UC Davis
UC Davis Previously Published Works

Title

Substrate and Process Engineering for Biocatalytic Synthesis and Facile Purification of Human Milk Oligosaccharides

Permalink
https://escholarship.org/uc/item/26t447sz

Journal
ChemSusChem, 15(9)

ISSN
1864-5631

Authors
Bai, Yuanyuan
Yang, Xiaohong
Yu, Hai
et al.

Publication Date
2022-05-06

DOI
10.1002/cssc.202102539

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

HHS Public Access

Author manuscript

ChemSusChem. Author manuscript; available in PMC 2023 May 06.

Published in final edited form as:

ChemSusChem. 2022 May 06; 15(9): e202102539. doi:10.1002/cssc.202102539.

Substrate and process engineering for biocatalytic synthesis and facile purification of human milk oligosaccharides (HMOs)

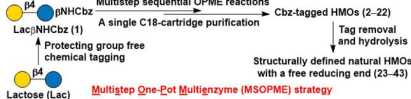
Yuanyuan Bai[a], **Xiaohong Yang**[a], **Hai Yu**[a], **Xi Chen**[a]

[a]Department of Chemistry, University of California, Davis, One Shields Avenue, Davis, California 95616, United States

Abstract

Innovation in process development is essential for applying biocatalysis in industrial and laboratory productions of organic compounds including beneficial carbohydrates such as human milk oligosaccharides (HMOs). HMOs have attracted increasing attention for exploring their potential applications as key ingredients in products that can improve human health. To access HMOs via biocatalysis in an efficient manner, we have developed a combined substrate and process engineering strategy namely the multistep one-pot multienzyme (MSOPME) design. The strategy allows the access to a pure tagged HMO in a single reactor with a single C18-cartridge purification process despite the length of the target. Its efficiency was demonstrated in high-yield (71–91%) one-pot synthesis of twenty tagged HMOs (83–155 mg) including long-chain oligosaccharides with or without fucosylation or sialylation up to nonaoses from a lactoside without the isolation of the intermediate oligosaccharides. Gram-scale synthesis of an important HMO derivative, a tagged lacto-*N*-fucopentaose-I (LNFP-I), succeeded in 84% yield. Tag removal was achieved in high efficiency (94–97%) without the need for column purification to produce the desired natural HMOs with a free reducing end. We envision that the method can be readily adapted for large-scale synthesis and automation to allow quick access to HMOs, other glycans, and glycoconjugates.

Graphical Abstract



Got milk sugar? A highly efficient multistep one-pot multienzyme (MSOPME) strategy is developed for carbohydrate synthesis by integrating biocatalysis substrate and process engineering concepts. A single C18-cartridge purification process is sufficient for any target. The strategy is demonstrated for preparative and gram-scale synthesis of numerous human milk oligosaccharides (HMOs), which are attractive targets for industrial biocatalytic production.

xiichen@ucdavis.edu .

Conflict of Interest

The authors declare no conflict of interest.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Keywords

biocatalysis; carbohydrate; chemoenzymatic synthesis; glycosylation; human milk oligosaccharides (HMOs)

Introduction

Industrial production of chemicals has benefited tremendously and will continue to benefit from the incorporation of biocatalysis into its synthetic processes.[1] Biocatalysis is environmentally friendly, sustainable, highly efficient, and selective for the desired molecular transformation.[2] It has an especially high potential in industrial production of carbohydrates and glycoconjugates that are soluble in aqueous solution, a preferred medium for enzyme-catalyzed reaction.[1, 3] The enzyme stabilization power of high concentration carbohydrate substrates has also been shown,[1] adding another advantage for the production of carbohydrates by biocatalysis. With the growing demands for carbohydrates as pharmaceuticals, nutraceuticals, and ingredients for foods and health beneficial products,[1, 4] developing biocatalytic strategies for their synthesis is becoming increasingly important especially if their natural sources are limited and their chemical synthesis is challenging. Human milk oligosaccharides (HMOs) are among the well suited attractive carbohydrate targets for industrial biocatalytic production.[1, 5]

In addition to lactose, lipids, and proteins, human milk oligosaccharides (HMOs) are a major component of human milk which is the sole food for breast-fed infants in the first several months of their lives.[6] HMOs are not digested directly by infants but they contribute significantly to the health benefits of breastfeeding.[7] The roles of HMOs as prebiotics, decoys to pathogenic bacteria, bacteriostatic agents, gut epithelial cell maturation and immune regulators, nutrients for brain development and their current and potential applications as infant formula supplements, nutraceuticals, and therapeutics are being continuously explored.[5–8]

More than 150 HMO structures are known.[5, 6c, 9] They are extended from lactose, a free reducing end disaccharide, with β1–3/6-linked *N*-acetyl-D-glucosamine (GlcNAc) and β1–3/4-linked D-galactose (Gal), and with or without α2–3/6-linked *N*-acetyl-D-neuraminic acid (Neu5Ac) and/or α1–2/3/4-linked L-fucose (Fuc).[6c, 10]

The progress in exploring the biological functions of HMOs and their applications is highly dependent on the access to structurally defined HMOs in sufficient quantities.[5, 6c, 11] However, only some of the structurally characterized HMOs have been synthesized by chemical, enzymatic and chemoenzymatic, and fermentation approaches.[5, 6c, 9a, 10, 12] Long-chain HMOs and their sialylated and/or fucosylated derivatives are challenging targets for chemical synthesis or fermentation but are well suited for biocatalytic production using purified or partially purified enzymes. With the identification and characterization of an increasing number of glycosyltransferases and the advances in the development of enzymatic and chemoenzymatic methods, the number of HMOs that have been produced in preparative and multigram scales by biocatalysis is increasing. Product purification remains

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

a bottleneck process to quick access HMOs in even multigram quantities. Continuous innovation is necessary to overcome the challenges for large-scale production of HMOs.

We reported previously the development of a carboxybenzyl (Cbz) group-based glycosyltransferase acceptor substrate tagging strategy which allows facile purification of the more complex glycoside products from the reaction mixtures of one-pot multienzyme (OPME) glycosylation systems with a single C18-cartridge-based process.[13] The Cbz group is easy to install from an inexpensive commercially available material in a protection group-free fashion[13a] and is readily removable from the products[13b] to form oligosaccharides with a primary amine-terminated linker that is ready for further conjugation[13b] or with a free reducing end.[14] We report herein the development of a process engineering strategy that can be combined with the substrate-tagging (or substrate engineering) strategy and OPME glycosylation systems[15] for facile chemoenzymatic synthesis and purification of HMOs including those with long chain structures and with or without L-fucose or Neu5Ac. Examples for preparative-scale synthesis of HMOs up to nonaoses are shown and gram-scale synthesis is demonstrated for lacto-*N*-fucopentaose-I (LNFP-I). We envision that the method can be readily adapted for automation and is suitable for introducing isotope-labeled or modified monosaccharides to HMOs, other glycans, glycoconjugates, and their derivatives.

Results and Discussion

HMO targets

Twenty-one tagged HMOs (**2–22**) including long-chain oligosaccharides up to nonaoses with or without L-fucose or Neu5Ac (Figure 1), and their counterparts with a free reducing end after the tag is removed (**23–43**) were chosen as the targets for the synthesis.

Overall design of the substrate and process engineering strategies for efficient biocatalytic synthesis and facile purification of HMOs

As shown in Scheme 1, the overall design for the synthesis of target HMOs involves a chemical process for the formation of carboxybenzyl (Cbz) protected β-lactosylamine (LacβNHCbz, **1**) from inexpensive lactose (Lac) via a protecting group free tagging approach.[13a] Glycosyltransferase-based sequential one-pot multienzyme (OPME) systems are then used to extend LacβNHCbz (**1**) to form the desired glycosidic linkages, in a regio- and stereospecific manner, to obtain Cbz-tagged HMOs with *in situ* generation of the corresponding glycosyltransferase sugar nucleotide donor from a simple and inexpensive monosaccharide.[15] Deactivation of enzymes after the completion of each OPME glycosylation step by incubating the reaction mixture in a boiling water bath prevents uncontrolled glycosylation processes that may form multiple oligosaccharide products in the downstream reaction steps. Using this multistep one-pot multienzyme (MSOPME) design, the desired tagged HMOs, even those with long-chain structures, can be obtained in high yields in a true one-pot fashion without the need of purifying the intermediate oligosaccharides. A final step single C18-cartridge purification procedure is efficient to provide the pure product containing the hydrophobic Cbz tag. The Cbz tag is easily removed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by catalytic hydrogenation followed by spontaneous hydrolysis in water to provide the desired HMOs with a free reducing end.

Glycosyltransferase acceptor substrate engineering by Cbz tagging

All common HMOs are extended from lactose. Lactose is a low-cost starting material that can be used directly for enzymatic synthesis of HMOs by reactions catalyzed by glycosyltransferases[12d, 12f, 12g, 16] or mutant glycosidases.[17] Nevertheless, it is a disaccharide with a free reducing end and its reducing end glucose residue forms a mixture of α and β-anomers in its six- and five-membered ring structures as well as a small percentage of the open-chain structure in aqueous solutions.[18] This property leads to complications in HMO product characterization and purification. To simplify the reaction monitoring and product purification processes, carboxybenzyl (Cbz) protected β-lactosylamine (LacβNHCbz, **1**) was designed and synthesized.[13] The Cbz tag is an ultraviolet (UV)-active hydrophobic tag that is easy to install from a commercially available material. It is also easy to remove by simple catalytic hydrogenation[13b] and hydrolysis[14] without the need for chromatography purification processes. As shown in Scheme 2, LacβNHCbz (**1**) was synthesized from lactose in two steps as we reported previously.[13a] Briefly, lactose was incubated with ammonium bicarbonate in ammonium hydroxide to form lactosylamine (LacβNH$_2$). The solvent was removed and the dried residue was dissolved in anhydrous methanol and then coupled with benzyl chloroformate (CbzCl) to obtain LacβNHCbz (**1**) which was readily purified by a C18 cartridge. The procedure was robust and LacβNHCbz (**1**) was consistently produced in high yields (75–80%) in multigram (3–5 g)-scale reactions.

One-pot multienzyme (OPME) systems and selection of biocatalysts

HMOs are complex glycans that are extended from lactose with *N*-acetyl-D-glucosamine (GlcNAc), D-galactose (Gal), L-fucose (Fuc), and/or *N*-acetyl-D-neuraminic acid (Neu5Ac) units. One-pot multienzyme (OPME) systems,[15, 19] each containing a suitable glycosyltransferase and a monosaccharide sugar activation (SA) enzymes and reagents, are well suited to form the target HMOs with *in situ* generation of glycosyltransferase sugar nucleotide donor from inexpensive monosaccharide and nucleoside triphosphates such as adenosine 5'-triphosphate (ATP), uridine 5'-triphosphate (UTP), and/or cytidine 5'-triphosphate (CTP). As shown in Scheme 3, four SA components are needed for the activation of GlcNAc (SA1), Gal (SA2), Fuc (SA3), and Neu5Ac (SA4), respectively. Nine glycosyltransferases are used together with these SAs in the OPME systems for the synthesis of target HMOs.

The GlcNAc activation (SA1) component involves three enzymes including *Bifidobacterium longum* strain ATCC55813 *N*-acetylhexosamine-1-kinase (BLNahK),[20] *Pasteurella multocida N*-acetylglucosamine uridylyltransferase (PmGlmU),[21] and *Pasteurella multocida* inorganic pyrophosphatase (PmPpA).[21–22] BLNahK is highly active in catalyzing the direct phosphorylation of GlcNAc (k_{cat}/K_m = 18.3 s^{-1} mM^{-1}) using ATP for the formation of GlcNAc-1-phosphate.[20] PmGlmU[21] catalyzes high efficient synthesis of UDP-GlcNAc from GlcNAc-1-P and UTP. PmPpA[21–22] is used to break down the pyrophosphate formed in the PmGlmU reaction to inorganic phosphate to shift the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

equilibrium of the coupled enzymatic reactions towards the formation of UDP-GlcNAc. Two β1–3-*N*-acetylglucosaminyltransferases (β3GlcNAcTs)[23] from *Neisseria meningitidis* (NmLgtA) and *Helicobacter pylori* (Hpβ3GlcNAcT) are selected for OPME1a and OPME1b systems, respectively. NmLgtA (10 mg L^{-1}, 10 mg purified protein from 5.6–7.0 g wet cells collected from one liter of *E. coli* culture) with a better expression level than Hpβ3GlcNAcT (1 mg L^{-1})[23] is used to add β1–3-linked GlcNAc to LacβNHCbz (**1**), a short acceptor substrate, while Hpβ3GlcNAcT is used to add β1–3-linked GlcNAc to longer acceptor substrates which were less efficient for NmLgtA-catalyzed reactions.[24]

The Gal activation (SA2) component involves three enzymes including *Streptococcus pneumoniae* TIGR4 galactokinase (SpGalK),[25] *Bifidobacterium longum* UDP-sugar pyrophosphorylase (BLUSP),[26] and PmPpA. Two bacterial galactosyltransferases β1–4-galactosyltransferase (NmLgtB)[22–23] and *Chromobacterium violaceum* β1–3-galactosyltransferase (Cvβ3GalT)[12g] are chosen for OPME2a and OPME2b systems for the addition of a Gal with β1–4- and β1–3-linkages, respectively, to the GlcNAc-terminated glycans. Cvβ3GalT is a highly efficient β1–3-galactosyltransferase which was used previously for synthesizing LNT successfully in multigram-scale.[12g] NmLgtB is a β1–4-galactosyltransferase that has been used for efficient synthesis of LacNAc-containing structures.[27]

The L-fucose activation (SA3) component contains two enzymes including a *Bacteroides fragilis* NCTC9343 bifunctional L-fucokinase/GDP-fucose pyrophosphorylase (BfFKP)[28] and PmPpA. The bifunctional enzyme BfFKP[28] contains a C-terminal fucokinase domain and an N-terminal GDP-fucose pyrophosphorylase domain. Three bacterial fucosyltransferases including *Helicobacter mustelae* α1–2-fucosyltransferase (Hm2FT),[29] *Thermosynechococcus elongatus* α1–2-fucosyltransferase (Te2FT),[12f] and *Helicobacter pylori* α1–3/4 fucosyltransferase (Hp3/4FT)[30] are chosen for OPME3a, OPME3b, and OPME3c systems, respectively, for the addition of an α1–2-, or one or more α1–3/4-linked L-fucose residues. Hm2FT is able to catalyze the transfer of α1–2-linked L-fucose to both β1–3 and β1–4-linked galactosides while Te2FT[12f] has high selectivity towards α1–2-fucosylation of β1–3-linked galactosides. Te2FT is selected for the synthesis of LNFP-I which is an α1–2-fucosylated β1–3-galactoside. Hp3/4FT[30] can catalyze the transfer of L-fucose to form α1–3 and α1–4-fucosylated products and is a powerful enzyme for α1–3/4-fucosylation of the GlcNAc and Glc residues in HMOs.

Sialic acid can be activated by *Neisseria meningitidis* CMP-sialic acid synthetase (NmCSS)[31] (SA4) and transferred by *Pasteurella multocida* α2–3-sialyltransferase 3 (PmST3)[32] or *Photobacterium damselae* α2–6-sialyltransferase (Pd2,6ST_A200Y/S232Y),[33] respectively, in OPME4a and OPME4b systems. PmST3[32] prefers β1–4- over β1–3-linked galactoside acceptors and is chosen for sialylating β1–4-linked terminal galactose. Pd2,6ST_A200Y/S232Y[33] is highly selective in sialylating the terminal galactose in HMOs with an α2–6-linkage.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The multistep one-pot multienzyme (MSOPME) strategy and its application in synthesizing desired NHCbz-tagged HMOs

To assemble the desired HMOs by glycosyltransferase-based enzymatic or chemoenzymatic synthesis with or without *in situ* generation of the corresponding sugar nucleotide donor, a common practice has been to perform a product purification process after every glycosyltransferase-catalyzed reaction and the purified product is used for another glycosyltransferase-catalyzed reaction for the formation of a more complex product.[10, 12a–g, 34] Numerous methods have been developed to facilitate the product purification of these enzyme-catalyzed glycosylation reactions, including acceptor solid phase immobilization,[35] as well as acceptor tagging with a thermosensitive polymer,[36] a hydrophobic,[10, 12c, 14, 37] or an anionic[34] tail, etc. To synthesize a complex oligosaccharide target that requires reactions involving multiple glycosyltransferases, precipitation or solid phase extraction has been commonly used to purify every oligosaccharide intermediate before it is used as the acceptor substrate for the next glycosylation reaction. These processes allow the access to intermediate oligosaccharides which themselves can be important probes for protein-binding studies or enzyme substrate specificity studies. Nevertheless, multiple purification processes are needed to obtain a long-chain oligosaccharide target. The practice is not a one-pot design and comes with the expense of increased time, labor, solvents, reagents, process complexity, and cost needed for the overall synthetic process.[38]

To access structurally complex HMO targets involving the actions of multiple glycosyltransferases in a time and process efficient manner, we envision that purification of the oligosaccharide product after each intermediate glycosyltransferase-catalyzed reaction is not necessary and can be bypassed. This will allow the access to the desired compound from LacβNHCbz (**1**) in one pot with a single purification process for the final product despite its length and structure complexity. Such a process will greatly increase the time and pot economy[38] for chemoenzymatic total synthesis of HMOs and can be readily adaptable for automation. It is a step forward towards practical industrial biocatalytic production of HMOs using purified or partially purified enzymes.

