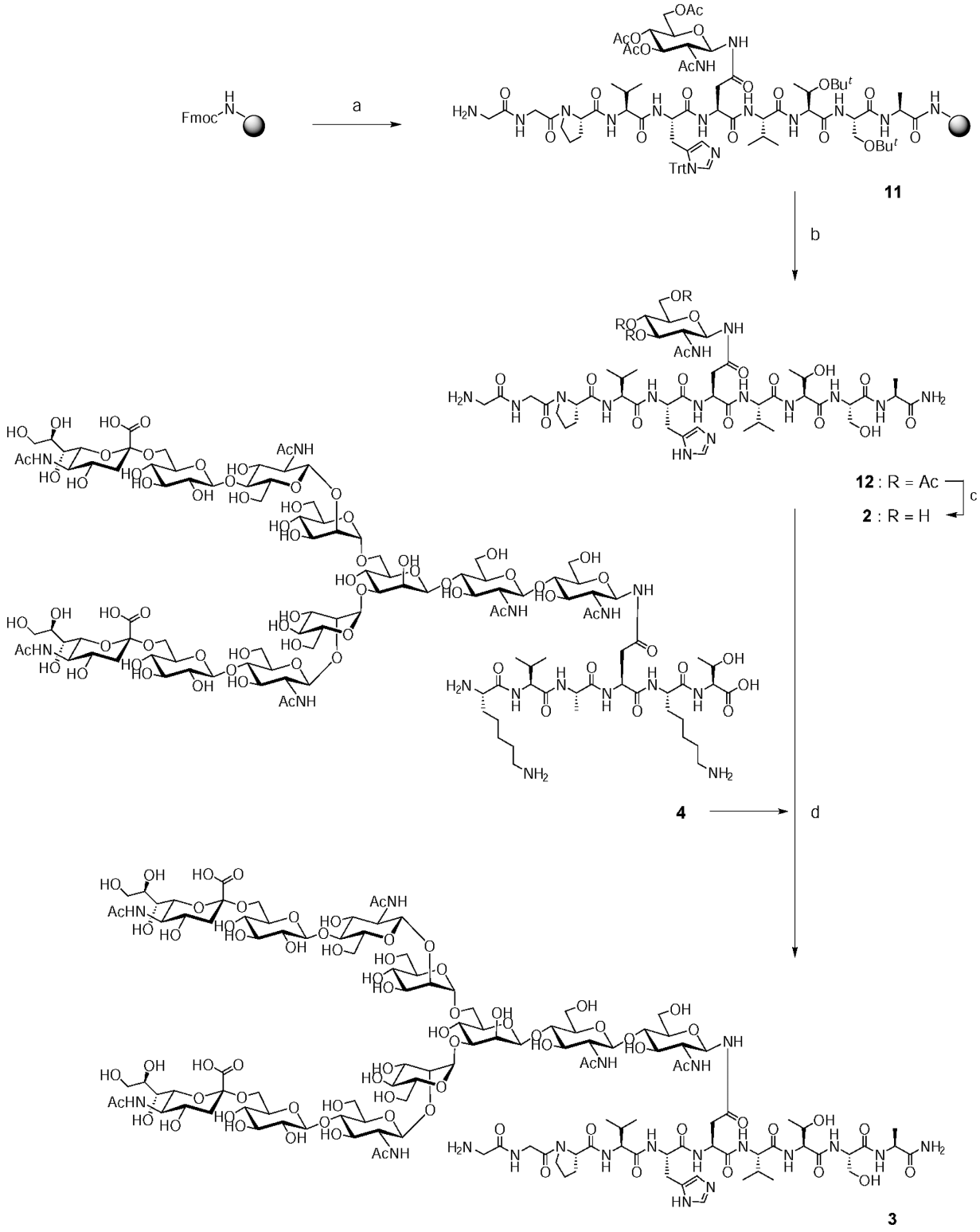


FIGURE 2: A synthetic strategy for the glycopeptide **1** having *O*-linked and *N*-linked type glycans on the basis of polymer-supported chemical and enzymatic protocols: (a, c) microwave-assisted solid-phase synthesis; (b) transglycosylation by endo-M; (d) polymer blotting; (e) sugar elongation by β1,4-GalT, α2,3-*N*-SiaT, and α2,3-*O*-SiaT; (f) segment ligation by SrtA on polymer; (g) cleavage by BLase.

solution of **2** (25.1 mg, 22.0 μmol) in water (440 μL) was added a solution of 150 mM purified SGP **4** (*60*, *61*) (Supporting Information Figure S-3) in water (1.1 mL), and the solution was lyophilized to give a white amorphous solid. The residue was dissolved in a solution of water (1584 μL), 1 unit/mL endo-M (396 μL), and 600 mM potassium phosphate buffer (pH 6.3) (220 μL). After incubation at 37 °C for 1 h, the crude material was purified by a preparative RP-HPLC to give **3** (t_R = 41.0 min, 30.0 mg, 9.5 μmol) in 44% yield (Supporting Information Figure S-4). **2** (t_R = 45.4 min, 13.0 mg, 11.4 μmol) and **4** (t_R = 22.7 min, 161.6 mg, 56.4 μmol) were also recovered in the same HPLC purification procedure. The purified **3** was analyzed by RP-HPLC and ESI-HRMS (Supporting Information Figure S-5 and Figure S-18). ESI-HRMS: $C_{123}H_{204}N_{20}O_{74}$ $[M + 4H]^{4+}$ calcd (*m*/*z*) 1048.0907, found (*m*/*z*) 1048.0903. The 600 MHz ^{1}H NMR spectrum of **3** in H_2O/D_2O (90/10) at 300 K is indicated in Supporting Information Figure S-11. The 600 MHz ^{1}H and 150 MHz ^{13}C NMR chemical shifts of **3** in H_2O/D_2O (90/10) at 300 K are listed in Supporting Information [Table S-1 (peptide moiety) and Table S-2 (carbohydrate moiety)]. Amino acid ratios (numbers in parentheses are theoretical values): Ala(1) 1.0, Asx(1) 1.0, Gly(2) 1.8, His(1) 0.9, Pro(1) 1.0, Ser(1) 0.8, Thr(1) 0.9, Val(2) 1.8.

N-(*5-Oxohexanoyl*)-L-*valyl*-L-*threonyl*-L-*seryl*-L-*alanyl*-L-*prolyl*-L-*aspartyl*-*O*-{*2-acetamido-3,4,6-O-acetyl-2-deoxy-β*-D-*glucopyranosyl*-(*1*→*6*)-[*2,3,4,6-tetra-O-acetyl-β*-D-*galactopyranosyl*-(*1*→*3*)]-*2-acetamido-2-deoxy*-α-D-*galactopyranosyl*}-L-*threonyl*-L-*arginyl*-L-*prolyl*-L-*alanyl*-L-*leucinyl*-L-*prolyl*-L-*lysyl*-L-*threonyl*-L-*glycyl*-L-*leucinyl*-L-*arginyl-amide* (***14***). For the synthesis of *O*-linked glycopeptide **14**, TentaGel S RAM resin (0.26 mmol/g, 192 mg, 50 μmol), Fmoc-amino acids (150 μmol, 3.0 equiv), Fmoc-Thr(Ac$_7$core2)-OH (75 μmol, 1.5 equiv), and 5-oxohexanoic acid (150 μmol, 3.0 equiv) were used. Fmoc-amino acids were described below: Fmoc-Ala-OH, Fmoc-Asp(O*t*Bu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Gly-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Thr(*t*Bu)-OH, and Fmoc-Val-OH. TentaGel S RAM resin was placed in a 5 mL LibraTube and allowed to swell in DMF for a period of 30 min. The resin was filtered, and the Fmoc-removal reaction was conducted under microwave irradiation. Piperidine (20%)/DMF (2 mL) was added, and the mixture was shaken under microwave irradiation for 3 min. Following filtration and washing with DMF (4 mL, five times), the coupling reaction was performed under microwave irradiation.

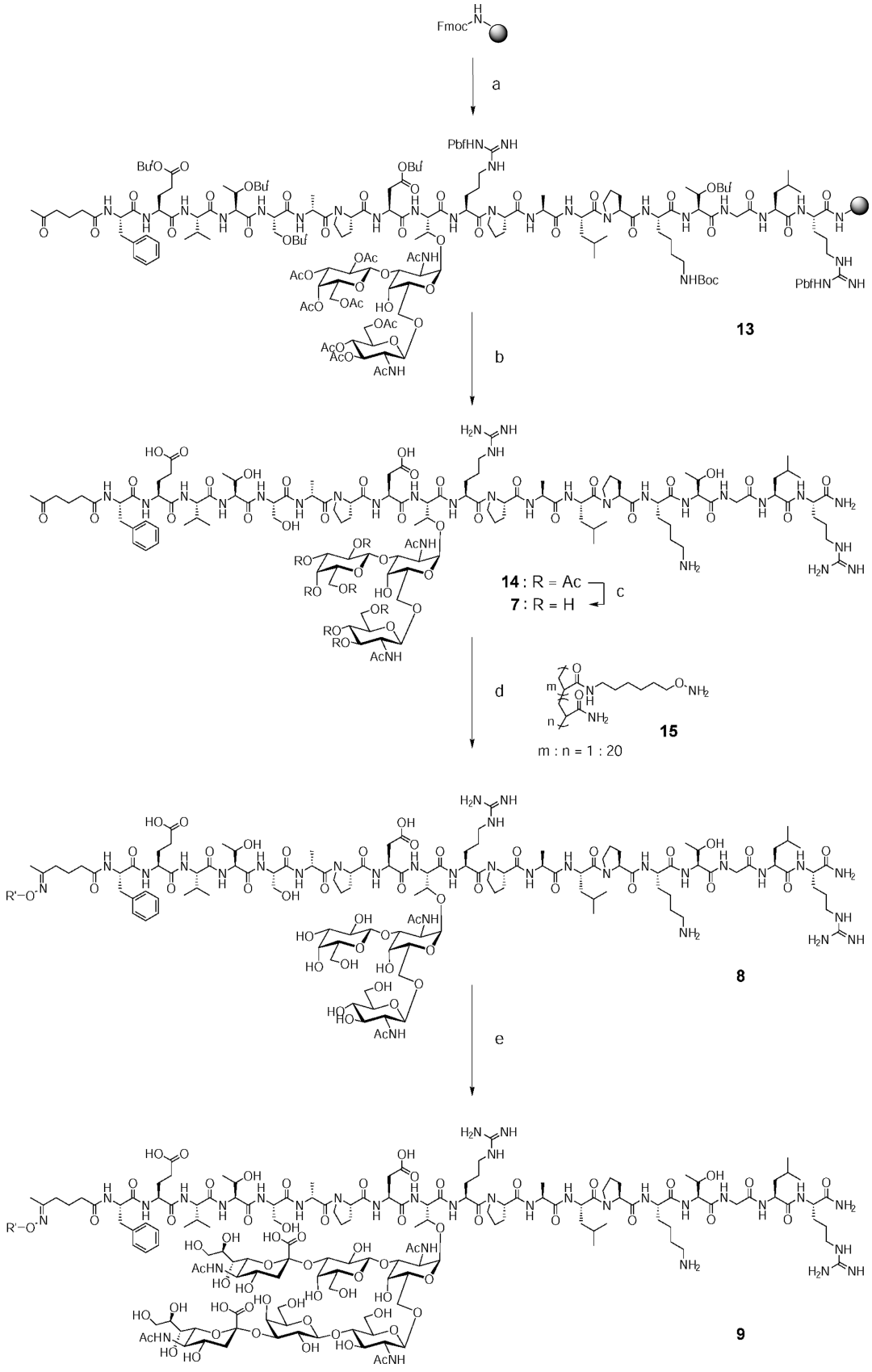
Scheme 1: [a]



