CHEMO-ENZYMATIC SYNTHESIS OF COMPLEX HUMAN MILK OLIGOSACCHARIDES

by

KUN YUAN

(Under the Direction of GEERT-JAN BOONS)

ABSTRACT

Human milk oligosaccharides (HMOs) are structurally diversified carbohydrates rich in the breast milk. They are the third largest component of human milk following lactose and lipids. Decades ago, these components were thought to play a role only as nutrient for intestinal microbiota with health benefits. With growing research into HMOs, they are found to exert additional beneficial properties. The structural similarities between cell surface glycans and HMOs enable pathogen recognition, preventing bacterial from attaching the epithelial cell surface. Although the biological function has been elucidated, multiple cases of breast-fed infants were observed to develop diarrhea in different stage of lactation. Whether variation in the HMO composition accounts for the loss of protective factors remains to be established.

To establish an integrated system of understanding functional and metabolic properties of individual HMOs, accessing well-defined structures is crucial. To address this challenge, chemo-enzymatic strategies have been developed to realize efficient synthetic strategy. For the first time a panel of defined HMOs are built with type I structure on β3 branch, displaying asymmetry and highly complex glycol-epitopes proposed to be

baits for glycan-binding proteins. Extensions of HMOs was achieved by engaging a range of glycosyl transferases based on specific regio- and stereo- selectivity. With the growing interest in how fucosylation patterns maneuver binding affinity, the substrate specificities of diverse FUTs (fucosyltransferases) have been thoroughly investigated on type I structures, elucidating the synthetic potential and limitation of such enzyme category. The modular enzymatic synthesis delineates patterns of substrates glycosyltransferases recognize, supplementing information of biosynthetic pathway of complex HMOs.

Our synthetic strategy was fulfilled with an auxiliary anomeric linker which could be used in HPLC purification and potential microarray printing. The linker is composed of a coumarin functional group which gives UV signal for HPLC purification and hydrophobicity, a lysine chain allowing the conjugated glycan to be mobilized on glass slides for subsequent microarray analysis.

INDEX WORDS: Glycans, Human Milk Oligosaccharides, Glycosyl Transferases, Chemoenzymatic Synthesis, Anomeric Linker, Carbohydrate

CHEMO-ENZYMATIC SYNTHESIS OF COMPLEX HUMAN MILK OLIGOSACCHARIDES

by

KUN YUAN

B.S., China Agricultural University, China, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfilment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

CHEMO-ENZYMATIC SYNTHESIS OF COMPLEX HUMAN MILK OLIGOSACCHARIDES

by

KUN YUAN

Major Professor: Geert-Jan Boons

Committee: Robert Phillips
Eric Ferreira

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
May 2020

DEDICATION

To my family and friends who have supported me along the way

through the bitter and the sweet.

ACKNOWLEDGEMENT

My profound gratitude goes first to Dr. Geert-Jan Boons, who has given me the opportunity to do research in such a highly motivated group, encouraged me to think out of the box and to jump outside of comfort zone. I would not have maintained belief in intrinsic worth of the carbohydrate research without his guidance. I would also like to extend my deepest thanks to Dr. Robert Phillips and Dr. Eric Ferreira for being my committee members. Both of them expanded my view of science and lighted the sparks of innovation.

There are numerous group members who have been extremely helpful throughout my research. Dr. Anthony Prudden played a key role in instructing me with all synthetic and analytical skills. He conveyed perceptions of multitasking and creative thinking which is proved to be a life-long fortune to me. Dr. Tiehai Li, who is a master of enzymatic synthesis, provided insightful suggestions for improvement of my project. Dr. Pradeep Chopra taught me to raise awareness and responsibilities of public affairs when working as a group member. Dr. John Glushka is always a magical NMR wizard who provided assistance on the equipment and spectrum interpretation. I am also grateful to Dr. Lin Liu for the help of HMO project.

Many thanks to all Boons group members who have shaped the group into a place where persistence and innovation are possible. Memory and knowledge I have acquired here will hold me in a long run.

Last but not the least, I am extremely grateful to my family and friends for their unparalled support and understanding for the past six years. They always back me up through up and down.

TABLE OF CONTENTS

LIST OF FIGURES

Page

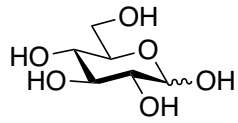
LIST OF SCHEMES

Page

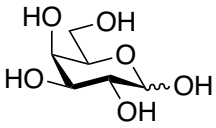
LIST OF TABLES

Page

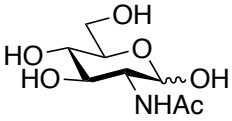
GLYCOBIOLOGY SYMBOLS



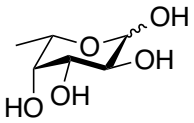
Glucose (Glc)



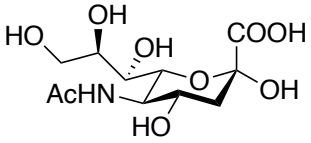
Galactose (Gal)



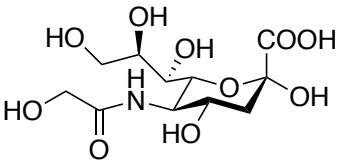
N-Acetylglucosamine (GlcNAc)



Fucose (Fuc)



Sialic Acid (Neu5Ac)



Sialic Acid (Neu5Ac)

LIST OF ABBREVIATIONS

Ac	Acetyl
Bn	Benzyl
Bz	Benzoyl
DCM	Dichloromethane
DMF	Dimethylformamide
DSLNT	Disialylated lacto-N-tetraose
ESI	Electrospray Ionization
EtOAc	Ethyl Acetate
EtOH	Ethanol
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Coherence Spectroscopy
LNT	Lacto-N-tetraose
LNnT	Lacto-N-neotetraose
M	Molarity
MALDI	Matrix-Assisted Laser Desorption Ionization
MSLNT	Mono-sialylated lacto-N-tetraose
MSLNnT	Mono-sialylated lacto-N-neotetraose
NMR	Nuclear Magnetic Resonance
NOSEY	Nuclear Overhauser Effect Spectroscopy
Ph	Phenyl

ppm	Parts per million
rt	Room temperature
TBDPS	tert-butyldiphenylsilyl
TDS	Dimethylthexylsilane
TFA	Trifluoroacetic Acid
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
TOCSY	Total Correlation Spectroscopy
Troc	Trichloroethoxycarbony

CHAPTER 1

HUMAN MILK OLIGOSACCHARIDES

Research Origin

The origin of human milk oligosaccharides research starts from the motivation of filling the gap of understanding the correlation between breast feeding and infant health. Not long after Escherich[1] discovered the close relationship between infant digestion and intestinal flora, Moro[2] and Tissier[3] spotted the differences in the bacterial composition in the feces of breast-fed and bottle-fed infants, although the functional components of breast milk haven't been defined in the next half century. In 1930, Polonowski and Lespagnol[4] identified an unknown nitrogen containing carbohydrate in human milk. This multi-component carbohydrate species named "gynolactose" has weak solubility in methanol. Followed by the discovery, the seperation of "gynolactose" was achieved using two-dimensional paper chromatography. Inspired by their work, Paul György, one of Moro's former students, collaborated with Richard Kuhn, confirming N-acetylglucosamine to be the growth-promoting factor for Bifidobacterium bifidus.[5,6]

With the help of purification technology, Richard Kuhn and Montreuil, together with their colleagues successfully extracted, isolated and characterized a series of HMOs, which later known as H and Lewis (Le) blood group epitopes.[7,8,9,10,11,12,13] By applying combinational tools of gel filtration and paper chromatography, Victor Ginsburg[14,15,16] and Akira Kobata[17,18,19,20,21] separated and characterized additional HMOs such as lacto-N-

fucopentaose, lacto-N-hexaose and lacto-N-neohexaose, contributing to the discovery of fingerprint of blood group glycans.

Human Milk Oligosaccharide structures

The patterns of oligosaccharides being investigated suggest the existence of genotype diversity between women and over the course of lactation.[22,23] The absolute and relative concentrations of HMOs also varied substantially among individuals. With the fact that more than a hundred of different HMOs being identified so far, the HMO composition appears to follow a defined pattern. HMOs are a set of carbohydrates composed of five momosaccharides, glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), fucose (Fuc) and sialic acid (Sia). All HMOs have lactose (Galβ1-4Glc) as their reducing end while the non-reducing end could be extended with either Galβ1-3GlcNAc (type I chain) or Galβ1-4GlcNAc (type II chain), making linear tetrasaccharides LNT (lacto-N-tetraose) and LNnT (lacto-N-neotraose). A β1-6-GlcNAc linkage could be installed at penultimate galactose of both tetrasaccharides, introducing chain branching to HMO structures. Lactose and elongated structures can be further fucosylated in α1-2, α1-3 or α1-4 fashion and/or sialylated in α2-3 or α2-6 fashion (Figure1.1).

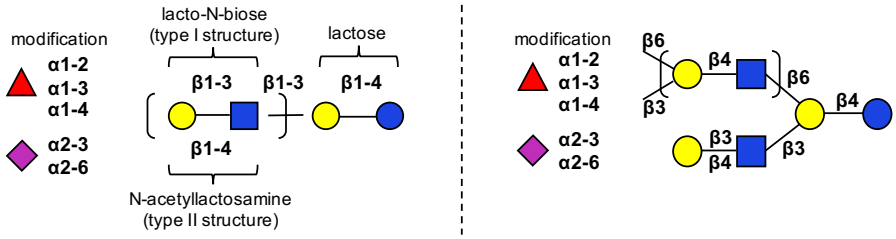


Figure 1.1 Linear and Branched Structures of HMOs

Diverse composition of HMOs also depicts the epitopes of Lewis (Le) blood group system. Four human milk groups can be built upon secretor types and Le blood system,

reflecting the activity of two gene loci which encoded two fucosyltransferase, FUT2 and FUT3.[24,25,26,27,28,29,30] Women with active secretors (Se) will express FUT2, producing α1-2 fucosylated HMOs rich milk. Lewis (Le) positive individuals express FUT3, giving α1-4 fucosylation. With research went deeper, more Se- and Le- independent FUTs have been spotted. Despite distinct function in fucose transfer, they also mutually affect each other by competing[24,26] or modular facilitation, creating a complicated scenario of HMO structures. (Figure 1.2)

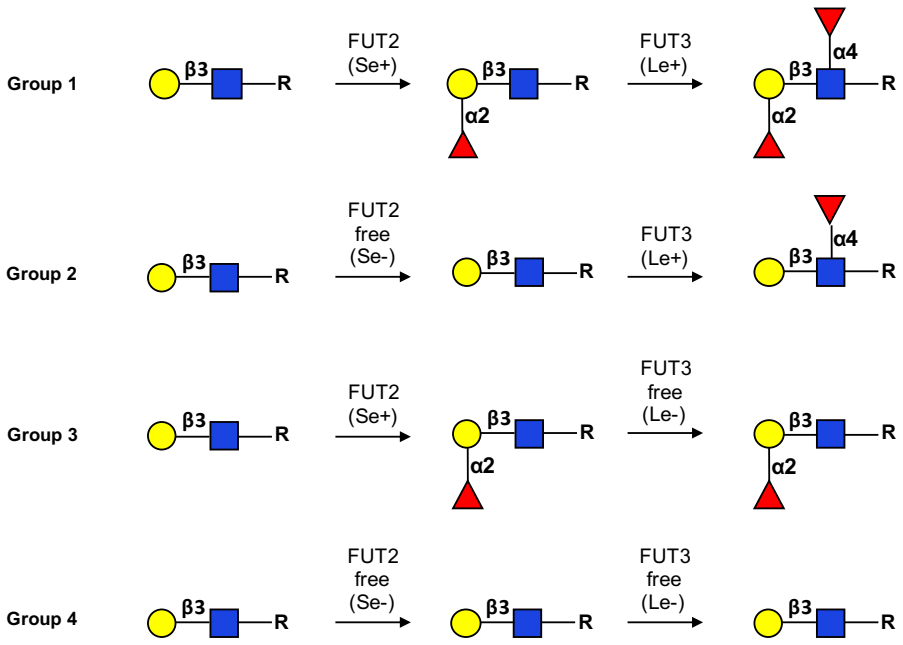


Figure 1.2 Fucosylation Patterns of HMOs based on Secrector(Se) and Lewis(Le) Status

Biosynthesis of HMOs

Although multiple HMO structures could be successfully isolated and characterized through past decades, the biosynthetic pathways of HMOs are proposed without solid proof on experimental data. Since all HMOs contains lactose at reducing end, the disaccharide is assumed to be the starting point. Lactose is synthesized in the Golgi apparatus. Under the catalytic reaction of the lactose synthase complex, UDP-

Gal is attached to glucose-1-P. The lactose synthase complex is composed of two proteins: β1-4 galactosyltransferase (β4GalT1) and α-lactalbumin. Only with the help of α-lactalbumin, which is only expressed under the control of lactation hormones in the mammary gland, could UGP-Gal be efficiently transferred to the glucose acceptor by β4GalT1.[31]

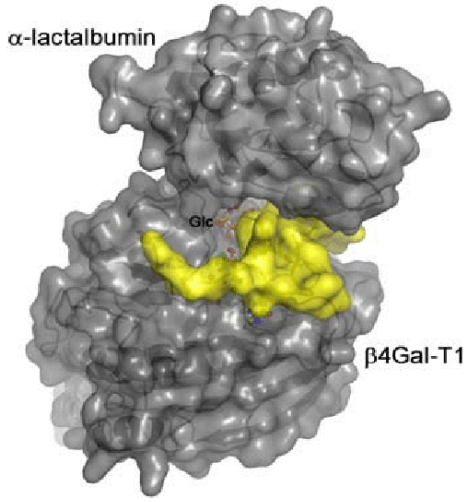


Figure 1.3 Crystal Structure of Lactose Synthase Complex (α-lactalbumin and β4GalT1)[81]

Hypothesis of HMO elongation and branching has been continuously proposed that concerted action of N-acetylglucosaminyltransferases (GlcNAcT) and GalT provide extension on lactose moiety.[32] A β3GlcNAcT (iGnT) commences linear chain elongation with GlcNAc while β6GlcNAcT (IGnT) initiate branching. Capping a Gal unit after GlcNAc moiety using β1-3-galactosyltransferase (β3GalT) will generate type 1 chain (Galβ1-3GlcNAc) while β1-4-galactosyltransferase would create type 2 chain (Galβ1-4GlcNAc). β6 GlcNAc moiety at the branching point also could be transglycosylated with Gal using β4GalT enzymes.

HMO fucosylation are regulated by the fucosyltransferases (FUTs) determined by Secretor (Se) and Le blood group system.[24,26,23,29] Women with positive secretor (Se+)

will produce FUT2, adding fucose to Gal of lactose in α1-2 fashion or to terminal Gal of LNT forming LNFP I. Women with positive Le (Le+) will create FUT3, giving α1-3- or α1-4- to GlcNAc. Four groups of fucosylated HMOs could be assigned base on different combinations of secretor and Le type. Cases could be further complicated with later research discovering more FUTs that is Se- and Le- independent but giving α1-3-/α1-4-fucosylation.[27,30,23,29,33]Secretor negative women have found to be capable of producing α1-2-fucosylation, suggesting FUT1 may also participate in HMO fucosylation.[34]

More questions remain unsettled as to the sialylation of HMOs. The GlcNAc in LNT could be sialylated in α2-6 fashion, which is rarely found in human tissue but some tumors and central nerve system.[35] In certain colon cancer cell lines, ST6GalNAc5 transfer sialic acid to penultimate GlcNAc.[35] However, the function of ST6GalNAc5 depends on the initial α2-3 sialylation on terminal Gal. Asialo-LNT cannot serve as substrate for STGalNAc5. Disialyl-LNT, one of HMOs, demonstrated the pattern. However, the existence of LST b in human body, which is composed of LNT with no α2-3 sialylation but with α2-6 sialylation, would indicate no involvement of ST6GalNAc5, suggesting other sialyltransferase (s) contributing to the biosynthesis of aforementioned structure. With the complete route of HMO biosynthesis not being clearly depicted, the synthetic pattern could be only rationalized by the information of enzymes obtained from in-vitro synthesis of limited scope of HMOs.

Intestinal microbiota shaping

Despite the fact that infants take in large amounts of HMOs during breastfeeding, human body cannot readily process such sugars. They dodge from degradation in the

upper parts of intestine, arriving at distal intestine where health-promoting bacteria[36,37] can consume HMOs through a unique, effective central fermentative pathway called bifid shunt.[38,39,40] Such dietary glycans are termed as prebiotics which, by modifying intestinal microbes, provide health benefits to neonates.[41,42] *Genus Bifidobacterium* was one of the most studied gastrointestinal microbiota species.[43,44] The gene sequencing revealed that Bifidobacterium can metabolically adapt to the complex carbohydrate-rich gastrointestinal tract environment as they encode a series of glycosyl hydrolases, sugar transporters and PEP-PTS(PEP—phosphoenolpyruvate; PTS—phosphotransferase system) components which are required for metabolism of plant- and host-derived carbohydrates.[45] The metabolic activities influence physiology of host as metabolites and end products of fermentation are absorbed by epithelial layer or transit to liver through portal vein. Most Bifidobacteria spp. induced α-L-fucosidase activity, producing abundant short-chain fatty acids, lactate and lactodifucotetraose.[46] Moderate growth of a series of *Bifidobacterium longum* species were stimulated when offering 3'-sialyllactose (3'-SL) and 6'-SL.

For the following decades, more microbiota have been found to grow upon feeding of HMOs. Across bacteria tested, the correlation between induction of fucosidase (2'-FL utilization) and production of lactic acid were significant. Specific *Lactobacillus delbrueckii*, *Enterococcus faecalis* and *Streptococcus thermophilus* species exhibit slight growth on 2'-FL or 3-FL instead of LDFT. *Bacteroides vulgatus* shows great neuraminidase activity, producing abundant lactate and SCFA. All *Bifidobacteria* spp. and *Bacteroides* spp. induce considerable amount of α-L-fucosidase, consuming 40%

3-FL. With the ascending counting of *Bifidobacteria* spp. colony, significant reduction in pH was observed.

A 9:1 ratio of the galactosyloligosaccharides and fructosyloligosaccharides comprises current prebiotic in infant formula.[47] The amount of both bifidobacteria and lactobacilli in the microbiota boost when fed to animals and humans. The prebiotic activity, however, is different from that shown by HMOs. 2'-FL, 3-FL and 6'-FL stimulate bifids but not effective on lactobacilli. Same condition applies to *Enterococcus faecalis* and *S. thermophilus*. Pioneering work has conclusively driven to an evolutionary adaptation through which a metabolic system could be established by providing advantageous growth condition on certain microbiota over the others competing for the same HMOs complex.

Cell-surface carbohydrate receptor analogue

To invade and infect the host, bacterial or viral bind to mucosal surfaces through adhesion proteins, lectins, from pathogen binding to specific carbohydrate receptors of cell surface.[48,34] The glycan-binding interaction has been described for viruses like noroviruses and rotaviruses by Hu et al.[49] Both streams of viruses are notorious for being the reasons of infant diarrhea. Through mimicking mucosal cell surface glycans, HMOs act as soluble decoys preventing pathogens from binding intestinal epithelial cells and reducing infections.

One of the most thoroughly studied case of HMO as antiadhesive agent is *Campylobacter jejuni* infection.[50,51] At the stage of mice study, α1-2 fucosylated HMOs diminish *C. jejuni* colonization. Similar conditions are also observed with human intestinal mucosa *ex vivo*. A cohort study carried out in Mexico City indicates that

mothers who excreted milk with high concentrations of 2'FL endow kids with immune potential, contributing to less cases of *C. jejuni* inducing diarrhea.[51] Further investigations witness the reducing cases of calicivirus diarrhea also correlates with abundance of α1-2 fucosylated HMOs, lacto-N-difucohexaose I.

Many microorganisms, instead of using proteins to bind to host glycol-epitopes, express glycans that bind to proteins on host cells. Gp120, envelope protein of human immunodeficiency virus (HIV) binds to DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin) on dendritic cells distributed on mucosal surface which sense external stimuli and mount appropriate response. This pathway is believed to be an essential initial route of HIV infection in mother-to-child transmission.[52,53] Subsequent credentials were also given that DC-SIGN binds to high-mannose glycans on gp120 whereas Naarding and Van Liempt deifined that DC-SIGN has higher affinity to Le blood group antigen.[54,55,56] Since HMOs contains Le blood group antigens, they may compete with gp120 binding to DC-SIGN. Investigation into HIV-infected mothers in Lusaka, Zambia showed that the non-sialylated, fucosylated HMOs (2'-FL, LNFP I), as well as non-2-linked fucosylated HMOs (3-FL, LNFP II/III) are found to be protective against mortality during breastfeeding (not after weaning) in HIV-exposed uninfected (HEU) children.[57] This may give a postulation as to why mother-to-child HIV transmission through breast feeding is inefficient even infants are constantly exposed to virus-containing breast milk. However, 3'FL in breast milk shows no sign of benefit for HEU mortality. Pioneering work even indicated that HIV-infected women with low CD4 counts have higher concentration of non-fucosylated HMOs, including LNT and 3'-SL.

Higher percentage of 3' FL was associated with increased risk of postnatal HIV transmission.

Immuno-Modulation:

Beside modulating the infant health through gastrointestinal beneficial microbiota shaping and pathogen binding prevention, HMOs also stimulate innate and adaptive immune development directly. Sialylated HMOs are reported to be internalized into blood stream where they systemically affect the maturation and binding of monocytes, lymphocytes to endothelial cells, as well as the formation of platelet-neutrophil complexes. Pioneering work by Kunz reported HMOs such as LNFP I, LNFP II and difucosyllacto-N-hexose are found in the urine of preterm breastfed infants, suggesting around 1% of HMOs are absorbed and reached systemic circulation.[58,59] A pilot study by Jantscher-Krenn et al. investigated individual variation of HMOs composition and concentration over time in maternal serum during pregnancy.[60] HMO concentration ascended with gestational age. The carbohydrate profile altered from sialylated species dominate at V1 (gestational weeks 10-14) to a balance between fucosylated and sialylated species at V3 (gestational weeks 30-35 weeks) owing to the increase of 2'-FL. The result confirmed the presence of HMOs in maternal serum, suggesting biological potential for HMOs during pregnancy. In 2019, more inspiring observation came as HMOs are confirmed to exist in cord blood. In order to study maternal-to-fetal transport pathway, Hirschmugl et al. carried out perfusion of isolated placental cotyledons from term pregnancies with 2'-FL. The result showed 18 oligosaccharides exist in maternal serum in all cord serum samples.[61] Moreover, the cord blood HMO concentration kept at the same level with that in maternal serum. 2'-FL and LDFT stands out in the HMO composition of both cord blood sample

and maternal serum. Such detail overhauled the image of human “milk” oligosaccharides, recapitulating the critical health-related potential HMOs present to both mothers and fetus.

Current research lacks the relevant evidence on the receptors and signaling pathways of HMO-induced effect on macrophage and lymphocyte cytokine production. Lectins are carbohydrate-binding proteins which are classified according to their carbohydrate recognition domains (CRD).[62,63] Three types of lectins are suspected to be related to the influence of HMO on immune response. C-type lectin contains DC-SIGN which has a CRD specific for fucose moiety, mediating leukocyte trafficking and platelet-neutrophil complex (PNC) formation.[64] Siglecs (sialic acid-binding lectins) are commonly expressed by diverse leukocyte populations such as CD22 and CD33. Through binding to siglecs, sialic acid containing species have access to macrophages. Bounding to specific siglecs stimulates production of immunoregulatory cytokine interleukin-10 (IL-10). Galectins are reported to engage in cell turnover and immunoregulation.[65] Many HMOs species are β-galactosides containing β-1,3- or β-1,4 Gal at non-reducing end, which target the galectin binding. More investigations are required as to whether HMOs directly modulate siglec or galectin-engaged immune response.

Sustainability of HMOs

Based on the pioneering research that breast-fed infants have lower cases of infection in gastrointestinal, respiratory and urogenital tracts compared with formula-fed infants, HMOs must have been served as either decoy for pathogens or block binding for harmful bacteria. To maintain the bioactivity, HMOs must resist degradation in the upper gastrointestinal tract. Engfer et al. conducted in vitro digestion studies of HMOs through which fractions of carbohydrate species were incubated using human pancreatic juice

and intestinal brush border membranes (BBMs) prepared from human or porcine intestinal tissue sample.[66] HMO species are purified through gel permeation chromatography and separated into three categories: neutral HMOs, acidic HMOs and lactose fractions. When subjected to digestion, FL (2'-FL and 3-FL), LNT, LNnT, LNFP and LNDFH were recovered intact with no hydrolysis products detected by MALDI-MS. As a positive control for determination of hydrolysis, same batch of HMOs were subject to digestion using a porcine pancreatic tissue homogenate containing zymogens and intracellular lysosomal and enzymes. With the help of full profile of digestive agents, HMOs were digested rapidly within 4 hours. Another control of digestibility was also included through which maltodextrin was rapidly and completely hydrolyzed. Free glucose was detected using enzymatic determination after the preincubation with Human duodenal aspirates (HDAs). Complete cleavage was observed in 2 hours after BBMs addition, indicated by the boost of glucose concentration. The result excluded the possibility of bacteria-origin enzymatic activity in HMOs hydrolysis model as they could have used glucose.

Beside the stability of HMOs in human gut, these glycans are proved to be sustainable during different types of pasteurization and lyophilization.[67] HMOs profile remain unaffected based on the analysis of MALDI TOF/TOF mass spectrometry. However, the length of storage time requires more investigation in further studies.

General methods for obtaining HMOs

Given the physiological benefit of HMOs, growing research and industrial interest has been shown. Although more than 200 milk glycans have been described to date, data from mass spectrometry suggests up to 900 structures may exist. Driven by the need of

thorough investigation over structures and potential of infant formula development, synthetic groups have dedicated towards chemical and enzymatical synthesis of defined HMOs structures. Biological groups have designed engineered bacteria to produce HMOs on larger scale for biological and modulation investigation.

Milk extraction: General strategies to extract HMOs in large quantities are using charcoal column chromatography and nanofiltration. Both methods require fat and protein being removed in the first place through centrifugation and precipitation. Then lactose was enzymatically converted into glucose and galactose to ease separation. Brand-Miller et al. used charcoal column to retain oligosaccharides while monosaccharide elute out.[68] Sarney et al. compared normal gel filtration and nanofiltration. The speed of separation is fast and engage no organic solvent. However, it doesn't allow further separation of oligosaccharides fractions.[69]

Academic lab acquired more HMOs through filtration strategies during a study of galectin binding microarray of 247 human milk glycans.[70] The breast milk of secretor positive and Lewis negative mothers was pooled. Lipids and lactose were removed through centrifugation and precipitation, followed by gel filtration through G-25 Sephadex size exclusion chromatography. To further divide HMOs into groups by charge, DEAE-cellulose chromatography was also employed. HPLC purification and mass spectrometry verification managed to isolate individual glycan for subsequent microarray slides which is readily arranged for probing the binding interaction and affinity of specific HMO species.

Chemical and Enzymatic Synthesis: With the development of chemical synthetic strategies and the production of efficient recombinant enzymes, a pool of structurally defined HMOs is anticipated to be constructed either through combinational

(chemoenzymatic strategy) or separate act. Furuike et al. chemically synthesized a branched tetrasaccharide with an 1,6-anhydro-β-N-acetyl lactosamine reducing terminus with Gal as branching point extended by β1,6 and β1,3 GlcNAc.[71] The aforementioned unit was recognized by bovine β1,4 galactosyltransferase which added two Gal. The octasaccharide was obtained by final addition of two sialic acid in α2,3 fashion on both branches utilizing α2,3 sialyltransferase.

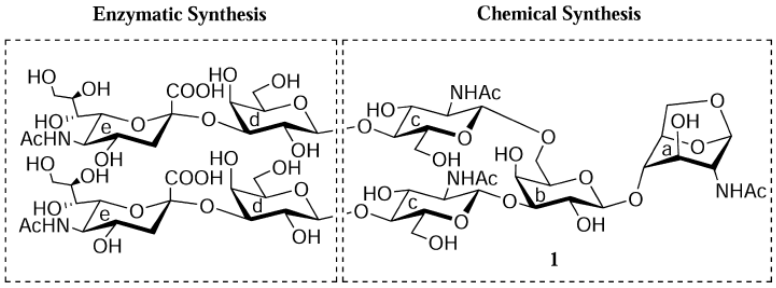


Figure 1.4 Chemoenzymatic Synthesis of a Biantennary Sialyllactosaminoglycan Model

By using a specific enzyme with reversible nature of catalysis, D-galactosyl-β1,3-D-hexosamine phosphorylase (BiGalHexNAc-P) from Bifidobacterium infantis, Yao et al. have designed a strategy in which α-Gal-1-phosphate could be transglycosylated onto ethyl-β-D-2-deoxy-trifluoroacetamido-1-thio-glucoside acceptor.[72] The obtained protected Lacto-N-biose(LNB) moiety will be transferred to a lactose acceptor exposing a free 3-OH chemically, giving a fully protected tetrasacchaide, which after removal of all protecting groups, affording Lacto-N-tetraose (LNT). Subsequent α2,3 sialylation produced sialyllacto-N-tetraose (LSTa) in considerable amount.

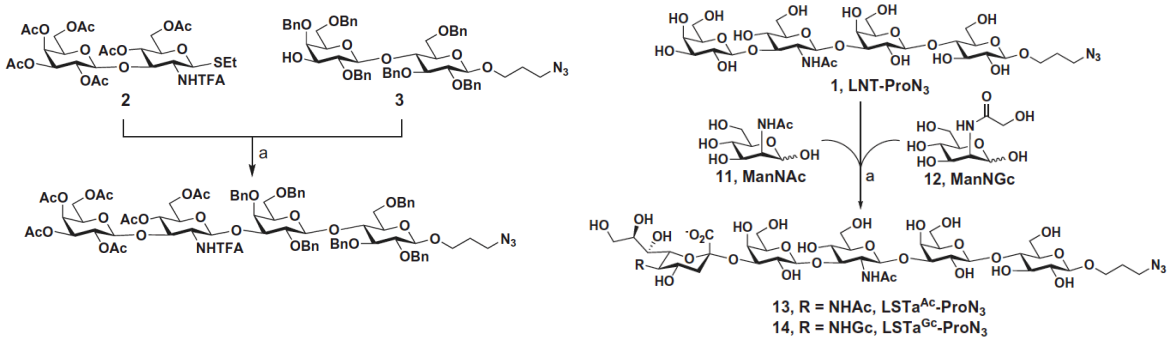


Figure 1.5 Synthesis of LNT-ProN_3 and LSTa-ProN_3

Chen et al. designed a one-pot multienzyme sequential glycan building strategy.[73] The synthesis began with chemically synthesized lactoside bearing an azido linker, starting from which multiple kinases, epimerases, uridylyltransferases, and glycosyltransferases were employed to synthesize lacto-N-neotetraose (LNnT) over two steps. Subsequent α2,3 and α2,6 sialylation engaged enzyme from *Pasteurella multocida* and *Photobacterium damselae* respectively.

Pudden et al. described synthesis of a library of HMOs bearing a cleavable anomeric linker.[74] This hydrophobic linker not only served as a handle for rapid purification of enzymatic reaction products through reverse-phase HPLC, but also could be immobilized on glass chip for subsequent microarray study of protein-glycan binding interaction.

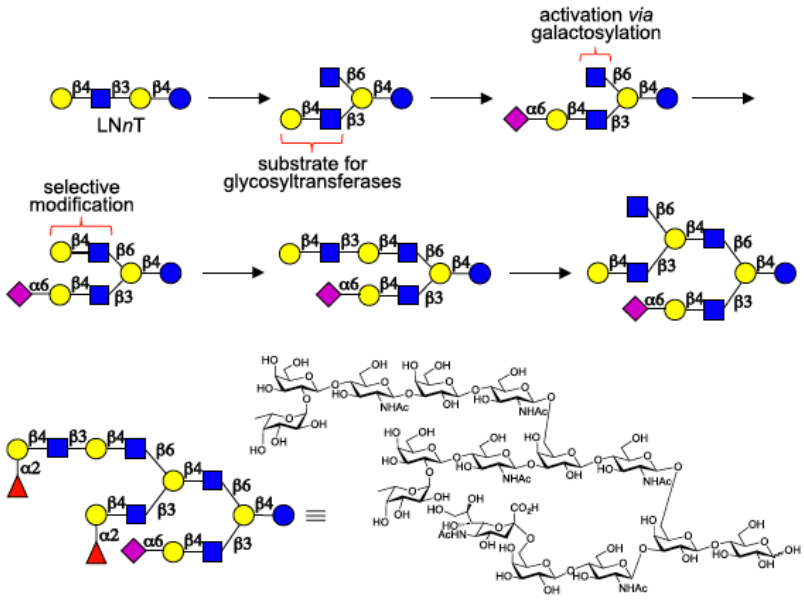


Figure 1.6 Enzymatic Synthesis of Asymmetric Multiantennary HMOs

Chemoenzymatic synthesis is an efficient strategy developed to circumvent the setback of pure chemical or enzymatic synthesis. Both yield and efficiency of synthesis would be largely promoted through combinational method. With the help of chemoenzymatic synthesis, an expanding library of structure-defined HMOs could be obtained with clear patterns of regio- and stereo-selectivity. Moreover, unnatural chemically modified carbohydrate moieties may have potential of being incorporated into certain glycosylation pattern using enzymes. The tolerance of synthetic sugar nucleotides by natural or engineering glycosyltransferases is also a boosting topic in antibody-drug conjugate (ADC) research.

