Published in final edited form as:

Bioorg Med Chem. 2013 August 15; 21(16): 4778–4785. doi:10.1016/j.bmc.2013.02.040.

One-pot multi-enzyme (OPME) chemoenzymatic synthesis of sialyl-Tn-MUC1 and sialyl-T-MUC1 glycopeptides containing natural or non-natural sialic acid[†]

Hamed Malekan^a, Gabriel Fung^b, Vireak Thon^{a,c}, Zahra Khedri^{a,d}, Hai Yu^a, Jingyao Qu^a, Yanhong Li^a, Li Ding^{a,e}, Kit S. Lam^b, and Xi Chen^{a,*}

^aDepartment of Chemistry, University of California, One Shields Avenue, Davis, CA 95616, USA

^bDepartment of Biochemistry and Molecular Medicine, University of California, 2700 Stockton Blvd, Sacramento, CA 95817, USA

Abstract

A series of STn-MUC1 and ST-MUC1 glycopeptides containing naturally occurring and non-natural sialic acids have been chemoenzymatically synthesized from Tn-MUC1 glycopeptide using one-pot multienzyme (OPME) approaches. *In situ* generation of the sialyltransferase donor cytidine 5'-monophosphate-sialic acid (CMP-Sia) using a CMP-sialic acid synthetase in the presence of an extra amount of cytidine 5'-triphosphate (CTP) and removal of CMP from the reaction mixture by flash C18 cartridge purification allow the complete consumption of Tn-MUC1 glycopeptide for quantitative synthesis of STn-MUC1. A *Campylobacter jejuni* β 1-3GalT (CjCgtB Δ 30-His₆) mutant has been found to catalyze the transfer of one or more galactose residues to Tn-MUC1 for the synthesis of T-MUC1 and galactosylated T-MUC1. Sialylation of T-MUC1 using *Pasteurella multocida* α 2-3-sialyltransferase 3 (PmST3) with *Neisseria meningitidis* CMP-sialic acid synthetase (NmCSS) and *Escherichia coli* sialic acid aldolase in one pot produced ST-MUC1 efficiently. These glycopeptides are potential cancer vaccine candidates.

Keywords

carbohydrate; chemoenzymatic synthesis; enzymatic synthesis; glycopeptide; T antigen; Tn antigen; ST antigen; STn antigen

1. Introduction

Mucins are a family of high molecular weight extensively O-glycosylated membrane-bound or secreted proteins produced by epithelial tissues. They have various numbers of peptide tandem repeats of diverse lengths and are involved in the protection and lubrication of epithelia surfaces,¹ signaling transduction regulation, bacterial and viral binding, and are

[†]Electronic Supplementary Information (ESI) available: High resolution mass spectrometry (HRMS) spectra of glycopeptide products.

© 2013 Elsevier Ltd. All rights reserved.

*Corresponding author. Tel: +1 530 754 6037; Fax: +1 530 752 8995; address: chen@chem.ucdavis.edu.

^cCurrent Address: Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA

^dCurrent Address: Department of Pharmaceutical Chemistry, University of Kansas, Lawrence, KS 66047, USA

^eCurrent Address: College of Life Sciences, Northwest University, Xian, Shaanxi, China

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

associated with embryogenesis, carcinogenesis, and cancer metastasis.² Among at least 19 mucins discovered to date,³ mucin 1 (MUC1) is the most extensively studied one. Human MUC1 is a transmembrane-bound mucin whose extracellular region contains 25–125 of 20-amino acid tandem repeat (GSTAPPAHGVTSAPDTRPAP).⁴ The O-glycans of MUC1 expressed on normal epithelial cell surfaces are commonly core 2-type (Gal β 1–3[GlcNAc β 1–6]GalNAc α Ser/Thr) with extended *N*-acetylactosamine (LacNAc).⁵ Overexpression of MUC1 is associated with many types of cancers. Aberrant expression of truncated O-GalNAc glycans including Tn (GalNAc α Ser/Thr), STn (Sia α 2–3GalNAc α Ser/Thr), and Thomsen-Friedenreich (TF or T-antigen, Gal β 1–3GalNAc α Ser/Thr) antigens on mucin MUC1 are associated with epithelial cancers.^{6–8} For example, STn-MUC1 has been found on up to 80% of breast and ovarian cancer cells. Its expression is associated with poor prognosis and is related to metastasis.⁹ Sialylated T (ST, Sia α 2–3 Gal β 1–3GalNAc α Ser/Thr) antigen has also found in numerous tumor tissues and cancer cell lines.^{10–12} Tn- and STn-MUC1 antigens have both been investigated as cancer vaccine candidates. The STn-MUC1 vaccines studied,^{9, 13–15} however, contain exclusively *N*-acetylneuraminic acid (Neu5Ac), the most common sialic acid form. *N*-Glycolylneuraminic acid (Neu5Gc), a sialic acid form commonly found in animals but cannot be synthesized in humans due to the lack of an active CMP-Neu5Ac hydroxylase (CMAH)¹⁶ and the absent of an alternative biosynthetic pathway,¹⁷ has been found to be overexpressed on cancer cells by metabolic incorporation.^{18–20} Low levels of circulating anti-Neu5Gc antibodies in human can facilitate tumor progression via chronic inflammation^{20, 21} while high level of anti-Neu5Gc antibodies can be used to kill cancer cells.²² Using a glycans microarray containing a library of sialosides, elevated levels of antibodies against STn antigen containing Neu5Gc (Neu5Gc-STn) have been recently found in breast, prostate, ovary, lung, and endometrial cancer patients.²² The cancer-associated Neu5Gc-STn and anti-Neu5Gc-STn antibodies can be potentially applied for cancer vaccine and immunotherapeutic studies.

Metabolic engineering-based incorporation of C5-modified non-natural sialic acids to cell surface glycoconjugates, combined with vaccination using tumor associated carbohydrate antigen (GM3, polysialic acid, or STn) analogs containing the same terminal non-natural sialic acid,^{23–26} has been shown to be effective in eliciting specific antibodies for immunotherapy.