[a]Reagents: (a) (i) piperidine, DMF, (ii) Fmoc-amino acid or Fmoc-Asn(Ac_3AcNH-β-Glc)-OH, HBTU, HOBt, DIPEA, NMP, DMF; (iii) Ac_2O, HOBt, DIPEA, DMF, (iv) repeating (i)−(iii); (b) 90% TFA aqueous; (c) NaOH in MeOH; (d) **4** (*60*, *61*), endo-M, potassium phosphate buffer (pH 6.3).

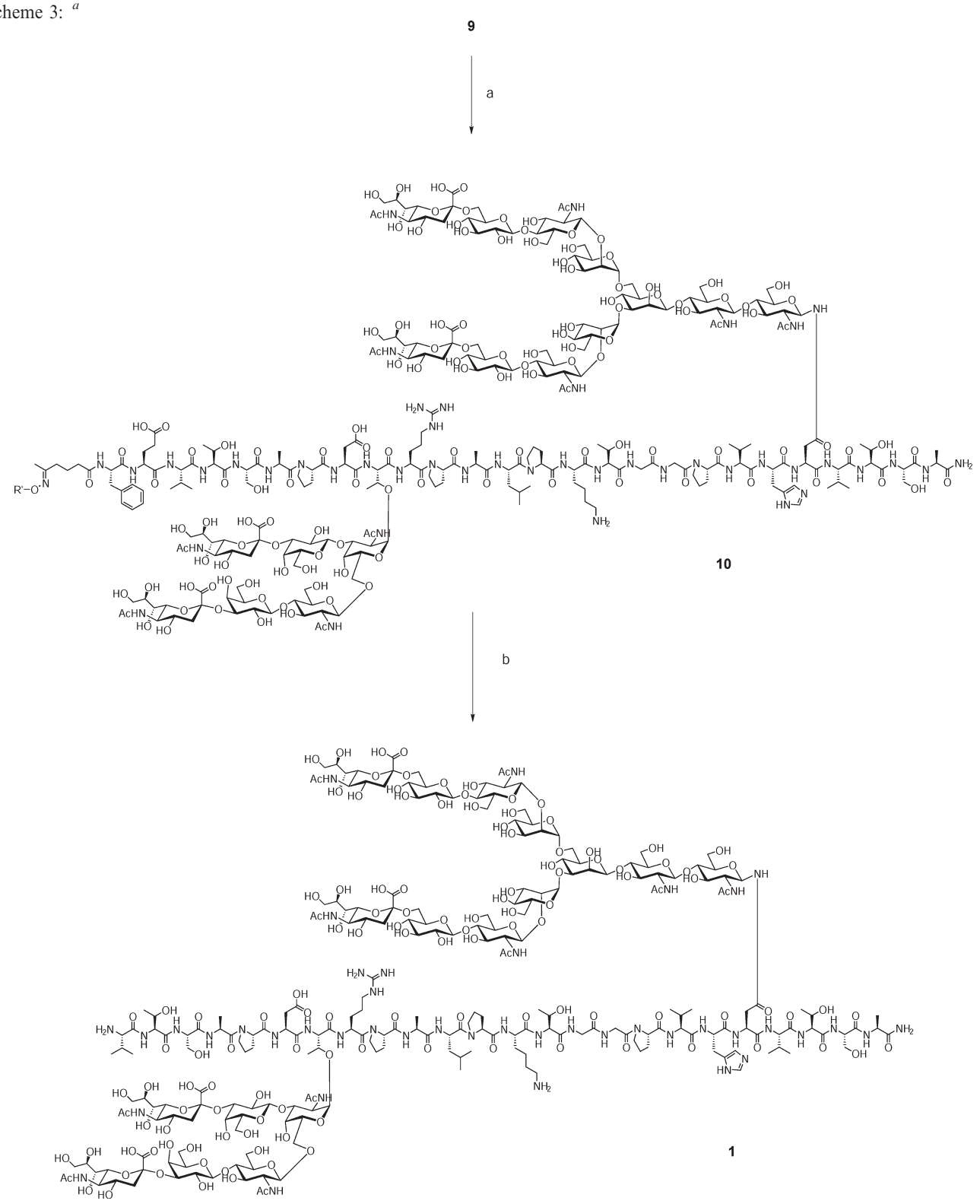
The corresponding Fmoc-amino acid (150 μmol, 3.0 equiv) dissolved in HBTU (150 μmol, 3.0 equiv), HOBt (150 μmol, 3.0 equiv), and DIPEA (300 μmol, 6 equiv) in DMF (1.5 mL) was added to the resin, and the mixture was shaken under microwave irradiation for 10 min. For introduction of Fmoc-glycosylated amino acid, Fmoc-Thr(Ac_7core2)-OH (75 μmol, 1.5 equiv) dissolved in HBTU (75 μmol, 1.5 equiv), HOBt (75 μmol, 1.5 equiv), and DIPEA (150 μmol, 3 equiv) in DMF (1.5 mL) was treated to the resin under microwave irradiation for 20 min. Following filtration and washing with DMF (4 mL, five times), unreacted amino group on the resin were acetylated with a solution of Ac_2O (4.75%, v/v), DIPEA (2.25%, v/v), and HOBt

Scheme 2: [a]



[a]Reagents: (a) (i) piperidine, DMF, (ii) Fmoc-amino acid or Fmoc-Thr(Ac_7core2)-OH or 5-oxohexanoic acid, HBTU, HOBt, DIPEA, NMP, DMF, (iii) Ac_2O, HOBt, DIPEA, DMF, (iv) repeating (i)−(iii); (b) 90% TFA aqueous; (c) NaOH in MeOH; (d) **15** (*40*), sodium acetate buffer (50 mM, pH 5.5); (e) β1,4-GalT, α2,3-*N*-SiaT, α2,3-*O*-SiaT, UDP-Gal, CMP-NANA, $MnCl_2$, BSA, HEPES buffer (50 mM, pH 7.0). R′ = $(CH_2)_6$-polyacrylamide.

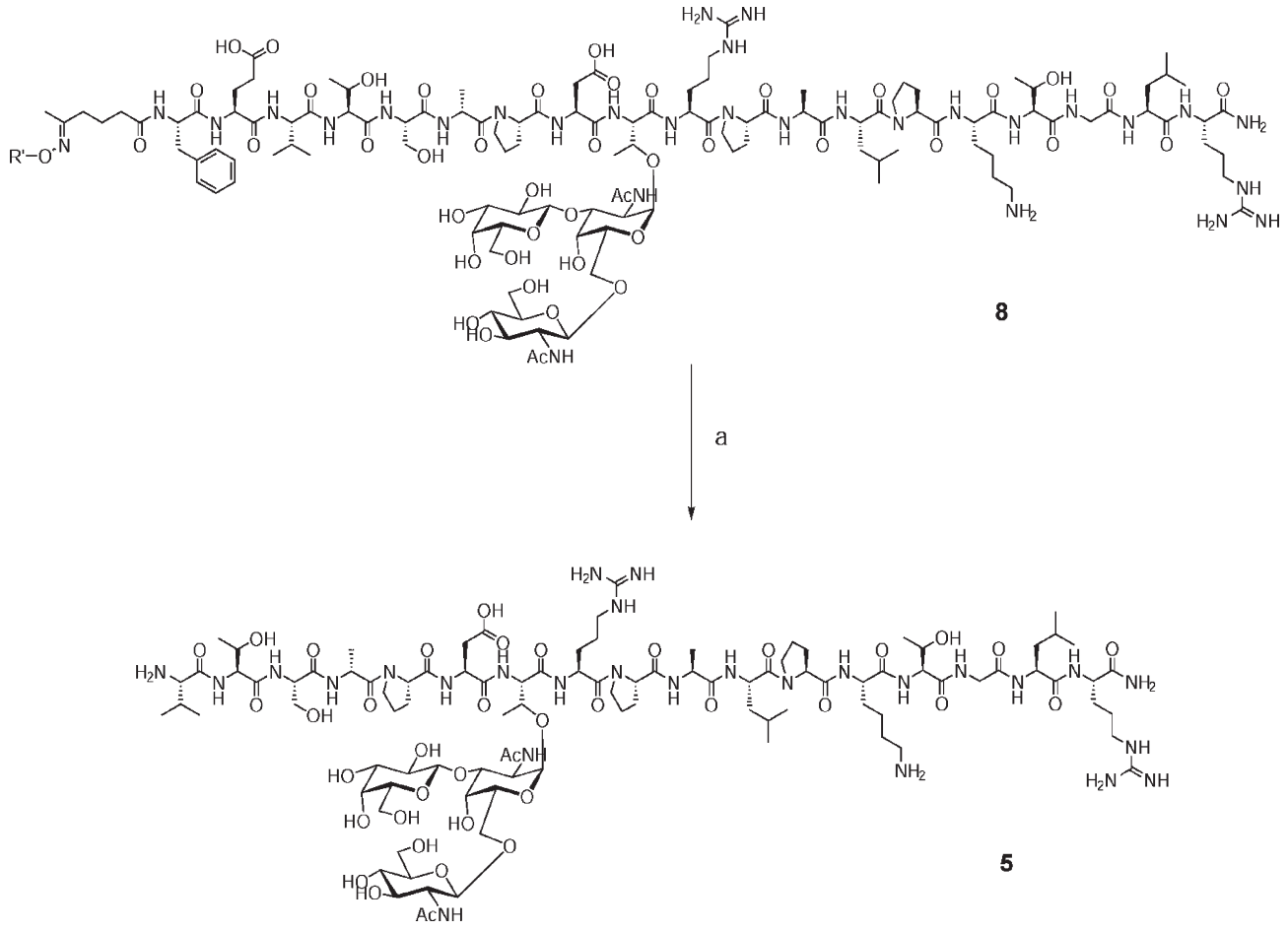
Scheme 3: [a]



[a]Reagents: (a) Sortase A, **3**, β-mercaptoethanol, $CaCl_2$, NaCl, Tris-HCl buffer (50 mM, pH 7.5); (b) BLase, ammonium acetate buffer (25 mM, pH 6.5). R′ = $(CH_2)_6$-polyacrylamide.