Metabolic Pathway Engineering: As an alternative to extraction and large-scale chemical synthesis, biosynthetic pathway of HMO could be re-engineered in proper recombinant microbial organisms. Sugar nucleotides which are fundamental for glycosyltransfer reactions are reported to be provided in organism like non-pathogenic E. coli strains. In E. coli, the glycosyltransferase substrate, UDP-GlcNAc and UDP-Gal are

constantly produced by cell since they are essential precursors of peptidoglycan and lipopolysaccharides-core respectively. However, GDP-fucose and sialic acid show no sign of not constantly exist in E. coli strains.[75]

2'-FL is one of the most abundant HMOs in secretor-positive mothers. With the potential role 2'-FL played in infants health uncovered, several groups have reported 2'-FL production in recombinant E. coli.[76,77,78,79] The prerequisites for production of 2'-FL are the presence of intracellular lactose, GDP-fucose and corresponding FUT. E. coli has a nature of taking up lactose by lactose permease, LacY. However, another enzyme also degrades lactose, termed β-galactosidase, LacZ. To elevate the concentration of lactose, LacY+LacZ- strains are preferred. Even GDP-fucose is a natural intermediate of *E. colis* colanic acid biosynthesis, the diversity of biosynthetic pathways still needs to be explored. Expression of *fkp* gene from *B. fragilis* in *E. coli* results in the uptake of exogeneous L-fucose, creating GDP-fucose. Through enhanced expression of the guanosine-inosine kinase or by increasing the supply of NADPH, GDP-fucose production could also be boosted. Since a native α1,2 or α1,3 FUT which takes lactose hasn't been reported in *E. coli* K-12 strains, the exogenous FUT from *helicobacter pylori* is introduced. Hereby a complete metabolic pathway to 2'-FL could be established and decent yields have been reported.

The production of LNT and LNnT require GlcNAc and Gal moieties to be transfer sequentially onto lactose in certain pattern. This glycosyl-transfering reaction needs activated sugar species, UDP-GlcNAc and UDP-Gal, however, E. coli strains lacks the ability to express β1,3-N-acetylglucosaminyltransferase, β1,3- and β1,4-galactosyltransferases. Therefore, *lgtA* genes from N. meningitis which encode the

enzyme activity of transferring GlcNAc were expressed using plasmids in lacZ- strain. This enzyme affords trisaccharide LNT II structure. Subsequent expression of *lgtB* from same species encoding β1,4-galactosyltransferase gives the formation of LNnT structure. For LNT biosynthesis, Baumgärtner *et al.* reported a whole-cell biotransformation in which both genes, lgtA and wbgO (encoding regio- and stereo-selective β1,3-galactosyltransferase from E. coli O55:H7) were chromosomally integrated simultaneously, avoiding the potential of instability of strains proposed by *lgtA* and *lgtB* two plasmid system.[77] Compared with the high titer of LNT II in *lgtA* system, the introduction of *wbgO* gene doesn't give decent transformation from trisaccharide LNT II to tetrasaccharide LNT. The outcome suggests either a limitation of WbgO activity or insufficient UDP-Gal supply. Since LNT II still dominates the extracellular product by ~50%, more investigation should be conducted on whether WbgO activity might be the reason of poor conversion.

In all, the biosynthetic methodology with recombinant cells serve as a supplement to chemical or chemoenzymatic synthesis scale-wise, delivering large quantity of samples for biological study. However, establishing the profile of complex HMO structures are currently hampered by the method since the limitation of cooperation between enzymes has not been clearly depicted. Branched HMOs are only available through milk extraction, chemical or chemoenzymatic synthesis.

Conclusion:

Ever since the discovery of human milk oligosaccharides (HMOs), they are believed to be engaged in biological activities, providing nutrition to infants. Subsequent progress in biological studies indicated HMOs serve as prebiotics, supporting the growth of health-

beneficial gut microbiota. With the increasing interest in HMO research, HMO species and structures are expected to be defined and classified into categories. Contemporary advanced analytical technology allows for the extraction and separation of HMOs from crude breast milk. However, the quantity doesn't guarantee further research into scale-up biological studies. Chemoenzymatic synthesis strategy and re-engineering of recombinant bacteria streams have fulfilled the routes for quantitative production of structure-defined HMOs with decent yields. The optimization of HMOs synthetic ways will provide us with the potential use in infant formula, preventing babies from developing gastrointestinal disease.

References:

(1) Escherich, T. *Die Darmbakterien Des Säuglings Und Ihre Beziehungen Zur Physiologie Der Verdauung*; F. Enke, 1886.

(2) Moro, E. Morphologie Und Bakteriologische Untersuchungen Über Die Darmbakterien Des Säuglings: Die Bakterien-Flora Des Normalen Frauenmilchstuhls. *Jahrb. Kinderh* **1900**, *61*, 686–734.

(3) Tissier, H. Recherches Sur La Flora Intestinale de Nourissons (État Normal et Pathologique). *Paris, Fr.* **1900**.

(4) Polonowski, M.; Lespagnol, A. Sur l'existence de Plusieurs Glucides Dans Le Lactoserum de Femme. *CR Soc Biol* **1930**, *104*, 555–557.

(5) György, P.; Kuhn, R.; Rose, C. S.; Zilliken, F. Bifidus Factor. II. Its Occurrence in Milk from Different Species and in Other Natural Products. *Arch. Biochem. Biophys.* **1954**, *48* (1), 202–208.

(6) György, P.; Norris, R. F.; Rose, C. S. Bifidus Factor. I. A Variant of Lactobacillus Bifidus Requiring a Special Growth Factor. *Arch. Biochem. Biophys.* **1954**, *48* (1), 193–201.

(7) Kuhn, R.; Baer, H. H. Die Konstitution Der Lacto-N-tetraose. *Chem. Ber.* **1956**, *89* (2), 504–511.

(8) Kuhn, R.; Baer, H. H.; Gauhe, A. Kristallisation Und Konstitutionsermittlung Der Lacto-N-fucopentaose I. *Chem. Ber.* **1956**, *89* (11), 2514–2523.

(9) Kuhn, R.; Baer, H. H.; Gauhe, A. Die Konstitution Der Lacto-N-fucopentaose II. Ein

Beitrag Zur Spezifität Der Blutgruppensubstanz Lea. *Chem. Ber.* **1958**, *91* (2), 364–374.

(10) Kuhn, R.; Gauhe, A. Über Ein Kristallisiertes, Lea-aktives Hexasaccharid Aus Frauenmilch. *Chem. Ber.* **1960**, *93* (3), 647–651.

(11) Kuhn, R.; Gauhe, A. Über Die Lacto-Difuco-Tetraose Der Frauenmilch. Ein Beitrag Zur Strukturspezifität Der Blutgruppensubstanz Le. *Justus Liebigs Ann. Chem.* **1958**, *611* (1), 249–253.

(12) Kuhn, R.; Gauhe, A. Über Drei Saure Pentasaccharide Aus Frauenmilch. *Chem. Ber.* **1962**, *95* (2), 513–517.

(13) Montreuil, J. The Glucides of Milk. *Bull. Soc. Chim. Biol. (Paris).* **1960**, *42*, 1399–1427.

(14) Grollman, E. F.; Ginsburg, V. Correlation between Secretor Status and the Occurrence of 2'-Fucosyllactose in Human Milk. *Biochem. Biophys. Res. Commun.* **1967**, *28*, 50–53.

(15) Grollman, E. F.; Kobata, A.; Ginsburg, V. An Enzymatic Basis for Lewis Blood Types in Man. *J. Clin. Invest.* **1969**, *48* (8), 1489–1494.

(16) Grollman, E. F.; Kobata, A.; Ginsburg, V. Enzymatic Basis of Blood Types in Man. *Ann. N. Y. Acad. Sci.* **1970**, *169* (1), 153–158.

(17) Kobata, A.; Ginsburg, V. Oligosaccharides of Human Milk II. Isolation and Characterization of a New Pentasaccharide, Lacto-N-Fucopentaose III. *J. Biol. Chem.* **1969**, *244* (20), 5496–5502.

(18) Kobata, A.; Ginsburg, V. Oligosaccharides of Human Milk III. Isolation and Characterization of a New Hexasaccharide, Lacto-N-Hexaose. *J. Biol. Chem.* **1972**,

247 (5), 1525–1529.

(19) Kobata, A.; Ginsburg, V. Oligosaccharides of Human Milk: IV. Isolation and Characterization of a New Hexasaccharide, Lacto-N-Neohexaose. *Arch. Biochem. Biophys.* **1972**, *150* (1), 273–281.

(20) Kobata, A.; Ginsburg, V.; Tsuda, M. Oligosaccharides of Human Milk: I. Isolation and Characterization. *Arch. Biochem. Biophys.* **1969**, *130*, 509–513.

(21) Kobata, A. Structures and Application of Oligosaccharides in Human Milk. *Proc. Japan Acad. Ser. B* **2010**, *86* (7), 731–747.

(22) Xu, G.; Davis, J. C.; Goonatilleke, E.; Smilowitz, J. T.; German, J. B.; Lebrilla, C. B. Absolute Quantitation of Human Milk Oligosaccharides Reveals Phenotypic Variations during Lactation. *J. Nutr.* **2017**, *147* (1), 117–124. https://doi.org/10.3945/jn.116.238279.

(23) Chaturvedi, P.; Warren, C. D.; Altaye, M.; Morrow, A. L.; Ruiz-Palacios, G.; Pickering, L. K.; Newburg, D. S. Fucosylated Human Milk Oligosaccharides Vary between Individuals and over the Course of Lactation. *Glycobiology* **2001**, *11* (5), 365–372. https://doi.org/10.1093/glycob/11.5.365.

(24) Kumazaki, T.; Yoshida, A. Biochemical Evidence That Secretor Gene, Se, Is a Structural Gene Encoding a Specific Fucosyltransferase. *Proc. Natl. Acad. Sci.* **1984**, *81* (13), 4193–4197.

(25) Viverge, D.; Grimmonprez, L.; Cassanas, G.; Bardet, L.; Solere, M. Discriminant Carbohydrate Components of Human Milk According to Donor Secretor Types. *J. Pediatr. Gastroenterol. Nutr.* **1990**, *11* (3), 365–370.

(26) Johnson, P. H.; Watkins, W. M. Purification of the Lewis Blood-Group Gene

Associated α-3/4-Fucosyltransferase from Human Milk: An Enzyme Transferring Fucose Primarily to Type 1 and Lactose-Based Oligosaccharide Chains. *Glycoconj. J.* **1992**, *9* (5), 241–249.

(27) Thurl, S.; Henker, J.; Siegel, M.; Tovar, K.; Sawatzki, G. Detection of Four Human Milk Groups with Respect to Lewis Blood Group Dependent Oligosaccharides. *Glycoconj. J.* **1997**, *14* (7), 795–799.

(28) Chaturvedi, P.; Warren, C. D.; Buescher, C. R.; Pickering, L. K.; Newburg, D. S. Survival of Human Milk Oligosaccharides in the Intestine of Infants. In *Bioactive components of human milk*; Springer, 2001; pp 315–323.

(29) Stahl, B.; Thurl, S.; Henker, J.; Siegel, M.; Finke, B.; Sawatzki, G. Detection of Four Human Milk Groups with Respect to Lewis-Bloodgroup-Dependent Oligosaccharides by Serologic and Chromatographic Analysis. In *Bioactive Components of Human Milk*; Springer, 2001; pp 299–306.

(30) Thurl, S.; Munzert, M.; Henker, J.; Boehm, G.; Müller-Werner, B.; Jelinek, J.; Stahl, B. Variation of Human Milk Oligosaccharides in Relation to Milk Groups and Lactational Periods. *Br. J. Nutr.* **2010**, *104* (9), 1261–1271.

(31) Ramakrishnan, B.; Boeggeman, E.; Qasba, P. K. β-1, 4-Galactosyltransferase and Lactose Synthase: Molecular Mechanical Devices. *Biochem. Biophys. Res. Commun.* **2002**, *291* (5), 1113–1118.

(32) Kobata, A. Possible Application of Milk Oligosaccharides for Drug Development. *Chang Gung Med. J.* **2003**, *26* (9), 621–636.

(33) Gabrielli, O.; Zampini, L.; Galeazzi, T.; Padella, L.; Santoro, L.; Peila, C.; Giuliani, F.; Bertino, E.; Fabris, C.; Coppa, G. V. Preterm Milk Oligosaccharides during the

First Month of Lactation. *Pediatrics* **2011**, *128* (6), e1520–e1531.

(34) Newburg, D. S.; Ruiz-Palacios, G. M.; Morrow, A. L. Human Milk Glycans Protect Infants against Enteric Pathogens. *Annu. Rev. Nutr.* **2005**, *25*, 37–58.

(35) Tsuchida, A.; Okajima, T.; Furukawa, K.; Ando, T.; Ishida, H.; Yoshida, A.; Nakamura, Y.; Kannagi, R.; Kiso, M.; Furukawa, K. Synthesis of Disialyl Lewis a (Lea) Structure in Colon Cancer Cell Lines by a Sialyltransferase, ST6GalNAc VI, Responsible for the Synthesis of α-Series Gangliosides. *J. Biol. Chem.* **2003**, *278* (25), 22787–22794.

(36) Gibson, G. R.; Probert, H. M.; Van Loo, J.; Rastall, R. A.; Roberfroid, M. B. Dietary Modulation of the Human Colonic Microbiota: Updating the Concept of Prebiotics. *Nutr. Res. Rev.* **2004**, *17* (2), 259–275.

(37) Roberfroid, M. Prebiotics: The Concept Revisited. *J. Nutr.* **2007**, *137* (3), 830S-837S.

(38) Fushinobu, S. Unique Sugar Metabolic Pathways of Bifidobacteria. *Biosci. Biotechnol. Biochem.* **2010**, 1010282237.

(39) Devika, N. T.; Raman, K. Deciphering the Metabolic Capabilities of Bifidobacteria Using Genome-Scale Metabolic Models. *Sci. Rep.* **2019**, *9* (1), 1–9.

(40) González-Rodríguez, I.; Gaspar, P.; Sánchez, B.; Gueimonde, M.; Margolles, A.; Neves, A. R. Catabolism of Glucose and Lactose in Bifidobacterium Animalis Subsp. Lactis, Studied by 13C Nuclear Magnetic Resonance. *Appl. Environ. Microbiol.* **2013**, *79* (24), 7628–7638.

(41) NEWBURG, D. S.; HE, Y. Neonatal Gut Microbiota and Human Milk Glycans Cooperate to Attenuate Infection and Inflammation. *Clin. Obstet. Gynecol.* **2015**, *58*

(4), 814–826.

(42) Barile, D.; Rastall, R. A. Human Milk and Related Oligosaccharides as Prebiotics. *Curr. Opin. Biotechnol.* **2013**, *24* (2), 214–219.

(43) Sgorbati, B.; Biavati, B.; Palenzona, D. The Genus Bifidobacterium. In *The genera of lactic acid bacteria*; Springer, 1995; pp 279–306.

(44) Biavati, B.; Castagnoli, P.; Trovatelli, L. D. Species of the Genus Bifidobacterium in the Feces of Human Adults. *Microbiologica* **1986**, *9* (1), 39–45.

(45) Turroni, F.; Strati, F.; Foroni, E.; Serafini, F.; Duranti, S.; van Sinderen, D.; Ventura, M. Analysis of Predicted Carbohydrate Transport Systems Encoded by Bifidobacterium Bifidum PRL2010. *Appl. Environ. Microbiol.* **2012**, *78* (14), 5002–5012.

(46) Yu, Z.-T.; Chen, C.; Newburg, D. S. Utilization of Major Fucosylated and Sialylated Human Milk Oligosaccharides by Isolated Human Gut Microbes. *Glycobiology* **2013**, *23* (11), 1281–1292.

(47) Bruzzese, E.; Volpicelli, M.; Salvini, F.; Bisceglia, M.; Lionetti, P.; Cinquetti, M.; Iacono, G.; Guarino, A. Early Administration of GOS/FOS Prevents Intestinal and Respiratory Infections in Infants. *J. Pediatr. Gastroenterol. Nutr.* **2006**, *42* (5), E95.

(48) Kunz, C.; Rudloff, S.; Baier, W.; Klein, N.; Strobel, S. Oligosaccharides in Human Milk: Structural, Functional, and Metabolic Aspects. *Annu. Rev. Nutr.* **2000**, *20* (1), 699–722.

(49) Hu, L.; Crawford, S. E.; Czako, R.; Cortes-Penfield, N. W.; Smith, D. F.; Le Pendu, J.; Estes, M. K.; Prasad, B. V. V. Cell Attachment Protein VP8* of a Human Rotavirus Specifically Interacts with A-Type Histo-Blood Group Antigen. *Nature*

2012, *485* (7397), 256–259.

(50) Ruiz-Palacios, G. M.; Cervantes, L. E.; Ramos, P.; Chavez-Munguia, B.; Newburg, D. S. Campylobacter Jejuni Binds Intestinal H (O) Antigen (Fucα1, 2Galβ1, 4GlcNAc), and Fucosyloligosaccharides of Human Milk Inhibit Its Binding and Infection. *J. Biol. Chem.* **2003**, *278* (16), 14112–14120.

(51) Morrow, A. L.; Ruiz-Palacios, G. M.; Altaye, M.; Jiang, X.; Guerrero, M. L.; Meinzen-Derr, J. K.; Farkas, T.; Chaturvedi, P.; Pickering, L. K.; Newburg, D. S. Human Milk Oligosaccharides Are Associated with Protection against Diarrhea in Breast-Fed Infants. *J. Pediatr.* **2004**, *145* (3), 297–303.

(52) Baan, E.; de Ronde, A.; Luchters, S.; Vyankandondera, J.; Lange, J. M.; Pollakis, G.; Paxton, W. A. HIV Type 1 Mother-to-Child Transmission Facilitated by Distinctive Glycosylation Sites in the Gp120 Envelope Glycoprotein. *AIDS Res. Hum. Retroviruses* **2012**, *28* (7), 715–724.

(53) Scarlatti, G. Mother-to-Child Transmission of HIV-1: Advances and Controversies of the Twentieth Centuries. *AIDS Rev* **2004**, *6* (2), 67–78.

(54) Naarding, M. A.; Ludwig, I. S.; Groot, F.; Berkhout, B.; Geijtenbeek, T. B. H.; Pollakis, G.; Paxton, W. A. Lewis X Component in Human Milk Binds DC-SIGN and Inhibits HIV-1 Transfer to CD4+ T Lymphocytes. *J. Clin. Invest.* **2005**, *115* (11), 3256–3264.

(55) van Liempt, E.; Bank, C. M. C.; Mehta, P.; Garcı´ a-Vallejo, J. J.; Kawar, Z. S.; Geyer, R.; Alvarez, R. A.; Cummings, R. D.; Kooyk, Y. van; van Die, I. Specificity of DC-SIGN for Mannose-and Fucose-containing Glycans. *FEBS Lett.* **2006**, *580* (26), 6123–6131.

(56) Hong, P.; Ninonuevo, M. R.; Lee, B.; Lebrilla, C.; Bode, L. Human Milk Oligosaccharides Reduce HIV-1-Gp120 Binding to Dendritic Cell-Specific ICAM3-Grabbing Non-Integrin (DC-SIGN). *Br. J. Nutr.* **2008**, *101* (4), 482–486.

(57) Kuhn, L.; Kim, H.-Y.; Hsiao, L.; Nissan, C.; Kankasa, C.; Mwiya, M.; Thea, D. M.; Aldrovandi, G. M.; Bode, L. Oligosaccharide Composition of Breast Milk Influences Survival of Uninfected Children Born to HIV-Infected Mothers in Lusaka, Zambia. *J. Nutr.* **2015**, *145* (1), 66–72.

(58) Rudloff, S.; Pohlentz, G.; Diekmann, L.; Egge, H.; Kunz, C. Urinary Excretion of Lactose and Oligosaccharides in Preterm Infants Fed Human Milk or Infant Formula. *Acta Paediatr.* **1996**, *85* (5), 598–603.

(59) Bode, L. Human Milk Oligosaccharides: Every Baby Needs a Sugar Mama. *Glycobiology* **2012**, *22* (9), 1147–1162. https://doi.org/10.1093/glycob/cws074.

(60) Jantscher-Krenn, E.; Aigner, J.; Reiter, B.; Köfeler, H.; Csapo, B.; Desoye, G.; Bode, L.; Van Poppel, M. N. M. Evidence of Human Milk Oligosaccharides in Maternal Circulation Already during Pregnancy-a Pilot Study. *Am. J. Physiol. Metab.* **2018**.

(61) Hirschmugl, B.; Brandl, W.; Csapo, B.; van Poppel, M.; Köfeler, H.; Desoye, G.; Wadsack, C.; Jantscher-Krenn, E. Evidence of Human Milk Oligosaccharides in Cord Blood and Maternal-to-Fetal Transport across the Placenta. *Nutrients* **2019**, *11* (11), 2640.

(62) Crocker, P. R.; Paulson, J. C.; Varki, A. Siglecs and Their Roles in the Immune System. *Nat. Rev. Immunol.* **2007**, *7* (4), 255–266.

(63) Koliwer-Brandl, H.; Siegert, N.; Umnus, K.; Kelm, A.; Tolkach, A.; Kulozik, U.; Kuballa, J.; Cartellieri, S.; Kelm, S. Lectin Inhibition Assays for the Analysis of

Bioactive Milk Sialoglycoconjugates. *Int. dairy J.* **2011**, *21* (6), 413–420.

(64) Osborn, L. Leukocyte Adhesion to Endothelium in Inflammation. *Cell* **1990**, *62* (1), 3–6.

(65) Rabinovich, G. A.; Baum, L. G.; Tinari, N.; Paganelli, R.; Natoli, C.; Liu, F.-T.; Iacobelli, S. Galectins and Their Ligands: Amplifiers, Silencers or Tuners of the Inflammatory Response? *Trends Immunol.* **2002**, *23* (6), 313–320.

(66) Engfer, M. B.; Stahl, B.; Finke, B.; Sawatzki, G.; Daniel, H. Human Milk Oligosaccharides Are Resistant to Enzymatic Hydrolysis in the Upper Gastrointestinal Tract. *Am. J. Clin. Nutr.* **2000**, *71* (6), 1589–1596.

(67) Daniels, B.; Coutsoudis, A.; Autran, C.; Amundson Mansen, K.; Israel-Ballard, K.; Bode, L. The Effect of Simulated Flash Heating Pasteurisation and Holder Pasteurisation on Human Milk Oligosaccharides. *Paediatr. Int. Child Health* **2017**, *37* (3), 204–209.

(68) Brand-Miller, J. C.; McVeagh, P.; McNeil, Y.; Messer, M. Digestion of Human Milk Oligosaccharides by Healthy Infants Evaluated by the Lactulose Hydrogen Breath Test. *J. Pediatr.* **1998**, *133* (1), 95–98.

(69) Sarney, D. B.; Hale, C.; Frankel, G.; Vulfson, E. N. A Novel Approach to the Recovery of Biologically Active Oligosaccharides from Milk Using a Combination of Enzymatic Treatment and Nanofiltration. *Biotechnol. Bioeng.* **2000**, *69* (4), 461–467.

(70) Yu, Y.; Lasanajak, Y.; Song, X.; Hu, L.; Ramani, S.; Mickum, M. L.; Ashline, D. J.; Prasad, B. V. V.; Estes, M. K.; Reinhold, V. N. Human Milk Contains Novel Glycans That Are Potential Decoy Receptors for Neonatal Rotaviruses. *Mol. Cell. Proteomics* **2014**, *13* (11), 2944–2960.

(71) Furuike, T.; Yamada, K.; Ohta, T.; Monde, K.; Nishimura, S.-I. An Efficient Synthesis of a Biantennary Sialooligosaccharide Analog Using a 1, 6-Anhydro-β-Lactose Derivative as a Key Synthetic Block. *Tetrahedron* **2003**, *59* (27), 5105–5113.

(72) Yao, W.; Yan, J.; Chen, X.; Wang, F.; Cao, H. Chemoenzymatic Synthesis of Lacto-N-Tetrasaccharide and Sialyl Lacto-N-Tetrasaccharides. *Carbohydr. Res.* **2015**, *401*, 5–10.

(73) Chen, C.; Zhang, Y.; Xue, M.; Liu, X.; Li, Y.; Chen, X.; Wang, P. G.; Wang, F.; Cao, H. Sequential One-Pot Multienzyme (OPME) Synthesis of Lacto-N-Neotetraose and Its Sialyl and Fucosyl Derivatives. *Chem. Commun.* **2015**, *51* (36), 7689–7692.

(74) Prudden, A. R.; Liu, L.; Capicciotti, C. J.; Wolfert, M. A.; Wang, S.; Gao, Z.; Meng, L.; Moremen, K. W.; Boons, G.-J. Synthesis of Asymmetrical Multiantennary Human Milk Oligosaccharides. *Proc. Natl. Acad. Sci.* **2017**, *114* (27), 6954–6959.

(75) Sprenger, G. A.; Baumgärtner, F.; Albermann, C. Production of Human Milk Oligosaccharides by Enzymatic and Whole-Cell Microbial Biotransformations. *J. Biotechnol.* **2017**, *258*, 79–91.

(76) Chin, Y.-W.; Kim, J.-Y.; Lee, W.-H.; Seo, J.-H. Enhanced Production of 2′-Fucosyllactose in Engineered Escherichia Coli BL21star (DE3) by Modulation of Lactose Metabolism and Fucosyltransferase. *J. Biotechnol.* **2015**, *210*, 107–115.

(77) Baumgärtner, F.; Seitz, L.; Sprenger, G. A.; Albermann, C. Construction of Escherichia Coli Strains with Chromosomally Integrated Expression Cassettes for the Synthesis of 2′-Fucosyllactose. *Microb. Cell Fact.* **2013**, *12* (1), 40.

(78) Lee, W.-H.; Pathanibul, P.; Quarterman, J.; Jo, J.-H.; Han, N. S.; Miller, M. J.; Jin,

Y.-S.; Seo, J.-H. Whole Cell Biosynthesis of a Functional Oligosaccharide, 2′-Fucosyllactose, Using Engineered Escherichia Coli. *Microb. Cell Fact.* **2012**, *11* (1), 48.

(79) Drouillard, S.; Driguez, H.; Samain, E. Large-Scale Synthesis of H-Antigen Oligosaccharides by Expressing Helicobacter Pylori A1, 2-Fucosyltransferase in Metabolically Engineered Escherichia Coli Cells. *Angew. Chem. Int. Ed.* **2006**, *45* (11), 1778–1780.

(80) Bode, L. Human Milk Oligosaccharides: Every Baby Needs a Sugar Mama. *Glycobiology* **2012**, *22* (9), 1147–1162.

(81) Ramakrishnan, B.; Qasba, P. UDP-Gal: BetaGlcNAc Beta 1,4-Galactosyltransferase, Polypeptide 1 (B4GALT1). In *Handbook of Glycosyltransferases and Related Genes, Second Edition*; 2014; pp 51–62. https://doi.org/10.1007/978-4-431-54240-7_110.

CHAPTER 2

CHEMOENZYMATIC SYNTHESIS OF HMO GLYCOEPITOPES AND SUBSTRATE SPECIFICITIES OF DIVERSE FUCOSYLTRANSFERASES

Introduction:

Human milk oligosaccharides (HMOs) are a structurally diverse class of unconjugated glycans. They have been recognized for their role in infant health since the late nineteenth century.[1] Observational studies indicate that certain HMO from secretors are associated with various preventive effects, such as reducing diarrhea and promoting intestinal maturation in preterm neonates. Human milk oligosaccharides (HMOs) consumed by infants serve as substrates for potentially beneficial bacteria which may reduce the inflammation and GI infections and by modulating the immune epithelial cell responses.[2] HMOs function as a prebiotic helping to establish commensal bacteria and as anti-adhesives that help prevent the attachment of microbial pathogens to mucosal surfaces.

Of all carbohydrate epitopes that have been implicated in cell-cell recognition and membrane attachment, the Lewis glycoepitope family plays an important role in facilitating such manner.[3,4] Bacterial colonization and following inflammation could partially be impacted by Lewis (a) blood group antigens. Incidence of diarrhea was observed comparatively higher among children with Le (a+b-) blood group infected with ETEC (*Enterotoxigenic Escherichia coli*) expressing the CFA/I (colonization factor antigen I) than those with Le (a-b+), suggesting capability of bacteria binding to Lewis a terminated

glycoconjugate. Presumably the expression of carbohydrates modified by Lewis structure is associated with different risks of enteric infections.

Being main ligands for selectins, Lewis antigens allow leukocytes tethering and rolling along the endothelial wall awaiting transmembrane trafficking. Thus, upon inflammatory stimulation, pro-inflammatory factors will activate epithelial cells and mobilize selectins to the cell surface, helping to recruit leukocytes for the following transmigration. Concerning the biological significance of Lewis glycoepitopes, biological synthetic pathways have been developed. Catalyzed by widely available β3GNT2 and βGalT1, lactose can efficiently be converted into Lacto-N-neotetraose (LNnT) with type II LacNAc (Galβ1-4GlcNAc) residing at the non-reducing end. A panel of LNnT based structure can subsequently be fabricated by applying diverse glycosyltransferases with corresponding sugar nucleotides affording Le (x), Le (y) and sLe (x) terminated structures. Compared with the highly efficient enzymatic synthesis of type I chain, efficacy of enzymes required for constructing structures terminated with type I LacNAc (Galβ1-3GlcNAc) in the synthesis lacto-N-tetraose (LNT) is low (20% yield), which hampered the expansion of enzymatic synthesis pathway. LNT is the core structure of type I (Galβ1-3GlcNAc) HMOs which is abundant in human breast milk.[5] Chemical strategy for LNT notoriously employs laborious protecting and decaging processes.[6, 7]Although emerging chemoenzymatic synthesis gave entry to convergent strategies, using one-pot two-enzyme has suspicious repeatability towards thiol species, resulting in an overall discouraging methodology. Thus, there is an unmet need of exploiting alternative synthetic route for LNT which will be employed in the late-stage modification into diversified oligosaccharides.

To address aforementioned difficulties, we envisaged the nature of enzyme to be employed during chemoenzymatic synthesis and adopted TFA derivatized glycosyl donor for chemical glycosylation. The LNT backbone could be obtained by sequential removal of orthogonal protecting groups of tetrasaccharide **7**. With LNT at hand, it was anticipated that LNT can be further selectively modified with different patterns of fucosylation and/or sialylation using various fucosyltransferases and sialyltransferases, giving a series of Lewis structures. Guillermo and colleagues[6] have illustrated that α1,2-fucosylated HMO moieties containing H2 blood group epitope (Fucα1-2-Galβ1-4-GlcNAc) were capable of preventing the adherence of *C. jejuni* [7] to epithelial cells *in vitro*.

With a table of fucosyltransferases readily accessible, synthetic attempts were made towards LNT-based structures affording anticipated Le^a and sLe^a, meanwhile obtaining unexpected difucosylation at 3-OH of glucose at non-reducing end in compounds terminated with both type I and type II LacNAc. Over-fucosylation suggests that glucose within the chain (LNT and LNnT) also serves as a substrate for relevant fucosyltransferases. A reasonable deduction could be drawn that monofucosylation at GlcNAc position brought conformational change to the compounds which also fits in the active sites and led to the subsequent catalysis overfucosylation on Glucose. Screenings of LNT, LNnT and corresponding sialylated backbones with an array of fucosyltransferases[8] give insights into substrate specificities and uncover how glycan similarities and primary products controls catalytical activities in unanticipated manners.

Results and discussion

Synthesis of Lacto-N-Tetraose

Previously established microbial and coupled enzymatic approach only provides limited access to sufficient quantities of LNT. Yao *et al.*[9] have described a practical pathway in which lacto-N-biose donor and lactoside acceptor building blocks are used for LNT synthesis. Upon the earlier stage of reconfirming the aforementioned route, a thioglycoside was found not favored by the glycosyltrasferase BiGalHexNacP which catalyze the galactosylation reaction. Thus, we were compelled to explore whether an overall synthesis of LNT can be optimized.