Most anti-MUC1, anti-Tn-MUC1, or anti-STn-MUC1 antibodies target the PDTRP region of MUC1.^{5, 27} Patients with antibodies to the GSTA region of MUC1 have also been identified. Recently, monoclonal antibodies specifically recognizing Tn/STn-antigens linked to the GSTA region of the MUC1 tandem repeats have been identified and are believed to have potential in diagnosis and therapeutic applications.^{5, 28}

Here we describe solid phase-supported synthesis of Tn-MUC1 GSTA region APG(GalNAc α)STAPPA and the application of efficient one-pot multienzyme (OPME) chemoenzymatic systems in high-yield synthesis of T-MUC1 glycopeptide as well as STn- and ST-MUC1 glycopeptides containing Neu5Ac and its C5- or C9-modified sialic acid derivatives such as Neu5Gc, *N*-azidoacetylneuraminic acid (Neu5AcN₃), and 9-azido-9-deoxy-*N*-acetylneuraminic acid (Neu5Ac9N₃).

2. Results and Discussion

2.1. Cloning, expression, and purification of CjCgtBA30-His₆ mutant

A C-His₆-tagged C-terminal 30 amino acid residue truncated *Campylobacter jejuni* OH4384 β 1-3-galactosyltransferase (CjCgtBA30-His₆) mutant (amino acid changes in the mutant are: N26K, L68I, K151E, L227G, and E234D)^{29, 30} was cloned using a synthetic gene with codon optimized for *Escherichia coli* expression systems (Figure 1). Expression in Luria-

Bertani (LB) medium with isopropyl-1-thio- β -D-galactopyranoside (IPTG, 0.1 mM) induction followed by nickel-nitrilotriacetic acid-agarose (Ni^{2+} -NTA-agarose) affinity column routinely produced 20 mg of purified protein per liter of bacterial cell culture with > 95% purity.

2.2. Synthesis of per-*O*-acetylated GalNAc α -linked to Fmoc-protected serine as a building block for solid phase glycopeptide synthesis of Tn-MUC1

Per-*O*-acetylated GalNAc(Fmoc)Ser containing a fluorenylmethyloxycarbonyl (Fmoc)-protected amino group was synthesized as a key building block for the solid-phase-supported synthesis of Tn-MUC1 nonapeptide APG(Per-OAc-GalNAc α)STAPPA. For the synthesis of the per-*O*-Ac-GalNAc(Fmoc)Ser building block, trimethylsilyl trifluoromethanesulfonate (TMSOTf) was used as the promoter. Peracetylated 2-azido-2-deoxygalactosyl trichloroacetimidate was used as the glycosyl donor to minimizing the formation of β -GalNAc linkage by neighboring group participation of GalNAc-based donors. Upon the completion of glycosylation, the 2-azido group was easily converted to *N*-acetyl by Zinc reduction of azide to amine followed by *N*-acetyl protection using acetic anhydride.³¹

2.3. Solid phase synthesis of Tn-MUC1 glycopeptide APG(GalNAc α)STAPPA

Nonapeptide APG(Per-OAc-GalNAc α)STAPPA containing per-*O*-acetylated GalNAc was produced by solid phase peptide synthesis (Scheme 1). Each peptide bond formation reaction was performed in one and a half hours by adding 5 equivalents of each amino acid, HOBt, and DIC coupling reagents except for GalNAc α .Ser (1.5 eq.) which was carried out for 8 hours. A dansyl chromophore was introduced as a UV/Vis probe to the N-terminus of the nonapeptide to facilitate high-performance liquid chromatography (HPLC) analysis.³² The resulting glycopeptide was released from the solid support by TFA treatment. Removal of the *O*-acetyl groups on the GalNAc for the formation of target APG(GalNAc α)STAPPA was achieved by treatment with an aqueous solution of NaOH. To minimize the base-catalyzed cleavage of the glycosidic bond of GalNAc α .Ser by β -elimination during *O*-acetyl deprotection, closely monitoring of the reaction and controlling the pH of the reaction in the range of 8.5–9.5 were critical. The produced Tn-MUC1 glycopeptide APG(GalNAc α)STAPPA was used as a starting materials for one-pot multienzyme (OPME) synthesis of T-MUC1 glycopeptide as well as STn- and ST-MUC1 glycopeptides containing different sialic acid forms (Scheme 1).

2.4. One-pot three-enzyme synthesis of STn-MUC1 glycopeptides APG(Sia α 2–6GalNAc α)STAPPA

Two bacterial α 2–6-sialyltransferases including *Photobacterium damsela* (Pd2,6ST)³³ and *Photobacterium* sp. (Psp2,6ST)⁸ α 2–6-sialyltransferases were tested in a one-pot three-enzyme α 2–6-sialylation system containing the sialyltransferase, *Escherichia coli* sialic acid aldolase (EcNanA),³⁴ and *N. meningitidis* CMP-sialic acid synthetase (NmCSS)³³ for the synthesis of STn-MUC1 glycopeptides from chemically synthesized Tn-MUC1 glycopeptide (Scheme 2). Both α 2–6-sialyltransferases worked well and led to high efficient sialylation with comparable yields. Due to its higher expression level, Pd2,6ST was used for preparative-scale synthesis of STn-MUC1.

Donor hydrolysis activity for which water molecules compete with acceptors for glycosylation is common to glycosyltransferase-catalyzed reactions (Scheme 2).³⁵ In order to remain sufficient amount of CMP-Sia in the one-pot three-enzyme chemoenzymatic sialylation system shown in Scheme 2, CTP was added periodically during the reaction to allow continuous generation of sialyltransferase donor CMP-Sia from sialic acid formed by the donor hydrolysis activity of the sialyltransferase. This strategy worked well and allowed

high-yield production of the desired product. In addition, flash C18 cartridge purification of the reaction mixture to separate glycosylated product and remaining acceptor from other components followed by an additional OPME reaction allowed the complete consumption of the acceptor. The combination of adding CTP periodically and C18 cartridge purification followed by an additional OPME reaction was successfully exploited for the synthesis of STn-MUC1 containing different sialic acid forms including natural Neu5Ac and Neu5Gc as well as non-natural Neu5AcN₃ and Neu5Ac9N₃ in excellent yields (89–91%) (Scheme 1, step c). The purity of the products was demonstrated by high-performance liquid chromatography (HPLC) chromatograms and high-resolution mass spectrometry (HRMS) spectra (Figure 2 and Supporting Information).