(13 mM) in NMP (4 mL). After bering shaken for 5 min at ambient temperature, the resin was filtered and washed with DMF (4 mL, five times). Fmoc-removal, coupling, and capping procedures as described above were carried out repeatedly. At the final step of glycopeptide elongation on the resin, extra Fmoc-Glu(O*t*Bu)-OH, Fmoc-Phe-OH, and 5-oxohexanoic acid were assembled under microwave irradiation to the N-terminus of glycopeptidyl-resin as a component of the bifunctional linker. After completion of the synthesis, the glycopeptidyl-resin **13** was treated with 90% TFA aqueous at room temperature for 2 h, and the resin was filtered. The resin was washed twice with the same cocktail; the filtrates were combined and concentrated by streaming of nitrogen gas. The glycopeptide was precipitated by addition of cold *tert*-butyl methyl ether in an ice bath to give a colorless solid. The solid was washed twice with cold *tert*-butyl methyl ether and dried by streaming of nitrogen gas to give crude

Scheme 4: [a]



[a]Reagents: (a) BLase, ammonium acetate buffer (25 mM, pH 6.5). R′ = $(CH_2)_6$-polyacrylamide.

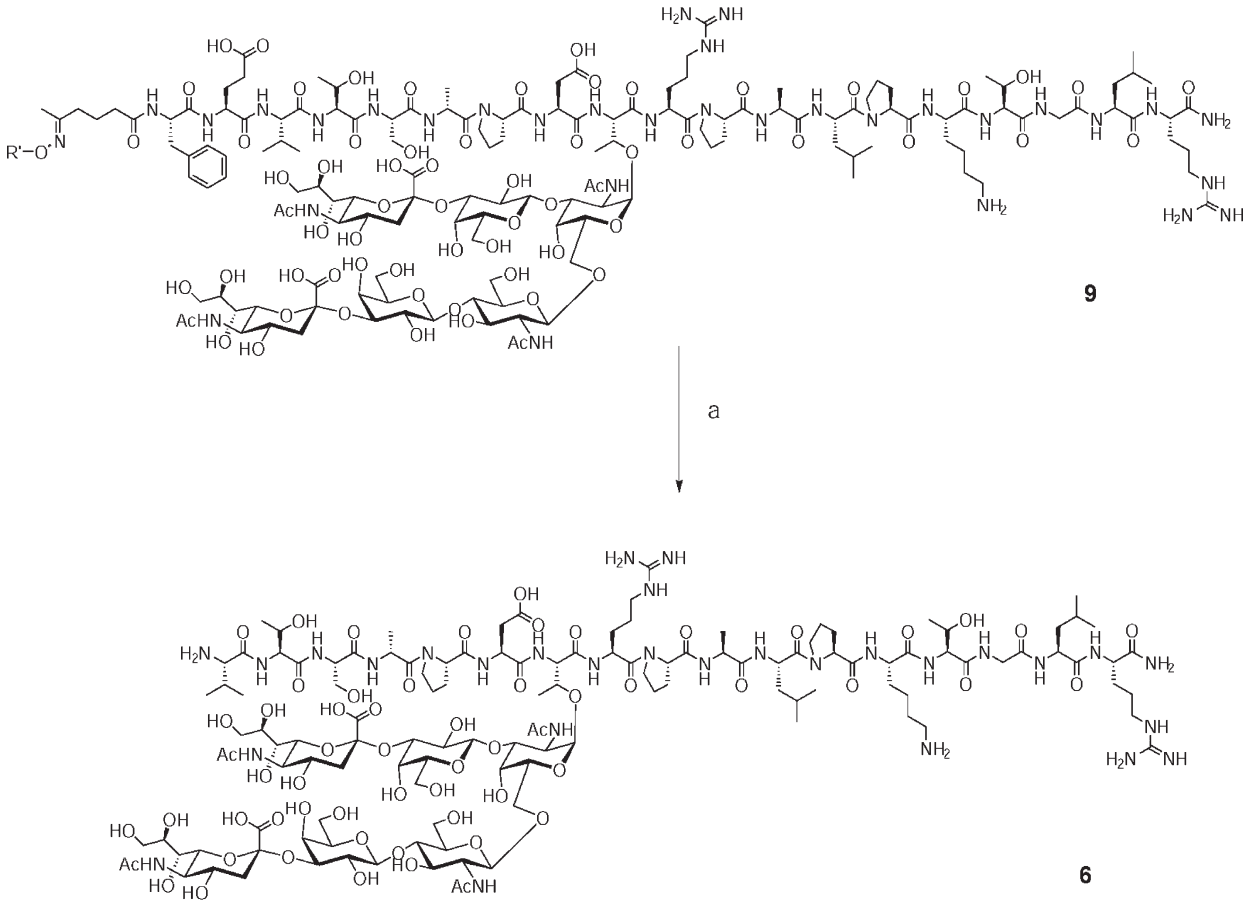
14. MALDI-TOFMS: $C_{113}H_{210}N_{29}O_{51}$ $[M + H]^+$ calcd (*m*/*z*) 3029.5, found (*m*/*z*) 3029.5.

N-(*5-Oxohexanoyl*)-L-*valyl*-L-*threonyl*-L-*seryl*-L-*alanyl*-L-*prolyl*-L-*aspartyl*-*O*-{*2-acetamido-2-deoxy*-β-D-*glucopyranosyl*-(*1*→*6*)-[β-D-*galactopyranosyl*-(*1*→*3*)]-*2-acetamido-2-deoxy*-α-D-*galactopyranosyl*}-L-*threonyl*-L-*arginyl*-L-*prolyl*-L-*alanyl*-L-*leucinyl*-L-*prolyl*-L-*lysyl*-L-*threonyl*-L-*glycyl*-L-*leucinyl*-L-*arginylamide* (*7*). The crude **14** in methanol solution (10 mL) was adjusted and kept at pH 12.4 (as indicated by pH meter) with 1 N sodium hydroxide for 2 h at ambient temperature. The mixture was then neutralized with 1 N acetic acid and evaporated *in vacuo*. The residue was dissolved in 50 mM sodium acetate buffer (pH 5.5) (5 mL) as the stock solution of **7** (10 mM; theoretical concentration based on starting TentaGel resin). An analytical amount of the stock solution (20 μL) was purified by RP-HPLC (t_R = 17.3 min) to give **7**. RP-HPLC purification condition was described below: column, Inertsil ODS-3 (250 × 4.6 mm); eluent A, 25 mM ammonium acetate, pH 5.8; eluent B, acetonitrile containing 10% eluent A. Eluent (A/B = 90/10) was employed; then the ratio of eluent B was increased lineally from 10% to 50% over 30 min with a flow rate of 1.0 mL/min. The purified **7** was analyzed by RP-HPLC and ESI-HRMS (Supporting Information Figure S-6 and Figure S-15). ESI-HRMS: $C_{119}H_{198}N_{29}O_{44}$ $[M + 3H]^{3+}$ calcd (*m*/*z*) 912.4716, found (*m*/*z*) 912.4710.

Chemoselective Polymer Blotting of Compound 7 (*Polymer **8***). The stock solution of **7** (4980 μL, 10 mM; theoretical concentration) was added to the solution of 10 mM polymer **15**(*40*) (60 μmol for aminooxy group) in H_2O (5 mL). After being stirred for 7 h at ambient temperature, completion of the blotting reaction was confirmed by RP-HPLC analyses (Supporting Information Figure S-7). The reaction mixture was passed through a centrifugal ultrafiltration (UF) unit. The retentate was washed with 25 mM HEPES buffer (pH 7.0) three times, and the volume of was adjusted to 2.5 mL with 25 mM HEPES buffer (pH 7.0) as the stock solution of polymer **8** (20 mM; theoretical concentration based on the starting TentaGel resin).

One-Pot Enzymatic Sugar Elongation Catalyzed by Glycosyltransferases (*Polymer **9***). The stock solution of **8** (1.25 mL, 20 mM; theoretical concentration) was added to the mixture of 500 mM HEPES buffer (pH 7.0) (188 μL), 100 mM $MnCl_2$ (500 μL), 2 wt % bovine serum albumin (BSA) (250 μL), 100 mM UDP-galactose (500 μL), 100 mM CMP-NANA (500 μL), 4 units/mL β1,4-GalT (125 μL), 875 milliunits/mL α2,3-*O*-SiaT (200 μL), 3.7 units/mL α2,3-*N*-SiaT (50 μL), and water (1438 μL). After incubation at 25 °C for 24 h, the reaction mixture was passed through a centrifugal UF unit. The retentate was washed with 25 mM HEPES buffer (pH 7.0) three times and 25 mM Tris-HCl buffer (pH 7.5) three times. The volume of the retentate was adjusted to 2.5 mL with 25 mM Tris-HCl buffer (pH 7.5) as the stock solution of polymer **9** (10 mM; theoretical concentration based on the starting TentaGel resin). An analytical amount of the stock solution of **9** (20 μL) was sampled and was treated with 1.74 mg/mL BLase (1 μL) for 1 h. The reaction

Scheme 5: [a]



[a]Reagents: (a) BLase, ammonium acetate buffer (25 mM, pH 6.5). R′ = $(CH_2)_6$-polyacrylamide.

mixture was analyzed by RP-HPLC and MALDI-TOFMS (Supporting Information Figure S-8).

Enzymatic Segment Ligation with SrtA on a Water-Soluble Polymer Support (***10***). *N*-Glycopeptide **3** (18.8 μmol) was added to the mixture of 500 mM Tris-HCl buffer (125 μL), 1.5 M NaCl (250 μL), 100 mM $CaCl_2$ (125 μL), 20 mM 2-mercaptoethanol (250 μL), the stock solution of polymer **9** (1.25 mL, 10 mM; theoretical concentration), 1.0 mM sortase A (260 μL), and water (240 μL). After incubation at 37 °C for 1 h, an analytical amount of the solution (20 μL) was sampled and was treated with 1.74 mg/mL BLase (1 μL) for 1 h. The reaction mixture was analyzed by RP-HPLC and MALDI-TOFMS (Supporting Information Figure S-9). After incubation at 37 °C for 2 h, the mixture was washed with 25 mM Tris-HCl buffer (pH 7.5) three times and 25 mM ammonium acetate buffer (pH 6.5) three times by using a centrifugal UF unit. The volume of the retentate was adjusted to 1.56 mL with 25 mM ammonium acetate buffer (pH 6.5) as the stock solution of **10** (8 mM; theoretical concentration based on the starting TentaGel resin).