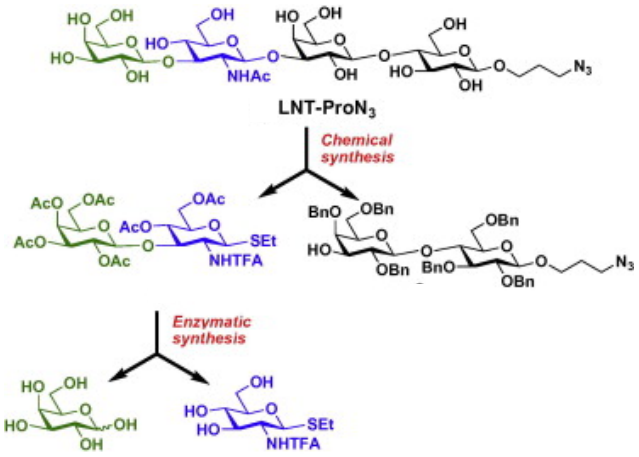
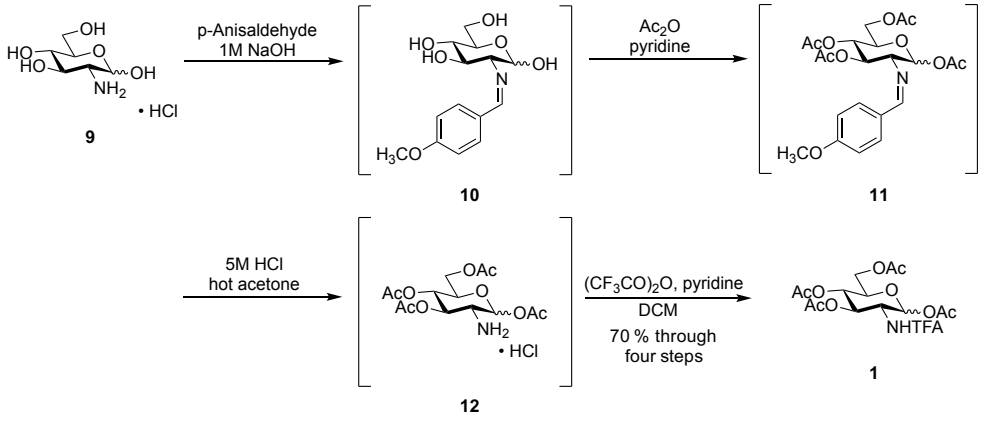


Figure 2.1 Retrosynthetic Analysis of LNT-ProN39

Herein a convergent synthetic route for LNT has been designed without a linker for the concern of late-stage enzymatic manipulation. The lacto-N-biose donor will be created in the form of a TFA oxazoline instead of traditional thioglycoside. In contrast to the sole exposure of 3-OH on galactose in acceptor moiety in traditional strategy, both 3-OH and 4-OH of galactose were set free for chemical glycosylation aiming to achieve regio- and stereoselectivities.

Synthesis of oxazoline glycosyl donor

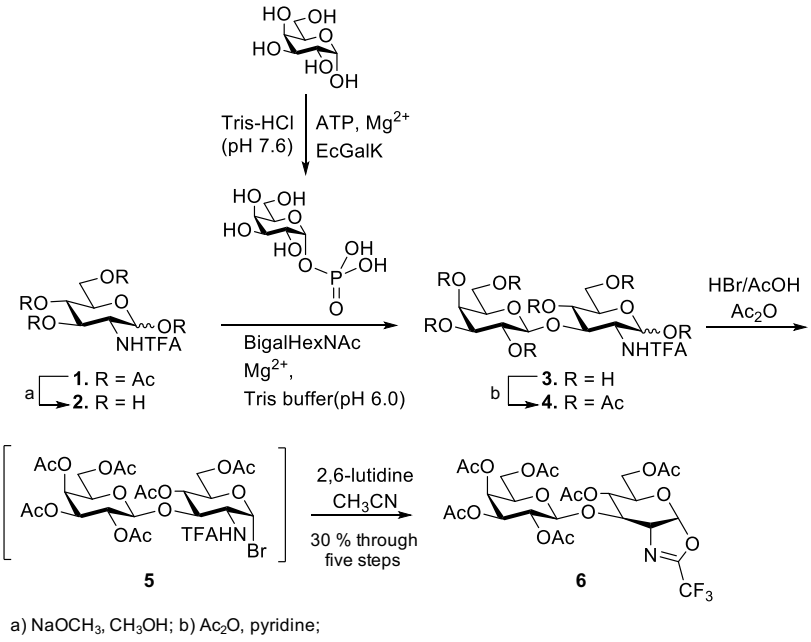
Disaccharide donor **6** consists of two monosaccharide units. Since galactose unit was introduced in the form of gal-1-phosphate which was enzymatically synthesized by EcGalK in situ, N-trifluoroacetylglucosamine **2** is the only building block chemically synthesized in the first stage. Upon reacting with p-anisaldehyde, the free amine at C-2 position in glucosamine hydrochloride **9** was protected to afford 4-methoxybenzimine, giving **10**. All hydroxyl groups of **10** were protected following routine global peracetylation methods using Ac_2O and pyridine. Allowing for the following orthogonal manipulation on amino groups, 4-methoxybenzimine on **11** was disarmed by stirring peracetylated imine in hot acetone with addition of 5M HCl. After amine group exposed, TFA group can be installed under the condition of trifluoroacetic anhydride pyridine/DCM to give **1**.



Scheme 2.1 Synthesis of TFA-amino Monosaccharide

Before entering the enzymatic step, building block **1** needs to be globally deacetylated using $NaOCH_3/CH_3OH$ providing **2** (Scheme 2.2). With galactose and water-soluble **2** at hand, sequential one-pot two-enzyme approach was carried out. Galactose was converted to Gal-1-phosphate *in situ* assisted by galactokinase EcGalK

at pH 7.6 in the presence of ATP and divalent metal cation Mg^{2+}. Once the reaction was completed, the pH of the mixture was adjusted to 6.0 to accommodate the second enzymatic glycosylation catalyzed by BiGalHexNAc-P to produce disaccharide **3**. The latter compound was peracetylated using Ac_2O and pyridine to give **4**, which was further activated into bromoglycoside. Due to the instability of the bromide, only a simple workup of the reaction was performed before moving to the next step. TFA oxazoline disaccharide donor **6** was obtained in 2 hours by directly applying 2,6-lutidine to bromo sugar intermediate **5**. It is anticipated that oxazoline conformation that blocks the α side of the ring, which offers a better regioselectivity for β linkage formation in chemical glycosylation.

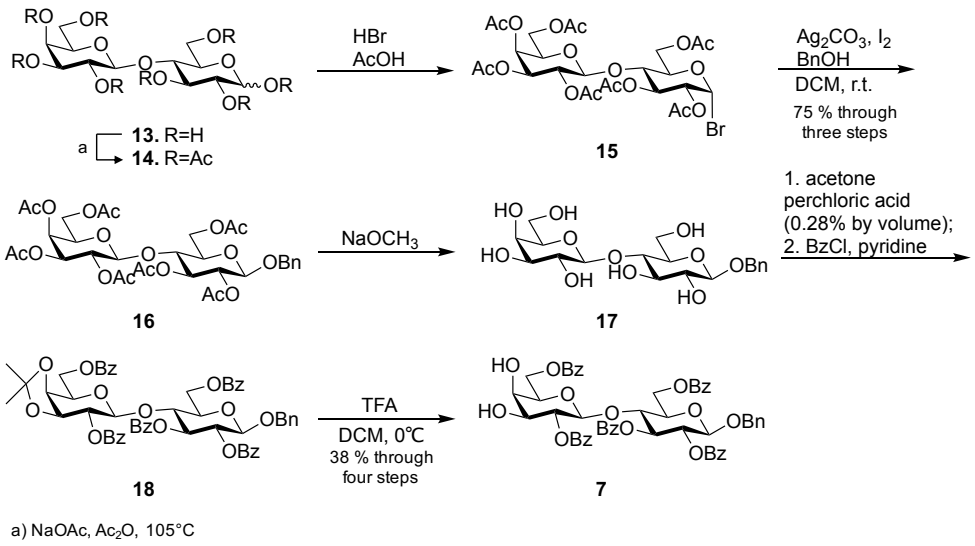


Scheme 2.2 Chemo-enzymatic Synthesis of Oxazoline Glycosyl Donor

Synthesis of glycosyl acceptor

Synthesis of glycosyl acceptor **7** started from lactose. Peracetate β lactoside **13** was formed using sodium acetate under refluxing condition. To install benzyl group at

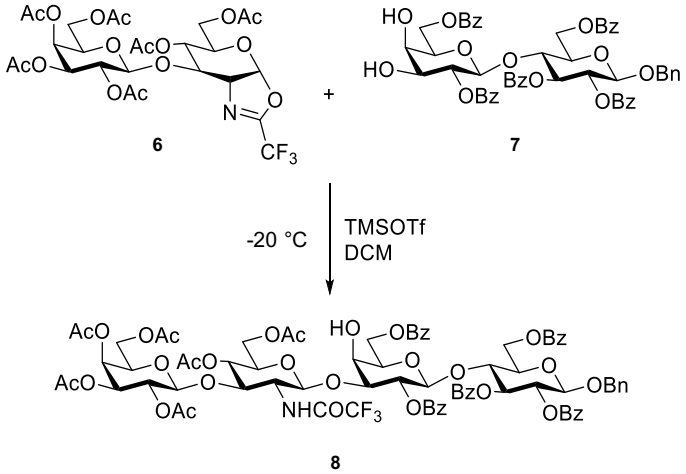
the anomeric center of glucose unit, synthesis of bromoglycoside **15** is the route chosen. The anomeric acetate of **14** was converted into α-bromolactoside **15** using HBr, AcOH and Ac_2O. The latter compound would be directly activated using freshly prepared Ag_2CO_3 and catalytic amount of I_2 together with benzylation reagent BnOH to give **16** as product.[7] Global deacetylation of **16** using sodium methoxide provided **17** which could allow further selective orthogonal protection. In contrast to the traditional tedious and low-efficient protecting strategy using anhydrous acetone and CSA as catalyst, a faster way of installing isopropyl group on C-3 and C-4 position was developed, which utilized catalytic amount (0.28% by volume) of perchloric acid to motivate the reaction in the presence of acetone. The advantage of this strategy is that reaction rate will be largely accelerated and the reaction accomplished in 15 mins with the sign of forming a homogenous yellow solution. The virtue of this fast route is that no anhydrous reagent was engaged. With the aforementioned method, acetone ketal would successfully be introduced onto C-3 and C-4 position of **17** yielding a crosslinked protected glycoside. Taking necessity of orthogonal protection into account, benzoyl was chosen as the global protecting group for the remaining hydroxyl groups on lactoside. Perbenzoylation of the isopropylidene-protected benzyl lactoside could be achieved through the use of benzoyl chloride and pyridine to give **18**. The removal of isopropyl crosslinking protection using TFA under 0°C gave the readily prepared glycosyl acceptor **7**.



Scheme 2.3 Synthesis of Glycosyl Acceptor

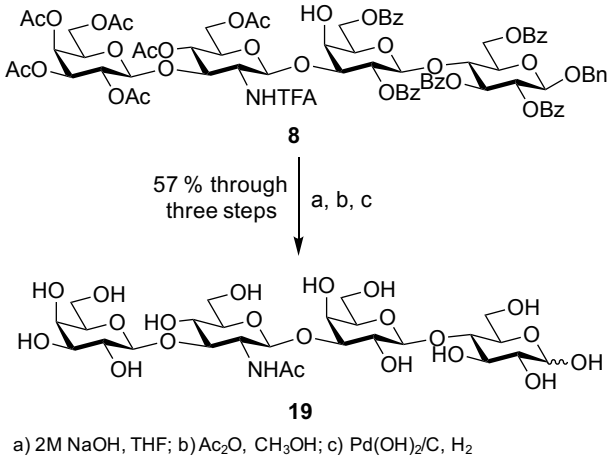
Synthesis of Lacto-N-tetraose

Having prepared oxazoline glycosyl donor **5** and glycosyl acceptor **6**, the chemical glycosylation was performed using the promoter TMSOTf in DCM. Various glycosylation conditions were investigated and it was found that -20 °C together with 0.7 equiv. of TMSOTf would give the highest yield of LNT (87%).



Scheme 2.4 Synthesis of Lacto-N-tetraose(LNT)

Global deprotection of tetrasaccharide



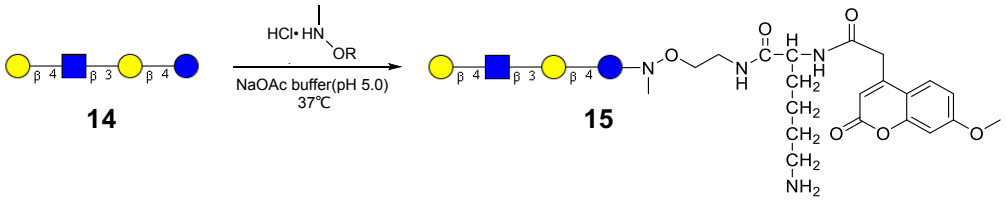
*Scheme 2.5 Global Deprotection of Tetrasaccharide **8***

The chemically synthesized tetrasaccharide 8 was globally deprotected in sequential manners. By subjecting to 2M NaOH until pH reached 7, all ester groups (acetyl and benzoyl groups) of **8** have been discharged. Acetyl group could be re-installed on free amine by employing acetic anhydride in methanol, after which the benzyl group at the anomeric center of glucose could be deprotected using $Pd(OH)_2$ on carbon under hydrogen atmosphere, giving fully deprotected tetrasaccharide 19. (Scheme 2.5)

Multi-functional linker installation of Lacto-N-tetraose

With Lacto-N-tetraose in hand, concern was that with the increasing size of oligosaccharides during enzymatic extension, purification would be tedious. Prudden *et. al*[10] have developed a multi-functional linker with fluorescent coumarin which allows for purification and efficient UV detection by reverse phase HPLC. Moreover, the linker contains a lysine chain having a free amine which can be used for further functionalization or attachment to a succinimide activated glass slide for micro-array

analysis. The coumarin linker was installed at the reducing end of the globally deprotected LNT **14** under acidic condition (pH 5.0) to give **15**.

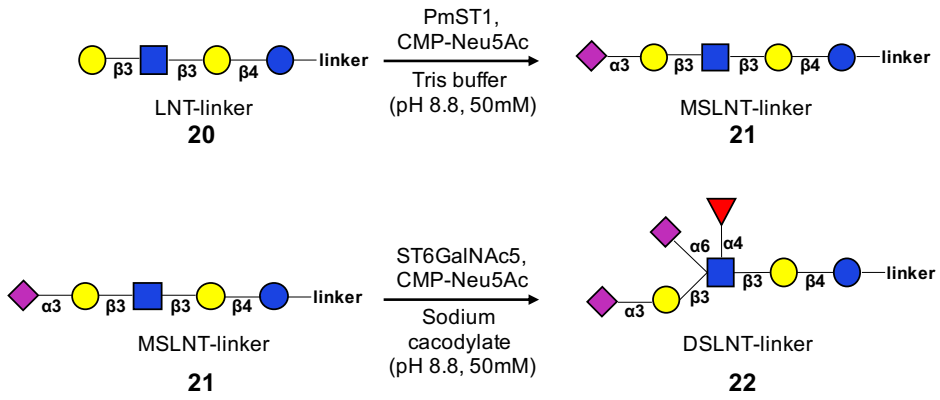


Scheme 2.6 Linker Ligation

Substrate specificity of different FUTs against LNT based library

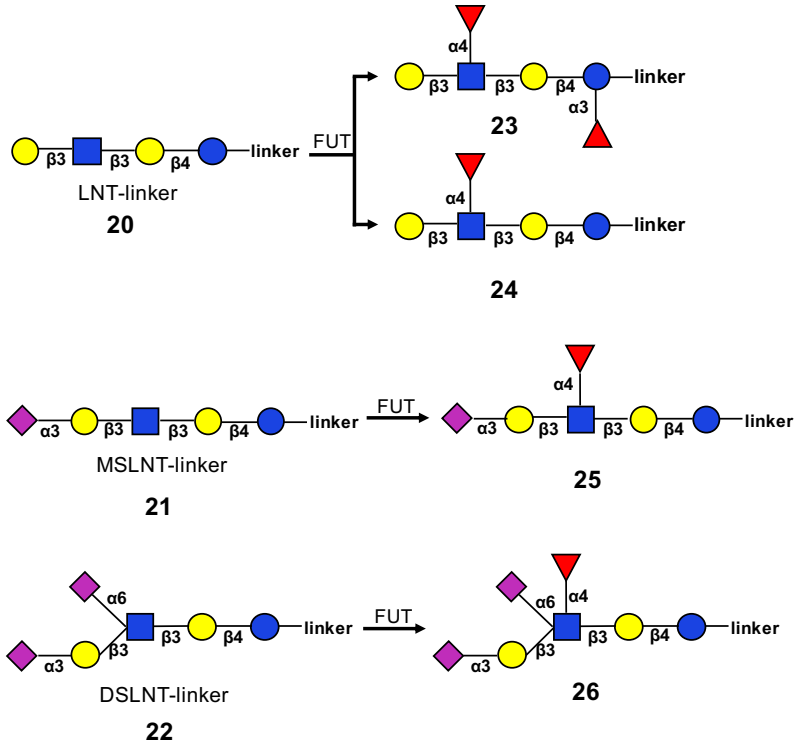
Research by Newburg *et. al*[11] in 2001 collected diverse milk samples from 12 women and looked into the HMO composition. It revealed that the concentration and fucosylated pattern of human milk oligosaccharides vary between individuals and also over the course of lactation. One of the 12 milk donors produced fucosyloligosaccharides devoid of α1,2- linkage but contained α1,3- and α1,4- linked fucose until the later stage in lactation. This observation is intriguing, motivating us to research the substrate specificities of different mammalian fucosyltransferases (FUTs).

To understand how different fucosyltransferases (FUTs) modify LNT based structures, LNT with sialic acid(s) modification were enzymatically synthesized as substrates for FUTs enzymatic screenings (Scheme 2.7).



Scheme 2.7 Enzymatic Synthesis of Sialylated LNT

MSLNT and DSLNT, they were exposed to a collection of FUTs (FUT3, FUT4, FUT5, FUT6, FUT9) of small-scale (Scheme 2.8). The products were analyzed by ESI-MS spectra. It was found that all fucosyltransferases gave difucosylation to LNT-linker without providing preference between the suspected structures **23** and **24**. However, when LNT-linker has one or two sialic acid on them, namely MSLNT-linker **21** or DSLNT-linker **22**, only monofusosylation patterns were observed within certain species of enzymes. (Table 2.1)



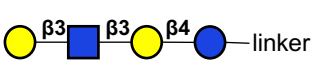
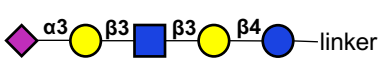
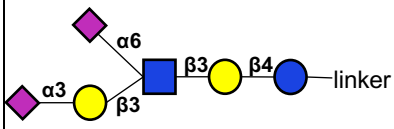
Scheme 2.8 Fucosyltransferases (FUTs) Enzymatic Screenings of LNT Series

After the screening, 0.5 mg-scale enzymatic reaction was performed, allowing the product to be purified and structures characterized by NMR. Upon completion of purification, it was confirmed with the pilot studies that difucosylated product were obtained. Compared with the handbook of glycotransferases written by Naoyuki *et.al*[12], some readouts are surprisingly quite the opposite. For example, FUT 3 is supposed to add on only one fucose at GlcNAc residue in all LNT derivatives, while in our cases, ESI results show that FUT 3 has no selectivity and will add on two fucose at the same time (the other fucose is suspected to be added on the glucose residue), giving structures **23** (Table 2.1). This discrepancy led us into re-define the enzymatic reaction condition in a more subtle way that the quantity of enzyme being applied is proper (0.05mg/ml)[13] without pushing reaction into full completion generating unnatural

oligosaccharides. Moreover, different FUTs reaction against certain substrate were performed in parallel to standardize condition for all FUTs reaction.

To monitor the enzymatic reaction progress more rigorously, high performance liquid chromatography (HPLC) is applied into analyzing and purifying products. In the given reaction time (12hrs), the ratio of either mono-fucosylated or di-fucosylated product will be defined by elution peak area. Fractions will be collected and checked by ESI-MS, after which the product will be lyophilized and characterized by NMR.

Table 2.1 Substrate Specificities of Different FUTs against LNT-based Oligosaccharides

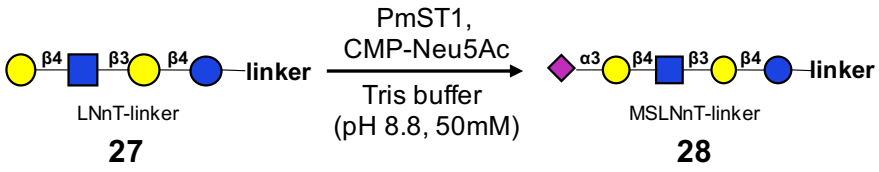
	LNT-linker 	**MSLNT-linker** 	**DSLNT-linker** 
FUT 3	**23**	**25**	**26**
FUT 4	**23**	N/A	N/A
FUT 5	**23**	N/A	**26**
FUT 6	**23**	**25**	N/A
FUT 9	**23**	N/A	N/A

After subjecting to corresponding enzymatic condition, LNT-linker showed no sign of specificity towards the panel of fucosyltransferases, all yielding di-fucosylated compounds corresponding to **23**. When LNT-linker was mono-silalylated at terminal galactose in α2,3 fashion giving **21**, only FUT3 and FUT6 recognize the MSLNT-linker as substrate. Instead of giving di-fucosylation, only mono-fucosylated compound **25** was harvested. When two sialic acid were added to LNT-linker, disialyllacto-N-tetraose (DSLNT), a crucial HMO species which has been identified to be associated with lower risk of necrotizing enterocolitis (NEC) in both neonatal rat and human models, are

yielded. Only FUT3 and FUT5 identify DSLNT as substrate and provide mono-fucosylated product **26**. Driven by the reaction patterns of MSLNT and DSLNT, it is safe to conclude that sialylation could effectively block the glucose from being fucosylated, giving only mono-fucosylation at GlcNAc moiety.

Synthesis and substrate specificity of LNnT-based oligosaccharides

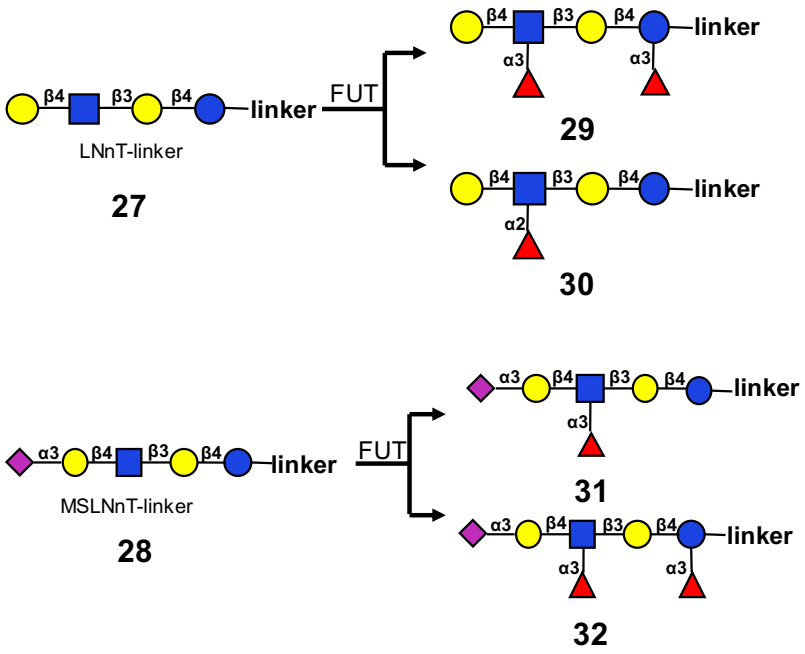
To make comparison with the structure bearing type I LacNAc, a panel of compounds containing type II LacNAc were synthesized. Enzymatic synthesis proceeded with coumarin-linked lactose **27**, which consecutively subject to β3GNT2 and β4GalT1, adding GlcNAc and galactose, respectively. The synthesized Lacto-N-Neotetraose (LNnT)-linker could be further sialylated with PmST1 which adds sialic acid in α2,3 fashion, giving mono-sialylated LNnT (MSLNnT) **28**.



Scheme 2.9 Enzymatic Synthesis of Sialylated LNnT

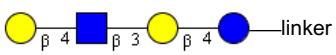
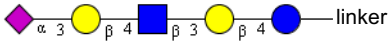
Since di-sialylated LNnT species are not among the natural HMOs that has been discovered so far, it was not included in the research. Subsequent substrate specificities investigation of diverse fucosyltransferases could be carried out based on LNnT and MSLNnT species. Similar to the result of enzymatic reactions carried on LNT-linker, all fucosyltransferase species showed capability of adding two fucose to LNnT-linker, giving structures **29**, with one case having minor amount of mono-fucosylated compound which is not in sufficient amount for NMR study. One could assume that di-fucosylation could be achieved with more substrate provided. The result

suggests that all enzymes lack substrate specificities on LNnT-linker as to fucosylation patterns. In another case where MSLNnT-linker **28** are subjected to fucosyltransferases, FUT3, FUT4 and FUT5 give di-fucosylated products **32** while FUT6 and FUT9 yield mono-fucosylated compound **31**, confirming that FUT3 and FUT5, FUT6 and FUT9 may pose similar carbohydrate recognition domain (CRD), enabling the transglycosylation.



Scheme 2.10 Fucosyltransferases (FUTs) Enzymatic Screenings of LNnT Series

Table 2.2 Substrate Specificities of Different FUTs against LNnT-based Glycans

	LNnT 	**MSLNnT** 
FUT 3	**29**	**32**
FUT 4	**29 (<10% product 30)**	**32**
FUT 5	**29**	**32**
FUT 6	**29**	**31**
FUT 9	**29**	**31**

CONCLUSION

Although human milk oligosaccharides have been discussed over past decades, major concerns fall on the progress certain intestinal flora has made with the influence of carbohydrates. Due to structural complexity and vast volume of the oligosaccharides library, pioneering work was mainly directed to drawing conclusion of pathogenesis of various gastrointestinal diseases from observed and purified HMO species. To address the importance of correlation between functionality and framework, as well as to expand upon the idea that enzymes responsible for structure fabrication are mutually regulated, we describe here a chemoenzymatic methodology that makes it possible to make an array of HMOs taking type I LacNAc as backbone. Based on the work that Schwab and coworkers observed that various bifidobacterial species have a capability of utilizing fucosyllactose as nutrient source, LNT was employed as synthetic precursor. Subsequently, relevant mammalian fucosyltransferases could recognize the sugar moiety in a sequential manner that could be traced with mass spectrometry, affording mono- and bi-fucosylated structures which can be purified through reverse phase

HPLC with a pre-installed aromatic linker. The strategy made it possible, for the first time, to gain a detailed overview of how different fucosyltransferases performance upon type I HMO core structures.

EXPERIMENTAL

General Methods and Reagents: ^{1}H and ^{13}C NMR spectra were recorded on Varian Inova 300, 500 or 600 MHz. Chemical shifts are reported in parts per million (ppm) relative to H1 of reducing glucose which was set to δ 4.00. NMR data is represented as follows: Chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, dd = doublet of doublets, m = multiplet and/or multiple resonances), J coupling, integration, and peak identity. All NMR signals were assigned on the basis of ^{1}H NMR, gCOSY, gHSQC, zTOCSY. Mass spectra were recorded on an Applied Biosystems SCIEX MALDI TOF / TOF 5800 mass spectrometer using 2,5-dihydroxybenzoic acid (DHB) as a matrix or a high-resolution Shimadzu LCMS-IT-TOF mass spectrometer. Reagents were purchased from Sigma-Aldrich (unless otherwise noted) and used without further purification. HPLC purification of compounds was performed on an Agilent Technologies 1200 Series HPLC equipped with a Halo® RP-Amide 4.6 x 100 mm, 2.7 µm column. HPLC grade acetonitrile and water were purchased from Fisher. Neutral and sialylated compounds were eluted in a 50 mM solution of ammonium bicarbonate with a linear gradient of 0~50%, 0~20% or 0~25% acetonitrile over 30 mins, respectively at 1 mL/min. Spectra were monitored at a wavelength of 324 nm using an Agilent 1200 Series diode array with a multiple wavelength detector. Column chromatography was performed on silica gel G60 (Silicycle, 60 -200 µm, 60 Å). Thin layer chromatography (TLC) was conducted on Silica gel 60 F_{254} (EMD Chemicals Inc.) with detection by UV light absorption (254 nm) where applicable.

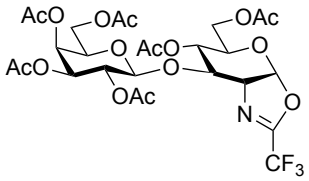
Visualization was accomplished by immersing the plate into 10% sulfuric acid in ethanol or Ninhydrin, followed by charring using heat gun. All reactions were carried out under argon atmosphere unless specified. Molecular sieves were flame-dried under vacuum immediately prior to use.

General biochemical reagents

Calf intestine alkaline phosphate (CIAP) was purchased from BioLabs® Inc. Pasteurella multocida α-2,3-sialyltransferase 1 (PmST1) was purchased from Chemily Glycoscience. Cytidine-5'-monophospho-N-acetylneuraminic acid (CMP-Neu5Ac), uridine 5'-diphosphogalactose (UDP-Gal) and uridine 5'-diphospho-N-acetylglucosamine (UDP-GlcNAc) were purchased from Roche Life Science. Fucosyltransferases were provided by Dr. Kelley Moremen's group.

Chemical Synthesis Procedures

2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-2-trifluoromethyl-(2-acetamido-4,6-di-O-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-d]-2-oxazoline (6)

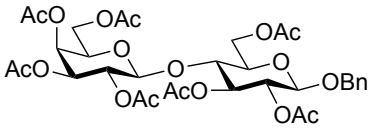


Compound **2** (0.5 g, 1.82 mmol), α-Galactose-1-phosphate (55% purity, 1.03 g, 2.23 mmol) were dissolved in an aqueous buffer containing Tris-HCl (50 mM, pH 6.0, 45mL) and $MgCl_2$ (20 mM, 1. 8 mL). After the addition of BigalHexNAc (1.3 mL, 16 mg), the reaction mixture was incubated at 37 °C for 48 h. Progress of the reaction was monitored by silica gel-coated TLC (1:5 H_2O/CH_3CN). The reaction was quenched by the addition of the same volume of ice-cold ethanol, and after 2 h the precipitates enzyme were

removed by centrifugation (5,000 rpm). The supernatant was collected and concentrated *in vacuo* and the residue purified by silica gel chromatography (H_2O/CH_3CN, 1:6, v:v) to give **3** as an α and β mixture as a white powder. Compound **3** (530 mg, 1.22 mmol) was dissolved in anhydrous pyridine (10 ml). The reaction vessel was cooled (0 °C) in ice bath, after which acetic anhydride (1.1 mL, 12.11 mmol) was added in portions. The reaction mixture was stirred overnight at room temperature. Progress of the reaction was monitored by silica gel-coated TLC (ethyl acetate/hexane, 3:2, v:v). The pyridine was evaporated and the residue was diluted with EtOAc (50 mL) and washed with 1M HCl (3 x 60 mL) and brine (2 x 60 mL). Organic layers were combined and dried ($MgSO_4$), filtered, and the filtrate was concentrated *in vacuo* to give **4** as a white foam solid which was directly used in the next step. Compound **4** was dissolved in a mixed of acetic acid (1.8 mL) and HBr (33% in acetic acid, 2.4 mL). The reaction mixture was stirred at room temperature for 10 min, after which silica gel TLC (ethyl acetate/hexane, 3:2 v/v) showed the completion of reaction. The reaction mixture was co-evaporated with toluene (3 x 4 mL). The residue was dried *in vacuo* to give a crude product as dark brown sticky oil. After drying in vacuo for 1 h, the bromosugar was dissolved in anhydrous acetonitrile (6.6 mL), to which 2,6-lutidine (112.4 µL, 1.4 mmol) was added under argon protection. The reaction mixture was stirred at room temperature for 1.5 h. Upon the completion of the reaction, the reaction mixture was diluted with ethyl acetate (50 mL) and washed with a copper sulfate solution (5%, 3 x 50 mL) and brine (2 x 50 mL). The organic layers were combined, dried ($MgSO_4$), and filtered. The filtrate was concentrated *in vacuo* to give a residue which was purified by silica gel column chromatography (EtOAc/hexane, 1:1, v:v), obtaining oxazoline **6** (366 mg, 30% through 5 steps). ^1H NMR (500 MHz, $CDCl_3$): δ 6.31

(d, *J* = 7.6 Hz, 1H, H-1, Glc), 5.40 (dd, *J* = 3.4 Hz, 0.7 Hz, 1H, H-4, Glc), 5.23 (m, 1H, H-3, Glc), 5.19 (dd, *J* = 8.0, 10.43 Hz, 1H, H-2, Gal), 5.04 (dd, *J* = 3.5, 10.43 Hz, 1H, H-3, Gal), 4.76 (d, *J* = 8.0 Hz, 1H, H-1, Gal), 4.29 (dd, *J* = 7.5, 7.6 Hz, 1H, H-2, Glc), 4.24 (t, *J* = 2.2 Hz, 1H, H-3, Glc), 4.22 (m, 2H, H-6a, H-6b Glc), 4.20 (m, 1H, H-6a, Gal), 4.11 (dd, *J* = 6.6, 11.2 Hz, 1H, H-6b, Gal), 4.00 (t, *J* = 6.5 Hz, 1H, H-5, Gal), 3.67 (m, 1H, H-5, Glc), 1.95-2.18 (m, 18H, COC*H*$_3$). ^{13}C NMR (75 MHz, $CDCl_3$): δ 102.50, 101.72, 75.76, 71.32, 70.73, 68.78, 68.73, 67.40, 67.04, 64.35, 63.50, 61.45, 61.40, 20.74, 20.70, 20.65, 20.57.

Benzyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-acetyl-β-D-glucopyranoside (16)

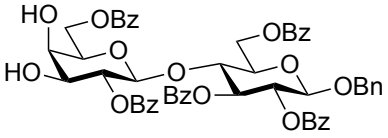


A suspension of sodium acetate (2.4 g, 29.2 mmol) and acetic anhydride (55.2 mL, 0.58 mol) was brought to boil (~105 °C) with stirring. The reaction vessel was temporarily removed from the heat source and D-Lactose (10.0 g, 29.2 mmol) was added slowly in small portions. The reaction mixture was brought to boil again and the suspension became clear in 10 min. The reaction mixture was poured into ice water (300 mL) after cooling down to 50 °C and vigorous mechanic stirring was employed to facilitate product precipitation. After stirring for 10 min a sticky gum-like compound was formed, which gradually dispersed into powder-like compound within 1.5 h under mechanical stirring. DCM (3 x 300 mL) was added to the suspension. The mixture was turned into a clear double-layer solution, which was subsequently extracted with, sat. $NaHCO_3$ (2 x 100 mL), brine (200 mL). The organic layers were combined, dried ($MgSO_4$), and filtered. The filtrate was concentrated *in vacuo* and the residue was recrystallized from ethanol to give

14 which was directly used in the next step. Lactose octaacetate **4** was added to a solution of glacier acetic acid (6 mL) and acetic anhydride (6 mL). The mixture was cooled (0 °C) in an ice bath and HBr/AcOH (33 wt %, 20 mL) was added dropwise over a period of 15 min with the aid of a dropping funnel. The reaction mixture was stirred for additional 1 h at room temperature, yielded a yellow homogenous solution. The reaction mixture was diluted with DCM (100 mL) and washed with water (2 x 100 mL) and $NaHCO_3$ (3 x 100 mL), brine (50 mL). The combined organic phases were dried ($MgSO_4$) and filtered. The filtrate was concentrated *in vacuo* to give **15** as a white foam. Crude bromosugar **15** (19.8 g, 28.3 mmol) in anhydrous DCM (80 mL) was added to an anhydrous DCM (80 mL) suspension of Ag_2CO_3 (20.0 g, 72.7 mmol), 4Å flame-dried molecular sieves (3 g) and I_2 (one crystal). The mixture was stirred for 5 min followed by addition of benzyl alcohol (14.6 mL, 141.5 mmol). The reaction mixture was stirred in the dark overnight and was filtered through a celite plug. The filtrate was concentrated *in vacuo* and purified by flash column chromatography (EtOAc/hexane, 1:1, v/v) to give **16** (15.4 g, 75 % over 3 steps) as a white foam. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.37-7.21 (m, 5H, aromatic), 5.32 (dd, *J* = 0.9, 3.3 Hz, 1H, H-4, Gal), 5.14 (dd, *J* = 9.3, 9.2 Hz, 1H, H-3, Glc), 5.08 (dd, *J* = 7.3, 9.9 Hz, 1H, H-2, Gal), 4.98 (dd, *J* = 7.9, 9.4 Hz, 1H, H-2, Glc), 4.94 (dd, *J* = 3.4, 10.4 Hz, 1H, H-3, Gal), 4.84 (dd, *J* = 12.3 Hz, 1H, PhC*H*), 4.57 (dd, *J* = 12.3 Hz, 1H, PhC*H*), 4.66 (br, 1H, H-6a, Gal), 4.50 (dd, *J* = 7.9 Hz, 1H, H-1, Glc), 4.49 (dd, *J* = 3.1, 10.9 Hz, 1H, H-6a, Glc), 4.46 (dd, *J* = 7.7 Hz, 1H, H-1, Gal), 4.09 (m, 1H, H-6b, Glc), 3.85 (ddd, *J* = 0.9, 6.9, 7.8 Hz, 1H, Gal), 3.80 (dd, *J* = 9.2, 9.7 Hz, 1H, H-4, Glc), 3.56 (ddd, *J* = 2.0, 5.0, 7.0 Hz, 1H, H-5, Glc), 2.12 (s, 3H, COC*H*$_3$), 2.12 (s, 3H, COC*H*$_3$), 2.03 (s, 3H, COC*H*$_3$), 2.02 (s, 3H, COC*H*$_3$), 2.02 (s, 3H, COC*H*$_3$), 1.98 (s, 3H, COC*H*$_3$), 1.94 (s, 3H, COC*H*$_3$); ^{13}C NMR

(75 MHz, $CDCl_3$): δ 128.00, 100.94, 99.07, 76.22, 72.66, 72.66, 71.30, 70.61, 70.61, 70.61, 69.15, 66.51, 65.92, 65.15, 61.92, 61.07, 20.71, 20.66, 20.66, 20.66, 20.61, 20.47. . MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{80}H_{80}F_3NNaO_{32}$, 1646.4513 found 1646.1632.

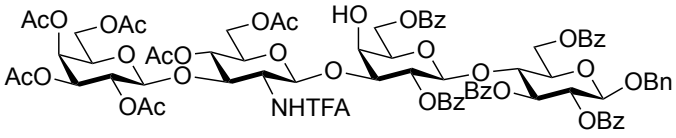
Benzyl 2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-glucopyranoside (7)



Compound **16** (2.7 g, 3.7 mmol) was dissolved in a mixture of CH_3OH/DCM (57 mL, 3:1, v/v). To this mixture was added $NaOCH_3$ in CH_3OH (500 mM) dropwise untill the pH reached 10. The reaction was stirring at room temperature overnight. The reaction was quenched by adding Dowex-H^+ resin until pH reached 6. The resin was filtered off and the filtrate was concentrated *in vacuo* to give **17** as crude product which was used directly in the next step. Benzyl lactoside **17** was dissolved in acetone (20 mL). To this solution was added 2,2-dimethoxypropane (4 mL, 32.5 mmol). The solution was cooled (0 °C) and perchloric acid (20 µL, 0.33 mmol) was added dropwise. The reaction was stirred for another 5 min (0 °C) before it was gradually warmed up to room temperature and kept stirring for another 10 min, giving a homogenous yellow solution. The reaction was quenched with Et_3N and the resulting mixture was concentrated *in vacuo*. The residue was dissolved in pyridine (10 mL) and the solution was brought to -40 °C (to the point where pyridine almost freeze). To the solution was added benzoyl chloride (1.23 mL, 10.6 mmol) and the reaction mixture was stirred gradually warmed up to 0 °C in a period of 1 h. After stirring at 0 °C for 4 h, the reaction was diluted with EtOAc (100 mL) and washed with 5% HCl (3 x 70 mL), water (2 x 70 mL), brine (100 mL). The combined organic layers

were dried ($MgSO_4$) and filtered. The filtrate was concentrated *in vacuo* to give **18** as crude product. Compound **16** was dissolved in DCM (80 mL) and the solution was cooled to 0 °C. The TFA (6 mL, 77.3mmol) was diluted with DCM (40 mL) and the solution was added into the solution of **18** in portions. The reaction mixture was stirred (0 °C) for 1 h after which it was quenched with sat. $NaHCO_3$. The reaction was diluted with DCM (100 mL) and washed with sat. $NaHCO_3$ (2 x 50 mL), brine (50 mL). The organic layers were combined, dried ($MgSO_4$) and filtered. The filtrate was concentrated *in vacuo*. Purification of the residue by silica gel chromatography using EtOAc/hexane (1:2, v/v) as eluent gave the disaccharide acceptor **7** as a white foam. (1.34 g, 1.4 mmol, 38% over four steps). ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.10-7.07 (m, 30H, aromatic), 5.58 (dd, *J* = 9.2, 9.2 Hz, 1H, H-3, Glc), 5.50 (dd, *J* = 7.6, 9.6 Hz, 1H, H-2, Gal), 5.29 (dd, *J* = 7.9, 9.6 Hz, 1H, H-2, Glc), 4.81 (d, *J* = 12.6 Hz, 1H, PhC*H*), 4.64 (d, *J* = 7.6 Hz, 1H, H-1, Gal), 4.58 (d, *J* = 6.2 Hz, 1H, H-1, Glc), 4.57 (d, *J* = 12.8 Hz, 1H, PhC*H*), 4.61-4.47 (m, 2H, H-6a, H-6b, Glc), 4.14 (dd, *J* = 9.4, 9.4 Hz, 1H, H-4, Glc), 4.01 (dd, *J* = 6.1, 10.9 Hz, 1H, H-6a, Gal), 3.78 (m, 1H, H-5, Gal), 3.75 (m, 1H, H-5, Glc), 3.67 (dd, *J* = 3.6, 9.8 Hz, 1H, H-4, Gal), 3.52 (m, 1H, H-6, Gal), 3.46 (dd, *J* = 6.2 Hz, 1H, H-3, Gal); ^{13}C NMR from HSQC (75 MHz, $CDCl_3$): δ 133.31, 133.25, 129.93, 129.85, 129.65, 128.60, 128.46, 128.07, 128.07,100.89, 98.96, 76.18, 73.73, 73.14, 73.06, 72.64, 72.64, 71.65, 70.44, 70.44, 68.53, 62.61, 61.80, 61.80. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{54}H_{48}O_{16}Na$, 975.96 found 976.13.

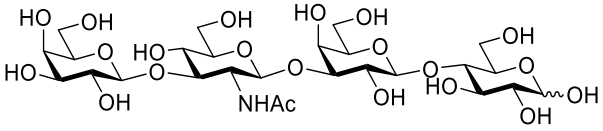
Benzyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-4,6-di-O-acetyl-2-trifluoroacetamido-β-D-glucopyranosyl-(1→3)-2,6-di-O-benzoyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-glucopyranoside (8)



A mixture of donor **6** (107.3 mg, 0.16 mmol), acceptor **7** (126.9 mg, 0.13 mmol), and 4Å flame-dried molecular sieves (100 mg), was stirred in anhydrous DCM (3 mL) for 10 min. The resulting solution was cooled (-20°C) upon which TMSOTf (10 μL, 55.1 μmol) was added, and stirring was continued for 5 min. After TLC showed the full consumption of donor, the reaction was quenched with Et_3N. The molecular sieves were filtered off and the filtrate was concentrated *in vacuo*. The obtained residue was applied to silica gel chromatography and purified using toluene/acetone (4:1, v/v) as the eluent to give tetrasaccharide **8** as a white foam (93.5 mg, 43%). ^{1}H NMR (300MHz, $CDCl_3$): δ 8.07-7.08 (m, 30H, Ar-H), 5.62 (dd, *J* = 9.4, 9.4 Hz, 1H, H-3, Glc), 5.49 (dd, *J* = 7.8, 9.6 Hz, 1H, H-2, Glc), 5.39 (dd, *J* = 7.9, 9.6 Hz, 1H, H-2, Gal-1), 5.32 (dd, *J* = 0.7, 3.3 Hz, 1H, H-4, Gal-2), 5.04 (dd, *J* = 7.9, 10.3 Hz, 1H, H-2, Gal-2), 4.97 (d, *J* = 7.8 Hz, 1H, H-1, GlcNTFA), 4.88 (dd, *J* = 8.8, 8.8 Hz, 1H, H-4, GlcNTFA), 4.83 (dd, *J* = 3.2, 10.4 Hz, 1H, H-3, Gal-2), 4.78 (d, *J* = 12.6 Hz, 1H, C*H*HPh), 4.64 (d, *J* = 7.8 Hz, 1H, H-1, Glc), 4.57 (d, *J* = 7.7 Hz, 1H, H-1, Gal-1), 4.54 (d, *J* = 13.2 Hz, 1H, CH*H*Ph), 4.48 (d, *J* = 11.5 Hz, 1H, H-6a, Glc), 4.42 (dd, *J* = 4.8,12.3 Hz, 1H, H-6b, Glc), 4.34 (d, *J* = 7.9 Hz, 1H, H-1, Gal-2), 4.23 (dd, *J* = 9.2, 9.2 Hz, 1H, H-3, GlcNTFA), 4.15 (m, 1H, H-6a, Gal-1), 4.13 (m, 1H, H-4, Glc), 4.08 (m, 1H, H-6a, GlcNTFA), 4.04 (d, *J* = 6.6 Hz, 2H, H-6b, GlcNTFA, Gal), 3.92 (m, 1H, H-4, Gal-1), 3.77 (ddd, 1H, H-5, Gal-1), 3.75 (m, 1H, H-3, Gal-1), 3.68 (ddd, 1H,

H-5, Glc), 3.63 (ddd, 1H, H-5, GlcNTFA), 3.50 (t, *J* = 6.0 Hz, 1H), 3.41 (dd, *J* = 8.0, 16.3 Hz, 1H, H-2, GlcNTFA), 2.10 (s, 3H, COC*H*$_3$), 2.02 (s, 3H, COC*H*$_3$), 1.97 (s, 3H, COC*H*$_3$), 1.96 (s, 3H, COC*H*$_3$), 1.95 (s, 3H, COC*H*$_3$), 1.93 (s, 3H, COC*H*$_3$); ^{13}C NMR from HSQC (75 MHz, $CDCl_3$): δ 129.44, 129.65, 129.44, 129.76, 133.36, 133.14, 133.04, 128.56, 128.35, 128.14, 127.75, 127.75, 100.52, 100.41, 98.90, 98.40, 80.10, 75.51, 75.20, 72.73, 72.90, 72.20, 72.20, 71.60, 70.93, 70.93, 70.72, 70.40, 70.28, 68.71, 68.00, 67.72, 66.52, 62.50, 62.50, 62.50, 62.32, 61.90, 60.81, 57.21, 20.52, 20.42, 20.42, 20.42. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{33}H_{42}O_{18}Na$, 749.68 found 749.93.

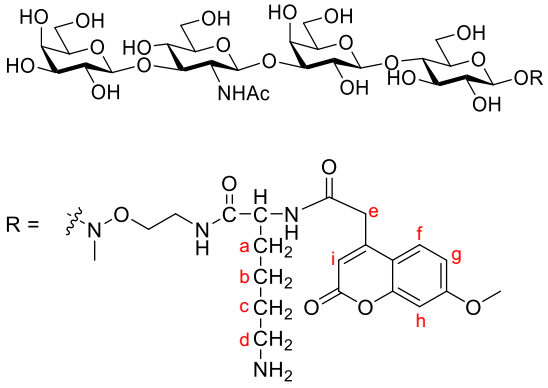
Lacto-N-Tetraose (19)



THF (2 mL) and CH_3OH (2 mL) were added to tetrasaccharide **8** (79.7 mg, 49.06 μmol) and aqueous 5M NaOH solution (100 μL) was added. The reaction mixture was left stirring overnight at room temperature. After MALDI-TOF MS confirmed cleavage of all esters, the reaction was quenched with acetic acid. The reaction mixture was concentrated *in vacuo* and the residue was re-dissolved in CH_3OH (3 mL) followed by addition of Ac_2O (1 mL). The reaction was stirred at room temperature for 2 h and then concentrated *in vacuo*. The residue was applied to P2 biogel column to remove the accumulated salts of the previous steps. The fractions containing the product were combined and lyophilized. The dried compound was dissolved in a solvent mixture of CH_3OH/H_2O (2 mL, 1:1, v/v), after which $Pd(OH)_2$ (7.4 mg) was added. The reaction mixture was stirred overnight at room temperature under an atmosphere of hydrogen gas. After MALDI-TOF MS confirmed the cleavage of benzyl group, the $Pd(OH)_2$ was filtered

off using a celite plug. The filtrate was lyophilized and purified using P2 biogel size exclusion chromatography (deionized water as eluent) to give **19** as white powder. (20 mg, 57% through 3 steps). ^{1}H NMR (900 MHz, D_2O): δ 4.63 (d, *J* = 8.6 Hz, 1H, H-1, GlcNAc), 4.55 (d, *J* = 8.0 Hz, 1H, H-1, Glc), 4.33 (d, *J* = 7.8 Hz, 1H, H-1, Gal-1, Gal-2), 3.17 (dd, *J* = 8.5, 8.5 Hz, 1H, H-2, Glc), 3.54 (m, 1H, H-2, Glc), 3.80 (m, 1H, H-3, Glc), 3.67 (m, 2H, H-4 Glc, H-4 Gal-2), 3.49 (m, 1H, H-2, Gal-1), 3.42 (m, 1H, H-2, Gal-2), 3.61 (m, 1H, H-3, Gal-1), 4.04 (d, *J* = 3.4 Hz, 1H, H-4, Gal-1), 3.53 (m, 1H, H-3, Gal-2). ^{13}C NMR from HSQC (75 MHz, D_2O): δ 103.15,102.49, 95.69, 91.75, 82.01, 81.93, 78.29, 75.11, 75.04, 74.56, 73.80, 72.44, 71.40, 71.06, 70.65, 70.02, 69.94, 68.51, 68.44, 68.37, 60.94, 60.38, 60.04, 59.92, 54.68, 54.61, 46.42. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{26}H_{45}NNaO_{21}$, 730.2382 found 730.2808.

Multifunctional linked lacto-N-tetraose (20)



Lacto-N-tetraose **19** (1 mg, 1.41 mmol) and anomeric linker (17.9 mg, 35.33 mmol) were dissolved in sodium acetate buffer (400 µl, 1M, pH = 5.0) and incubated overnight at 37 °C. The reaction was monitored by MALDI-TOF MS. When reaction did not progress further, it was filtered using Nylon membrane syringe filter (13 mm, pore size 0.2 µm) and lyophilized. The dried product was purified by reverse phase HPLC with the following conditions:

A=50mM ammonium bicarbonate, B=acetonitrile

Time (mins)	%A	%B	Flow Rate (mL/min)
0	100	0	1
20	50	50	1

Collected fractions were confirmed by MALDI-TOF, combined and lyophilized to give product **15** as white powder.

^{1}H NMR (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	NHAc
Glc	3.99	3.35	3.45	3.56	-	-	-
Gal(1)	4.27 (d, *J*=7.9 Hz, 1H)	3.43	3.53	3.98	-	-	-
GlcNAc	4.57 (d, *J*=8.5 Hz, 1H)	3.74	3.32	3.41	3.65	3.74	1.91 (s, 3H)
Gal(2)	4.28 (d, *J*=7.5Hz, 1H)	3.37	3.48	3.75	-	-	-

Linker part NMR assignment

N-CH3	2.51
N-O-CH2	3.73
CH2-NH	3.25
C(O)-CH-NH	4.13 (t, *J* = 7.38 Hz, 1H)
A	1.66, 1.61
B	1.23
C	1.48
D	2.75
E	3.81
F	7.55 (d, *J* = 8.74 Hz, 1H)
G	6.88
H	6.90
I	6.17
O-CH3	3.76

^{13}C NMR from HSQC(150 MHz, D_2O): δ125.93, 113.05, 112.45, 101.01, 103.07, 102.41, 92.90, 82.00, 81.90, 77.99, 75.92, 75.43, 75.23, 75.12, 72.39, 70.73, 69.93, 69.07, 68.38,

68.17, 60.80, 60.42, 60.04, 55.82, 54.65, 53.85, 39.12, 38.20, 38.20, 37.34, 30.39, 30.39, 26.38, 22.17, 22.05, 22.05. MALDI-TOF-MS(m/z): $[M+Na]^+$ calcd for $C_{47}H_{73}N_5NaO_{26}$, 1146.4441 found 1146.5583.

General Enzymatic Procedures:

General procedure for the addition of α1,4 (and α1,3) Fuc using FUT3

HMO acceptor and GDP-fucose (1.5 eq per fucose) were dissolved in a water buffer solution containing Tris-HCl (50 mM, pH 7.5), $MnCl_2$ (10 mM) buffer at a final concentration of 5 mM. To this solution was added CIAP (10 U μL^{-1}) and FUT3 (10 μg/μmol substrate). The reaction mixture was incubated overnight at 37 °C. Progress of the reaction was monitored by MALDI-TOF MS or ESI. In case the remaining starting material was detected, additional GDP-fucose, FUT3 and CIAP was added until the substrate was consumed. The reaction mixture was centrifuged using a Nanosep® Omega ultrafiltration device (3 kDa MWCO) to remove enzymes and the filtrate was lyophilized, followed by purification using reverse-phase HPLC equipped with Halo® RP-Amide 4.6 x 100mm, 2.7 μm column to give the desired product.

General procedure for the addition of α1,4 (and α1,3) Fuc using FUT4

HMO acceptor and GDP-fucose (1.5 eq per fucose) were dissolved in a water buffer solution containing Tris-HCl (50 mM, pH 7.2), $MnCl_2$ (10 mM) buffer at a final concentration of 5 mM. To this solution was added CIAP (10 U μL^{-1}) and FUT4 (10 μg/μmol substrate). The reaction mixture was incubated overnight at 37 °C. Progress of the reaction was monitored by MALDI-TOF MS or ESI. In case the remaining starting material was detected, additional GDP-fucose, FUT4 and CIAP was added until the substrate was consumed. The reaction mixture was centrifuged using a Nanosep®

Omega ultrafiltration device (3 kDa MWCO) to remove enzymes and the filtrate was lyophilized, followed by purification using reverse-phase HPLC equipped with Halo® RP-Amide 4.6 x 100mm, 2.7 µm column to give the desired product.

General procedure for the addition of α1,4 (and α1,3) Fuc using FUT5

HMO acceptor and GDP-fucose (1.5 eq per fucose) were dissolved in a water buffer solution containing Tris-HCl (50 mM, pH 7.5), $MnCl_2$ (10 mM) buffer at a final concentration of 5 mM. To this solution was added CIAP (10 U μL^{-1}) and FUT5 (10 µg/µmol substrate). The reaction mixture was incubated overnight at 37 °C. Progress of the reaction was monitored by MALDI-TOF MS or ESI. In case the remaining starting material was detected, additional GDP-fucose, FUT5 and CIAP was added until substrate was consumed. The reaction mixture was centrifuged using a Nanosep® Omega ultrafiltration device (3 kDa MWCO) to remove enzymes and the filtrate was lyophilized, followed by purification using reverse-phase HPLC equipped with Halo® RP-Amide 4.6 x 100mm, 2.7 µm column to give the desired product.

General procedure for the addition of α1,4 (and α1,3) Fuc using FUT6

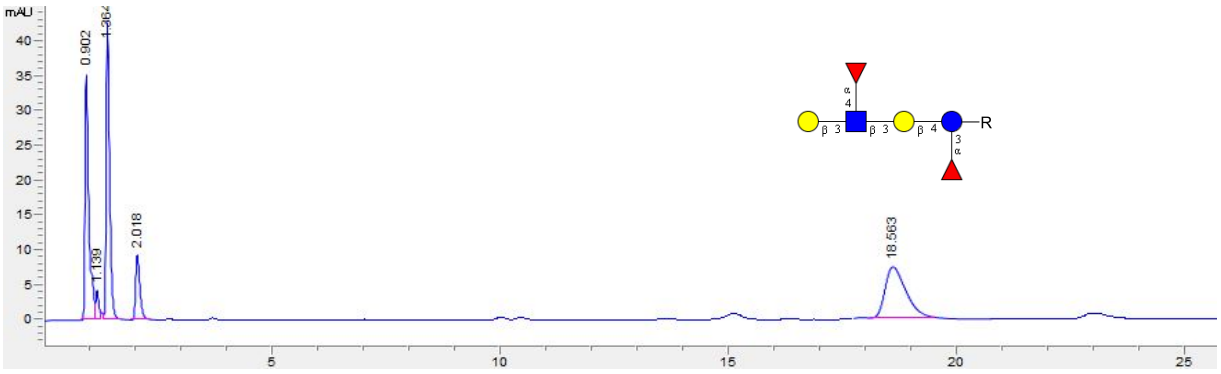
HMO acceptor and GDP-fucose (1.5 eq per fucose) were dissolved in a water buffer solution containing MES (80 mM, pH 6.5), $MnCl_2$ (10 mM) buffer at a final concentration of 5 mM. To this solution was added CIAP (10 U μL^{-1}) and FUT6 (10 µg/µmol substrate). The reaction mixture was incubated overnight at 37 °C. Progress of the reaction was monitored by MALDI-TOF MS or ESI. In case the remaining starting material was detected, additional GDP-fucose, FUT6 and CIAP was added until the substrate was consumed. The reaction mixture was centrifuged using a Nanosep® Omega ultrafiltration device (3 kDa MWCO) to remove enzymes and the filtrate was lyophilized, followed by

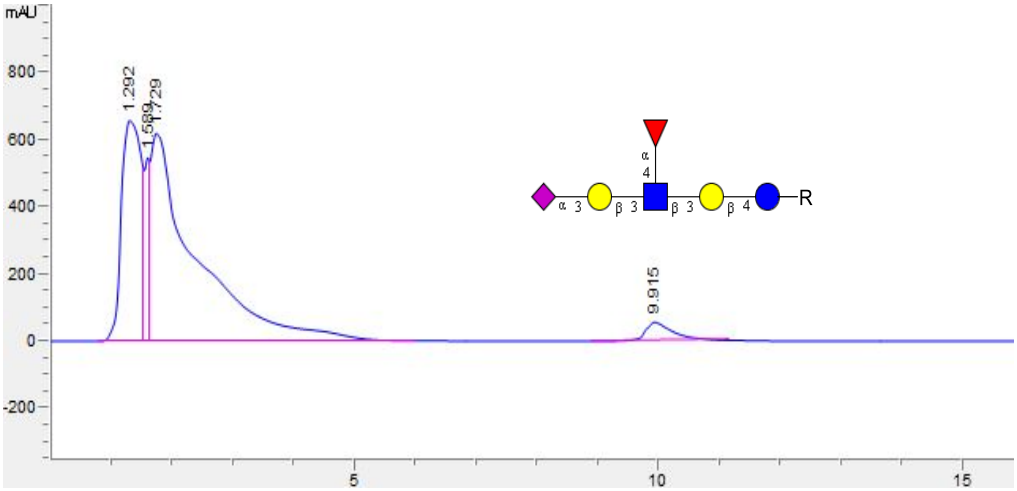
purification using reverse-phase HPLC equipped with Halo® RP-Amide 4.6 x 100mm, 2.7 µm column to give the desired product.

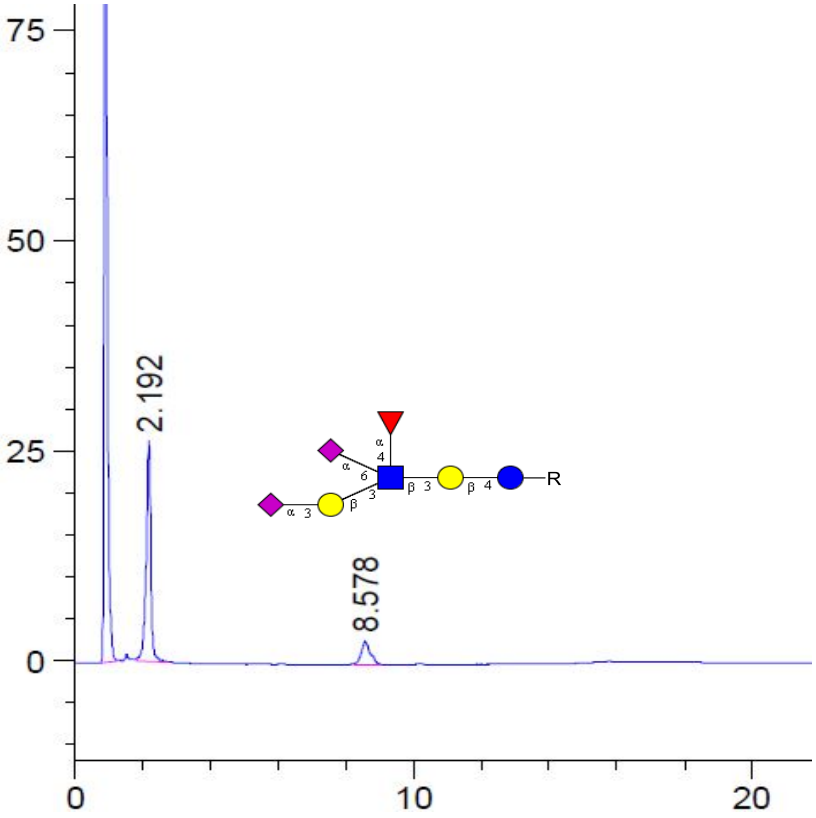
General procedure for the addition of α1,4 (and α1,3) Fuc using FUT9

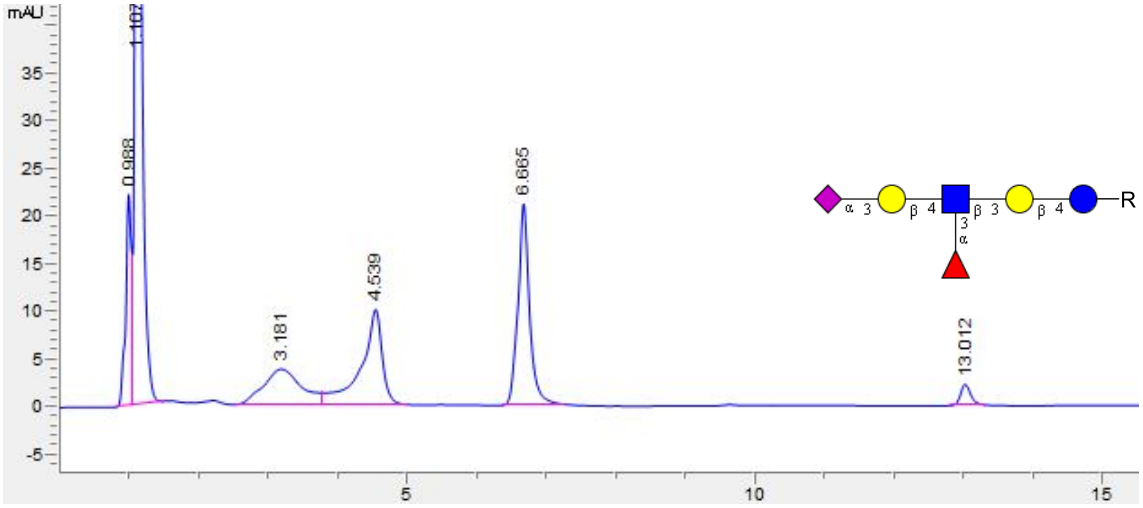
HMO acceptor and GDP-fucose (1.5 eq per fucose) were dissolved in a water buffer solution containing Tris-HCl (50 mM, pH 7.5), $MnCl_2$ (10 mM) buffer at a final concentration of 5 mM. To this solution was added CIAP (10 U μL^{-1}) and FUT9 (10 µg/µmol substrate). The reaction mixture was incubated overnight at 37 °C. Progress of the reaction was monitored by MALDI-TOF MS or ESI. In case the remaining starting material was detected, additional GDP-fucose, FUT9 and CIAP was added until the substrate was consumed. The reaction mixture was centrifuged using a Nanosep® Omega ultrafiltration device (3 kDa MWCO) to remove enzymes and the filtrate was lyophilized, followed by purification using reverse-phase HPLC equipped with Halo® RP-Amide 4.6 x 100mm, 2.7 µm column to give the desired product.

HPLC Traces




mAU
800
600
400
200
0
-200
1.292
1.589
1.729
9.915
α 4
α 3
β 3
β 3
β 4
R
5
10
15

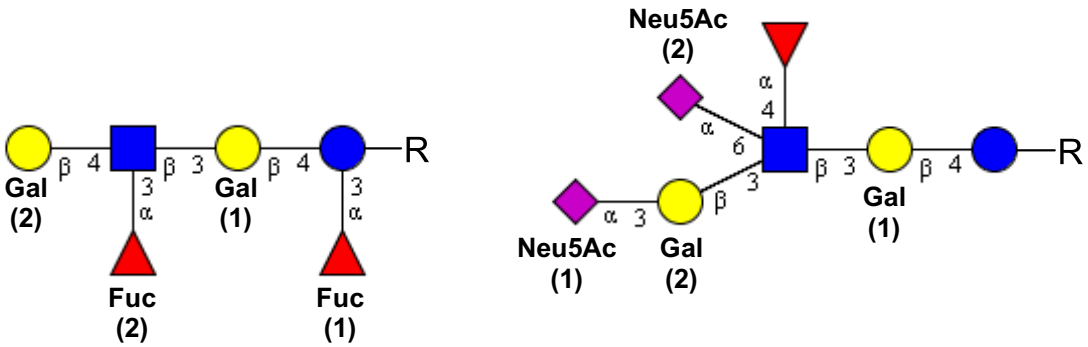

75
50
25
0
2.192
8.578
α 4
α 6
α 3
β 3
β 3
β 4
β
R
0
10
20

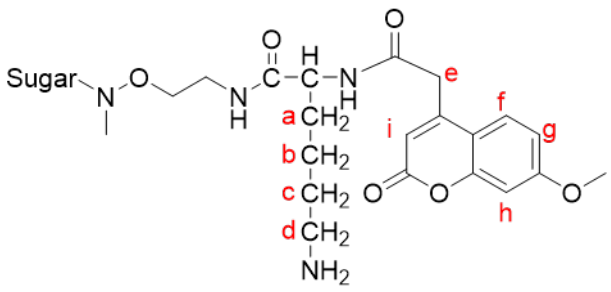


3. NMR and MS

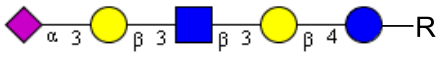
a. NMR nomenclature

Glycan assignments were made by numbering each monosaccharide starting fron reducing terminus and continuing in sequential order. Numbering for target difucosyl LNT is illustrated below in addition to the numbering of the anomeric linker.