Elimination of the product purification processes for intermediate reactions will work well if these intermediate products are selective substrates for the designed downstream reactions. This is, however, not the case for the synthesis of some long-chain HMO targets. For example, pentasaccharide GlcNAc-LNnTβNHCbz (**4**) can be formed from LacβNHCbz (**1**) by NmLgtA-containing OPME1a, followed by NmLgtB-containing OPME2a, and then Hpβ3GlcNAc-containing OPME1b. If the reactions were carried out in one-pot in a step-wise OPME fashion without the purification of oligosaccharide intermediates, pentasaccharide GlcNAc-LNnTβNHCbz (**4**) formed would be a suitable acceptor substrate for the NmLgtB in OPME2a for the formation of a hexasaccharide which will be an acceptor for the Hpβ3GlcNAcT in OPME1b for the formation of a heptasaccharide. Longer-chain oligosaccharides containing a poly-LacNAc repeat can also be formed.[24] The formation of multiple oligosaccharide products in such uncontrolled glycosylation reactions would lower the yield for the synthesis of the target, complicate the product purification process, and is undesirable.

Author Manuscript

The complications of uncontrolled glycosylation in the synthesis of some HMO targets can be prevented by including a simple enzyme deactivation step after each OPME glycosylation is completed as indicated by the disappearance of the acceptor substrate by mass spectrometry analyses and/or thin-layer chromatography (TLC) assays. The enzyme deactivation can be achieved readily by incubating the reaction mixture in a boiling water bath for 5 minutes. After the reaction mixture is cooled down, it can then be used as the acceptor substrate for the glycosyltransferase in the next OPME reaction without any workup or purification processes. Such a simple process engineering strategy will allow the target HMO be synthesized in one-pot using multiple OPME reactions carried out sequentially in multiple steps without the need for the purification of intermediate oligosaccharides. We name this time and pot-efficient biocatalytic process the multistep one-pot multienzyme (MSOPME) strategy.

Such a strategy worked very well for synthesizing HMOs including those with a long-chain structure. As shown in Table 1, trisaccharide LNT-IIβNHCbz (**2**) was synthesized using OPME1a and purified in a 91% yield from 100 mg LacβNHCbz (**1**). Two sequential OPME reactions followed by a single C18-cartridge purification produced 152 mg LNnTβNHCbz (**3**) or 155 mg LNTβNHCbz (**7**) from 100 mg of LacβNHCbz (**1**) in 86–88% yield. Three sequential OPME reactions with a single C18-cartridge purification produced 83–91 mg (80–83% yield) of each of the three pure pentasaccharides GlcNAc-LNnTβNHCbz (**4**), LNnFP-VβNHCbz (**5**), and LNFP-VβNHCbz (**9**), as well as 94 mg pure hexasaccharide LNnDFH-IIβNHCbz (**6**) in 80% yield from 50 mg LacβNHCbz (**1**). Four sequential OPME reactions with a single C18-cartridge purification produced 95–101 mg of each of the three pure hexasaccharides LNDFH-IIβNHCbz (**10**), *p*LNnHβNHCbz (**11**), and *p*LNHβNHCbz (**17**) in 79–80% yield from 50 mg LacβNHCbz (**1**). Similar processes also worked well for more complex longer chain glycans. Five-step sequential OPME reactions with a single C18-cartridge purification produced 105–123 mg of each of the pure heptasaccharides including GlcNAc-*p*LNnHβNHCbz (**12**), F-*p*LNnH-IβNHCbz (**13**), Neu5Acα2–3*p*LNnHβNHCbz (**15**), Neu5Acα2–6*p*LNnHβNHCbz (**16**), and F-*p*LNH-IβNHCbz (**18**) in 74–78% yield, as well as 124–131 mg of each of the two pure nonasaccharides TF-*p*LNnHβNHCbz (**14**) and TF-*p*LNH-IIIβNHCbz (**19**) in 72–76% yield from 50 mg of LacβNHCbz (**1**). Six-step sequential OPME reactions with a single C18-cartridge purification produced 117–119 mg of each of the two pure octasaccharides *p*LNnOβNHCbz (**20**) and *p*LNOβNHCbz (**21**) in 71–72% yield from 50 mg of LacβNHCbz (**1**).

Purified octasaccharide *p*LNOβNHCbz (**21**) was used as the acceptor substrate for Hm2FT in OPME3a for the synthesis of Fucα1–2*p*LNOβNHCbz (**22**) which was obtained as a pure compound in 92% yield after a C18-cartridge purification.

It is worth to point out that mono-fucosylated pentasaccharide LNnFP-VβNHCbz (**5**) was formed from LacβNHCbz (**1**) by adding a β1–3-linked GlcNAc using OPME1a, followed by α1–3-fucosylation with OPME3c and β1–4-galactosylation with OPME2a. Altering the order of the latter two OPME steps similar to that reported previously[12d] led to the formation of difucosylated hexasaccharide LNnDFH-IIβNHCbz (**6**) (OPME1a, OPME2a, and OPME3c) when 2.86 equivalents of L-fucose and 3.05 equivalents of ATP

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and UTP were used in OPME3c. Mono-fucosylated pentasaccharide LNFP-VβNHCbz (**9**) was formed similarly by extending LacβNHCbz (**1**) with a β1–3-linked GlcNAc using OPME1a, followed by α1–3-fucosylation with OPME3c and β1–3-galactosylation with OPME2b. On the other hand, to ensure the installation of two L-fucose residues to form the difucosylated hexasaccharide LNDFH-IIβNHCbz (**10**), a four-step OPME process was used by adding an additional α1–3/4-fucosylation OPME3c to the three-step OPME preparation of LNFP-VβNHCbz (**9**). Installation of all three L-fucose residues in trifucosylated nonasaccharides TF-*p*LNnHβNHCbz (**14**) and TF-*p*LNH-IIIβNHCbz (**19**) was achieved in one-step OPME3c α1–3/4-fucosylation process from *p*LNnHβNHCbz (**11**) and *p*LNHβNHCbz (**17**) intermediates, respectively, formed from four-step OPME glycosylation of LacβNHCbz (**1**).

If the products of the intermediate OPME reactions are selective substrates for the glycosyltransferases in the designed downstream OPME reactions without the complication of uncontrolled glycosylation, the MSOPME can be carried out without the heat inactivation of the enzymes after the intermediate OPME steps. Such an example was demonstrated for gram-scale synthesis of LNFP-IβNHCbz (**8**). The LNFP-IβNHCbz (**8**) was successfully synthesized in 1.74 grams in 84% yield with three consecutive OPME steps without heat inactivation of the reaction mixture after intermediate OPME steps.

Tag removal and hydrolysis of glycosylamines

The Cbz tag in the obtained glycosides (**2–22**) was readily removed by catalytic hydrogenation using palladium and charcoal in methanol and H_2O. The glycosylamine formed was partially converted to the target HMO with a free reducing end during the hydrogenation process. After the palladium and charcoal were removed by filtering the mixture using a 0.45 um syringe filter, the solvent was removed by rotavap and the complete conversion of the glycosylamine to the desired HMO with a free reducing end was achieved by dissolving the residue in water and incubating the mixture at 37 °C. As shown in Scheme 4, the NHCbz group on the glycosides (**2–22**) was converted to toluene, CO_2, and NH_3 which were removed by rotary evaporation and lyophilization without purification. Excellent yields (94–97%) were achieved.

Conclusion

In conclusion, we have developed an efficient process engineering strategy that has been combined with a substrate engineering strategy and OPME systems for multistep one-pot multienzyme (MSOPME) synthesis of tagged glycan targets from a simple lactoside in one pot without the purification of intermediate oligosaccharides. A single C18-cartridge-based purification of the final tagged product followed by catalytic hydrogenation and spontaneous hydrolysis led to the formation of the desired naturally occurring HMO with a free reducing end. The one-pot strategy has been demonstrated for high-yield facile synthesis of twenty-one pure HMOs including those with or without sialic acid or L-fucose and with a size up to nonasaccharide using only a single C18-cartridge purification process for any given target. Gram-scale synthesis has also been demonstrated for an important HMO, lacto-*N*-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

fucopentaose-I (LNFP-I). The method can be readily adapted for automation to allow quick access of HMOs as well as other glycans and glycoconjugates.

Experimental section

Material

All chemicals were obtained from commercial suppliers and used without further purification. ^{1}H NMR (600 or 800 MHz) and ^{13}C NMR (150 or 200 MHz) spectra were recorded on a Bruker Avance-600 Spectrometer or a Avance-800 Spectrometer. High-resolution electrospray ionization (ESI) mass spectra were recorded using a Thermo Scientific Q Exactive HF Orbitrap Mass Spectrometer at the Mass Spectrometry Facilities in The University of California, Davis. Thin-layer chromatography (TLC, Sorbent Technologies) was performed on silica gel plates using anisaldehyde sugar stain for detection.

Bacterial strains, plasmids, and reagents

Escherichia coli DH5α and BL21(DE3) chemically competent cells were from Invitrogen (Carlsbad, CA, USA). AccuPrep® PCR/Gel purification kit was from BIONEER Corporation. LB Broth, isopropyl-1-thio-β-D-galactopyranoside (IPTG), ampicillin, GeneJET plasmid spin kit, and FastDigest DpnI were from Fisher Scientific (Chicago, IL, USA). Phusion® HF DNA polymerase was from New England Biolabs, Inc. (Ipswich, MA, USA).

Enzyme expression

Recombinant enzymes were expressed and purified as described previously for NmLgtA,[23] Hpβ3GlcNAcT,[23] NmLgtB,[22] Cvβ3GalT,[12g] PmST3,[32] Pd2,6ST_A200Y/S232Y,[33] Hm2FT,[29] Hp3/4FT,[30] Te2FT,[12f] SpGalK,[25] BLUSP,[26] BLNahK,[20] PmGlmU,[21] NmCSS,[31] BfFKP,[28] PmPpA.[22] Briefly, *E. coli* BL21 (DE3) strains harboring the recombinant plasmid with the target gene was cultured in 50 mL Luria-Bertani (LB) media (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, and 10 g L^{-1} NaCl) containing 0.1 mg mL^{-1} ampicillin with rapid shaking (220 rpm) at 37 °C for overnight. Then 15 mL of the overnight cell culture was transferred into 1 L of LB media containing 0.1 mg mL^{-1} ampicillin and incubated at 37 °C. When the $OD_{600\ nm}$ of the cell culture reached 0.6–0.8, isopropyl-1-thio-β-D-galactopyranoside (IPTG, 0.1 mM) was added to induce the expression of the recombinant enzyme. The culture was then incubated at 20 °C with shaking (220 rpm) for 20 h. Cells were collected by centrifugation at 4392 × g for 50 min at 4 °C. The cell pellet was re-suspended with lysis buffer (100 mM Tris-HCl buffer, pH 8.0, containing 0.1% Triton X-100). The cells were lysed by sonication with the following conditions: amplitude at 65%, 2 s pulse on and 3 s pulse off for 120 cycles. Cell lysate was obtained by centrifugation at 9016 × g for 1 hour at 4 °C and the collection of the supernatant.

Enzyme purification

Purification was carried out using 5 mL Bio-Scale Mini Profinity IMAC Cartridges in a Bio-Rad NGC 100 Medium-Pressure Chromatography System with a flow rate of 5 mL min^{-1}. The supernatant was loaded to the column pre-equilibrated with 10 column volumes

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of a binding buffer (5 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl buffer, pH 7.5). The column was washed with 10 column volumes of the binding buffer followed by 10 column volumes of a washing buffer (10 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl buffer, pH 7.5). The target protein was eluted with 10 column volumes of a elution buffer (200 mM imidazole, 0.5 M NaCl, 50 mM Tris-HCl buffer, pH 7.5). Fractions containing the target protein were collected. PmST3, SpGalK, BLUSP, BLNahK, PmGlmU, NmCSS and PmPpA were dialyzed against a dialysis buffer (10% glycerol, 20 mM Tris-HCl buffer, pH 7.5) and stored at –20 °C. NmLgtA, Hpβ3GlcNAcT, NmLgtB, Cvβ3GalT, Pd2,6ST_A200Y/S232Y, Hm2FT, Hp3/4FT, Te2FT and BfFKP were added with 10% glycerol without dialysis and stored at –20 °C.

Mutagenesis to obtain Pd2,6ST_A200Y/S232Y

Plasmid in pET15b vector for expressing an N-terminal His_6-tagged truncated Pd2,6ST (16–497 aa)[39] was used as the template for polymerase chain reactions (PCRs) to introduce mutation sites following the QuikChange™ site-directed mutagenesis method developed by Stratagene (La Jolla, CA, USA). The single mutant Pd2,6ST_A200Y was obtained first which was then used as the template to obtain the double mutant Pd2,6ST_A200Y/S232Y. Primers used were: forward 5'-TATACAAATACA<u>TAT</u>CATGTTTTTAATAATTTACCACCT-3', reverse 5'- ATTAAAAACATG<u>ATA</u>TGTATTTGTATAGAACTCTATATT-3' for generating the A200Y mutation and the mutation sites are underlined; forward 5'-GATGATGGTTCT<u>TAT</u>GAATATGTAAGTTTATATCAATGG-3' and reverse 5'-ACTTACATATTC<u>ATA</u>AGAACCATCATCATACAAACTAAT-3' for generating the S232Y mutation and the mutation sites are underlined. Polymerase chain reactions (PCRs) were each performed in a 50 μL reaction mixture containing 5 ng of template DNA, 1 μM for each of the forward and reverse primers, 5 μL of 10 × Phusion® HF buffer, 1 mM dNTP mixture, and 5 units (1 μL) of Phusion® HF DNA polymerase. The reaction mixtures were subjected to 30 cycles of amplifications with an annealing temperature of 55 °C. The resulting PCR products were digested with DpnI, purified, and transformed into chemically competent *E. coli* DH5α cells. The plasmids of selected colonies were purified and sent to Genewiz (South Plainfield, NJ, USA) for sequencing. Positive plasmids were transformed into BL21 (DE3) chemically competent cells. Selected clones were grown for protein expression, purification, characterization, and application in chemoenzymatic synthesis.

Multigram-scale synthesis and purification of LacβNHCbz (1) from lactose

This was carried out similarly to that was reported previously.[13a] Briefly, lactose (5 g, 14.62 mmol) and ammonium bicarbonate (1.3 g, 16.45 mmol) was dissolved in ammonium hydroxide (25 mL) and the solution was incubated at 45–50 °C for 24 hours to form lactosylamine (LacβNH_2). Solvent was removed *in vacuo* and the residue was dried under vacuum for 3–4 h. To obtain LacβNHCbz, LacβNH_2 (5 g, 14.62 mmol) was dissolved in MeOH (650 mL) and benzyl chloroformate (CbzCl, 9.5 g, 55.69 mmol) was added to the reaction mixture in a 1 L round bottom flask submerged in an ice-water bath. The pH of the mixture was adjusted and kept at 8.0–10.0 by adding *N,N*-diisopropylethylamine. The reaction mixture was then stirred at room temperature for 22 hours. LacβNHCbz was purified by ODS-SM column (140 g, 50 μm, 120 Å, Yamazen) on a CombiFlash® Rf 200i

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

system. Consistently, 75–80% yields were obtained for 3–5 g-scale syntheses. The product formation was monitored by thin-layer chromatography (TLC) assays and by high-resolution mass spectrometry (HRMS) analyses. Ethyl acetate (EtOAc):methanol (MeOH):H_2O = 5:1.6:1 (by volume) was used as the developing solvent for TLC assays. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{20}H_{29}NNaO_{12}$ 498.1587; found 498.1585.

Single C18-cartridge purification process for MSOPME-synthesized βNHCbz-tagged HMOs (2–22)

After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes. The mixture was then cooled down to room temperature and centrifuged at 9016 × g for 30 min at 4 °C. The supernatant was collected. The precipitate was washed twice, each time with H_2O (3 mL), and the supernatants were combined. The combined supernatant was concentrated by rotavap to reduce the volume to about 3–5 mL which was purified by passing through a ODS-SM column (51 g, 50 μm, 120 Å, Yamazen) pre-equilibrated with three column volumes of mobile phase A (water) on a CombiFlash® Rf 200i system and monitored at 214 nm. The product was eluted with a mixed solvent of acetonitrile and water with a flow rate of 30 mL/min. The eluting program used was the following: Mobile phase A: water (v/v); Mobile phase B: acetonitrile (v/v); 0% B for 8 min followed by gradient 0% to 60% B over 25 min, gradient 60% to 100% B over 3 min, 100% B for 2 min, then 100% to 80% B over 2 min.