2.5. One-pot four-enzyme synthesis of T-MUC1 glycopeptide APG(Galβ1–3GalNAcα)STAPPA

To obtain T-MUC1 glycopeptide APG(Galβ1–3GalNAcα)STAPPA, two OPME systems were tested including a one-pot two-enzyme system containing *Bifidobacterium infantis* D-galactosyl-β1–3-*N*-acetyl-D-hexosamine phosphorylase (BiGalHexNAcP) and *E. coli* galactokinase (EcGalK),³⁶ and a one-pot four-enzyme galactosylation system containing *Campylobacter jejuni* β1–3-galactosyltransferase (CjCgtBA30-His₆) mutant,³⁰ *E. coli* galactokinase (EcGalK), *Bifidobacterium longum* UDP-sugar pyrophosphorylase (BLUSP), and *Pasteurella multocida* inorganic pyrophosphatase (PmPpA).³⁷ Although BiGalHexNAcP was shown to have broad acceptor substrate specificity and can use GalNAcαSer and GalNAcαThr as substrates for high-yield (91–92%) synthesis of T-antigens,³⁶ the Tn-MUC1 glycopeptide APG(GalNAcα)STAPPA was not an acceptor substrate for the enzyme.

CjCgtBA30-His₆ mutant in one-pot four-enzyme system (Scheme 3) produced the desired T-MUC1 glycopeptide. Quite interestingly, after adding one galactose (Gal) residue to the Tn-MUC1 glycopeptide for the formation of T-MUC1 glycopeptide, the CjCgtBA30-His₆ mutant was able to add additional Gal residues to T-MUC1 glycopeptide for the formation of poly-Gal-glycopeptides as demonstrated by extra peaks in HPLC chromatogram and confirmed by HRMS analysis. The additional galactose residues are mostly likely linked through β1–3-linkage but this needs to be confirmed in the future by a larger scale preparation of the products followed by NMR characterization of the product. OPME galactosylation with periodic addition of ATP and UTP combined with flash C18-cartridge purification allow the total consumption of the Tn-MUC1 glycopeptide and produced T-MUC1 glycopeptide in 55% yield and Gal-T-MUC1 glycopeptide with an additional Gal residue in 34% yield.

2.6. One-pot three-enzyme synthesis of ST-MUC1 glycopeptides APG(Siaα2–3Galβ1–3GalNAcα)STAPPA

Enzymatic synthesis of ST-MUC1 glycopeptides APG(Siaα2–3Galβ1–3GalNAcα)STAPPA containing different sialic acid forms from T-MUC1 glycopeptide was achieved by a one-pot three-enzyme α2–3-sialylation system similar to that described in Scheme 2 for the synthesis of STn-MUC1 glycopeptides except that an α2–3-sialyltransferase, instead of an α2–6-sialyltransferase, was used. Among several bacterial α2–3-sialyltransferases and mutants with good expression levels including *Pasteurella multocida* α2–3-sialyltransferase 1 (PmST1),³⁸ PmST2,³⁹ PmST3,⁴⁰ and PmST1 mutants,^{35, 41} PmST3 was chosen due to its good expression level and excellent acceptor substrate tolerance towards oligosaccharides and glycolipids. ST-MUC1 glycopeptides containing Neu5Ac, Neu5Ac9N₃, Neu5Gc, and Neu5AcN₃ were obtained in excellent yields (85–89%) (Scheme 1). The purity of the products was demonstrated by high-performance liquid chromatography (HPLC) chromatograms and high-resolution mass spectrometry (HRMS) spectra (Figure 2 and

Supporting Information). The lower yields obtained for the synthesis of ST-MUC1 glycopeptides compared to STn-MUC1 glycopeptides were most likely due to the lower amount (10 fold less) of the starting material used which led to a larger percentage of material lost during reaction and purification.

3. Conclusion

Several OPME chemoenzymatic approaches were shown to be highly effective in synthesizing T-MUC1 as well STn- and ST-MUC1 glycopeptides containing natural and non-natural sialic acid forms. Periodic addition of nucleoside triphosphates during OPME reactions for the regeneration of sugar nucleotides was shown to improve the yields of the enzymatic reactions. This can be used as a general strategy for glycosylating glycopeptides and glycoproteins. Flash C18-cartridge purification can be used to further improve the glycosylation yields of OPME reactions for glycopeptides. A newly reported OPME UDP-Gal *in situ* generation system was successfully used with a β 1–3-galactosyltransferase (CjCgtB Δ 30-His₆ mutant) in one pot for efficient synthesis of T-MUC1 glycopeptide. CjCgtB Δ 30-His₆ mutant was shown to be able to add one or more Gal residues to Tn-MUC1 glycopeptide for the formation of T-MUC1 and Gal-T-MUC1 glycopeptides. STn-MUC1 and ST-MUC1 glycopeptides containing different sialic acid forms can be useful probes for diagnosis. They can also be vaccine candidates for therapeutic applications.

4. Experimental

4.1. Materials

Vector plasmid pET22b(+) was purchased from Novagen (EMD Biosciences, Inc. Madison, WI). Ni²⁺-NTA agarose (nickel-nitrilotriacetic acid agarose), QIAprep spin miniprep kit, and QIAEX II gel extraction kit were from Qiagen (Valencia, CA). Herculase enhanced DNA polymerase was from Stratagene (La Jolla, CA). Electrocompetent *E. coli* DH5 α cells and chemically competent *E. coli* BL21 (DE3) cells were from Invitrogen (Carlsbad, CA). T4 DNA ligase, 1 kb DNA ladder, and restriction enzymes were from Promega (Madison, WI). *N,N'*-Diisopropylcarbodiimide (DIC), *N*-hydroxybenzotriazole (HOBt), MBHA-Rink amide resin, and amino acids were purchased from Chem-Impex International (Wood Dale, IL). Triisopropylsilane (TIPS) and *N,N*-diisopropylethylamine (DIPEA) were purchased from Fisher Scientific (Chicago, IL). Dabsyl chloride was purchased from AnaSpec (Fremont, CA). D-*N*-Acetylmannosamine (ManNAc) was purchased from Sigma (St. Louis, MO). *N*-Acetylmannosamine derivatives including *N*-glycolylmannosamine (ManNGc),⁴² *N*-azidoacetylmannosamine (ManNAcN₃),⁴² and 6-azido-6-deoxy-*N*-acetylmannosamine (ManNAc6N₃)³⁸ were chemically synthesized as reported previously. All other materials were obtained from commercial suppliers and used as received with or without further purification.