L-Valyl-L-threonyl-L-seryl-L-alanyl-L-prolyl-L-aspartyl-O-{(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→6)-[(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→3)-β-D-galactopyranosyl-(1→3)]-2-acetamido-2-deoxy-α-D-galactopyranosyl}-L-threonyl-L-arginyl-L-prolyl-L-alanyl-L-leucinyl-L-prolyl-L-lysyl-L-threonyl-L-glycyl-L-glycyl-L-prolyl-L-valyl-L-histidyl-{N^4-(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→6)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl-mac_opb;1→6)-[(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→6)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl-(1→3)]-β-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl}-L-asparaginyl-L-valyl-L-threonyl-L-seryl-L-alanylamide (***1***). The stock solution of polymer **10** (1.56 mL, 8 mM; theoretical concentration) was passed through a centrifugal UF unit, and the retentate was washed three times with 25 mM ammonium acetate buffer (pH 6.5). The volume of the retentate was adjusted to 1.56 mL with the same washing buffer, and then 1.74 mg/mL BLase solution (8 μL) was added. After incubation of the solution at 25 °C for 2 h, the reaction mixture was passed through a centrifugal UF unit. The retentate was washed twice with 25 mM ammonium acetate buffer (pH 6.5), and the filtrate was lyophilized. RP-HPLC purification was performed to give pure compound **1** (3.3 mg, 0.55 μmol, from amino acid analysis) in 4.4% overall yield calculated from the first step in the solid-phase synthesis started from Fmoc (sugar) amino acids. The purified glycopeptide **1** was analyzed by RP-HPLC (Supporting Information Figure S-10) and ESI-HRMS (Supporting Information Figure S-18). ESI-HRMS: $C_{236}H_{383}N_{42}O_{130}$ $[M - 3H]^{3-}$ calcd (*m*/*z*) 1961.8217, found (*m*/*z*) 1961.8214 (Supporting Information Figure S-15). ^{1}H NMR spectrum of **1** in H_2O/D_2O (90/10) at 300 K is indicated in Supporting Information (Figure S-17). The 600 MHz ^{1}H and 150

MHz ^{13}C NMR chemical shifts of **1** in H_2O/D_2O (90/10) at 300 K are listed in Supporting Information [Table S-5 (peptide moiety), Table S-6 (*O*-glycan moiety), and Table S-7 (*N*-glycan moiety)].

L-Valyl-L-threonyl-L-seryl-L-alanyl-L-prolyl-L-aspartyl-O-{2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→6)-[β-D-galactopyranosyl-(1→3)]-2-acetamido-2-deoxy-α-D-galactopyranosyl}-L-threonyl-L-arginyl-L-prolyl-L-alanyl-L-leucinyl-L-prolyl-L-lysyl-L-threonyl-L-glycyl-L-leucinyl-L-arginylamide (***5***). The stock solution of polymer **8** (1.25 mL, 20 mM; theoretical concentration) was passed through a centrifugal UF unit and washed three times with 25 mM ammonium acetate buffer (pH 6.5). The volume of the retentate was adjusted to 1.25 mL with the same washing buffer, and then 1.74 mg/mL BLase solution (16 μL) was added. After incubation of the solution at 25 °C for 2 h, the reaction mixture was passed through a centrifugal UF unit. The retentate was washed three times with 25 mM ammonium acetate buffer (pH 6.5), and the filtrate was lyophilized. The crude material was purified by a preparative RP-HPLC (t_R = 36.0 min) to give **5** (3.92 μmol) in 15.7% overall yield from the first step in the solid-phase *O*-glycopeptide synthesis (calculated by amino acid analysis). RP-HPLC purification condition was described below: column, Inertsil ODS-3 (250 × 20 mm); eluent A, 25 mM ammonium acetate, pH 5.8; eluent B, acetonitrile containing 10% eluent A. Eluent (A/B = 95/5) was employed from 0 to 5 min; then the ratio of eluent B was increased lineally from 5% to 30% over 60 min with a flow rate of 5.0 mL/min. The purified **5** was analyzed by RP-HPLC and ESI-HRMS (Supporting Information Figure S-11 and Figure S-18). ESI-HRMS: $C_{99}H_{172}N_{27}O_{38}$ $[M + H]^+$ calcd (*m*/*z*) 2347.2351, found (*m*/*z*) 2347.2380. The 600 MHz ^{1}H NMR spectrum of **5** in H_2O/D_2O (90/10) at 300 K is indicated in Supporting Information (Figure S-15). The 600 MHz ^{1}H and 150 MHz ^{13}C NMR chemical shifts of **5** in H_2O/D_2O (90/10) at 300 K are listed in Supporting Information [Table S-3 (peptide moiety) and Table S-4 (carbohydrate moiety)].

L-Valyl-L-threonyl-L-seryl-L-alanyl-L-prolyl-L-aspartyl-O-{(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→6)-[(5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosylonic acid)-(2→3)-β-D-galactopyranosyl-(1→3)]-2-acetamido-2-deoxy-α-D-galactopyranosyl}-L-threonyl-L-arginyl-L-prolyl-L-alanyl-L-leucinyl-L-prolyl-L-lysyl-L-threonyl-L-glycyl-L-leucinyl-L-arginylamide (***6***). The stock solution of polymer **9** (1.25 mL, 10 mM; theoretical concentration) was passed through a centrifugal UF unit and washed three times with 25 mM ammonium acetate buffer (pH 6.5). The volume of the retentate was adjusted to 1.25 mL with the same washing buffer, and then 1.74 mg/mL BLase solution (8 μL) was added. After incubation of the solution at 25 °C for 2 h, the reaction mixture was passed through a centrifugal UF unit. The retentate was washed with 25 mM ammonium acetate buffer (pH 6.5) three times, and the filtrate was lyophilized. The crude material was purified by a preparative RP-HPLC (t_R = 40.9 min) to give **6** (1.24 μmol) in 9.9% overall yield from the first step in the solid-phase *O*-glycopeptide synthesis (calculated by amino acid analysis). RP-HPLC purification condition was described below: column, Inertsil ODS-3 (250 × 20 mm); eluent A, 25 mM ammonium acetate, pH 5.8; eluent B, acetonitrile containing 10% eluent A. Eluent (A/B = 95/5) was employed from 0 to 5 min; then the ratio of eluent B was increased lineally from 5% to 20% over 60 min with a flow rate of 7.0 mL/min. The purified **6** was analyzed by RP-HPLC and ESI-HRMS (Supporting Information Figure S-12 and Figure S-18). ESI-HRMS: $C_{99}H_{172}N_{27}O_{38}$ $[M + H]^+$ calcd (*m*/*z*) 3091.4788, found (*m*/*z*) 3091.4778. The 600 MHz ^{1}H NMR spectrum of **6** in H_2O/D_2O (90/10) at 300 K is indicated in Supporting Information (Figure S-16). The 600 MHz ^{1}H and 150 MHz ^{13}C NMR chemical shifts of **6** in H_2O/D_2O (90/10) at 300 K are listed in Supporting Information [Table S-3 (peptide moiety) and Table S-4 (carbohydrate moiety)].

Competition ELISA. The microtiter plate (Pierce, reactibind neutravidin-coated high binding capacity (HBC) clear 96-well plate with superblock blocking buffer) was coated with 100 μL/well of synthetic *N*-biotinylated glycopeptide Pro-Asp-Thr(Galβ1→3GalNAcα)-Arg-Pro-Ala (MALDI-TOFMS: *m*/*z* calcd for $C_{61}H_{101}N_{15}O_{25}S$ $[M + H]^+$ 1475.7, found 1477.220), and 20-mer glycopeptide used as a positive control (MALDI-TOFMS: *m*/*z* calcd for $C_{102}H_{165}N_{28}O_{42}$ $[M + H]^+$ 2454.2, found 2454.5) was synthesized by the procedure reported in our previous paper (*43*). The wells were washed three times and blocked with PBS containing 0.1% BSA at room temperature for 1 h. Anti-MUC1 mAb (clone B421 from GenTex Inc.) and the inhibitor glycopeptides were added with 100 μL/well. Inhibitors were used at concentrations ranging from 0.2 to 500 μM. After incubation at room temperature for overnight, the wells were washed three times, and the second antibody, peroxidase-labeled affinity-purified antibody to mouse IgG (H+L) (Kirkegaard Perry Laboratries (KPL), Guildford, U.K.), was added at a dilution of 1/1000 at 100 μL/well. After incubation for 1 h and washing three times, TMB substrate solution (KPL kit) was added (100 μL/well). The reaction was terminated with addition of 100 μL/well of stop solution 1 M H_3PO_4. The rates of reaction in each well were measured for absorbance at 450 nm.

NMR Spectroscopy. The glycopeptides **1**, **2**, **3**, **5**, and **6** were dissolved in 300 μL of either 10% D_2O in H_2O or 99.9% D_2O at the concentration of 1.8−2.0 mM. The pH was adjusted to 5.0 (as indicated by pH meter) by addition of HCl and NaOH. The Shigemi NMR sample tubes matched with D_2O (BMS-005B, i.d. 4.2 mm; Shigemi Co., Ltd.) were used for NMR experiments. NMR spectra were collected at 300 K with a Bruker Avance 600 spectrometer at 600.03 MHz for proton frequency equipped with cryoprobe. For the complete assignments and the structure determination of glycopeptides, two-dimensional homonuclear DQF-COSY(*62*), TOCSY with MLEV-17 sequence(*63*, *64*), and NOESY(*65*) spectra were recorded in the indirect dimension using States−TPPI phase cycling. Additionally, two-dimensional heteronuclear ^{13}C-edited HSQC and HSQC-TOCSY measurements were also recorded with echo−antiecho mode for sensitivity enhancement. TOCSY experiments were applied for a spin-locking time of 60 ms, and NOESY experiments were carried out with mixing times of 100, 150, 200, and 400 ms. The suppression of the water signal was performed by presaturation during the relaxation delay (1 s) and by a 3-9-19 WATERGATE pulse sequence with field gradient (*66*). TOCSY and NOESY spectra were acquired with 2048 by 512 frequency data points and were zero-filled to yield 2048 by 2048 data matrices. DQF-COSY with 16384 by 512 frequency data points was also recorded and zero-filled to yield a 16384 by 16384 matrix in order to measure the coupling constants. The sweep widths of 8389.26 Hz were applied. Time domain data in both dimensions were multiplied by a sine bell window function with a 90° phase shift prior to Fourier transformation. All NMR data were processed by NMRPipe software (*67*) and analyzed using the Sparky program (*68*). Sequence-specific resonance assignments were

achieved according to the standard methods for small proteins established by Wüthrich and co-workers (*69*). Stereospecific assignments for methylene protons were carried out by analyzing the intensities of intraresidue NOE between amide and β protons.