Preparative-scale synthesis of GlcNAcβ3Galβ4GlcβNHCbz (LNT-IIβNHCbz) (2) by OPME1a

LacβNHCbz (**1**, 100 mg, 0.21 mmol), GlcNAc (0.32 mmol), ATP (0.32 mmol), UTP (0.32 mmol) were dissolved in water in a 50 mL centrifuge tube containing Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM). After the addition of BLNahK (1.5 mg), PmGlmU (1.2 mg), NmLgtA (1.8 mg), and PmPpA (1 mg), water was added to bring the final volume to 21 mL and the concentration of LacβNHCbz (**1**) was 10 mM. The reaction mixture was incubated at 30 °C with agitation at 180 rpm in an incubator shaker for overnight. The product formation was monitored by thin-layer chromatography (TLC) assays and by high-resolution mass spectrometry (HRMS) analyses. Ethyl acetate (EtOAc):methanol (MeOH):H_2O = 5:1.6:1 (by volume) was used as the developing solvent for TLC assays. After the reaction was completed (22 h), the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes. The reaction mixture was then cooled down and the product was purified by following the single C18-cartridge purification procedures described above. LNT-IIβNHCbz (**2**) was obtained as a white powder (130 mg, 91% yield in 1 step from 100 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.46–7.38 (m, 5H), 5.17 (t, J = 9.4 Hz, 2H), 4.81 (d, J = 9.3 Hz, 1H), 4.67 (d, J = 8.5 Hz, 1H), 4.43 (d, J = 7.9 Hz, 1H), 4.13 (d, J = 3.3 Hz, 1H), 3.93–3.84 (m, 2H), 3.80–3.52 (m, 12H), 3.49–3.34 (m, 3H), 2.02 (s, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.92, 158.08, 135.99, 128.76, 128.46, 127.80, 102.82, 102.81, 81.85, 81.62, 77.67, 76.41, 76.11, 75.62, 74.99, 74.85, 73.51, 71.37, 69.98, 69.65, 69.20, 68.33, 67.38, 60.94, 60.44, 59.83, 55.62, 22.12. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{28}H_{42}N_2NaO_{17}$ 701.2381; found 701.2383.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

MSOPME preparative-scale synthesis of Galβ4GlcNAcβ3Galβ4GlcβNHCbz (LNnTβNHCbz) (3)

The same procedure was carried out as described above for the preparation of LNT-IIβNHCbz (**2**) from LacβNHCbz (**1**). After confirming the completion of the reaction for the formation of LNT-IIβNHCbz (**2**), the reaction mixture was incubated in a boiling water bath for 5 min and was then cooled down to room temperature. To the reaction mixture in the same tube, Gal (0.32 mmol), ATP (0.32 mmol), and UTP (0.32 mmol) were added and the *p*H of the reaction was adjusted to 7.5 by adding 4 M NaOH. Enzymes including SpGalK (2.5 mg), BLUSP (2 mg), NmLgtB (3 mg), and PmPpA (1 mg) were then added without increasing the volume of the reaction mixture significantly. The reaction mixture was incubated at 30 °C in an incubator shaker with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed (14 h), the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes and then cooled down to room temperature before the product was purified. LNnTβNHCbz (**3**) was obtained as a white powder (152 mg, 86% yield for two steps from 100 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.45–7.38 (m, 5H), 5.17 (t, J = 9.3 Hz, 2H), 4.81 (s, 1H), 4.69 (d, J = 8.4 Hz, 1H), 4.46 (d, J = 7.9 Hz, 1H), 4.43 (d, J = 7.9 Hz, 1H), 4.14 (d, J = 3.3 Hz, 1H), 3.96–3.87 (m, 3H), 3.86–3.61 (m, 16H), 3.60–3.50 (m, 3H), 3.43–3.33 (m, 1H), 2.02 (s, 3H).^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.94, 158.15, 136.05, 128.83, 128.53, 127.87, 102.89, 102.78, 82.01, 81.70, 78.19, 77.74, 76.48, 76.18, 75.38, 75.06, 74.91, 74.59, 72.53, 72.20, 71.44, 71.00, 70.01, 69.27, 68.58, 68.39, 67.45, 61.06, 61.01, 59.89, 55.22, 22.21. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{34}H_{52}N_2NaO_{22}$ 863.2909; found 863.2896.

MSOPME preparative-scale synthesis of GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (GlcNAc-LNnTβNHCbz) (4)

LacβNHCbz (**1**, 50 mg, 0.105 mmol), GlcNAc (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were dissolved in water in a 50 mL centrifuge tube containing Tris-HCl buffer (200 mM, pH 7.5) and $MgCl_2$ (40 mM). After the addition of BLNahK (1.2 mg), PmGlmU (1 mg), NmLgtA (1.5 mg), and PmPpA (0.5 mg), water was added to bring the final volume to 5 mL, resulting in a solution containing 20 mM LacβNHCbz (**1**). The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. Reaction progress was monitored by TLC assays and by HRMS analyses similar to that described above for LNT-IIβNHCbz (**2**). After incubating the reaction mixture in a boiling water bath for 5 min and then cooled down, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), NmLgtB (1.8 mg), and PmPpA (0.5 mg) were then added and the reaction mixture was incubated at 30 °C for overnight with agitation at 180 rpm. Reaction progress was monitored by TLC assays and by HRMS analyses. Once the reaction was completed (14 h), the mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, GlcNAc (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BLNahK (1 mg), PmGlmU (0.8 mg), Hpβ3GlcNAcT

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(1.5 mg), and PmPpA (0.5 mg) were then added. The final concentration of the acceptor in the reaction mixture was kept at around 10 mM and the final volume of the reaction mixture was kept at around 10.5 mL by adding Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM). The reaction mixture was incubated at 30 °C for overnight with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.4:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed (16 h), the reaction mixture was incubated in a boiling water bath for 5 min, cooled down, and purified. GlcNAc-LNnTβNHCbz (**4**) was obtained as a white powder (91 mg, 83% yield for three steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.39 (m, 5H), 5.18 (d, J = 3.5 Hz, 2H), 4.82 (d, J = 9.2 Hz, 1H), 4.69 (dd, J = 11.7, 8.4 Hz, 2H), 4.47 (d, J = 7.9 Hz, 1H), 4.44 (d, J = 7.9 Hz, 1H), 4.15 (t, J = 3.0 Hz, 2H), 3.97–3.87 (m, 3H), 3.85–3.65 (m, 18H), 3.58 (ddt, J = 11.8, 10.4, 5.9 Hz, 4H), 3.51–3.40 (m, 3H), 2.04 (s, 3H), 2.03 (s, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.93, 174.87, 158.08, 135.99, 128.77, 128.47, 128.41, 127.81, 127.67, 102.87, 102.83, 102.71, 81.97, 81.94, 81.62, 78.18, 77.71, 76.13, 75.64, 75.01, 74.86, 74.84, 74.53, 73.54, 72.14, 69.98, 69.95, 69.66, 68.34, 68.29, 67.39, 60.95, 60.93, 60.46, 59.85, 55.64, 55.13, 22.15. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{42}H_{65}N_3NaO_{27}$ 1066.3703; found 1066.3708.

MSOPME preparative-scale synthesis of Galβ4GlcNAcβ3Galβ4(Fucα3)GlcβNHCbz (LNnFP-VβNHCbz) (5)

LacβNHCbz (**1**, 50 mg, 0.105 mmol), GlcNAc (0.16 mmol), ATP (0.16 mmol), UTP (0.16 mmol) were dissolved in water in a 50 mL centrifuge tube containing Tris-HCl buffer (200 mM, pH 7.5) and $MgCl_2$ (40 mM). After the addition of BLNahK (1.2 mg), PmGlmU (1 mg), NmLgtA (1.5 mg), and PmPpA (0.5 mg), water was added to bring the final volume to 5 mL with the concentration of LacβNHCbz (**1**) of 20 mM. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. Reaction was monitored as described above for preparation of LNT-IIβNHCbz (**2**). After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.16 mmol), ATP (0.16 mmol), and GTP (0.16 mmol) were added, and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BfFKP (1.5 mg), Hp3/4FT (1 mg), and PmPpA (0.5 mg) were then added. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2:1 (by volume) as the developing solvent and by HRMS analyses. The reaction mixture was incubated at 30 °C for 14 h with agitation at 180 rpm. After reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), NmLgtB (1.8 mg), and PmPpA (0.5 mg) were then added. The final concentration of the acceptor was kept at around 10 mM by adding Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM) to bring the final volume of the reaction to around 10.5 mL. The reaction mixture was incubated at 30 °C for 16 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.4:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and

Author Manuscript

purified. LNnFP-VβNHCbz (**5**) was obtained as a white powder (83 mg, 80% yield for three steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.39 (m, 5H), 5.44 (d, J = 4.0 Hz, 1H), 5.18 (s, 2H), 4.82 (d, J = 5.7 Hz, 1H), 4.80 (s, 1H), 4.70 (d, J = 8.3 Hz, 1H), 4.48 (d, J = 7.8 Hz, 1H), 4.42 (d, J = 7.7 Hz, 1H), 4.09 (d, J = 3.5 Hz, 1H), 3.98–3.90 (m, 4H), 3.90–3.42 (m, 23H), 2.03 (s, 3H), 1.16 (d, J = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.91, 158.18, 136.09, 128.84, 128.53, 127.85, 127.74, 102.91, 102.74, 101.74, 98.53, 81.88, 81.54, 78.25, 77.67, 76.80, 75.39, 74.57, 74.53, 73.54, 73.15, 72.55, 72.17, 72.07, 71.96, 71.01, 70.68, 69.26, 68.59, 68.31, 68.02, 67.44, 66.57, 61.52, 61.06, 59.91, 59.66, 55.20, 22.20, 15.25. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{40}H_{62}N_2NaO_{26}$ 1009.3488; found 1009.3410.

MSOPME preparative-scale synthesis of Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)GlcβNHCbz (LNnDFH-IIβNHCbz) (6)

LNnTβNHCbz (**3**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) similar to that described above for the preparation of GlcNAc-LNnTβNHCbz (**4**). After the reaction for the formation of LNnTβNHCbz (**3**) was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.3 mmol), ATP (0.32 mmol), and GTP (0.32 mmol) were added. The pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BfFKP (4 mg), Hp3/4FT (1.5 mg), and PmPpA (1 mg) were then added. The final concentration of the acceptor was kept at around 10 mM, and the final volume of the reaction was kept at around 10.5 mL by adding Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM). The reaction mixture was incubated at 30 °C in an incubator shaker for 14 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.6:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. LNnDFH-IIβNHCbz (**6**) was obtained as a white powder (94 mg, 80% yield for three steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.39 (m, 5H), 5.44 (d, J = 4.0 Hz, 1H), 5.18 (s, 2H), 5.14 (d, J = 4.0 Hz, 1H), 4.84 (d, J = 7.2 Hz, 2H), 4.71 (d, J = 8.4 Hz, 1H), 4.47 (d, J = 7.8 Hz, 1H), 4.42 (d, J = 7.8 Hz, 1H), 4.09 (d, J = 3.6 Hz, 1H), 4.01–3.55 (m, 29H), 3.50 (ddd, J = 10.6, 5.4, 3.3 Hz, 2H), 2.02 (s, 3H), 1.20–1.12 (m, 6H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.64, 158.11, 136.02, 128.78, 128.47, 127.78, 127.68, 102.48, 101.73, 101.67, 98.56, 98.46, 81.82, 81.48, 77.60, 76.73, 75.06, 74.89, 74.73, 74.47, 73.02, 72.46, 71.99, 71.89, 71.03, 70.64, 69.18, 68.33, 68.24, 67.96, 67.68, 67.38, 66.67, 66.50, 61.49, 61.46, 59.62, 55.93, 22.22, 15.29, 15.18. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{46}H_{72}N_2NaO_{30}$ 1155.4068; found 1155.4071.

MSOPME preparative-scale synthesis of Galβ3GlcNAcβ3Galβ4GlcβNHCbz (LNTβNHCbz) (7)

LNT-IIβNHCbz (**2**) was formed as an intermediate from 100 mg LacβNHCbz (**1**) similar to that described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.32 mmol), ATP (0.32 mmol), and UTP (0.32 mmol) were added, and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. SpGalK (2.5 mg), BLUSP (2 mg), Cvβ3GalT (4 mg), and PmPpA (1 mg) were then added and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the volume of the reaction mixture was kept at around 21 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 14 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. LNTβNHCbz (**7**) was obtained as a white powder (155 mg, 88% yield for two steps from 100 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.45–7.39 (m, 5H), 5.17 (t, *J* = 9.4 Hz, 2H), 4.81 (d, *J* = 9.1 Hz, 1H), 4.72 (d, *J* = 8.5 Hz, 1H), 4.43 (dt, *J* = 7.8, 1.3 Hz, 2H), 4.14 (d, *J* = 3.3 Hz, 1H), 3.92–3.87 (m, 4H), 3.82–3.61 (m, 14H), 3.60–3.50 (m, 3H), 3.47 (ddd, *J* = 10.0, 5.1, 2.3 Hz, 1H), 3.42–3.33 (m, 1H), 2.01 (s,, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 175.00, 158.15, 136.06, 128.83, 128.52, 127.87, 103.52, 102.88, 102.60, 82.07, 81.95, 81.84, 81.70, 77.76, 77.33, 76.48, 76.18, 75.31, 75.22, 75.06, 74.92, 72.49, 71.44, 70.71, 70.05, 69.27, 68.56, 68.48, 68.38, 67.45, 61.07, 61.01, 60.53, 59.90, 54.73, 22.26. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{34}H_{52}N_2NaO_{22}$ 863.2909; found 863.2911.

MSOPME gram-scale synthesis of Fucα2Galβ3GlcNAcβ3Galβ4GlcβNHCbz (LNFP-IβNHCbz) (8)

LacβNHCbz (**1**) (1 g, 2.1 mmol), GlcNAc (3.2 mmol), ATP (3.2 mmol), UTP (3.2 mmol) were dissolved in water in a 250 mL wide-mouth bottle. Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM) were then added. After BLNahK (8 mg), PmGlmU (9 mg), NmLgtA (12 mg), and PmPpA (3 mg) were added, water was added to bring the final concentration of LacβNHCbz (**1**) to 25 mM. The reaction mixture was incubated at 30 °C in an incubator shaker for 3 days with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:1.6:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, Gal (3.2 mmol), ATP (3.2 mmol), UTP (3.2 mmol) were added and the pH was adjusted to 7.5. SpGalK (16 mg), BLUSP (16.5 mg), Cvβ3GalT (14 mg), and PmPpA (3 mg) were added and the concentration of the acceptor was kept at 25 mM. The reaction mixture was incubated at 30 °C for 2 days with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, Fuc (3.2 mmol), ATP (3.2 mmol), and GTP (3.2 mmol) were added and the pH was adjusted to 7.5. BfFKP (30 mg), Te2FT (34 mg), and PmPpA (3 mg) were then added and the concentration of the acceptor was kept at 25 mM. The reaction mixture was incubated at 30 °C for 5 days with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.4:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and then centrifuge at 9016 × g for 30 min at 4 °C. The supernatant was concentrated and purified by ODS-SM column (140 g, 50 μm, 120 Å, Yamazen) using CombiFlash® Rf 200i system, with a flow rate of 40 mL/min and a gradient elution of 0–100% acetonitrile in water for 60 min. Mobile phase A: water (v/v); Mobile phase B: acetonitrile (v/v); Gradient: 0% B for 15 min, 0% to 100% B over 40 min, 100% B for 2 min, then 100% to 80% B over 3 min. LNFP-IβNHCbz (**8**) was obtained as a white powder (1.73 g, 84% yield for three steps from 1 g of LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.53–7.48 (m, 5H), 5.25 (m, *J* =

5.7 Hz, 3H), 4.92–4.83 (m, 1H), 4.70 (d, J = 7.7 Hz, 1H), 4.69 (d, J = 8.4 Hz, 1H), 4.49 (d, J = 7.9 Hz, 1H), 4.35 (q, J = 6.7 Hz, 1H), 4.21–4.17 (m, 1H), 4.05 (dd, J = 10.5, 8.6 Hz, 1H), 4.00–3.68 (m, 21H), 3.66–3.54 (m, 4H), 3.47 (dt, J = 10.0, 8.1 Hz, 1H), 2.12 (s, 3H), 1.29 (d, J = 6.6 Hz, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.27, 159.29, 136.08, 128.85, 128.54, 128.48, 127.88, 127.75, 103.27, 102.95, 100.30, 99.56, 81.65, 81.59, 78.10, 77.76, 77.35, 77.28, 76.71, 76.51, 76.22, 75.32, 75.13, 75.07, 74.87, 73.56, 71.91, 71.48, 70.26, 69.51, 69.33, 69.19, 68.64, 68.63, 68.55, 68.12, 67.46, 66.54, 62.56, 61.19, 61.01, 60.62, 60.50, 59.96, 55.02, 22.22, 15.32. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{40}H_{62}N_2NaO_{26}$ 1009.3488; found 1009.3459.