MSLNT-linker (21)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	3.99	3.35	3.45	3.64	3.25	-[a]	n/a	n/a	n/a[b]	n/a
Gal (1)	4.27 (d, *J*=7.5 Hz, 1H)	3.44	3.56	3.98	3.51	-	n/a	n/a	n/a	n/a
GlcNAc	4.58 (d, *J*=8.5 Hz, 1H)	3.73	3.65	3.32	3.41	-	n/a	n/a	n/a	1.87 (s, 6H)
Gal (2)	4.35 (d, *J*=7.7 Hz, 1H)	3.38	3.92 (d, *J*=9.9 Hz, 1H)	3.78	3.57	-	n/a	n/a	n/a	n/a
Neu5Ac	n/a	n/a	2.60-eq (dd, *J*=12.4, 4.2 Hz) 1.62-axial	3.52	-	-	-	-	-	1.87 (s, 6H)

^{13}C from HSQC (150 MHz, D_2O): δ (ppm)

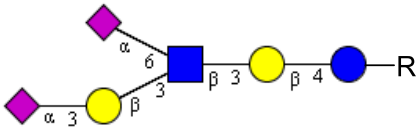
	C1
Glc	92.91
Gal(1)	102.72
GlcNAc	102.31
Gal(2)	103.22

[a] Not assigned

[b] Not applicable

ESI(*m/z*): $[M-H]^-$ calcd for $C_{58}H_{90}N_6O_{34}$, 1413.5420 found 1413.1875.

DSLNT-linker (22)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.00	3.37	3.46	3.61	3.55	-[a]	n/a	n/a	n/a[b]	n/a
Gal (1)	4.27 (d, *J*=7.8 Hz, 1H)	3.44	3.55	4.01	3.50	-	n/a	n/a	n/a	n/a
GlcNAc	4.55 (d, *J*=8.5 Hz, 1H)	3.74	3.64	3.47	3.40	-	n/a	n/a	n/a	1.88 (s, 3H)
Gal (2)	4.35 (d, *J*=7.8 Hz, 1H)	3.39	3.93 (d, *J*=9.9 Hz, 1H)	3.78	-	-	n/a	n/a	n/a	-
Neu5Ac (1)	n/a	n/a	2.59-eq 1.63-axial	3.52	-	-	n/a	n/a	n/a	1.87 (s, 6H)
Neu5Ac (2)	n/a	n/a	2.59-eq 1.63-axial	3.52	-	-	-	-	-	1.87 (s, 6H)

^{13}C NMR from HSQC (150 MHz, D_2O):

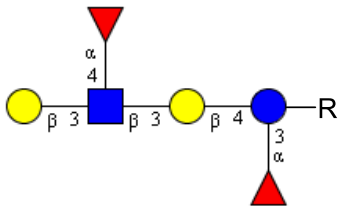
	C1
Glc	92.94
Gal(1)	102.79
GlcNAc	102.42
Gal(2)	103.33

[a] Not assigned

[b] Not applicable

ESI (*m/z*): $[M-2H]^{2-}$ calcd for $C_{69}H_{105}N_7O_{42}$, 851.8148 found 851.5860.

Bisfucosylated LNT-linker (23)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	3.99 (d, *J*=8.8 Hz, 1H)	3.57	3.64	3.34	-[a]	-	n/a[b]	n/a
Gal(1)	4.25 (d, *J*=7.2 Hz, 1H)	3.34	3.53	3.93	-	-	n/a	n/a
GlcNAc	4.53 (d, *J*=8.1 Hz, 1H)	3.79	3.91	3.60	3.38	-	n/a	1.87 (s, 3H)
Gal(2)	4.35 (d, *J*=6.9 Hz, 1H)	3.33	3.46	3.72	-	-	n/a	n/a
Fuc(1)	5.27	3.63	3.78	-	-	-	1.02 (d, J=6.5 Hz, 3H)	n/a
Fuc(2)	4.87	3.63	3.73	-	-	-	1.00 (d, J=6.4 Hz, 3H)	n/a

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

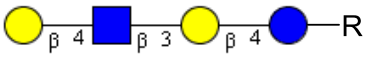
	C1
Glc	93.00
Gal(1)	101.62
GlcNAc	102.54
Gal(2)	102.75
Fuc(1)	98.40
Fuc(2)	97.85

[a] Not assigned

[b] Not applicable

MALDI-TOF-MS(*m/z*): $[M+Na]^+$ calcd for $C_{59}H_{93}N_5NaO_{34}$, 1438.5600 found 1438.9646.

LNnT-linker (27)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	3.99	3.36	3.46	3.64	-[a]	-	n/a[b]	n/a
Gal(1)	4.27 (d, *J*=7.9 Hz, 1H)	3.44	3.55	3.98	-	-	n/a	n/a
GlcNAc	4.55 (d, *J*=8.3 Hz, 1H)	3.64	3.58	3.43	-	-	n/a	1.90 (s, 3H)
Gal(2)	4.32 (d, *J*=7.7 Hz, 1H)	3.39	3.51	3.77	-	-	n/a	n/a

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

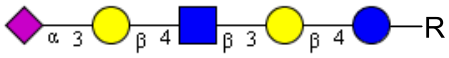
	C1
Glc	92.92
Gal(1)	101.81
GlcNAc	102.66
Gal(2)	102.79

[a] Not assigned

[b] Not applicable

MALDI-TOF-MS(*m/z*): $[M+Na]^+$ calcd for $C_{47}H_{73}N_5NaO_{26}$, 1146.4441 found 1146.4773.

MSLNnT-linker (28)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	3.99	3.36	3.46	-[a]	-	-	n/a[b]	n/a	n/a	n/a
Gal (1)	4.27 (d, *J*=7.7 Hz, 1H)	3.44	3.55	3.99	-	-	n/a	n/a	n/a	n/a
GlcNAc	4.54 (d, *J*=8.3 Hz, 1H)	3.64	3.58	3.43	-	-	n/a	n/a	n/a	1.85-1.90 6H
Gal (2)	4.40 (d, *J*=7.5 Hz, 1H)	3.42	3.96	3.80	-	-	n/a	n/a	n/a	-
Neu5Ac	n/a	n/a	2.60-eq 1.62-axial (dd, *J*=12.1 Hz, 1H)	3.53	3.70	3.53	-	-	3.47	1.85-1.90 6H

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

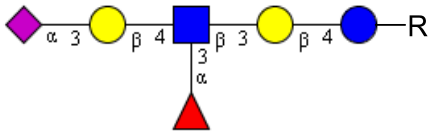
	C1
Glc	92.89
Gal(1)	102.77
GlcNAc	102.72
Gal(2)	102.44

[a] Not assigned

[b] Not applicable

ESI (*m/z*): $[M-H]^-$ calcd for $C_{58}H_{89}N_6O_{34}$, 1413.5420 found 1413.1384.

Monofucosylayed MSLNnT-linker (31)



^{1}H NMR (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	CH_3	NHAc
Glc	3.99	3.36	3.46	-[a]	-	-	n/a[b]	n/a	n/a	n/a	n/a
Gal (1)	4.27 (d, *J*=7.8 Hz, 1H)	3.43	3.55	4.00	-	-	n/a	n/a	n/a	n/a	n/a
GlcNAc	4.55 (d, *J*=7.7 Hz, 1H)	3.81	3.72	3.43	-	-	n/a	n/a	n/a	n/a	1.85-1.90 6H
Gal (2)	4.38 (d, *J*=7.5 Hz, 1H)	3.38	3.93	3.78	-	-	n/a	n/a	n/a	n/a	n/a
Fuc	4.97 (d, *J*=3.6 Hz, 1H)	3.53	3.75	3.63	-	-	n/a	n/a	n/a	1.02 (d, *J*=6.4 Hz, 3H)	n/a
Neu5Ac	n/a	n/a	2.61-eq 1.64-axial	3.53	3.70	-	-	-	3.47	n/a	1.85-1.90 6H

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

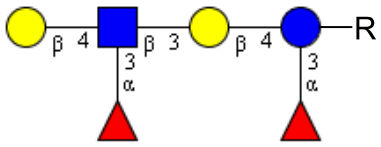
	C1
Glc	93.02
Gal(1)	102.69
GlcNAc	102.51
Gal(2)	101.46
Fuc	98.47

[a] Not assigned

[b] Not applicable

ESI(*m/z*): $[M-H]^-$ calcd for $C_{64}H_{99}N_6O_{38}$, 1559.5999 found 1559.1613.

Bisfucosylated LNnT-linker (29)



^{1}H NMR (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	3.99 (d, *J*=9.0 Hz, 1H)	3.57	3.64	3.35	-[a]	-	n/a[b]	n/a
Gal(1)	4.25 (d, *J*=7.7 Hz, 1H)	3.34	3.53	3.93	-	-	n/a	n/a
GlcNAc	4.54 (d, *J*=8.3 Hz, 1H)	3.81	3.72	3.42	3.58	-	n/a	1.87 (s, 3H)
Gal(2)	4.31 (d, *J*=7.8 Hz, 1H)	3.34	3.50	3.74	-	-	n/a	n/a
Fuc(1)	5.27 (d, *J*=3.7 Hz, 1H)	3.63	3.78	-	-	-	1.02 (d, J=6.5 Hz, 3H)	n/a
Fuc(2)	4.98 (d, *J*=3.8 Hz, 1H)	3.54	3.75	3.64	-	-	1.00 (d, J=6.4 Hz, 3H)	n/a

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

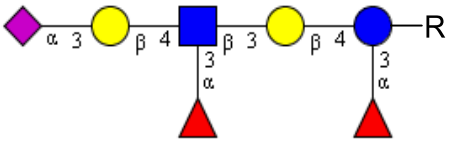
	C1
Glc	93.05
Gal(1)	101.64
GlcNAc	102.43
Gal(2)	101.68
Fuc(1)	98.39
Fuc(2)	98.47

[a] Not assigned

[b] Not applicable

ESI(*m/z*): $[M+H]^+$ calcd for $C_{59}H_{93}N_5O_{34}$, 1415.5702 found 1416.0019.

MSLNnT-bisfucosylated (32)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	CH_3	NHAc
Glc	4.00 (d, *J*=8.9 Hz, 1H)	3.57	3.35	3.67	-[a]	-	n/a[b]	n/a	n/a	n/a	n/a
Gal (1)	4.25 (d, *J*=7.8 Hz, 1H)	3.33	3.53	3.93	-	-	n/a	n/a	n/a	n/a	n/a
GlcNAc	4.53 (d, *J*=8.3 Hz, 1H)	3.81	3.71	3.43	-	-	n/a	n/a	n/a	n/a	1.85-1.90 6H
Gal (2)	4.38 (d, *J*=7.8 Hz, 1H)	3.37	3.77	3.93	-	-	n/a	n/a	n/a	n/a	n/a
Fuc(1)	5.27 (d, *J*=3.7 Hz, 1H)	3.63	3.78	3.60	-	-	n/a	n/a	n/a	1.02 (d, *J*=7.0 Hz, 3H)	n/a
Fuc(2)	4.97 (d, *J*=3.6 Hz, 1H)	3.52	3.74	3.62	-	-	n/a	n/a	n/a	1.00 (d, *J*=7.5 Hz, 3H)	n/a
Neu5Ac	n/a	n/a	2.61-eq 1.64-axial	3.51	3.70	-	-	-	-	n/a	1.85-1.90 6H

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

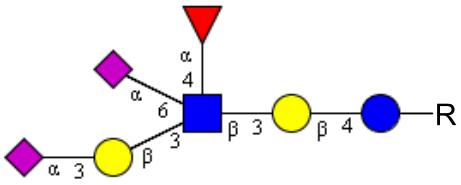
	C1
Glc	92.94
Gal(1)	102.78
GlcNAc	102.62
Gal(2)	102.72
Fuc(1)	98.41
Fuc(2)	98.47

[a] Not assigned

[b] Not applicable

ESI (m/z): $[M-2H]^{2-}$ calcd for $C_{70}H_{109}N_6O_{42}$, 852.8289 found 852.0721.

Monofucosylated DSLNT (26)



^{1}H NMR (600 MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	CH_3	NHAc
Glc	4.00	3.37	3.45	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a	n/a
Gal (1)	4.26 (d, *J*=7.8 Hz, 1H)	3.44	3.52	4.00	-	-	n/a	n/a	n/a	n/a	n/a
GlcNAc	4.53 (d, *J*=8.2 Hz, 1H)	3.79	3.91	3.63	3.47	-	n/a	n/a	n/a	n/a	1.85-1.90 9H
Gal (2)	4.38 (d, *J*=7.5 Hz, 1H)	3.36	3.76	3.89	-	-	n/a	n/a	n/a	n/a	n/a
Fuc(1)	5.02 (d, *J*=3.5 Hz, 1H)	3.65	3.70	3.62	-	-	n/a	n/a	n/a	1.02 (d, *J*=7.0 Hz, 3H)	n/a
Neu5Ac (1)	n/a	n/a	2.61-eq 1.57-axial	3.47	3.67	-	-	-	-	n/a	1.85-1.90 9H
Neu5Ac (2)	n/a	n/a	2.61-eq 1.61-axial	3.52	3.68	-	-	-	-	n/a	1.85-1.90 9H

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

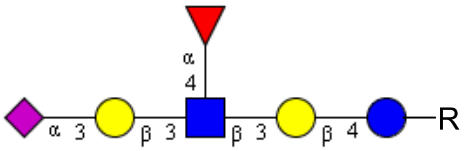
	C1
Glc	93.02
Gal(1)	102.69
GlcNAc	102.51
Gal(2)	101.46
Fuc	98.47

[a] Not assigned

[b] Not applicable

ESI(*m/z*): $[M-H]^-$ calcd for $C_{17}H_{115}N_7O_{46}$, 1849.6875 found 924.5865(2).

Monofucosylated MSLNT (25)



^{1}H NMR (600 MHz, D_2O): δ

	H1	H2	H3	H4	H5	H6	H7	H8	H9	CH_3	NHAc
Glc	3.99	3.33	3.43	3.52	3.62	-[a]	n/a[b]	n/a	n/a	n/a	n/a
Gal (1)	4.25 (d, *J*=7.7 Hz, 1H)	3.40	3.51	3.95	-	-	n/a	n/a	n/a	n/a	n/a
GlcNAc	4.52 (d, *J*=8.6 Hz, 1H)	3.76	3.90	3.57	3.35	-	n/a	n/a	n/a	n/a	1.85-1.90 6H
Gal (2)	4.37 (d, *J*=7.6 Hz, 1H)	3.33	3.86	3.72	-	-	n/a	n/a	n/a	n/a	n/a
Fuc	4.83 (d, *J*=3.5 Hz, 1H)	3.60	3.69	-	-	-	n/a	n/a	n/a	1.02 (d, *J*=7.0 Hz, 3H)	n/a
Neu5Ac	n/a	n/a	2.58-eq 1.59-axial	3.48	3.65	-	-	-	-	n/a	1.85-1.90 6H

^{13}C NMR from HSQC (150 MHz, D_2O): δ (ppm)

	C1
Glc	92.88
Gal(1)	102.75
GlcNAc	102.44
Gal(2)	102.67
Fuc	97.86

[a] Not assigned

[b] Not applicable

ESI(*m/z*): $[M-H]^-$ calcd for $C_{64}H_{99}N_6O_{38}$, 1559.5999 found 1559.1613.

References

1. Bode, L., Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **2012,** *22* (9), 1147-1162.

2. Triantis, V.; Bode, L.; Van Neerven, R., Immunological effects of human milk oligosaccharides. *Frontiers in pediatrics* **2018,** *6*, 190.

3. Schumacher, G.; Bendas, G.; Stahl, B.; Beermann, C., Human milk oligosaccharides affect P-selectin binding capacities: in vitro investigation. *Nutrition* **2006,** *22* (6), 620-627.

4. Barthel, S. R.; Gavino, J. D.; Descheny, L.; Dimitroff, C. J., Targeting selectins and selectin ligands in inflammation and cancer. *Expert opinion on therapeutic targets* **2007,** *11* (11), 1473-1491.

5. De Leoz, M. L. A.; Gaerlan, S. C.; Strum, J. S.; Dimapasoc, L. M.; Mirmiran, M.; Tancredi, D. J.; Smilowitz, J. T.; Kalanetra, K. M.; Mills, D. A.; German, J. B., Lacto-N-tetraose, fucosylation, and secretor status are highly variable in human milk oligosaccharides from women delivering preterm. *Journal of proteome research* **2012,** *11* (9), 4662-4672.

6. Bandara, M. D.; Stine, K. J.; Demchenko, A. V. The Chemical Synthesis of Human Milk Oligosaccharides: Lacto-N-Tetraose (Galβ1→ 3GlcNAcβ1→ 3Galβ1→ 4Glc). *Carbohydr. Res.* **2019**, *486*, 107824.

7. Craft, K. M.; Townsend, S. D. Synthesis of Lacto-N-Tetraose. *Carbohydr. Res.* **2017**, *440*, 43–50.

8. Ruiz-Palacios, G. M.; Cervantes, L. E.; Ramos, P.; Chavez-Munguia, B.; Newburg, D. S., Campylobacter jejuni binds intestinal H (O) antigen (Fuca1, 2Galb1, 4GlcNAc), and fucosyloligosaccharides of human milk ihibit its binding and infection. *Journal of Biological Chemistry* **2003**.

9. Guerrant, R. L.; Hughes, J. M.; Lima, N. L.; Crane, J., Diarrhea in developed and developing countries: magnitude, special settings, and etiologies. *Reviews of infectious diseases* **1990,** *12* (Supplement_1), S41-S50.

10. Holmes, E. H.; Macher, B. A.; Knegtel, R. M. A.; de Vries, T., Fucosyltransferases: structure/function studies. *Glycobiology* **2001,** *11* (10), 119R-128R.

11. Yao, W.; Yan, J.; Chen, X.; Wang, F.; Cao, H., Chemoenzymatic synthesis of lacto-N-tetrasaccharide and sialyl lacto-N-tetrasaccharides. *Carbohydrate research* **2015,** *401*, 5-10.

12. Prudden, A. R.; Liu, L.; Capicciotti, C. J.; Wolfert, M. A.; Wang, S.; Gao, Z.; Meng, L.; Moremen, K. W.; Boons, G.-J., Synthesis of asymmetrical multiantennary human milk oligosaccharides. *Proceedings of the National Academy of Sciences* **2017,** *114* (27), 6954-6959.

13. Chaturvedi, P.; Warren, C. D.; Altaye, M.; Morrow, A. L.; Ruiz-Palacios, G.; Pickering, L. K.; Newburg, D. S., Fucosylated human milk oligosaccharides vary between individuals and over the course of lactation. *Glycobiology* **2001,** *11* (5), 365-372.

14. Taniguchi, N.; Honke, K.; Fukuda, M., *Handbook of glycosyltransferases and related genes*. Springer Science & Business Media: 2011.

15. Calderon, A. D.; Liu, Y.; Li, X.; Wang, X.; Chen, X.; Li, L.; Wang, P. G., Substrate specificity of FUT8 and chemoenzymatic synthesis of core-fucosylated asymmetric N-

glycans. *Organic & biomolecular chemistry* **2016,** *14* (17), 4027-4031.

CHAPTER 3

CHEMOENZYMATIC SYNTHESIS OF ASYMMETRICAL BI-ANTENNARY HUMAN MILK OLIGOSACCHARIDES

Introduction:

Human milk oligosaccharides (HMOs) are structurally diverse, unconjugated glycans that are found in human breast milk and have been recognized for their role in infant health since the late nineteenth century.[1,2] Instead of providing nutrient directly to neonates, HMOs are utilized by beneficial bacteria, promoting the gastrointestinal health.[3] Studies indicate that certain human milk oligosaccharides (HMOs) from secretor are associated with various preventive effects, such as serving as soluble glycan receptor decoys for pathogen to prevent attachment to the mucosal surface.[2,4] HMOs also modulate the immune epithelial cell response. When incubated with certain HMO, 3'-sialyllactose, the human intestinal epithelial cell line reduced the expression of several sialyltransferases, leading to decreased binding of enteropathogenic E. coli which utilize sialylated glycans to attach cells.[5] As to immune modulation, Lewis glycan family, specifically sialylated Le blood group epitopes, commonly serves as a binding target for selectins.[6] By reducing leukocyte rolling and adhesion in an *in vitro* flow model with human endothelial cells, the count of perinucleolar compartment (PNC) formation, a prognostic marker of cancer cell malignancy, dropped. These properties suggest profound potential of HMOs as health-related molecules.

The state-of-the-art analytical methods have shown that HMOs containing the type I lacto-N-biose (Galβ1-3GlcNAc) predominate over those containing the type II lacto-N-biose (Galβ1-4GlcNAc).[7] Lactose can be extended with type I or type II structures, affording lacto-N-tetraose (LNT) and lacto-N-neotetraose (LNnT). LNT is one of the most abundant core structures of type I HMOs. Further variations of structure occur due to the attachment of N-acetyllactosamine (LacNAc), fucose and/or Neu5Ac residues at different positions on LNT or elongation of the chain. Among LNT-based structures, fucosylated species are of great interest in serving as free floating receptor analogs competing for bacterial binding.[8] The successful inhibition of binding suggested that HMOs and glycans on cell surface are synthesized by similar fucosyltransferases (FUTs) and thus having common epitopes. FUTs that have been studied were encoded by FUT gene family. Among the enzymes that catalyze the transfer of fucose to substrate in α-3 or α-4 manner, three FUTs (FUT3, FUT5 and FUT6) have high sequence similarity, whereas FUT4 and FUT9 are less similar to each other.[9] Diverse expression patterns endow differences whereas overlapping enzymatic properties, thus discriminating various enzymatic specificities requires delicate study to assess the substrate specificity patterns.

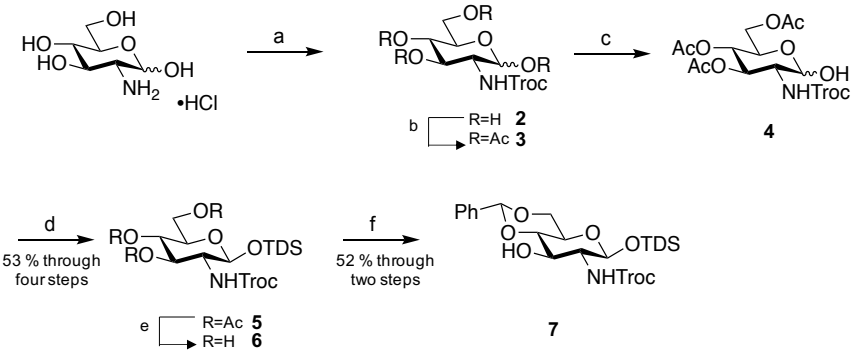
Current strategies to obtain substantial quantities of HMOs are redundant, such as pooled human milk extraction or metabolically engineering bacteria production. These methods either give poor yields or lack the ability to provide complex glycans.[10] In order to obtain in-depth knowledge of properties of HMOs at the molecular level, like glycan-protein binding specificities, as well as to investigate the metabolic pathways of intestinal microflora, well-defined HMO structures are required.

Pioneering research using HPLC-Chip/TOF MS have elucidated that HMO biosynthesis can provide a wide range of complex symmetrical and asymmetrical HMOs, which suggests that nature has employed an array of mammalian glycosyl transferases to construct structurally diverse glycans.[11] Inferred from the case of branched LNnT-based structures, the linear tetrasaccharide LNT may serve as substrate for branching enzyme GCNT2, giving a modification where β1,6-GlcNAc could be selectively added to the penultimate galactose of LNT.[12] However, in the pilot study, the branching enzyme failed to modify LNT, thus diverting the project to pursue an alternative path through which the branched LNT-based pentasaccharide with β1,6-GlcNAc was chemically synthesized. Given that the two branches are different, the β3-arm could be modified using diverse range of glycosyltransferases. The β6-arm could subsequently be extended with galactose using β4GalT1. The newly formed structure could further be elongated with repeated type II structures or type I structures, which could then diversify into different patterns of fucosylation and/or sialylation, affording complex asymmetric bi-antennary glycans. The aforementioned strategy allows the establishment of a panel of bi-antennary structures which will be used for future glycan-protein microarray binding.

Results and Discussion:

HMO Library Synthesis:

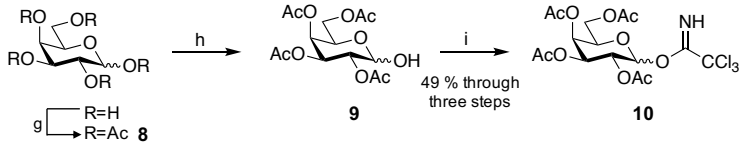
Branched penta-saccharide: The synthesis of penta-saccharide **1** (Scheme 3.7) began with retrosynthetic analysis of the target compound. The target compounds for saccharide building blocks are **7**, **10**, **11** and **22**. Through multi-step protection and deprotection strategy, each monosaccharide was chemically synthesized for subsequent glycosylation.



a) $NaHCO_3$, H_2O; TrocCl, r.t. b) Ac_2O, pyridine; c) hydrazine acetate, DMF; d) TDSCl, imidazole, DMF; e) Guanidinium chloride, $NaOCH_3$; CH_3OH/DCM; f) benzaldehyde dimethyl acetal, CSA, CH_3CN.

Scheme 3.1 Building Block Synthesis of GlcNAc Moiety I

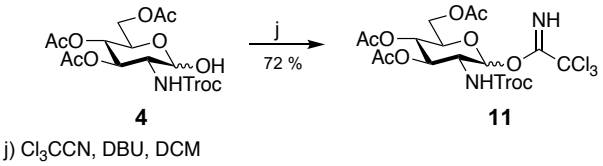
Modification of GlcNAc moiety started with Troc- group protection. Glucosamine hydrochloride was treated with TrocCl under mild base condition to give **2**, which was further per-acetylated using Ac_2O in pyridine providing **3**. The anomeric acetyl ester was selectively removed using hydrazine acetate and reprotected using TDSCl yielding **5**. Next, all remaining acetyl esters were hydrolyzed using sodium methoxide buffering with guanidinium chloride. The 4-OH and 5-OH of resulting **6** were simultaneously protected forming benzylidene protected **7**.



g) sodium acetate, Ac_2O; h) hydrazine acetate, DMF; i) Cl_3CCN, DBU, DCM

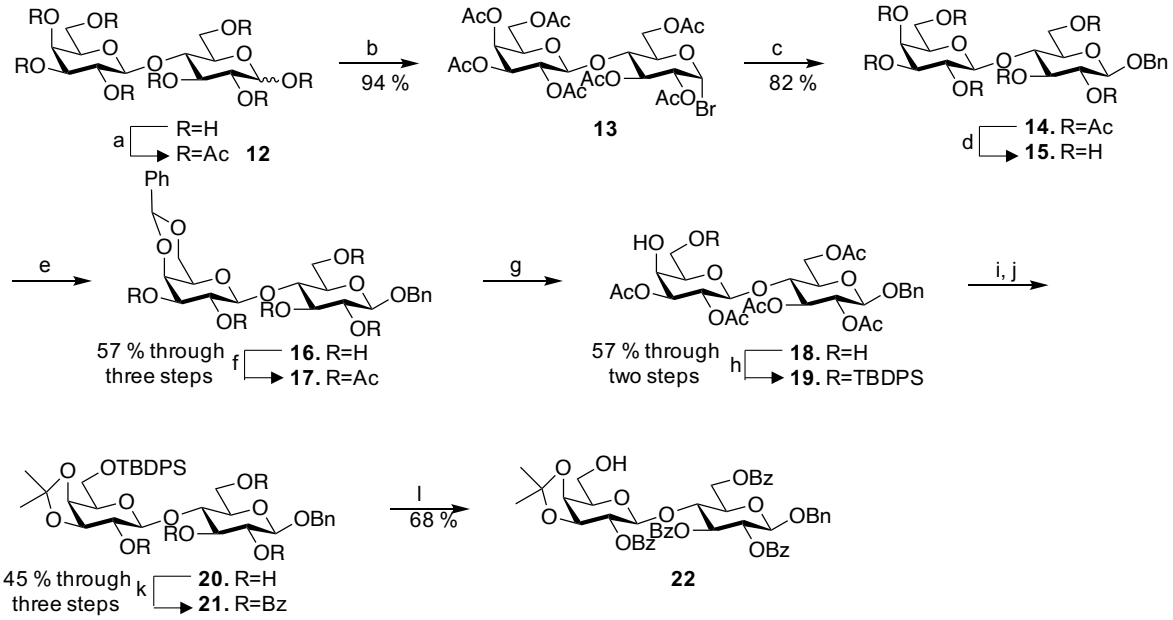
Scheme 3.2 Building Block Synthesis of Glucoside Moiety

Modification of galactose moiety began with per-acetylation, giving **8**. The anomeric acetyl was deprotected using hydrazine acetate. Then the resulting anomeric center was activated as a trichloroacetimidate species by reaction with trichloroacetonitrile in the presence of DBU, affording **10**.



Scheme 3.3 Building Block Synthesis of GlcNAc Moiety II

Another GlcNAc building block was obtained by subjecting intermediate **4** from the aforementioned process to trichloroacetonitrile and DBU giving activated glycosyl donor **11**.



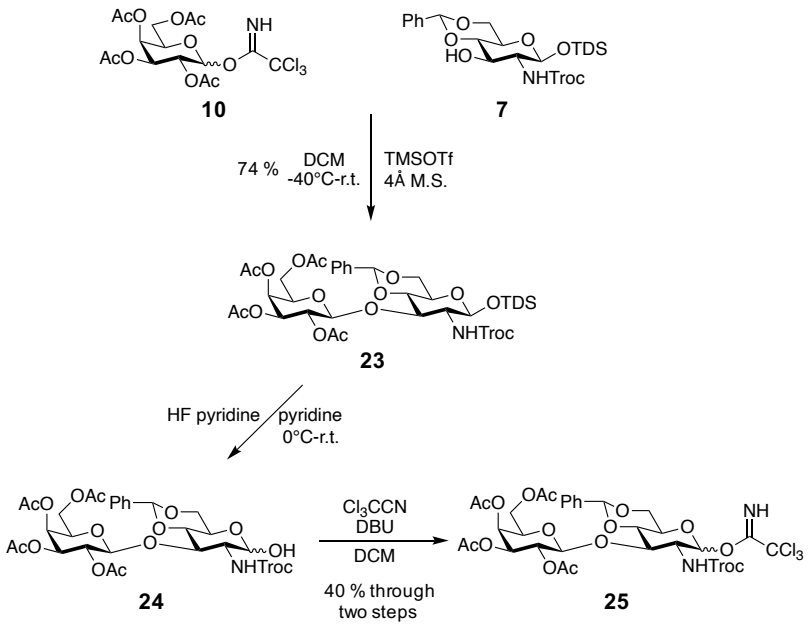
a) Ac_2O, pyridine; b) AcOH, Ac_2O, 33%HBr in AcOH; c) Ag_2CO_3, cat I_2, BnOH, DCM; d) DCM/CH_3OH1:3, $NaOCH_3$;
e) Benzaldehyde dimethyl acetal, p-TsOH, DMF, 45°C; f) Ac_2O, pyridine; g) DCM/TFA/H_2O 9:1:0.1; h) TBDPS-Cl, imidazole, DMF;
i) $NaOCH_3$, CH_3OH; j) 2,2-dimethoxypropane perchloric acid(cat.), acetone; k) BzCl, pyridine, 0°C-r.t.; l) HF pyridine, r.t.

Scheme 3.4 Building Block Synthesis of Lactoside Moiety

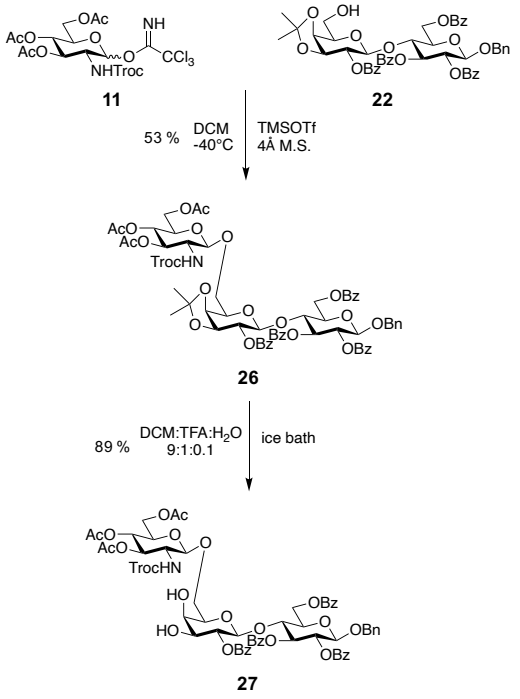
Modification of the lactose began with per-acetylation to give **12**, which was turned into a sugar halide through treatment with 33% HBr in acetic acid. An anomeric benzyl ether was then installed by reaction of bromide **13** with benzyl alcohol in the presence of silver carbonate forming **14**. Hydrolysis of the acetyl ester was carried out using sodium methoxide giving **15**, which was then subject to benzylidene acetal installation by reaction with benzaldehyde dimethyl acetal and p-TsOH in DMF to give **16**. The compound was

then re-acetylated, followed by selective deprotection of benzylidene using aqueous TFA. TBDPSCl was applied to the molecule to protect the 6-OH of the galactose moiety, after which the lactoside was de-acetylated. By using 2,2-dimethoxypropane under mild acid condition, iso-propylidene group was installed at 3-OH and 4-OH giving **20**. The orthogonal protection strategy between 3-, 4- and 6-OH allowed for sequential glycosylation reaction. The hydroxyls of compound **20** were protected as benzoyl esters using BzCl in pyridine, followed by deprotection of TBDPS group using HF/pyridine, to afford glycosyl acceptor **22**.

With all building blocks in hand, glycosylations could be carried out. The β-3 arm of the molecule consists of type I LacNAc moiety, was established through glycosylation using building block **10** and **7** in the presence of TMSOTf in DCM at -40 °C. (Scheme 3.5)



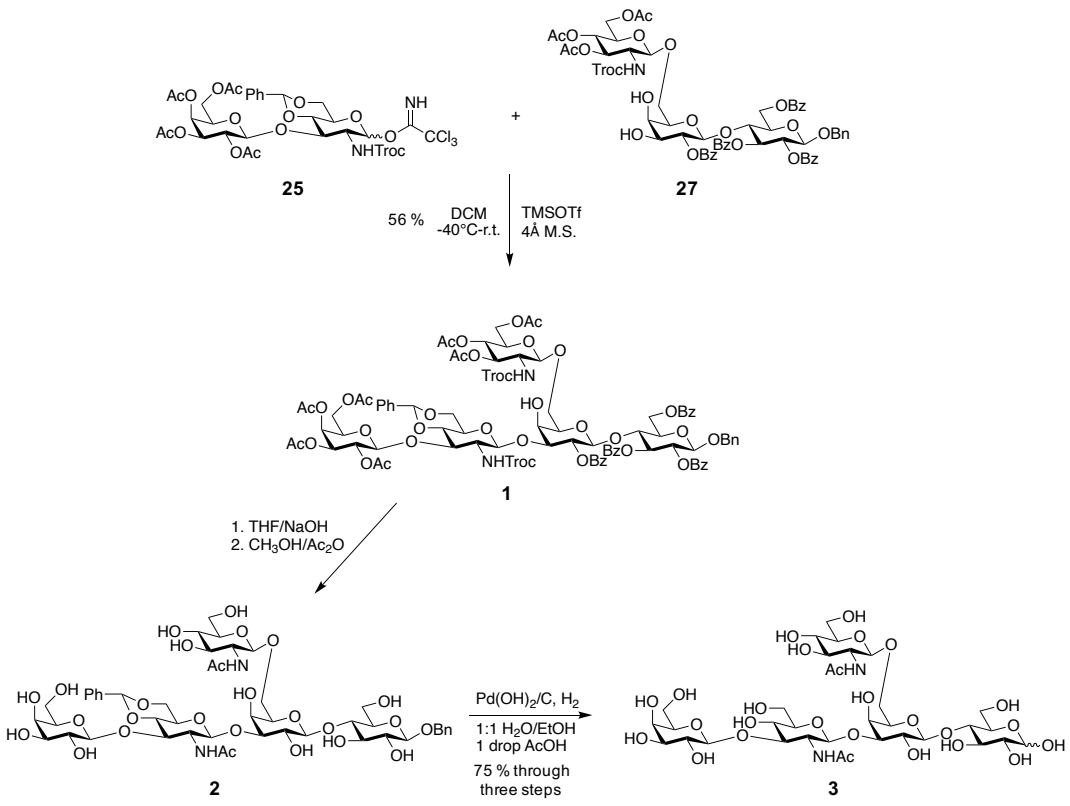
Scheme 3.5 Modular Glycosylation I for Building Disaccharide



Scheme 3.6 Modular Glycosylation II for Building Trisaccharide

The newly formed disaccharide **23** has a TDS ether at the anomeric center, which could be selectively removed using HF/pyridine, followed by installing a trichloroacetimidate, giving glycosyl donor disaccharide **25**.

The trisaccharide module was assembled through glycosylation of monosaccharide **11** and lactoside **22**. The isopropylidene protection was removed using TFA, giving **27** as glycosyl acceptor for final assembly of penta-saccharide **1**.

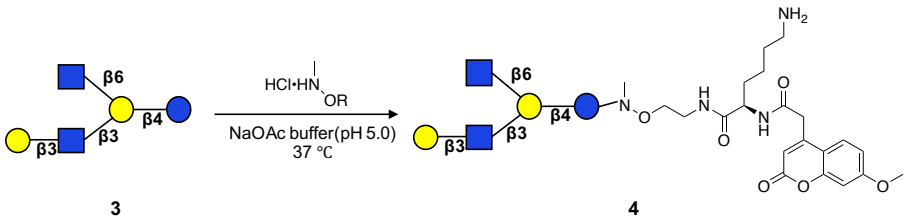


Scheme 3.7 Modular Glycosylation III for Building Pentasaccharide

The pentasaccharide was subject to strong base condition to remove all esters and Troc groups, followed selective acylation of the free amine of the GlcNAc moieties using Ac_2O in methanol. The benzyl esters were deprotected using hydrogenation strategy, affording pentasaccharide **3**.

Linker synthesis and conjugation:

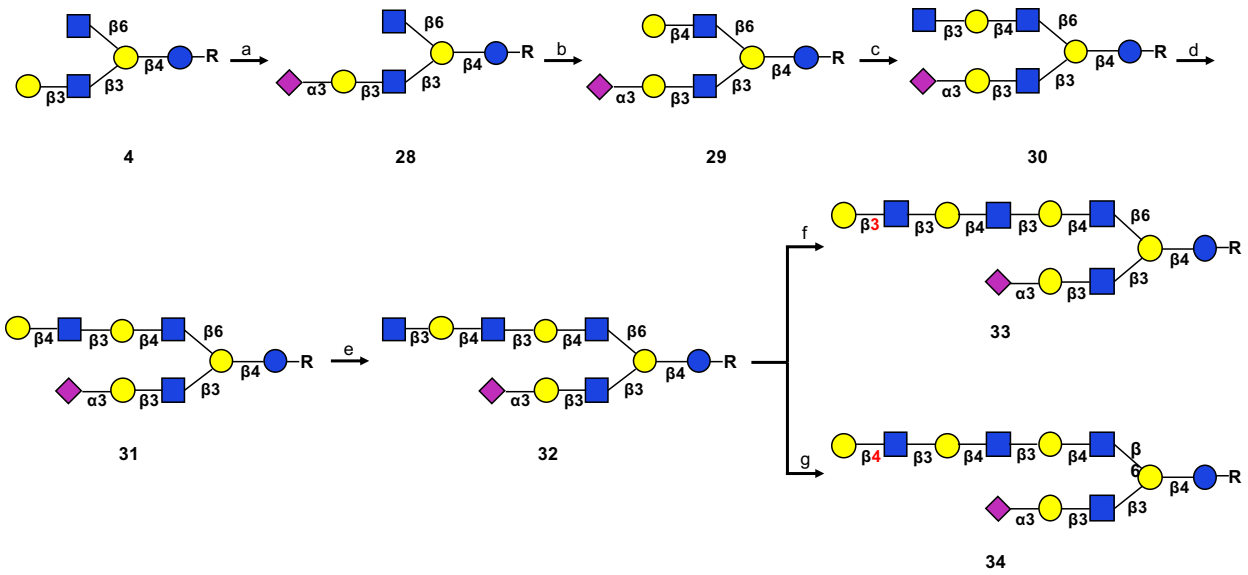
The anomeric Coumarin-linker was described by Prudden *et al.*[13] for the synthesis of type II containing complex HMOs. The linker was synthesized by the reported approach, and its conjugation with **3** was optimized in ammonium acetate buffer. 15 equivalents of linker had to be used to drive the reaction to completion. The product was purified *via* Bio-Gel P2 size exclusion chromatography, C-18 fast liquid chromatography and RP-amide HPLC purification.



Scheme 3.8 Linker ligation

Bi-antennary β3 sialylated HMOs:

As depicted in Scheme 3.8, the pentasaccharide has a type I LacNAc unit on the bottom β3 arm and a GlcNAc moiety on the top β6 arm. This difference made it possible to enzymatically modify each arm. Since many glycosyltransferases recognize the LacNAc but not GlcNAc, enzymatic modification started at the β3 arm. We classified the target HMO library into three groups based on different modifications of the β3 arm: α3 sialylated HMOs, α4 fucosylated HMOs, α2 fucosylated HMOs.



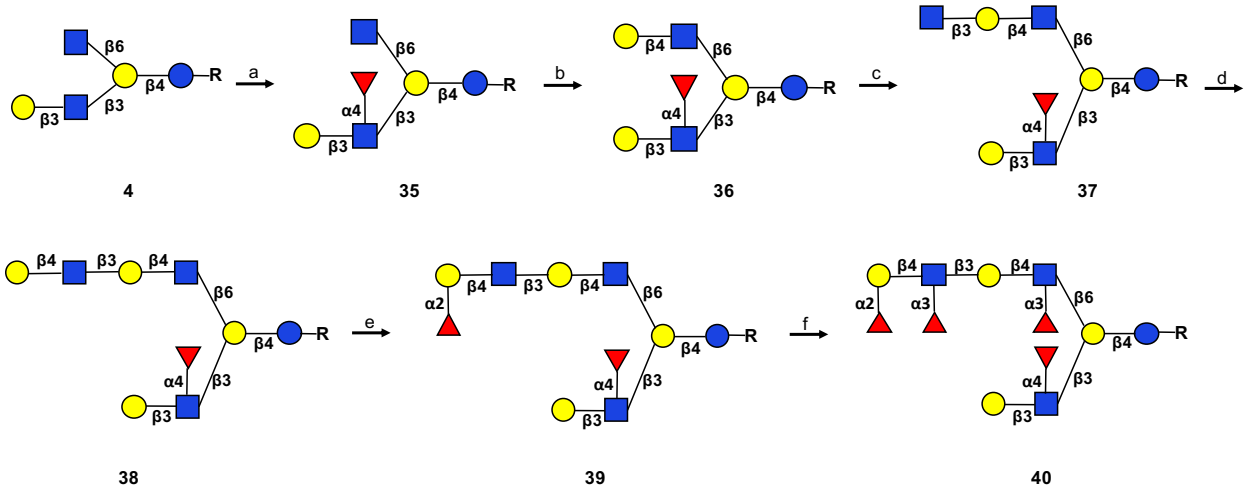
Scheme 3.9 Synthesis of Biantennary HMOs Bearing β3 3'sialyl type I LacNAc

a) PmST1, CMP-Neu5Ac; b) β4GalT1, UDP-Gal; c) β3GnT2, UDP-GlcNAc; d) β4GalT1, UDP-Gal; e) β3GnT2, UDP-GlcNAc; f) β3GalT5, UDP-Gal; g) β4GalT1, UDP-Gal.

For the first series, we aimed at modifying the β3 arm with α2,3-Neu5Ac to yield 3' sialylLacNAc, which blocks the bottom arm from further enzymatic modification and allowed selective elongation of the top β6 arm. Scheme 3.9 illustrates the enzymatic synthesis of bi-antennary library members with β3-3'sialylLacNAc. Starting with pentasaccharide **4**, pmST1 and CMP-Neu5Ac were used to selectively sialylate the β3 arm giving hexasaccharide **28** bearing a 3'sialylLacNAc terminus. Following the sialylation, the β6 arm was extended by treatment with β4GalT1 and UDP-Gal to form heptasaccharide **29** bearing a type II structure, which could be extended into a dimeric LacNAc on β6 arm by subjecting **29** sequentially with β3GnT2/UDP-GlcNAc and β4GalT1/UDP-Gal to give **31**. The dimeric arm could be further elongated into trimeric LacNAc by sequential treatment with either β3GnT2/β4GalT1 or β3GnT2/β3GalT5, affording type II **34** or type I LacNAc **33**, respectively.

Bi-antennary β3-α4 fucosylated HMOs:

By applying a similar modular synthetic strategy as outlines above, a collection of HMOs containing Lewis a epitope could be established on the β3 arm with LacNAc units and multi-fucosylation patterns on the β6 antenna.



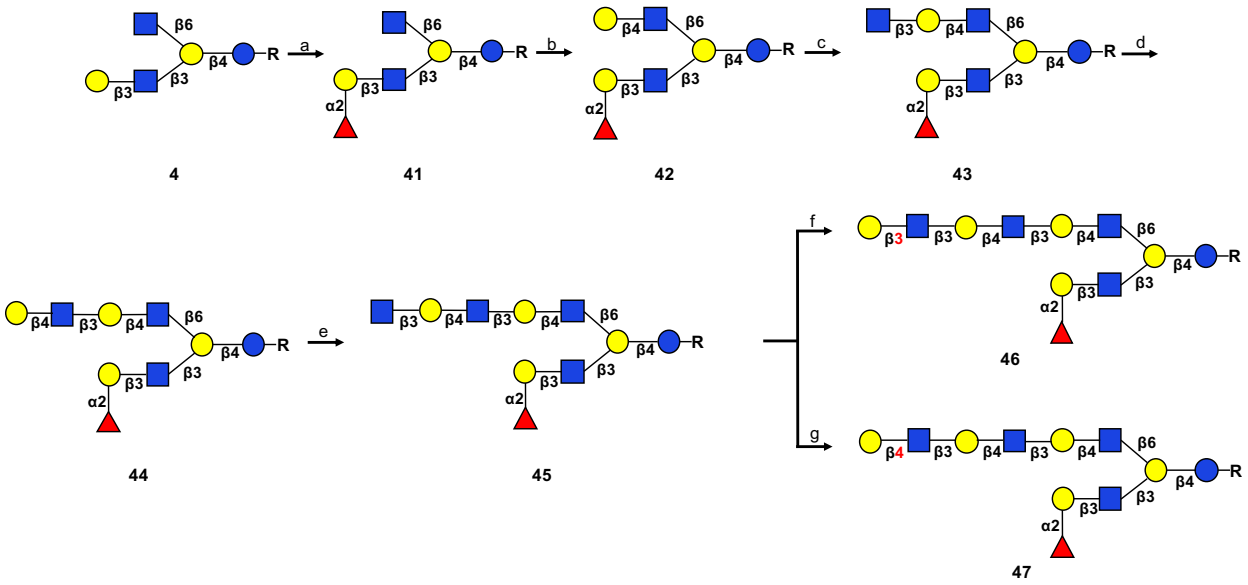
Scheme 3.10 Synthesis of Biantennary HMOs Bearing β3 Lewis a Epitope

a) FUT5, GDP-fucose; b) β4GalT1, UDP-Gal; c) β3GnT2, UDP-GlcNAc; d) β4GalT1, UDP-Gal; e) FUT1, GDP-fucose; f) FUT5, GDP-fucose.

Starting with pentasaccharide **4**, FUT3 and GDP-Fucose installed α1,4-fucose to GlcNAc moiety on β3 arm yielding **35**. The presence of the fucoside prohibited further extension of the β3 arm so that the β6 branch could be selectively converted into type II LacNAc by treatment with β4GalT1 and UDP-Gal giving **36**. The β6 arm could then be extended into a dimeric LacNAc by subjecting **36** sequentially to β3GnT2/UDP-GlcNAc and β4GalT1/UDP-Gal to give **38**. Treatment with FUT1 and GDP-fucose endowed a H2 epitope at type II terminus of the β6 dimeric arm. To obtain a highly fucosylation pattern of HMO targets, FUT5 and GDP-Fucose were employed to install two fucosyl moieties at the terminal and penultimate GlcNAc on β6 branch, giving target compound **40**.

Bi-antennary β3-α2 fucosylated HMOs:

Starting with pentasaccharide **4**, FUT1 and GDP-Fucose installed an α1,2-fucoside at the Gal moiety of the β3 arm yielding **41**. The presence of terminal fucose blocked the extension of β3 branch. Following the same modular strategy, the β6 position was converted into type II LacNAc moiety with β4GalT1 and UDP-Gal giving **42**. Sequential reaction with β3GnT2/UDP-GlcNAc and β4GalT1/UDP-Gal gave a dimeric LacNAc motif on β6 arm of **44**. The β6 antenna could be capped with either type I or type II structures at the dimeric LacNAc moiety using β3GnT2/β3GalT5 or β3GnT2/β4GalT1, respectively.

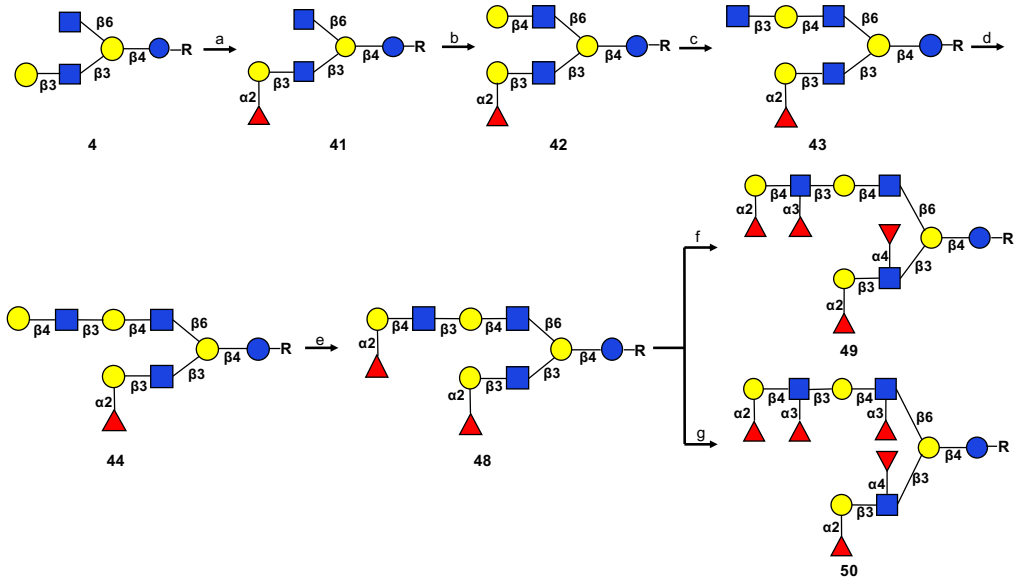


Scheme 3.11 Synthesis of Biantennary HMOs Bearing β3 H1 Epitope

a) FUT1, GDP-fucose; b) β4GalT1, UDP-Gal; c) β3GnT2, UDP-GlcNAc; d) β4GalT1, UDP-Gal; e) β3GnT2, UDP-GlcNAc; f) β3GalT5, UDP-Gal; g) β4GalT1, UDP-Gal.

With dimeric structure on β3 branch, high-fucosylation pattern could be achieved by capping the dimer with α1,2-fucose installed at terminal Gal on β3 arm and as such, exposure of FUT3 delivered two fucosylation to ultimate GlcNAc on β3 branch and GlcNAc on β3 arm to give **49**. FUT5 install three fucose residues on both GlcNAcs on β6

arm and GlcNAc on β3 arm to give **50**. Both targets have Lewis b structures on β3 branches. For the β6 branch, FUT3 gave a Lewis^y structure while FUT5 gave a Lewis^y and Lewis^b structures.



Scheme 3.12 Synthesis of Biantennary HMOs Bearing H1 and Lewis Epitopes

a) FUT1, GDP-fucose; b) β4GalT1, UDP-Gal; c) β3GnT2, UDP-GlcNAc; d) β4GalT1, UDP-Gal; e) FUT1, GDP-fucose; f) FUT3, GDP-fucose; g) FUT5, GDP-fucose.

Conclusion:

A panel of bi-antennary HMOs was constructed using a chemoenzymatic synthetic strategy. These structures are of high binding potential based on the pioneering library collection of HMOs built with type II LacNAc moiety on the β3 branch. In the latter case, 60 HMOs were screened against bacterial and viral adhesion protein together with two human galectins. Glycans having the strongest binding were bi-antennary structures. The fact that type I containing structures are predominantly found in human breast milk, we focus on synthesizing an array of well-defined HMO structures bearing different glycol-epitopes on β3 and β6 antenna, which serves as a valuable collection of compounds for future binding studies with a diverse set of proteins.

Experimental

General Methods and Reagents. All reagents, unless otherwise stated, were purchased from Sigma-Aldrich. ^{1}H and ^{13}C NMR spectra were recorded on Varian Inova 300, 500 or 600MHz. Chemical shifts are reported in parts per million (ppm) relative to H1 of reducing glucose which was set to δ4.00. NMR data is represented as follows: Chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, dd = doublet of doublets, m = multiplet and/or multiple resonances), *J* coupling, integration, and peak identity. All NMR signals were assigned on the basis of ^{1}H NMR, gCOSY, gHSQC, zTOCSY and NOESY experiments. Mass spectra were recorded on an Applied Biosystems SCIEX MALDI TOF / TOF 5800 mass spectrometer using 2,5-dihydroxybenzoic acid (DHB) as a matrix or a high-resolution Shimadzu 20AD LCMS-IT-TOF mass spectrometer. Reagents were purchased from Sigma-Aldrich (unless otherwise noted) and used without further purification. HPLC purification of compounds was performed on an Agilent Technologies 1200 Series HPLC equipped with a Halo® RP-Amide 4.6x100 mm, 2.7 µm column. HPLC grade acetonitrile and water were purchased from Fisher. RP column chromatography was accomplished using Bondapak® C18 125Å 37-55 µm bulk packing material (Waters Corp) Neutral and sialylated compounds were eluted in a 50mM solution of ammonium bicarbonate with a linear gradient of 0~50%, 0~20% or 0~25% acetonitrile over 30 min, respectively at 1 mL/min. Spectra were monitored at a wavelength of 324nm using an Agilent 1200 Series diode array with a multiple wavelength detector. Column chromatography was performed on silica gel G60 (Silicycle, 60-200 µm, 60 Å). Thin layer chromatography (TLC) was conducted on Silica gel 60 F_{254} (EMD Chemicals Inc.) with detection by UV light absorption (254 nm) where applicable. Visualization was

accomplished by immersed the plate into 10% sulfuric acid in ethanol or Ninhydrin, followed by charring using heat gun. All reactions were carried out under argon atmosphere unless specified. Molecular sieves were flame-dried *in vacuo* immediately prior to use.

General biochemical reagents

Calf intestine alkaline phosphate (CIAP) was purchased from BioLabs® Inc. Pasteurella multocida α-2,3-sialyltransferase 1 (PmST1) was purchased from Chemily Glycoscience. Cytidine-5'-monophospho-N-acetylneuraminic acid (CMP-Neu5Ac), uridine 5'-diphosphogalactose (UDP-Gal) and uridine 5'-diphospho-N-acetylglucosamine (UDP-GlcNAc) were purchased from Roche Life Science. Fucosyltransferases were provided by Kelley Dr. Moremen's (CCRC, UGA).

General Enzymatic Reaction Procedures:

Addition of β1,3 GlcNAc:

HMO acceptor (10 μmol, 1 eq) was dissolved in HEPES buffer (pH 7.3, 50 mM) containing UDP-GlcNAc (1.5 eq), KCl (25 mM), $MgCl_2$ (2 mM), DTT (1 mM), and Calf Intestine Alkaline Phosphotase (CIAP, 10 units). The volume was adjusted to give a final acceptor concentration of 10 mM. To this solution was added 83 μg of beta-1,3-N-acetylglucosaminyltransferase 2 (β3GNT2). The reaction mixture was incubated overnight at 37 °C. Reaction progress was monitored by MALDI-TOF MS or ESI. In case of remaining starting material was detected, additional UDP-GlcNAc, β3GNT2 and CIAP was added until all starting material was consumed.

Addition of β1,4 Galactoside:

HMO acceptor (10 µmol, 1 eq) was added Tris buffer (pH 7.3, 50 mM) containing UDP-Gal (1.5 eq), $MnCl_2$ (10 mM), BSA (40 µg) and Calf Intestine Alkaline Phosphotase (CIAP, 10 units). The volume was adjusted to give a final acceptor concentration of 10 mM. To this solution was added 50 µg of beta-1,4-galactosyl transferase from Bovine Milk (β4GalT1). The reaction mixture was incubated overnight at 37 °C. Reaction progress was monitored by MALDI-TOF MS or ESI. In case of remaining starting material, additional UDP-Gal, β4GalT1 and CIAP was added until all starting material was consumed.

Addition of β1,3 Galactoside:

HMO acceptor (10 µmol, 1 eq) was added to sodium cacodylate buffer (pH 7.5, 150 mM) containing UDP-Gal (1.5 eq), $MnCl_2$ (10 mM), BSA (40 µg) and Calf Intestine Alkaline Phosphotase (CIAP, 10 units). The volume was adjusted to give a final acceptor concentration of 10 mM. To this solution was added 82 µg of beta-1,3-galactosyl transferase, Polypeptide 5 (β3GalT5). The reaction mixture was incubated overnight at 37 °C. Reaction progress was monitored by MALDI-TOF MS or ESI. In case of remaining starting material, additional UDP-Gal, β3GalT5 and CIAP was added until all starting material was consumed.

Addition of α1,2-fucoside:

HMO acceptor (10 µmol, 1 eq) was added Tris buffer (pH 7.5, 150 mM) containing GDP-fucose (1.5 eq), $MnCl_2$ (10 mM) and Calf Intestine Alkaline Phosphotase (CIAP, 10 units). The volume was adjusted to give a final acceptor concentration of 10 mM. To this solution was added 100 µg of α-1,2-fucosyltransferase (FUT1). The reaction mixture was

incubated overnight at 37 °C. Reaction progress was monitored by MALDI-TOF MS or ESI. In case of remaining starting material, additional GDP-fucose, FUT1 and CIAP was added until all starting material was consumed.

Addition of α1,3-fucoside

HMO acceptor (10 µmol, 1 eq) was added Tris buffer (pH 7.3, 50 mM) containing GDP-fucose (1.5 eq), $MnCl_2$ (10 mM) and Calf Intestine Alkaline Phosphotase (CIAP, 10 units). The volume was adjusted to give a final acceptor concentration of 10 mM. To this solution was added 100 µg of α-1,3/4-fucosyltransferase (FUT3 or FUT5). The reaction mixture was incubated overnight at 37°C. Reaction progress was monitored by MALDI-TOF MS or ESI. In case of remaining starting material, additional GDP-fucose, FUT3 or FUT5 and CIAP was added until all starting material was consumed.

Addition of α1,4-fucoside

HMO acceptor (10 µmol, 1 eq) was added to a Tris buffer (pH 7.3, 50 mM) containing GDP-fucose (1.5 eq), $MnCl_2$ (10 mM) and Calf Intestine Alkaline Phosphotase (CIAP, 10 units). The volume was adjusted to give a final acceptor concentration of 10 mM. To this solution was added 100 µg of α-1,3/4-Fucosyltransferase (FUT3). The reaction mixture was incubated overnight at 37 °C. Reaction progress was monitored by MALDI-TOF MS or ESI. In case of remaining starting material was detected, additional GDP-fucose, FUT3 and CIAP was added until all starting material was consumed.

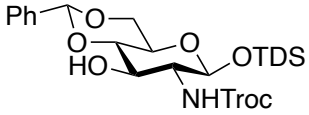
Addition of α2,3-Neu5Ac

HMO acceptor (10 µmol, 1 eq) was added to Tris buffer (pH 8.5, 100 mM) containing CMP-Neu5Ac (1.5 eq). The volume was adjusted to give a final acceptor concentration of 5 mM. To this solution was added 100 µg PmST1. The reaction mixture was incubated

at 37 °C for 30 min. After ESI result showed that the optimal conversion of substrate has achieved. The reaction was quenched by heating up the reaction to 65 °C for 5 min to inactive PmST1.

Chemical Synthesis Procedures

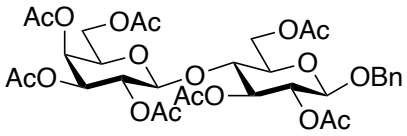
Dimethylthexylsilyl 4,6-O-benzylidene-2-[(2,2,2-trichloroethoxy) carbonylamino]-2-deoxy-β-D-glucopyranoside (7)



1H-imidazole (4.0 g, 58.4 mmol) was added to a solution **4** (14.0 g, 29.2 mol) in DMF (70 mL) followed by dropwise addition of TDSCl (8.6 mL, 43.8 mol). The reaction mixture was stirred at room temperature overnight. Upon completion of the reaction indicated by TLC (ethyl acetate/hexane, 3:1, v/v), the DMF was evaporated *in vacuo* and residue was dissolved in DCM (120 mL) and washed with saturated NH_4Cl solution (2 x 50 mL). The organic layers were combined, dried ($MgSO_4$) and concentrated *in vacuo*, and the residue was purified by silica gel chromatography (ethyl acetate/hexane, 2:3, v/v) to give **5** as white solid. The obtained compound was dissolved in DCM (21 mL). Guanidinium chloride (1.87 g, 19.6 mmol) was added to sodium methoxide/methanol solution (3 M, 10 mL), after which the mixture was added to the DCM solution of **5**. The reaction mixture was stirred for 30 mins and progress of the reaction was monitored by silica gel coated TLC (CH_3OH/DCM, 1:10, v/v). The reaction was quenched by adding Dowex-H^+ resin. DCM was evaporated and the residue dried *in vacuo* to give **6**, which was dissolved in anhydrous acetonitrile (19 mL) and benzaldehyde dimethyl acetal (1 mL, 6.5 mmol) and CSA (275 mg, 1.2 mmol) were added. The reaction mixture was stirred overnight at room

temperature, after which TLC (EtOAc/hexane, 1:3, v/v) analysis showed completion of the reaction. After neutralization with saturated $NaHCO_3$ solution, the organic layer was separated and concentrated *in vacuo*. The residue was applied to silica gel chromatography for purification to give **7** as white solid. ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.45-7.52 (m, 2H, aromatic), 7.32-7.42 (m, 3H, aromatic), 5.54 (s, 1H, *CH*Ph), 4.87 (d, *J* = 7.7 Hz, 1H, H-1), 4.63-4.73 (m, 2H, O*CH*$_2$CCl$_3$), 4.08 (m, 1H, H-3), 3.78 (m, 1H, H-5), 3.57 (m, 1H, H-4), 3.56 (d, *J* = 8.92 Hz, 1H, H-6a), 3.46 (d, *J* = 9.25 Hz, 1H, H-6b), 3.36 (m, 1H, H-2), 1.24-1.26 (m, 1H, TDS), 0.16 (m, 6H, TDS), 0.80-0.90 (m, 12 H, TDS); ^{13}C NMR (75 MHz, $CDCl_3$): δ 3.6, 2.0, 19.6, 20.5, 21.0, 25.8, 35.6, 61.8, 67.5, 69.7, 72.0, 76.2, 83.4, 95.8, 97.5, 103.2, 128.3, 128.4, 131.0, 137.9, 155.1. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{24}H_{36}Cl_3NO_7Na$, 606.1224 found 606.1092.

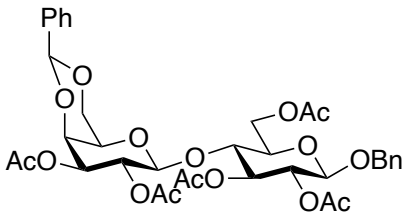
Benzyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-acetyl-β-D-glucopyranoside (14)



Sodium acetate (12 g, 146 mmol) was added to Ac_2O (165.7 mL, 1.7 mol) and solution was brought to reflux in oil bath. The flask was temporarily removed from heat and lactose (50 g, 146 mmol) was added in portion. The reaction mixture was kept refluxing for additional 30 min. Progress of the reaction was monitored by silica gel-coated TLC (7:3 ethyl acetate/hexane). The mixture was poured into ice water (500 mL) upon completion. Gum-like sticky compound formed with mechanical stir. Bulky mass gradually dispersed after stirring for additional 20 min. The reaction mixture was stirred for additional 1.5 h. Reaction mixture was diluted with DCM (200 mL) and washed with brine (1 x 100 mL),

saturated $NaHCO_3$ (2 x 100 mL). The organic layer was combined, dried ($MgSO_4$) and concentrated *in vacuo*, followed by recrystallization from ethanol, giving per acetylated lactose **12**. **12** was dissolved in glacial acetic acid (6.3 mL) and acetic anhydride (6.3 mL). The mixture was brought to 0°C. HBr (33%wt in AcOH, 21 mL) was added dropwise over a period of 10 min with the aid of dropping funnel. The reaction mixture was kept stirring for 1 h at 0°C, presenting a yellow homogeneous solution. Silica gel TLC (7:3 ethyl acetate/hexane) indicated the completion of reaction. The reaction mixture was concentrated, co-evaporated 3 times with toluene and dried on vacuum line to give bromosugar **13**. The crude bromosugar (19.3 g, 27.6 mol) was dissolved in anhydrous DCM (80 mL). Freshly made Ag_2CO_3 (19.1 g, 69.1 mol), iodine (a crystal) and 4Å activated dry molecular sieves (3 g) were suspended in DCM (80 mL) to give promoter suspension. The bromosugar solution was added into the suspension of promoters followed by addition of BnOH (15 mL, 138 mmol). The reaction mixture was stirred overnight at room temperature. Silica gel TLC (1:1 ethyl acetate/hexane) indicated reaction completion. Ag_2CO_3, molecular sieves and iodine were filtered out. Resulting mixture was concentrated *in vacuo*, applied to silica gel chromatography and purified using 2:1 EtOAc and hexane as eluent affording **14**. The NMR was in accordance with previous obtained data (Chapter 2).

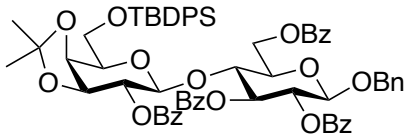
Benzyl 4,6-O-benzylidene-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (17)



Compound **14** (5.2 g, 7.1 mmol) was dissolved in a methanol/DCM solvent mixture (106 mL, 3:1, v/v). To the solution was added a 5 M $NaOCH_3$ methanol solution until pH~10. The reaction mixture was stirred overnight at room temperature. MALDI-TOF MS indicated completion of the reaction. The reaction was quenched using Dowex-H^+ resin, which was filtered afterwards, and the filtrate was concentrated *in vacuo* to give crude deprotected benzyl-lactoside **15**. The latter was dissolved in DMF (25 mL). To this solution was added benzaldehyde dimethyl acetal (2 mL, 13.4 mmol) and p-TsOH (160 mg, 0.8 mmol). The reaction mixture was heated (45 °C) and kept stirring overnight with a venting system attached. Silica gel TLC (EtOAc/CH_3OH, 9:1, v/v) indicated completion of the reaction. The reaction mixture was concentrated *in vacuo* to give crude **16**. Compound **16** was dissolved in pyridine (5.6 mL, 69.2 mmol), followed by dropwise addition of Ac_2O (3.6 mL, 38.4 mmol). The reaction mixture was stirred overnight at room temperature. Silica gel TLC (EtOAc/hexane, 1:1, v/v) indicated completion of the reaction. The reaction mixture was diluted with EtOAc (80 mL) and washed with water (3 x 50 mL) and brine (2 x 50 mL). The organic layers were combined, dried ($MgSO_4$) and concentrated *in vacuo*. The residue was applied to silica gel chromatography (EtOAc/hexane, 2:3, v/v) purification to give **17** (2.98 g, 57% through three steps). ^1H NMR (300 MHz, $CDCl_3$): 7.49-7.22 (m, 10H, aromatic), 5.24(dd, *J* = 8.0, 10.3 Hz, 1H, H-2, Gal), 5.18 (dd, *J* = 9.4, 9.4 Hz, 1H, H-3, Glc), 4.99 (dd, *J* = 7.9, 9.7 Hz, 1H, H-2, Glc), 4.88 (dd, *J* = 3.7, 10.3 Hz, 1H, H-3, Gal), 4.54 (m, 1H, H-6a, Glc), 4.52 (d, *J* = 7.9 Hz, 1H, H-1, Glc), 4.47 (d, *J* = 7.9 Hz, 1H, H-1, Gal), 4.32 (dd, *J* = 0.7, 3.7 Hz, 1H, H-4, Gal), 3.80 (dd, *J* = 9.3, 9.8 Hz, 1H, H-4, Glc), 3.58 (ddd, *J* = 1.9, 4.9, 6.9 Hz, 1H, H-5, Glc), 2.13 (s, 3H, COC*H*$_3$), 2.03 (s, 3H, COC*H*$_3$), 2.02 (s, 6H, COC*H*$_3$), 1.98 (s, 3H, COC*H*$_3$). ^{13}C NMR

(75 MHz, $CDCl_3$): δ 128.26, 126.46, 101.10, 100.79, 99.00, 73.12, 73.34, 72.81, 72.09, 71.78, 70.62, 70.62, 68.87, 68.74, 68.74, 62.52, 62.42, 21.27, 20.11. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{36}H_{42}O_{16}Na$, 753.2371 found 753.2361.

Benzyl 2-O-benzoyl-3,4-O-isopropylidene-6-O-tert-Butyldiphenylsilyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-glucopyranoside (21)

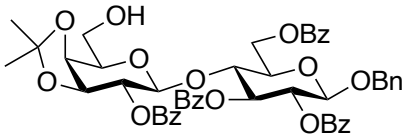


Compound **16** (1.2 g, 1.9 mmol) was dissolved in pyridine (2.8 mL), followed by the addition of Ac_2O (1.8 mL, 19.2 mmol). The reaction mixture was stirred overnight at room temperature. Silica gel TLC confirmed completion of the reaction (EtOAc/hexane, 1:1, v/v). The solvents were evaporated and the residue was diluted with ethyl acetate (80 mL), washed with 1M HCl (2 x 150 mL), brine (2 x 80 mL) and water (2 x 100 mL). The organic layers were combined, dried ($MgSO_4$) and concentrated *in vacuo*. The residue was dissolved in a mixture of DCM (11.7 mL) and H_2O (130 µL), after which it was brought to 0 °C. To the solution was added TFA (1.3 mL, 17.0 mmol) and the reaction mixture was kept stirring for another 1 h at 0 °C. Silica gel TLC indicated completion of the reaction (EtOAc/hexane, 3:2, v/v). The reaction was quenched by the addition of Et_3N and was concentrated *in vacuo* to give compound **18**. **18** (903 mg, 1.4 mmol) and 1H-imidazole (239 mg, 3.5 mmol) was dissolved in DMF (9.3 mL). To the solution was added TBDPS-Cl (402 µL, 1.5 mmol) under argon protection. The reaction was stirred overnight at room temperature. Silica gel TLC indicated completion of the reaction (EtOAc/hexane, 2:3, v/v). The organic solvent was evaporated and the residue was diluted with ethyl acetate, washed with sat. NH_4Cl (2 x 80 mL), brine (2 x 80 mL). The organic layers were combined,

dried ($MgSO_4$) and concentrated *in vacuo*. The residue was purified by silica gel chromatography using 1:2 EtOAc/hexane as eluent to give **19** (2.0 g, 57% through two steps). Compound **19** was dissolved in CH_3OH/DCM solvent mixture (9.3 mL, 3:1, v/v). To the solution was added 1M $NaOCH_3$ methanol solution until pH~8.5, after which the reaction mixture was stirred overnight at room temperature. MALDI-TOF MS indicated completion of the reaction. The reaction was quenched using Dowex-H^+ resin, which was filtered off after neutralization. The filtrate was concentrated *in vacuo*, giving a crude product to proceed to next step without further purification. To an acetone (15 mL) solution of the aforementioned crude compound was added 2,2-dimethoxypropane (3.8 mL, 33.2 mmol) forming a cloudy suspension. The mixture was brought to 0 °C, followed by dropwise addition of catalytic amount of perchloric acid (18 µL). After 5 min, silica gel-coated TLC indicated full consumption of the starting material. The reaction was quenched by the addition of Et_3N. The solvents were evaporated giving **20** as crude product used immediately in the next step. It was dissolved in pyridine (1.1 mL, 13.1 mmol) and cooled to 0 °C. BzCl (1.2 g, 11.2 mmol) was added dropwise, after which the reaction mixture was gradually warmed to room temperature and stirring was continued overnight. Silica gel TLC analysis indicated completion of the reaction (EtOAc/hexane, 1:3, v/v). The reaction was quenched with methanol, followed by concentration *in vacuo*. The mixture was washed with 1M HCl (2 x 100 mL), brine (2 x 100 mL). The organic layers were combined, dried ($MgSO_4$) and concentrated *in vacuo*. The residue was applied to silica gel chromatography using 1:6 EtOAc/hexane as eluent, providing **21** as white foam. (1.15 g, 45%) ^{1}H NMR (300 MHz, $CDCl_3$): δ 6.78-8.06 (m, 35H, aromatic), 5.53 (dd, *J* = 9.2, 9.5 Hz, 1H, H3, Gal), 5.42 (dd, *J* = 7.7, 9.7 Hz, 1H, H-2, Gal), 5.09 (dd, *J* = 7.5, 7.5 Hz, 1H,

H-2, Glc), 4.64 (d, *J* = 7.7 Hz, 1H, H-1, Gal), 4.52 (m, 1H, H-6a, Gal), 4.47 (d, *J* = 8.0 Hz, 1H, H-1, Glc), 4.19 (dd, *J* = 5.7, 6.7 Hz, 1H, H-3, Glc), 4.06 (dd, *J* = 9.4*,* 9.4 Hz, 1H, H-4, Gal), 3.70 (m, 1H, H-5, Gal), 1.21-1.29 (m, 6H, $(CH_3)_2CO_2$), 0.90-1.00 (s, 9H, $(CH_3)_3CSi$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 135.48, 129.50, 129.64, 129.29, 128.34, 128.20, 127.86, 127.72, 100.59, 98.91, 76.82, 75.80, 73.06, 72.89, 72.71, 72.39, 71.87, 70.49, 70.32, 62.62, 60.93, 29.00, 28.01, 27.75, 26.42, 26.34. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{66}H_{66}O_{15}SiNa$, 1149.4069 found 1149.4115.

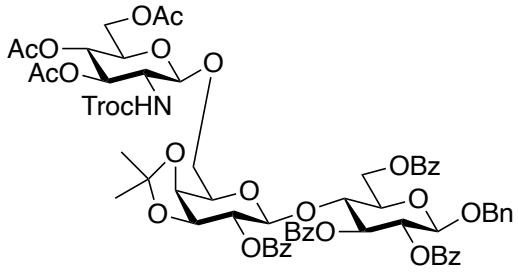
Benzyl 2-O-benzoyl-3,4-O-isopropylidene-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzoyl-β-D-glucopyranoside (22)



Compound **21** (300 mg, 0.3 mmol) was dissolved in pyridine (2.4 mL) and cooled (0 °C). HF pyridine (200 µL) was added dropwise and the reaction mixture was gradually warmed to room temperature, after which the stirring was continued overnight. TLC analysis indicated strating material remaining and another portion of HF pyridine (100 µL) was added and stirring was continued for additional 6 hours. The reaction mixture was diluted with EtOAc and washed with 5M $CuSO_4$ solution (2 x 60 mL), brine (2 x 80 mL). The organic layers were combined, dried ($MgSO_4$) and concentrated *in vacuo*. The residue was purified by silica gel chromatography (EtOAc/hexane, 2:3, v/v) to give **22** (160 mg, 68%). ^1H NMR (300 MHz, $CDCl_3$): δ 7.05-8.11 (m, 25H, aromatic), 5.65 (dd, *J* = 9.1, 9.1 Hz, 1H, H-3, Gal), 5.54 (dd, *J* = 7.9, 8.9 Hz, 1H, H-2, Gal), 5.18 (dd, *J* = 7.4. 7.4 Hz, 1H, H-2, Glc), 4.83 (d, *J* = 12.5 Hz, 1H, CH*H*Ph), 4.73 (d, *J* = 7.6 Hz, 1H, H-1, Gal), 4.67 (d, *J* = 11.8 Hz, 1H, H-6b, Gal), 4.63 (d, *J* = 7.9 Hz, 1H, H-1, Glc), 4.60 (d, *J* = 12.6 Hz, 1H,

C*H*HPh), 4.50 (dd, *J* = 4.6, 4.6 Hz, 1H, H-6a, Gal), 4.24 (m, 2H, Gal H-4, Glc H-3), 4.02 (d, *J* = 5.6 Hz, 1H, H-4, Glc), 3.83 (dd, *J* = 3.2, 9.0 Hz, 1H, H-5, Gal), 3.60 (m, 1H, H-5, Glc), 3.38 (dd, *J* = 4.0, 11.9 Hz, 1H, H-6b, Glc), 3.24 (dd, *J* = 7.6, 11.8 Hz, 1H, H-6a, Glc), 1.49-1.59 (s, 3H, *CH*$_3$CH3CO$_2$), 1.21-1.32 (s, 3H, CH3*CH*$_3$CO$_2$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 133.32, 133.21, 133.21, 129.85, 129.69, 128.50, 128.39, 128.39, 128.07, 127.95, 100.51, 98.99, 76.70, 73.91, 73.55, 73.55, 73.40, 73.05, 71.83, 70.43, 70.43, 62.75, 62.75, 61.71, 27.53, 26.25, 26.06. MALDI-TOF-MS (*m/z*): [M+Na] $^+$ calcd for $C_{50}H_{48}O_{15}Na$, 911.2891 found 911.2931.

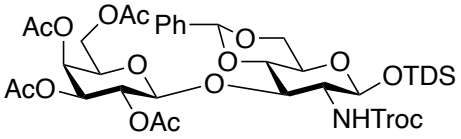
3,4,6-tri-O-acetyl-2-deoxy-2-trichloroethoxyacetamido-β-D-glucopyranosyl-(1→6)-2-O-benzoyl-3,4-O-isopropyl-β-D-galactopyransoyl-(1→4)-β-2,3,6-tri-O-benzoyl-1-O-benzyl-D-glucopyranoside (26)



A mixture of donor **11** (78.7 mg, 126.0 μmol), acceptor **22** (56 mg, 63.0 μmol) and 4Å activated molecular sieves (50 mg) were stirred in anhydrous DCM (1.1 mL) for 5 min. The suspension was cooled to -40°C followed by TMSOTf (1.5 μL, 8.5 μmol) addition. The reaction mixture was kept stirring at 40°C for 10 min, after which it was slowly warmed to room temperature. Silica gel TLC indicated completion of the reaction, which was quenched with Et_3N. The molecular sieves were filtered off. The mixture was concentrated *in vacuo* and purified by silica gel column chromatography using 3:7 EtOAc/Toluene as eluent to give **26** as white foam (44 mg, 53 %). ^{1}H NMR (500 MHz,

$CDCl_3$): δ 7.10-8.02 (m, 25H, aromatic), 5.57 (dd, *J* = 8.6, 8.6 Hz, 1H, H-3, Gal), 5.49 (m, 1H, H-2, Gal), 5.29 (dd, *J* = 9.7, 9.7 Hz, 1H, H-3, GlcNAc), 5.10 (dd, *J* = 7.1, 7.1 Hz, 1H, H-2, Glc), 5.06 (dd, *J* = 9.6, 9.6 Hz, 1H, H-4, GlcNAc), 4.79 (d, *J* = 12.4 Hz, 1H, C*H*HPh), 4.78 (m, 1H, C*H*HCCl$_3$), 4.74 (d, *J* = 7.0 Hz, 1H, H-1, Gal), 4.62 (m, 2H, Gal H-6b, Glc H1), 4.56 (d, J = 12.6 Hz, 1H, CH*H*Ph), 4.53 (m, 1H, CH*H*CCl$_3$), 4.42 (dd, *J* = *4.7,* 4.8 Hz, 1H, H-6a, Gal), 4.28 (d, *J* = 7.9 Hz, 1H, H-1, GlcNAc), 4.27 (m, 1H, H-6a, GlcNAc), 4.23 (dd, *J* = 8.9, 8.9 Hz, 1H, H-4, Gal), 4.18 (m, 1H, H-3, Glc), 4.10 (m, 1H, H-6b, GlcNAc), 3.99 (dd, *J* = 1.4, 5.1 Hz, 1H, H-4, Glc), 3.81 (m, 1H, H-5, Gal), 3.74 (m, 1H, H-6a, Glc), 3.67 (m, 1H, H-2, GlcNAc), 3.58 (1H, H-5, GlcNAc), 3.21 (dd, *J* = 7.1, 7.1 Hz, 1H, H6b, Glc), 2.08 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.02 (s, 3H, $COCH_3$), 1.50 (s, 3H, $CH_3CH_3CO_2$), 1.25 (s, 3H, $CH_3CH_3CO_2$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 133.37, 133.18, 133.08, 129.72, 128.80, 128.32, 128.32, 128.05, 100.93, 100.23, 98.99, 76.73, 75.00, 74.11, 74.11,73.54, 73.54, 73.10, 73.10, 72.08, 71.94, 71.79, 71.67, 70.51, 70.51, 68.75, 68.33, 68.19, 62.98, 62.98, 61.97, 61.97, 56.02, 29.45, 27.49, 26.12, 20.82, 20.70, 20.70. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{65}H_{66}Cl_3NO_{24}Na$, 1372.2938 found 1374.1135.

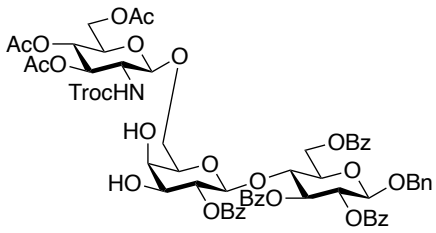
Thexyl dimethylsilyl O-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-(1→3)-4,6-O-benzylidene-2-deoxy-2-(2,2,2-trichloroethoxycarbonylamino)-β-D-glucopyranoside (23)



A mixture of donor **10** (84 mg, 170.9 μmol), acceptor **7** (50 mg, 85.5 μmol) and 4Å activated molecular sieves (80 mg) were suspended in anhydrous DCM (1.7 mL) and stirred for 5 min. The mixture was cooled to -40°C and TMSOTf (1.5 μL, 8.5 μmol) was

added and reaction was continued stirring for 10 min, after which it was slowly warmed to room temperature. The completion of reaction was indicated by silica gel TLC (EtOAc/toluene, 1:5, v/v) after which the reaction was quenched with Et_3N, followed by filtering off molecular sieves. The residue was concentrated *in vacuo* and purified by silica gel column chromatography using 1:5 EtOAc/hexane as eluent to give **23** (78.2 mg, 74%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.50-7.24 (m, 5H, aromatic), 5.52 (s, 1H, PhC*H*), 5.28 (d, *J* = 2.2 Hz, 1H, H-4, Gal), 5.18 (dd, *J* = 8.8, 10.0 Hz, 1H, H-2, Gal), 5.02 (d, *J* = 7.8 Hz, 1H, H-1, GlcN), 4.91 (dd, *J* = 3.4, 10.5 Hz, 1H, H-3, Gal), 4.70 (m, 1H, CH*H*CCl_3), 4.67 (m, 2H, H-1, Gal, C*H*$HCCl_3$), 4.73-4.63 (m, 2H, NHCOOC*HH*), 4.30 (d, *J* = 9.4, 9.4 Hz, 1H, H-3, GlcN), 4.25 (dd, *J* = 5.1, 10.7 Hz, 1H, H6a, GlcN), 4.04 (dd, *J* = 9.6, 9.6 Hz, 1H, H6b, Gal), 3.83 (dd, *J* = 5.3, 10.5 Hz, 1H, H6a, Gal), 3.77 (dd, *J* = 10.3, 10.3 Hz, 1H, H6b, GlcN), 3.70 (dd, *J* = 9.1, 9.1 Hz, 1H, H-4, GlcN), 3.62 (m, 1H, H-5, Gal), 3.45 (ddd, *J* = 4.7, 9.4, 14.3 Hz, 1H, H-5, GlcN), 3.25 (dd, *J* = 8.1, 17.2 Hz, 1H, H-2, GlcN), 1.88-2.14 (m, 12H, CO*CH*$_3$), 0.79-0.89 (m, 13H, CH and CH_3 of thexyl), 0.09-0.15 (m, 6H, Si*CH*$_3$*CH*$_3$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 128.55, 126.04, 125.78, 101.38, 100.80, 95.29, 80.02, 77.96, 75.04, 74.50, 70.93, 70.51, 69.26, 68.86, 68.58, 66.80, 66.09, 61.13, 60.73, 60.31, 33.99, 29.25, 20.77, 20.77, 20.60, 19.99, 18.52. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{38}H_{54}Cl_3NO_{16}SiNa$, 936.2175 found 938.2997.

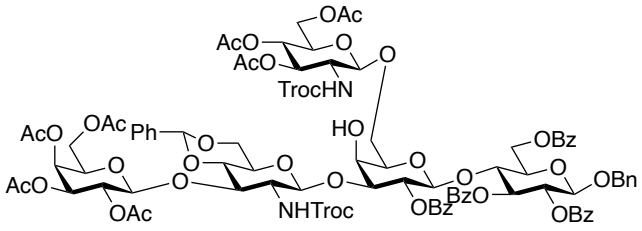
3,4,6-tri-O-acetyl-2-deoxy-2-trichloroethoxyacetamido-β-D-glucopyranosyl-(1→6)]-2-O-benzoyl-β-D-galactopyransoyl-(1→4)-β-2,3,6-tri-O-benzoyl-1-O-benzyl-D-glucopyranoside (27)



Compound **26** (44 mg, 32.6 μmol) was dissolved in a mixed system (V_{total} = 666 μL) consisting of DCM/TFA/H_2O (9:1:0.1). The reaction mixture was stirred at 0 °C for 1.5 h, after which silica gel TLC indicate completion of the reaction (EtOAc/toluene, 3:7, v/v). The reaction was quenched by the addition of Et_3N and the mixture was then concentrated *in vacuo*. The residue was purified by silica gel column chromatography (EtOAc/hexane, 1:3, v/v), providing **27** as white foam.(38 mg, 89%) ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.09-8.07 (m, 25H, aromatic), 5.58 (dd, *J* = 9.1, 9.1 Hz, 1H, H-3, Gal), 5.47 (dd, *J* = 8.5, 8.5 Hz, 1H, H-2, Gal), 5.26 (dd, *J* = 9.8, 9.8 Hz, 1H, H-3, GlcN), 5.20 (dd, *J* = 8.2, 9.6 Hz, 1H, H-2, Glc), 4.97 (dd, *J* = 9.7, 9.7 Hz, 1H, H-4, GlcN), 4.82 (d, *J* = 12.0 Hz, 1H, C*H*HCCl$_3$), 4.80 (d, *J* = 12.2 Hz, 1H, C*H*HPh), 4.70 (d, *J* = 7.7 Hz, 1H, H-1, Gal), 4.63 (d, *J* = 12.3 Hz, 1H, CH*H*CCl$_3$), 4.57 (m, 1H, CH*H*Ph), 4.49 (dd, *J* = 4.5, 11.9 Hz, 1H, H-6a, Gal), 4.58 (m, 1H, H-6b, Gal), 4.54 (d, *J* = 7.9 Hz, 1H, H-1, Glc), 4.22 (d, *J* = 7.9 Hz, 1H, H-1, GlcN), 4.18 (m, 2H, H6a, H6b, GlcN), 4.08 (dd, *J* = 9.2, 9.2 Hz, 1H, H-4, Gal), 3.80 (m, 1H, H-4, Glc), 3.78 (m, 1H, H-5, Gal), 3.63 (m, 1H, H-5, GlcN), 3.58 (m, 1H, H-3, Glc), 3.49 (dd, *J* = 9.1, 18.3 Hz, 1H, H-2, GlcN), 3.26 (dd, *J* = 7.0 Hz, 1H, H-5, Glc), 3.09 (m, 2H, H6a, H6b, Glc), 2.00-2.20 (m, 9H, COOCH$_3$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 1133.28, 133.25, 133.22, 132.99, 129.80, 129.80, 129.80, 128.55, 128.41, 128.41, 128.03, 127.96,

101.11, 100.57, 98.83, 76.44, 74.41, 74.41, 73.61, 73.61, 72.94, 72.52, 72.10, 71.85, 71.85, 71.58, 70.35, 70.35, 68.60, 67.52, 66.15, 62.65, 62.65, 61.95, 55.75, 53.78, 31.74, 29.29, 20.66, 20.66, 20.66. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{62}H_{62}Cl_3NO_{24}Na$, 1332.2625 found 1334.1396.

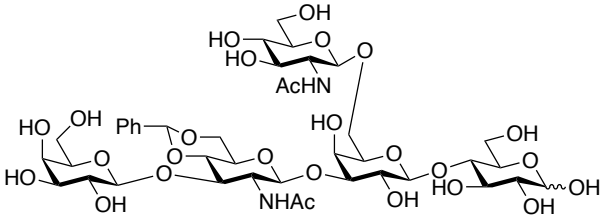
2,3,4,6-tetra-O-acetyl-D-galactopyranosyl-(1→3)-β-4,6-O-benzylidene-2-trichloroethoxyacetamido-2-deoxy-D-glucopyranosyl-(1→3)-[3,4,6-tri-O-acetyl-2-deoxy-2-trichloroethoxyacetamido-β-D-glucopyranosyl-(1→6)]-2-O-benzoyl-β-D-galactopyransoyl-(1→4)-β-2,3,6-tri-O-benzoyl-1-O-benzyl-D-glucopyranoside (2)



A mixture of donor **25** (90 mg, 98.1 μmol), acceptor **27** (99 mg, 75.5 μmol) and 4Å activated molecular sieves (90 mg) were suspended in DCM (1.5 mL) and stirred for 5 min. The mixture was cooled to -40°C and TMSOTf (0.3 μL, 1.5 μmol) was added and the reaction mixture was continued stirring for 10 mins, after which it was slowly warmed to room temperature. Silica gel TLC indicated completion of the reaction, which was quenched with Et_3N, followed by filtering off molecular sieves. The mixture was concentrated *in vacuo* and purified by silica gel column chromatography using 1:2 EtOAc/toluene as eluent giving **2** as white foam.(87 mg, 56%) ^1H NMR (600 MHz, $CDCl_3$): δ 7.04-8.09 (m, 30H, aromatic), 5.55 (dd, *J* = 8.6, 8.6 Hz, 1H, H-3, Gal-2), 5.49 (d, *J* = 7.8 Hz, 1H, H-1, GlcN-1), 5.45 (d, *J* = 8.0, 8.0 Hz, 1H, H-2, Gal-2), 5.43 (m, 1H, H-3, GlcN-2), 5.37 (dd, *J* = 8.8, 8.8 Hz, 1H, H-2, Glc), 5.13 (dd, *J* = 9.2, 9.2 Hz, 1H, H-3, Gal-1), 4.94 (d, *J* = 8.1 Hz, 1H, H-1, Gal-1), 4.69 (d, *J* = 7.3 Hz, 1H, H-1, Gal-2), 4.54 (m, 1H, H-1,

Glc), 4.11 (m, 1H, H-1, GlcN-2), 4.09 (m, 1H, H-4, Gal-2), 3.87 (m, 1H, H-3, Glc), 3.72 (dd, *J* = 9.3, 9.3 Hz, 1H, H-3, Glc), 3.70 (m, 2H, Gal-2 H-5, GlcN-2 H2), 3.57 (dd, *J* = 9.1 Hz, 1H, H-2, GlcN-2), 3.14 (dd, *J* = 7.7, 15.7 Hz, 1H, H2, Gal-1), 1.85-2.20 (m, 35H, CO*CH*$_3$); ^{13}C NMR (75 MHz, $CDCl_3$): δ 133.11, 133.17, 129.81, 129.81, 129.81, 128.51, 128.51, 128.07, 101.09, 101.09, 101.00, 100.38, 98.89, 79.34, 78.95, 77.35, 75.74, 74.45, 74.45, 73.85, 73.75, 72.85, 72.74, 71.84,71.74, 71.74, 71.66, 71.25, 70.75, 70.24, 70.24, 68.74, 68.74, 68.35, 68.14, 67.96, 67.34, 67.34, 67.24, 67.24, 66.85, 62.76, 62.76, 62.04, 61.34, 61.34, 60.85, 58.45, 55.86, 30.00, 20.60, 20.60. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{92}H_{96}Cl_6N_2O_{39}Na$, 2085.3619 found 2089.8164.

D-galactopyranosyl-(1→3)-β-2-acetamido-2-deoxy-D-glucopyranosyl-(1→3)-[2-deoxy-2-acetamido-β-D-glucopyranosyl-(1→6)]-β-D-galactopyransoyl-(1→4)-β-D-glucopyranoside (3)



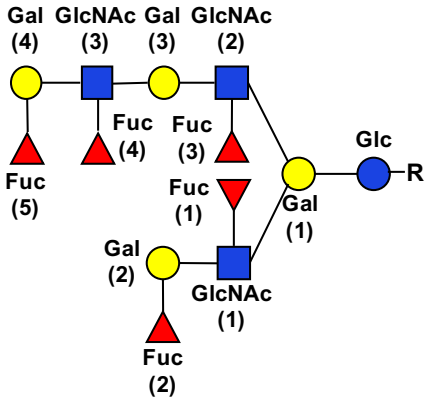
Compound **2** (27.2 mg, 8.2 μmol) was dissolved in THF (2.7 mL), followed by addition of 1M NaOH solution (165 μL, 164.5 μmol). The reaction mixture was stirred overnight at room temperature after which MALDI-TOF MS confirmed completion of the reaction. The reaction was quenched with glacial acetic acid adjusting pH to 4. THF was evaporated *in vacuo* and the residue was co-evaporated with toluene (2 x 2 mL) and dried *in vacuo*. The residue was dissolved in methanol (1 mL) followed by addition of 15 μL acetic anhydride. The reaction mixture was continued stirring overnight at room temperature. MALDI-TOF MS indicated completion of reaction, after which it was concentrated *in vacuo*. The

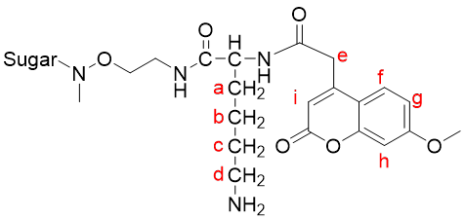
residue was applied to BioGel P-2 column and elute with nano pure water to remove salt. The product containing fractions were collected and lyophilized to give **2**, which was dissolved in a mixed of methanol/H_2O (1 mL, 1:1, v/v), followed by $Pd(OH)_2$ on carbon (1 mg) added to the aforementioned solution of **2**. The reaction mixture was left stirring overnight at room temperature under an atmosphere of hydrogen. After confirming the full product formation by MALDI-TOF MS, the $Pd(OH)_2$ on carbon was filtered off using a celite plug. The filtrate was concentrated *in vacuo* and the residue was filtered by BioGel P-2 column eluted with nano pure water to give the desired pentasaccharide **3** as an α and β mixture (5.6 mg, 75%). ^{1}H NMR (600 MHz, $CDCl_3$): δ 5.06 (d, *J* = 3.7 Hz, 1H, H-1, α-Glc), 4.57 (m, 1H, H1, GlcNAc-1), 4.53 (m, 1H, H-1, GlcNAc-1), 4.52 (m, 1H, H-1, β-Glc), 4.46 (m, 2H, H-1, GlcNAc-2), 4.27 (m, 4H, Gal-1 H-1 and Gal-2 H-1), 3.76 (m, 1H, H-3, Glc), 3.74 (m, 1H, H-2, GlcNAc-1), 3.66 (m, 4H, Gal-1 H-3 and Gal-2 H-3), 3.65 (m, 1H, H-3, Glc), 3.60 (m, 1H, H-2, GlcNAc-1), 3.59 (m, 1H, H-2, Glc), 3.54 (m, 2H, H-2, GlcNAc-2), 3.43 (m, 5H, Gal-1 H-2 and Glc H-2), 3.41 (m, 2H, H-3, GlcNAc-2), 3.31 (m, 4H, GlcNAc-1 H-3 and GlcNAc-2 H-4), 1.86-1.92 (m, 12H, COOCH3); ^{13}C NMR (75 MHz, $CDCl_3$): δ 102.92, 102.61, 102.41, 100.95, 95.58, 91.62, 81.92, 81.58, 78.76, 75.64, 74.96, 74.48, 73.64, 73.64, 73.47, 72.01, 70.99, 70.51, 69.83, 69.70, 68.51, 68.51, 68.51, 68.38, 60.61, 60.61, 55.48, 55.48, 54.50, 22.12, 22.12. MALDI-TOF-MS (*m/z*): $[M+Na]^+$ calcd for $C_{34}H_{58}N_2O_{26}Na$, 933.3176 found 933.1783.

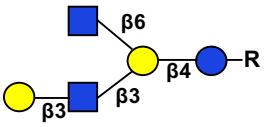
NMR and MS

NMR nomenclature of complex oligosaccharides

Glycan assignments were made by numbering each monosaccharide starting from the reducing terminus and continuing in sequential order. Numbering for target difucosylated LNT is illustrated below in addition to the numbering of the anomeric linker.