Structure Calculations. Three-dimensional structures of glycopeptides **1**, **5**, and **6** were calculated using the CNS 1.1(*70*) program with standard protocols for distance geometry-simulated annealing and refinement. Distance restraints for calculations were estimated from the cross-peak intensities in NOESY spectra with a mixing time of 150 ms for **1** and 200 ms for **5** and **6**. The estimated restraints were classified into four categories: strong (1.6−2.6 Å), medium (1.6−3.5 Å), weak (1.6−5.0 Å), and very weak (1.6−6.0 Å). In the first stage of structure determination, the structures of glycopeptides were calculated using only interproton distance information. After the validation of fulfilling distance restraints for the obtained structure, the restraints of the dihedral angle ϕ and χ_1 were adopted for structure calculation. When the coupling constant $^3J_{HN\alpha}$ was more than 8.0 Hz and less than 6.0 Hz, the dihedral angle ϕ was constricted to $-120 \pm 30°$ and $-60 \pm 30°$, respectively. The conformation of the sugar rings was held fixed to the chair conformation. All analyses of rmsd values and the solution structures of glycopeptides were performed with PROCHECK (*71*) and MOLMOL(*72*) programs.

*Structural Characterization by ESI-(ECD/CID)-TOFMS of Glycopeptide **1** Carrying N- and O-Glycans*. Experiments on the structural characterization of glycopeptide **1** 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 (*73*, *74*). Ions in the TOF were additionally accelerated by 4 kV before the microchannel plate (MCP) detection. Output signals from MCP were processed using a high-speed digitizer (AP240; Agilent Technologies). Obtained ECD spectra were further analyzed by an in-house software tool covering peak picking, peak assignment, and sequence validation. An uncoated SilicaTips (tip i.d. 15 μm; New Objective, Woburn, MA) was used as an electrospray tip. Samples dissolved in water (ca. 10 μM) were diluted by the addition of an equal volume of acetonitrile containing 0.2% formic acid and infused into the microESI source at a flow rate of 1 μL/min with a syringe pump (Harvard Apparatus, Holliston, MA). Conditions for ESI-LITTOFMS were as follows: capillary voltage, 1.5 kV; flow rate of curtain gas, 0.8−1.2 L/min; scan mass range, m/z 200−2000. ECD parameters for MS^2 spectral acquisitions were as follows: isolation time, 5 ms; isolation widths, 20; electron energy, less than 2 eV nominally and current 0.7 μA; electron irradiation time, 40 ms. CID parameters for MS^2 spectral acquisitions were as follows: CID time, 10 ms; CID gain, 2.06.

RESULTS AND DISCUSSION

We designed MUC1-related glycopeptide **1** (MW = 5981.8) as a plausible glycoprotein model which has *O*- and *N*-glycans concurrently (Figure 1). Compound **1** is a hybrid neoglycopeptide composed of two characteristic segments, a cancer-relevant *O*-glycosylation site in the tandem repeating sequence in MUC1 glycoprotein [Val-Thr-Ser-Ala-Pro-Asp-*Thr*(*O*-glycan)-Arg-Pro-Ala] (*75*−*78*) and a partial structure involving the *N*-glycosylation site in the C-terminal region [^{953}Pro-Val-His-*Asn*(*N*-glycan)-Val-Thr-Ser-^{961}Ala] of membrane glycoprotein MUC1 (mucin 1, polymorphic epitherial mucin from OGlycBase 6.00) (*79*−*83*). We thought that glycopeptide **1** is a convenient model to discuss the effect of the neighboring *N*-glycan moiety on the conformation and immunological property of the distinct peptide epitope PDTR motif locating in the C-terminus of the tandem repeats. In fact, Parry et al. reported that *N*-glycan structures of this region differ markedly between membrane-bound and secreted forms of native MUC1 molecules (*84*).

Figure 2 illustrates a synthetic strategy of the target glycopeptide **1** bearing *O*-glycan (core 2-based sialyl hexasaccharide) and biantennary sialyl *N*-glycan (complex-type undecasaccharide) chains. Glycopeptide **1** was divided into an amino component **3** with Gly-Gly and a carboxyl component **9** with Leu-Pro-Lys-Thr-Gly-Leu-Arg, and further synthetic route is summarized as follows: (i) microwave-assisted solid-phase (glyco)peptide synthesis (*42*, *56*) of key intermediates **2** and **7** from glycosylated amino acid building blocks (*85*−*87*), (ii) chemoselective blotting of **7** by an aminooxy-functionalized polymer, (iii) sugar elongation of **2** and **8** by endoglycosidase (*57*−*59*) and glycosyltransferases (*32*−*44*), respectively, (vi) SrtA-mediated ligation of segment **3** with segment **9** on a polymer support, and (v) release of the target glycopeptide **1** by treating with BLase (*41*).

Compound **2** was synthesized by the Fmoc-based SPPS strategy employing HBTU/HOBt/DIEA coupling (*88*, *89*) on a TentaGel resin functionalized with Rinkamide linker (*90*). Incorporation of a glycosylated Fmoc-asparagine was performed by means of an off-line microwave irradiation coupling protocol (*42*, *56*), and other procedures were conducted on an automated peptide synthesizer. Cleavage from the resin with 90% TFA aqueous solution afforded the protected glycopeptide **12** (62%) and followed by hydrolysis of the acetyl protecting groups under alkaline condition gave an intermediate **2** (93%). Subsequently, compound **2** was used for the acceptor substrate in the *trans*-glycosylation by endo-M (*57*−*59*) with a known donor substrate **4** (*60*, *61*), a disialylated biantennary *N*-glycopeptide, to afford segment **3** in 60% isolated yield.

On the other hand, segment **9** bearing the LPXTG recognition motif (X = K in this study) at the C-terminus was synthesized by combined use of microwave-assisted SPPS and polymer-supported enzymatic sugar elongation (*40*, *41*). At the final step of the synthesis of compound **7**, extra Fmoc-Glu(O*t*-Bu)-OH, Fmoc-Phe-OH, and 5-oxohexanoic acid were sequentially attached to the N-terminus in order to function as a molecular shuttle carrying a heterobifunctional linker with an aminooxy-functionalized polymer, and the Phe-Glu moiety becomes a specific cleavage site by BLase to release the final product from the polymer support. One-pot synthesis of segment **9** from intermediate **8** bearing core 2 trisaccharide was carried out according to the condition described in our previous papers (*40*−*44*) using β1,4-GalT, α2,3-*N*-SiaT, and α2,3-*O*-SiaT in the presence of sugar nucleotides UDP-Gal and CMP-Neu5Ac as glycosyl donor substrates. Conversion of **8** into **9** proceeded quantitatively as determined by HPLC analysis of compound **6** released by BLase treatment (Supporting Information Figure S-8), and the resulting polymer **9** was readily obtained by using only simple centrifugal ultrafiltration (UF) without any tedious chromatographic purification.

The mixture of **3** and polymer **9** (molar ratio = 10:1) was incubated with SrtA at 37 °C for 4 h and afforded polymer **10**. Judging from the HPLC profile of the products released by treating polymer **10** with BLase, SrtA-mediated ligation proceeded smoothly and gave the target glycopeptide **1** in 77% yield

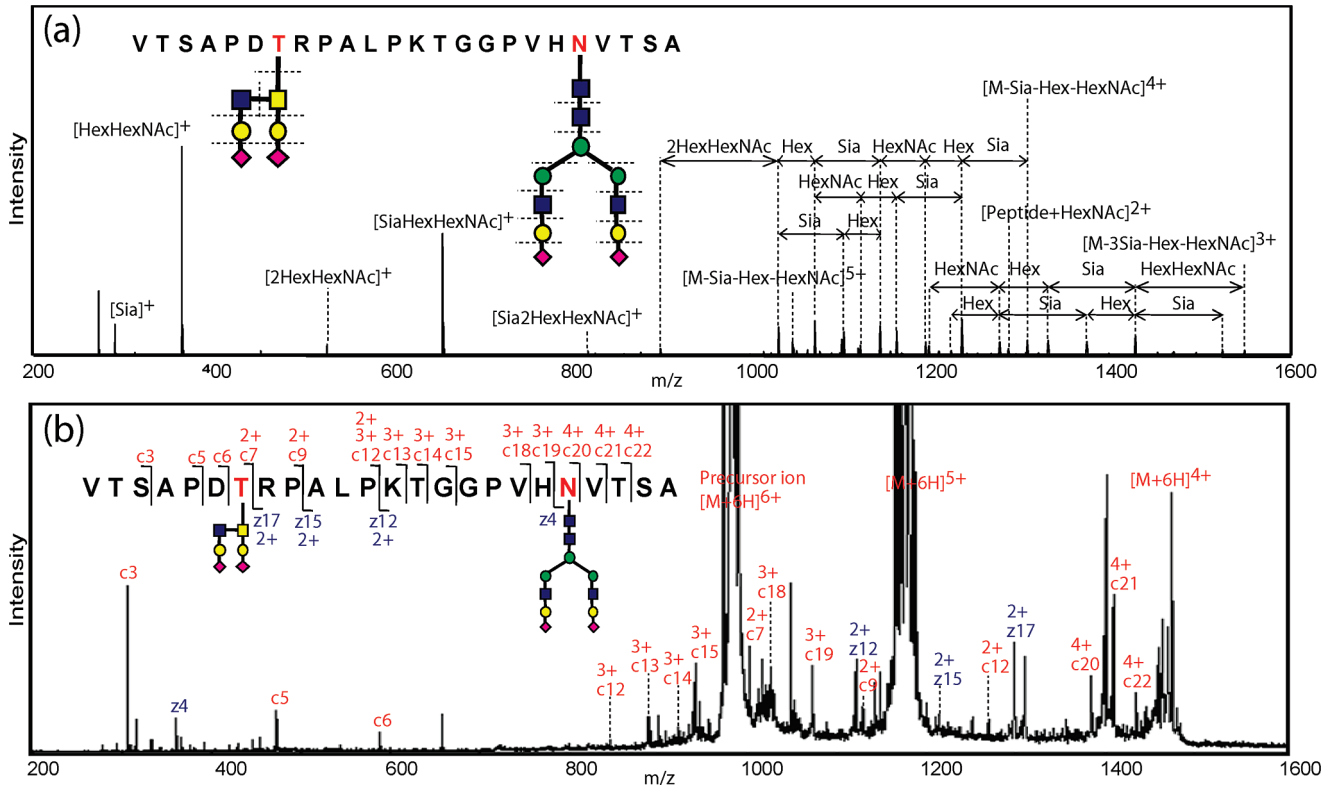


FIGURE 3: MS/MS spectra of the charge state 6+ (m/z 982.43) of glycopeptide **1** carrying *N*- and *O*-glycans. (a) CID and (b) ECD spectra.