MSOPME preparative-scale synthesis of Galβ3GlcNAcβ3Galβ4(Fucα3)GlcβNHCbz (LNFP-VβNHCbz) (9)

LacβNHCbz (**1**, 50 mg, 0.105 mmol), GlcNAc (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were dissolved in water in a 50 mL centrifuge tube containing Tris-HCl buffer (200 mM, pH 7.5) and $MgCl_2$ (40 mM). BLNahK (1.2 mg), PmGlmU (1 mg), NmLgtA (1.5 mg), and PmPpA (0.5 mg) were then added and water was added to bring the total volume to 5 mL and the concentration of LacβNHCbz (**1**) to 20 mM. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. Reaction was monitored similar to that described for preparation of LNT-IIβNHCbz (**2**). After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.16 mmol), ATP (0.16 mmol), and GTP (0.16 mmol) were added and the pH was adjusted to 7.5 by adding 4 M NaOH. BfFKP (1.5 mg), Hp3/4FT (1 mg), and PmPpA (0.5 mg) were then added and the acceptor concentration was around 15 mM. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2:1 (by volume) as the developing solvent and by HRMS analyses. The reaction mixture was incubated at 30 °C for 14 h with agitation at 180 rpm. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added and the pH was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), Cvβ3GalT (2.5 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL with the final concentration of the acceptor at around 10 mM. The reaction mixture was incubated at 30 °C for 16 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.4:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. LNFP-VβNHCbz (**9**) was obtained as a white powder (84 mg, 81% yield for three steps from 50 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.45–7.41 (m, 5H), 5.44 (d, J = 4.0 Hz, 1H), 5.18 (d, J = 5.5 Hz, 2H), 4.82 (s, 1H), 4.80 (s, 1H), 4.72 (d, J = 8.5 Hz, 1H), 4.44 (d, J = 7.7 Hz, 1H), 4.42 (d, J = 7.8 Hz, 1H), 4.09 (d, J = 3.5 Hz, 1H), 3.95–3.68 (m, 19H), 3.65–3.46 (m, 8H), 2.02 (s, 3H), 1.16 (d, J = 6.7 Hz, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.98, 158.17, 136.08, 128.84, 128.83, 128.53, 127.84, 127.74, 103.50, 102.56, 101.72, 98.53, 82.01, 81.87, 81.48, 77.67, 76.80, 76.49, 75.32, 75.19, 74.54, 73.14, 72.50, 72.07, 71.95, 70.73, 70.72, 69.28, 69.25, 68.57, 68.48, 68.30, 68.02, 67.44, 66.56, 61.52, 61.08, 60.53, 59.67, 54.73, 29.64, 22.25,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

15.25. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{40}H_{62}N_2NaO_{26}$ 1009.3488; found 1009.3409.

MSOPME preparative-scale synthesis of Galβ3(Fucα4)GlcNAcβ3Galβ4(Fucα3)GlcβNHCbz (LNDFH-IIβNHCbz) (10)

LNFPVβNHCbz (**9**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.16 mmol), ATP (0.16 mmol), and GTP (0.16 mmol) were added and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BfFKP (1.5 mg), Hp3/4FT (0.8 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 14 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.6:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. LNDFH-IIβNHCbz (**10**) was obtained as a white powder (95 mg, 80% yield for four steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.39 (m, 5H), 5.44 (d, *J* = 4.0 Hz, 1H), 5.18 (s, 2H), 5.03 (d, *J* = 4.0 Hz, 1H), 4.88 (q, *J* = 6.7 Hz, 1H), 4.83–4.80 (m, 1H), 4.76 (s, 1H), 4.69 (d, *J* = 8.5 Hz, 1H), 4.51 (d, *J* = 7.7 Hz, 1H), 4.42 (d, *J* = 7.7 Hz, 1H), 4.10–4.05 (m, 2H), 3.98–3.67 (m, 22H), 3.65–3.46 (m, 8H), 2.03 (s, 3H), 1.17 (dd, *J* = 9.1, 6.6 Hz, 6H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.71, 158.11, 136.01, 128.77, 128.46, 127.77, 127.67, 102.81, 102.56, 101.68, 98.46, 97.97, 81.80, 81.50, 77.59, 76.73, 75.89, 75.16, 74.78, 74.47, 73.07, 72.28, 72.09, 72.00, 71.92, 71.89, 70.62, 70.55, 70.48, 69.19, 69.11, 68.33, 68.23, 67.95, 67.77, 67.37, 66.82, 66.50, 61.63, 61.46, 59.58, 55.84, 22.25, 15.34, 15.18. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{46}H_{72}N_2NaO_{30}$ 1155.4068; found 1155.4080.

MSOPME preparative-scale synthesis of Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (*p*LNnHβNHCbz) (11)

GlcNAc-LNnTβNHCbz (**4**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), NmLgtB (1.8 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 16 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.8:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. *p*LNnHβNHCbz (**11**) was obtained as a white powder (100 mg, 79% yield for four steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.39 (m, 4H), 5.18 (d, *J* = 3.6 Hz, 2H), 4.82 (d, *J* = 9.5 Hz, 1H), 4.70 (d, *J* = 8.4 Hz, 2H), 4.46 (dt, *J* = 15.5, 7.8 Hz, 3H), 4.15 (dd, *J* = 5.2, 3.3 Hz, 2H), 3.94 (tt, *J* = 10.9,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

4.7 Hz, 4H), 3.87–3.63 (m, 24H), 3.61–3.50 (m, 5H), 3.42 (t, *J* = 8.9 Hz, 1H), 2.03 (s, 6H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.87, 158.08, 136.00, 128.78, 128.47, 127.82, 127.68, 102.87, 102.84, 102.73, 102.71, 82.05, 81.95, 81.65, 78.18, 78.15, 78.04, 77.71, 76.12, 75.33, 75.01, 74.85, 74.53, 72.49, 72.16, 72.14, 72.03, 71.40, 70.94, 69.96, 69.94, 68.53, 68.33, 68.29, 67.40, 63.02, 62.46, 61.00, 60.95, 60.93, 59.84, 56.34, 55.17, 55.13, 22.16. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{48}H_{75}N_3NaO_{32}$ 1228.4231; found 1228.4226.

MSOPME preparative-scale synthesis of GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (GlcNAc-*p*LNnHβNHCbz) (12)

*p*LNnHβNHCbz (**11**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, GlcNAc (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BLNahK (1 mg), PmGlmU (0.8 mg), Hpβ3GlcNAcT (1.5 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:3.2:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. GlcNAc-*p*LNnHβNHCbz (**12**) was obtained as a white powder (114 mg, 77% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.46–7.39 (m, 5H), 5.18 (d, *J* = 6.1 Hz, 2H), 4.69 (d, *J* = 8.4 Hz, 2H), 4.67 (d, *J* = 8.4 Hz, 1H), 4.46 (d, *J* = 7.9 Hz, 2H), 4.43 (d, *J* = 7.9 Hz, 1H), 4.14 (t, *J* = 3.6 Hz, 3H), 3.92 (ddd, *J* = 43.6, 12.5, 2.2 Hz, 4H), 3.85–3.64 (m, 26H), 3.60–3.54 (m, 6H), 3.48–3.39 (m, 3H), 2.03 (s, 3H), 2.02 (s, 6H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.92, 174.87, 158.07, 136.00, 128.77, 128.46, 128.41, 127.81, 127.67, 102.86, 102.83, 102.72, 102.71, 82.03, 81.96, 81.94, 81.64, 78.14, 77.69, 76.12, 75.63, 75.00, 74.86, 74.84, 74.52, 73.53, 72.14, 71.39, 69.97, 69.95, 69.93, 69.65, 68.33, 68.29, 67.39, 60.94, 60.93, 60.45, 59.83, 55.63, 55.12, 29.57, 22.15. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{56}H_{88}N_4NaO_{37}$ 1431.5025; found 1431.5033.

MSOPME preparative-scale synthesis of Fucα2Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (F-*p*LNnH-IβNHCbz) (13)

*p*LNnHβNHCbz (**11**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.16 mmol), ATP (0.16 mmol), and GTP (0.16 mmol) were added and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BfFKP (1.5 mg), Hm2FT (1.5 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:3:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. F-*p*LNnH-IβNHCbz (**13**) was obtained as a white powder (108 mg, 76% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.38 (m, 5H), 5.31 (d, *J* = 3.0 Hz, 1H), 5.18 (d, *J* = 3.8 Hz, 2H), 4.82 (s, 1H), 4.70 (d, *J* = 8.4 Hz, 2H), 4.55 (d, *J* = 7.7 Hz, 1H), 4.47 (d, *J* = 8.1 Hz, 1H), 4.44 (d, *J* = 7.9 Hz, 1H), 4.22 (d, *J* = 6.7 Hz, 1H), 4.15 (d, *J* = 3.3 Hz, 2H), 4.01–3.62 (m, 33H), 3.62–3.52 (m, 3H), 3.49–3.38 (m, 2H), 2.04 (s, 3H), 2.03 (s, 3H), 1.23 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.88, 158.07, 135.98, 128.77, 128.47, 128.40, 127.81, 127.67, 102.86, 102.83, 102.75, 102.70, 100.22, 99.39, 81.98, 81.94, 78.16, 77.69, 76.43, 76.12, 75.84, 75.23, 75.07, 75.00, 74.84, 74.82, 74.52, 73.50, 72.14, 72.04, 71.64, 71.37, 69.96, 69.59, 69.09, 68.34, 68.28, 68.18, 67.39, 66.92, 61.10, 60.95, 60.89, 59.84, 55.36, 55.13, 22.18, 22.15, 15.29. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{54}H_{85}N_3NaO_{36}$ 1374.4810; found 1374.4807.

MSOPME preparative-scale synthesis of Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)GlcβNHCbz (TF-*p*LNnHβNHCbz) (14)

*p*LNnHβNHCbz (**11**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.53 mmol), ATP (0.53 mmol), and GTP (0.53 mmol) were added and the pH of the reaction was adjusted to 7.5 by adding 4 M NaOH. BfFKP (5.5 mg), Hp3/4FT (2.5 mg), and PmPpA (1 mg) were then added and the final volume was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by HRMS. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. TF-*p*LNnHβNHCbz (**14**) was obtained as a white powder (131 mg, 76% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.49–7.39 (m, 5H), 5.44 (d, *J* = 4.0 Hz, 1H), 5.18 (s, 2H), 5.13 (dd, *J* = 8.3, 4.0 Hz, 2H), 4.84 (d, *J* = 6.6 Hz, 1H), 4.82 (d, *J* = 8.1 Hz, 2H), 4.81–4.80 (m, 1H), 4.70 (d, *J* = 8.4 Hz, 2H), 4.49–4.39 (m, 3H), 4.09 (dd, *J* = 6.0, 3.4 Hz, 2H), 3.99–3.55 (m, 43H), 3.49 (ddt, *J* = 9.6, 7.9, 4.9 Hz, 3H), 2.07–1.98 (m, 6H), 1.20–1.12 (m, 9H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.64, 174.62, 158.10, 136.01, 128.77, 128.46, 127.77, 127.67, 102.46, 101.72, 101.67, 98.66, 98.56, 98.45, 81.80, 81.59, 81.45, 77.59, 76.72, 75.04, 74.89, 74.73, 74.45, 74.41, 73.08, 73.00, 72.74, 72.45, 71.97, 71.88, 71.83, 71.03, 70.64, 70.50, 69.17, 69.14, 68.33, 68.22, 67.95, 67.68, 67.61, 67.37, 66.66, 66.49, 61.48, 61.44, 59.61, 55.92, 23.24, 22.21, 15.29, 15.25, 15.18. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{66}H_{105}N_3NaO_{44}$ 1666.5969; found 1666.5840.

MSOPME preparative-scale synthesis of Neu5Acα3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (Neu5Acα2–3*p*LNnHβNHCbz) (15)

*p*LNnHβNHCbz (**11**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Neu5Ac (0.16 mmol) and CTP (0.26 mmol) were added and the pH

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was adjusted to 8.5 by adding 4 M NaOH. NmCSS (0.5 mg) and PmST3 (1.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 14 h with agitation at 180 rpm. The product formation was monitored by HRMS. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. Neu5Acα2–3*p*LNnHβNHCbz (**15**) was obtained as a white powder (123 mg, 78% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.48–7.39 (m, 5H), 5.18 (d, *J* = 3.7 Hz, 2H), 4.82 (s, 1H), 4.70 (dd, *J* = 8.5, 2.3 Hz, 2H), 4.56 (d, *J* = 7.9 Hz, 1H), 4.45 (dd, *J* = 15.4, 7.8 Hz, 2H), 4.15 (dd, *J* = 5.6, 3.3 Hz, 2H), 4.12 (dd, *J* = 9.9, 3.1 Hz, 1H), 3.99–3.38 (m, 40H), 2.76 (dd, *J* = 12.4, 4.6 Hz, 1H), 2.03 (s, 9H), 1.80 (t, *J* = 12.1 Hz, 1H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 175.06, 174.94, 174.92, 173.90, 158.15, 136.06, 128.84, 128.54, 127.88, 127.74, 102.92, 102.85, 102.79, 102.60, 99.86, 82.12, 82.02, 81.72, 78.24, 78.05, 77.79, 76.19, 75.54, 75.22, 75.08, 74.94, 74.92, 74.60, 72.94, 72.21, 72.19, 71.82, 71.47, 70.02, 69.43, 68.39, 68.35, 68.14, 67.53, 67.46, 62.63, 61.08, 61.05, 61.02, 59.91, 59.69, 55.23, 55.19, 54.41, 51.73, 42.59, 39.69, 22.23, 22.09, 12.19. HRMS (ESI-Orbitrap) m/z: [M-H]$^-$ calculated for $C_{59}H_{91}N_4O_{40}$ 1495.5215; found 1495.5212.

MSOPME preparative-scale synthesis of Neu5Acα6Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (Neu5Acα2–6*p*LNnHβNHCbz) (16)

*p*LNnHβNHCbz (**11**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Neu5Ac (0.13 mmol) and CTP (0.21 mmol) were added and the pH of the reaction was adjusted to 8.5 by adding 4 M NaOH. NmCSS (0.5 mg) and Pd2,6ST_A200Y/S232Y (1.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 14 h with agitation at 180 rpm. The product formation was monitored by HRMS. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. Neu5Acα2–6*p*LNnHβNHCbz (**16**) was obtained as a white powder (118 mg, 75% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.49–7.38 (m, 5H), 5.23–5.15 (m, 2H), 4.83 (s, 1H), 4.73 (d, *J* = 7.7 Hz, 1H), 4.70 (d, *J* = 8.3 Hz, 1H), 4.50–4.40 (m, 3H), 4.15 (t, *J* = 4.0 Hz, 2H), 4.05–3.36 (m, 40H), 2.72–2.62 (m, 1H), 2.05 (s, 3H), 2.03 (s, 6H), 1.72 (t, *J* = 12.2 Hz, 1H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.89, 173.50, 158.09, 136.00, 128.77, 128.47, 127.81, 103.44, 102.84, 102.73, 102.56, 100.11, 82.00, 81.96, 81.64, 80.46, 78.17, 77.72, 76.12, 75.00, 74.86, 74.53, 74.24, 73.68, 72.52, 72.40, 72.21, 72.15, 71.69, 71.39, 70.71, 69.95, 68.38, 68.34, 68.28, 68.19, 67.39, 63.33, 62.64, 60.95, 60.12, 59.83, 59.29, 55.11, 54.92, 51.86, 40.05, 22.27, 22.15, 22.01. HRMS (ESI-Orbitrap) m/z: [M-H]$^-$ calculated for $C_{59}H_{91}N_4O_{40}$ 1495.5215; found 1495.5140.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

MSOPME preparative-scale synthesis of Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (*p*LNHβNHCbz) (17)

GlcNAc-LNnTβNHCbz (**4**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), Cvβ3GalT (2.5 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:2.8:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. *p*LNHβNHCbz (**17**) was obtained as a white powder (101 mg, 80% yield for four steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.50–7.39 (m, 5H), 5.19 (t, *J* = 3.4 Hz, 2H), 4.82 (d, *J* = 9.8 Hz, 1H), 4.73 (d, *J* = 8.4 Hz, 1H), 4.70 (d, *J* = 8.3 Hz, 1H), 4.47 (d, *J* = 7.8 Hz, 1H), 4.44 (d, *J* = 7.8 Hz, 2H), 4.15 (t, *J* = 3.8 Hz, 2H), 3.98–3.88 (m, 4H), 3.87–3.33 (m, 30H), 1.97–2.08 (m, 6H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.94, 174.88, 158.09, 136.00, 128.78, 128.47, 127.82, 127.68, 103.47, 102.84, 102.73, 102.71, 102.55, 82.05, 82.00, 81.95, 81.65, 78.18, 78.15, 77.71, 76.12, 75.33, 75.26, 75.16, 75.01, 74.85, 74.53, 72.49, 72.45, 72.14, 71.40, 70.94, 70.66, 69.96, 68.51, 68.42, 68.34, 68.28, 67.40, 63.02, 61.01, 60.95, 60.93, 60.47, 59.84, 55.17, 55.13, 54.68, 22.21, 22.15. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{48}H_{75}N_3NaO_{32}$ 1228.4231; found 1228.4233.

MSOPME preparative-scale synthesis of Fucα2Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (F-*p*LNH-IβNHCbz) (18)

*p*LNHβNHCbz (**17**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.16 mmol), ATP (0.16 mmol), and GTP (0.16 mmol) were added and the pH was adjusted to 7.5 by adding 4 M NaOH. BfFKP (1.5 mg), Hm2FT (1.5 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 5:3:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. F-*p*LNH-IβNHCbz (**18**) was obtained as a white powder (105 mg, 74% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.44 (d, *J* = 4.4 Hz, 5H), 5.18 (dd, *J* = 5.7, 3.7 Hz, 3H), 4.81 (d, *J* = 9.4 Hz, 1H), 4.69 (d, *J* = 8.3 Hz, 1H), 4.64 (d, *J* = 7.7 Hz, 1H), 4.61 (d, *J* = 8.4 Hz, 1H), 4.44 (dd, *J* = 7.9, 6.2 Hz, 2H), 4.29 (q, *J* = 6.7 Hz, 1H), 4.14 (q, *J* = 3.3 Hz, 2H), 4.01–3.86 (m, 5H), 3.85–3.63 (m, 24H), 3.61–3.33 (m, 9H), 1.98–2.07 (m, 6H), 1.23 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.87, 174.20, 158.09, 135.99, 128.76, 128.46, 127.80, 103.22, 102.88, 102.83,

102.70, 100.21, 99.48, 81.93, 81.76, 81.58, 78.04, 77.67, 77.26, 77.12, 76.63, 76.41, 76.11, 75.21, 75.04, 74.99, 74.83, 74.78, 74.53, 73.45, 72.10, 71.82, 71.70, 71.38, 70.16, 69.94, 69.40, 69.20, 69.09, 68.53, 68.42, 68.33, 68.01, 67.38, 66.46, 61.12, 60.91, 60.50, 60.35, 59.82, 55.11, 54.94, 22.14, 22.11, 15.23. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{54}H_{85}N_3NaO_{36}$ 1374.4810; found 1374.4860.