4

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	NHAc
Glc	4.03 (d, *J*=9.3Hz, 1H)	3.37	3.49	3.43	-[a]	-	n/a[b]
Gal(1)	4.29	3.45	3.59	4.00	-	-	n/a
GlcNAc(1)	4.60 (d, *J*=8.7Hz, 1H)	3.76	3.35	3.43	-	-	1.91 (s, 3H)
Gal(2)	4.31	3.39	3.51	3.78	-	-	n/a
GlcNAc(2)	4.46 (d, *J*=8.6Hz, 1H)	3.55	3.42	3.32	-	-	1.89 (s, 3H)

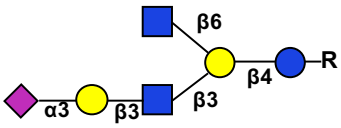
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.78
Gal(1)	103.20
GlcNAc(1)	102.34
Gal(2)	103.34
GlcNAc(2)	101.00

[a] Not assigned

[b] Not applicable

MALDI TOF-MS m/z calcd for $C_{55}H_{86}N_6O_{31}Na(M+Na)^+$ exact 1349.5235 , found 1349.3809



28

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.04 (d, *J*=9.1Hz, 1H)	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.30 (d, *J*=7.9Hz, 1H)	3.46	3.58	3.99	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.61 (d, *J*=8.2Hz, 1H)	3.75	3.34	3.43	3.67	-	n/a	n/a	n/a	1.91 (s, 6H)
Gal(2)	4.38 (d, *J*=7.8Hz, 1H)	3.40	3.95	3.80	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.47 (d, *J*=8.4Hz, 1H)	3.55	3.41	3.31	-	-	n/a	n/a	n/a	1.91 (s, 6H)
Neu5Ac	n/a	n/a	2.62-eq 1.64-axial	3.54	3.69	3.47	-	-	-	1.89 (s, 3H)

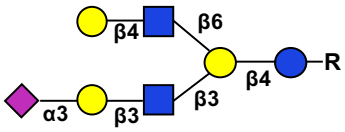
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.90
Gal(1)	102.99
GlcNAc(1)	102.45
Gal(2)	103.31
GlcNAc(2)	100.97

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{66}H_{102}N_7O_{39}Na_2(M+2Na)^{2+}$ exact 831.3004, found 831.0041



29

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.03 (d, *J*=9.0Hz, 1H)	3.38	3.43	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.29	3.46	3.58	3.99	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.60 (d, *J*=8.5Hz, 1H)	3.75	3.34	3.43	3.68	-	n/a	n/a	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.37 (d, *J*=7.6Hz, 1H)	3.40	3.94	3.80	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.59	3.44	3.55	-	-	n/a	n/a	n/a	1.91 (m, 9H)
Gal(3)	4.30	3.40	3.52	3.78	-	-	n/a	n/a	n/a	n/a
Neu5Ac	n/a	n/a	2.62-eq 1.64-axial	3.54	3.69	3.48	-	-	-	1.89 (m, 9H)

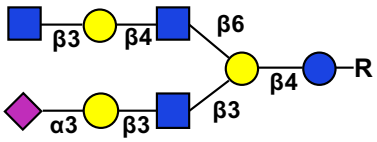
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.83
Gal(1)	102.90
GlcNAc(1)	102.33
Gal(2)	103.25
GlcNAc(2)	100.84
Gal(3)	102.98

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{72}H_{108}N_7O_{44}(M\text{-}2H)^{2-}$ exact 887.3214, found 889.0412



30

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.03 (d, *J*=9.2Hz, 1H)	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.29 (d, *J*=8.0Hz, 1H)	3.45	3.58	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.60 (d, *J*=8.3Hz, 1H)	3.75	3.35	3.44	3.68	-	n/a	n/a	n/a	1.87-1.92 (m, 12H)
Gal(2)	4.37 (d, *J*=7.6Hz, 1H)	3.41	3.95	3.80	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=8.0Hz, 1H)	3.59	3.45	3.54	-	-	n/a	n/a	n/a	1.91 (m, 12H)
Gal(3)	4.29 (d, *J*=8.0Hz, 1H)	3.45	3.58	4.00	-	-	n/a	n/a	n/a	- n/a
GlcNAc(3)	4.54 (d, *J*=8.4Hz, 1H)	3.62	3.43	3.32	-	-	n/a	n/a	n/a	1.89 (m, 12H)
Neu5Ac	n/a	n/a	2.62-eq 1.64-axial	3.55	3.69	3.48	-	-	-	1.89 (m, 12H)

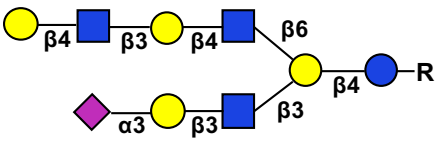
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.73
Gal(1)	102.91
GlcNAc(1)	102.20
Gal(2)	103.20
GlcNAc(2)	101.22
Gal(3)	102.91
GlcNAc(2)	102.65

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{80}H_{127}N_8O_{49}(M+2H)^{2+}$ exact 991.8846, found 991.0999



31

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.03 (d, *J*=9.2Hz, 1H)	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.29 (d, *J*=7.4Hz, 1H)	3.45	3.57	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.60 (d, *J*=8.4Hz, 1H)	3.75	3.34	3.43	3.67	-	n/a	n/a	n/a	1.85-1.93 (m, 12H)
Gal(2)	4.37 (d, *J*=7.5Hz, 1H)	3.41	3.96	3.80	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.6Hz, 1H)	3.59	3.45	3.54	-	-	n/a	n/a	n/a	1.85-1.93 (m, 12H)
Gal(3)	4.29 (d, *J*=7.4Hz, 1H)	3.45	3.57	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.4Hz, 1H)	3.66	3.59	3.44	-	-	n/a	n/a	n/a	1.85-1.93 (m, 12H)
Gal(4)	4.34 (d, *J*=7.6Hz, 1H)	3.41	3.53	3.79	-	-	n/a	n/a	n/a	n/a
Neu5Ac	n/a	n/a	2.62-eq 1.64-axial	3.53	3.69	3.48	-	-	-	1.85-1.93 (m, 12H)

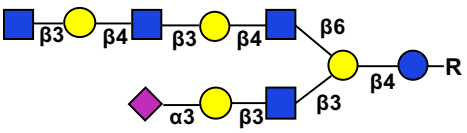
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.83
Gal(1)	102.84
GlcNAc(1)	102.39
Gal(2)	103.16
GlcNAc(2)	100.99
Gal(3)	102.84
GlcNAc(3)	102.59
Gal(4)	102.82

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{86}H_{137}N_8O_{54}(M+2H)^{2+}$ exact 1072.9110, found 1072.1387



32

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.03	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.29 (d, *J*=7.7Hz, 1H)	3.44	3.57	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.60 (d, *J*=8.4Hz, 1H)	3.75	3.35	3.43	3.67	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(2)	4.37 (d, *J*=7.7Hz, 1H)	3.41	3.95	3.80	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.47 (d, *J*=8.1Hz, 1H)	3.59	3.45	3.54	-	-	n/a	n/a	n/a	1.85-1.93 (m, 12H)
Gal(3)	4.29 (d, *J*=7.7Hz, 1H)	3.44	3.57	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.4Hz, 1H)	3.65	3.59	3.44	-	-	n/a	n/a	n/a	1.85-1.93 (m, 12H)
Gal(4)	4.33 (d, *J*=7.9Hz, 1H)	3.41	3.53	3.79	-	-	n/a	n/a	n/a	n/a
GlcNAc(4)	4.54 (d, *J*=8.4Hz, 1H)	3.61	3.43	3.32	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Neu5Ac	n/a	n/a	2.62-eq 1.64-axial	3.53	3.69	3.48	-	-	-	1.85-1.93 (m, 12H)

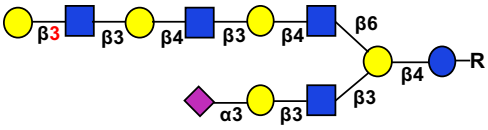
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.86
Gal(1)	102.88
GlcNAc(1)	102.38
Gal(2)	103.27
GlcNAc(2)	100.94
Gal(3)	102.88
GlcNAc(3)	102.71
Gal(4)	102.82
GlcNAc(4)	102.77

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{94}H_{150}N_9O_{59}(M+2H)^{2+}$ exact 1174.4507, found 1173.7062



33

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.04	3.39	3.44	3.51	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.30	3.46	3.59	4.01	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.60 (d, *J*=8.0Hz, 1H)	3.77	3.36	3.45	3.69	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(2)	4.38 (d, *J*=7.3Hz, 1H)	3.42	3.96	3.81	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.9Hz, 1H)	3.60	3.46	3.55	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(3)	4.30	3.44	3.57	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.3Hz, 1H)	3.67	3.60	3.46	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(4)	4.34 (d, *J*=7.6Hz, 1H)	3.47	3.61	4.03	-	-	n/a	n/a	n/a	n/a
GlcNAc(4)	4.56 (d, *J*=8.3Hz, 1H)	3.67	3.60	3.46	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(5)	4.31	3.46	3.59	4.01	-	-	n/a	n/a	n/a	n/a -
Neu5Ac	n/a	n/a	2.62-eq 1.64-axial	3.53	3.69	3.48	-	-	-	1.85-1.93 (m, 15H)

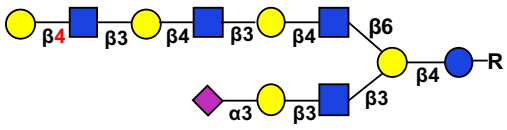
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.90
Gal(1)	103.12
GlcNAc(1)	102.55
Gal(2)	103.36
GlcNAc(2)	100.94
Gal(3)	103.12
GlcNAc(3)	102.70
Gal(4)	103.04
GlcNAc(4)	102.70
Gal(5)	103.12

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{100}H_{154}N_9O_{64}(M-2H)^{2-}$ exact 1252.4536, found 1254.7075



34

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	NHAc
Glc	4.04	3.38	3.43	3.50	-[a]	-	n/a[b]	n/a	n/a	n/a
Gal(1)	4.29 (d, *J*=7.6Hz, 1H)	3.45	3.58	4.01	-	-	n/a	n/a	n/a	n/a
GlcNAc(1)	4.60 (d, *J*=8.3Hz, 1H)	3.75	3.35	3.44	3.68	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(2)	4.37 (d, *J*=7.7Hz, 1H)	3.41	3.95	3.81	-	-	n/a	n/a	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.59	3.45	3.54	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(3)	4.29 (d, *J*=7.6Hz, 1H)	3.44	3.57	4.00	-	-	n/a	n/a	n/a	n/a
GlcNAc(3)	4.56	3.67	3.60	3.45	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(4)	4.33 (d, *J*=7.7Hz, 1H)	3.46	3.59	4.02	-	-	n/a	n/a	n/a	n/a
GlcNAc(4)	4.56	3.67	3.60	3.46	-	-	n/a	n/a	n/a	1.85-1.93 (m, 15H)
Gal(5)	4.34 (d, *J*=8.3Hz, 1H)	3.41	3.53	3.79	-	-	n/a	n/a	n/a	n/a
Neu5Ac	-	-	2.62-eq 1.64-axial	3.53	3.69	3.48	-	-	-	1.85-1.93 (m, 15H)

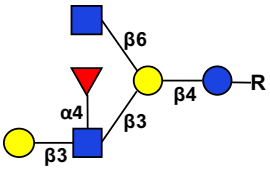
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.90
Gal(1)	102.92
GlcNAc(1)	102.37
Gal(2)	103.36
GlcNAc(2)	100.95
Gal(3)	102.92
GlcNAc(3)	102.64
Gal(4)	102.86
GlcNAc(4)	102.64
Gal(5)	102.86

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{100}H_{154}N_9O_{64}(M\text{-}2H)^{2-}$ exact 1252.4536, found 1254.7469



35

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03	3.39	3.45	3.51	-[a]	-	n/a[b]	n/a
Gal(1)	4.29 (d, *J*=7.6Hz, 1H)	3.46	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.57 (d, *J*=8.3Hz, 1H)	3.82	3.42	3.63	3.96	-	n/a	1.85-1.93 (m, 6H)
Gal(2)	4.37 (d, *J*=7.7Hz, 1H)	3.36	3.49	3.75	-	-	n/a	n/a
GlcNAc(2)	4.47 (d, *J*=7.8Hz, 1H)	3.56	3.43	3.32	-	-	n/a	1.85-1.93 (m, 6H)
Fuc(1)	4.90	3.68	3.76	-	-	-	1.05 (d, J=5.8Hz, 3H)	n/a

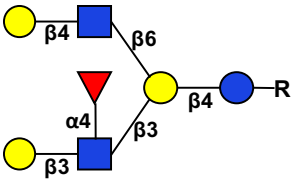
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.88
Gal(1)	102.95
GlcNAc(1)	102.52
Gal(2)	102.79
GlcNAc(2)	101.04
Fuc(1)	97.97

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{61}H_{98}N_6O_{35}(M+2H)^{2+}$ exact 737.3036, found 737.4596



36

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.04	3.39	3.45	3.51	-[a]	-	n/a[b]	n/a
Gal(1)	4.31 (d, *J*=7.6Hz, 1H)	3.41	3.53	3.79	-	-	n/a	n/a
GlcNAc(1)	4.57	3.82	3.42	3.63	3.96	-	n/a	1.85-1.93 (m, 6H)
Gal(2)	4.38 (d, *J*=7.7Hz, 1H)	3.36	3.49	3.75	-	-	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.61	3.56	3.46	-	-	n/a	1.85-1.93 (m, 6H)
Fuc(1)	4.90	3.68	3.76	-	-	-	1.05 (d, J=5.8Hz, 3H)	n/a
Gal(3)	4.31	3.41	3.53	3.79	-	-	n/a	n/a

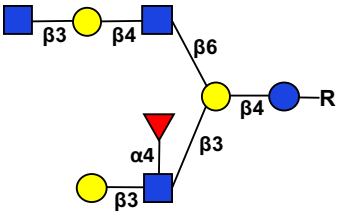
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.82
Gal(1)	102.85
GlcNAc(1)	102.54
Gal(2)	102.80
GlcNAc(2)	100.97
Fuc(1)	98.00
Gal(3)	102.85

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{67}H_{108}N_6O_{40}(M+2H)^{2+}$ exact 818.3300, found 818.4984



37

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.04	3.39	3.45	3.51	-[a]	-	n/a[b]	n/a
Gal(1)	4.30 (d, *J*=7.6Hz, 1H)	3.45	3.59	4.01	-	-	n/a	n/a
GlcNAc(1)	4.57	3.82	3.42	3.64	3.95	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.38 (d, *J*=7.7Hz, 1H)	3.36	3.49	3.75	-	-	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.61	3.56	3.46	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	4.90	3.68	3.76	-	-	-	1.05 (d, J=5.8Hz, 3H)	n/a
Gal(3)	4.30	3.45	3.59	4.01	-	-	n/a	n/a
GlcNAc(3)	4.55	3.62	3.44	3.32	-	-	n/a	1.85-1.93 (m, 9H)

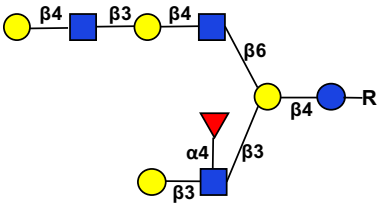
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.82
Gal(1)	102.85
GlcNAc(1)	102.54
Gal(2)	102.80
GlcNAc(2)	100.97
Fuc(1)	98.00
Gal(3)	102.85
GlcNAc(3)	102.78

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{75}H_{121}N_7O_{45}(M+2H)^{2+}$ exact 919.8698, found 920.5586



38

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03	3.39	3.44	3.51	-[a]	-	n/a[b]	n/a
Gal(1)	4.30 (d, *J*=7.6Hz, 1H)	3.46	3.57	4.02	-	-	n/a	n/a
GlcNAc(1)	4.57	3.82	3.42	3.67	3.96	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.37 (d, *J*=7.7Hz, 1H)	3.36	3.49	3.75	-	-	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.60	3.55	3.46	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	4.90	3.68	3.76	n/a	-	-	1.05 (d, J=5.8Hz, 3H)	n/a
Gal(3)	4.30	3.45	3.59	4.01	-	-	n/a	n/a
GlcNAc(3)	4.57	3.82	3.45	3.67	3.96	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.35	3.41	3.55	3.80	-	-	n/a	n/a

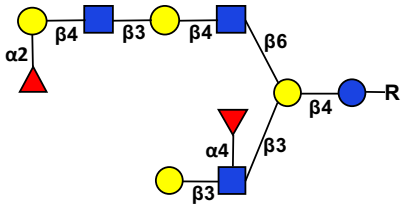
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.82
Gal(1)	102.85
GlcNAc(1)	102.54
Gal(2)	102.80
GlcNAc(2)	100.97
Fuc(1)	98.00
Gal(3)	102.85
GlcNAc(3)	102.78
Gal(4)	102.81

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{81}H_{131}N_7O_{50}(M+2H)^{2+}$ exact 1000.8961, found 1001.5835



39

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.04	3.39	3.44	3.51	-[a]	-	n/a[b]	n/a
Gal(1)	4.30	3.46	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.56	3.82	3.42	3.67	3.96	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.37 (d, *J*=7.7Hz, 1H)	3.36	3.49	3.75	-	-	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.60	3.55	3.46	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	4.90	3.68	3.76	n/a	-	-	1.05 (d, J=5.9Hz, 3H)	n/a
Gal(3)	4.30	3.45	3.59	4.01	-	-	n/a	n/a
GlcNAc(3)	4.56	3.82	3.45	3.67	3.96	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.41	3.41	3.55	3.80	-	-	n/a	n/a
Fuc(2)	5.17	3.68	-	-	-	-	1.09 (d, J=6.1Hz, 3H)	n/a

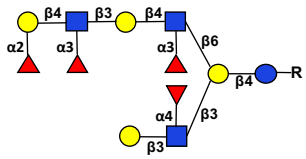
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.82
Gal(1)	102.85
GlcNAc(1)	102.54
Gal(2)	102.80
GlcNAc(2)	100.97
Fuc(1)	98.00
Gal(3)	102.85
GlcNAc(3)	102.78
Gal(4)	102.81
Fuc(2)	99.36

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{87}H_{141}N_7O_{54}(M+2H)^{2+}$ exact 1073.9251, found 1074.6246



40

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.04	3.38	3.43	3.51	-[a]	-	n/a[b]	n/a
Gal(1)	4.30	3.46	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.58	3.83	3.32	3.73	-	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.38 (d, *J*=7.7Hz, 1H)	3.37	3.50	3.73	-	-	n/a	n/a
GlcNAc(2)	4.48 (d, *J*=7.8Hz, 1H)	3.76	3.46	-	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	4.90	3.68	3.77	-	-	-	1.05 (d, J=5.8Hz, 3H)	n/a
Gal(3)	4.30	3.37	3.55	-	-	-	n/a	n/a
GlcNAc(3)	4.57	3.83	3.42	3.64	-	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.37	3.37	3.52	3.74	-	-	n/a	n/a
Fuc(2)	5.14	3.67	-	-	-	-	1.10 (d, J=6.2Hz, 3H)	n/a
Fuc(α3)(1)	4.95	3.55	3.75	3.64	-	-	1.01 (d, J=5.6Hz, 3H)	n/a
Fuc(α3)(2)	4.99	3.56	3.78	3.68	-	-	1.13 (d, J=6.1Hz, 3H)	n/a

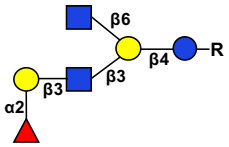
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.87
Gal(1)	103.05
GlcNAc(1)	102.40
Gal(2)	102.82
GlcNAc(2)	100.80
Fuc(1)	97.96
Gal(3)	101.74
GlcNAc(3)	102.40
Gal(4)	102.82
Fuc(2)	99.36
Fuc(α3)(1)	98.67
Fuc(α3)(2)	98.53

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{99}H_{161}N_7O_{62}(M+2H)^{2+}$ exact 1219.9830, found 1220.6918



41

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03 (d, *J*=9.1Hz, 1H)	3.38	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.27 (d, *J*=7.9Hz, 1H)	3.43	3.57	3.99	-	-	n/a	n/a
GlcNAc(1)	4.51 (d, *J*=7.8Hz, 1H)	3.46	3.69	3.75	-	-	n/a	1.85-1.93 (m, 6H)
Gal(2)	4.49 (d, *J*=8.5Hz, 1H)	3.68	3.68	3.37	-	-	n/a	n/a
GlcNAc(2)	4.46 (d, *J*=8.5Hz, 1H)	3.55	3.41	3.31	-	-	n/a	1.85-1.93 (m, 6H)
Fuc(1)	5.06	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a

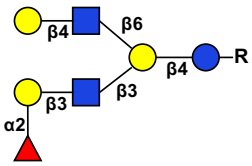
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.79
Gal(1)	102.91
GlcNAc(1)	100.11
Gal(2)	103.04
GlcNAc(2)	100.94
Fuc(1)	99.35

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{61}H_{97}N_6O_{35}Na(M+H+Na)^{2+}$ exact 748.2946, found 748.4629



42

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03 (d, *J*=9.1Hz, 1H)	3.38	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.28 (d, *J*=8.0Hz, 1H)	3.43	3.57	3.99	-	-	n/a	n/a
GlcNAc(1)	4.52	3.46	3.70	3.75	-	-	n/a	1.85-1.93 (m, 6H)
Gal(2)	4.49	3.68	3.55	3.45	-	-	n/a	n/a
GlcNAc(2)	4.48	3.60	3.67	3.36	-	-	n/a	1.85-1.93 (m, 6H)
Fuc(1)	5.06 (d, *J*=3.9Hz, 1H)	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a
Gal(3)	4.31 (d, *J*=8.0Hz, 1H)	3.39	3.52	3.78	-	-	n/a	n/a

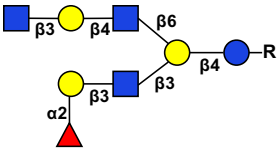
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.82
Gal(1)	102.93
GlcNAc(1)	99.89
Gal(2)	103.13
GlcNAc(2)	100.84
Fuc(1)	99.37
Gal(3)	102.84

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{67}H_{108}N_6O_{40}(M+2H)^{2+}$ exact 818.3300, found 818.5140



43

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29 (d, *J*=8.3Hz, 1H)	3.44	3.58	4.00	-	-	n/a	n/a
GlcNAc(1)	4.50	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.49	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.48	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	5.05	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a
Gal(3)	4.27 (d, *J*=8.1Hz, 1H)	3.43	3.57	3.99	-	-	n/a	n/a
GlcNAc(3)	4.54 (d, *J*=8.4Hz, 1H)	3.61	3.31	3.43	-	-	n/a	1.85-1.93 (m, 9H)

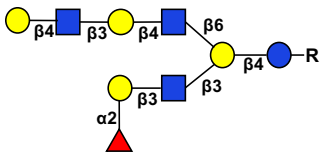
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{75}H_{121}N_7O_{45}(M+2H)^{2+}$ exact 919.8697, found 918.0658



44

^{1}H(600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.02	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29 (d, *J*=8.3Hz, 1H)	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.50	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.49	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.48	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	5.05	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a
Gal(3)	4.27 (d, *J*=8.1Hz, 1H)	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.4Hz, 1H)	3.66	3.44	3.59	-	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.34 (d, *J*=7.7Hz, 1H)	3.40	3.53	3.79	-	-	n/a	n/a

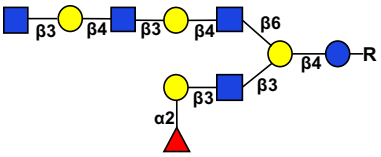
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{81}H_{131}N_7O_{50}(M+2H)^{2+}$ exact 1000.8961, found 1001.6201



45

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.02	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29 (d, *J*=8.3Hz, 1H)	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.50	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 12H)
Gal(2)	4.48	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.47	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 12H)
Fuc(1)	5.05	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a
Gal(3)	4.28	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.56	3.62	3.43	3.33	-	-	n/a	1.85-1.93 (m, 12H)
Gal(4)	4.33 (d, *J*=7.7Hz, 1H)	3.46	3.59	4.02	-	-	n/a	n/a
GlcNAc(4)	4.55	3.62	3.43	3.33	-	-	n/a	1.85-1.93 (m, 12H)

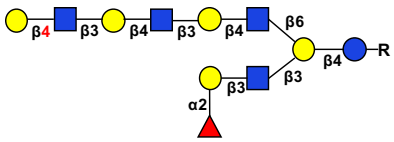
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72
GlcNAc(4)	102.69

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{89}H_{144}N_8O_{55}(M+2H)^{2+}$ exact 1102.4358, found 1102.6741



47

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.02	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29 (d, *J*=8.3Hz, 1H)	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.50	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 12H)
Gal(2)	4.48	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.47	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 12H)
Fuc(1)	5.05	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a
Gal(3)	4.28	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.57	3.66	3.44	3.59	-	-	n/a	1.85-1.93 (m, 12H)
Gal(4)	4.33 (d, *J*=7.7Hz, 1H)	3.41	3.59	3.79	-	-	n/a	n/a
GlcNAc(4)	4.57	3.66	3.46	3.60	-	-	n/a	1.85-1.93 (m, 12H)
Gal(5)	4.34	3.41	3.54	3.79	-	-	n/a	n/a

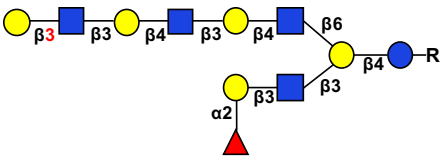
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72
GlcNAc(4)	102.69
Gal(5)	102.72

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{95}H_{154}N_8O_{60}(M+2H)^{2+}$ exact 1183.4622, found 1184.2076



46

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.02	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.50	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 12H)
Gal(2)	4.49	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.47	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 12H)
Fuc(1)	5.06	3.63	3.53	-	-	-	1.10 (d, J=6.6Hz, 3H)	n/a
Gal(3)	4.28	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.56	3.66	3.44	3.59	-	-	n/a	1.85-1.93 (m, 12H)
Gal(4)	4.33	3.41	3.59	3.79	-	-	n/a	n/a
GlcNAc(4)	4.60	3.77	3.35	3.44	3.69	-	n/a	1.85-1.93 (m, 12H)
Gal(5)	4.34	3.47	3.61	4.03	-	-	n/a	n/a

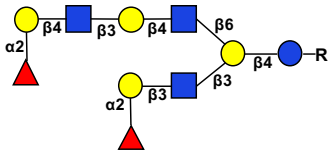
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72
GlcNAc(4)	102.69
Gal(5)	102.83

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{95}H_{154}N_8O_{60}(M+2H)^{2+}$ exact 1183.4622, found 1184.2076



48

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.51	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.49	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.48	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	5.05	3.63	3.53	-	-	-	1.10	n/a
Gal(3)	4.27	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.4Hz, 1H)	3.66	3.44	3.59	-	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.41 (d, *J*=7.8Hz, 1H)	3.54	3.75	-	-	-	n/a	n/a
Fuc(2)	5.17	3.66	-	-	-	-	1.09	n/a

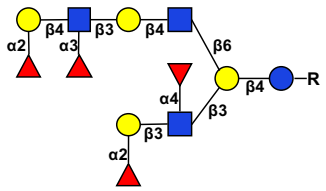
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72
Fuc(2)	99.30

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{87}H_{141}N_7O_{54}(M+2H)^{2+}$ exact 1073.9251, found 1074.6509



49

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.51	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.49	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.46	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 9H)
Fuc(1)	5.01	3.63	3.53	-	-	-	1.08-1.16	n/a
Gal(3)	4.27	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.4Hz, 1H)	3.66	3.44	3.59	-	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.45 (d, *J*=7.8Hz, 1H)	3.54	3.75	-	-	-	n/a	n/a
Fuc(2)	5.13	3.66	-	-	-	-	1.08-1.16	n/a
Fuc(3)(α4)	4.89	3.67	3.78	-	-	-	1.08-1.16	n/a
Fuc(4)(α3)	4.98	3.55	3.77	3.66	-	-	1.08-1.16	n/a

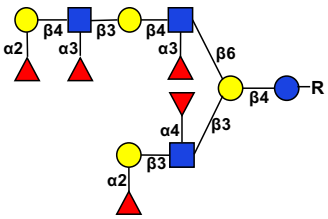
^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Fuc(1)	99.37
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72
Fuc(2)	99.30
Fuc(3)(α4)	97.68
Fuc(4)(α3)	98.47

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{99}H_{161}N_7O_{62}(M+2H)^{2+}$ exact 1219.9830, found 1220.7468



50

^{1}H (600MHz, D_2O): δ (ppm)

	H1	H2	H3	H4	H5	H6	CH_3	NHAc
Glc	4.03	3.37	3.43	3.50	-[a]	-	n/a[b]	n/a
Gal(1)	4.29	3.44	3.57	4.01	-	-	n/a	n/a
GlcNAc(1)	4.51	3.45	3.69	3.74	-	-	n/a	1.85-1.93 (m, 9H)
Gal(2)	4.49	3.68	3.44	3.83	-	-	n/a	n/a
GlcNAc(2)	4.46	3.59	3.53	3.44	-	-	n/a	1.85-1.93 (m, 9H)
Gal(3)	4.27	3.42	3.56	3.98	-	-	n/a	n/a
GlcNAc(3)	4.56 (d, *J*=8.4Hz, 1H)	3.66	3.44	3.59	-	-	n/a	1.85-1.93 (m, 9H)
Gal(4)	4.45 (d, *J*=7.8Hz, 1H)	3.54	3.75	-	-	-	n/a	n/a
Fuc(1)(α4)	4.89	3.67	3.78	-	-	-	1.08-1.16	n/a
Fuc(2)(α2)	5.01	3.61	3.55	-	-	-	1.08-1.16	n/a
Fuc(3)(α3)	4.94	3.54	3.73	3.63	-	-	1.08-1.16	n/a
Fuc(4)(α3)	4.98	3.55	3.77	3.67	-	-	1.08-1.16	n/a
Fuc(5)(α2)	5.13	3.66	-	-	-	-	1.08-1.16	n/a

^{13}C from HSQC (150MHz, D_2O): δ (ppm)

	C1
Glc	92.80
Gal(1)	102.96
GlcNAc(1)	100.16
Gal(2)	103.11
GlcNAc(2)	100.76
Gal(3)	102.86
GlcNAc(3)	102.74
Gal(4)	102.72
Fuc(1)(α4)	97.69
Fuc(2)(α2)	99.52
Fuc(3)(α3)	98.65
Fuc(4)(α3)	98.51
Fuc(5)(α2)	99.29

[a] Not assigned

[b] Not applicable

LC-MS/MS m/z calcd for $C_{105}H_{171}N_7O_{66}(M+2H)^{2+}$ exact 1293.0120, found 1293.7709

References:

(1) PALACIOS, A. L. M.; RUIZ, G. M.; ALTAYE, M.; JIANG, X.; MEINZEN-DERR, J. K.; GUERRERO, M. L.; FARKAS, T.; CHATURVEDI, P.; PICKERING, L. K.; NEWBURG, D. S. Protection Against Diarrhea in Breast-Fed Infants. *J. Pediatr.* **2004**, 297–303.

(2) Donovan, S. M.; Comstock, S. S. Human Milk Oligosaccharides Influence Neonatal Mucosal and Systemic Immunity. *Ann. Nutr. Metab.* **2017**, *69* (2), 42–51. https://doi.org/10.1159/000452818.

(3) Engfer, M. B.; Stahl, B.; Finke, B.; Sawatzki, G.; Daniel, H. Human Milk Oligosaccharides Are Resistant to Enzymatic Hydrolysis in the Upper Gastrointestinal Tract. *Am. J. Clin. Nutr.* **2000**, *71* (6), 1589–1596.

(4) Andersson, B.; Porras, O.; Hanson, L. A.; Lagergfird, T.; Svanborg-eden, C. Inhibition of Attachment of Streptococcus Pneumoniae and Haemophilus Influenzae by Human Milk and Receptor Oligosaccharides Author (s): Bengt Andersson , Oscar Porras , Lars Å . Hanson , Teresa Lagergård and Catharina Published by : Oxford University Pr. *J. Infect. Dis.* **1986**, *153* (2), 232–237.

(5) Angeloni, S.; Ridet, J. L.; Kusy, N.; Gao, H.; Crevoisier, F.; Guinchard, S.; Kochhar, S.; Sigrist, H.; Sprenger, N. Glycoprofiling with Micro-Arrays of Glycoconjugates and Lectins. *Glycobiology* **2005**, *15* (1), 31–41.

(6) Trinchera, M.; Aronica, A.; Dall'Olio, F. Selectin Ligands Sialyl-Lewis a and Sialyl-Lewis x in Gastrointestinal Cancers. *Biology (Basel).* **2017**, *6* (1), 16.

(7) Urashima, T.; Asakuma, S.; Leo, F.; Fukuda, K.; Messer, M.; Oftedal, O. T. The Predominance of Type I Oligosaccharides Is a Feature Specific to Human Breast Milk. *Adv. Nutr.* **2012**, *3* (3), 473S-482S.

(8) Newburg, D. S.; Ruiz-Palacios, G. M.; Morrow, A. L. Human Milk Glycans Protect Infants against Enteric Pathogens. *Annu. Rev. Nutr.* **2005**, *25*, 37–58.

(9) de Vries, T.; Knegtel, R. M. A.; Holmes, E. H.; Macher, B. A. Fucosyltransferases: Structure/Function Studies. *Glycobiology* **2001**, *11* (10), 119R-128R.

(10) Bode, L. Human Milk Oligosaccharides: Every Baby Needs a Sugar Mama. *Glycobiology* **2012**, *22* (9), 1147–1162.

(11) Wu, S.; Tao, N.; German, J. B.; Grimm, R.; Lebrilla, C. B. Development of an Annotated Library of Neutral Human Milk Oligosaccharides. *J. Proteome Res.* **2010**, *9* (8), 4138–4151.

(12) Pras, E.; Raz, J.; Yahalom, V.; Frydman, M.; Garzozi, H. J.; Pras, E.; Hejtmancik, J. F. A Nonsense Mutation in the Glucosaminyl (N-Acetyl) Transferase 2 Gene (GCNT2): Association with Autosomal Recessive Congenital Cataracts. *Invest. Ophthalmol. Vis. Sci.* **2004**, *45* (6), 1940–1945.

(13) Prudden, A. R.; Liu, L.; Capicciotti, C. J.; Wolfert, M. A.; Wang, S.; Gao, Z.; Meng, L.; Moremen, K. W.; Boons, G.-J. Synthesis of Asymmetrical Multiantennary Human Milk Oligosaccharides. *Proc. Natl. Acad. Sci.* **2017**, *114* (27), 6954–6959.