4.2. Cloning, expression, and purification of *Campylobacter jejuni* CgtB

4.2.1. Cloning—A synthetic gene encoding *Campylobacter jejuni* CgtB mutant (N26K, L68I, K151E, L227G, and E234D) with codons optimized for *E. coli* expression (designed based on GenBank Accession no. AAF31770) was custom synthesized by Biomatik (Wilmington, DE, USA) and provided in a pGH vector. It was cloned into pET22b(+) vector for a C-His₆-tagged C-terminal 30 amino acid residue truncated fusion protein. The primers used were: forward primer 5'-GATCCATATGTTCAAATTTCTATCATCCTGCCG-3' (*Nde*I restriction site is underlined) and reverse primer 5'-CAGCGTCGACCTTAATTTTGTAGATCTGAATATAC-3' (*Sa*I restriction site is underlined). Polymerase chain reactions (PCRs) for amplifying the target gene were performed in a 50 μ L reaction mixture containing plasmid DNA (10 ng), forward and

reverse primers (0.2 μ M each), 1 $\times$ Herculase buffer, dNTP mixture (0.2 mM), and 5 U (1 μ L) of Herculase-enhanced DNA polymerase. The reaction mixture was subjected to 30 cycles of amplification at an annealing temperature of 55°C. The resulted PCR product was purified and double digested with *Nde*I and *Sa*I restriction enzymes. The purified and digested PCR product was ligated with the predigested pET22b(+) vector and transformed into electrocompetent *E. coli* DH5 α cells. Selected clones were grown for minipreps and characterized by restriction mapping. DNA sequencing was performed by the Davis Sequencing Facility in the University of California-Davis.

4.2.2. Expression—Positive plasmids were selected and transformed into *E. coli* BL21 (DE3) chemical competent cells. The plasmid-bearing *E. coli* strains were cultured in LB-rich medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) supplemented with ampicillin (100 μ g/mL). Overexpression of the target protein was achieved by inducing the *E. coli* culture with 0.1 mM of isopropyl-1-thio- β -D-galactopyranoside (IPTG) when OD_{600 nm} reached 0.8–1.0 followed by incubating at 20 °C for 20 h with vigorous shaking at 250 rpm in a C25KC incubator shaker (New Brunswick Scientific, Edison, NJ).

4.2.3. Purification—The C-His₆-tagged protein was purified from cell lysate. To obtain the cell lysate, cell pellet harvested by centrifugation at 2,683 $\times$ g for 2 h was resuspended in lysis buffer (pH 8.0, 100 mM Tris-HCl containing 0.1% Triton X-100) (20 mL/L cell culture). Lysozyme (50 μ g/mL) and DNaseI (3 μ g/mL) were then added and the mixture was incubated at 37 °C for 60 min with vigorous shaking. Cell lysate was obtained by centrifugation at 11,270 $\times$ g for 30 min as the supernatant. Purification of His-tagged proteins from the lysate was achieved using a Ni²⁺-resin column. The column was pre-equilibrated with 10 column volumes of binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl, pH 7.5) before the lysate was loaded. After washing with 8 column volumes of binding buffer and 10 column volumes of washing buffer (40 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl, pH 7.5), the protein was eluted with an elute buffer (200 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl, pH 7.5). The fractions containing the purified enzymes were collected, dialyzed against Tris-HCl buffer (20 mM, pH 7.5) containing 10% glycerol and stored at 4 °C.

4.3. General methods

HPLC analysis and purification were carried out using Shimadzu LC-2010_{HT} (analytical run) and Waters 2525 binary gradient module (preparative run). HPLC columns used for analysis and purification were reverse-phase C18 columns: Waters Spherisorb ODS-2 (250 $\times$ 4.6 mm i.d., 5 μ m particle size) and Waters XTerra OBD (150 $\times$ 19 mm i.d., 5 μ m particle size), respectively. High resolution electrospray ionization (ESI) mass spectra were obtained in positive and negative modes using Thermo Electron LTQ-Orbitrap mass spectrometer. Thin-layer chromatography (TLC) was performed on silica gel plates 60 GF254. Flash C18 cartridge purifications were performed using Discovery DSC-18 3 mL tubes. Glycopeptides were concentrated and dried using Thermo Scientific SAVANT ISS110 SpeedVac concentrator.

4.4. Synthesis of per-O-acetylated GalNAc α -linked to Fmoc-protected serine as a building block for solid phase glycopeptide synthesis of Tn-MUC1

Synthesis of the Tn building block, per-OAc-GalNAc α (Fmoc)Ser, was carried out similar to that reported.³⁶ Briefly, glycosylation of 2-azido-2-deoxygalactosyl trichloroacetimidate donor (400 mg, 844 μ mol) with Fmoc (at amino group) and *t*-butyl (at the carboxyl group) protected serine (290 mg, 760 μ mol) using TMSOTf (20 μ L, 84 μ mol) as a promoter provided glycosylated product with a good yield (469 mg, 89%). Subsequent conversion of the azido group to an *N*-acetyl group using Zn/THF/HOAc/Ac₂O (472 mg, 80%) and

removal of the *t*-butyl group using TFA/H₂O produced the glycosylated Fmoc-protected amino acid building block per-*O*Ac-GalNAcα(Fmoc)Ser in quantitative yield (442 mg, 100%).

4.5. Solid phase synthesis of Tn-MUC1 glycopeptide APG(GalNAcα)STAPPA

Tn-MUC1 glycopeptide APG(Per-*O*Ac-GalNAcα)STAPPA containing per-*O*-acetylated GalNAc was synthesized on Rink Amide-MBHA resin (100 mg, 0.55 mmol/g) employing standard Fmoc solid phase peptide synthetic procedures. The hydroxyl group of threonine was protected by *t*-butyl group. Fmoc was removed from the resin (Scheme 1) by treating the resin twice for 20 minutes each time with 5 mL of 20% piperidine solution in DMF. Before and after coupling of each amino acid, the resin was washed twice with 5 mL of DMF, twice with 5 mL of methanol, and twice again with 5 mL of DMF. Reactions were monitored by Kaiser Test before and after each coupling. To add the first amino acid residue alanine (Ala) to the resin, Fmoc-Ala-OH (85 mg, 275 μmol, 5 equiv.), HOBt (42 mg, 275 μmol, 5 equiv.), and DIC (43 μL, 275 μmol, 5 equiv.) were dissolved in 5 mL of DMF and added to the resin in a capped peptide synthesis column. The reaction was mixed constantly at room temperature for more than 90 min using a Rotamix rotator from Appropriate Technical Resources (ATR) until Kaiser Test assured the completion of the reaction. The similar procedure was applied for all peptide bond formation reactions with other non-glycosylated amino acid residues. The coupling of per-*O*Ac-GalNAcα(Fmoc)Ser was achieved by dissolving the building block (70 mg, 84 μmol, 1.5 equiv.), HOBt (13 mg, 84 μmol, 1.5 equiv.), and DIC (13 μL, 84 μmol, 1.5 equiv.) in 5 mL of DMF and then mixed with the resin. The reaction was carried out for more than 8 h until Kaiser Test assured the completion of the reaction. To add a dabsyl group, dabsyl chloride (36 mg, 111 μmol, 2 equiv.) was added to resin-bound peptide in 5 mL of DMF/DIPEA (95:5 by volume). The resulting glycopeptide was then removed from the resin by treatment with TFA:TIPS:H₂O (95:2.5:2.5 by volume, 5 mL) for 2 h. The *t*-butyl protection group of threonine was simultaneously removed during this step. The TFA was evaporated by rotary evaporation and cold diethyl ether (40 mL) was added to precipitate the glycopeptide. The suspended mixture was transferred to a 50 mL tube and centrifuged at 3835 ×g for 5 minutes at 4 °C. The ether was discarded by decanting. The glycopeptide was washed twice by repeating the procedure (Crude yield = 65%). Deacetylation of the *O*-acetylated glycopeptide was carried out by dissolving in NaOH/MeOH (5 mL, pH = 9.5) followed by incubation at room temperature for 16 h. The deprotection process was monitored by MALDI-TOF carefully. The pH was adjusted to 9.5 every hour by adding NaOH to avoid the cleavage of labile α-GalNAc linkage under acidic conditions caused by pH decrease due to acetic acid formed. Upon completion, the reaction mixture was neutralized using acidic Amberlite resin. After filtration and lyophilization, the synthesized Tn glycopeptide were purified by reverse-phase HPLC (Yield = 90%) and confirmed by high-resolution mass spectrometry (HRMS).

4.6. One-pot three-enzyme synthesis of STn-MUC1 glycopeptides APG(Siaα2-6GalNAcα)STAPPA

Tn-MUC1 glycopeptide **2** (4 mg, 3 μmol, 1.0 equiv.), sodium pyruvate (5 equiv.), and CTP (2 equiv.) were dissolved in water and in a 0.5 mL centrifuge tube containing Tris-HCl buffer (100 mM, pH 8.5) and MgCl₂ (20 mM). The resulting mixture was split equally to 4 tubes and ManNAc or its derivative (1.5 equiv.) was added to each tube. After the addition of EcNanA (25 μg/mL), NmCSS (30 μg/mL), and Pd2,6ST (50 μg/mL), ddH₂O was added to bring the volume of the reaction mixture to 300 μL. The reaction mixture was incubated at 37 °C for 6 h with shaking (140 rpm). Additional CTP (1 equiv.) was added and the reaction was continued for 10 h. The product formation was monitored by thin-layer chromatography (TLC) developed with EtOAc:MeOH:H₂O:HOAc (4:2:1:0.2) as the developing solvent. After TLC indication that the reaction reached 90% yield, the reaction

mixture was loaded to a C18 cartridge for flash purification. The cartridge was washed six times with 1 mL of water each time and the glycopeptide was eluted with 2 mL of methanol and dried using SpeedVap concentrator. The obtained glycopeptide mixture was subjected to another round of OPME reaction to allow the complete consumption of the Tn-MUC1 glycopeptide acceptor. The resulting STn-MUC1 glycopeptides were dissolved in a 1:1 mixture of CH₃CN:H₂O (100 μ L) and purified by reverse-phase HPLC column as the final purification step.

4.7. One-pot four-enzyme synthesis of T-MUC1 glycopeptide APG(Gal β 1–3GalNAc α)STAPPA

Tn-MUC1 glycopeptide **2** (1 mg, 0.8 μ mol, 1.0 equiv.), galactose (1.5 equiv.), ATP (2.0 equiv.), and UTP (2.0 equiv.) were dissolved in water in a 0.5 mL centrifuge tube containing Tris-HCl buffer (100 mM, pH 7.5), MgCl₂ (10 mM), and MnCl₂ (10 mM). After the addition of EcGalK (40 μ g/mL), BLUSP (50 μ g/mL), PmPpA (30 μ g/mL), and CjCgtBA30-His₆ mutant (55 μ g/mL), ddH₂O was added to bring the total volume of the reaction mixture to 300 μ L. The reaction mixture was incubated at 37 °C for 6 h with shaking at 1.5 $\times$ g. Additional ATP and CTP (equiv. each) were added and the reaction was continued for 10 h. The product formation was monitored by TLC with EtOAc:MeOH:H₂O:HOAc (4:2:1:0.2 by volume) as the developing solvent. The C18 cartridge purification and the repeat of OPME were carried out similar to those described above for the synthesis of STn-MUC1 glycopeptides.

4.8. One-pot three-enzyme synthesis of ST-MUC1 glycopeptides APG(Sia α 2–3Gal β 1–3GalNAc α)STAPPA

T-MUC1 glycopeptide **7** (0.4 mg, 0.3 μ mol, 1.0 equiv.), sodium pyruvate (5 equiv.), and CTP (2 equiv.) were dissolved in water in a 0.5 mL centrifuge tube containing Tris-HCl buffer (100 mM, pH 8.5) and MgCl₂ (20 mM). The resulting mixture was split equally to 4 tubes and ManNAc or its derivative (1.5 equiv.) was added to each tube. After the addition of EcNanA (25 μ g/mL), NmCSS (30 μ g/mL), and PmST3 (50 μ g/mL), millipore water was added to bring the volume to 80 μ L. The reaction and product purification were carried out similarly to those described above for the synthesis of STn-MUC1 glycopeptides.

4.9. HRMS and HPLC analysis and purification of Tn-, T-, STn, and ST MUC1

4.9.1. HPLC analysis and purification—The Tn-MUC1 glycopeptide (10 mg, retention time 13.6 min) was purified by reverse-phase C18 column (Waters XTerra OBD, 5 μ m, 19 $\times$ 150 mm) with a flow rate of 10 mL/min and a gradient elution of 0–50% of A from 0–30 min and monitored at 480 nm [(A) CH₃CN containing 0.1% TFA and (B) H₂O containing 0.1% TFA]. STn-MUC1 and ST-MUC1 glycopeptides were purified by reverse-phase C18 column (Waters Spherisorb ODS, 2.5 μ m, 4.6 $\times$ 250 mm) with a flow rate of 0.8 mL/min and a gradient elution of 30–50% of C from 0–17 min and monitored at 480 nm [(C) CH₃CN containing 0.1% formic acid and (D) H₂O containing 0.1% formic acid]. HPLC purification of the T-MUC1 glycopeptide was achieved similarly using 30–40% of C from 0–50 minutes. The yields of OPME reactions were calculated by comparing the HPLC peaks of the purified glycopeptides to the HPLC peak of a known concentration of Tn-MUC1 glycopeptide standard. The values were further confirmed using the molar extinction coefficient (ϵ = 32,000) of dabsyl measured at 480 nm. All glycopeptides were analyzed by HPLC assays (Table 1) using an analytic C18 column (Waters Spherisorb ODS, 2.5 μ m, 4.6 $\times$ 250 mm) with the same conditions as for STn-MUC1 and ST-MUC1 purifications.

4.9.2. High-resolution LTQ-Orbitrap hybrid MS analysis—The mass spectrometric analysis of the obtained glycopeptides was performed on LTQ-Orbitrap hybrid with

nanospray source in the positive mode for Tn-MUC1 and T-MUC1 glycopeptides **2** and **7**. HRMS spectra for STn-MUC1 and ST-MUC1 glycopeptides **3–6** and **8–11** were obtained in the negative mode.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grant R01HD065122, NSF grant CHE-1012511, and the Camille & Henry Dreyfus Foundation. X. Chen is a Camille Dreyfus Teacher-Scholar and a UC-Davis Chancellor's Fellow.

References

1. Forstner JF. Digestion. 1978; 17:234. [PubMed: 25218]
2. Hollingsworth MA, Swanson BJ. Nat Rev Cancer. 2004; 4:45. [PubMed: 14681689]
3. Perez-Vilar, J.; Hill, RL. Encyclopedia of Biological Chemistry. Lennarz, WJ.; Lane, MD., editors. Vol. 2. Oxford: Academic Press/Elsevier; 2004. p. 758
4. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, Duhig T, Peat N, Burchell J, Pemberton L, Lalani EN, Wilson D. J Biol Chem. 1990; 265:15286. [PubMed: 1697589]
5. Tarp MA, Sorensen AL, Mandel U, Paulsen H, Burchell J, Taylor-Papadimitriou J, Clausen H. Glycobiology. 2007; 17:197. [PubMed: 17050588]
6. Springer GF. Science. 1984; 224:1198. [PubMed: 6729450]
7. Taylor-Papadimitriou J, Epenetos AA. Trends Biotechnol. 1994; 12:227. [PubMed: 7765069]
8. Ding L, Yu H, Lau K, Li Y, Muthana S, Wang J, Chen X. Chem Commun. 2011; 47:8691.
9. Holmberg, LA.; Guthrie, KA.; Sandmaier, BM. American Chemical Society Symposium Series Chemical Glycobiology. Chen, X.; Halcomb, RL.; Wang, PG., editors. Vol. 990. Oxford University Press; 2008. p. 197
10. Chandrasekaran EV, Xue J, Xia J, Locke RD, Patil SA, Neelamegham S, Matta KL. J Proteome Res. 2012; 11:2609. [PubMed: 22329400]
11. Storr SJ, Royle L, Chapman CJ, Hamid UM, Robertson JF, Murray A, Dwek RA, Rudd PM. Glycobiology. 2008; 18:456. [PubMed: 18332077]
12. Brockhausen I. EMBO Rep. 2006; 7:599. [PubMed: 16741504]
13. Sorensen AL, Reis CA, Tarp MA, Mandel U, Ramachandran K, Sankaranarayanan V, Schwientek T, Graham R, Taylor-Papadimitriou J, Hollingsworth MA, Burchell J, Clausen H. Glycobiology. 2006; 16:96. [PubMed: 16207894]
14. Westerlind U, Hobel A, Gaidzik N, Schmitt E, Kunz H. Angew Chem Int Ed. 2008; 47:7551.
15. Zhu J, Wan Q, Lee D, Yang G, Spassova MK, Ouerfelli O, Ragupathi G, Damani P, Livingston PO, Danishefsky SJ. J Am Chem Soc. 2009; 131:9298. [PubMed: 19518111]
16. Varki A. Proc Natl Acad Sci U S A. 2010; 107(Suppl 2):8939. [PubMed: 20445087]
17. Hedlund M, Tangvoranuntakul P, Takematsu H, Long JM, Housley GD, Kozutsumi Y, Suzuki A, Wynshaw-Boris A, Ryan AF, Gallo RL, Varki N, Varki A. Mol Cell Biol. 2007; 27:4340. [PubMed: 17420276]
18. Malykh YN, Schauer R, Shaw L. Biochimie. 2001; 83:623. [PubMed: 11522391]
19. Tangvoranuntakul P, Gagneux P, Diaz S, Bardor M, Varki N, Varki A, Muchmore E. Proc Natl Acad Sci U S A. 2003; 100:12045. [PubMed: 14523234]
20. Padler-Karavani V, Yu H, Cao H, Chokhawala H, Karp F, Varki N, Chen X, Varki A. Glycobiology. 2008; 18:818. [PubMed: 18669916]
21. Pham T, Gregg CJ, Karp F, Chow R, Padler-Karavani V, Cao H, Chen X, Witztum JL, Varki NM, Varki A. Blood. 2009; 114:5225. [PubMed: 19828701]

22. Padler-Karavani V, Hurtado-Ziola N, Pu M, Yu H, Huang S, Muthana S, Chokhawala HA, Cao H, Secrest P, Friedmann-Morvinski D, Singer O, Ghaderi D, Verma IM, Liu YT, Messer K, Chen X, Varki A, Schwab R. *Cancer Res.* 2011; 71:3352. [PubMed: 21505105]
23. Pan Y, Chefalo P, Nagy N, Harding C, Guo Z. *J Med Chem.* 2005; 48:875. [PubMed: 15689172]
24. Liu T, Guo Z, Yang Q, Sad S, Jennings HJ. *J Biol Chem.* 2000; 275:32832. [PubMed: 10976100]
25. Wang Q, Zhang J, Guo Z. *Bioorg Med Chem.* 2007; 15:7561. [PubMed: 17892942]
26. Wu J, Guo Z. *Bioconjug Chem.* 2006; 17:1537. [PubMed: 17105234]
27. Singh R, Bandyopadhyay D. *Cancer Biol Ther.* 2007; 6:481. [PubMed: 18027437]
28. Huang ZH, Shi L, Ma JW, Sun ZY, Cai H, Chen YX, Zhao YF, Li YM. *J Am Chem Soc.* 2012; 134:8730. [PubMed: 22587010]
29. Bernatchez S, Gilbert M, Blanchard MC, Karwaski MF, Li J, Defrees S, Wakarchuk WW. *Glycobiology.* 2007; 17:1333. [PubMed: 17766267]
30. Yang G, Rich JR, Gilbert M, Wakarchuk WW, Feng Y, Withers SG. *J Am Chem Soc.* 2010; 132:10570. [PubMed: 20662530]
31. Brocke C, Kunz H. *Bioorg Med Chem.* 2002; 10:3085. [PubMed: 12150854]
32. Teo CF, Hwang TS, Chen PH, Hung CH, Gao HS, Chang LS, Lin CH. *Adv Synth Catal.* 2005; 347:967.
33. Yu H, Huang S, Chokhawala H, Sun M, Zheng H, Chen X. *Angew Chem Int Ed.* 2006; 45:3938.
34. Li Y, Yu H, Cao H, Lau K, Muthana S, Tiwari VK, Son B, Chen X. *Appl Microbiol Biotechnol.* 2008; 79:963. [PubMed: 18521592]
35. Sugiarto G, Lau K, Qu J, Li Y, Lim S, Mu S, Ames JB, Fisher AJ, Chen X. *ACS Chem Biol.* 2012; 7:1232. [PubMed: 22583967]
36. Yu H, Thon V, Lau K, Cai L, Chen Y, Mu S, Li Y, Wang PG, Chen X. *Chem Commun.* 2010; 46:7507.
37. Muthana MM, Qu J, Li Y, Zhang L, Yu H, Ding L, Malekan H, Chen X. *Chem Commun.* 2012; 48:2728.
38. Yu H, Chokhawala H, Karpel R, Yu H, Wu B, Zhang J, Zhang Y, Jia Q, Chen X. *J Am Chem Soc.* 2005; 127:17618. [PubMed: 16351087]
39. Thon V, Lau K, Yu H, Tran BK, Chen X. *Glycobiology.* 2011; 21:1206. [PubMed: 21515586]
40. Thon V, Li Y, Yu H, Lau K, Chen X. *Appl Microbiol Biotechnol.* 2012; 94:977. [PubMed: 22075637]
41. Sugiarto G, Lau K, Li Y, Khedri Z, Yu H, Le DT, Chen X. *Mol Biosyst.* 2011; 7:3021. [PubMed: 21858283]
42. Yu H, Yu H, Karpel R, Chen X. *Bioorg Med Chem.* 2004; 12:6427. [PubMed: 15556760]

ATGTTCAAATTTCTATCATCCTGCCGACCTATAACGTGGAACAGTATATCGCTCGCGCA 60
 M F K I S I I L P T Y N V E Q Y I A R A 20
 ATCGAGAGCTGCATCAAACAGACCTTCAAGGACATCGAGATTATCGTGGTGGACGACTGC 120
 I E S C I **K** Q T F K D I E I I V V D D C 40
 GGCAACGATAACTCCATCAACATTGCTAAGGAATACAGCAAGAAGGACAAGCGTATCAAA 180
 G N D N S I N I A K E Y S K K D K R I K 60
 ATCATTACACAACGAAAAGAACATCGGCCTGCTGCGTGCACGTTACGAAGGTGTTAAAGTG 240
 I I H N E K N **I** G L L R A R Y E G V K V 80
 GCGAACAGCCCGTACATCATGTTCTCTGGA_cCCGGACGATTATCTGGAGCTGAACGCGTGC 300
 A N S P Y I M F L D P D D Y L E L N A C 100
 GAAGAGTGCATTAAGATTCTGGACGAGCAGGACGAAGTGGACCTGGTGTCTTCAACGCT 360
 E E C I K I L D E Q D E V D L V F F N A 120
 ATTGTGGAGTCCAACGTGATCAGCTACAAGAAATTCGATTTTAACTCCGGCTTCTACTCC 420
 I V E S N V I S Y K K F D F N S G F Y S 140
 AAAAAGGAGTTCGTGAAAAAGATCATCGCAGAGAAGAACCTGTACTGGACCATGTGGGGC 480
 K K E F V K K I I A **E** K N L Y W T M W G 160
 AAAGTATTCGCAAGAACTGTACCTGGAGGCGTTTCGTTCCCTGCGTCTGGAAGGAT 540
 K L I R K K L Y L E A F A S L R L E K D 180
 GTTAAGATCAACATGGCGGAAGACGTTCTGCTGTACTACCCGATGCTGTCTCAGGCACAG 600
 V K I N M A E D V L L Y Y P M L S Q A Q 200
 AAAATTGCGTACATGAACCTGCAACCTGTACCACTATGTGCCGAACAACAACCTCTATTTGC 660
 K I A Y M N C N L Y H Y V P N N N S I C 220
 AACACCAAAAACGAGGTGGGCGTTAAAAACAACATTCAGGACCTGCAGCTGGTGCTGAAC 720
 N T K N E V **G** V K N N I Q **D** L Q L V L N 240
 TACCTGCGCCAGAACTACATCCTGAACAAGTACTGCAGCGTTCTGTACGTGCTGATTAAG 780
 Y L R Q N Y I L N K Y C S V L Y V L I K 260
 TACCTGCTGTATATTAGATCTACAAAATTAAGCGCACCAAACTGATGGTGACCTGCTG 840
 Y L L Y I Q I Y K I K R T K L M V T L L 280
 GCGAAGATTAACATTCTGACCCTGAAGATCCTGTTCAAGTATAAGAAGTTCTGAAACAG 900
 A K I N I L T L K I L F K Y K K F L K Q 300
TGCTAA 903
C 301

Figure 1. Synthetic gene and protein sequences of full length CjCgtB mutant
 Five mutation sites (N26K, L68I, K151E, L227G, and E234D) are in bold and underlined.
 Truncated C-terminal sequence in CjCgtBΔ30-His₆ mutant is italicized.

Figure 2.

(a) High-performance liquid chromatography (HPLC) chromatograms and (b) representative high-resolution mass spectrometry (HRMS) spectra of purified Tn-MUC1 (compound **2**) as well as T-MUC1 (compound **7**), STn-MUC1 (compounds **3–6** containing Neu5Ac, Neu5Ac9N₃, Neu5Gc, and Neu5AcN₃, respectively), and ST-MUC1 (compounds **8–11** containing Neu5Ac, Neu5Ac9N₃, Neu5Gc, and Neu5AcN₃, respectively) glycopeptides synthesized by one-pot multienzyme (OPME) systems.

Scheme 1.

Solid phase synthesis of Tn-MUC1 glycopeptide APG(GalNAc)STAPPA (**1**) and one-pot multienzyme (OPME) chemoenzymatic synthesis of T-MUC1 glycopeptide as well as STn- and ST-MUC1 glycopeptides containing different sialic acid forms. a) Standard solid phase glycopeptide synthesis using Fmoc-amino acid-OH, HOBt, DIC in DMF. Dabsyl group was introduced using dabsyl chloride in DIPEA/DMF; b) cleavage of glycopeptide from beads using TFA:TIPS:H₂O cocktail for 2 hrs followed by de-*O*-acetylation using NaOH/MeOH (pH was controlled in the range of 8.5–9.5) for 16 hrs; c) one-pot three-enzyme α 2–6-sialylation using EcNanA, NmCSS, and Pd2,6ST in Tris-HCl buffer (100 mM, pH = 8.5) containing MgCl₂ (20 mM); d) one-pot four-enzyme β 1–3-galactosylation using EcGalK, BLUSP, PmPpA, and CjCgtB Δ 30-His₆ mutant in Tris-HCl buffer (100 mM, pH = 7.5) containing (10 mM) and (10 mM); e) one-pot three-enzyme α 2–3-sialylation using MgCl₂MnCl₂ EcNanA, NmCSS, and PmST3 in Tris-HCl buffer (100 mM, pH = 8.5) containing MgCl₂ (20 mM). Enzymes used: EcNanA, *Escherichia coli* sialic acid aldolase; NmCSS, *Neisseria meningitidis* CMP-sialic acid synthetase; Pd2,6ST, *Photobacterium damsela* α 2–6-sialyltransferase; EcGalK, *Escherichia coli* galactokinase; BLUSP, *Bifidobacterium longum* UDP-sugar pyrophosphorylase; PmPpA, *Pasteurella multocida* inorganic pyrophosphatase; CjCgtB Δ 30-His₆ mutant, *Campylobacter jejuni* β 1–3-galactosyltransferase mutant; PmST3, *Pasteurella multocida* α 2–3-sialyltransferase 3.

Scheme 2.

Strategies for optimizing synthetic yields of the one-pot multienzyme (OPME) sialylation reactions by adding excess amount of CTP and removal of CMP byproduct by flash C18 cartridge purification (OPME synthesis of STn-MUC glycopeptide is used as an example).

Scheme 3.

One-pot four-enzyme galactosylation system for the synthesis of T-MUC1 glycopeptide from Tn-MUC1. Enzymes used: EcGalK, *Escherichia coli* galactokinase; BLUSP, *Bifidobacterium longum* UDP-sugar pyrophosphorylase; PmPpA, *Pasteurella multocida* inorganic pyrophosphatase; CjCgtBΔ30-His₆ mutant, *Campylobacter jejuni* β1–3-galactosyltransferase mutant.

Table 1

HPLC retention times (RT) and HRMS data of Tn-, STn-, T-, and ST-MUC1 glycopeptides synthesized.

Glycopeptides	RT (min)	Exact Mass (m/z)	Observed Mass (m/z)
2	15.87	1257.5568 ^a	1257.5649 ^a
3	10.85	1546.6377 ^b	1546.6521 ^b
4	11.42	1571.6442 ^b	1571.6598 ^b
5	10.29	1562.6326 ^b	1562.6472 ^b
6	10.53	1587.6391 ^b	1587.6539 ^b
7	13.57	1419.6097 ^a	1419.6118 ^a
8	8.90	1708.6905 ^b	1708.7031 ^b
9	9.57	1733.6970 ^b	1733.7067 ^b
10	8.71	1724.6855 ^b	1724.6964 ^b
11	9.37	1749.6919 ^b	1749.7059 ^b

The masses given are for ^a[M + H] (positive mode) or ^b[M – H] (negative mode).