MSOPME preparative-scale synthesis of Galβ3(Fucα4)GlcNAcβ3Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)GlcβNHCbz (TF-*p*LNH-IIIβNHCbz) (19)

*p*LNHβNHCbz (**17**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, L-fucose (0.53 mmol), ATP (0.53 mmol), and GTP (0.53 mmol) were added and the pH was adjusted to 7.5 by adding 4 M NaOH. BfFKP (5.5 mg), Hp3/4FT (2.5 mg), and PmPpA (1 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by HRMS. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. TF-*p*LNH-IIIβNHCbz (**19**) was obtained as a white powder (124 mg, 72% yield for five steps from 50 mg LacβNHCbz). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.46–7.39 (m, 5H), 5.43 (d, J = 4.0 Hz, 1H), 5.18 (d, J = 6.4 Hz, 2H), 5.12 (d, J = 4.0 Hz, 1H), 5.02 (d, J = 4.0 Hz, 1H), 4.87 (q, J = 6.7 Hz, 1H), 4.84–4.79 (m, 3H), 4.69 (t, J = 7.8 Hz, 2H), 4.51 (d, J = 7.6 Hz, 1H), 4.43 (dd, J = 16.3, 7.8 Hz, 2H), 4.11–4.04 (m, 3H), 3.98–3.66 (m, 33H), 3.65–3.44 (m, 12H), 2.02 (s, 3H), 2.01 (s, 3H), 1.19–1.12 (m, 9H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.69, 174.64, 158.11, 136.01, 128.76, 128.46, 127.77, 127.67, 102.82, 102.55, 102.45, 101.69, 98.66, 98.45, 97.97, 81.79, 81.63, 81.44, 77.58, 76.72, 75.89, 75.15, 75.03, 74.78, 74.72, 74.45, 74.42, 73.07, 72.76, 72.27, 72.07, 71.96, 71.92, 71.87, 71.83, 70.63, 70.46, 69.18, 69.14, 69.10, 68.32, 68.22, 67.94, 67.75, 67.61, 67.37, 66.82, 66.65, 66.49, 61.63, 61.44, 59.59, 55.92, 55.84, 22.24, 22.20, 15.34, 15.25, 15.17. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{66}H_{105}N_3NaO_{44}$ 1666.5969; found 1666.5909.

MSOPME preparative-scale synthesis of Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (*p*LNnOβNHCbz) (20)

GlcNAc-*p*LNnHβNHCbz (**12**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), NmLgtB (1.8 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 16 h with agitation at 180 rpm. The product formation was monitored by HRMS. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. *p*LNnOβNHCbz (**20**) was obtained as a white powder (119 mg, 72% yield for six steps from

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

50 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.46–7.38 (m, 5H), 5.18 (dd, *J* = 11.8, 6.8 Hz, 2H), 4.81 (d, *J* = 10.4 Hz, 1H), 4.69 (dd, *J* = 8.4, 2.5 Hz, 3H), 4.48–4.44 (m, 3H), 4.43 (d, *J* = 7.8 Hz, 1H), 4.14 (dt, *J* = 5.3, 2.9 Hz, 3H), 3.95–3.85 (m, 5H), 3.85–3.61 (m, 32H), 3.60–3.50 (m, 7H), 3.36–3.44 (m, 1H), 2.00–2.03 (m, 9H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.86, 135.98, 128.76, 128.46, 128.40, 127.80, 127.66, 102.85, 102.82, 102.73, 102.71, 82.03, 81.94, 81.63, 78.11, 78.01, 77.67, 76.11, 75.32, 74.99, 74.83, 74.51, 72.47, 72.14, 72.13, 71.38, 70.93, 69.94, 69.92, 68.51, 68.32, 68.28, 67.39, 61.00, 60.94, 60.92, 59.81, 55.15, 55.11, 22.14. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{62}H_{98}N_4NaO_{42}$ 1593.5553; found 1593.5419.

MSOPME preparative-scale synthesis of Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (*p*LNOβNHCbz) (21)

GlcNAc-*p*LNnHβNHCbz (**12**) was prepared as an intermediate from 50 mg LacβNHCbz (**1**) as described above. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min and then cooled down. In the same reaction tube without workup or purification, Gal (0.16 mmol), ATP (0.16 mmol), and UTP (0.16 mmol) were added, and the pH was adjusted to 7.5 by adding 4 M NaOH. SpGalK (1.5 mg), BLUSP (1.2 mg), Cvβ3GalT (2.5 mg), and PmPpA (0.5 mg) were then added and the final volume of the reaction was kept at around 10.5 mL. The reaction mixture was incubated at 30 °C in an incubator shaker for 18 h with agitation at 180 rpm. The product formation was monitored by HRMS. After the reaction was completed, the reaction mixture was incubated in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. *p*LNOβNHCbz (**21**) was obtained as a white powder (117 mg, 71% yield for six steps from 50 mg LacβNHCbz). ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 7.47–7.38 (m, 5H), 5.17 (d, *J* = 6.2 Hz, 2H), 4.81 (d, *J* = 9.1 Hz, 1H), 4.69 (dd, *J* = 8.4, 2.2 Hz, 3H), 4.46 (dd, *J* = 11.2, 7.9 Hz, 3H), 4.43 (d, *J* = 7.9 Hz, 1H), 4.14 (dt, *J* = 6.0, 2.6 Hz, 3H), 3.95–3.87 (m, 6H), 3.85–3.74 (m, 12H), 3.73–3.68 (m, 16H), 3.66–3.62 (m, 3H), 3.59–3.51 (m, 7H), 3.40 (t, *J* = 9.1 Hz, 1H), 2.00–2.03 (m, 9H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.86, 158.08, 135.98, 128.75, 128.46, 127.80, 103.45, 102.85, 102.82, 102.72, 102.71, 102.54, 82.02, 81.93, 81.62, 78.11, 77.66, 76.11, 75.31, 75.24, 75.14, 74.99, 74.83, 74.51, 72.46, 72.43, 72.12, 71.37, 70.93, 70.64, 69.96, 69.94, 69.92, 68.51, 68.49, 68.40, 68.32, 68.28, 67.39, 60.99, 60.93, 60.92, 60.45, 59.81, 55.15, 55.11, 22.19, 22.14. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{62}H_{98}N_4NaO_{42}$ 1593.5553; found 1593.5554.

Preparative-scale synthesis of Fucα2Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcβNHCbz (Fucα2*p*LNOβNHCbzβNHCbz) (22) from 21 by OPME3a

*p*LNOβNHCbz (**21**, 20 mg, 0.0127 mmol), L-fucose (0.019 mmol), ATP (0.019 mmol), and GTP (0.019 mmol) were dissolved in a small amount of water. Tris-HCl buffer (100 mM, pH 7.5) and $MgCl_2$ (20 mM) were added. After the addition of BfFKP (0.8 mg), Hm2FT (0.5 mg), and PmPpA (0.1 mg), water was added to bring the final concentration of *p*LNOβNHCbz (**21**) to 10 mM. The reaction mixture was incubated at 30 °C in an incubator shaker for 12 h with agitation at 180 rpm. The product formation was monitored by TLC assays with EtOAc:MeOH:H_2O = 4:3:1 (by volume) as the developing solvent and by HRMS analyses. After the reaction was completed, the reaction mixture was incubated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in a boiling water bath for 5 min to denature the enzymes, cooled down, and purified. Fucα2*p*LNOβNHCbzβNHCbz (**22**) was obtained as a white powder (20 mg, 92% yield from 20 mg of *p*LNOβNHCbz **21**). ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 7.52–7.36 (m, 5H), 5.19 (d, *J* = 4.0 Hz, 3H), 4.84–4.81 (m, 1H), 4.70 (d, *J* = 8.4 Hz, 2H), 4.63 (dd, *J* = 15.8, 8.0 Hz, 2H), 4.48–4.42 (m, 3H), 4.29 (q, *J* = 6.5 Hz, 1H), 4.15 (q, *J* = 3.6 Hz, 3H), 4.01–3.88 (m, 6H), 3.86–3.64 (m, 34H), 3.61–3.48 (m, 8H), 3.41 (q, *J* = 8.2, 7.5 Hz, 1H), 2.06 (s, 3H), 2.03 (s, 3H), 1.23 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.95, 174.27, 136.07, 128.84, 128.54, 128.48, 127.88, 127.74, 103.29, 102.93, 102.90, 102.82, 102.79, 100.29, 99.56, 99.46, 82.10, 82.03, 78.21, 78.10, 77.77, 77.20, 76.70, 76.49, 76.18, 75.91, 75.29, 75.13, 75.07, 74.91, 74.59, 73.57, 72.21, 72.11, 71.90, 71.71, 71.46, 70.23, 70.02, 70.00, 69.66, 69.48, 69.16, 68.60, 68.51, 68.40, 68.36, 68.25, 68.09, 67.46, 66.99, 66.53, 61.16, 60.99, 60.43, 60.04, 59.90, 55.42, 55.19, 55.02, 22.24, 22.22, 22.19, 15.36, 15.30. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{68}H_{108}N_4NaO_{46}$ 1739.6132; found 1739.6113.

Hydrogenation procedures

To remove the Cbz tag from the obtained glycosides (**2–22**), a catalytic amount (10%) of palladium on charcoal (Pd/C) (2 mg) was added to a solution containing 20 mg of each compound in H_2O:MeOH = 1:1 (by volume) (3 mL). The mixture was stirred at room temperature under a hydrogen atmosphere with a balloon. The reaction was monitored by HRMS analyses. When the reaction was completed (12 h), the mixture was passed through a 0.45 μm syringe filter to remove palladium and charcoal. The solvent was removed *in vacuo*. The residue obtained was dissolved in H_2O (3 mL) and the mixture was incubated at 37 °C with 120 rpm agitation in an incubator shaker. The reaction was monitored by HRMS analyses until the β-glycosylamine was completely converted to the target HMO with a free reducing end (4–5 days for HMOs containing an L-fucose linked to the reducing end Glc residue or 20–48 hours for other HMOs). The pure HMO was obtained by lyophilization without further purification.

GlcNAcβ3Galβ4Glc (LNT-II, 23).—White powder, 15.4 mg, 95% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.21 (d, *J* = 3.7 Hz, 0.4H), 4.67 (d, *J* = 8.5 Hz, 1H), 4.65 (d, *J* = 8.0 Hz, 0.6 H), 4.42 (dd, *J* = 7.9, 1.4 Hz, 1H), 4.14 (d, *J* = 3.4 Hz, 1H), 3.95–3.68 (m, 9H), 3.68–3.22 (m, 7H), 2.02 (s, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.92, 102.89, 102.85, 102.81, 95.70, 91.77, 81.91, 81.89, 78.32, 78.22, 75.65, 75.62, 74.86, 74.76, 74.31, 73.97, 73.75, 73.52, 71.36, 71.09, 70.08, 70.00, 69.97, 69.66, 69.65, 68.34, 68.31, 60.94, 60.93, 60.44, 60.04, 59.91, 55.68, 55.62, 22.14, 22.12. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{20}H_{35}NNaO_{16}$ 568.1854; found 568.1889.

Galβ4GlcNAcβ3Galβ4Glc (LNnT, 24).—White powder, 16.5 mg, 95% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.22 (d, *J* = 3.8 Hz, 0.4H), 4.73–4.71 (d, 1H), 4.67 (d, *J* = 7.9 Hz, 0.6H), 4.48 (d, *J* = 7.8 Hz, 1H), 4.44 (d, *J* = 7.8 Hz, 1H), 4.16 (d, *J* = 3.3 Hz, 1H), 3.99–3.92 (m, 3H), 3.90–3.71 (m, 13H), 3.69–3.53 (m, 7H), 3.28 (t, *J* = 8.5 Hz, 0.6H), 2.04 (s, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.88, 102.91, 102.87, 102.84, 102.72, 95.72, 91.79, 82.01, 78.36, 78.26, 78.16, 75.33, 74.86, 74.78, 74.54, 74.33, 73.77, 72.49, 72.16, 71.38, 71.11, 70.95, 70.10, 69.95, 68.53, 68.32, 61.01, 60.95, 60.06, 59.93, 59.85, 55.18,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

22.16. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{26}H_{45}NNaO_{21}$ 730.2382; found 730.2371.

GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (25).—White powder, 17.2 mg, 95% yield. ^1H NMR (800 MHz, D_2O, 30 °C) δ 5.21 (d, *J* = 3.8 Hz, 0.4H), 4.69 (dd, *J* = 8.4, 3.0 Hz, 1H), 4.67 (d, *J* = 8.5 Hz, 1H), 4.65 (d, *J* = 8.0 Hz, 0.6H), 4.45 (d, *J* = 7.9 Hz, 1H), 4.42 (d, *J* = 7.9, 1.4 Hz, 1H), 4.14 (d, *J* = 3.3 Hz, 2H), 3.91 (ddd, *J* = 44.3, 12.1, 2.0 Hz, 3H), 3.84–3.54 (m, 23H), 3.47–3.26 (m, 2H), 2.04–2.00 (m, 6H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.91, 174.86, 102.89, 102.85, 102.83, 102.70, 95.70, 91.77, 81.99, 81.97, 81.96, 78.32, 78.21, 78.13, 75.62, 74.85, 74.83, 74.76, 74.51, 74.31, 73.75, 73.52, 72.13, 71.36, 71.09, 70.08, 69.96, 69.93, 69.63, 68.33, 68.30, 68.28, 60.93, 60.92, 60.43, 60.03, 59.90, 59.82, 55.62, 55.11, 22.14, 22.12. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{34}H_{58}N_2NaO_{26}$ 933.3175; found 933.3238.

Galβ4GlcNAcβ3Galβ4(Fucα3)Glc (LNnFP-V, 26).—White powder, 17.1 mg, 96% yield. ^1H NMR (800 MHz, D_2O, 30 °C) δ 5.41 (d, *J* = 4.0 Hz, 0.4H), 5.36 (d, *J* = 4.0 Hz, 0.4H), 5.17 (d, *J* = 3.8 Hz, 0.3H), 4.80 (d, *J* = 6.8 Hz, 1H), 4.69 (dd, *J* = 8.4, 2.9 Hz, 1H), 4.64 (d, *J* = 8.0 Hz, 0.4H), 4.47 (d, *J* = 7.8 Hz, 1H), 4.40 (d, *J* = 8.2 Hz, 1H), 4.08 (d, *J* = 3.4 Hz, 1H), 3.93 (tdd, *J* = 12.1, 6.2, 2.7 Hz, 4H), 3.86–3.65 (m, 18H), 3.58–3.43 (m, 5H), 2.01–2.03 (m, 3H), 1.15 (dd, *J* = 6.7, 4.6 Hz, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.83, 102.82, 102.66, 101.69, 98.52, 98.41, 95.80, 92.08, 81.53, 81.46, 78.14, 76.96, 75.50, 75.33, 75.31, 74.65, 74.48, 74.46, 72.64, 72.46, 72.34, 72.24, 72.09, 71.89, 71.88, 70.93, 70.91, 70.63, 70.58, 69.24, 69.18, 68.51, 68.25, 68.22, 67.99, 67.96, 66.46, 66.42, 61.46, 61.43, 60.99, 59.82, 59.75, 59.68, 55.13, 22.12, 15.19, 15.17. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{32}H_{55}NNaO_{25}$ 876.2961; found 876.2956.

Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)Glc (LNnDFH-II, 27).—White powder, 17.3 mg, 96% yield. ^1H NMR (800 MHz, D_2O, 30 °C) δ 5.41 (d, *J* = 4.0 Hz, 0.4H), 5.36 (d, *J* = 4.0 Hz, 0.4H), 5.17 (d, *J* = 3.8 Hz, 0.4H), 5.12 (d, *J* = 4.1 Hz, 1H), 4.82 (d, *J* = 6.7 Hz, 1H), 4.80 (d, 1H), 4.69 (dd, *J* = 8.5, 2.8 Hz, 1H), 4.64 (d, *J* = 8.0 Hz, 0.4H), 4.45 (d, *J* = 7.8 Hz, 1H), 4.40 (d, *J* = 7.7 Hz, 1H), 4.08 (d, *J* = 3.4 Hz, 1H), 3.97–3.83 (m, 11H), 3.81–3.41 (m, 20H),1.99–2.02 (m, 3H), 1.18–1.13 (m, 6H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.62, 102.47, 101.71, 101.69, 98.55, 98.51, 98.40, 95.80, 92.09, 81.52, 81.45, 76.95, 75.50, 75.33, 75.04, 74.88, 74.72, 74.64, 74.45, 72.99, 72.88, 72.64, 72.43, 72.31, 72.21, 71.89, 71.87, 71.01, 70.91, 70.66, 70.61, 69.23, 69.18, 69.15, 68.31, 68.24, 68.21, 67.99, 67.96, 67.66, 66.65, 66.45, 66.41, 61.47, 61.43, 59.76, 59.69, 59.59, 55.91, 22.20, 15.27, 15.18, 15.17. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{38}H_{65}NNaO_{29}$ 1022.3540; found 1022.3510.