(calculated from compound **9**; see also Supporting Information). When the *trans*-glycosylation by endo-M was conducted after the SrtA-mediated ligation between polymer **9** and compound **2**, no significant *trans*-glycosylation at the GlcNAc residue was observed (data not shown), suggesting that the active site of the endo-M appears to be strongly influenced by the steric hindrance due to the adjacent bulky *O*-glycan chain of the acceptor glycopeptide. Purification by centrifugal UF and reversed-phase HPLC afforded the desired glycopeptide **1** in 4.4% isolated overall yield (from the initial step in the solid-phase peptide synthesis of compound **7**). Structural analysis by ESI-TOFMS (MS/MS) revealed that the ECD device (*73*, *74*) gives highly informative *c* and *z* ions allowing for the identification of both *O*- and *N*-glycosylation sites of glycopeptide **1** without any significant degradation in the glycan chains, while typical fragmentation at every glycoside linkage in the carbohydrate moieties occurred predominantly under common CID conditions (Figure 3). Compounds **5** (16% from **7**) and **6** (10% from **7**) were also derived via intermediates **8** and **9**, respectively. Further characterization of these compounds was carried out by MALDI-TOFMS, ESI-HRMS, ^{1}H and ^{13}C NMR, RP-HPLC, and amino acid analysis (Supporting Information).

Next, we tested the immunological property of compounds **1**−**6** by employing competitive ELISA assay using anti-MUC1 monoclonal antibody established by human breast cancer cell line ZR-75-1 (anti-MUC1 mAb, clone B421; GenTex, Inc.) that reacts with the microtiter plate displaying antigenic MUC1 glycopeptide, the *N*-biotinylated Pro-Asp-Thr-(Galβ1→3GalNAcα1)-Arg-Pro-Ala (*43*, *91*). As indicated in Figure 4, compounds **1**, **5**, and **6** exhibited a strong inhibitory effect on the interaction of anti-MUC1 mAb with this microtiter plate, while the other *N*-glycopeptides **2**, **3**, and **4** did not show any inhibition. These results clearly indicate that the anti-MUC1 mAb (clone B421) reacts with glycopeptides involving core 2-based

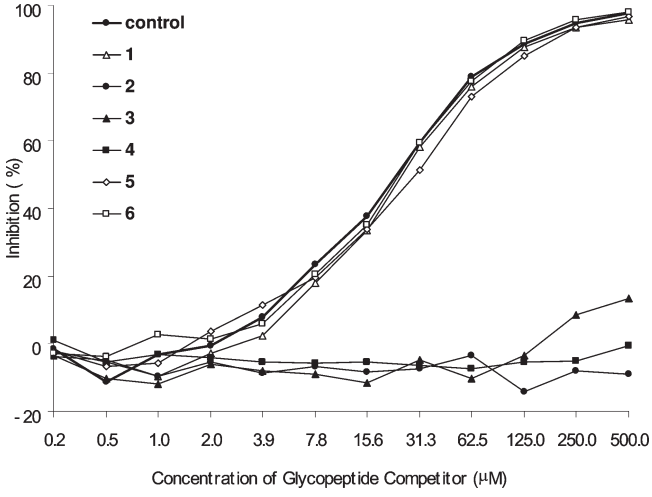


FIGURE 4: Competitive ELISA assay by means of microtiter plate precoated by *N*-biotinylated-Pro-Asp-Thr(Galβ1→3GalNAcα1)-Arg-Pro-Ala and anti-MUC1 mAb (clone B421) in the presence of synthetic glycopeptides **1**−**6**. The 20-mer synthetic MUC1 glycopeptide, His-Gly-Val-Thr-Ser-Ala-Pro-Asp-Thr[Galβ1→3(GlcNAcβ1→6)GalNAcα1]-Arg-Pro-Ala-Pro-Gly-Ser-Thr-Ala-Pro-Pro-Ala (*43*), was employed as a tentative positive control recognized by the above anti-MUC1 mAb.

trisaccharide (**5**) and disialylated hexasaccharide (**1**) in the tandem repeating PDTR sequence. In addition, it is clear that the disialylated biantennary *N*-glycan chain of the C-terminal region in glycopeptide **1** does not disturb the interaction between anti-MUC1 mAb with the antigenic *O*-glycopeptide region. On the contrary, it has been well documented that extra (multiple) *O*-glycosylation(s) near the immunodominant *O*-glycosylation site greatly influence(s) the antigenicity of the cancer epitopes of the MUC1 glycoproteins (*43*, *91*). It is considered that *O*-glycan(s) substituted at other potential *O*-glycosylation sites

would cover the crucial epitope structure or provide some conformational impact due to the additional GalNAcα1→Thr(Ser) linkage(s), denoted a high level of molecular organization mechanism exists within the mucin glycoproteins (*19*−*21*). Considering the similarity in terms of molecular size and structure of *O*-glycan (core 2-based disialylated hexasaccharide) and *N*-glycan (disialylated biantennary undecasaccharide) of MUC1 glycopeptide **1**, we assumed that the C-terminal region [Pro-Val-His-Asn(*N*-glycan)-Val-Thr-Ser-Ala] seems to be much more flexible and random structure in solution than the N-terminal immunodominant *O*-glycosylated region. As a result, it is likely that anti-MUC1 mAb could access to this epitope sequence without any steric hindrance due to the bulky C-terminal *N*-glycopeptide. This seems to be the reason why attachment of the C-terminal *N*-glycopeptide moiety did not interrupt the binding of anti-MUC1 mAb with the neighboring *O*-glycopeptide epitope.

To assess the above-mentioned assumption, our attention was directed toward the solution structures of compound **1** and two related glycopeptides, **5** and **6**, recognized by the anti-MUC1 mAb. We determined the solution structures by means of distance geometry-stimulated annealing calculations with restraints from various NMR experiments such as a series of proton two-dimensional NMR spectra, DQF-COSY, TOCSY, NOESY, ^{13}C-edited HSQC, and HSQC-TOCSY (all chemical shifts and summary of structure statistics are given in the Supporting Information). Figure 5a represents the lowest energy structure of this 24 amino acid glycopeptide **1**, and Figure 5b and Figure 5c indicate the models constructed by superimpositions with the 30 lowest energy structures of the partial sequence [Thr-Ser-Ala-Pro-Asp-Thr(*O*-glycan)-Arg] in compounds **1** (left), **6** (middle), and **5** (right).

As anticipated, NMR study of these glycopeptides revealed that the pentapeptide [Ala-Pro-Asp-Thr(*O*-glycan)-Arg] within the N-terminal immunodominant region forms an inverse γ-turn-like structure (*92*−*94*), while the C-terminal region composed of *N*-glycopeptide and linker peptide was proved to be random structure. Interestingly, the well-converged peptide backbone represented as a superimposition of 30 lowest energy structures of compound **5** (N-terminal segment carrying core 2 trisaccharide derived from intermediate **8**) appears to be a more extended conformation than those of **1** and **6** (Figure 5c). Moreover, the inverse γ-turn-like structure detected in the case of **6** did not change by converting into **1** through a series of sugar extension reactions with glycosyltransferases, suggesting that there might be an entropy merit allowing for a protein folding-like mechanism of the conformational stabilization of the artificial glycoproteins during SrtA-mediated ligation between the two macromolecular glycopeptide segments. In summary, NMR study revealed that (i) the C-terminal segment of glycopeptide **1** is random structure, (ii) the N-terminal immunodominant region [Ala-Pro-Asp-Thr(*O*-glycan)-Arg] of three glycopeptides, **1**, **5**, and **6**, exhibits a well-converged peptide backbone structure, and (iii) SrtA-mediated conjugation with the C-terminal *N*-glycopeptide moiety gives a detectable conformational influence to the inverse γ-turn-like structure observed in the immunodominant pentapeptide region, though this conformational alteration did not affect the nature of the essential epitope sequence, *O*-glycosylated PDTR motif.

CONCLUSION

In the present study, we established an efficient protocol for the construction of neoglycopeptides by exemplifying the rapid synthesis of MUC1-related glycoprotein models having complex *O*-glycan and *N*-glycan chains based on the integration of chemical and enzymatic approaches on the functional polymer platforms. We demonstrated feasibility of SrtA-mediated ligation between two different glycopeptide segments by tagging with signal peptides, LPKTGLR and GG, at each C- or N-terminal position. Merit of the use of SrtA-mediated ligation is evident because our approach made large-scale synthesis, precise structural analyses by MS and NMR, and immunological characterization of macromolecular glycopeptides possible. Moreover, the present results clearly showed high potentials of the intervening artificial peptides having a possible effect of separating two functional glycopeptide moieties in molecular design of highly complicated glycoprotein models. It is our belief that the present strategy will greatly facilitate rapid and combinatorial synthesis of the multiply functionalized glycopeptide library as convenient tools or models for the investigation of the structure−function relationship of glycoproteins and development of a novel class of

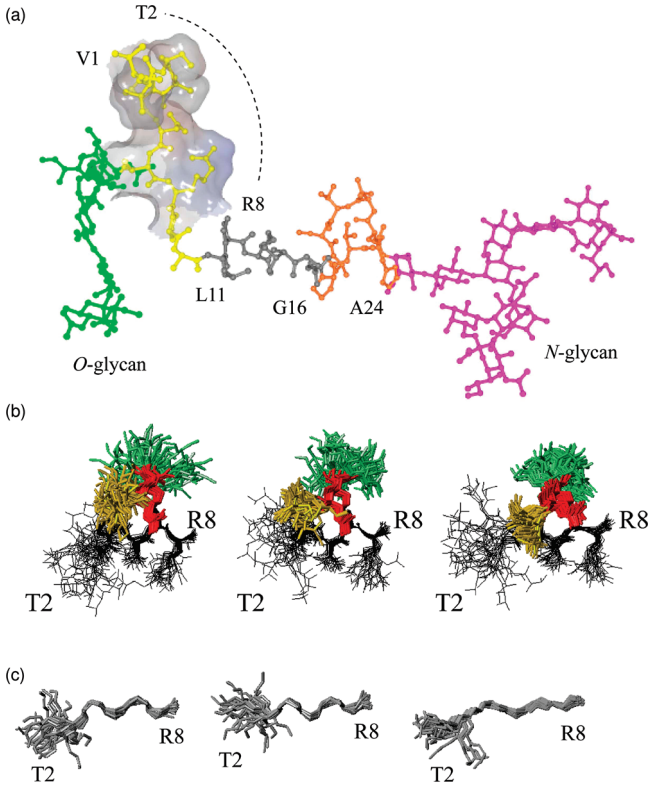


FIGURE 5: NMR structures of synthetic MUC1 glycopeptides recognized by anti-MUC1 mAb (clone B421). (a) The lowest energy structure of compound **1** shown by using ball-and-stick representation with heavy atoms and the Connolly surface (1.4 Å) for the immunodominant epitope region [^{2}TSAPDTR^{8}]. N-Terminal partial sequence of the MUC1 tandem repeat with *O*-glycan [^{1}VTSAPDTRA^{10}] is shown by yellow, the sequence required for SrtA-mediated ligation [^{11}LPKTGG^{16}] is in gray, and the C-terminal segment involving the *N*-glycosylation site [^{17}PVHNVTSA^{24}] is in orange. The *O*-glycan chain is represented in green, and *N*-glycan is in magenta, respectively. (b) Superpositions of 30 lowest energy structures of [^{2}TSAPDTR^{8}] in **1** (left), **6** (middle), and **5** (right). The core 2 trisaccharide moiety [GalNAc (red, neon), Gal (gold, neon), and GlcNAc (pale green, neon)] is shown in all structures, and black lines represent the peptide backbone. The Thr^{7} residue with the core 2 trisaccharide is superimposed with the heavy atoms of Asp^{6}-Arg^{8} and the GalNAc residue. (c) Structures of the [^{2}TSAPDTR^{8}] moiety in **1** (left), **6** (middle), and **5** (right) represented only by the peptide backbone. They are superimposed with the backbone atoms of Ala^{4}-Arg^{8}.

glycoprotein-based biopharmaceuticals, drug delivery systems, and biomedical materials.

ACKNOWLEDGMENT

We appreciate Ms. Y. Takahashi for technical assistance for the synthesis of some glycopeptides. We thank Ms. M. Kikuchi, Ms. S. Oka, and Mr. T. Hirose at the Center for Instrumental Analysis, Hokkaido University, for ESI-MS measurement.

SUPPORTING INFORMATION AVAILABLE

HPLC, MALDI-TOFMS, ESI-HRMS, and ^{1}H and ^{13}C NMR data for all new compounds; tables of chemical shifts, summary of dihedral angles, $^{1}H-^{13}C$-HSQC, NOESY, energy statistics, and structure statistics for compounds **1**, **5**, and **6**; sortase motif found in human genome. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Zachara, N. E., and Hart, G. W. (2002) The emerging significance of *O*-GlcNAc in cellular regulation. *Chem. Rev. 102*, 431–438.
2. Shriver, Z., Raguram, S., and Sasisekharan, R. (2004) Glycomics: a pathway to a class of new and improved therapeutics. *Nat. Rev. Drug Discov. 3*, 863–873.
3. Ludwig, J. A., and Weinstein, J. N. (2005) Biomarkers in cancer staging, prognosis and treatment selection. *Nat. Rev. Cancer 5*, 845–856.
4. Dube, D. H., and Bertozzi, C. R. (2005) Glycans in cancer and inflammation—potential for therapeutics and diagnostics. *Nat. Rev. Drug Discov. 4*, 477–488.
5. Molinari, M. (2007) N-glycan structure dictates extension of protein folding or onset of disposal. *Nat. Chem. Biol. 3*, 313–320.
6. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M., and Dwek, R. A. (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol. 25*, 21–50.
7. Lee, Y. C., and Lee, R. T. (1995) Carbohydrate-protein interactions: basis of glycobiology. *Acc. Chem. Res. 28*, 321–327.
8. Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem., Int. Ed. 37*, 2754–2794.
9. Lundquist, J. J., and Toone, E. J. (2002) The cluster glycoside effect. *Chem. Rev. 102*, 555–578.
10. Lee, Y. C., Stowell, C. P., and Krantz, M. J. (1976) 2-Imino-2-methoxyethyl 1-thioglycosides: new reagents for attaching sugars to proteins. *Biochemistry 15*, 3956–3963.
11. Dam, T. K., and Brewer, C. F. (2002) Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev. 102*, 387–429.
12. Gestwicki, J. E., Cairo, C. W., Strong, L. E., Oetjen, K. A., and Kiessling, L. L. (2002) Influencing receptor-ligand binding mechanisms with multivalent ligand architecture. *J. Am. Chem. Soc. 124*, 14922–14933.
13. Ohta, T., Miura, N., Fujitani, N., Nakajima, F., Niikura, K., Sadamoto, R., Guo, C. T., Suzuki, T., Suzuki, Y., Monde, K., and Nishimura, S.-I. (2003) Glycotentacles: synthesis of cyclic glycopeptides, toward a tailored blocker of influenza virus hemagglutinin. *Angew. Chem., Int. Ed. 42*, 5186–5189.
14. Elliott, S., Lorenzini, T., Asher, S., Aoki, K., Brankow, D., Buck, L., Busse, L., Chang, D., Fuller, J., Grant, J., Hernday, N., Hokum, M., Hu, S., Knudten, A., Levin, N., Komorowski, R., Martin, F., Navarro, R., Osslund, T., Rogers, G., Rogers, N., Trail, G., and Egrie, J. (2003) Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nat. Biotechnol. 21*, 414–421.
15. Sato, M., Sadamoto, R., Niikura, K., Monde, K., Kondo, H., and Nishimura, S.-I. (2004) Site-specific introduction of sialic acid into insulin. *Angew. Chem., Int. Ed. 43*, 1516–1520.
16. Sato, M., Furuike, T., Sadamoto, R., Fujitani, N., Nakahara, T., Niikura, K., Monde, K., Kondo, H., and Nishimura, S.-I. (2004) Glycoinsulins: dendritic sialyloligosaccharide-displaying insulins showing a prolonged blood-sugar-lowering activity. *J. Am. Chem. Soc. 126*, 14013–14022.
17. Ueda, T., Tomita, K., Notsu, Y., Ito, T., Fumoto, M., Takakura, T., Nagatome, H., Takimoto, A., Mihara, S., Togame, H., Kawamoto, K., Iwasaki, T., Asakura, K., Oshima, T., Hanasaki, K., Nishimura, S.-I., and Kondo, H. (2009) Chemoenzymatic synthesis of glycosylated glucagon-like peptide 1: effect of glycosylation on proteolytic resistance and in vivo blood glucose-lowering activity. *J. Am. Chem. Soc. 131*, 6237–6245.
18. Wormald, M. R., Petrescu, A. J., Pao, Y.-L., Glithero, A., Elliott, T., and Dwek, R. A. (2002) Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem. Rev. 102*, 371–386.
19. 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.
20. 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.
21. 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.
22. Friedrich-Bochnitschek, S., Waldmann, H., and Kunz, H. (1989) Allyl esters as carboxy protecting groups in the synthesis of *O*-glycopeptides. *J. Org. Chem. 54*, 751–756.
23. Sears, P., and Wong, C.-H. (2001) Toward automated synthesis of oligosaccharides and glycoproteins. *Science 291*, 2344–2350.
24. Warren, J. D., Miller, J. S., Keding, S. J., and Danishefsky, S. J. (2004) Toward fully synthetic glycoproteins by ultimately convergent routes: a solution to a long-standing problem. *J. Am. Chem. Soc. 126*, 6576–6578.
25. Liu, L., Bennett, C. S., and Wong, C.-H. (2006) Advances in glycoprotein synthesis. *Chem. Commun.,* 21–33.
26. Galonić, D. P., and Gin, D. Y. (2007) Chemical glycosylation in the synthesis of glycoconjugate antitumour vaccines. *Nature 446*, 1000–1007.
27. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science 266*, 776–779.
28. Yamamoto, N., Tanabe, Y., Okamoto, R., Dawson, P. E., and Kajihara, Y. (2008) Chemical synthesis of a glycoprotein having an intact human complex-type sialyloligosaccharide under the Boc and Fmoc synthetic strategies. *J. Am. Chem. Soc. 130*, 501–510.
29. Piontek, C., Varón Silva, D., Heinlein, C., Pöhner, C., Mezzato, S., Ring, P., Martin, A., Schmid, F. X., and Unverzagt, C. (2009) Semisynthesis of a homogeneous glycoprotein enzyme: ribonuclease C: part 2. *Angew. Chem., Int. Ed. 48*, 1941–1945.
30. Zhang, Z., Gildersleeve, J., Yang, Y. Y., Xu, R., Loo, J. A., Uryu, S., Wong, C.-H., and Schultz, P. G. (2004) A new strategy for the synthesis of glycoproteins. *Science 303*, 371–373.
31. Gamblin, D. P., Scanlan, E. M., and Davis, B. G. (2009) Glycoprotein synthesis: an update. *Chem. Rev. 109*, 131–163.
32. Unverzagt, C., Kunz, H., and Paulson, J. C. (1990) High-efficiency synthesis of sialyloligosaccharides and sialoglycopeptides. *J. Am. Chem. Soc. 112*, 9308–9309.
33. Meldal, M., Auzanneau, F.-I., Hindsgaul, O., and Palcic, M. M. (1994) A PEGA resin for use in the solid-phase chemical−enzymatic synthesis of glycopeptides. *J. Chem. Soc., Chem. Commun.,* 1849–1850..
34. Unverzagt, C. (1996) Chemoenzymatic synthesis of a sialylated undecasaccharide-asparagine conjugate. *Angew. Chem., Int. Ed. 35*, 2350–2353.
35. Seitz, O., and Wong, C.-H. (1997) Chemoenzymatic solution- and solid-phase synthesis of *O*-glycopeptides of the mucin domain of MAdCAM-1. A general route to *O*-LacNAc, *O*-sialyl-LacNAc, and *O*-sialyl-Lewis-X peptides. *J. Am. Chem. Soc. 119*, 8766–8776.
36. Nishimura, S.-I., and Yamada, K. (1997) Transfer of ganglioside GM3 oligosaccharide from a water soluble polymer to ceramide by ceramide glycanase. A novel approach for the chemical-enzymatic synthesis of glycosphingolipids. *J. Am. Chem. Soc. 119*, 10555–10556.
37. Koeller, K. M., and Wong, C.-H. (2000) Synthesis of complex carbohydrates and glycoconjugates: enzyme-based and programmable one-pot strategies. *Chem. Rev. 100*, 4465–4494.
38. George, S. K., Schwientek, T., Holm, B., Reis, C. A., Clausen, H., and Kihlberg, J. (2001) Chemoenzymatic synthesis of sialylated glycopeptides