Galβ3GlcNAcβ3Galβ4Glc (LNT, 28).—White powder, 15.6 mg, 95%. ^1H NMR (800 MHz, D_2O, 30 °C) δ 5.28 (d, *J* = 3.8 Hz, 0.3 H), 4.80 (d, *J* = 4.9 Hz, 1H), 4.72 (d, *J* = 7.9 Hz, 0.5 H), 4.50 (dd, *J* = 7.8, 3.0 Hz, 2H), 4.21 (d, *J* = 3.3 Hz, 1H), 4.03–3.75 (m, 15H), 3.72–3.32 (m, 8H), 2.09 (s, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.99, 103.52, 102.97, 102.93, 102.56, 95.79, 91.86, 82.18, 82.03, 82.00, 81.82, 78.70, 78.52, 78.42, 75.33, 75.25, 74.94, 74.85, 74.41, 73.85, 72.54, 71.46, 71.20, 70.74, 70.18, 70.09, 70.06,

68.59, 68.53, 68.39, 68.37, 61.07, 61.01, 60.99, 60.59, 60.18, 60.05, 54.76, 22.30. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{26}H_{45}NNaO_{21}$ 730.2382; found 730.2361.

Fucα2Galβ3GlcNAcβ3Galβ4Glc (LNFP-I, 29).—White powder, 16.6 mg, 96% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.22 (d, J = 3.6 Hz, 0.4H), 5.19 (d, J = 4.1 Hz, 1H), 4.68–4.61 (m, 2.6H), 4.42 (d, J = 7.9 Hz, 1H), 4.29 (q, J = 6.6 Hz, 1H), 4.14 (d, J = 3.3 Hz, 1H), 4.02–3.45 (m, 27H), 3.28 (t, J = 8.5 Hz, 0.5H), 2.06 (s, 3H), 1.23 (d, J = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.22, 103.21, 102.93, 102.89, 100.22, 99.49, 95.87, 95.71, 92.05, 91.78, 81.54, 78.21, 78.12, 77.13, 76.64, 75.91, 75.71, 75.22, 75.04, 74.79, 74.30, 74.09, 73.76, 73.46, 72.72, 71.83, 71.34, 71.10, 70.18, 70.12, 69.60, 69.56, 69.40, 69.10, 68.56, 68.45, 68.02, 66.47, 61.12, 60.93, 60.70, 60.54, 60.37, 60.05, 59.92, 54.95, 22.11, 15.24.

Galβ3GlcNAcβ3Galβ4(Fucα3)Glc (LNFP-V, 30).—White powder, 16.8 mg, 95% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.41 (d, J = 4.0 Hz, 0.5H), 5.36 (d, J = 4.0 Hz, 0.4H), 5.17 (d, J = 3.8 Hz, 0.4H), 4.81 (d, J = 1.3 Hz, 1H), 4.71 (dd, J = 8.5, 3.2 Hz, 1H), 4.64 (d, J = 8.0 Hz, 0.5H), 4.43 (d, J = 7.7 Hz, 1H), 4.40 (d, J = 7.9 Hz, 1H), 4.08 (d, J = 3.4 Hz, 1H), 3.97–3.67 (m, 24H), 3.63–3.37 (m, 9H), 2.00–2.02 (m, 3H), 1.15 (dd, J = 6.7, 4.6 Hz, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.91, 103.42, 102.48, 101.68, 98.52, 98.41, 95.85, 95.80, 92.08, 81.93, 81.47, 81.39, 76.96, 75.70, 75.50, 75.33, 75.24, 75.11, 74.65, 74.47, 72.64, 72.41, 72.34, 72.24, 71.89, 71.87, 71.82, 70.91, 70.69, 70.63, 69.54, 69.23, 69.18, 68.49, 68.39, 68.24, 68.20, 67.99, 67.96, 66.45, 66.41, 61.45, 61.43, 61.00, 60.69, 60.44, 59.76, 59.69, 54.66, 22.17, 15.18. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{32}H_{55}NNaO_{25}$ 876.2961; found 876.2968.

Galβ3(Fucα4)GlcNAcβ3Galβ4(Fucα3)Glc (LNDFH-II, 31).—White powder, 17.5 mg, 97% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.42 (d, J = 4.0 Hz, 0.4H), 5.36 (d, J = 4.0 Hz, 0.4H), 5.17 (d, J = 3.8 Hz, 0.4H), 5.02 (d, J = 4.0 Hz, 1H), 4.87 (q, J = 6.8 Hz, 1H), 4.83–4.80 (m, 1H), 4.68 (d, J = 8.4 Hz, 1H), 4.64 (d, J = 8.0 Hz, 0.5H), 4.51 (d, J = 7.7 Hz, 1H), 4.41 (d, J = 7.8 Hz, 1H), 4.10–4.03 (m, 2H), 3.99–3.65 (m, 22H), 3.65–3.42 (m, 8H), 2.05–2.00 (m, 3H), 1.16 (dd, J = 10.9, 6.8 Hz, 6H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.71, 102.81, 102.55, 101.71, 98.52, 98.41, 97.97, 95.81, 92.09, 81.56, 81.49, 76.96, 75.89, 75.51, 75.34, 75.15, 74.77, 74.65, 74.47, 72.65, 72.35, 72.27, 72.08, 71.91, 70.92, 70.64, 70.59, 70.46, 69.24, 69.18, 69.10, 68.32, 68.20, 68.00, 67.97, 67.75, 66.81, 66.46, 66.42, 61.62, 61.44, 59.76, 59.56, 55.83, 22.24, 15.33, 15.17. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{38}H_{65}NNaO_{29}$ 1022.3540; found 1022.3551.

Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (*p*LNnH, 32).—White powder, 17.4 mg, 96% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.21 (d, J = 3.8 Hz, 0.3H), 4.69 (d, J = 8.4 Hz, 2H), 4.65 (d, J = 8.0 Hz, 0.5H), 4.46 (dd, J = 10.2, 7.9 Hz, 2H), 4.42 (d, J = 7.9 Hz, 1H), 4.14 (t, J = 3.8 Hz, 2H), 3.95–3.90 (m, 4H), 3.88–3.69 (m, 21H), 3.67–3.51 (m, 9H), 2.02 (s, 6H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.86, 102.89, 102.85, 102.82, 102.72, 102.70, 95.70, 91.77, 82.03, 81.99, 81.97, 78.32, 78.21, 78.12, 78.10, 75.31, 74.83, 74.76, 74.51, 74.31, 73.74, 72.46, 72.14, 72.13, 71.36, 71.09, 70.92, 70.08, 69.95, 69.92, 68.51,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

68.30, 68.27, 60.99, 60.93, 60.92, 60.03, 59.81, 55.15, 55.11, 22.14. HRMS (ESI-Orbitrap) m/z: [M+Na]$^+$ calculated for $C_{40}H_{68}N_2NaO_{31}$ 1095.3704; found 1095.3761.

GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (GlcNAc-*p*LNnH, 33).— White powder, 17.7 mg, 96% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.21 (d, *J* = 3.8 Hz, 0.3H), 4.70–4.66 (m, 3H), 4.65 (d, *J* = 7.9 Hz, 0.6H), 4.45 (d, *J* = 7.9 Hz, 2H), 4.42 (dd, *J* = 7.9, 1.4 Hz, 1H), 4.14 (t, *J* = 2.7 Hz, 3H), 3.95–3.68 (m, 29H), 3.64–3.54 (m, 8H), 3.47–3.41 (m, 2H), 3.29–3.24 (m, 0.5H), 2.00–2.05 (m, 9H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.91, 174.86, 102.89, 102.85, 102.83, 102.72, 102.70, 95.70, 91.77, 82.02, 81.99, 81.97, 81.95, 78.32, 78.21, 78.11, 75.61, 74.85, 74.83, 74.76, 74.51, 74.31, 73.75, 73.52, 72.13, 71.36, 71.09, 70.08, 69.96, 69.92, 69.63, 68.30, 68.28, 60.92, 60.43, 60.03, 59.90, 59.81, 55.62, 55.11, 22.14. HRMS (ESI-Orbitrap) m/z: [M+Na]$^+$ calculated for $C_{48}H_{81}N_3NaO_{36}$ 1298.4497; found 1298.4452.

Fucα2Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (F-*p*LNnH-I, 34).—White powder, 17.6 mg, 96% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.29 (d, *J* = 3.4 Hz, 1H), 5.20 (d, *J* = 3.7 Hz, 0.4H), 4.70–4.67 (m, 2H), 4.65 (d, *J* = 8.0 Hz, 0.6H), 4.53 (d, *J* = 7.7 Hz, 1H), 4.45 (d, *J* = 7.8 Hz, 1H), 4.42 (dd, *J* = 7.9, 1.4 Hz, 1H), 4.20 (q, *J* = 6.7 Hz, 1H), 4.13 (t, *J* = 3.5 Hz, 2H), 3.94 (ddd, *J* = 12.5, 8.6, 2.2 Hz, 3H), 3.88–3.43 (m, 35H), 3.26 (dd, *J* = 9.2, 7.9 Hz, 0.6H), 2.05–2.00 (m, 6H), 1.21 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.88, 174.87, 102.89, 102.85, 102.84, 102.75, 102.70, 100.21, 99.37, 95.70, 91.77, 81.99, 81.97, 78.32, 78.21, 78.12, 76.40, 75.82, 75.22, 75.05, 74.83, 74.81, 74.76, 74.51, 74.31, 73.75, 73.49, 72.13, 72.03, 71.63, 71.36, 71.09, 70.08, 69.95, 69.93, 69.57, 69.08, 68.32, 68.30, 68.27, 68.16, 66.91, 61.09, 60.93, 60.92, 60.89, 60.03, 59.96, 59.90, 59.81, 55.34, 55.12, 22.16, 22.13, 15.28. HRMS (ESI-Orbitrap) m/z: [M+Na]$^+$ calculated for $C_{46}H_{78}N_2NaO_{35}$ 1241.4283; found 1241.4262.

Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)Glc (TF-*p*LNnH, 35).—White powder, 17.9 mg, 95% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.41 (d, *J* = 4.0 Hz, 0.4H), 5.36 (d, *J* = 4.0 Hz, 0.3H), 5.17 (d, *J* = 3.8 Hz, 0.3H), 5.12 (d, *J* = 4.0 Hz, 1H), 5.11 (d, *J* = 4.2 Hz, 1H), 4.82 (d, *J* = 6.7 Hz, 1H), 4.81 (d, *J* = 6.8 Hz, 2H), 4.69 (d, *J* = 8.4 Hz, 2H), 4.64 (d, *J* = 8.0 Hz, 0.4H), 4.45 (d, *J* = 7.8 Hz, 1H), 4.43 (d, *J* = 7.8 Hz, 1H), 4.40 (d, *J* = 7.8 Hz, 1H), 4.08 (t, *J* = 4.1 Hz, 2H), 3.98–3.81 (m, 18H), 3.79–3.44 (m, 28H), 2.02–1.98 (m, 6H), 1.17–1.11 (m, 9H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.64, 174.61, 102.46, 101.70, 98.65, 98.55, 81.58, 76.95, 75.51, 75.33, 75.03, 74.88, 74.72, 74.45, 74.41, 72.98, 72.88, 72.72, 72.43, 71.87, 71.82, 71.62, 71.01, 70.49, 69.23, 69.15, 69.13, 68.31, 68.21, 67.99, 67.66, 67.60, 66.66, 66.45, 61.47, 61.43, 59.59, 55.92, 22.20, 15.47, 15.28, 15.24, 15.17. HRMS (ESI-Orbitrap) m/z: [M+Na]$^+$ calculated for $C_{58}H_{98}N_2NaO_{43}$ 1533.5441; found 1533.5476.

Neu5Acα3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (Neu5Acα2–3*p*LNnH, 36). —White powder, 17.1 mg, 94% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.28 (d, *J* = 3.8 Hz, 0.4H), 4.76 (d, *J* = 4.2 Hz, 2H), 4.72 (d, *J* = 8.0 Hz, 0.7H), 4.61 (d, *J* = 7.8 Hz, 1H), 4.53 (d, *J* = 7.9 Hz, 1H), 4.50 (d, *J* = 7.9 Hz, 1H), 4.21 (t, *J* = 4.1 Hz, 2H), 4.17 (dd, *J* = 9.9, 3.2 Hz, 1H), 4.04–3.99 (m, 4H), 3.96–3.75 (m, 25H), 3.73–3.61 (m, 11H), 3.34 (t,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

J = 8.5 Hz, 0.6H), 2.82 (dd, *J* = 12.5, 4.7 Hz, 1H), 2.12–2.07 (m, 9H), 1.86 (t, *J* = 12.1 Hz, 1H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 175.08, 174.92, 173.86, 102.98, 102.94, 102.81, 102.74, 102.63, 99.88, 95.79, 91.86, 88.76, 82.12, 82.08, 82.05, 78.51, 78.41, 78.34, 78.15, 75.57, 75.23, 74.95, 74.93, 74.85, 74.62, 74.41, 73.86, 72.95, 72.23, 72.21, 71.81, 71.45, 71.20, 70.18, 70.03, 69.43, 68.41, 68.38, 68.35, 68.18, 67.55, 62.67, 61.08, 61.01, 61.00, 60.17, 60.04, 59.94, 59.63, 55.25, 55.22, 51.76, 39.72, 22.24, 22.10. HRMS (ESI-Orbitrap) m/z: [M-H]$^-$ calculated for $C_{51}H_{84}N_3O_{39}$ 1362.4687; found 1362.4706.

Neu5Acα6Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (Neu5Acα2–6*p*LNnH, 37).—White powder, 17.1 mg, 94% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.33 (d, *J* = 3.8 Hz, 0.4H), 4.85 (d, *J* = 7.6 Hz, 1H), 4.83 (dd, *J* = 8.4, 3.0 Hz, 1H), 4.77 (s, 0.5H), 4.61–4.54 (m, 3H), 4.26 (t, *J* = 3.6 Hz, 2H), 4.12–4.03 (m, 5H), 4.01–3.81 (m, 22H), 3.79–3.64 (m, 14H), 3.39 (dd, *J* = 9.0, 7.9 Hz, 0.6H), 2.79 (dd, *J* = 12.4, 4.7 Hz, 1H), 2.18–2.13 (m, 9H), 1.82 (t, *J* = 12.2 Hz, 1H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 175.05, 174.99, 174.98, 173.57, 103.53, 103.05, 103.01, 102.78, 102.63, 100.30, 95.86, 91.93, 82.14, 82.12, 80.56, 78.64, 78.54, 78.48, 74.99, 74.91, 74.69, 74.48, 74.43, 73.94, 73.82, 72.67, 72.59, 72.35, 72.31, 71.84, 71.52, 71.28, 70.88, 70.24, 70.13, 70.10, 68.57, 68.52, 68.44, 68.41, 68.32, 63.47, 62.83, 61.06, 60.34, 60.26, 60.05, 55.28, 55.13, 54.51, 52.03, 40.22, 22.43. HRMS (ESI-Orbitrap) m/z: [M-H]$^-$ calculated for $C_{51}H_{84}N_3O_{39}$ 1362.4687; found 1362.4661.

Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (*p*LNH, 38).—White powder, 17.4 mg, 96% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.22 (d, *J* = 3.8 Hz, 0.3H), 4.74–4.68 (m, 2H), 4.47 (d, *J* = 7.9, 2.2 Hz, 0.5H), 4.44 (dd, *J* = 7.8, 2.4 Hz, 2H), 4.15 (t, *J* = 2.6 Hz, 2H), 3.95 (d, *J* = 12.3 Hz, 2H), 3.92–3.87 (m, 3H), 3.85–3.69 (m, 20H), 3.65–3.46 (m, 9H), 3.28 (t, *J* = 8.5 Hz, 0.5H), 2.01–2.04 (m, 6H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.94, 174.88, 103.46, 102.90, 102.86, 102.71, 102.55, 95.71, 91.79, 82.04, 82.00, 78.35, 78.24, 78.16, 75.26, 75.16, 74.85, 74.77, 74.52, 74.32, 73.76, 72.44, 72.15, 71.37, 71.10, 70.65, 70.09, 69.97, 68.50, 68.42, 61.00, 60.93, 60.46, 60.05, 59.83, 55.13, 54.67, 54.33, 22.20, 22.15. HRMS (ESI-Orbitrap) m/z: [M+Na]$^+$ calculated for $C_{40}H_{68}N_2NaO_{31}$ 1095.3704; found 1095.3770.