derived from mucins and T-cell stimulating peptides. *J. Am. Chem. Soc. 123*, 11117–11125.
39. Bézay, N., Dudziak, G., Liese, A., and Kunz, H. (2001) Chemoenzymatic-chemical synthesis of a (2−3)-sialyl T threonine building block and its application to the synthesis of the N-terminal sequence of leukemia-associated leukosialin (CD 43). *Angew. Chem., Int. Ed. 40*, 2292–2295.
40. Fumoto, M., Hinou, H., Matsushita, T., Kurogochi, M., Ohta, T., Ito, T., Yamada, K., Takimoto, A., Kondo, H., Inazu, T., and Nishimura, S.-I. (2005) Molecular transporter between polymer platforms: highly efficient chemoenzymatic glycopeptide synthesis by the combined use of solid-phase and water-soluble polymer supports. *Angew. Chem., Int. Ed. 44*, 2534–2537.
41. 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.
42. 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.
43. 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.* (in press).
44. 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.
45. Mazmanian, S. K., Ton-That, H., and Schneewind, O. (2001) Sortase-catalysed anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *Mol. Microbiol. 40*, 1049–1057.
46. Paterson, G. K., and Mitchell, T. J. (2004) The biology of Gram-positive sortase enzymes. *Trends Microbiol. 12*, 89–95.
47. Perry, A. M., Ton-That, H., Mazmanian, S. K., and Schneewind, O. (2002) Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. III. Lipid II is an *in vivo* peptidoglycan substrate for sortase-catalyzed surface protein anchoring. *J. Biol. Chem. 277*, 16241–16248.
48. Huang, X., Aulabaugh, A., Ding, W., Kapoor, B., Alksne, L., Tabei, K., and Ellestad, G. (2003) Kinetic mechanism of *Staphylococcus aureus* sortase SrtA. *Biochemistry 42*, 11307–11315.
49. Mao, H., Hart, S. A., Schink, A., and Pollok, B. A. (2004) Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc. 126*, 2670–2671.
50. Parthasarathy, R., Subramanian, S., and Boder, E. T. (2007) Sortase A as a novel molecular "stapler" for sequence-specific protein conjugation. *Bioconjugate Chem. 18*, 469–476.
51. Pritz, S., Wolf, Y., Kraetke, O., Klose, J., Bienert, M., and Beyermann, M. (2007) Synthesis of biologically active peptide nucleic acid-peptide conjugates by sortase-mediated ligation. *J. Org. Chem. 72*, 3909–3012.
52. Popp, M. W., Antos, J. M., Grotenbreg, G. M., Spooner, E., and Ploegh, H. L. (2007) Sortagging: a versatile method for protein labeling. *Nat. Chem. Biol. 3*, 707–708.
53. Antos, J. M., Miller, G. M., Grotenbreg, G. M., and Ploegh, H. L. (2008) Lipid modification of proteins through sortase-catalyzed transpeptidation. *J. Am. Chem. Soc. 130*, 16338–16343.
54. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol. 215*, 403–410.
55. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389–3402.
56. 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.
57. Haneda, K., Inazu, T., Yamamoto, K., Kumagai, H., Nakahara, Y., and Kobata, A. (1996) Transglycosylation of intact sialo complex-type oligosaccharides to the *N*-acetylglucosamine moieties of glycopeptides by *Mucor hiemalis* endo-*β*-*N*-acetylglucosaminidase. *Carbohydr. Res. 292*, 61–70.
58. Mizuno, M., Haneda, K., Iguchi, R., Muramoto, I., Kawakami, T., Aimoto, S., Yamamoto, K., and Inazu, T. (1999) Synthesis of a glycopeptide containing oligosaccharides: chemoenzymatic synthesis of eel calcitonin analogues having natural N-linked oligosaccharides. *J. Am. Chem. Soc. 121*, 284–290.
59. O'Connor, S. E., Pohlmann, J., Imperiali, B., Saskiawan, I., and Yamamoto, K. (2001) Probing the effect of the outer saccharide residues of *N*-linked glycans on peptide conformation. *J. Am. Chem. Soc. 123*, 6187–6188.
60. Seko, A., Koketsu, M., Nishizono, M., Enoki, Y., Ibrahim, H. R., Juneja, L. R., Kim, M., and Yamamoto, T. (1997) Occurence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. *Biochim. Biophys. Acta 1335*, 23–32.
61. Yamamoto, K., Kadowaki, S., Watanabe, J., and Kumagai, H. (1994) Transglycosylation activity of *Mucor hiemalis* endo-*β*-*N*-acetyl-glucosaminidase which transfers complex oligosaccharides to the *N*-acetylglucosamine moieties of peptides. *Biochem. Biophys. Res. Commun. 203*, 244–252.
62. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) Improved spectral resolution in cosy ^{1}H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun. 117*, 479–485.
63. Braunschweiler, L., and Ernst, R. R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson. 53*, 521–528.
64. Bax, A., and Davis, D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson. 65*, 355–360.
65. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys. 71*, 4546–4563.
66. Piotto, M., Saudek, V., and Sklenár, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR 2*, 661–665.
67. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.
68. Godderd, T. D., and Kneller, D. G. SPARKY 3, University of California, San Francisco.
69. Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley and Sons, New York.
70. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905–921.
71. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477–486.
72. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics 14* (51−55), 29–32.
73. 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.
74. 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.
75. Müller, S., Goletz, S., Packer, N., Gooley, A., Lawson, A. M., and Hanisch, F.-G. (1997) Localization of *O*-glycosylation sites on glycopeptide fragments from lactation-associated MUC1. All putative sites within the tandem repeat are glycosylation targets in vivo. *J. Biol. Chem. 272*, 24780–24793.
76. Karsten, U., Diotel, C., Klich, G., Paulsen, H., Goletz, S., Müller, S., and Hanisch, F.-G. (1998) Enhanced binding of antibodies to the DTR motif of MUC1 tandem repeat peptide is mediated by site-specific glycosylation. *Cancer Res. 58*, 2541–2549.
77. Ryuko, K., Schol, D. J., Snijdewint, F. G, von Mensdorff-Pouilly, S., Poort-Keesom, R. J., Karuntu-Wanamarta, Y. A., Verstraeten, R. A., Miyazaki, K., Kenemans, P., and Hilgers, J. (2000) Characterization of a new MUC1 monoclonal antibody (VU-2-G7) directed to the glycosylated PDTR sequence of MUC1. *Tumor Biol. 21*, 197–210.
78. Karsten, U., Serttas, N., Paulsen, H., Danielczyk, A., and Goletz, S. (2004) Binding patterns of DTR-specific antibodies reveal a glycosylation-conditioned tumor-specific epitope of the epithelial mucin (MUC1). *Glycobiology 14*, 681–692.
79. Taylor-Papadimitriou, J., Burchell, J., Miles, D. W., and Dalziel, M. (1999) MUC1 and cancer. *Biochim. Biophys. Acta 1455*, 301–313.

80. Hanisch, F.-G., and Müller, S. (2000) MUC1: the polymorphic appearance of a human mucin. *Glycobiology 10*, 439–449.
81. Hollingsworth, M. A., and Swanson, B. J. (2004) Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer 4*, 45–60.
82. 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.
83. Hattrup, C. L., and Gendler, S. J. (2008) Structure and function of the cell surface (tethered) mucins. *Annu. Rev. Physiol. 70*, 431–457.
84. Parry, S., Hanisch, F.-G., Leir, S. H., Sutton-Smith, M., Morris, H. R., Dell, A., and Harris, A. (2006) N-Glycosylation of the MUC1 mucin in epithelial cells and secretions. *Glycobiology 16*, 623–634.
85. Bernatowicz, M. S., Daniels, S. B., and Köster, H. (1989) A comparison of acid labile linkage agents for the synthesis of peptide C-terminal amides. *Tetrahedron Lett. 30*, 4645–4648.
86. Mathieux, N., Paulsen, H., Meldal, M., and Bock, K. (1997) Synthesis of glycopeptide sequences of repeating units of the mucins MUC 2 and MUC 3 containing oligosaccharide side-chains with core 1, core 2, core 3, core 4 and core 6 structure. *J. Chem. Soc., Perkin Trans. 1,* 2359–2368.
87. Meinjohanns, E., Meldal, M., Schleyer, A., Paulsen, H., and Bock, K. (1996) Efficient syntheses of core 1, core 2, and core 3 and core 4 building blocks for SPS of mucin O-glycopeptides based on the N-Dts-method. *J. Chem. Soc., Perkin Trans. 1,* 985–993.
88. Knorr, R., Trzeciak, A., Bannwarth, W., and Gillessen, D. (1989) New coupling reagents in peptide chemistry. *Tetrahedron Lett. 30*, 1927–1930.
89. Köenig, W., and Volk, A. (1977) Succinimidbildung bei der synthese des insulin-A-ketten(14−21)-octapeptids. *Chem. Ber. 110*, 1–11.
90. Rink, H. (1987) Solid-phase synthesis of protected peptide fragments using a trialkoxy-diphenyl-methylester resin. *Tetrahedron Lett. 28*, 3787–3790.
91. Lloyd, K. O., Burchell, J., Kudryashov, V., Yin, B. W., and Taylor-Papadimitriou, J. (1996) Comparison of *O*-linked carbohydrate chains in MUC-1 mucin from normal breast epithelial cell lines and breast carcinoma cell lines. Demonstration of simpler and fewer glycan chains in tumor cells. *J. Biol. Chem. 271*, 33325–33334.
92. Bystrov, V. F., Portnova, S. L., Tsetlin, V. I., Ivanov, V. T., and Ovchinnikov, Yu. A. (1969) Conformational studies of peptide systems: the rotational states of the NH—CH fragment of alanine dipeptides by nuclear magnetic resonance. *Tetrahedron 25*, 493–515.
93. Dziadek, S., Griesinger, C., Kunz, H., and Reinscheid, U. M. (2006) Synthesis and structural model of an α(2,6)-sialyl-T glycosylated MUC1 eicosapeptide under physiological conditions. *Chem.—Eur. J. 12*, 4981–4993.
94. Kuhn, A., and Kunz, H. (2007) Saccharide-induced peptide conformation in glycopeptides of the recognition region of LI-Cadherin. *Angew. Chem., Int. Ed. 46*, 454–458.