Fucα2Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (F-*p*LNH-I, 39).—White powder, 17.1 mg, 95% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.22 (d, *J* = 3.6 Hz, 0.4H), 5.19 (d, *J* = 4.1 Hz, 1H), 4.70 (dd, *J* = 8.3, 2.2 Hz, 1H), 4.64 (ddd, *J* = 15.9, 10.7, 8.2 Hz, 2.6H), 4.44 (t, *J* = 8.2 Hz, 2H), 4.29 (q, *J* = 6.5 Hz, 1H), 4.14 (dd, *J* = 6.3, 3.3 Hz, 2H), 4.01–3.87 (m, 6H), 3.85–3.70 (m, 21H), 3.69–3.63 (m, 3H), 3.61–3.38 (m, 8H), 3.28 (dd, *J* = 9.1, 8.0 Hz, 0.6H), 2.09–2.00 (m, 6H), 1.23 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.88, 174.21, 103.23, 102.89, 102.70, 100.22, 99.49, 95.87, 95.71, 92.05, 91.78, 82.00, 81.60, 78.34, 78.24, 78.05, 77.13, 76.63, 75.91, 75.71, 75.22, 75.05, 74.84, 74.78, 74.54, 74.32, 74.09, 73.76, 73.46, 72.72, 72.12, 71.83, 71.44, 71.37, 71.10, 70.16, 70.09, 69.94, 69.56, 69.41, 69.10, 68.54, 68.44, 68.32, 68.02, 66.47, 61.12, 60.92, 60.70, 60.36, 60.05, 59.83, 55.13, 54.95, 22.15, 22.12, 15.23. HRMS (ESI-Orbitrap) m/z: [M+Na]$^+$ calculated for $C_{46}H_{78}N_2NaO_{35}$ 1241.4283; found 1241.4254.

Galβ3(Fucα4)GlcNAcβ3Galβ4(Fucα3)GlcNAcβ3Galβ4(Fucα3)Glc (TF-*p*LNH-III, 40).—White powder, 17.8 mg, 97% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.43 (d, *J* = 4.0 Hz, 0.5H), 5.37 (d, *J* = 4.0 Hz, 0.4H), 5.18 (d, *J* = 3.8 Hz, 0.4H), 5.12 (d, *J* = 4.0 Hz, 1H), 5.03 (d, *J* = 4.0 Hz, 1H), 4.88 (q, *J* = 6.7 Hz, 1H), 4.81 (d, *J* = 6.6 Hz, 2H), 4.70 (t, *J* = 8.0 Hz, 2H), 4.65 (d, *J* = 8.0 Hz, 0.5H), 4.51 (d, *J* = 7.7 Hz, 1H), 4.45 (d, *J* = 7.9 Hz, 1H), 4.41 (d, *J* = 7.8 Hz, 1H), 4.10 (t, *J* = 4.0 Hz, 2H), 4.06 (d, *J* = 9.7 Hz, 1H), 3.98–3.83 (m, 16H), 3.82–3.67 (m, 18H), 3.64–3.44 (m, 11H), 2.03 (s, 3H), 2.02 (s, 3H), 1.21–1.13 (m, 9H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.69, 174.64, 102.82, 102.55, 102.45, 101.70, 98.66, 97.97, 95.81, 92.10, 81.64, 81.52, 76.97, 75.90, 75.52, 75.34, 75.16, 75.04, 74.79, 74.73, 74.43, 72.77, 72.66, 72.34, 72.29, 72.08, 71.92, 71.84, 70.93, 70.63, 70.47, 69.55, 69.20, 69.15, 69.11, 68.33, 68.21, 67.97, 67.76, 67.62, 66.82, 66.66, 66.46, 61.63, 61.44, 59.58, 55.93, 55.84, 22.25, 22.21, 15.34, 15.25, 15.18. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{58}H_{98}N_2NaO_{43}$ 1533.5441; found 1553.5425.

Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (*p*LNnO, 41).—White powder, 17.6 mg, 96% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.21 (d, *J* = 3.8 Hz, 0.3H), 4.68 (dd, *J* = 8.3, 2.3 Hz, 3H), 4.65 (d, *J* = 8.0 Hz, 0.7H), 4.46 (dd, *J* = 10.8, 7.8 Hz, 3H), 4.42 (d, *J* = 7.9 Hz, 1H), 4.14 (q, *J* = 3.2, 2.4 Hz, 3H), 3.93 (td, *J* = 13.1, 12.7, 2.8 Hz, 5H), 3.85–3.69 (m, 29H), 3.68–3.61 (m, 3H), 3.59–3.51 (m, 8H), 3.29–3.24 (m, 0.5H), 2.02 (s, 9H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.86, 102.89, 102.85, 102.82, 102.72, 102.70, 95.70, 91.77, 82.03, 81.99, 81.97, 78.32, 78.21, 78.11, 75.31, 74.83, 74.76, 74.51, 74.31, 73.75, 72.46, 72.14, 72.13, 71.36, 71.09, 70.93, 70.08, 69.95, 69.92, 68.51, 68.30, 68.28, 60.99, 60.93, 60.92, 60.03, 59.90, 59.81, 55.15, 55.11, 22.14. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{54}H_{91}N_3NaO_{41}$ 1460.5026; found 1460.5006.

Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (*p*LNO, 42).—White powder, 17.8 mg, 96% yield. ^{1}H NMR (800 MHz, D_2O, 30 °C) δ 5.28 (d, *J* = 3.8 Hz, 0.3H), 4.80 (s, 1H), 4.76 (d, *J* = 6.4 Hz, 2H), 4.72 (d, *J* = 8.0 Hz, 0.7H), 4.53 (dd, *J* = 9.7, 7.8 Hz, 3H), 4.50 (dd, *J* = 7.8, 1.7 Hz, 1H), 4.21 (d, *J* = 3.1 Hz, 3H), 4.02–3.76 (m, 34H), 3.73–3.53 (m, 11H), 3.34 (t, *J* = 8.5 Hz, 0.6H), 2.07–2.11 (m, 9H). ^{13}C NMR (200 MHz, D_2O, 30 °C) δ 174.93, 103.52, 102.98, 102.95, 102.93, 102.75, 102.74, 102.58, 95.79, 91.86, 85.43, 82.18, 82.11, 82.07, 78.50, 78.40, 78.32, 75.40, 75.33, 75.25, 74.92, 74.85, 74.61, 74.41, 73.86, 72.58, 72.54, 72.23, 71.45, 71.20, 71.02, 70.74, 70.18, 70.05, 70.02, 68.61, 68.52, 68.38, 68.36, 61.07, 60.99, 60.17, 59.96, 55.26, 55.22, 54.75, 22.29, 22.24. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{54}H_{91}N_3NaO_{41}$ 1460.5026; found 1460.4964.

Fucα2Galβ3GlcNAcβ3Galβ4GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glc (Fucα1–2*p*LNO, 43).—White powder, 17.5 mg, 95% yield. ^{1}H NMR (600 MHz, D_2O, 30 °C) δ 5.21 (d, *J* = 3.8 Hz, 0.3H), 5.18 (d, *J* = 4.1 Hz, 1H), 4.69 (d, *J* = 8.3 Hz, 2H), 4.65 (dd, *J* = 9.9, 7.8 Hz, 1.5H), 4.61 (d, *J* = 8.4 Hz, 1H), 4.47–4.42 (m, 3H), 4.28 (q, *J* = 6.7 Hz, 1H), 4.14 (dd, *J* = 7.4, 3.2 Hz, 3H), 4.01–3.86 (m, 7H), 3.85–3.69 (m, 32H), 3.68–3.62 (m, 4H), 3.61–3.46 (m, 10H), 3.27 (t, *J* = 8.5 Hz, 0.5H), 2.05 (s, 3H), 2.02 (s, 6H), 1.23 (d, *J* = 6.6 Hz, 3H). ^{13}C NMR (150 MHz, D_2O, 30 °C) δ 174.87, 174.20, 103.22, 102.88, 102.86, 102.71, 100.22, 99.48, 98.39, 82.03, 81.58, 78.02, 77.12, 76.63, 75.21, 75.04, 74.83, 74.78, 74.76, 74.52, 74.32, 73.75, 73.45, 72.14, 71.37, 71.10, 69.93, 69.40, 69.12, 69.09,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

68.53, 68.43, 68.31, 68.01, 66.85, 66.46, 62.44, 61.12, 60.92, 60.35, 59.81, 55.12, 54.94, 22.14, 22.11, 15.23. HRMS (ESI-Orbitrap) m/z: $[M+Na]^+$ calculated for $C_{60}H_{101}N_3NaO_{45}$ 1606.5605; found 1606.5542.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Research reported in this publication was partially supported by the United States (U.S.) National Institutes of Health (NIH) Common Fund Glycoscience Program under Award Number U01GM120419. The Thermo Scientific Q Exactive HF Orbitrap Mass Spectrometer was purchased with a United States NIH Shared Instrumentation Grant (grant no. S10OD025271). The Bruker AVANCE-800 NMR spectrometer was purchased with a grant funded by the United States National Science Foundation (grant no. DBI-0722538).

References

[1]. Wu S, Snajdrova R, Moore JC, Baldenius K, Bornscheuer UT, Angew. Chem. Int. Ed 2021, 60, 88–119.

[2]. a)Sheldon RA, Woodley JM, Chem. Rev 2018, 118, 801–838; [PubMed: 28876904] b)Wohlgemuth R, Curr. Opin. Biotechnol 2010, 21, 713–724. [PubMed: 21030244]

[3]. Wohlgemuth R, Biocataly. Biotransform 2008, 26, 42–48.

[4]. a)Rakhra G, Jaiswal SK, Rakhra G, in Bioactive Natural Products in Drug Discovery (Eds.: Singh J, Meshram V, Gupta M), Springer Singapore, Singapore, 2020, pp. 335–353;b)Pallela R, Adv. Food Nutr. Res 2014, 73, 183–195. [PubMed: 25300547]

[5]. Xu LL, Townsend SD, J. Am. Chem. Soc 2021, 143, 11277–11290. [PubMed: 34296874]

[6]. a)Smilowitz JT, Lebrilla CB, Mills DA, German JB, Freeman SL, Annu. Rev. Nutr 2014, 34, 143–169; [PubMed: 24850388] b)Jenness R, Semin. Perinatol 1979, 3, 225–239; [PubMed: 392766] c)Chen X, Adv. Carbohydr. Chem. Biochem 2015, 72, 113–190; [PubMed: 26613816] d)Zhang S, Li T, Xie J, Zhang D, Pi C, Zhou L, Yang W, Microb. Cell. Fact 2021, 20, 108. [PubMed: 34049536]

[7]. a)German JB, Freeman SL, Lebrilla CB, Mills DA, Nestle Nutr. Workshop Ser. Pediatr. Program 2008, 62, 205–218; discussion 218–222;b)Kunz C, Adv. Nutr 2012, 3, 430S–439S. [PubMed: 22585922]

[8]. a)Bode L, Front. Pediatr 2018, 6, 385; [PubMed: 30564564] b)Sanchez C, Franco L, Regal P, Lamas A, Cepeda A, Fente C, Nutrients 2021, 13, 1026; [PubMed: 33810073] c)Hill DR, Chow JM, Buck RH, Nutrients 2021, 13, 3364; [PubMed: 34684364] d)Zhang B, Li LQ, Liu F, Wu JY, Carbohydr. Polym 2022, 276, 118738; [PubMed: 34823774] e)Morozov V, Hansman G, Hanisch FG, Schroten H, Kunz C, Mol. Nutr. Food Res 2018, 62, e1700679; [PubMed: 29336526] f)Newburg DS, Walker WA, Pediatr. Res 2007, 61, 2–8; [PubMed: 17211132] g)Bode L, Glycobiology 2012, 22, 1147–1162; [PubMed: 22513036] h)Ayechu-Muruzabal V, van Stigt AH, Mank M, Willemsen LEM, Stahl B, Garssen J, Van't Land B, Front. Pediatr 2018, 6, 239; [PubMed: 30250836] i)Holscher HD, Bode L, Tappenden KA, J. Pediatr. Gastroenterol. Nutr 2017, 64, 296–301; [PubMed: 28114245] j)Kirmiz N, Robinson RC, Shah IM, Barile D, Mills DA, Annu. Rev. Food Sci. Technol 2018, 9, 429–450. [PubMed: 29580136]

[9]. a)Yu H, Chen X, in Synthetic Glycomes, The Royal Society of Chemistry, 2019, pp. 254–280;b)Urashima T, Hirabayashi J, Sato S, Kobata A, Trends Glycosci. Glycotechnol 2018, 30, SE51–SE65.

[10]. Prudden AR, Liu L, Capicciotti CJ, Wolfert MA, Wang S, Gao Z, Meng L, Moremen KW, Boons GJ, Proc. Natl. Acad. Sci. U. S. A 2017, 114, 6954–6959. [PubMed: 28630345]

[11]. Bode L, Contractor N, Barile D, Pohl N, Prudden AR, Boons GJ, Jin YS, Jennewein S, Nutr. Rev 2016, 74, 635–644. [PubMed: 27634978]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[12]. a)Xiao Z, Guo Y, Liu Y, Li L, Zhang Q, Wen L, Wang X, Kondengaden SM, Wu Z, Zhou J, Cao X, Li X, Ma C, Wang PG, J. Org. Chem 2016, 81, 5851–5865; [PubMed: 27305319] b)Fang J-L, Tsai T-W, Liang C-Y, Li J-Y, Yu C-C, Adv. Synth. Catal 2018, 360, 3213–3219;c)Huang Y-T, Su Y-C, Wu H-R, Huang H-H, Lin EC, Tsai T-W, Tseng H-W, Fang J-L, Yu C-C, ACS Catal. 2021, 11, 2631–2643;d)Yu H, Li Y, Wu Z, Li L, Zeng J, Zhao C, Wu Y, Tasnima N, Wang J, Liu H, Gadi MR, Guan W, Wang PG, Chen X, Chem. Commun 2017, 53, 11012–11015;e)Wen L, Edmunds G, Gibbons C, Zhang J, Gadi MR, Zhu H, Fang J, Liu X, Kong Y, Wang PG, Chem. Rev 2018, 118, 8151–8187; [PubMed: 30011195] f)Zhao C, Wu Y, Yu H, Shah IM, Li Y, Zeng J, Liu B, Mills DA, Chen X, Chem. Commun 2016, 52, 3899–3902;g)McArthur JB, Yu H, Chen X, ACS Catal. 2019, 9, 10721–10726; [PubMed: 33408950] h)Bandara MD, Stine KJ, Demchenko AV, Org. Biomol. Chem 2020, 18, 1747–1753; [PubMed: 32048706] i)Bandara MD, Stine KJ, Demchenko AV, J. Org. Chem 2019, 84, 16192–16198; [PubMed: 31749363] j)Bandara MD, Stine KJ, Demchenko AV, Carbohydr. Res 2019, 486, 107824; [PubMed: 31585319] k)Bandara MD, Stine KJ, Demchenko AV, Carbohydr. Res 2019, 483, 107743; [PubMed: 31319351] l)Craft KM, Townsend SD, Carbohydr. Res 2017, 440–441, 43–50;m)Priem B, Gilbert M, Wakarchuk WW, Heyraud A, Samain E, Glycobiology 2002, 12, 235–240; [PubMed: 12042246] n)Zhang A, Sun L, Bai Y, Yu H, McArthur JB, Chen X, Atsumi S, Metab. Eng 2021, 66, 12–20; [PubMed: 33812022] o)Baumgartner F, Conrad J, Sprenger GA, Albermann C, Chembiochem 2014, 15, 1896–1900; [PubMed: 25044565] p)Baumgartner F, Sprenger GA, Albermann C, Enzyme Microb. Technol 2015, 75–76, 37–43;q)Baumgartner F, Jurzitza L, Conrad J, Beifuss U, Sprenger GA, Albermann C, Bioorg. Med. Chem 2015, 23, 6799–6806; [PubMed: 26481658] r)Dumon C, Samain E, Priem B, Biotechnol. Prog 2004, 20, 412–419. [PubMed: 15058985]

[13]. a)Li W, Ghosh T, Bai Y, Santra A, Xiao A, Chen X, Carbohydr. Res 2019, 479, 41–47; [PubMed: 31132641] b)Kooner AS, Diaz S, Yu H, Santra A, Varki A, Chen X, J. Org. Chem 2021, 86, 14381–14397. [PubMed: 34636559]

[14]. Norberg T, Kallin E, Blixt O, Carbohydr. Res 2021, 502, 108272. [PubMed: 33711724]

[15]. Yu H, Chen X, Org. Biomol. Chem 2016, 14, 2809–2818. [PubMed: 26881499]

[16]. Nidetzky B, Gutmann A, Zhong C, ACS Catal. 2018, 8, 6283–6300.

[17]. a)Ruzic L, Bolivar JM, Nidetzky B, Biotechnol. Bioeng 2020, 117, 1597–1602; [PubMed: 32017022] b)Schmölzer K, Weingarten M, Baldenius K, Nidetzky B, ACS Catal. 2019, 9, 5503–5514.

[18]. Sanchez-Garcia YI, Gutierrez-Mendez N, Salmeron I, Ramos-Sanchez VH, Leal-Ramos MY, Sepulveda DR, Food Res. Int 2021, 142, 110204. [PubMed: 33773679]

[19]. Li W, McArthur JB, Chen X, Carbohydr. Res 2019, 472, 86–97. [PubMed: 30529493]

[20]. Li Y, Yu H, Chen Y, Lau K, Cai L, Cao H, Tiwari VK, Qu J, Thon V, Wang PG, Chen X, Molecules 2011, 16, 6396–6407. [PubMed: 21799473]

[21]. Chen Y, Thon V, Li Y, Yu H, Ding L, Lau K, Qu J, Hie L, Chen X, Chem. Commun 2011, 47, 10815–10817.

[22]. Lau K, Thon V, Yu H, Ding L, Chen Y, Muthana MM, Wong D, Huang R, Chen X, Chem. Commun 2010, 46, 6066–6068.

[23]. Li Y, Xue M, Sheng X, Yu H, Zeng J, Thon V, Chen Y, Muthana MM, Wang PG, Chen X, Bioorg. Med. Chem 2016, 24, 1696–1705. [PubMed: 26968649]

[24]. a)Vasiliu D, Razi N, Zhang Y, Jacobsen N, Allin K, Liu X, Hoffmann J, Bohorov O, Blixt O, Carbohydr. Res 2006, 341, 1447–1457; [PubMed: 16650392] b)Peng W, Pranskevich J, Nycholat C, Gilbert M, Wakarchuk W, Paulson JC, Razi N, Glycobiology 2012, 22, 1453–1464. [PubMed: 22786570]

[25]. Chen M, Chen LL, Zou Y, Xue M, Liang M, Jin L, Guan WY, Shen J, Wang W, Wang L, Liu J, Wang PG, Carbohydr. Res 2011, 346, 2421–2425. [PubMed: 21903203]

[26]. Muthana MM, Qu J, Li Y, Zhang L, Yu H, Ding L, Malekan H, Chen X, Chem. Commun 2012, 48, 2728–2730.

[27]. Wakarchuk W, Martin A, Jennings MP, Moxon ER, Richards JC, J. Biol. Chem 1996, 271, 19166–19173. [PubMed: 8702594]

[28]. Yi W, Liu X, Li Y, Li J, Xia C, Zhou G, Zhang W, Zhao W, Chen X, Wang PG, Proc. Natl. Acad. Sci. U. S. A 2009, 106, 4207–4212. [PubMed: 19251666]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[29]. Xiao Z, Guo Y, Liu Y, Li L, Zhang Q, Wen L, Wang X, Kondengaden SM, Wu Z, Zhou J, Cao X, Li X, Ma C, Wang PG, J. Org. Chem 2016, 81, 5851–5865. [PubMed: 27305319]

[30]. Yu H, Li Y, Wu Z, Li L, Zeng J, Zhao C, Wu Y, Tasnima N, Wang J, Liu H, Gadi MR, Guan W, Wang PG, Chen X, Chem. Commun 2017, 53, 11012–11015.

[31]. Yu H, Karpel R, Chen X, Bioorg Med Chem 2004, 12, 6427–6435. [PubMed: 15556760]

[32]. Thon V, Li Y, Yu H, Lau K, Chen X, Appl. Microbiol. Biotechnol 2012, 94, 977–985. [PubMed: 22075637]

[33]. Xu Y, Fan Y, Ye J, Wang F, Nie Q, Wang L, Wang PG, Cao H, Cheng J, ACS Catal. 2018, 8, 7222–7227.

[34]. Li T, Liu L, Wei N, Yang J-Y, Chapla DG, Moremen KW, Boons G-J, Nat. Chem 2019, 11, 229–236. [PubMed: 30792508]

[35]. Zhang J, Liu D, Saikam V, Gadi MR, Gibbons C, Fu X, Song H, Yu J, Kondengaden SM, Wang PG, Wen L, Angew. Chem. Int. Ed 2020, 59, 19825–19829.

[36]. Zhang J, Chen C, Gadi MR, Gibbons C, Guo Y, Cao X, Edmunds G, Wang S, Liu D, Yu J, Wen L, Wang PG, Angew. Chem. Int. Ed 2018, 57, 16638–16642.

[37]. Hwang J, Yu H, Malekan H, Sugiarto G, Li Y, Qu J, Nguyen V, Wu D, Chen X, Chem. Commun 2014, 50, 3159–3162.

[38]. a)Hayashi Y, Chem. Sci 2016, 7, 866–880; [PubMed: 28791118] b)Hayashi Y, Acc. Chem. Res 2021, 54, 1385–1398. [PubMed: 33617234]

[39]. a)Yu H, Huang S, Chokhawala H, Sun M, Zheng H, Chen X, Angew. Chem. Int. Ed. Engl 2006, 45, 3938–3944; [PubMed: 16721893] b)Huynh N, Li Y, Yu H, Huang S, Lau K, Chen X, Fisher AJ, FEBS Lett. 2014, 588, 4720–4729. [PubMed: 25451227]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

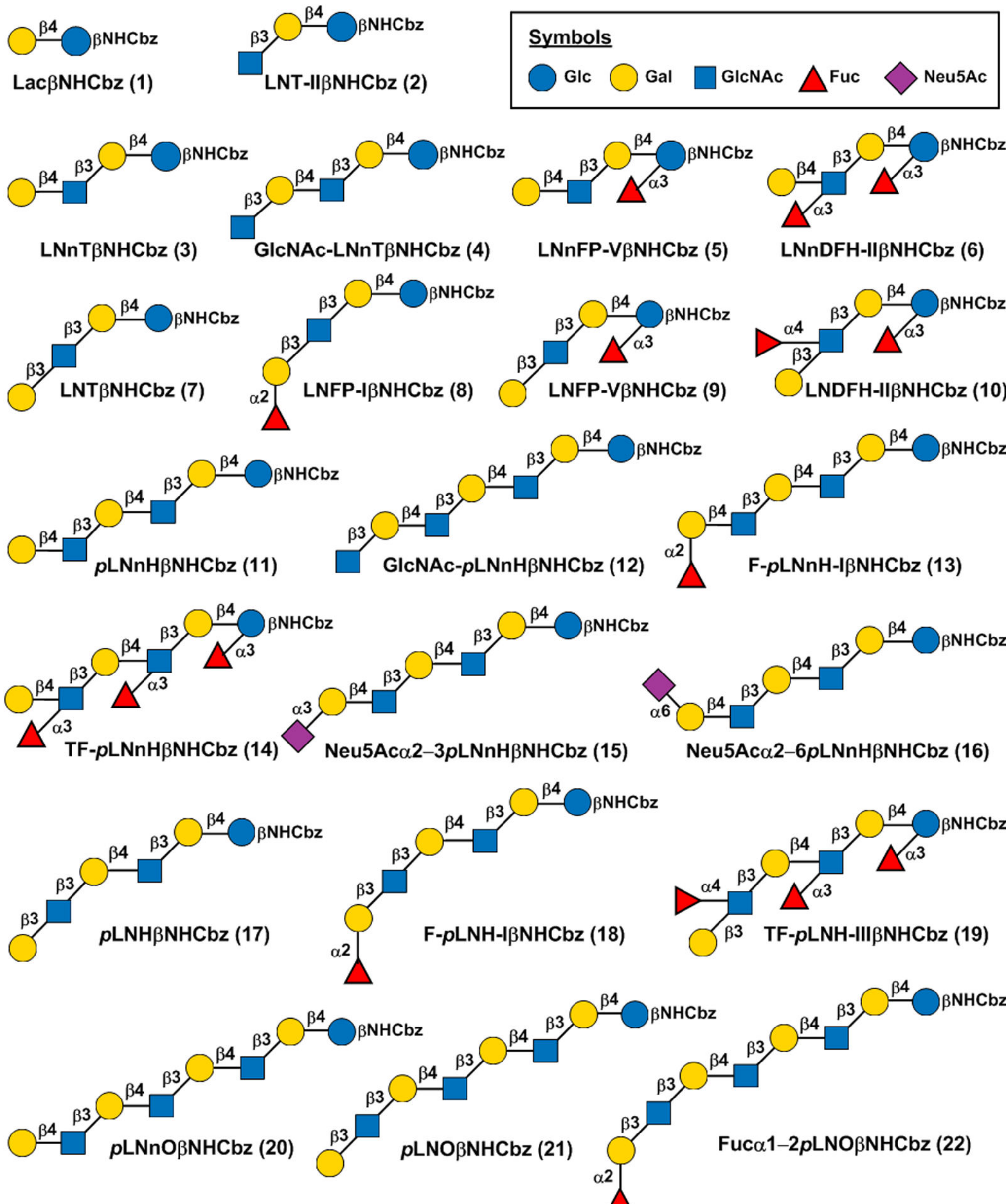


Figure 1.

The symbol nomenclatures for LacβNHCbz (**1**) and the βNHCbz-tagged HMO targets (**2–22**). Abbreviations: Glc, D-glucose; Gal, D-galactose; GlcNAc, *N*-acetyl-D-glucosamine; Fuc, L-fucose; Neu5Ac, *N*-acetyl-D-neuraminic acid.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Scheme 1.

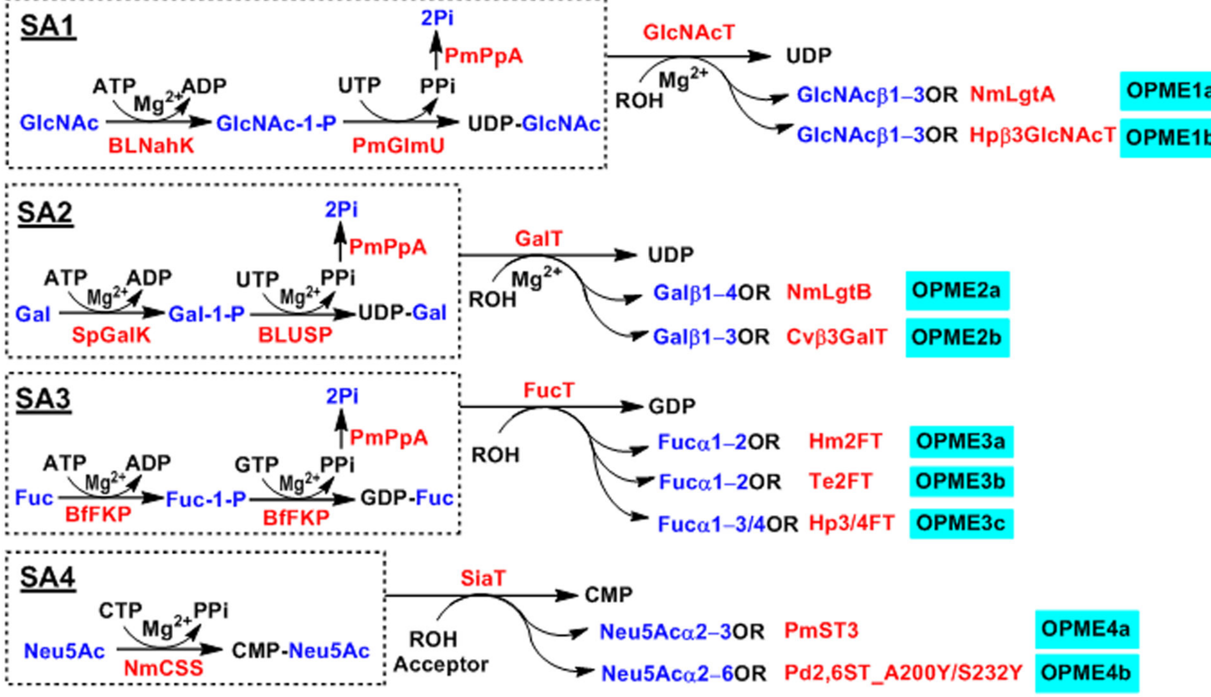
Schematic illustration of the combined substrate and process engineering strategies for chemoenzymatic synthesis and purification of complex human milk oligosaccharides (HMOs) with a free reducing end.

Scheme 2.

Glycosyltransferase acceptor substrate engineering by protecting group-free chemical tagging of lactose for the formation of LacβNHCbz (**1**) via a β-lactosylamine intermediate.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Scheme 3.

One-pot multienzyme (OPME) systems for synthesizing HMOs. These involve four monosaccharide sugar activation (SA) component (SA1–4).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Scheme 4.

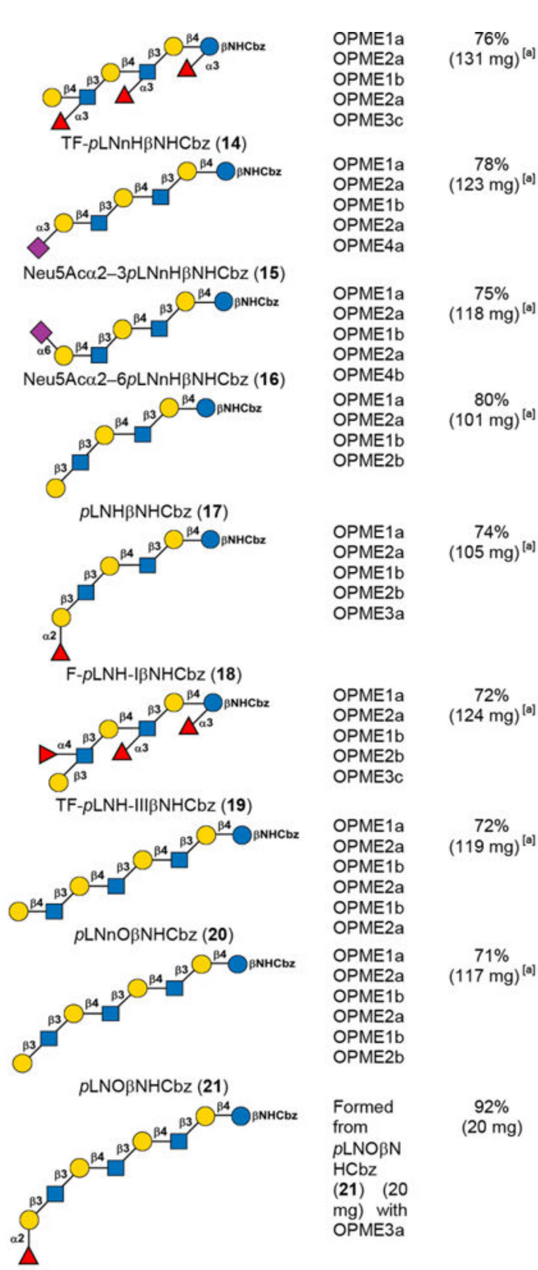
Catalytic hydrogenation of glycan-βNHCbz (**2–22**) followed by hydrolysis to form target HMOs (**23–43**) with a free reducing end.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Reaction processes and yields for the synthesis of target glycanβNHCbz (**2–22**).

Product	Process	Yield (amount)
LNT-IIβNHCbz (**2**)	OPME1a	91% (130 mg) [b]
LNnTβNHCbz (**3**)	OPME1a OPME2a	86% (152 mg) [b]
GlcNAc-LNnTβNHCbz (**4**)	OPME1a OPME2a OPME1b	83% (91 mg) [a]
LNnFP-VβNHCbz (**5**)	OPME1a OPME3c OPME2a	80% (83 mg)
LNnDFH-IIβNHCbz (**6**)	OPME1a OPME2a OPME3c	80% (94 mg) [a]
LNTβNHCbz (**7**)	OPME1a OPME2b	88% (155 mg) [b]
LNFP-IβNHCbz (**8**)	OPME1a OPME2b OPME3b	84% (1.73 g)[c]
LNFP-VβNHCbz (**9**)	OPME1a OPME3c OPME2b	81% (84 mg) [a]
LNDFH-IIβNHCbz (**10**)	OPME1a OPME3c OPME2b OPME3c	80% (95 mg) [a]
*p*LNnHβNHCbz (**11**)	OPME1a OPME2a OPME1b OPME2a	79% (100 mg) [a]
GlcNAc-*p*LNnHβNHCbz (**12**)	OPME1a OPME2a OPME1b OPME2a OPME1b	77% (114 mg) [a]
F-*p*LNnH-IβNHCbz (**13**)	OPME1a OPME2a OPME1b OPME2a OPME3a	76% (108 mg) [a]



Product	OPME modules	Yield
TF-*p*LNnHβNHCbz (**14**)	OPME1a OPME2a OPME1b OPME2a OPME3c	76% (131 mg) [a]
Neu5Acα2–3*p*LNnHβNHCbz (**15**)	OPME1a OPME2a OPME1b OPME2a OPME4a	78% (123 mg) [a]
Neu5Acα2–6*p*LNnHβNHCbz (**16**)	OPME1a OPME2a OPME1b OPME2a OPME4b	75% (118 mg) [a]
*p*LNHβNHCbz (**17**)	OPME1a OPME2a OPME1b OPME2b	80% (101 mg) [a]
F-*p*LNH-IβNHCbz (**18**)	OPME1a OPME2a OPME1b OPME2b OPME3a	74% (105 mg) [a]
TF-*p*LNH-IIIβNHCbz (**19**)	OPME1a OPME2a OPME1b OPME2b OPME3c	72% (124 mg) [a]
*p*LNnOβNHCbz (**20**)	OPME1a OPME2a OPME1b OPME2a OPME1b OPME2a	72% (119 mg) [a]
*p*LNOβNHCbz (**21**)	OPME1a OPME2a OPME1b OPME2a OPME1b OPME2b	71% (117 mg) [a]
Fucα1–2*p*LNOβNHCbz (**22**)	Formed from *p*LNOβNHCbz (**21**) (20 mg) with OPME3a	92% (20 mg)

Reactions were carried out using [a] 50 mg, [b] 100 mg, or [c] 1 gram of LacβNHCbz (**1**).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript