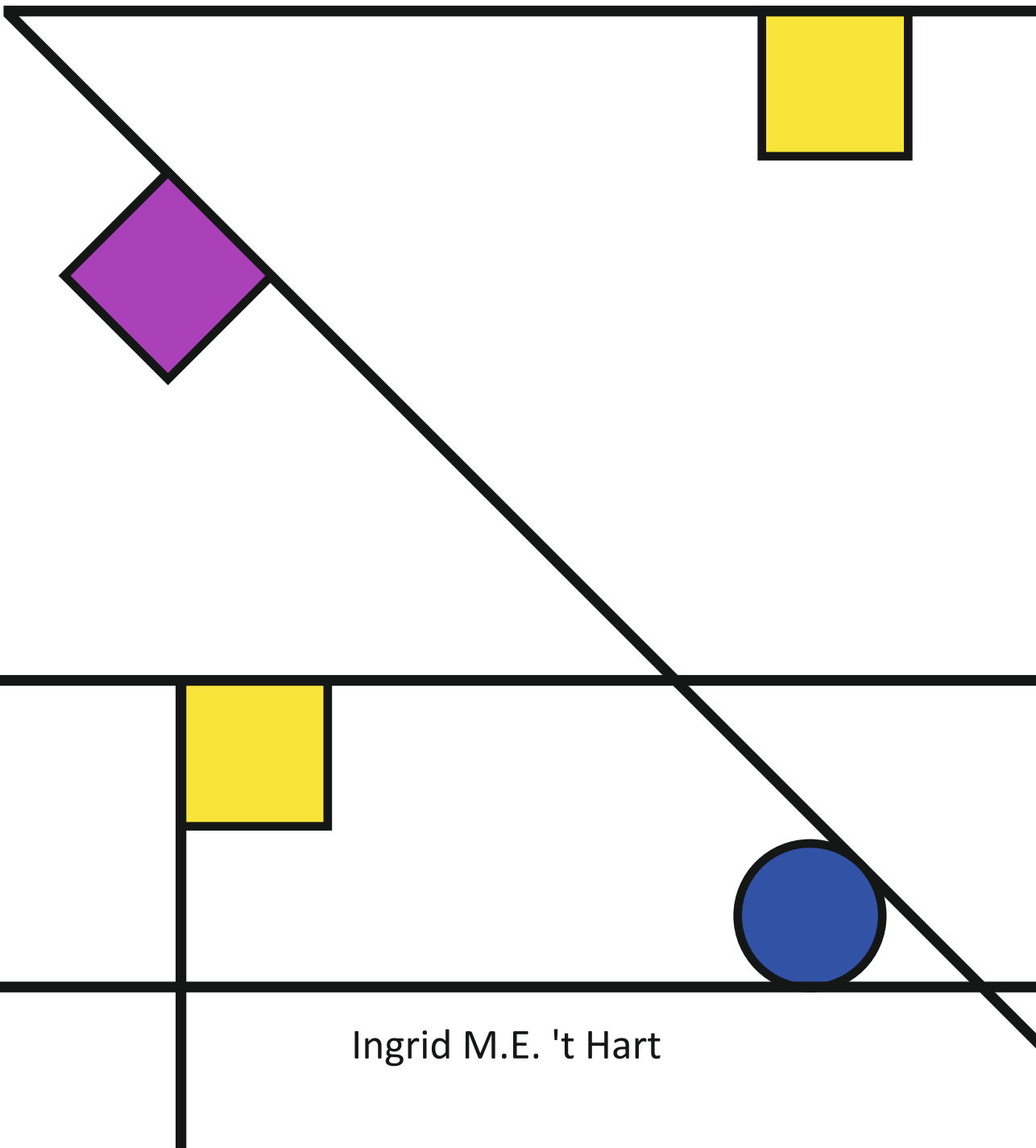
Chemical and chemoenzymatic synthesis of ganglioside(mimic)s and sulfoglycolipids



Ingrid M.E. 't Hart

Chemical and chemoenzymatic synthesis of ganglioside(mimic)s and sulfoglycolipids

Chemische en chemo-enzymatische synthese van ganglioside(achtige)n en sulfoglycolipiden (met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op maandag 19 oktober 2020 des middags

te 4:15 uur

door

Ingrid Maria Engelina 't Hart

geboren op 10 april 1992
te Krimpen aan den IJssel

Promotor
Prof. dr. G.J.P.H. Boons

Printed by: Ridderprint | www.ridderprint.nl

Table of contents

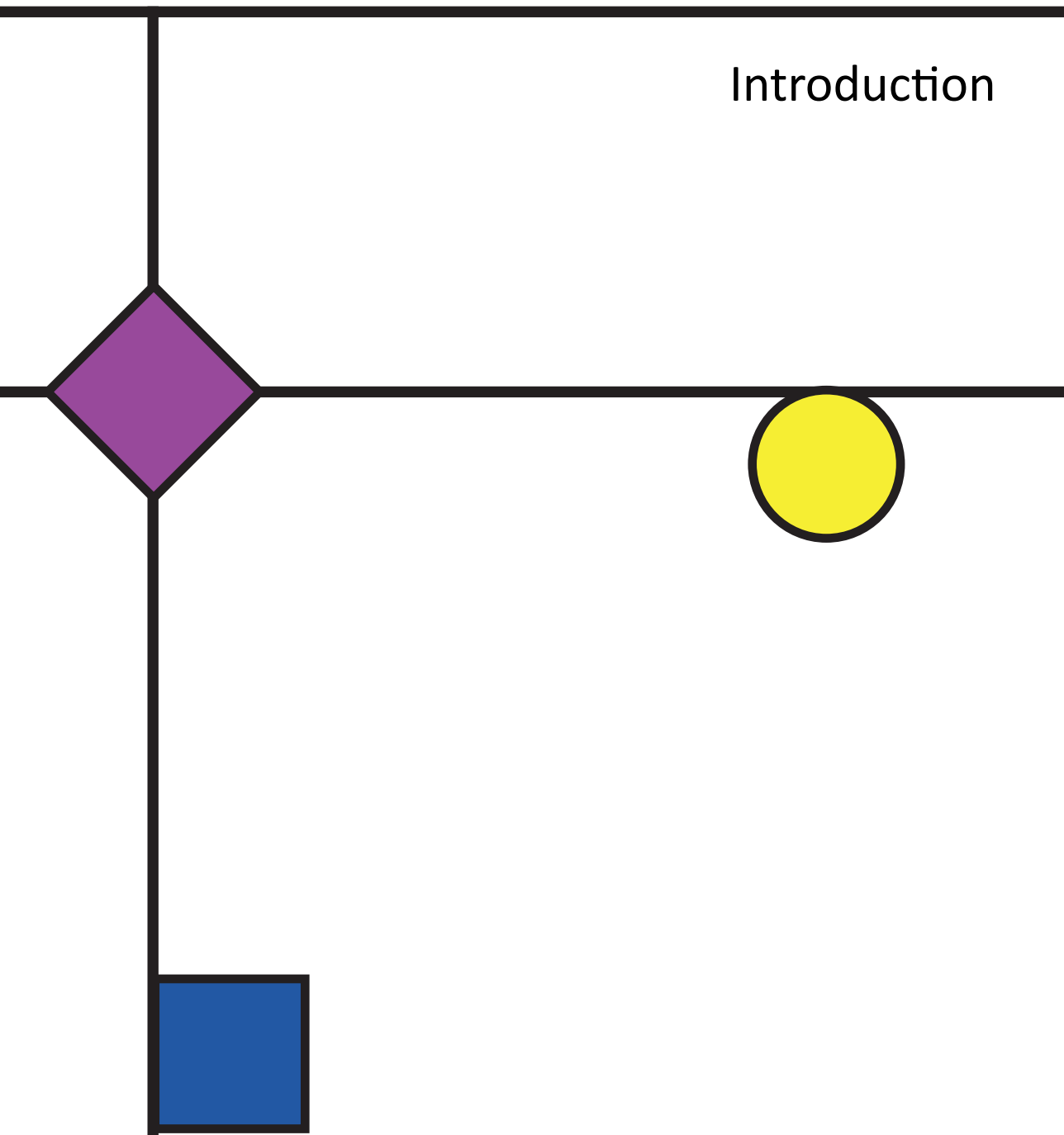
List of abbreviations

Ac	acetyl
AE3 (HAE3)	anti-epiglycanin antibody
AMAN	acute motor axonal degeneration neuropathy
AMSAN	acute motor and sensory axonal neuropathy
Bn	benzyl
Bz	benzoyl
CAN	Cerium ammonium nitrate
Cbz	carboxybenzyl
Cer	ceramide
CgtA	β(1,4)*N*-acetylgalactosaminyltransferase from *Campylobacter jejuni*
CgtB	β(1,3)galactosyltransferase from *Campylobacter jejuni*
CIAP	calf intestinal alkaline phosphatase
CMP	cytidine monophosphate
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DMAP	2,6-di-*tert*-butyl-4-methylpyridine
DMF	dimethylformamide
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EGCase	endoglycoceramidase
ESI	electrospray ionization
Et_3N	triethylamine
Et_3SiH	triethylsilane
Gal	galactose
GalNAc	*N*-acetylgalactosamine
GBS	Guillain-Barré syndrome
Glc	glucose
GlcNAc	*N*-acetyl-D-glucosamine
GSLs	glycosphingoglipids
GT	glycosyltransferase
Hep	L-*glycero*-D-manno-heptose
HILIC	hydrophilic interaction liquid chromatography
HOAc	acetic acid
HOBt	hydroxylbenzotriazole
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Lac	lactose
Lev	levulinoyl
LG	leaving group
LgtD	β(1,3)*N*-acetylgalactosaminyltransferase from *Haemophilus influenzae*
LOS	lipooligosaccharides
mAb	monoclonal antibody
MAL-II	*maackia amurensis* lectin II

MALDI-TOF	matrix-assisted laser desorption/ionization – time of flight
MeOH	methanol
MS	mass spectrometry
4 Å MS	4 Angstrom molecular sieves
Nap	2-naphthyl
Neu5Ac	*N*-acetylneuraminic acid
NHS	*N*-hydroxysuccinimide
NIS	*N*-iodosuccinimide
NMR	nuclear magnetic resonance
NPhth	*N*-phthaloyl
PG	protecting group
Ph	phenyl
pMP	*para*-methoxybenzyl
PmST1	α(2,3)sialyltransferase from *Pasteurella multocida*
SBA	soybean agglutinin
SEC	size-exclusion chromatography
Ser	serine
SGLs	sulfoglycolipids
Siglec	sialic acid-binding Ig-like lectin
SPE	solid-phase extraction
Sph	sphingosine
ST3Gal1	mammalian α(2,3)-sialyltransferase
ST6GaNAc5	mammalian α(2,6)-sialyltransferase
TDS	*tert*-hexyl-*di*-methyl-silyl
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
TfOH	triflic acid
THF	tetrahydrofuran
Thr	threonine
TLC	thin layer chromatography
TMSOTf	trimethylsilyl trifluoromethanesulfonate
Troc	2,2,2-trichloroethoxycarbonyl
UDP	uridine diphosphate
WGA	wheat germ agglutinin

Chapter 1

Introduction



1.1 Glycosphingolipids: gangliosides and sulfoglycolipids

Glycosphingolipids (GSLs) are glycans conjugated to a lipid (ceramide), which is covalently linked to the reducing end of the carbohydrate. The first discovered GSL was galactosylceramide (Gal-Cer) in 1884 from brain tissue and its structure proved so difficult to determine that it was named after the equally enigmatic Egyptian Sphinx. Currently, hundreds of different oligosaccharide moieties have been discovered among GSLs and various ceramide subtypes increase the diversity of GSLs. The most common sphingoid base of ceramide in mammals is D-*erythro*-sphingosine (d18:1) (sphingosine). The fatty acid of the ceramide can vary in chain length, contain an α-hydroxyl group or can be (poly)unsaturated.[1,2,3,4] Furthermore, GSLs can be divided in neutral and negatively charged subclasses. The negatively charged subclass covers GSLs with one or more sialic acids (gangliosides) and sulfated glycolipids (sulfoglycolipids or sulfatides; SGLs). Notably, sialylation or sulfation often occurs at the same position of a glycolipid, as is the case for ganglioside GM1a and sulfoglycolipid SM1a (Fig. 1).

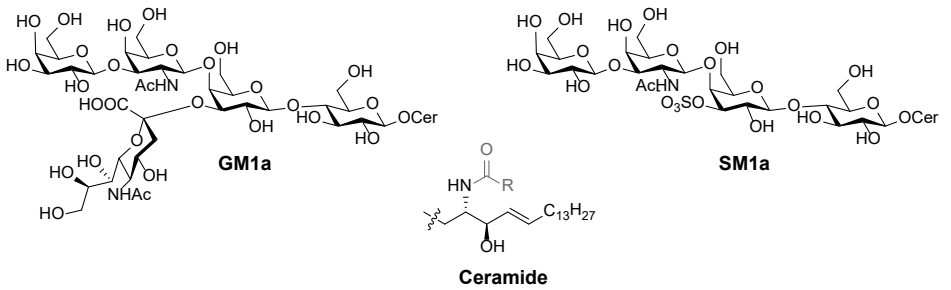


Figure 1. Ganglioside GM1a (left) and sulfoglycolipid SM1a (right). Ceramide (middle) mainly contains (D-*erythro*-)sphingosine (d18:1) in mammals (black). The carbon chain length of the fatty acid (grey) of the ceramide can vary in length (R = 13 to >29), can be unsaturated or bear an α-hydroxyl group.

Glycosphingolipids contain a polar par (the sugar) and an apolar part (the lipid) and are mostly found in the cell membrane with the glycan part facing the extracellular environment (Fig. 2).[1,2,3,5] GSLs form clusters in the membrane, called lipid rafts, together with other lipids such as cholesterol and sphingomyelin and proteins such as GPI-anchored proteins and receptor tyrosine kinases.[3,5,6] Interactions of GSLs can take place with structures residing in the same membrane (*cis* interactions) or with extracellular components (*trans* interactions).[3,4]

All cells in our body synthesize GSLs to a different extent.[3] Still, GSLs are mainly found in nervous tissues where they are important for the integrity and regeneration of nerves.[5] Besides nervous tissue, GSLs are also found on different types of innate and adaptive immune cells such as lactosylceramide (Lac-Cer) on neutrophils and a-series and 0-series gangliosides on CD4+ and CD8+ T cells.[1,6] Different studies have shown that GSLs are important for proper development in multicellular organisms, where deficiencies or alterations can cause severe neurological, immunological and epithelial malfunctions.[3,7]

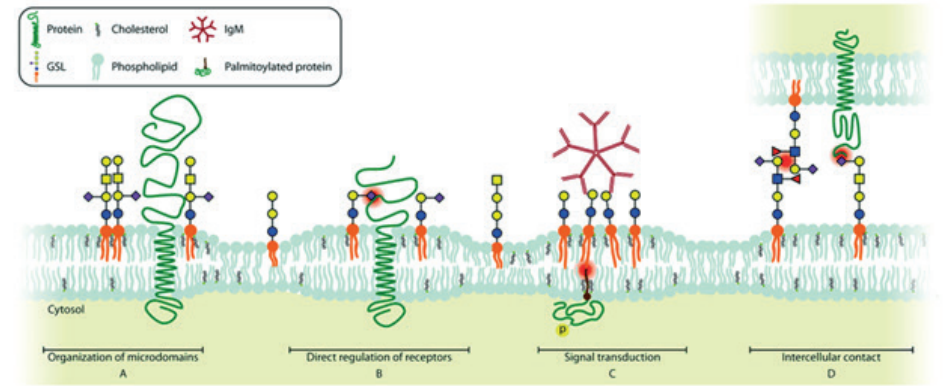


Figure 2. GSLs in the membrane and the types of interactions. (**A**) GSLs cluster in membrane microdomains with other proteins and lipids. (**B**) cis interactions of GSLs with components in the same cell membrane. (**C**) crosslinking of several GSLs can induce transmembrane signal transduction. (D) trans interactions with components on other cell membranes. (Image from Zhang et al., 2019[1])

GSLs are divided into different subclasses based on their glycan structures, namely the: ganglio-, lacto-, neolacto-, globo-, isoglobo-, mollu- and arthroseries. A common component of these GSL subtypes is glucosylceramide (Glc-Cer) (Fig. 3).[3] For GSLs, the minimal glycolipid component is either Gal-Cer or Lac-Cer, which are 3-O-sulfated by galactose-3-O-sulfotransferases, which use PAPS (3'-phosphoadenosine-5'-phosphosulfate) as donor substrate. 3-O-sulfation of Gal-Cer results in sulfatide (SM4), which is highly abundant in nervous tissues, but will not be further discussed in this thesis.[8,9] GSL(-mimic)s of globo- and ganglio-series members represent the main targets of this thesis, therefore the biosynthesis of these subclasses will be discussed in more detail. Furthermore, their prevalence in healthy and diseased environments will be briefly discussed.

The first step in the biosynthesis of GSLs is the attachment of glucose to ceramide by UDP-glucose ceramide glucosyltransferase or glucosylceramide synthase (GlcCer-S) on the cytoplasmic face of the early (or cis-)Golgi membrane. Glc-Cer is next flipped to the non-cytoplasmic face by a lipid flippase to enable further extension of the glycosphingolipid saccharide moiety by luminal glycosyltransferases (GTs).[10] Two glycosyltransferases, B4GalT5 and B4GalT6, can add galactose to Glc-Cer to form lactosylceramide (Lac-Cer).[2,3,11]

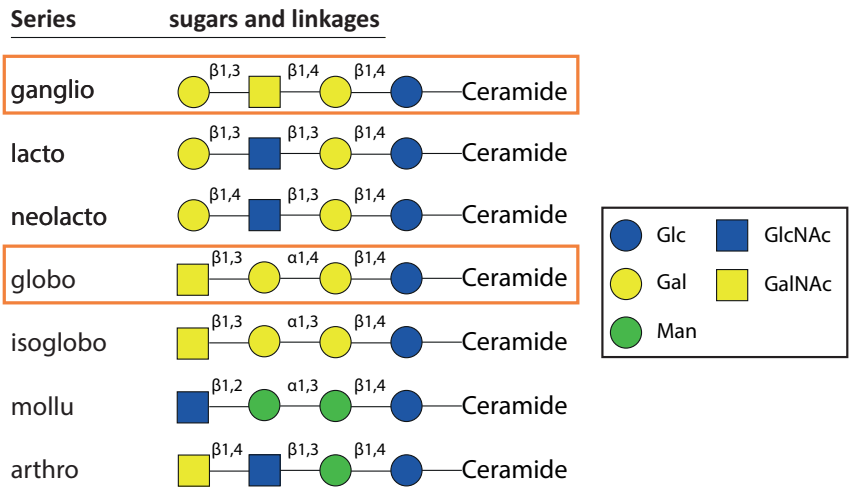


Figure 3. GSL subclasses, based on their glycan composition, adapted from: Schnaar et al., 2017.[3] Highlighted are globo- and ganglio-series GSLs.

1.1.1 Globo-series GSLs

The first extension towards globo-series glycosphingolipids is performed by the α-1,4 galactosyltransferase A4GalT to form Gb3-Cer (also known as ceramide trihexoside, CD77 or P^K antigen) from Lac-Cer (Fig. 4).[12] Gb3-Cer can be further extended by β1,3-acetylgalactosaminyltransferase B3GalNT1 to form Gb4-Cer (P antigen). Gb4-Cer forms the acceptor substrate for β1,3-galactosyltransferase B3GalT5, providing Gb5-Cer (SSEA-3).[13,14] Gb5 can be decorated with fucose by the enzymes FUT1 or FUT2 to form α1,2-fucosylated Globo-H.[15] Apart from that, Gb5-Cer can be extended by one or more sialic acids. Addition of α2,3-Neu5Ac to the reducing galactose residue of Gb5-Cer by ST3GAL2 leads to the formation of MSGb5 (SSEA-4).[16] An additional 2,6-Neu5Ac can be transferred onto the internal GalNAc of MSGb5-Cer by ST6GalNAc6 to form DSGb5-Cer.[17] Although little is known of globo-series SGLs, gal-3-*O*-sulfotransferases were studied for their ability to install 3-*O*-sulfates on globo-series oligosaccharides. From these studies it was found that both Gal3ST2 and Gal3ST4 are able to install a sulfate on the non-reducing trisaccharide of Gb5 (Galβ1,3GalNAcβ1,3GalαOMe).[18,19]

The smallest glycolipid of the globo-series GSLs, Gb3-Cer, is known to be involved in a variety of diseases. Gb3-Cer, and to a lesser extend Gb4-Cer, are known to be a receptor for *Shiga* toxins (also called verotoxin).[20] Furthermore, accumulation of Gb3 is a biomarker for Fabry Disease: a lyosomal storage disorder, caused by a deficiency in α-galactosidase.[21] Next to that, Gb3-Cer on monocytes (mainly $CD4^+$ T-cells) is proposed to bind gp120 on HIV, however whether this interaction induces or inhibits HIV is not fully known.[22] Gb4-Cer is a major glycolipid component of erythrocytes and has been reported to bind epidermal growth factor receptor (EGFR), also for this GSL little is known about the exact function.[5,23] GSLs of the globo-series are known to be the first to appear during embryogenesis; in particular Gb5-Cer (SSEA-3) and MSGb5-Cer (SSEA-4) have been identified as embryonic stem cell markers.[24]

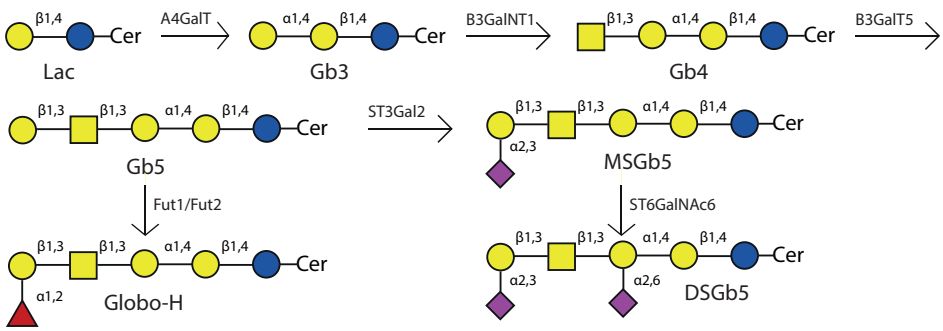


Figure 4. Biosynthesis of globo-series GSLs by mammalian glycosyltransferases.

1.1.2 Ganglio-series GSLs

The ganglio-series GSL synthetic pathway starts with the addition of an α2,3-linked sialic acid to Lac-Cer by ST3Gal5 to form GM3-Cer (Fig. 5).[24] The synthesis of ganglio-series SGLs commences with the 3-O-sulfation of Lac-Cer, providing SM3-Cer. SM3-Cer can be used for further enzymatic modifications in a similar manner as for GM3-Cer.[8] A high variety of ganglio-series GSL subclasses can be synthesized starting from Lac-Cer or GM3-Cer: the 0-, a-, b- and c- series.[10] Here, we will focus on a subset of gangliosides derived from the a- and b-series. In the a-series GM3-Cer can be extended with an β1,4-acetyl galactosamine by B4GalNT1 to form GM2-Cer.[25] The β1,3-galactosyltransferase B3GalT4 extends GM2-Cer to form GM1(a)-Cer.[26] GM1(a)-Cer can be further decorated with sialic acids by the sialyltransferases ST3Gal2 or ST3Gal3 to form GD1a-Cer.[27] GD1a-Cer can be further sialylated by the α2,8-sialyltransferase ST8Sia5 to form GT1a-Cer or by the α2,6-sialyltransferase ST6GalNAc5 (ST6GalNT5) to form GT1aα-Cer.[28] The b-series start with the addition of an α2,8-linked sialic acid by ST8Sia1 to GM3-Cer to form GD3-Cer.[29] GD3-Cer can then be extended by the same enzymes as for the a-series to form GD2-Cer, GD1b-Cer, GT1b-Cer, GQ1b-Cer and GQ1bα-Cer.[10]

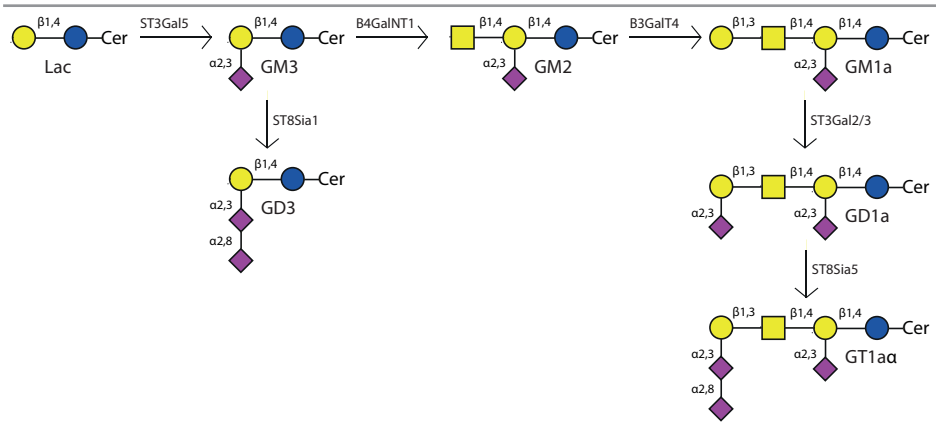


Figure 5. The full biosynthesis of the a-series ganglio-series GSLs and the initial step for the b-series GSLs. GD3 can be extended by the glycosyltransferases described for a-series.

Ganglio-series GSLs are mainly found in the nervous system, where they cover 10-12% of the lipid content. GM3-Cer is found in almost all tissues in vertebrates and is the major ganglioside in serum. Comparison of a standard diet (SD) and a high-fat diet (HFD) in mice resulted in more GM3-Cer and a higher expression of ST3Gal5 in the HFD group. It was found that GM3-Cer downregulates the response of the insulin receptor to insulin and showed higher serum levels in diabetes type 2 patients.[3,31] Accumulation of GM2-Cer, as a result of deficiencies in the degradation enzymes beta-hexosaminidase A and B, can cause severe disorders such as Tay Sachs and Sandhoff disease.[32] GM1-Cer forms a main component in lipid rafts and can therefore easily be detected by anti-GM1 antibodies and cholera toxin B subunit.[33] Recent findings suggest that treatment with GM1-Cer can reduce symptoms Parkinson's disease (PD).[34] Furthermore, GM1-Cer is strongly expressed on CD4+ T-cells and is known to be an important component in lipid rafts involved in the T-cell interaction with antigen presenting cells (APCs).[31] While most gangliosides reside on the outer membranes of the cells, GD3-Cer has also been found on mitochondrial membranes, where it regulates apoptosis.[35] Four ganglio-series GSLs, GM1-Cer, GD1a-Cer, GD1b-Cer and GT1b-Cer, cover up to 97% of gangliosides in the brain.[36] GD1a-Cer and GT1b-Cer on axons form ligands for myelin-associated glycoprotein (MAG, Siglec-4). This interaction is important for myelin stability, axon regeneration and protection from toxic damage.[28] Overall, the major SGLs in animals are SM4, SM3, SM2 and SB1a, which are the sulfated analogs of GM4, GM3, GM2 and GD1a. SGLs are thought to act as ion barriers or traps on the membrane, thereby protecting the cell against high extracellular osmotic concentrations.[8]

Mutations in ST3Gal5 can cause deficiencies of a- and b-series gangliosides, resulting in e.g. slow brain growth, growth failure, blindness and deafness.[31,37,38] Maintenance of ST3Gal5 expression and knock-down of B4GalNT1 (GM2-cer/GD2-Cer synthase) did not cause loss of hearing and led to the suggestion that GM3-Cer is important for organization and function of auditory hair cells.[38] Mice with disrupted B4GalNT1 initially showed normal development, but as they aged male infertility, motor deficits and defects in myelination and axonal degeneration were observed.[2,3,39] Furthermore, B4GalNT1 has been found to positively regulate β-site cleavage by inhibiting BACE1 (β-site APP cleaving enzyme 1) protein degradation. An increase in amyloid precursor protein (APP) cleavage leads to amyloid-β aggregation and the formation of plaques: a main event in Alzheimer's disease.[40] Mice deficient in ST8Sia1, lacking GD3-Cer and b-series gangliosides, have a normal life span and do not show developmental defects. However, disruption of both ST8Sia1 and B4GalNT1 result in GM3-only-mice, which showed extreme sound sensitivities, lethal seizures or even sudden death.[41] Also, a genetic study indicated a correlation between ST8Sia1 variants and the development of multiple sclerosis (MS).[42] Knockdown of B3GalT4 reduces the main brain gangliosides, such as GM1a-Cer, resulting in different neurodegenerative diseases such as Parkinson's.[43] While ST3Gal2 deficient mice appear normal, ST3Gal3 deficient mice are partially impaired and a double knockout of ST3Gal2 and -3 resulted in severely impaired mice at weaning.[28]

1.2 GSLs and SGLs in cancer

Carcinogenesis is the transformation of normal cells into tumor cells and is characterized by changes in biological processes such as increased proliferation, resistance of cell death, evasion of growth suppression and increased migration.[44] Most of these biological processes involve glycans and during cellular transformation, alterations in the glycosylation have been observed. In early stages of cancer a decrease in glycosylation is commonly seen, leading to truncated structures such as sialyl Tn (STn). The formation of neo-epitopes, such as sialyl Lewis a (SLe^a) and sialyl Lewis x (SLe^x), has been correlated with more advanced stages of cancer. Overall, an increase in sialylation is highly associated with carcinogenesis.[45] Furthermore, higher Gal-3-O-sulfotransferase activity was seen in breast and colon cancer cells compared to healthy cells.[17] Both neutral and negatively charged GSLs (gangliosides or SGLs) are up- or downregulated in different types of cancer: a selection of globo- and ganglio-series glycans are listed in Table 1.[46,47]

Table 1. Globo- and ganglio-series GSLs and SGLs and their prevalence in different types of cancer. Adapted from Zhuo et al., 2018.[16,18,46,47,48,49]

GSL	Structure	Cancer type
Gb3	Galα1,4Lac-Cer	lung↑, colorectal↑, gastric↑
Gb4	GalNAcβ1,3Galα1,4Lac-Cer	colorectal↑
Gb5	Galβ1,3GalNAcβ1,3Galα1,4Lac-Cer	breast↑
Globo-H	Fucα1,2Galβ1,3GalNAcβ1,3Galα1,4Lac-Cer	breast↑, prostate↑,
MSGb5	Neu5Acα2,3Galβ1,3GalNAcβ1,3Galα1,4Lac-Cer	breast↑, renal↑, prostate↑,
DSGb5	Neu5Acα2,3Galβ1,3[Neu5Acα2,6]GalNAcβ1,3Galα1,4Lac-Cer	renal↑, prostate↓
GM3	Neu5Acα2,3Lac-Cer	colorectal↓, leukemia↑↓, ovarian↓, renal↑, bladder↓
SM3	3-O-SO_3-Lac-Cer	renal↑, liver↑
GM2	GalNAcβ1,4[Neu5Acα2,3]Lac-Cer	lung↑
SM2	GalNAcβ1,4[3-O-SO_3]Lac-Cer	renal↑
GM1	Galβ1,3GalNAcβ1,4[Neu5Acα2,3]Lac-Cer	colorectal↓
GD3	Neu5Acα2,8Neu5Acα2,3Lac-Cer	breast↑, melanoma↑, ovarian↑
GD2	GalNAcβ1,4[Neu5Acα2,8Neu5Acα2,3]Lac-Cer	breast↑, melanoma↑
GD1a	Neu5Acα2,3Galβ1,3GalNAcβ1,4[Neu5Acα2,3]Lac-Cer	colorectal↓, prostate↑
SB1a	3-O-SO_3-Galβ1,3GalNAcβ1,4[3-O-SO_3]Lac-Cer	renal↑
GD1α	Neu5Acα2,3Galβ1,3[Neu5Acα2,6]GalNAcβ1,4Lac-Cer	breast↑
GD1b	Galβ1,3GalNAcβ1,4[Neu5Acα2,8Neu5Acα2,3]Lac-Cer	breast↓

GSLs form a great pool of tumor markers, however their exact roles in cancer remains to be elucidated. Some pathways and interactions have been proposed in literature, mainly focusing on cell-cell interactions, cell adhesion, growth factor regulation and immune cell function. For convenience, GSL abbreviations in the discussion on GSLs in cancer and autoimmune diseases will be used without description of the lipid moiety: e.g. “Gb3” instead of “Gb3-Cer”.

Cell-cell interactions provide an important way to transduce signals to slow down cell growth if the cell density gets too high. The tumor suppressor Merlin/NF2 is involved in this contact inhibition and reduced expression of Merlin/NF2 has been associated with tumorigenesis. A study on the effect of GSLs on contact inhibition revealed that GSLs GD3 and Gb3 inhibited cell growth, possibly through interactions with Merlin/NF2. This study provides a possible role of GSLs in cell contact inhibition, however its exact role in cancer remains to be elucidated.[50] Another proposed mechanism for GSLs, such as GM3 and Gb4 are carbohydrate-carbohydrate interactions (CCIs) between cells. The role of CCI on cell adhesion in cancer remains to be further investigated.[51]

GSLs also modulate cancer related processes by cis-interactions or transmembrane signal transduction. Whereas Gb4 and Gb5 have been reported to promote activation of EGFR, GT1b, GD1a GM3 and GM1 are proposed to inhibit EGFR phosphorylation and thereby activation.[2,23,52] In addition, GD3 has been found to interact with various membrane molecules, including neogenin-1. It was found that neogenin-1 colocalizes with GD3 in lipid rafts and is involved in enhancing cell growth, invasion and migration.[53] MSGb5 was found to cluster with different signal transducing proteins, leading to downstream activation of extra-cellular-signal-regulated-kinase (ERK), resulting in a decreased expression of integrins. This pathway could be an underlying mechanism for MSGb5 in cancer cell invasion.[54] Another important detected interaction is between platelet-derived growth factor receptor α (PDGFRα) and GSLs GD2 and GD3, resulting in downstream activation of signals that increase invasion, cell growth and migration.[53] Furthermore, GSLs are suggested to be involved in the epithelial-to-mesenchymal transition (EMT).[55] A recent study revealed that loss of globo-series GSLs through CRISPR-Cas9 mediated deletion of the enzyme A4GalT induces EMT.[56]

Different immune cell interactions have been reported for GSLs, mainly gangliosides. For instance, ganglioside GD3 binds with high affinity to CD1d, which is a protein that presents lipid antigens to T cells. GD3 administration inhibited NKT immune cell activation *in vivo*.[57] Other gangliosides are known to modulate immune cells by binding sialic-acid binding Ig-like lectins (Siglecs). Siglecs are expressed on various immune cells and most of these proteins contain an immunoreceptor tyrosine-based inhibitory motif (ITIM). Upon ligand binding, these Siglecs transduce signals that result in the inhibition of an immune response. Therefore, if cancer cells express ligands for Siglecs, they can escape recognition and removal by immune cells. Siglec-7 -9, -10 and -15 all have been found to be involved in immune regulation in cancer.[58,59] As an example, gangliosides GD3, GT1b and DSGb5 are all (proposed) ligands for Siglec-7, an inhibitory receptor on NK cells (Fig. 6).[60,61,62] A co-crystal structure of a GT1b analog with Siglec-7, revealed that the terminal α2,8-sialic acid is the major determinant for ligand binding. A key interaction takes place between the sialic acid carboxylate and Arg-124 in Siglec-7 and removal of either the sialic acid or Arg-124 abolished binding. Furthermore, it was observed that the neutral glycan core interacts with the C-C' loop of Siglec-7. The GT1b analog used for these studies lacked the ceramide residue.[62] However, another (*in vitro*) study revealed that altered lipid compositions of GD3, such as phytosphingosine instead of sphingosine or an α-hydroxyl on the fatty acid, abolished Siglec-7 binding.[60]

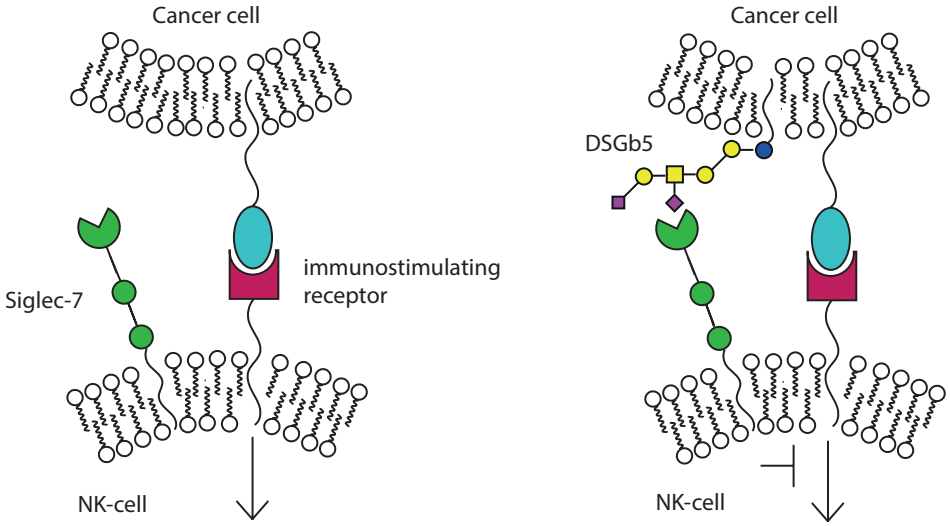


Figure 6. (**A**) Altered structures on cancer cell membranes are recognized by immune cells. (**B**) Co-expression of Siglec-7 ligands such as ganglioside DSGb5 on cancer cells, inhibit an immune response.

Overexpression of GSLs in cancer cells make them interesting targets for antibodies. These antibodies can be elicited for treatment purposes, e.g. to block the interactions of GSLs with immune receptors. The FDA approved antibody Dinutuximab targets GD2 and is an example of passive immunization. Other targets, such as GM2 or Globo-H, are being tested in the late phases of clinical trials for active immunization.[47,63,64] Since carbohydrates often trigger weak immune responses, immunogenicity is improved by conjugation to peptides or proteins such as keyhole limpet haemocyanin (KLH) or diphtheria toxoid cross-reactive material (CRM)197.[65,66,67] Some antibodies have not reached clinical trials, but show promising results in vitro. As an example, an antibody targeting Gb3 demonstrated inhibition of tumor development and angiogenesis.[68] Furthermore, GSLs on cancer cells could be targeted to deliver anticancer drugs more selectively. A recent study describes effective toxin delivery to breast cancer cells by an antibody for MSGb5.[69]

Immunization of A4GalT deficient mice by GSLs from renal cancer cells not only resulted in monoclonal antibodies (mAbs) against gangliosides, but also to SGLs. Anti-ganglioside mAbs bound both healthy and cancerous renal cells, while antibodies towards SGLs, such as SM2, SM3 and SM4, showed high specificity for renal cancer cells over healthy cells.[70] Other SGLs have also shown to be ligands for cancer-specific antibodies. An important cancer-specific antibody, the anti-epiglycanin antibody (AE3, later named HAE3), was raised against epiglycanin from mouse mammary carcinoma cells.[71,72] Sulfoglycolipid SM1a (Fig. 1) strongly bound HAE3 in a carbohydrate microarray study. This was a surprise finding, since HAE3 is known to bind mucin-type glycans and not SGLs.[73]

1.3 GSLs in the autoimmune disease Guillain-Barré Syndrome

Most antibodies are formed upon infection by bacteria, viruses or parasites to detect and remove the pathogens from the human body. Occasionally, antibodies can become cross-reactive towards endogenous antigens by molecular mimicking of the pathogen, which can result in auto-immune diseases.[74]

Guillain-Barré syndrome (GBS) is an autoimmune disease where the peripheral nervous system (PNS) is damaged. GBS is rare disease (0.6-4/100,000 persons/year) and symptoms vary from tingling sensation to complete muscular failure.[3,75,76] This auto-immune disease is in 25-33% of the cases preceded by *Campylobacter jejuni* infection.[75,76,77] However, not all *C. jejuni* infections lead to GBS, suggesting factors of the host are also of importance.[78] GBS can be divided in two subtypes: acute inflammat ory demyelinating polyneuropathy (AIDP) and axonal subtypes such as acute motor axonal degeneration neuropathy (AMAN) and acute motor and sensory axonal neuropathy (AMSAN).[77] AIDP is characterized by segmental demyelination and secondary axonal damage. Upon molecular mimicry, T-cells are activated and damage the myelin sheet together with produced cytokines and radicals. AMAN is the second most common subtype and the main feature is axonal degeneration. In AMAN, molecular mimicry also seems to play a key role. High similarity is found between *C. jejuni* ganglioside mimics and host gangliosides that are mainly present on motor axons, such as: GM1a, GM1b and GD1a.[78] Anti-ganglioside antibodies against GD1b, GT1a, GT1b and GQ1b have been detected in GBS as well.[79]

The gram-negative bacterium *C. jejuni* can carry ganglioside mimics on its lipooligosaccharide (LOS), which forms an important component of the bacterial outer membrane (Fig. 7).[80,81,82,83] Although the exact function of *C. jejuni*'s molecular mimicry is unknown, it plays a key role in GBS.[3,74,84] Different mechanisms have been proposed for the immune response on ganglioside mimics affecting myelin or axons, however much needs to be elucidated. It seems that humoral immunity is the major pathway in the AMAN subtype. Overall, anti-ganglioside antibodies are frequently found in serum of GBS patients: for AIDP up to 62% and for AMAN up to 80%.[78] Additionally, antiganglioside antibodies from GBS serum samples have not only found to be directed towards the single gangliosides, but also to complexes, such as GD1a/GD1b or GQ1b/GM1.[85] These gangliosides and antibodies are useful biomarkers to guide treatment and provide possible new therapeutic targets.

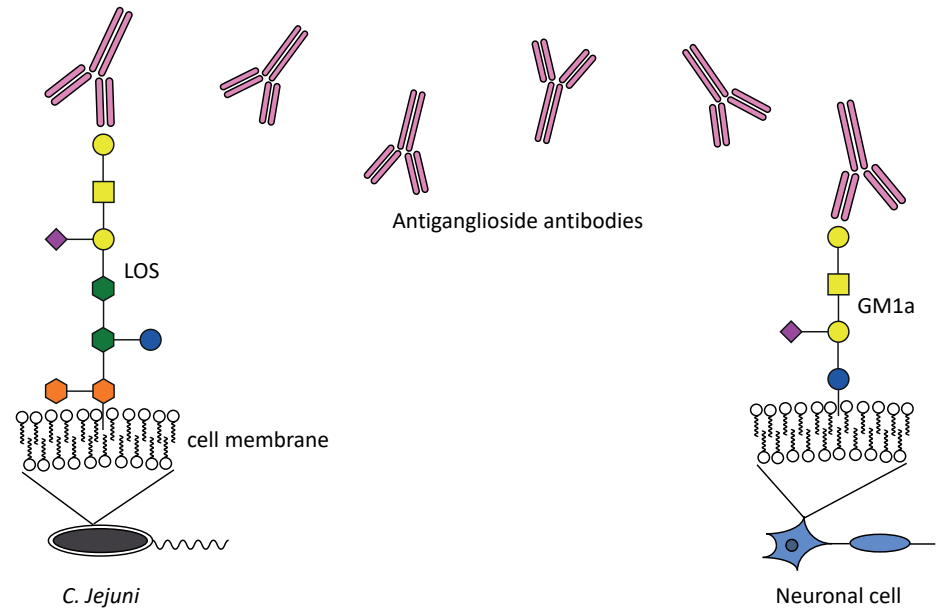


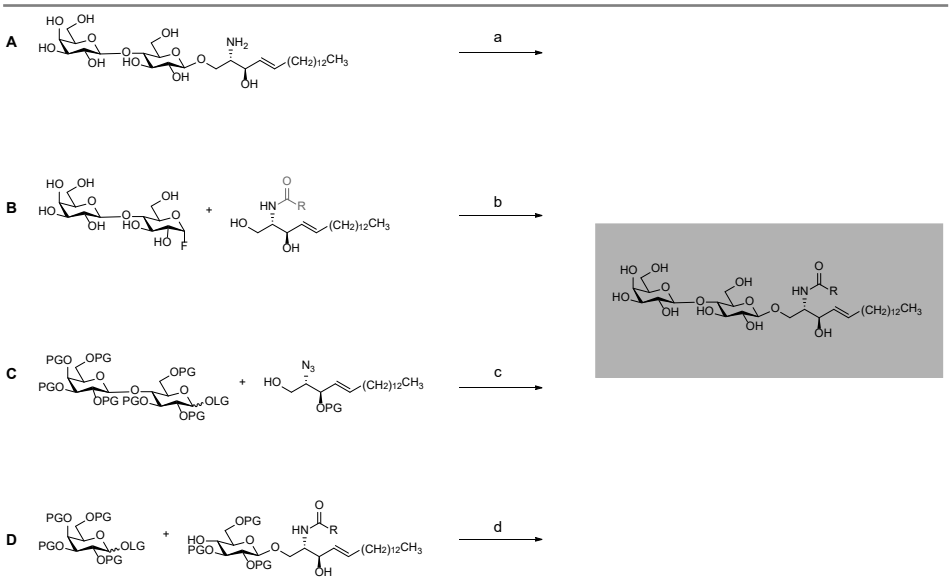
Figure 7. Molecular mimicry of the lipooligosaccharides (LOS) in the outer membrane of *C. jejuni* triggers an immune response where antiganglioside antibodies are formed. These antiganglioside antibodies can become cross-reactive towards endogenous gangliosides on myelin and axon neuronal cells.

1.4 Chemical and chemoenzymatic synthesis of GSL(mimic)s.

Studying GSL interactions more closely requires homogeneous pure glycolipids. However, isolation of pure GSLs is a tedious job and often results in heterogeneous mixtures.[86,87] Chemical or enzymatic synthesis would provide a way to obtain sufficient pure GSL glycolipids. Still, glycolipid synthesis offers quite some challenges over glycan synthesis. The lipid residue adds additional functional groups, such as an alkene, which can lead to side reactions during glycosylation or (de)protections. Furthermore, ceramide (and sphingosine to a lesser extend) is a glycosyl acceptor with poor reactivity.[88,89] Not unimportantly, introducing a lipid moiety will affect the solubility of the GSL substrate, especially in the deprotected form. Nonetheless, some strategies have been developed to overcome some of the aforementioned problems and will be discussed in this chapter and are summarized in Scheme 1.

Recently, a chemoenzymatic strategy was developed where lactosylsphingosine (Lac-Sph) is synthesized chemically, followed by enzymatic extensions (Scheme 1A).[90] With this strategy, a set of ganglio-series GSLs has been synthesized by bacterial glycosyltransferases. The sphingosine analogues were purified on C18 SPE cartridges followed by coupling of a fatty acid to the sphingosine to form the ceramide. It was found that commonly used α2,3-sialyltransferase PmST1 was not suitable for glycolipid substrates. Nonetheless, another bacterial enzyme (PmST3) was found to perform this reaction. Another enzyme, CgtA, showed reduced activity on glycolipids compared to the free glycans.[90] A second chemoenzymatic approach requires the chemical or

chemoenzymatic synthesis of α-fluoride glycan donors (Scheme 1B). These donors can be coupled with ceramide or sphingosine acceptors, catalyzed by a mutant endoglycoceramidase II (EGCase): EGC-E. Enzymatic coupling of the lipid overcomes the problem of lower reactivity of sphingosine or ceramide acceptors. However, the fluoride donors are prone to hydrolysis, particularly under acidic conditions. Also, the enzyme is still selective towards certain donor substrates.[91,92]



Scheme 1. Chemical and chemoenzymatic strategies to synthesize GSL-ceramides, where the synthesis of Lac-Cer is used as an example for each strategy. (**A**) This chemoenzymatic strategy starts with the chemical synthesis of Lac-Sph, followed by extension of the glycan moiety by bacterial glycosyltransferases (GTs). Coupling of an acyl chloride with the free amine provides the respective GSL-ceramides. (a) i. extension of glycan chain by bacterial GTs; ii. coupling with acyl chloride.87 (**B)** Another chemoenzymatic strategy starts with chemically synthesized α-fluoride glycosyl donors. EGC-E, a mutant EGC-II enzyme, glycosylates the α-F glycans with sphingosine or ceramide. (b) EGC-E, pH 5.0.88,89 (**C**) Full chemical synthesis of GSLs by chemical glycosylation of a protected saccharide donor and sphingosine acceptor. (c) i. chemical glycosylation; ii. coupling with fatty acid; iii. global deprotection.90,91 (**D**) Chemical synthesis of GSLs through the Glc-Cer cassette approach. (d) i. chemical glycosylation; ii. global deprotection.92 R = alkyl chain of fatty acid. Abbreviations: PG = protecting group, LG = leaving group.

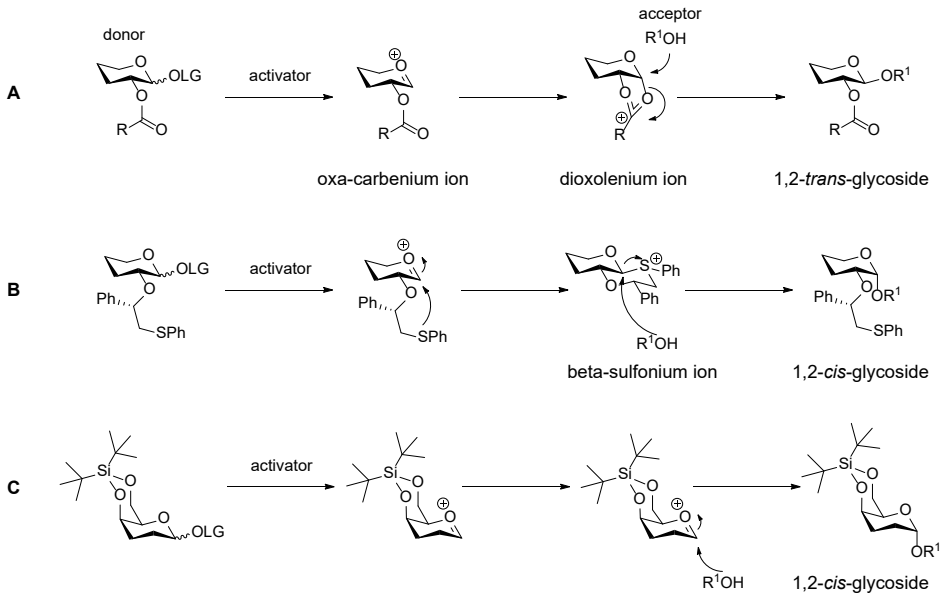
Few chemical synthetic strategies have been developed up till now. In the most straightforward approach, the lipid moiety is coupled to the glycan reducing end by chemical glycosylation (Scheme 1C). Since ceramide acceptors have poor reactivity, a protected form of sphingosine is preferred as the acceptor substrate and coupling with a fatty acid is performed in a later stage.[93,94] In a more recent "cassette" approach, a glucosyldonor is chemically glycosylated with ceramide (Glc-Cer) (Scheme 1D). Reducing the size of the glycan donor increases reactivity. The Glc-Cer acceptor is then glycosylated with a sugar donor to form the protected glycosylceramide. Still, the chemical synthesis of all building blocks requires a series of (de)protection steps and carefully chosen of orthogonal protecting groups.[95]

1.5 Chemical synthesis of GSL oligosaccharides

Whereas oligonucleotides (DNA and RNA) and peptides can be obtained by general (automated) chemical synthetic methods, no such general method is available for oligosaccharide synthesis. Therefore, individual synthetic strategies are required to obtain the desired oligosaccharide(s).[87,86]

Controlling stereo- and regioselectivity are key challenges in carbohydrate synthesis and various solutions have been developed over time. Glycosidic linkages are formed through chemical glycosylation, where a glycosyl donor and acceptor are coupled in the presence of an activator. Neighboring group participation, solvent systems and temperature can influence the stereoselective outcome of a chemical glycosylation.[87,96,97]

The most efficient and robust way to obtain 1,2-*trans* linkages is through neighboring group participation. The C-2 hydroxyl on the glycosyl donor is protected with an ester functionality and the *trans* glycoside is formed through the dioxolenium ion intermediate (Scheme 2A). This method provides β-linked products for sugars such as glucose or galactose and α-linked products for sugar such as mannose.[87]



Scheme 2. (**A**) Neighbouring group participation by the C-2 ester provides trans products. (**B**) A chiral auxiliary provides 1,2-*cis* products through a trans-decalin intermediate. (**C**) The formation of 1,2-*cis* linkages can be obtained by a bulky di-*tert*-butyl-silane on galactose C-4 and C-6 positions. Adapted from Boltje et al., 2009.[87]

Forming a 1,2-*cis* linkage has proved more challenging, although some effective methods have been developed. Incorporation of a chiral auxiliary on C-2 can provide a *trans*-decalin intermediate and results 1,2-*cis* glycosidic bonds (Scheme 2B). Furthermore, a bulky di-tert-butylsilyl group on the C-4 and C-6 hydroxyls of galactose also provides 1,2-cis linkages, most likely by steric hindrance of the β-face (Scheme 2C).[87] The latter *cis*-directing strategy has been used to synthesize Gal(α1,4)Lac in the globo-series oligosaccharides.[98,99]

Regioselective outcomes are usually obtained by use of orthogonal protecting groups, keeping one hydroxyl free as the nucleophile. Another option to control regioselectivity is by exploiting hydroxyl reactivity.[100,101] As an example, in the synthesis of ganglio-series oligosaccharides, the lactose acceptor can have both 3- and 4-hydroxyls as possible sites for glycosylation. Due to the low reactivity of the axial C-4 hydroxyl group, a (near) selective glycosylation with C-3 can be obtained.[102,103]

Furthermore, reactivity of both donor and acceptor should be taken into account by choosing protecting groups. It has been observed that benzyl compounds are more reactive ("armed") than the benzoylated or acetylated derivatives ("disarmed").[104] This does not promote the sole use of benzyl ether protecting groups, but provides handles to increase reactivity. Matching of donors, acceptors and suitable activators is of key importance in all chemical glycosylations.[105]

Glycosylations with sialic acid donors are a common challenge in the chemical synthesis of gangliosides. The anomeric center is crowded and contains an electron-withdrawing carboxylic acid. Furthermore, they lack a C-3 hydroxyl for neighboring group participation. Different optimizations, such as the oxazolidinone protecting group and phosphate leaving groups, have led to more reactive and α-selective donors.[106,107,108] However, the availability of sialyltransferases, such as pmST1, provide a good alternative method for sialylation.[109]

Protected synthetic glycans are usually purified by C18, silica column or size-exclusion chromatography (SEC) (e.g. LH-20, SX1), while deprotected synthetic glycans are mainly purified by SEC (e.g. Bio-Gel P-2) or hydrophilic interaction chromatography (HILIC). Next to various mass spectrometric methods (ESI, MADLI-TOF, ion mobility), 1D and 2D NMR analysis is highly relevant for the proper analysis of complex carbohydrates.[110,111]

1.6 Enzymatic synthesis of GSL oligosaccharides

All oligosaccharides in nature are assembled and cleaved by enzymes and over 500,000 glycosyltransferase (GT) sequences have been identified to far.[112] These natural catalysts provide an efficient regio- and stereoselective way to synthesize oligosaccharides in the lab. Therefore, expression and isolation of stable and active glycosyltransferases is of high importance. Next to GTs, (mutant) glycosyl hydrolases can be used for glycan synthesis as well under specific conditions.[87,92] The advantage of enzymatic oligosaccharide synthesis is the high regio- and stereoselectivity, which drastically reduces the amount of reaction steps compared to chemical synthesis. However, enzymes are usually expressed in small quantities and expensive sugar-nucleotides are required, which limits the scale of the reaction.

Glycosyltransferases can be derived from microbial or mammalian sources, each having their own advantages and disadvantages.[113] Expression of mammalian GTs requires expression in mammalian or insect cells, since prokaryotic systems often lack glycosylation, disulfide bonding and chaperones required for these proteins. Using eukaryotic cell systems usually limits the scale of mammalian GT expression.[114] Mammalian enzymes are often found to be selective to specific acceptor substrates, while bacterial enzymes are usually easier to express in higher yields and have a broader substrate scope.[115] It is hard to find bacterial GTs that synthesize the same glycan linkages as mammalian GTs based on their amino acid sequences.[113] Still, bacterial transferases have been identified that form various glycosidic linkages, including those found in globo- and ganglio-series oligosaccharides.

The bacterial enzymes LgtC and LgtD were identified to synthesize globo-series Gb3, Gb4 and Gb5 starting from lactose. α1,4-Galactosyltransferse LgtC efficiently couples donor substrate UDP-Gal with a lactose acceptor and selectively provides Galα1,4Lac (Gb3).[116] This enzyme accepts small substituents at the reducing end of lactose, such as benzyl (Bn).[117] LgtD is a β1,3-N-acetylgalactosaminyltransferase and has been used synthesize Gb4 from Gb3. Furthermore, a bifunctional activity was found for LgtD, since it also catalyzes the formation of Gb5 from Gb4.[118,109] Another β1,3-galactosyltransferase, CgtB from *C. jejuni*, has been used to synthesize Gb5 as well, but can cause polygalactosylation in case of excess UDP-Gal.[109]

Mimics of the ganglio-series have been found in the bacterium *C. jejuni* and the corresponding GTs have been identified and expressed.[80,109,119,120] Two α2,3-sialyltransferases have been isolated from from *C. jejuni* and have been used for the synthesis of GM3 (Neu5Acα2,3Lac): Cst-I and Cst-II. Cst-II is a bifunctional enzyme since it also has α2,8-sialyltransferase activity.[119] A *Pasteurella multocida* mutant M144D α2,3-sialyltransferase (PmST1) has frequently been used for α2,3-sialylation, including the synthesis of GM3.[109] PmST1 showed poor activity on Lac-Sph, therefore another α2,3-sialyltransferase, PmST3, was used to synthesize GM3-Sph.[121] CgtA is a β1,4-N-acetylgalactosyltransferase from *C. jejuni* and has been identified and used to synthesize GM2 (GalNAcβ1,4[α2,3Neu5Ac]-Lac) from GM3.[119,120] Aforementioned β1,3-galactosyltransferase CgtB was identified and used to catalyze the formation of GM1 (Galβ1,3GalNAcβ1,4[α2,3Neu5Ac]Lac) from GM2.[119,120]

Besides a suitable sugar nucleotide donor, glycan acceptor and glycosyltransferase, an optimal buffer pH and reaction temperature are required for efficient oligosaccharide synthesis. Furthermore, certain GTs require metal ions (Mg^{2+} or Mn^{2+}), which coordinate to the enzyme DXD motif and stabilize phosphate groups of the sugar nucleotide donor.[109,122] Enzymatically synthesized oligosaccharides are mainly in their natural ("deprotected") form and enzyme precipitation (by ice-cold EtOH), spin filtration, Bio-Gel P2 and HILIC are commonly used purification methods.[109] Meanwhile, enzymatic synthesis of glycolipids requires other purification methods due to the introduction of a lipophilic residue. Frequently used methods for glycolipid purifications are C18-based SPE or HPLC.[121]

1.7 Scope and outline

This thesis describes the chemical and chemoenzymatic syntheses of biologically relevant GSLs and SGLs and analogs thereof.

In **Chapter 2**, the chemoenzymatic synthesis of globo-series glycans Gb5, MSGb5 and DSGb5 is described. Chemical synthesis provided pentasaccharide Gb5 and sialic acids were attached by mammalian sialyltransferases ST3Gal1 and ST6GalNAc5. Gb5, MSGb5 and DSGb5 were equipped with an aminopentyl linker at their reducing ends, through which they were linked to NHS-activated microarray slides. Binding studies with plant lectins SBA and WGA indicated correct printing of all three oligosaccharides. Strong binding was observed between α2,8α2,3-disialylated ganglioside GD3 and Siglec-7, however no binding was seen between Siglec-7 and DSGb5.

Chemical and chemoenzymatic syntheses of globo-series glycosphingolipids Gb3-Cer, Gb4-Cer and Gb5-Cer are described in **Chapter 3**. The first part of this chapter focusses on the chemical synthesis of protected Gb5-Sph. Overall, chemical glycosylation of saccharide donors with a sphingosine acceptor resulted in low yields. Whereas benzoyl ester protecting groups drastically reduced donor reactivity, benzyl ether protecting groups proved difficult to remove by Birch reduction, which is required to maintain the unsaturated sphingosine residue. In the second part of this chapter a successful chemoenzymatic strategy for glucosylceramides is described. Chemically synthesized Gb3-Sph is extended by the bifunctional bacterial enzyme LgtD to obtain Gb4-Sph and Gb5-Sph. Peptide coupling with palmitic acid provides Gb3-Cer, Gb4-Cer and Gb5-Cer.

The chemical synthetic target of **Chapter 4** is the sulfated glycan "SM1-core3". SM1-core3 is a hybrid of sulfoglycolipid SM1a and the core3 (GlcNAcβ1,3GalNAc) disaccharide found in O-glycans. This sulfated glycolipid is proposed to be a ligand for cancer-specific antibody HAE3 and thereby a possible epithelial tumor marker. An optimized chemical synthetic strategy provided pure protected SM1-core3 glycan. SM1-core3 will be globally deprotected and immobilized through its reducing aminopentyl linker. The results of a glycan microarray binding study will provide us insight in the SM1-core3 – HAE3 interaction. This will promote the search for SM1-core3 and similar structures as epithelial cancer biomarkers.

In **Chapter 5**, the chemoenzymatic synthesis of *C. jejuni* core oligosaccharide fragments is described. Chemical glycosylation, followed by a two-step deprotection provided the Galβ1,3Hepα(CH2)5NH2 disaccharide. Ganglioside mimics GM3-Hep, GM2-Hep and GM1-Hep were then synthesized by bacterial enzymes pmST1, CgtA and CgtB. The low stability of CgtA and quick loss of activity could be overcome by using cell lysate and storage at -20 °C. All synthesized Hep-ganglioside mimics will be immobilized on NHS-activated slides alongside their human ganglioside analogs. A glycan microarray binding study with serum antibodies of GBS patients will elucidate the effect of heptose on antibody binding.

References

1. Zhang, T.; de Waard, A. A.; Wuhrer, M.; Spaapen, R. M., *Front Immunol* **2019**, *10*, 90.
2. D'Angelo, G.; Capasso, S.; Sticco, L.; Russo, D., *FEBS J* **2013**, *280* (24), 6338-53.
3. Schnaar, R. L.; Kinoshita, T., Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), **2015-2017**.
4. Sonnino, S.; Chiricozzi, E.; Grassi, S.; Mauri, L.; Prioni, S.; Prinetti, A., *Prog Mol Biol Transl Sci* **2018**, *156*, 83-120.
5. Furukawa, K.; Ohmi, Y.; Kondo, Y.; Bhuiyan, R. H.; Tajima, O.; Zhang, P.; Ohkawa, Y.; Furukawa, K., *Proc Jpn Acad Ser B Phys Biol Sci* **2019**, *95* (3), 136-149.
6. Nakayama, H.; Nagafuku, M.; Suzuki, A.; Iwabuchi, K.; Inokuchi, J. I., *FEBS Lett* **2018**, *592* (23), 3921-3942.
7. Allende, M. L.; Proia, R. L., *Glycoconj J* **2014**, *31* (9), 613-22.
8. Ishizuka, I., *Prog. Lipid. Res.* **1997**, *36* (4), 245-319.
9. Honke, K., *Proc Jpn Acad Ser B Phys Biol Sci* **2013**, *89* (4), 129-38.
10. Chalat, M.; Menon, I.; Turan, Z.; Menon, A. K., *J Biol Chem* **2012**, *287* (19), 15523-32.
11. Sandhoff, R.; Schulze, H.; Sandhoff, K., *Prog Mol Biol Transl Sci* **2018**, *156*, 1-62.
12. Keusch, J. J.; Manzella, S. M.; Nyame, K. A.; Cummings, R. D.; Baenziger, J. U., *J Biol Chem* **2000**, *275* (33), 25315-21.
13. Okajima, T.; Nakamura, Y.; Uchikawa, M.; Haslam, D. B.; Numata, S. I.; Furukawa, K.; Urano, T.; Furukawa, K., *J Biol Chem* **2000**, *275* (51), 40498-503.
14. Zhou, D.; Henion, T. R.; Jungalwala, F. B.; Berger, E. G.; Hennet, T., *J Biol Chem* **2000**, *275* (30), 22631-4.
15. Chang, W.-W.; Lee, C. H.; Lee, P.; Lin, J.; Hsu, C.-W.; Hung, J.-T.; Lin, J.-J.; Yu, J. C.; L., S.; Yu, J.; Wong, C.-H.; Yu, A. L., *Proc Natl Acad Sci U S A.* **2008**, *105* (33), 11667-11672.
16. Saito, S.; Aoki, H.; Ito, A.; Ueno, S.; Wada, T.; Mitsuzuka, K.; Satoh, M.; Arai, Y.; Miyagi, T., *J Biol Chem* **2003**, *278* (29), 26474-9.
17. Senda, M.; Ito, A.; Tsuchida, A.; Hagiwara, T.; Kaneda, T.; Nakamura, Y.; Kasama, K.; Kiso, M.; Yoshikawa, K.; Katagiri, Y.; Ono, Y.; Ogiso, M.; Urano, T.; Furukawa, K.; Oshima, S.; Furukawa, K., *Biochem J* **2007**, *402* (3), 459-70.
18. Chandrasekaran, E. V.; Xue, J.; Neelamegham, S.; Matta, K. L., *Carbohydr Res* **2006**, *341* (8), 983-94.
19. Chandrasekaran, E. V.; Lakhaman, S. S.; Chawda, R.; Piskorz, C. F.; Neelamegham, S.; Matta, K. L., *J Biol Chem* **2004**, *279* (11), 10032-41.
20. Legros, N.; Dusny, S.; Humpf, H. U.; Pohlentz, G.; Karch, H.; Muthing, J., *Glycobiology* **2017**, *27* (1), 99-109.
21. Heywood, W. E.; Doykov, I.; Spiewak, J.; Hallqvist, J.; Mills, K.; Nowak, A., *Biochim Biophys Acta Mol Basis Dis* **2019**, *1865* (10), 2726-2735.
22. Lingwood, C. A.; Branch, D. R., **2011**, *59*, 303-313.
23. Park, S. Y.; Kwak, C. Y.; Shayman, J. A.; Kim, J. H., *Biochim Biophys Acta* **2012**, *1820* (7), 1141-8.
24. Breimer, M. E.; Saljo, K.; Barone, A.; Teneberg, S., *Glycoconj J* **2017**, *34* (6), 713-723.
25. Ishii, A.; Ohta, M.; Watanabe, Y.; Matsuda, K.; Ishiyama, K.; Sakoe, K.; Nakamura, M.; Inokuchi, J.; Sanai, Y.; Saito, M., *J Biol Chem.* **1998**, *273* (48), 31652-31655.
26. Nagata, Y.; Yamashiro, S.; Yodoi, J.; Lloyd, K. O.; Shiku, H.; Furukawa, K., *J. Biol. Chem.* **1992**, *267* (17), 12082-12089.
27. Amado, M.; Almeida, R.; Carneiro, F.; Levery, S. B.; Holmes, E. H.; Nomoto, M.; Hollingsworth, M. A.; Hassan, H.; Schwientek, T.; Nielsen, P. A.; Bennett, E. P.; Clausen, H., *J. Biol. Chem.* **1998**, *273* (21), 12770-12778.
28. Sturgill, E. R.; Aoki, K.; Lopez, P. H.; Colacurcio, D.; Vajn, K.; Lorenzini, I.; Majic, S.; Yang, W. H.; Heffer, M.; Tiemeyer, M.; Marth, J. D.; Schnaar, R. L., *Glycobiology* **2012**, *22* (10), 1289-301.
29. Kono, M.; Yoshida, Y.; Kojima, N.; Tsuji, S., *J. Biol. Chem.* **1996**, *271* (46), 29366-29371.
30. Haraguchi, M.; Yamashiro, S.; Yamamoto, A.; Furukawa, K.; Takamiya, K.; Lloyd, K. O.; Shiku, H.; Furukawa, K., *Proc Natl Acad Sci U S A.* **1994**, *91* (22), 10455-10459.
31. Inokuchi, J. I.; Inamori, K. I.; Kabayama, K.; Nagafuku, M.; Uemura, S.; Go, S.; Suzuki, A.; Ohno, I.; Kanoh, H.; Shishido, F., *Prog Mol Biol Transl Sci* **2018**, *156*, 151-195.
32. Masingue, M.; Dufour, L.; Lenglet, T.; Saleille, L.; Goizet, C.; Ayrignac, X.; Ory-Magne, F.; Barth, M.; Lamari, F.; Mandia, D.; Caillaud, C.; Nadjar, Y., *Ann Neurol* **2020**, *87* (4), 609-617.
33. Ledeen, R. W.; Wu, G., *Trends Biochem Sci* **2015**, *40* (7), 407-18.
34. Schneider, J. S.; Cambi, F.; Gollomp, S. M.; Kuwabara, H.; Brasic, J. R.; Leiby, B.; Sendek, S.; Wong, D.

F., *J Neurol Sci* **2015**, *356* (1-2), 118-23.
35. Garofalo, T.; Tinari, A.; Matarrese, P.; Giammarioli, A. M.; Manganelli, V.; Ciarlo, L.; Misasi, R.; Sorice, M.; Malorni, W., *FEBS Lett* **2007**, *581* (21), 3899-903.
36. Tettamanti, G.; Bonali, F.; Marchesini, S.; Zambotti, V., *Biochim. Biophys. Acta.* **1973**, *296* (1), 160-170.
37. Bowser, L. E.; Young, M.; Wenger, O. K.; Ammous, Z.; Brigatti, K. W.; Carson, V. J.; Moser, T.; Deline, J.; Aoki, K.; Morlet, T.; Scott, E. M.; Puffenberger, E. G.; Robinson, D. L.; Hendrickson, C.; Salvin, J.; Gottlieb, S.; Heaps, A. D.; Tiemeyer, M.; Strauss, K. A., *Mol Genet Metab* **2019**, *126* (4), 475-488.
38. Yoshikawa, M.; Go, S.; Suzuki, S.; Suzuki, A.; Katori, Y.; Morlet, T.; Gottlieb, S. M.; Fujiwara, M.; Iwasaki, K.; Strauss, K. A.; Inokuchi, J., *Hum Mol Genet* **2015**, *24* (10), 2796-807.
39. Takamiya, K.; Yamamoto, A.; Furukawa, K.; Yamashiro, S.; Shin, M.; Okada, M.; Fukumoto, S.; Haraguchi, M.; Takeda, N.; Fujimura, K.; Sakae, M.; Kishikawa, M.; Shiku, H.; Furukawa, K.; Aizawa, S., *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10662-10667.
40. Yamaguchi, T.; Yamauchi, Y.; Furukawa, K.; Ohmi, Y.; Ohkawa, Y.; Zhang, Q.; Okajima, T.; Furukawa, K., *Sci Rep* **2016**, *6*, 34505.
41. Kolter, T., *ISRN Biochem* **2012**, *2012*, 506160.
42. Husain, S.; Yildirim-Toruner, C.; Rubio, J. P.; Field, J.; Southern, M. S. G. C.; Schwalb, M.; Cook, S.; Devoto, M.; Vitale, E., *PLoS One* **2008**, *3* (7), e2653.
43. Verma, M.; Schneider, J. S., *Mol Cell Neurosci* **2019**, *95*, 25-30.
44. Hanahan, D.; Weinberg, Robert A., *Cell* **2011**, *144* (5), 646-674.
45. Pinho, S. S.; Reis, C. A., *Nat Rev Cancer* **2015**, *15* (9), 540-55.
46. Krengel, U.; Bousquet, P. A., *Front Immunol* **2014**, *5*, 325.
47. Zhuo, D.; Li, X.; Guan, F., *Front Physiol* **2018**, *9*, 466.
48. Slovin, S. F.; Ragupathi, G.; Adluri, S.; Ungers, G.; Terry, K.; Kim, S.; Spassova, M.; Bornmann, W. G.; Fazzari, M.; Dantis, L.; Olkiewicz, K.; Lloyd, K. O.; Livingston, P. O.; Danishefsky, S. J.; Scher, H. I., *Proc Natl Acad Sci U S A.* **1999**, *96* (10), 5710-5715.
49. Chuang, P. K.; Hsiao, M.; Hsu, T. L.; Chang, C. F.; Wu, C. Y.; Chen, B. R.; Huang, H. W.; Liao, K. S.; Chen, C. C.; Chen, C. L.; Yang, S. M.; Kuo, C. W.; Chen, P.; Chiu, P. T.; Chen, I. J.; Lai, J. S.; Yu, C. T.; Wong, C. H., *Proc Natl Acad Sci U S A* **2019**, *116* (9), 3518-3523.
50. Huang, X.; Schurman, N.; Handa, K.; Hakomori, S., *FEBS Lett* **2017**, *591* (13), 1918-1928.
51. Handa, K.; Hakomori, S. I., *Glycoconj J* **2012**, *29* (8-9), 627-37.
52. Bremer, E. G.; Schlessinger, J.; Hakomori, S., *J. Biol. Chem.* **1986**, *261* (5), 2434-2440.
53. Furukawa, K.; Ohmi, Y.; Ohkawa, Y.; Bhuiyan, R. H.; Zhang, P.; Tajima, O.; Hashimoto, N.; Hamamura, K.; Furukawa, K., *Cancer Sci* **2019**, *110* (5), 1544-1551.
54. Van Slambrouck, S.; Steelant, W. F., *Biochem J* **2007**, *401* (3), 689-99.
55. Guan, F.; Handa, K.; Hakomori, S.-I., *Proc. Natl. Acad. Sci. U S A.* **2009**, *106* (18), 7461-7466.
56. Jacob, F.; Alam, S.; Konantz, M.; Liang, C. Y.; Kohler, R. S.; Everest-Dass, A. V.; Huang, Y. L.; Rimmer, N.; Fedier, A.; Schotzau, A.; Lopez, M. N.; Packer, N. H.; Lengerke, C.; Heinzelmann-Schwarz, V., *Cancer Res* **2018**, *78* (11), 2952-2965.
57. Webb, T. J.; Li, X.; Giuntoli, R. L.; Lopez, P. H. H.; Heuser, C.; Schnaar, R. L.; Tsuji, M.; Kurts, C.; Oelke, M.; Schneck, J. P., *Cancer Research* **2012**, *72* (15), 3744-3752.
58. Scott, E.; Elliott, D. J.; Munkley, J., *Clin Chim Acta* **2020**, *502*, 167-173.
59. Ledeen, R. W.; Kopitz, J.; Abad-Rodriguez, J.; Gabius, H. J., *Prog Mol Biol Transl Sci* **2018**, *156*, 289-324.
60. Hashimoto, N.; Ito, S.; Tsuchida, A.; Bhuiyan, R. H.; Okajima, T.; Yamamoto, A.; Furukawa, K.; Ohmi, Y.; Furukawa, K., *J Biol Chem* **2019**, *294* (28), 10833-10845.
61. Kawasaki, Y.; Ito, A.; Withers, D. A.; Taima, T.; Kakoi, N.; Saito, S.; Arai, Y., *Glycobiology* **2010**, *20* (11), 1373-1379.
62. Attrill, H.; Imamura, A.; Sharma, R. S.; Kiso, M.; Crocker, P. R.; van Aalten, D. M., *J Biol Chem* **2006**, *281* (43), 32774-83.
63. Yu, A. L.; Hung, J. T.; Ho, M. Y.; Yu, J., *Stem Cells Dev* **2016**, *25* (20), 1532-1548.
64. Keyel, M. E.; Reynolds, C. P., *Biologics* **2019**, *13*, 1-12.
65. Jin, K. T.; Lan, H. R.; Chen, X. Y.; Wang, S. B.; Ying, X. J.; Lin, Y.; Mou, X. Z., *Biotechnol Lett* **2019**, *41* (6-7), 641-650.
66. Lee, H. Y.; Chen, C. Y.; Tsai, T. I.; Li, S. T.; Lin, K. H.; Cheng, Y. Y.; Ren, C. T.; Cheng, T. J.; Wu, C. Y.; Wong, C. H., *J Am Chem Soc* **2014**, *136* (48), 16844-53.
67. O'Cearbhaill, R. E.; Deng, W.; Chen, L. M.; Lucci, J. A., 3rd; Behbakht, K.; Spirtos, N. M.; Muller, C.

Y.; Benigno, B. B.; Powell, M. A.; Berry, E.; Tewari, K. S.; Hanjani, P.; Lankes, H. A.; Aghajanian, C.; Sabbatini, P. J., *Gynecol Oncol* **2019**, *155* (3), 393-399.
68. Desselle, A.; Chaumette, T.; Gaugler, M. H.; Cochonneau, D.; Fleurence, J.; Dubois, N.; Hulin, P.; Aubry, J.; Birkle, S.; Paris, F., *PLoS One* **2012**, *7* (11), e45423.
69. Ruggiero, F. M.; Rodriguez-Walker, M.; Daniotti, J. L., *Immunol Cell Biol* **2020**, *98* (3), 187-202.
70. Ando, R.; Tokuda, N.; Yamamoto, T.; Ikeda, K.; Hashimoto, N.; Taguchi, R.; Fan, X.; Furukawa, K.; Niimura, Y.; Suzuki, A.; Goto, M.; Furukawa, K., *Glycoconj J* **2016**, *33* (2), 169-80.
71. Thingstad, T.; Vos, H. L.; Hilkens, J., *Biochem. J.* **2001**, *353*, 33-40.
72. Codington, J. F.; Haavik, S.; Nikrui, N.; Kuter, I.; Vassileva, C.; Zhang, H.; Matson, S.; Chen, X.; Wu, Z., *Cancer* **2002**, *94* (3), 803-13.
73. Palma, A. S.; Liu, Y.; Childs, R. A.; Herbert, C.; Wang, D.; Chai, W.; Feizi, T., *Biochem Biophys Res Commun* **2011**, *408* (4), 548-52.
74. Rivera-Correa, J.; Rodriguez, A., *Trends Immunol* **2018**, *39* (7), 515-522.
75. Israeli, E.; Agmon-Levin, N.; Blank, M.; Chapman, J.; Shoenfeld, Y., *Clin Rev Allergy Immunol* **2012**, *42* (2), 121-30.
76. Hughes, R. A. C.; Cornblath, D. R., *The Lancet* **2005**, *366* (9497), 1653-1666.
77. Wang, Y.; Sun, S.; Zhu, J.; Cui, L.; Zhang, H.-L., *Mediators of Inflammation* **2015**, *2015*, 1-12.
78. Ebrahim Soltani, Z.; Rahmani, F.; Rezaei, N., *Eur Cytokine Netw* **2019**, *30* (1), 1-14.
79. Wanleenuwat, P.; Iwanowski, P.; Kozubski, W., *J Neurol Sci* **2020**, *408*, 116576.
80. Parkhill, J.; Wren, B. W.; Mungall, K.; Ketley, J. M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R. M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A. V.; Moule, S.; Pallen, M. J.; Penn, C. W.; Quail, M. A.; Rajandream, M. A.; Rutherford, K. M.; van Vliet, A. H.; Whitehead, S.; Barrell, B. G., *Nature* **2000**, *403* (6770), 665-668.
81. Karlyshev, A. V.; Ketley, J. M.; Wren, B. W., *FEMS Microbiol Rev* **2005**, *29* (2), 377-90.
82. Moran, A. P.; Prendergast, M. M.; Appelmelk, B. J., *FEMS Immunol Med Microbiol.* **1996**, *16* (2), 105-115.
83. Kanipes, M. I.; Tan, X.; Akelaitis, A.; Li, J.; Rockabrand, D.; Guerry, P.; Monteiro, M. A., *J Bacteriol* **2008**, *190* (5), 1568-74.
84. van den Berg, B.; Walgaard, C.; Drenthen, J.; Fokke, C.; Jacobs, B. C.; van Doorn, P. A., *Nat Rev Neurol* **2014**, *10* (8), 469-82.
85. Kaida, K.; Kusunoki, S., *Future Lipidology* **2008**, *3* (4), 425-434.
86. Committee, N. R. C. U., *National Academies Press (US), Washington, D.C.* **2012**.
87. Boltje, T. J.; Buskas, T.; Boons, G. J., *Nat Chem* **2009**, *1* (8), 611-22.
88. Morales-Serna, J. A.; Diaz, Y.; Matheu, M. I.; Castillon, S., *Org Biomol Chem* **2008**, *6* (20), 3831-6.
89. Vankar, Y. D.; Schmidt, R. R., *Chemical Society Reviews* **2000**, *29* (3), 201-216.
90. Santra, A.; Li, Y.; Yu, H.; Slack, T. J.; Wang, P. G.; Chen, X., *Chem Commun (Camb)* **2017**, *53* (59), 8280-8283.
91. Vaughan, M. D.; Johnson, K.; DeFrees, S.; Tang, X.; Warren, R. A.; Withers, S. G., *J. Am. Chem. Soc.* **2006**, *128* (19), 6300-6301.
92. Rich, J. R.; Withers, S. G., *Angew Chem Int Ed Engl* **2012**, *51* (34), 8640-3.
93. Lassaletta, J. M.; Carlsson, K.; Garegg, P. J.; Schmidt, R. R., *J Org Chem* **1996**, *61* (20), 6873-6880.
94. Gege, C.; Schumacher, G.; Rothe, U.; Schmidt, R. R.; Bendas, G., *Carbohydr Res* **2008**, *343* (14), 2361-8.
95. Nakashima, S.; Ando, H.; Saito, R.; Tamai, H.; Ishida, H.; Kiso, M., *Chem Asian J* **2012**, *7* (5), 1041-51.
96. de Kleijne, F. F. J.; Moons, S. J.; White, P. B.; Boltje, T. J., *Org Biomol Chem* **2020**, *18* (6), 1165-1184.
97. Mong, K.-K. T.; Nokami, T.; Tran, N. T. T.; Nhi, P. B., *Selective Glycosylations: Synthetic Methods and Catalysts: Synthetic Methods and Catalysts* **2017**, *Chapter 3*.
98. Gold, H.; Boot, R. G.; Aerts, J. M. F. G.; Overkleeft, H. S.; Codée, J. D. C.; van der Marel, G. A., *European Journal of Organic Chemistry* **2011**, *2011* (9), 1652-1663.
99. t Hart, I. M. E.; Li, T.; Wolfert, M. A.; Wang, S.; Moremen, K. W.; Boons, G. J., *Org Biomol Chem* **2019**, *17* (31), 7304-7308.
100. Wang, C. C.; Lee, J. C.; Luo, S. Y.; Kulkarni, S. S.; Huang, Y. W.; Lee, C. C.; Chang, K. L.; Hung, S. C., *Nature* **2007**, *446* (7138), 896-9.
101. Ding, Y.; Vara Prasad, C. V. N. S.; Wang, B., *European Journal of Organic Chemistry* **2020**, *2020* (12), 1784-1801.
102. Duclos Jr., R. I., *Carb. Res.* **2000**, *328*, 489-507.
103. Komori, T.; Imamura, A.; Ando, H.; Ishida, H.; Kiso, M., *Carbohydr Res* **2009**, *344* (12), 1453-63.

104. Fraser-Reid, B.; López, J. C., *Top Curr Chem.* **2011**, *301*, 1-29.
105. Premathilake, H. D.; Demchenko, A. V., *Top Curr Chem* **2011**, *301*, 189-221.
106. De Meo, C.; Priyadarshani, U., *Carbohydr Res* **2008**, *343* (10-11), 1540-52.
107. Crich, D.; Li, W., *J. Org. Chem.* **2007**, *72* (7), 2387-2391.
108. Hsu, C. H.; Chu, K. C.; Lin, Y. S.; Han, J. L.; Peng, Y. S.; Ren, C. T.; Wu, C. Y.; Wong, C. H., *Chemistry* **2010**, *16* (6), 1754-60.
109. Yu, H.; Li, Y.; Zeng, J.; Thon, V.; Nguyen, D. M.; Ly, T.; Kuang, H. Y.; Ngo, A.; Chen, X., *J Org Chem* **2016**, *81* (22), 10809-10824.
110. Mutenda, K. E.; Matthiesen, R., *Mass Spectrometry Data Analysis in Proteomics* **2007**, *Chapter 17*, 289-301.
111. Duus, J. O.; Gotfredsen, C. H.; Bock, K., *Chem. Rev* **2000**, *100*, 4589-4614.
112. Moremen, K. W.; Haltiwanger, R. S., *Nat Chem Biol* **2019**, *15* (9), 853-864.
113. Brockhausen, I., *Front Immunol* **2014**, *5*, 492.
114. Moremen, K. W.; Ramiah, A.; Stuart, M.; Steel, J.; Meng, L.; Forouhar, F.; Moniz, H. A.; Gahlay, G.; Gao, Z.; Chapla, D.; Wang, S.; Yang, J. Y.; Prabhakar, P. K.; Johnson, R.; Rosa, M. D.; Geisler, C.; Nairn, A. V.; Seetharaman, J.; Wu, S. C.; Tong, L.; Gilbert, H. J.; LaBaer, J.; Jarvis, D. L., *Nat Chem Biol* **2018**, *14* (2), 156-162.
115. Johnson, K. F., *Glycoconjugate Journal* **1999**, *16* (2), 141-146.
116. Wakarchuk, W. W.; Cunningham, A.; Watson, D. C.; Young, N. M., *Protein Eng* **1998**, *11* (4), 295-302.
117. Zhou, G.; Liu, X.; Su, D.; Li, L.; Xiao, M.; Wang, P. G., *Bioorg Med Chem Lett* **2011**, *21* (1), 311-4.
118. Shao, J.; Zhang, J.; Kowal, P.; Lu, Y.; Wang, P. G., *Biochem Biophys Res Commun* **2002**, *295* (1), 1-8.
119. Gilbert, M.; Brisson, J.-R.; Karwaski, M.-F.; Michniewicz, J.; Cunningham, A.-M.; Wu, Y.; N.M., Y.; Wakarchuk, W. W., *J. Biol. Chem.* **2000**, *275*, 3896-3906.
120. Blixt, O.; Vasiliu, D.; Allin, K.; Jacobsen, N.; Warnock, D.; Razi, N.; Paulson, J. C.; Bernatchez, S.; Gilbert, M.; Wakarchuk, W., *Carbohydr Res* **2005**, *340* (12), 1963-72.
121. Yu, H.; Santra, A.; Li, Y.; McArthur, J. B.; Ghosh, T.; Yang, X.; Wang, P. G.; Chen, X., *Org Biomol Chem* **2018**, *16* (22), 4076-4080.
122. Schmid, J.; Heider, D.; Wendel, N. J.; Sperl, N.; Sieber, V., *Front Microbiol* **2016**, *7*, 182.

Chapter 2

Chemoenzymatic synthesis of the oligosaccharide moiety of the tumor-associated antigen disialosyl globopentaosylceramide

Abstract: Disialosyl globopentaosylceramide (DSGb5) is often expressed by renal cell carcinomas. To investigate properties of DSGb5, we have prepared its oligosaccharide moiety by chemically synthesizing Gb5 which was enzymatically sialylated using the mammalian sialyltransferases ST3Gal1 and ST6GalNAc5. Glycan microarray binding studies indicate that Siglec-7 does not recognize DSGb5, and preferentially binds Neu5Acα(2,8)Neu5Ac containing glycans.

Ingrid M.E. 't Hart, Tiehai Li, Margreet A. Wolfert, Shuo Wang, Kelley W. Moremen, and Geert-Jan Boons

Adapted from: *Org. Biomol. Chem.*, 2019, 17, 7304

Introduction

Glycosphingolipids (GSPs) are a diverse group of biomolecules that are composed of ceramide modified by a glycan. They decorate the cell surface of all vertebrate cells and play important roles in a variety of cellular processes such as cell signalling, trafficking, adhesion, proliferation, and immune modulation.[1] GSPs are also critically involved in embryogenesis and are expressed in stage dependent manner. For example, globopentaosyl ceramide (**1**, Gb5Cer or SSEA-3, Fig. 1) and monosialyl Gb5Cer (**2**, MSGb5Cer or SSEA-4), which belong to the globo-series of GSPs, are well established markers of pluripotent stem cells and their expression is often used for embryonic stem cell (ES) characterization. Recently, a number of other GSPs, including Gb4, Lc4, fucosyl Lc4Cer, Globo H, and disialyl Gb5 (**3**, DSGb5), were identified in undifferentiated human ES and iPS cells.[2] It was found that during ES differentiation, the biosynthesis of globo- and lacto-series GSPs declines and switches to the formation of ganglioside type GSPs.[3]

A number of GSPs that are expressed during early embryogenesis reoccur during oncogenesis. For example, Gb5 and MSGb5 are highly expressed in breast cancer,[4] testicular germ cell tumors,[5] and aggressive human renal cell carcinomas.[6] Furthermore, GSLs such as Globo-H are overexpressed by many epithelial cell cancers and occur on cancer stem cells. The overexpression of these glycolipids appears to promote tumorigenicity,[7] immune suppression,[8] and enhances cancer cell motility and invasiveness.[9] The chemical synthesis of the oligosaccharide moieties of these glycolipids has received considerable attention,[10] which has made it possible to examine biological properties of individual GSPs and opened the way to develop immune-therapeutic strategies such as experimental cancer vaccines.[11]

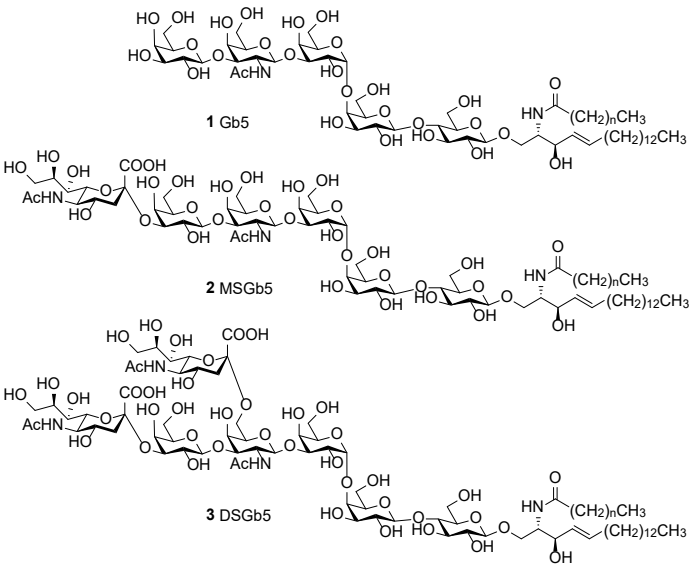


Figure 1. Chemical structures of Gb5 (**1**), MSGb5 (**2**) and DSGb5 (**3**) where the fatty acid chain length can vary between n = 14 – 22.

Here, we report a strategy for the preparation of the oligosaccharide moiety of DSGb5 by chemically synthesizing the pentasaccharide Gb5, which was modified by the mammalian glycosyltransferases ST3Gal1 and ST6GalNAc5. This compound and a number of glycans derived from other globosides and gangliosides were printed as a glycan microarray and their interaction with Siglec-7 was investigated.

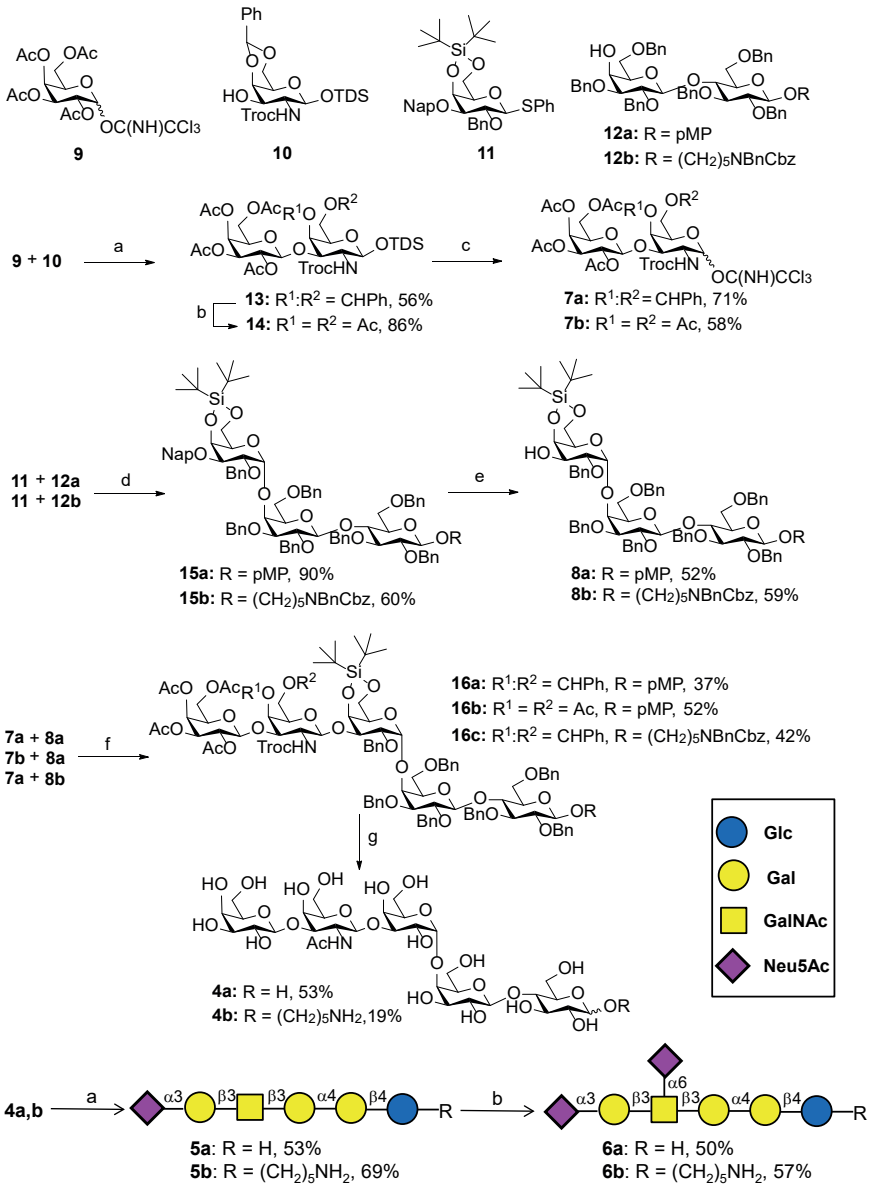
A challenging aspect of the preparation of DSGb5 is the regio- and stereoselective introduction of the α2,3- and α2,6-linked sialosides.[12] Recently, a number of microbial sialyltransferases have been described that make it possible to prepare gangliosides from the ganglio-, lacto-, and globo-series having α2,8-Neu5Ac-α2,3-Neu5Ac and/or α2,3-sialosides at the terminal galactose.[13] These enzymes (e.g. PmST1 and CstII) can readily be expressed in *E. coli* making it straightforward to install these sialosides. To date, no microbial sialyltransferase has been identified that can selectively install an α2,6-linked sialoside at GalNAc of a Gal-β1,3-GalNAc epitope. Recently, considerable progress has been made in the expression of human sialyltransferases,[14] and therefore we were compelled to investigate whether DSGb5 can be prepared by human ST3Gal1 and ST6GalNAc6, which are enzymes that can install an α2,3- and α2,6- sialoside at the terminal Gal and internal GalNAc, respectively.[15]

Results and Discussion

Chemical synthesis of core pentasaccharide Gb5

It was envisaged that the oligosaccharide moiety of Gb5 (**4a**) could be assembled by block coupling of disaccharide **7a** with trisaccharide **8a**, which in turn were expected to be available from building blocks **9**, **10**, **11** and **12** (Scheme 1). We anticipated that a chemical strategy to prepare Gb5 would be more attractive than reported chemoenzymatic[16] or enzymatic approaches using microbial enzymes because these give low conversions, requiring large quantities of enzyme and can be promiscuous to give unwanted side products such as Gb5 modified by additional β1,3-Gal that is difficult to remove.**13b**, [17] Thus, a TMSOTf catalyzed glycosylation of trichloroacetimidate **9** with galactosamine acceptor **10** gave disaccharide **13** in a yield of 56% as only the β-anomer. The anomeric TDS protecting group of **13** could easily be cleaved by treatment with HF-pyridine in pyridine to give a lactol, which was converted into trichloroacetimidate **7a** by reaction with trichloroacetonitrile in the presence of Cs_2CO_3. The use of DBU as the base resulted in substantially lower yield due to hydrolysis of the base sensitive Troc protecting group.

Trisaccharide acceptor **8a** was prepared by a glycosylation of thioglycosyl donor **11** with lactosyl acceptor **12a** using NIS/TMSOTf as the promoter system. The glycosylation proceeded with absolute α-anomeric selectivity due to the presence of the bulky 4,6-di-O-*tert*-butyl-silane protecting group that sterically blocks the β-face of the acceptor.[18] The trisaccharide was isolated in a yield of 90% after purification by silica column chromatography. Treatment of **15a** with DDQ in a mixture of DCM and PBS (24/1 v/v) gave, after purification by silica column chromatography, acceptor **8a** in a yield of 52%.



Scheme 1. Chemoenzymatic synthesis of DSGb5. (**A**) Chemical synthesis of Gb5. Reagents and conditions: (a) TMSOTf, CH_2Cl_2, 4 Å MS, −35 °C; (b) i. 80% AcOH, 80 °C, ii. Ac_2O, pyridine, DMAP; (c) i. HF·pyridine, pyridine, ii. Cs_2CO_3, Cl_3CCN, CH_2Cl_2, 0 °C; (d) NIS, TfOH, CH_2Cl_2, 4 Å MS, −30 °C; (e) DDQ, CH_2Cl_2/PBS buffer; (f) TMSOTf, CH_2Cl_2, 4 Å MS, −30 °C; (g) i. HF·pyridine, ii. NaOH, THF, 80 °C, iii. Ac_2O, pyridine, iv. CAN, CH_3CN/H_2O, 0 °C (for **16a** and **16b**), v. NaOMe, MeOH, vi. $Pd(OH)_2/C$, H_2, MeOH/H_2O/AcOH. (**B**) Enzymatic extension to produce DSGb5. Reagents and conditions: (a) ST3Gal1, CMP-Neu5Ac, CIAP, $MgCl_2$ (20 mM), sodium cacodylate buffer (50 mM, pH 7.5), 37 °C; (b) ST6GalNAc5, CMP-Neu5Ac, CIAP, $MgCl_2$ (20 mM), sodium cacodylate buffer (50 mM, pH 7.5), 37 °C.

Glycosylation of disaccharide **7a** and trisaccharide **8a** in the presence of TMSOTf in DCM at -30 °C resulted in the formation of pentasaccharide **16a** in a yield of 59% as a separable mixture of α/β anomers (β/α= 1.7). Optimization of the reaction conditions revealed that the overall yield of the glycosylation could be increased by lowering the reaction temperature, however, this did not affect the poor anomeric selectivity (SI, Table S1). Fortunately, the use of glycosyl donor **7b**, having acetyl esters instead of a benzylidene acetal at the 4,6-diol of GalNTroc,[19] gave in a TMSOTf mediated glycosylation with acceptor **8a** at -50 °C, pentasaccharide **16b** as only the β-anomer in an isolated yield of 52%.

Pentasaccharide **16a** was deprotected by the six-step procedure to give Gb5 (**4a**). Thus, the silyl protecting group was cleaved by treatment with HF·pyridine which was followed by hydrolysis of the acetyl esters and Troc protecting group with aqueous NaOH in THF with heating (80 °C). The resulting compound was acetylated with acetic anhydride in pyridine and then the anomeric methoxyphenyl (MP) protecting group was oxidatively removed by cerium ammonium nitrate (CAN) in a mixture of acetonitrile and H_2O. Finally, deacetylation under Zemplén conditions (*cat.* NaOMe in MeOH) followed by hydrogenation over $Pd(OH)_2/C$ in a mixture of $MeOH/H_2O$/HOAc afforded Gb5 (**4a**) in an overall yield of 53% after purification by Bio-Gel P-2 size exclusion chromatography followed by semi-preparative HPLC using a HILIC column (XBridge® Amide 5 µm, 4.6 mm x 250 mm, Waters).

Enzymatic synthesis of MSGb5 and DSGb5

Next, attention was focused on the enzymatic sialylation of **4a** to give the oligosaccharide moiety of DSGb5. Thus, **4a** was treated with ST3Gal1 in the presence of CMP-Neu5Ac (1.5 eq.) in sodium cacodylate buffer (pH = 7.5, 50 mM) containing $MgCl_2$ (20 mM) at 37 °C. The reaction was performed in the presence of CIAP to hydrolyse CMP which may cause product inhibition.[20] Analysis of the reaction mixture by TLC and MALDI-TOF MS indicated that after an incubation time of 4 days, all starting material had been converted into product. Interestingly, Gb5 proved to be a rather poor substrate for the microbial α2,3-sialyltransferase (PmST1) and even after prolonged incubation, only partial conversion was observed (~30%). Next, compound **5a** was treated with recombinant ST6GalNAc6 and surprisingly, no product formation was detected. Gratifyingly, the use of ST6GalNAc5 could readily add the second sialoside to provide DSGb5 (**6a**). We found this enzyme requires an α2,3-linked sialoside for activity and the use of Gal(β1,3) GalNAc did not give product whereas Neu5Ac(α2,3)Gal(β1,3)GalNAc was readily modified by ST6GalNAc5. These results indicate that the biosynthesis of DSGb5 involves an orchestrated attachment of the sialosides in which the 2,3-linked Neu5Ac is first installed, followed by the introduction of the 2,6-sialoside. MSGb5 (**5a**) and DSGb5 (**6a**) were purified by Bio-Gel P-2 size exclusion column chromatography followed HPLC using a HILIC column and the resulting compounds were fully characterized by high resolution mass spectrometry and multi-dimensional NMR. A ROESY experiment showed close proximity of H-4 and H-6 of GalNAc with H-3ax of branching Neu5Ac confirming proper connectivity of the α2,6-sialoside of DSGb5. We also prepared the oligosaccharides of Gb5, MSGb5 and DSGb5 modified by an amino pentyl linker (**4b**, **5b** and **6b**, respectively) in a similar fashion by using acceptor **12b** instead of **12a**. These compounds were printed as a microarray to explore ligand requirements of Siglec-7.

Microarray binding studies

The Siglecs are a family of transmembrane cell surface receptors expressed on hemopoietic cells that can bind specific sialic acid containing glycoconjugates.[21] Such binding events result in inhibitory signals that dampen innate and adaptive immune responses. Siglec-7 is predominantly expressed on natural killer (NK) cells, and its engagement with specific sialoglycans on target cells results in inhibition of NK cellular toxicity. Over-expression of Siglec-7 ligands on cancer cells is a proposed mechanism of immune escape, and reversal of such interactions may lead to a new class of checkpoint inhibitors.[22] *In vitro* binding studies have indicated that Siglec-7 has a preference for glycoconjugates bearing a Neu5Acα(2,8)Neu5Ac motif such as present in b-series of gangliosides including GD3, GD2, GT1b and GQ1b.[23] It has been shown that the expressing of GD3 on a target cells leads to suppression of NK mediated cytolytic activity in a Siglec-7 dependent manner.[24] There are also indications that glycans bearing an internal branching α(2,6)-linked sialic acid at GalNAc or GlcNAc, such as present in LSTb, disialyl Lewisa and DSGb5, can also be recognized by Siglec-7.[23, 25]

A glycan microarray was created by piezoelectric non-contact printing of compounds **4b**, **5b** and **6b** and GM3 (**17**), GM2 (**18**), GM1a (**19**), GD3 (**20**) and GT1b (**21**)[13b,26] on N-hydroxysuccinimide (NHS)-activated glass slides as replicates of 6. To validate proper printing, the microarray was examined for binding of the biotinylated lectins *Maackia amurensis* leukagglutinin (MAL-II), soybean agglutinin (SBA) and wheat germ agglutinin (WGA). These lectins were preincubated with Streptavidin-AlexaFluor635 and binding of immobilized oligosaccharides was established by measuring fluorescence intensity using a microarray scanner. As anticipated, MAL-II, which is known to recognize Neu5Ac(α2,3) Gal(β1,4)GlcNAc/Glc, did bind GM3 (**17**) and GD3 (**20**) having such an oligosaccharide fragment (Fig. 2). SBA recognized compounds **4b** and **18**, which contain a terminal Gal and GalNAc residue, respectively that are known to be ligands for this lectin. WGA, which binds the GlcNAc moiety but also some forms of sialic acid, bound compounds **5b**, **6b** and **21**, indicating it has a preference for Neu5Ac(α2,3)Gal(β1,3)GalNAc containing oligosaccharides. Next, the array was incubated with biotin-conjugated ganglioside GM1 polyclonal antibody, and as anticipated only binding to GM1 was detected. Interestingly, a similar binding experiment with recombinant human Siglec-7 comp showed binding to GD3 (**20**) and GT1b (**21**), however no recognition of DSGb5 (**6b**) was observed. These results indicate that Siglec-7 has a strong preference for α2,8-Neu5Ac-α2,3-Neu5Ac containing oligosaccharides, and has low or no affinity for DSGb5, which has a branching α2,6- and a terminal α2,3-sialoside.[27] The previously proposed interaction of DSGb5 with Siglec-7 was based on the observation that cells that express DSGb5 bind to Siglec-7 transfected COS-7 cells. Furthermore, knockdown of ST6GalNAc6 of renal cancer cells resulted in a reduced expression of DSGb5 and a substantial lower binding of a Siglec-7-Fc fusion protein.[25a,28] ST6GalNAc6 has been implicated in the biosynthesis of various other gangliosides including GM1b, GT1b and GD1a, and it is likely that knockdown of ST6GalNAc6 results in a lower expression of these gangliosides, which may affect Siglec-7 binding.[29] Furthermore, this transferase is also involved in the biosynthesis of disiayl Lewis[a], which is also a proposed ligand for Siglec-7.[30]

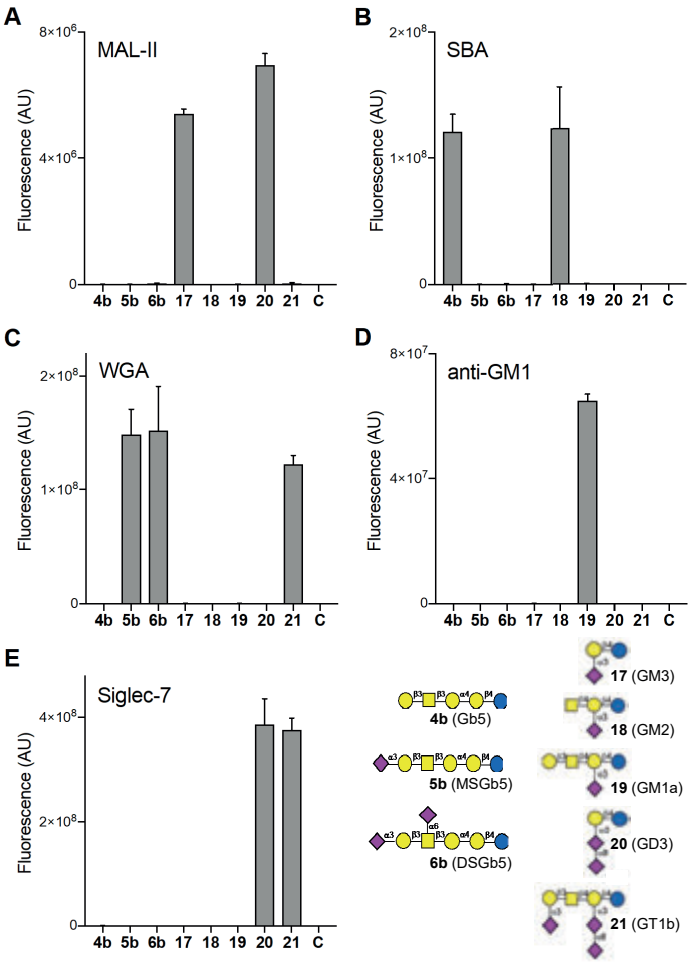


Figure 2. Microarray results of the synthetic compounds with (**A**) MAL-II, (**B**) SBA, (**C**) WGA, (**D**) anti-GM1 and (**E**) Siglec-7. C indicates buffer control. Bars represent the mean ± SD.

Although no binding was observed on the glycan microarray, it cannot be excluded that the ceramide moiety of DSGb5 can modulate Siglec-7 binding by organizing it into microdomains for low-affinity high-avidity multivalent interactions.[31] Therefore, our future studies will focus on the preparation of DSGb5 having a ceramide moiety for binding and cellular activation studies. Such a compound will also be valuable to explore other cellular roles of DSGb5 such as promoting tumor cell migration.[32] Furthermore, a very recent study, employing a library of isogenic HEK293 cells with combinatorially engineered glycosylation capacities, indicated that Siglec-7 recognizes core 2 *O*-glycans having α2,3 and 2,6-linked sialosides which will also be an important target for future synthesis.[33] It is clear that the ability of glycans having an internal α2,6-sialoside at GlcNAc and GalNAc to mediate cellular activation in Siglec-7 dependent manner needs further investigation.

Conclusions

In conclusion, DSGb5 was synthesized by a chemoenzymatic approach in which the oligosaccharide moiety of Gb5 was assembled chemically by a block coupling approach followed by enzymatic sialylation using the mammalian sialyltransferases, ST3Gal1 and ST6GalNAc5 to install an α2,3- and α2,6-linked sialoside, respectively. Glycan microarray binding studies indicate that the oligosaccharide moiety of DSGb5 is not recognized by Siglec-7, and it is likely this ganglioside promotes tumorigenesis through other mechanisms such as increasing cell migration and invasion.

Experimental Section

Chemical synthesis

General procedures

All chemicals were purchased from commercial sources. NMR spectra (^{1}H, ^{13}C, COSY, HSQC) were obtained on an Agilent 400-MR DD2 or Bruker 750 MHz. Chemical shifts are reported in part per million (ppm) relative to CDCl3 (7.26 ppm), TMS (0.00 ppm) or D2O (4.79 ppm). NMR data is presented as: chemical shift, multiplicity (where s = singlet, d = doublet, t = triplet, dd = doublet of doublets, m = multiplet) and the coupling constant in Hertz (Hz). Mass spectra were obtained on a Shimadzu ESI LC-MS QP8000 or Kratos Analytical Maxima-CFR MALDI-TOF system (using 2,5-dihydroxybenzoic acid matrix). Reported HRMS data was obtained on an Agilent technologies 6560 Ion mobility Q-TOF. Semi-preparative HPLC was performed on an Applied Biosystems 400 solvent delivery system and 757 Absorbance Detector (UV absorbance set on 214 nm) using HILIC column (XBridge® Amide 5 µm, 4.6 mm x 250 mm column, Waters). The mobile phase for analytical and semi-preparative HPLC runs consisted of buffers A and B. For C18 columns buffer A is 0.1 % TFA in H_2O and buffer B is 10 % A + 90 % CH_3CN and a gradient was used. For HILIC column chromatography buffer A is 10 mM NH_4COOH in H_2O (pH = 4) and B is 10 % A + 90 % CH_3CN at isocratic conditions. Size exclusion chromatography was performed on Bio-Gel P-2 (45-90 µm) with water as the eluent. Column chromatography was performed on silica gel G60 (Silicycle 60 – 200 µm, 60 Å). TLC analysis was conducted on silica gel 60 F254 (EMD Chemicals Inc.) with detection by UV light (254 nm) and staining by 10 % H2SO4 in EtOH or *p*-anisaldehyde solution, followed by heating for visualization. Molecular sieves (4 Å) were flame-dried prior to use.

NMR nomenclature

The monosaccharides of glycan DSGb5 have been labeled as shown in Fig. S1. Starting from the reducing end of the pentasaccharide core Gb5, these were labeled as Glc-I, Gal-II, Gal-III, GalNAc-IV, Gal-V. The sialosides were named Neu5Ac-VI for the α2,3-linked sialic acid and Neu5Ac-VII for the α2,6-linked sialic acid.

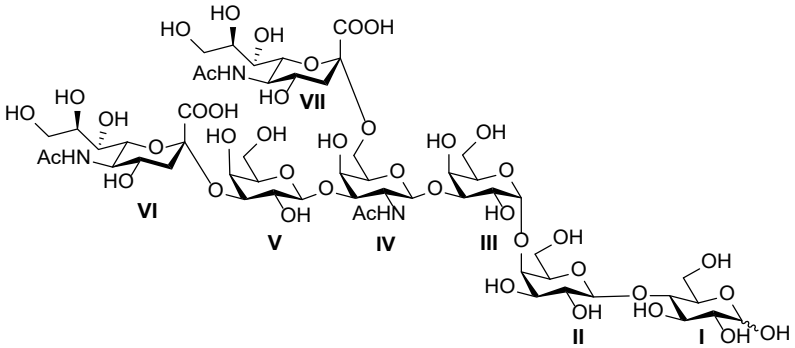
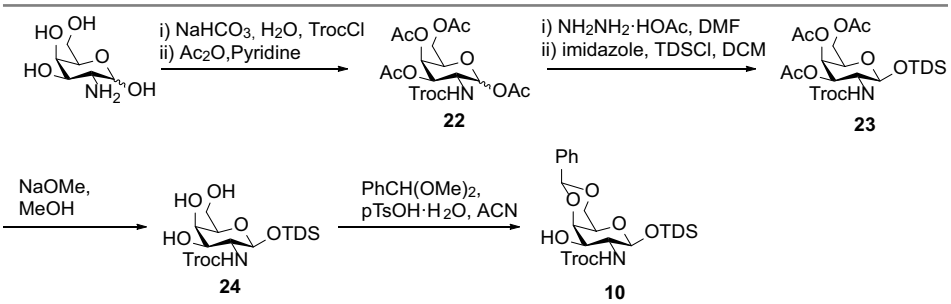


Figure S1. Monosaccharide labeling system for DSGb5

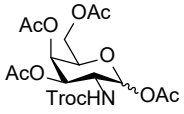
Experimental procedures

2,2,2-Trichloroacetimidate 2,3,4,6-O-acetyl-α-D-galactopyranoside (9). Compound **9** was synthesized according to previous synthesis.[34] NMRs of the α-anomer of the title compound are described below. 1H NMR (400 MHz, $CDCl_3$) δ 8.66 (1H, s, NH), 6.58 (1H, d, J = 3.1 Hz, H-1), 5.55 (1H, d, J = 2.6 Hz, H-4), 5.45 – 5.31 (2H, m, H-3, H-2), 4.43 (1H, t, J = 6.6 Hz, H-5), 4.20 – 4.01 (2H, m, H-6), 2.15 (3H, s, OAc), 2.03 – 1.98 (9H, m, 3x OAc). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.2 (C, OAc), 170.0 (C, OAc), 170.0 (C, OAc), 169.9 (C, OAc), 160.8 (C=NH), 93.4 (C-1), 68.9 (C-5), 67.40, 67.27, 66.80, 61.2 (C-6), 20.6 (CH_3, OAc), 20.5 (CH_3, OAc), 20.5 (CH_3, OAc), 20.4 (CH_3, OAc).

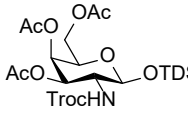


Scheme S1. Synthesis of GalNHTroc acceptor **10** (similar to the procedure as described for $GlcNH_2$).[35]

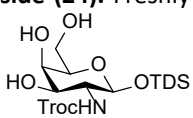
1,3,4,6-tetra-O-acetyl-2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galactopyranoside (22). $NaHCO_3$ (23 g, 278 mmol) was added to Galactosamine·HCl (20 g, 93 mmol) in H_2O (180 mL) and after 30 min 2,2,2-trichloroethyl chloroformate (15.3 mL, 111 mmol) was added. After overnight stirring the white solids were filtered off, washed with H_2O and dried under high vacuum overnight. The solids were dissolved in pyridine (100 mL) and acetic anhydride (80 mL) was added. The reaction mixture was stirred for 3 h and concentrated *in vacuo*. The obtained oil was dissolved in DCM, washed with 1M HCl (2x), H2O, sat. aq. $NaHCO_3$ and dried (Na_2SO_4), filtered and concentrated *in vacuo* to afford **22** (19 g, 40%, over two steps). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{17}H_{22}Cl_3NO_{11}$, 544.0156; found 544.0155. [α] 25/589 = 326.1° (C = 0.1; $CHCl_3$).



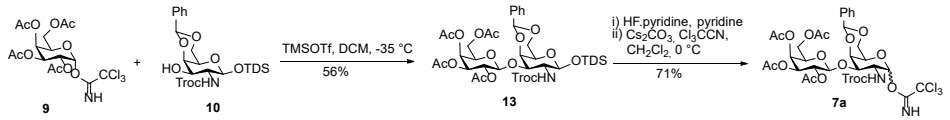
Dimethylthexylsilyl 3,4,6-tri-O-acetyl-2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galactopyranoside (23). Hydrazine acetate (3.7 g, 40 mmol) was added to a solution of compound **22** (19 g, 36 mmol) in DMF (60 mL). The mixture was stirred overnight, concentrated *in vacuo*, dissolved in EtOAc, washed with sat. aq. $NaHCO_3$, H_2O and dried (Na_2SO_4), filtered and concentrated *in vacuo*. The resulting crude was dissolved in DCM (80 mL) and imidazole (7.4 g, 109 mmol) was added. When all imidazole was dissolved, *tert*-hexyldimethylsilyl chloride (8.6 mL, 44 mmol) was added. The mixture was stirred overnight, washed with 1 M HCl, sat. aq. $NaHCO_3$, dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 6:4 v/v) as the eluent to afford **23** (17.2 g, 76%, over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.35 (1H, d, J = 3.2 Hz, H-4), 5.16 (1H, d, J = 9.9 Hz, H-3), 4.95 (1H, d, J = 8.3 Hz, NH), 4.84 – 4.68 (1H, m, H-1; CHH, Troc), 4.62 (1H, d, J = 11.7 Hz, CHH, Troc), 4.20 – 4.05 (2H, m, H-6), 3.89 (1H, t, J = 6.6 Hz, H-5), 3.78 (1H, dd, J = 18.4, 9.2 Hz, H-2), 2.16 (2H, s, OAc), 2.04 (3H, s, OAc), 1.99 (3H, s, OAc), 1.66 – 1.58 (1H, m, CH, TDS), 0.95 – 0.76 (12H, m, 4x CH_3, TDS), 0.17 (3H, s, CH_3-Si), 0.14 (3H, s, CH_3-Si). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.5 (C, OAc), 170.4 (C, OAc), 170.3 (C, OAc), 154.0 (C=O, Troc), 96.4 (C-1), 95.3 (CCl_3), 74.5 (CH_2, Troc), 70.7 (C-5), 69.9 (C-3), 66.9 (C-4), 61.8 (C-6), 54.7 (C-2), 33.9 (CH, TDS), 24.8 (C, TDS), 20.7 (CH_3, OAc), 20.6 (CH_3, OAc), 20.6 (CH_3, OAc), 19.9 (2x CH_3, TDS), 18.5 (2x CH_3, TDS), -1.9 (CH_3-Si), -3.4 (CH_3-Si). ESI HRMS (m/z): $[M + NH_4]^+$ calcd for $C_{23}H_{38}Cl_3NO_{10}Si$, 639.1669; found 639.1675. [α] 25/589 = -32.7° (C = 0.1; $CHCl_3$).



Dimethylthexylsilyl 2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galactopyranoside (24). Freshly prepared NaOMe was added to compound **23** (17.3 g, 28 mmol) in MeOH (50 mL). After 2 h the reaction was quenched by addition of Amberlite H^+ resin, filtered and concentrated *in vacuo* to afford **24** (13 g, 95%). This product was then used in the next step without additional purification. ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{17}H_{32}Cl_3NO_7Si$, 518.0911; found 518.0913. [α] 25/589 = -77.0° (C = 0.1; $CHCl_3$).



Dimethylthexylsilyl 4,6-O-benzylidene-2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galacto-pyranoside (10). Benzaldehyde dimethyl acetal (4.57 mL, 30.4 mmol) and pTsOH·H_2O (1.05 g, 5.5 mmol) were added to a solution of compound **24** (13 g, 26 mmol) in CH_3CN (90 mL). After 1 h the mixture was quenched with Et_3N, concentrated *in vacuo* and the obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 3:1 v/v) as the eluent to obtain compound **10** (4.6 g, 30%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 – 7.48 (2H, m, H-Ar), 7.43 – 7.35 (3H, m, H-Ar), 5.57 (1H, s, CH-C_6H_5), 5.05 (1H, s, NH), 4.80 (1H, d, J = 7.8 Hz, H-1), 4.69 (2H, s, CH_2, Troc), 4.28 (1H, d, J = 12.3, 1.3 Hz, H-6a), 4.20 (1H, d, J = 3.6 Hz, H-4), 4.07 (1H, d, J = 12.4, 1.8 Hz, H-6b), 3.92 (1H, d, J = 8.9 Hz, H-3), 3.63 (1H, d, J = 9.4 Hz, H-2), 3.47 (1H, s, H-5), 2.73 (1H, d, J = 9.0 Hz, OH), 1.66 – 1.58 (1H, m, CH, TDS), 0.94 – 0.78 (12H, m, 4x CH_3, TDS), 0.22 (3H, s, CH_3-Si), 0.17 (3H, s, CH_3-Si). ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.7 (C=O, Troc), 137.5 (C, Ar), 129.3, 128.5, 128.3, 126.4, 101.4 (CH-C_6H_5), 95.7 (C-1), 75.0 (C-4), 74.7 (CH_2, Troc), 70.4 (C-3), 69.3 (C-6), 66.5 (C-5), 57.8 (C-2), 34.0 (CH, TDS), 20.1 (CH_3, TDS), 20.1 (CH_3, TDS), 18.5 (CH_3, TDS), 18.5 (CH_3, TDS), -1.7 (CH_3-Si), -2.9 (CH_3-Si). ESI HRMS (m/z): [M + Na]$^+$ calcd for $C_{24}H_{36}Cl_3NO_7Si$, 606.1224; found 606.1227. [α] 25/589 = -22.5° (C = 1; $CHCl_3$).

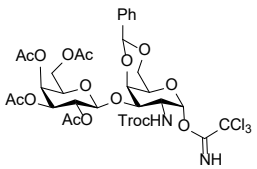


Scheme S2. Chemical glycosylation of donor **9** and acceptor **10** and formation of disaccharide donor **7a**.

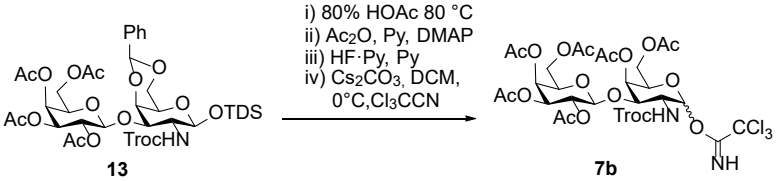
Dimethylthexylsilyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-4,6-O-benzylidene-2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galactopyranoside (13). A mixture of acceptor **10** (1.0 g, 1.7 mmol), donor **9** (1.3 g, 2.6 mmol) and 4 Å molecular sieves was stirred in DCM (5 mL) for 2 h. The reaction mixture was cooled to -35°C and TMSOTf (62 µL, 0.3 mmol) was added. After 30 min the reaction was quenched with Et_3N, filtered over a pad of Celite and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 8.5:1.5 v/v) as the eluent to afford compound **13** (878 mg, 56 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 – 7.52 (2H, m, H-Ar), 7.42 – 7.31 (3H, m, H-Ar), 5.55 (1H, s, CH-C_6H_5), 5.36 (1H, dd, J = 3.4, 0.8 Hz, H-4, Gal-V), 5.32 – 5.26 (1H, m, NH), 5.21 (1H, d, J = 7.9 Hz, H-1, GalNAc-IV), 5.13 (1H, d, J = 7.7 Hz, H-2, Gal-V), 4.96 (1H, dd, J = 10.4, 3.5 Hz, H-3, Gal-V), 4.78 (1H, d, J = 7.9 Hz, H-1, Gal-V), 4.75 – 4.60 (2H, m, CH_2, Troc), 4.46 (1H, dd, J = 11.1, 2.7 Hz, H-3, GalNAc-IV), 4.29 (1H, d, J = 3.3 Hz, H-4, GalNAc-IV), 4.25 (1H, d, J = 12.2, 1.1 Hz, H-6a, GalNAc-IV), 4.22 – 4.07 (2H, m, H-6, Gal-V), 4.03 (1H, d, H-6b, GalNAc-IV), 3.88 (1H, t, J = 6.3 Hz, H-5, Gal-V), 3.54 – 3.36 (2H, m, H-2, GalNAc-IV; H-5, GalNAc-IV), 2.15 (3H, s, OAc), 2.06 (3H, s, OAc), 2.04 (3H, s, OAc), 1.97 (3H, s, OAc), 1.63 (1H, p, J = 13.7, 6.9 Hz, CH, TDS), 0.92 – 0.78 (12H, m, 4x CH_3, TDS), 0.19 (3H, s, CH_3-Si), 0.13 (3H, s, CH_3-Si). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.2 (C, OAc), 170.0 (C, OAc), 169.3 (C, OAc), 153.8 (C=O, Troc), 138.0 (C, C_6H_5), 129.0, 128.2, 126.3, 101.6 (C-1, Gal-V), 100.7 (CH-C_6H_5), 95.20 (CCl_3), 94.5 (C-1, GalNAc-IV), 76.1 (C-4,

GalNAc-IV), 75.7 (C-3, GalNAc-IV), 74.5 (CH_2, Troc), 70.8 (C-5, Gal-V), 70.8 (C-3, Gal-V), 69.3 (C-6, GalNAc-IV), 68.8 (C-2, Gal-V), 67.0 (C-4, Gal-V), 66.4 (C-5, GalNAc-IV), 61.6 (C-6, Gal-V), 55.9 (C-2, GalNAc-IV), 34.0 (CH, TDS), 24.8 (C, TDS), 20.7 (CH_3, OAc), 20.7 (CH_3, OAc), 20.7 (CH_3, OAc), 20.5 (CH_3, OAc), 20.1 (CH_3, TDS), 20.0 (CH_3, TDS), 18.6 (CH_3, TDS), 18.5 (CH_3, TDS), -1.8 (CH_3-Si), -3.1 (CH_3-Si). ESI HRMS (m/z): $[M + NH_4]^+$ calcd for $C_{38}H_{54}Cl_3NO_{16}Si$, 931.2616; found 931.2634. [α] 25/589 = 223.8° (C = 0.1; $CHCl_3$).

2,2,2-Trichloroacetimidate 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-4,6-O-benzylidene-2-[(2,2,2-trichloromethoxy)carbonylamino]-α-D-galactopyranoside (7a).

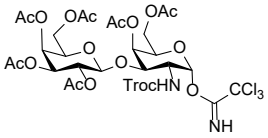


HF·Pyridine (70% HF, 1.2 mL) was added to a stirring solution of disaccharide **13** (1.2 g, 1.3 mmol) in pyridine (12 mL) in a plastic round bottom flask. After 2.5 h, the mixture was diluted with DCM and quenched by addition of sat. aq. $NaHCO_3$. The organic phase was washed with sat. aq. $NaHCO_3$ (2 x), dried (Na_2SO_4), filtered, concentrated *in vacuo* and 2 x co-evaporated with toluene. The obtained intermediate was dissolved in DCM, stirred with 4 Å molecular sieves for 30 min and 2,2,2-Trichloroethyl chloroformate (615 µL, 6.14 mmol) and Cs_2CO_3 (400 mg, 1.23 mmol) were added. After 3 h the reaction mixture was concentrated *in vacuo* and the obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 1:1 v/v) as the eluent to isolate the α-anomer of the title compound. (859 mg, 71 %, over 2 steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.72 (1H, s, C=NH), 7.59 – 7.47 (2H, m, Ar-H), 7.44 – 7.29 (3H, m, Ar-H), 6.67 (1H, d, J = 3.2 Hz, H-1, GalNAc-IV), 5.54 (1H, s, CH-C_6H_5), 5.41 (1H, d, J = 2.8 Hz, H-4, Gal-V), 5.32 – 5.20 (2H, m, H-2, Gal-V; NHTroc), 5.02 (1H, dd, J = 10.3, 3.3 Hz, H-3, Gal-V), 4.88 (1H, d, J = 8.1 Hz, H-1, Gal-V), 4.81 (1H, d, J = 12.1 Hz, CHH, Troc), 4.63 (1H, d, J = 12.1 Hz, CHH, Troc), 4.56 (1H, dd, J = 10.9, 7.4, 3.1 Hz, H-2, GalNAc-IV), 4.46 (1H, d, J = 2.8 Hz, H-4, GalNAc-IV), 4.34 (1H, d, J = 11.9 Hz, H-6a, GalNAc-IV), 4.27 (1H, dd, J = 11.2, 3.1 Hz, H-3, GalNAc-IV), 4.22 – 3.98 (4H, m, H-6, Gal-V; H-6b, GalNAc-IV; H-5, Gal-V), 3.88 (1H, s, H-5, GalNAc-IV), 2.18 (3H, s, OAc), 2.05 (3H, s, OAc), 2.04 (3H, s, OAc), 1.99 (3H, s, OAc). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.2 (C, OAc), 170.0 (C, OAc), 169.9 (C, OAc), 169.6 (C, OAc), 160.2 (C=NH), 154.0 (C=O, Troc), 137.4 (C, C_6H_5), 129.1, 128.2, 126.2, 101.0 (CH-C_6H_5), 100.0 (C-1, Gal-V), 96.3 (C-1, GalNAc-IV), 95.4 (CCl_3), 74.6 (C-4, GalNAc-IV), 74.5 (CH_2, Troc), 71.8 (C-3, GalNAc-IV), 71.3 (C-5, Gal-V), 70.8 (C-3, Gal-V), 68.9 (C-6, GalNAc-IV), 68.4 (C-2, Gal-V), 66.4 (C-4, Gal-V), 65.3 (C-5, GalNAc-IV), 60.9 (C-6, Gal-V), 49.9 (C-2, GalNAc-IV), 20.7 (CH_3, OAc), 20.7 (CH_3, OAc), 20.6 (CH_3, OAc), 20.5 (CH_3, OAc).

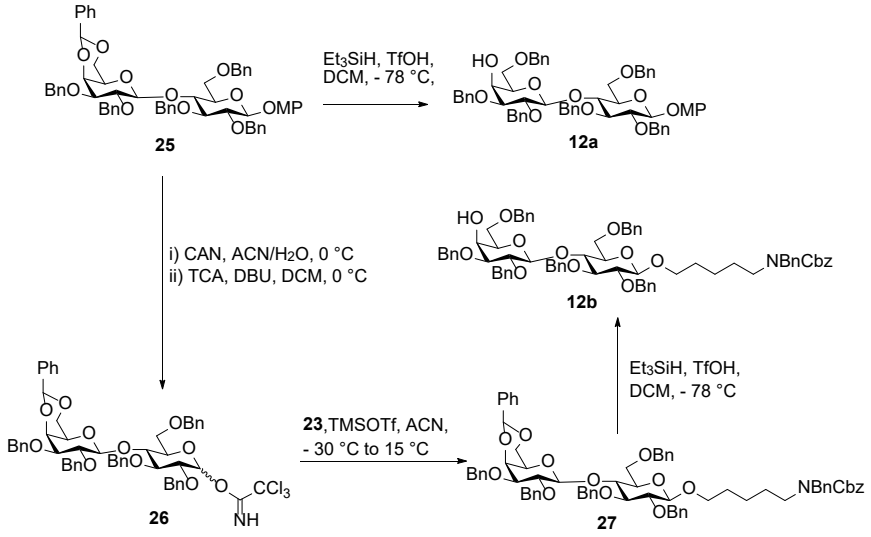


Scheme S3. Formation of disaccharide donor **7b** from disaccharide **13**.

2,2,2-Trichloroacetimidate 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-4,6-di-O-acetyl-2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galactopyranoside (7b). A



solution of disaccharide **13** (350 mg, 0.38 mmol) in 80% aq AcOH (4 mL) was heated to 80°C for 4 h. The mixture was allowed to cool to room temperature (RT) and was diluted with EtOAc, washed with H_2O, sat aq. $NaHCO_3$ (3x), dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained crude was dissolved in pyridine (3 mL) and Ac_2O (2 mL) was slowly added, followed by DMAP (cat.). After 1 h, the reaction showed complete conversion by TLC and the mixture was concentrated *in vacuo*. The crude was dissolved in pyridine (3.5 mL) and transferred to a plastic round bottom flask. HF·Pyridine (70% HF, 350 µL) was added and the mixture was stirred overnight. The reaction mixture was diluted with EtOAc, washed with sat. aq. $NaHCO_3$ (3x), dried (Na_2SO_4), filtered, concentrated *in vacuo* and co-evaporated with toluene. Quick silica column purification Hexane:EtOAc (1:0 to 1:3 v/v) provided the intermediate in 81% yield. The obtained intermediate (237 mg, 0.31 mmol) was dissolved in DCM and stirred with 4 Å molecular sieves at 0°C. 2,2,2-trichloroethyl chloroformate (297 µL, 2.96 mmol) and Cs_2CO_3 (301 mg, 0.92 mmol) were added after 30 min. After 22 h the reaction mixture was concentrated *in vacuo* and the obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 1:1 v/v) as the eluent to isolate the α-anomer of the title compound (143 mg, 53 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.75 (1H s, C=NH), 6.57 (1H, d, J = 3.3 Hz, H-1, GalNAc-IV), 5.50 – 5.34 (3H, m, H-4, GalNAc-IV; NHTroc; H-4, Gal-V), 5.28 – 5.20 (1H, m, H-2, Gal-V), 4.99 (1H, dd, J = 10.3, 3.2 Hz, H-3, Gal-V), 4.84 (1H, , J = 12.1 Hz, CHH, Troc), 4.76 (1H, d, J = 8.2 Hz, H-1, Gal-V), 4.62 (1H, d, J = 12.0 Hz, CHH, Troc), 4.41 – 4.34 (1H, m, H-2, GalNAc-IV), 4.31 (1H, t, J = 6.4 Hz, H-5, GalNAc-IV), 4.25 – 4.04 (4H, m, H-6a, GalNAc-IV; H-3, GalNAc-IV; H-6, Gal-V), 4.04 – 3.92 (2H, m, H-5, Gal-V; H-6b, GalNAc-IV), 2.20 (3H, s, OAc), 2.15 (3H, s, OAc), 2.10 (3H, s, OAc), 2.04 (6H, s, 2x OAc), 1.98 (3H, s, OAc).



Scheme S4. Synthesis of disaccharide acceptors **12a** and **12b** from protected lactose **25**.

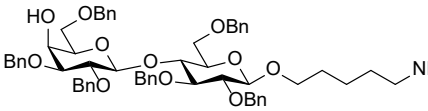
Para-methoxyphenyl 2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-gluco-pyranoside (12a). Compound **25**[36] (8.1 g, 8.2 mmol) was stirred in DCM (50 mL) with 4 Å molecular sieves for 3 h. The mixture was cooled to -78°C and after adding Et_3SiH (6.5 mL, 41 mmol) stirring was continued for another 30 min. TfOH (1.45 mL, 16.4 mmol) was introduced and after 2.5 h, the reaction mixture was quenched with Et_3N. The resulting mixture was filtered over Celite and the filtrate was washed with H_2O, sat. aq. $NaHCO_3$, dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 20:1 v/v) as the eluent to afford compound **12a** (2.28g, 64 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.54 – 7.14 (30H, m, Ar-H), 7.02 (2H, d, J = 9.1 Hz, OMP), 6.79 (2H, d, J = 9.1 Hz, OMP), 5.00 (2H, t, J = 10.7 Hz, CH_2, Bn), 4.85 (1H, d, J = 7.4 Hz, H-1, Glc-I), 4.83 – 4.64 (6H, m, 3x, CH_2, Bn), 4.53 – 4.35 (5H, m, 2x CH_2, Bn; H-1, Gal-II), 4.06 – 3.96 (2H, m, H-4, Gal-II; H-5 Gal-II), 3.82 – 3.74 (4H, m, CH_3, OMP; H-6 Glc-I), 3.72 – 3.57 (4H, m, H-4, Glc-I; H-6a, Gal-II; H-2, Glc-I; H-2, Gal-II), 3.52 – 3.45 (2H, m, H-5, Glc-I; H6-b, Gal-II), 3.40 (1H, dd, J = 9.3, 3.4 Hz, H-3, Gal-II), 3.35 (1H, t, H-3, Glc-I), 2.39 (1H, d, J = 2.0 Hz, OH). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.2 (C, OMP), 151.6 (C, OMP), 139.0 (C, OBn), 138.5 (C, OBn), 138.4 (C, OBn), 138.3 (C, OBn, 138.2 (C, OBn), 137.9 (C, OBn), 128.43, 128.34, 128.28, 128.25, 128.20, 128.08, 128.04, 127.83, 127.78, 127.73, 127.62, 127.59, 127.54, 127.40, 127.25, 118.4 (2x CH, OMP), 114.5 (2x CH, OMP), 102.8 (C-1, Glc-I), 102.6 (C-1, Gal-II), 82.9 (C-4, Glc-I), 81.6 (C-2, Glc-I), 81.1 (C-3, Gal-II), 79.4 (C-2, Gal-II), 77.2 (C-5, Gal-II), 75.4 (CH_2), 75.3 (2x CH_2), 75.1 (C-5, Glc-I), 73.5 (CH_2), 73.1 (CH_2), 72.8 (C-3, Glc-I), 72.0 (CH_2), 68.4 (C-6, Glc-I), 68.3 (C-6, Gal-II), 66.1 (C-4, Gal-II), 55.6 (CH_3, OMP). ESI HRMS (m/z): $[M + NH_4]^+$ calcd for $C_{61}H_{64}O_{12}$, 1006.4736; found 1006.4750.

N-(Benzyl)-benzyloxycarbonyl-5-aminopentan-1-ol (28). The protected aminopentanol linker was synthesized as described before.[37]

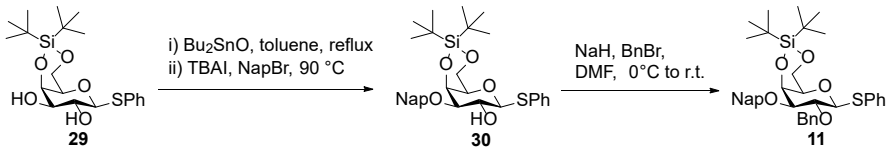
N-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2,3-di-O-benzyl-4,6-O-benzylidene-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (27). Oven-dried (90°C, 1.5 h) ceric ammonium nitrate (1.67 g, 3.0 mmol) was added to a stirring solution of compound **25**[36] (2.0 g, 2.0 mmol) in CH_3CN/H_2O (40/10 mL) at 0°C. After 30 min, the mixture was diluted with DCM, washed with sat. aq. $NaHCO_3$ (2 x), brine, dried (Na_2SO_4), filtered and concentrated *in vacuo*. Purification by silica column chromatography using Hexane:EtOAc (1:0 to 1:1 v/v) as the eluent provided the product, which was directly used in the next step. 2,2,2-Trichloroacetonitrile (848 µL, 8.5 mmol) and DBU (51 µL, 0.3 mmol) were added to the intermediate (1.49 g, 1.7 mmol) in DCM (3 mL) with 4 Å molecular sieves at 0°C. After 15 min the reaction mixture was concentrated *in vacuo* and the obtained crude was directly purified by silica column chromatography using Toluene:EtOAc (1:0 to 8:2 v/v) as the eluent. The obtained compound **26** was directly used in the next step. A mixture of acceptor **28** (100 mg, 0.3 mmol), donor **26** (468 mg, 0.46 mmol) and 4 Å molecular sieves was stirred in CH_3CN (5 mL) for 1 h. The mixture was cooled to -30°C and TMSOTf (11 µL, 0.06 mmol) was added. The reaction mixture was allowed to warm to 15°C over 2 h. The reaction mixture was

with Et_3N, filtered over a pad of Celite and concentrated *in vacuo*. The obtained residue was diluted with EtOAc, washed with $NaHCO_3$. The organic layer was dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained crude was purified by silica column chromatography using Toluene:EtOAc (1:0 to 7:3 v/v) as the eluent to afford compound **27** (305 mg, 84 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56 – 7.07 (40H, m, Ar-H), 5.45 (1H, s, CH-C_6H_5), 5.22 – 5.09 (3H, m), 4.93 – 4.67 (6H, m), 4.54 (1H, d), 4.50 – 4.41 (3H, m, H-1, Glc-I; CH_2, pentyl), 4.39 – 4.14 (4H, m, H-1, Gal-II), 4.01 (1H d, J = 3.5 Hz, H-4, Gal-II), 3.97 (1H, dd, J = 11.6, 7.1 Hz, H-5, Glc), 3.93 – 3.66 (5H, m, H-2, Glc-I), 3.66 – 3.58 (1H, m, H-3, Gal-II), 3.57 – 3.31 (4H, m, H-2, Gal-II; H-3, Glc-I; H-4, Glc-I), 3.28 – 3.10 (2H, m, CH_2, pentyl), 2.92 (1H, s, H-5, Gal-II), 1.74 – 1.42 (4H, m, 2x CH_2, pentyl), 1.41 – 1.17 (2H, m, CH_2, pentyl). ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.5 (C=O, Cbz), 138.9, 138.8, 138.7, 138.5, 138.3, 138.1, 129.0, 128.8, 128.6, 128.6, 128.5, 128.4, 128.4, 128.3, 128.3, 128.3, 128.2, 128.2, 128.2, 128.1, 128.0, 127.8, 127.8, 127.7, 127.6, 127.5, 127.5, 127.4, 126.6, 103.6 (C-1, Glc-I), 102.9 (C-1, Gal-II), 101.4 (CH-C_6H_5), 92.1, 83.0 (C-3, Gal-II), 81.8 (C-2, Gal-II), 79.6, 78.9 (C-2, Glc-I), 77.6, 7578, 75.3, 75.0, 75.0, 73.7 (C-4, Gal-II), 73.0, 71.7, 69.9, 69.0, 68.3, 67.2, 66.3 (C-5, Gal-II), 50.6, 50.3, 47.2, 46.3, 29.5, 28.0, 23.4.

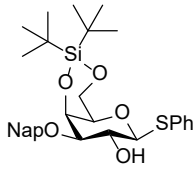
N-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (12b). Compound **27** (300 mg, 0.25 mmol)



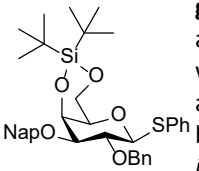
was stirred in DCM with 4 Å molecular sieves for 2 h. The mixture was cooled to -78°C and after adding Et_3SiH (201 µL, 1.26 mmol) stirring was continued for another 30 min. TfOH (45 µL, 0.50 mmol) was introduced. More Et_3SiH (200 µL) and TfOH (70 µL, 0.8 mmol) were added over time and after 2.5 h the reaction mixture was quenched with Et_3N. The quenched mixture was filtered over Celite, washed with H_2O, sat. aq. $NaHCO_3$, dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 5:1 v/v) as the eluent to afford compound **12b** (191 mg, 64 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.59 – 6.95 (40H, m, Ar-H), 5.15 (2H, d, J = 6.7 Hz), 4.97 (1H, d, J = 10.7 Hz), 4.89 – 4.61 (6H, m), 4.61 – 4.27 (8H, m, H-1 Glc-I; H-1, Gal-II), 4.01 (1H, s, H-4, Gal-II), 3.99 – 3.91 (1H, m), 3.91 – 3.28 (13H, m, H-2, Glc-I; H-2, Gal-II), 3.28 – 3.10 (2H, m, CH_2, pentyl), 2.40 (1H, s, OH), 1.71 – 1.42 (4H, m, 2x CH_2, pentyl), 1.42 – 1.17 (2H, m, CH_2, pentyl). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.1, 138.7, 138.6, 138.3, 138.2, 137.9, 128.5, 128.5, 128.4, 128.3, 128.0, 127.9, 127.8, 127.8, 127.8, 127.7, 127.6, 127.5, 127.5, 127.2, 103.6 (C-1, Glc-I), 102.5 (C-1, Gal-II), 82.9, 81.8 (C-2, Glc-I), 81.1, 79.4 (C-2, Gal-II), 76.6 (C-3, Gal-II), 75.3, 75.2, 75.1, 74.9, 73.5, 73.1, 72.7, 72.0, 68.4, 68.3, 67.1, 66.1 (C-4, Gal-II), 50.5, 50.2, 47.2, 46.2, 29.4, 28.0, 27.5, 23.4.



Scheme S5. Synthesis of galactosyl donor **11** from compound **29**.[38]

Phenyl 3-O-(2-naphthyl)methyl-4,6-O-di-tert-butylsilanediyl-1-thio-β-D-galacto-pyranoside (30). 

Bu_2SnO (5.78 g, 23 mmol) was added to compound **29**[38] (7.98 g, 19 mmol) in toluene (100 mL) and the suspension was heated under reflux for 3 h. The resulting clear solution was cooled to 90 °C and after the addition of 2-(bromomethyl)naphthalene (4.70 g, 21 mmol), tetrabutylammonium iodide (7.86 g, 21 mmol) was added portionwise over 1 h. After overnight stirring at 90 °C, the reaction mixture was concentrated *in vacuo*. The obtained crude was dissolved in DCM, washed with sat. aq. $NaHCO_3$, extracted with DCM (2x), washed with brine, dried (Na_2SO_4) and filtered over a pad of Celite. The concentrated crude was purified by silica column chromatography using Toluene:EtOAc (1:0 to 10:1 v/v) as the eluent to afford compound **30** (5.88 g, 55 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.88 – 7.75 (4H, m, Ar-H), 7.58 – 7.52 (3H, m, Ar-H), 7.51 – 7.43 (2H, m, Ar-H), 7.34 – 7.22 (3H, m, Ar-H), 4.96 (1H, d, J = 11.9 Hz, CHH, Nap), 4.81 (1H, d, J = 11.9 Hz, CHH, Nap), 4.61 – 4.54 (2H, m, H-1, H-4), 4.23 (2H, dd, J = 12.5, 1.9 Hz, H-6), 4.06 (1H, dd, J = 9.5, 1.9 Hz, H-2), 3.41 (1H, dd, J = 9.1, 3.0 Hz, H-3), 3.35 (1H, s, H-5), 2.60 (1H, d, J = 1.9 Hz, OH, H-2), 1.08 (9H, s, t-Bu), 1.07 (9H, s, t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 135.6, 133.60, 133.25, 133.09, 132.53, 128.85, 128.37, 127.87, 127.72, 127.67, 126.61, 126.19, 125.99, 125.85, 89.2 (C-1), 81.8 (C-3), 75.1 (C-5), 70.5 (CH_2, Nap), 69.5 (C-4), 68.6 (C-2), 67.4 (C-6), 27.7 (3x CH_3, t-Bu), 27.6 (3x CH_3, t-Bu), 23.4 (C, t-Bu), 20.7 (C, t-Bu). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{31}H_{40}O_5SSi$, 575,2263; found 575.2260. [α] 25/589 = -10.0° (C = 0.01; $CHCl_3$).

Phenyl 2-O-benzyl-3-O-(2-naphthyl)methyl-4,6-O-di-tert-butylsilanediyl-1-thio-β-D-galactopyranoside (11). 

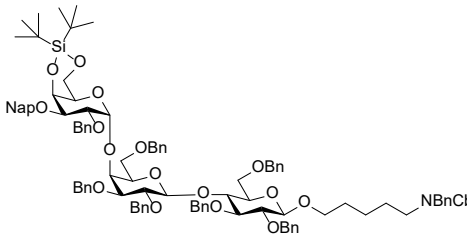
A mixture of compound **30** (1.9 g, 3.4 mmol) and NaH (60 % dispersion in oil; 275 mg, 6.9 mmol) in DMF (10 mL) was stirred at 0°C for 10 min. Benzyl bromide (612 µL, 5.2 mmol) was added dropwise and the mixture was stirred at RT for another 30 min before quenching with AcOH in MeOH. The mixture was concentrated *in vacuo*, diluted with DCM and washed with 1M HCl and water. The organic phase was dried (Na_2SO_4), filtered and concentrated *in vacuo*. Crystallization from MeOH provided compound **11** (1.39 g, 63 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.86 – 7.77 (3H, m, Ar-H), 7.76 – 7.69 (1H, m, Ar-H), 7.57 – 7.50 (3H, m, Ar-H), 7.50 – 7.40 (3H, m, Ar-H), 7.38 – 7.18 (7H, m, Ar-H), 4.98 – 4.82 (4H, m, 2x CH_2, Bn, Nap), 4.66 (1H, d, J = 9.8 Hz, H-1), 4.52 (1H, d, J = 2.7 Hz, H-4), 4.18 (2H, dd, J = 12.4, 1.9 Hz, H-6), 3.88 (3H, t, J = 9.4 Hz, H-2), 3.52 (1H, dd, J = 9.1, 3.0 Hz, H-3), 3.26 (1H, s, H-5), 1.15 (9H, s, t-Bu), 1.09 (9H, s, t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 138.37, 135.83, 134.82, 133.21, 133.01, 132.02, 128.70, 128.41, 128.29, 128.19, 127.84, 127.69, 127.66, 127.22, 126.44, 126.07, 125.86, 88.7 (C-1), 82.6 (C-3) 77.3 (C-2), 75.9, 74.7 (C-5), 71.1, 70.1 (C-4), 67.3 (C-6), 27.7 (CH_3, t-Bu), 27.6 (CH_3, t-Bu), 23.4 (C, t-Bu), 20.7 (C, t-Bu). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{38}H_{46}O_5SSi$, 665.2733; found 665.2731. [α] 25/589 = 4.5° (C = 0.1; $CHCl_3$).

Para-methoxyphenyl 2-O-benzyl-3-O-(2-naphthyl)methyl-4,6-O-di-tert-butylsilanediyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-gluco-pyranoside (15a). A mixture of acceptor **12a** (1.54 g, 1.56 mmol), donor **11** (1.20 g, 1.87 mmol) and 4 Å molecular sieves was stirred in DCM (15 mL) for 1 h. The reaction mixture was cooled to -30°C and N-iodosuccinimide (700 mg, 3.1 mmol) and triflic acid (14 µL, 0.16 mmol) were added. After 35 min the reaction was quenched with Et_3N, filtered over a pad of Celite and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 20:1 v/v) as the eluent to afford compound **15a** (2.13 g, 90 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.84 – 7.64 (4H, m, Ar-H), 7.49 – 7.39 (4H, m, Ar-H), 7.36 – 7.13 (34H, m, Ar-H), 7.01 (2H, d, J = 9.0 Hz, OMP), 6.78 (2H, d, J = 9.0 Hz, OMP), 5.12 – 5.07 (1H, d, J = 11.3 Hz, CHH), 5.01 – 4.92 (2H, m, H-1, Gal-III; CHH), 4.85 (1H, d, J = 7.3 Hz, H-1, Glc-I), 4.83– 4.64 (8H, m, 4x CH_2), 4.62 – 4.52 (2H, m, 2x CHH), 4.52 – 4.41 (3H, m, CHH; H-1, Gal-II; H-4, Gal-III), 4.40 – 4.24 (3H, m, CH_2; CHH), 4.17 – 3.87 (6H, m, H-6a, Gal-II; H-4, Glc-I; H-2, Gal-III, H-5, Gal-II; H-4, Gal-II; H-3, Gal-III) 3.83 – 3.78 (4H, m, H-6, Glc-I; H-6, Gal-III), 3.76 (3H, s, CH_3, OMP), 3.70 – 3.60 (2H, m, H-2, Glc-I; H-5, Gal-III), 3.59 – 3.43 (3H, m, H-2, Gal-II; H-6b, Gal-II; H-5, Glc-I), 3.35 (1H, dd, J = 8.1, 5.5 Hz, H-3, Glc-I), 3.29 (1H, dd, J = 10.0, 2.7 Hz, H-3, Gal-II), 1.01 (9H, s, t-Bu), 0.99 (9H, s, t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.2 (C, OMP), 151.6 (C, OMP), 139.2 (C, Ar), 138.7 (C, Ar), 138.5 (C, Ar), 138.4 (C, Ar), 138.3 (C, Ar), 138.2 (C, Ar), 136.7 (C, Ar), 133.2, 132.9, 129.0, 128.5, 128.4, 128.4, 128.3, 128.2, 128.2, 128.2, 128.1, 127.9, 127.8, 127.7, 127.7, 127.6, 127.6, 127.6, 127.5, 127.4, 127.4, 127.4, 127.4, 126.0, 125.9, 125.9, 125.7, 118.5 (2x CH, OMP), 114.5 (2x CH, OMP), 103.1 (C-1, Glc-I), 102.7 (C-1, Gal-II), 100.2 (C-1, Gal-III), 82.5 (C-5, Gal-III), 81.5 (C-2, Glc-I), 81.1 (C-3, Gal-II), 79.1 (C-2, Gal-II), 78.0 (C-3, Gal-III), 77.2 (C-5, Gal-II), 75.3 (C-5, Glc-I), 75.2, 75.1, 74.9, 74.3 (C-2, Gal-III), 73.7, 73.5 (C-4, Glc-I), 73.4 (C-3, Glc-I), 73.2, 73.1, 72.2, 71.2 (C-4, Gal-III), 70.6, 68.4 (C-6, Glc-I), 67.7 (C-6, Gal-II), 67.5 (C-4, Gal-II), 67.1 (C-6, Gal-III), 55.6 (CH_3, OMP), 27.7 (CH_3, t-Bu), 27.3 (CH_3, t-Bu), 23.3 (C, t-Bu), 20.7 (C, t-Bu). ESI HRMS (m/z): $[M + NH_4]^+$ calcd for $C_{93}H_{104}O_{17}Si$, 1538.7381; found 1538.7403. [α] 25/589 = 357.2° (C = 0.1; $CHCl_3$).

Para-methoxyphenyl 2-O-benzyl-4,6-O-di-tert-butylsilanediyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (8a). DDQ (377 mg, 1.7 mmol) was added to a stirring solution of compound **15a** (2.1 g, 1.4 mmol) in DCM (120 mL) and PBS buffer (pH 7.4, 5 mL) and the reaction mixture was kept in darkness. After 3 h, the mixture was diluted with DCM, washed with sat. aq. $NaHCO_3$ (2 x), H_2O, dried (Na_2SO_4), filtered and concentrated *in vacuo*. Purification by silica column chromatography using Toluene:EtOAc (1:0 to 4:1 v/v) as the eluent provided compound **8a** (987 mg, 52 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 (2H, d, J = 7.0

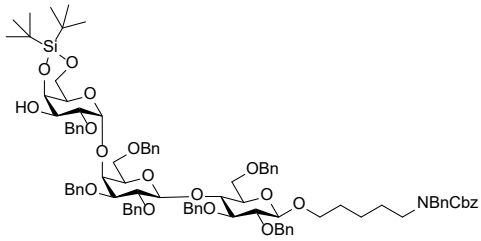
Hz, Ar-H), 7.36 – 7.17 (33H, m, Ar-H), 7.01 (2H, d, J = 9.0 Hz, OMP), 6.78 (2H, d, J = 9.1 Hz, OMP), 5.09 (1H, d, J = 11.6 Hz, CHH), 5.01 – 4.94 (2H, m, H-1, Gal-III; CHH), 4.86 (1H, d, J = 7.5 Hz, H-1, Glc-I), 4.81 – 4.56 (8H, m, 4x CH_2), 4.51 – 4.42 (2H, m, H-1, Gal-II; CHH), 4.40 – 4.27 (3H, m, CH_2; CHH), 4.24 (1H, d, J = 3.2 Hz, H-4, Gal-III), 4.09 – 3.95 (5H, m, H-4, Gal-II; H-4, Glc-I; H-6a, Gal-II; H-3, Gal-III; H-5, Gal-II), 3.79 (2H, d, J = 3.1 Hz, H-6, Glc-I), 3.76 (3H, s, CH_3, OMP), 3.72 (2H, s, H-6, Gal-III), 3.70 – 3.61 (3H, m, H-2, Glc-I; H-5, Gal-III; H-2, Gal-III), 3.56 (1H, dd, J = 9.9, 7.7 Hz, H-2, Gal-II), 3.51 – 3.43 (2H, m, H-6b, Gal-II; H-5, Glc-I), 3.34 (1H, dd, J = 8.3, 5.5 Hz, H-3, Glc-I), 3.29 (1H, dd, J = 10.0, 2.7 Hz, H-3, Gal-II), 2.40 – 2.31 (1H, m, OH), 0.97 (9H, s, t-Bu), 0.90 (9H, s, t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.2 (C, OMP), 151.6 (C, OMP), 139.3 (C, Ar), 138.4 (C, Ar), 138.4 (C, Ar), 138.4 (2 C, Ar), 138.2 (C, Ar), 138.1 (C, Ar), 128.4, 128.3, 128.3, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.7, 127.7, 127.6, 127.6, 127.6, 127.5, 127.4, 127.3, 127.1, 118.4 (2x CH, OMP), 114.5 (2x CH, OMP), 103.1 (C-1, Gal-II), 102.7 (C-1, Glc-I), 99.4 (C-1, Gal-III), 82.7 (C-5, Gal-III), 81.6 (C-2, Glc-I), 81.1 (C-3, Gal-II), 79.0 (C-2, Gal-II), 77.3 (C-5, Gal-II), 75.5 (C-2, Gal-III), 75.3 (C-5, Glc-I), 75.1, 75.0, 75.0, 73.9 (C-4, Gal-III), 73.2 (C-4, Glc-I), 73.1 (C-3, Glc-I), 73.0, 72.3, 70.1 (C-3, Gal-III), 68.4 (C-6, Glc-I), 67.5 (C-6, Gal, II), 67.1 (C-4, Gal-II), 66.7 (C-6, Gal-III), 55.6 (CH_3, OMP), 27.5 (CH_3, t-Bu), 27.2 (CH_3, t-Bu), 23.2 (C, t-Bu), 20.6 (C, t-Bu). ESI HRMS (m/z): $[M + NH_4]^+$ calcd for $C_{82}H_{96}O_{17}Si$, 1398.6755; found 1398.6760. [α] 25/589 = 54.2° (C = 0.05; $CHCl_3$).

N-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2-O-benzyl-3-O-(2-naphthyl)methyl-4,6-O-di-tert-butylsilanediyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (15b).



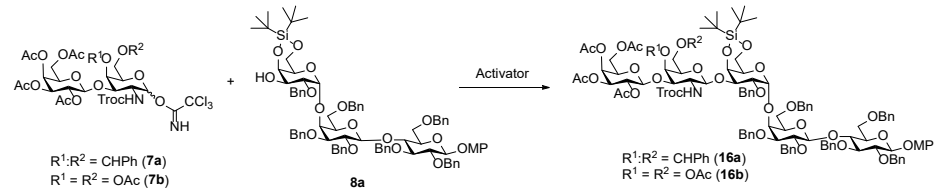
A mixture of acceptor **12b** (189 mg, 0.19 mmol), donor 11 (147 mg, 0.23 mmol) and 4 Å molecular sieves was stirred in DCM (1.5 mL) for 1 h. The reaction mixture was cooled to -30°C and N-iodosuccinimide (86 mg, 0.38 mmol) and triflic acid (1.7 µL, 0.02 mmol) were added. After 20 min the reaction was quenched with Et_3N, filtered over a pad of Celite and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 5:1 v/v) as the eluent to afford compound **15b** (197 mg, 60%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.95 – 7.64 (m, 4H, Ar-H), 7.54 – 6.79 (48H, m, Ar-H), 5.15 (2H, d, J = 6.3 Hz), 5.06 (1H, d, J = 11.1 Hz), 4.97 (1H, d, J = 3.4 Hz, H-1, Gal-III), 4.87 – 4.62 (7H, m), 4.62 – 4.38 (7H, m, H-1, Gal-II), 4.37 – 4.17 (4H, m, H-1, Glc-I), 4.16 – 4.07 (1H, m), 4.06 – 3.98 (2H, m, H-2, Gal-III), 3.98 – 3.63 (10H, m), 3.62 – 3.07 (10H, m, H-2, Gal-II, H-2, Glc-I), 1.68 – 1.40 (4H, m, 2x, CH_2, pentyl), 1.40 – 1.16 (2H, m, CH_2, pentyl), 1.00 (9H, s, t-Bu), 0.97 (9H, s, t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.3, 138.7, 138.7, 138.5, 138.3, 138.2, 136.7, 133.2, 132.9, 129.0, 128.5, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 127.9, 127.8, 127.7, 127.6, 127.6, 127.4, 127.4, 127.3, 126.0, 125.9, 125.9, 125.6, 125.3, 103.5 (C-1, Glc-I), 102.9 (C-1, Gal-II), 100.1 (C-1, Gal-III), 82.5, 81.7 (C-2, Glc-I), 81.1, 79.0 (C-2, Gal-II), 78.0, 77.2, 75.1, 75.0, 74.9, 74.9, 74.2, 73.6 (C-2, Gal-III), 73.1, 73.1, 72.1, 71.2, 70.6, 70.1, 69.7, 68.3, 67.6, 67.5, 67.1, 67.1, 50.2, 47.2, 46.2, 29.4, 27.6 (CH_3, t-Bu), 27.6, 27.3 (CH_3, t-Bu), 23.3 (C, t-Bu), 20.7 (C, t-Bu). [α] 25/589 = -23.5° (C = 0.02; $CHCl_3$).

N-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2-O-benzyl-4,6-O-di-tert-butylsilanediyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (8b).



β-Pinene (71 µL, 0.45 mmol) and DDQ (51 mg, 0.22 mmol) were added to a stirring solution of compound **15b** (2.1 g, 1.4 mmol) in DCM/H_2O (9/1 mL). The reaction mixture was kept in darkness and stirred overnight. The mixture was diluted with DCM, washed with sat. aq. $NaHCO_3$ (2 x), dried (Na_2SO_4), filtered and concentrated *in vacuo*. Purification by silica column chromatography using Toluene:EtOAc (1:0 to 10:1.2 v/v) as the eluent provided compound **8b** (105 mg, 59 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.09 (45H, m, Ar-H), 5.15 (2H, d, J = 7.0 Hz, CH_2), 5.05 (1H, d, J = 11.6 Hz, CHH), 4.98 (1H, d, J = 3.2 Hz, H-1, Gal-III), 4.89 – 4.80 (1H, m, CHH), 4.80 – 4.26 (16H, m, H-1, Gal-II; H-1, Glc-I), 4.23 (1H, d, J = 2.8 Hz, H-4, Gal-III), 4.09 – 3.68 (10H, m, H-3, Gal-III), 3.65 (1H, dd, J = 9.9, 3.2 Hz, H-2, Gal-III), 3.61 – 3.09 (10H, m, H-2, Gal-II; H-2, Glc-I), 2.39 – 2.31 (1H, m, OH), 1.78 – 1.42 (4H, m, 2x CH_2, pentyl), 1.38 – 1.18 (2H, m, CH_2, pentyl), 0.97 (9H, s, t-Bu), 0.90 (9H, s, t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.4, 138.6, 138.4, 138.4, 138.2, 128.5, 128.4, 128.3, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 127.4, 127.3, 127.1, 103.5 (C-1, Glc-I), 103.0 (C-1, Gal-II), 99.4 (C-1, Gal-III), 82.6, 81.7 (C-2, Glc-I), 81.0, 78.9 (C-2, Gal-II), 77.2, 75.5 (C-2, Gal-III), 75.1, 75.0, 74.9, 73.9 (C-4, Gal-III), 73.2, 73.1, 73.1, 73.0, 72.2, 70.1 (C-3, Gal-III), 68.4, 67.5, 67.1, 66.7, 29.4, 27.5 (CH_3, t-Bu), 27.2 (CH_3, t-Bu), 23.4, 23.2 (C, t-Bu), 20.6 (C, t-Bu). [α] 25/589 = -53° (C = 0.01; $CHCl_3$).

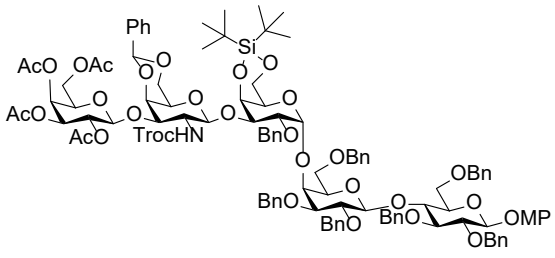
Table S1. Glycosylation conditions of trisaccharide acceptor **8a** and disaccharide donor **7a** or **7b** to afford protected Gb5 (**16a**, **16b**).



Donor	Activator	Temperature	Total product	Isolated β-product
7a	TMSOTf (0.2eq)	-30 °C	56-59%	37-42%
7a	TMSOTf (0.2-0.3eq)	-60 °C	72-76%	42-45%
7a	TfOH (0.1eq)	-10 °C	32%	32%
7a	TfOH (0.1eq)	-50 °C	43%	24-28%
7b	TfOH (0.1eq)	-10 °C	19-44%	19-44%
7b	TfOH (0.1eq)	-50 °C	52%	52%

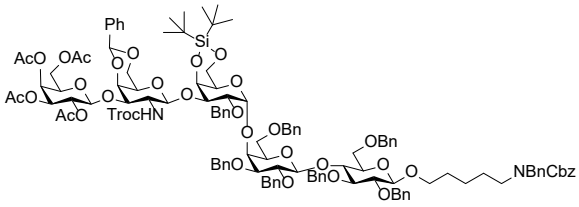
The highest overall yield of the glycosylation was obtained at the coldest activation temperature. However, changes in temperature or activator did not improve β/α selectivities when donor **7a** was used. The glycosylation proved most successful with donor **7b** and acceptor **8a** in DCM with molecular sieves at -50 °C, since no α-product was formed.

***Para*-methoxyphenyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-4,6-O-benzylidene-2-[(2,2,2-trichloromethoxy)carbonylamino]-β-D-galactopyranosyl-(1→3)-2-O-benzyl-4,6-O-di-tert-butyl-silanediyl-α-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (16a).** A



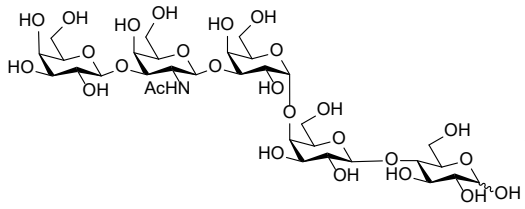
mixture of acceptor **8a** (987 mg, 0.71 mmol), donor **7a** (845 mg, 0.92 mmol) and 4 Å molecular sieves was stirred in DCM (9 mL) for 1 h. The mixture was cooled to -35°C and TMSOTf (26 µL, 0.14 mmol) was added. After 5 min the reaction was quenched by addition of Et_3N. The mixture was filtered over Celite and concentrated *in vacuo.* The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 8:2 v/v) as the eluent to give compound 16a as an oil. The β-anomer of the title pentasaccharide (562 mg, 37 %). (total isolated α/β yield: 59%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 – 7.09 (40H, m, H-Ar), 7.00 (2H, d, J = 8.8 Hz, OMP), 6.78 (2H, d, J = 8.8 Hz, OMP), 5.47 (1H, s, CH-C_6H_5), 5.36 (1H, s), 5.16 (1H, dd), 5.09 – 4.94 (3H, m), 4.93 – 3.95 (32H, m, 5x H-1), 3.95 – 3.70 (10H, m), 3.70 – 3.55 (3H, m), 3.53 – 3.42 (2H, m), 3.37 – 3.27 (2H, m), 2.90 (1H, s), 2.14 (3H, s, OAc), 2.06 (3H, s, OAc), 2.01 (3H, s, OAc), 1.95 (3H, s, OAc), 1.03 – 0.85 (18H, m, 2x t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.3 (C, OAc), 170.1 (C, OAc), 169.4 (C, OAc), 155.2 (C, OMP), 153.7 (C=O, Troc), 151.5 (C, OMP), 139.4 (C, Ar), 138.5 (C, Ar), 138.4 (C, Ar), 138.3 (C, Ar), 138.1 (C, Ar), 138.0 (C, Ar), 137.9 (C, Ar), 129.0, 129.0, 128.5, 128.4, 128.3, 128.2, 128.1, 128.0, 127.8, 127.7, 127.7, 127.6, 127.5, 127.5, 127.4, 126.5, 125.3, 118.4 (2x CH, OMP), 114.5 (2x CH, OMP), 103.1 (C-1, Gal-II), 102.7 (C-1, Glc-I), 101.5 (C-1, Gal-V), 101.4 (C-1, GalNAc-IV), 100.8 (C<u>H</u>-C_6H_5), 99.9 (C-1, Gal-III), 95.3 (CCl_3), 81.6, 81.2, 79.3, 78.9, 77.2, 75.9, 75.5, 75.3, 75.1, 74.9, 74.3, 74.2, 73.8, 73.5, 73.3, 73.1, 73.0, 72.1, 70.8, 68.9, 68.7, 68.4, 67.8, 67.5, 67.0, 66.0, 61.4, 55.6 (CH_3, OMP), 53.7, 27.5 (CH_3, t-Bu), 27.4 (CH_3, t-Bu), 23.30, 20.8 (CH_3, OAc), 20.7 (CH_3, OAc), 20.7 (CH_3, OAc), 20.5 (CH_3, OAc). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{112}H_{130}Cl_3NO_{32}Si$ 2156.7309; 2156.7454 found. [α] 25/589 = 55.0° (C = 0.05; $CHCl_3$).

N-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-4,6-O-benzylidene-2-[(2,2,2-trichloromethoxy) carbonylamino]-β-D-galactopyranosyl-(1→3)-2-O-benzyl-4,6-O-di-tert-butylsilanediyl-α-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-O-benzyl-β-D-glucopyranoside (16c). A mixture of acceptor **8b** (100 mg, 0.063 mmol), donor **7a** (75 mg, 0.082 mmol) and 4 Å molecular sieves was stirred in DCM (9 mL) for 1 h. The mixture was cooled to -30°C and TMSOTf (2 µL, 0.013 mmol) was added. After 20 min the reaction was quenched by addition of Et_3N. The mixture was filtered over Celite and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 8:2 v/v) as the



eluent to give compound 16c as an oil. The β-anomer of the title pentasaccharide (62 mg, 42%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.61 – 6.95 (50H, m, Ar-H), 5.47 (1H, s, CH-C_6H_5), 5.37 (1H, d, J = 3.1 Hz), 5.20 – 5.11 (3H, m), 5.00 (1H, d, J = 11.2 Hz), 4.94 – 4.41 (19H, m, H-1, Gal-III; H-1, Gal-V; H-1, GalNAc-IV; H-1, Gal-II), 4.40 – 3.63 (24H, m, H-1, Glc-I), 3.62 – 3.08 (10H, m), 2.91 (1H, s), 2.15 (3H, s, OAc), 2.06 (3H, s, OAc), 2.01 (3H, s, OAc), 1.96 (3H, s, OAc), 1.67 – 1.43 (4H, m, 2x CH_2, pentyl), 1.39 – 1.19 (2H, m, CH_2, pentyl), 0.97 – 0.92 (18H, m, 2x t-Bu). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.3, 170.1, 169.3, 153.6, 139.5, 139.5, 138.6, 138.5, 138.4, 138.3, 138.1, 138.1, 138.0, 137.9, 137.8, 133.8, 129.1, 129.0, 128.5, 128.5, 128.4, 128.3, 128.2, 128.2, 128.0, 127.9, 127.8, 127.7, 127.6, 127.5, 127.4, 126.5, 125.5, 125.3, 111.9, 103.5 (C-1, Glc-I), 102.9 (C-1, Gal-II), 101.5 (C-1, Gal-V), 101.3 (C-1, GalNAc-IV), 100.7 (CH-C_6H_5), 99.5 (C-1, Gal-III), 95.3, 88.0, 81.7, 81.6, 81.2, 79.3, 78.7, 77.2, 76.8, 76.2, 75.9, 75.5, 75.0, 74.9, 74.2, 74.2, 73.9, 73.4, 73.2, 73.1, 72.1, 70.8, 69.7, 68.9, 68.6, 68.4, 67.7, 67.5, 67.1, 67.0, 67.0, 66.0, 61.4, 57.4, 53.8, 53.4, 50.5, 50.2, 47.1, 46.2, 29.7, 29.4, 27.5 (CH_3, t-Bu), 27.4 (CH_3, t-Bu), 23.4, 23.3 (C, t-Bu), 20.8 (C, t-Bu), 20.8 (CH_3, OAc), 20.7 (CH_3, OAc), 20.6 (CH_3, OAc), 20.5 (CH_3, OAc). [α] 25/589 = -84° (C = 0.01; $CHCl_3$).

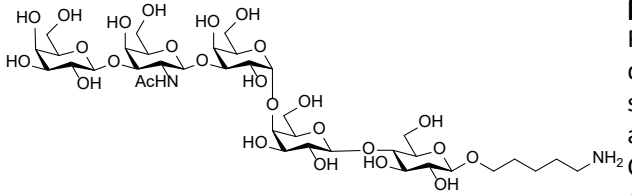
β-D-Galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(14)-β-D-galactopyranosyl-(1→4)-α/β-D-glucopyranose (4a).



Pentasaccharide **16a** was deprotected in a total of six steps, all steps were monitored by TLC and MALDI-TOF-MS. HF·Pyridine (300 µL of 70%) was added to a mixture of compound **16a** (590 mg, 0.28 mmol) in pyridine (6 mL). The mixture was stirred at RT overnight, diluted with EtOAc, washed with sat. aq. $NaHCO_3$ (3 x), dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained crude was co-evaporated with toluene (3 x) and dissolved in THF (10 mL). To this mixture 1 M NaOH (10 mL) was added and after heating to refluxed (80 °C) overnight, the mixture was concentrated *in vacuo* and co-evaporated with toluene (2 x). The resulting intermediate was dissolved in pyridine (10 mL) and Ac_2O (8 mL) was added. The mixture was stirred at RT for 6 h, diluted with EA, washed with 1 M HCl, H_2O, sat. aq. $NaHCO_3$ (4 x), dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 1:4 v/v) as the eluent to obtain the intermediate product (417 mg, 76 % over 3 steps). Ammonium cerium(IV) nitrate (587 mg, 1.07 mmol) was added to a solution of this intermediate in CH_3CN (10 mL)/H_2O (2.5 mL) at 0 °C. After 7 min the mixture was quenched by addition of sat. aq. $NaHCO_3$. The layers were separated and the organic layer was washed with sat. aq. $NaHCO_3$ (2x), H_2O, dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained crude was dissolved in MeOH (3 mL), freshly prepared NaOMe was added and the mixture was stirred at r.t for 2 h. The mixture was neutralized with Dowex H^+ resin, filtered, concentrated *in vacuo*. The crude intermediate was dissolved in a mixture of MeOH/H_2O/HOAc (3/3/1 mL), followed by the addition of $Pd(OH)_2$/C (560 mg, 20 %, Degussa type) and the reaction mixture was left stirring overnight under the atmosphere of hydrogen. The mixture was filtered over a pad of Celite, concentrated *in vacuo* and purified by Bio-Gel P-2 size exclusion chromatography

to give the title compound as a white amorphous solid (131 mg, 70 %, over three steps) Additional purification by HPLC with a semi-preparative HILIC column (XBridge® Amide 5 µm, 4.6 mm x 250 mm column, Waters) under isocratic conditions (74% B) with UV detection (210 nm) affords analytically pure glycan. ^{1}H NMR (400 MHz, D_2O) δ 5.21 (0.5H, d, J = 3.7 Hz, H-1α, Glc-I), 4.90 (1H, d, J = 3.8 Hz, H-1, Gal-III), 4.70 – 4.61 (1.5H, m, H-1, GalNAc-IV; H-1β Glc-I), 4.50 (1H, d, J = 7.7 Hz, H-1, Gal-II), 4.45 (1H, d, J = 7.7 Hz, H-1, Gal-V), 4.38 (1H, t, J = 6.3 Hz, H-5, Gal-III), 4.24 (1H, d, J = 2.3 Hz, H-4, Gal-III), 4.17 (1H, d, J = 2.9 Hz, H-4, GalNAc-IV), 4.10 – 4.01 (2H, m, H-2, GalNAc-IV; H-4, Gal-II), 3.99 – 3.54 (23H, m, H-2, Gal-III; H-2, Gal-II), 3.50 (1H, dd, J = 9.8, 7.8 Hz, H-2, Gal-V), 3.27 (0.5H, t, J = 8.4 Hz, H-2, Glc-Iβ), 2.02 (3H, s, NHAc). ^{13}C NMR (101 MHz, D_2O) δ 175.0 (C=O NHAc), 104.7 (C-1, Gal-V), 103.2 (C-1, Gal-II), 102.9 (C-1, GalNAc-IV), 100.3 (C-1, Gal-III), 95.6 (C-1β, Glc-I), 92.1 (C-1α, Glc-I), 79.5, 78.6, 78.6, 77.1, 75.4, 74.8, 74.5, 74.4, 73.8 (C-2, Glc-I), 72.4, 72.0, 71.4, 71.1 (C-2, Gal-III), 70.81 (C-2, Gal-II), 70.51 (C-2, Gal-V), 70.2, 70.1, 68.9, 68.5, 67.9, 67.5, 60.9, 60.9, 60.3, 60.3, 60.3, 51.4 (C-2, GalNAc-IV), 22.2 (CH_3, NHAc). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{32}H_{55}NO_{26}$, 892.2910; found 892.2912.

5-aminopentyl β-D-Galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (4b).



Pentasaccharide **16c** was deprotected in a total of 5 steps, all steps were the same as for Gb5-OMP, without the CAN reaction. (4.57 mg, 19 %, over 5 steps). ^{1}H NMR (400 MHz, D_2O) δ 4.89 (1H, d, J = 3.5 Hz, H-1, Gal-III), 4.67 (1H, d, J = 8.5 Hz, H-1, GalNAc-IV), 4.54 – 4.40 (3H, m, H-1, Gal-II; H-1, Glc-I; H-1, Gal-V), 4.37 (1H, t, J = 6.3 Hz, H-5, Gal-III), 4.23 (1H, s, H-4, Gal-III) 4.16 (1H, d, J = 2.4 Hz, H-4, GalNAc-IV), 4.10 – 3.46 (28H, m), 3.28 (1H, t, J = 8.2 Hz, H-2, Glc-I), 3.03 – 2.94 (2H, m, CH_2, pentyl), 2.01 (3H, s, NHAc), 1.76 – 1.58 (4H, m, 2x CH_2, pentyl), 1.50 – 1.38 (2H, m, CH_2 pentyl). ^{13}C NMR (101 MHz, D_2O) δ 175.0 (C=O, NHAc), 104.7 (C-1, Gal-V), 103.2 (C-1, Gal-II), 102.8 (C-1, GalNAc-IV), 101.9 (C-1, Glc-I), 100.3 (C-1, Gal-III), 79.5, 78.7, 78.6, 77.1, 75.3, 74.9, 74.7, 74.5, 74.4, 72.8 (C-2, Glc-I), 72.3, 72.0, 70.8 (C-2, Gal-II), 70.5 (C-2, Gal-V), 70.2, 70.0, 68.8 (C-2, Gal-III), 68.5, 67.9, 67.5, 60.9, 60.8, 60.2, 60.2, 59.9, 51.4 (C-2, GalNAc-IV), 39.2 (CH_2), 28.1 (CH_2), 26.3 (CH_2), 22.2 (CH_3, NHAc), 22.0 (CH_2). ESI HRMS (m/z): $[M + H]^+$ calcd for $C_{37}H_{66}N_2O_{26}$, 955.3977; found 955.3979.

Enzymatic synthesis

Human glycosyltransferase expression

The catalytic domains of human glycosyltransferases (see Table S2 below) were expressed as soluble, secreted fusion proteins by transient transfection of HEK293 suspension cultures.[14, 39] The coding regions were amplified from Mammalian Gene Collection clones using primers that appended a tobacco etch virus (TEV) protease cleavage site[14, 40] to the NH_2-terminal end of the coding region and attL1 and attL2 Gateway adaptor sites to the 5' and 3' terminal ends of the amplimer products. The amplimers were

recombined via BP clonase reaction into the pDONR221 vector and the DNA sequences were confirmed. The pDONR221 clone was then recombined via LR clonase reaction into a custom Gateway adapted version of the pGEn2 mammalian expression vector[14, 39, 41] to assemble a recombinant coding region comprised of a 25 amino acid NH_2-terminal signal sequence from the *T. cruzi* lysosomal α-mannosidase[42] followed by an 8xHis tag, 17 amino acid AviTag,[43] "superfolder" GFP,[44] the nine amino acid sequence encoded by attB1 recombination site, followed by the TEV protease cleavage site and the respective glycosyltransferase catalytic domain coding region.

Suspension culture HEK293 cells (Freestyle 293-F cells, Life Technologies, Grand Island, NY) were transfected as previously described[14, 39] and the culture supernatant was subjected to Ni-NTA superflow chromatography (Qiagen, Valencia, CA). Enzyme preparations eluted with 300 mM imidazole were concentrated to ~1 mg mL^{-1} using an ultrafiltration pressure cell membrane (Millipore, Billerica, MA) with a 10 kDa molecular weight cutoff.

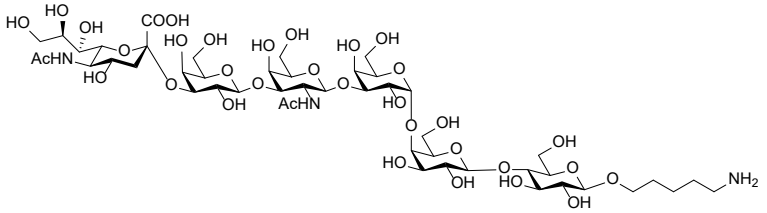
Table S2. Enzyme expression details.[14]

Enzyme	Amino Acid Residues	Uniprot ID
ST3GAL1	52 - 340	Q11201
ST6GALNAC5	50 - 336	Q9BVH7
ST6GALNAC6	31 - 333	Q969X2

Experimental procedures for enzymatic synthesis

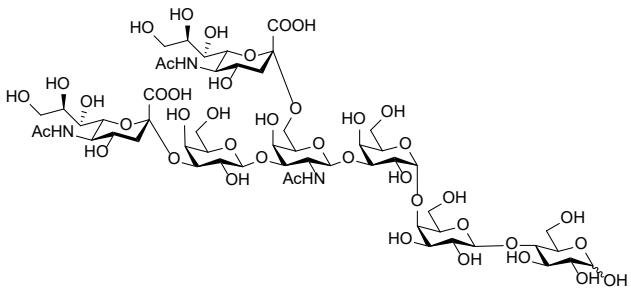
αNeu5Ac-(2→3)-β-D-Galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(14)-α/β-D-glucopyranose (5a). ST3Gal1 and CIAP were added to compound **4a** (6.5 mg, 10 mM final concentration) in H_2O with CMP-Neu5Ac (15 mM), $MgCl_2$ (20 mM), sodium cacodylate buffer (50 mM, pH 7.5). The mixture was shaken at 37°C for 94 h and monitored by TLC (EA:MeOH:H_2O:HOAc 4:3:2:1). More enzymes were added until no more starting material could be observed. Purification by Bio-Gel P-2 size exclusion chromatography and semi-preprarative HILIC column (XBridge® Amide 5 µm, 4.6 mm x 250 mm column, Waters, 70% B isocratic) provided compound **5a** (4.57 mg, 53 %). ^{1}H NMR (750 MHz, D_2O) δ 5.24 (0.5H, d, J = 3.7 Hz, H-1α, Glc-I), 4.93 (1H, d, J = 3.9 Hz, H-1, Gal-III), 4.70 (1H, d, J = 8.5 Hz, H-1, GalNAc-IV), 4.68 (0.5H, d, J = 8.0 Hz, H-1β, Glc-I), 4.55 – 4.51 (2H, m, H-1, Gal-V; H-1, Gal-II), 4.41 – 4.37 (1H, m, H-5, Gal-III), 4.26 (1H, s, H-4, Gal-III), 4.19 (1H, d, J = 3.0 Hz, H-4, GalNAc-IV), 4.08 (2H, dd, J = 9.9, 3.2 Hz, H-2, GalNAc-IV; H-3, Gal-V), 4.05 (1H, d, J = 2.9 Hz, H-4, Gal-II), 4.01 – 3.57 (29.5H, m), 3.55 (1H, dd, H-2, Gal-V), 3.29 (0.5H, t, J = 8.6 Hz, H-2, Glc-I), 2.76 (1H, dd, J = 12.4, 4.6 Hz, H-3eq, Neu5Ac-VI), 2.04 (6H, s, 2x NHAc), 1.79 (1H, t, J = 12.1 Hz, H-3ax, Neu5Ac-VI). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{43}H_{72}N_2O_{34}$, 1183.3864; found 1183.3861.

5-aminopentyl αNeu5Ac-(2→3)-β-D-Galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galacto-pyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranose (5b). ST3Gal1 and CIAP were added to compound **4b** (2.0



mg, 10 mM final concentration) in H_2O with CMP-Neu5Ac (15 mM), $MgCl_2$ (20 mM), sodium cacodylate buffer (50 mM, pH 7.5). The mixture was shaken at 37°C for 24 h and monotored by TLC (EA:MeOH:H_2O:HOAc 4:4:3.3:2). Purification by Bio-Gel P-2 size exclusion chromatography provided compound **5b** (1.8 mg, 69%). ^{1}H NMR (600 MHz, D_2O) δ 4.92 (1H, d, J = 3.9 Hz, H-1, Gal-III), 4.71 (1H, d, J = 8.5 Hz, H-1, GalNAc-IV), 4.56 – 4.49 (3H, m, H-1 Gal-V; H-1, Gal-II; H-1, Glc-I), 4.39 (1H, t, J = 6.4 Hz, H-5, Gal-III), 4.26 (1H, d, J = 2.4 Hz, H-4, Gal-III), 4.19 (1H, d, J = 2.8 Hz, H-4, GalNAc-IV), 4.13 – 3.48 (34H, m), 3.36 – 3.27 (1H, m, H-2, Glc-I), 3.02 (2H, t, J = 7.5 Hz, CH_2), 2.76 (2H, dd, J = 12.4, 4.6 Hz, H-3eq, Neu5Ac-VI), 2.04 (6H, s, 2x NHAc), 1.79 (1H, t, J = 12.1 Hz, H-3ax, Neu5Ac-VI), 1.75 – 1.65 (4H, m, 2x CH_2, pentyl), 1.52 – 1.43 (2H, m, CH_2, pentyl). ESI HRMS (m/z): $[M + H]^+$ calcd for $C_{48}H_{83}N_3O_{34}$, 1246.4858; found 1246.4970.

αNeu5Ac-(2→3)-β-D-Galactopyranosyl-(1→3)-[αNeu5Ac-(2→6)]-2-acetamido-2-deoxy-β-D-galacto-pyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-α/β-D-glucopyranose (6a). ST6GalNAc5 and CIAP were added to a mixture



of compound **5a** (4.2 mg, 10 mM final concentration) in H_2O with CMP-Neu5Ac (15 mM), $MgCl_2$ (20 mM), sodium cacodylate buffer (50 mM, pH 7.5). The mixture was shaken at 37°C overnight and monitored by TLC (EA:MeOH:H_2O:HOAc 3:3:3:2). More enzymes were added until no more starting material could be observed. Purification by Bio-Gel P-2 size exclusion chromatography and semi-preprarative HILIC column (XBridge® Amide 5 µm, 4.6 mm x 250 mm column, Waters, 70% B isocratic) provided compound **6a** (2.64 mg, 50 %). ^{1}H NMR (750 MHz, D_2O) δ 5.24 (0.5H, d, J = 3.7 Hz, H-1α, Glc-I, 4.93 (1H, d, J = 3.8 Hz, H-1, Gal-III), 4.70 – 4.65 (1.5H, m, H-1β, Glc-I; H-1, GalNAc-IV), 4.54 – 4.50 (2H, m, H-1, Gal-II, H-1, Gal-V), 4.43 – 4.39 (1H, m, H-5, Gal-III), 4.28 (1H, s, H-4, Gal-III), 4.20 (1H, d, J = 3.0 Hz, H-4, GalNAc-IV), 4.10 – 4.04 (3H, m, H-2, GalNAc-IV; H-3, Gal-V), 4.02 – 3.57 (36.5H, m), 3.55 (1H, dd, J = 9.6, 8.1 Hz, H-2, Gal-V), 3.29 (0.5H, dd, J = 9.0, 8.2 Hz, H-2, Glc-I), 2.79 – 2.70 (2H, m, H-3eq, Neu5Ac-VI; H-3eq, Neu5Ac-VII), 2.07 – 2.01 (9H, m, 3x NHAc), 1.80 (1H, t, J = 12.2 Hz, H-3ax, Neu5Ac-VI), 1.66 (1H, t, J = 12.2 Hz, H-3ax, Neu5Ac-VII). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{54}H_{89}N_3O_{42}$, 1474.481

5-aminopentyl αNeu5Ac-(2→3)-β-D-Galactopyranosyl-(1→3)-[αNeu5Ac-(2→6)]-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(14)-β-D-glucopyra-noside (6b). ST6GalNAc5 and CIAP were added to a mixture of compound **5b** (1.0 mg, 10 mM final concentration) in H_2O with CMP-Neu5Ac (15 mM), $MgCl_2$ (20 mM), sodium cacodylate buffer (50 mM, pH 7.5). The mixture was shaken at 37°C for 4 h and monitored by TLC (EA:MeOH:H_2O:HOAc 3:3:3:2). Purification by Bio-Gel P-2 size exclusion chromatography provided compound **6b** (0.7 mg, 57%). ^{1}H NMR (600 MHz, D_2O) δ 4.92 (1H, d, *J* = 4.0 Hz, H-1, Gal-III), 4.69 (1H, d, *J* = 8.5 Hz, H-1, GalNAc-IV), 4.55 – 4.49 (3H, m, H-1, Gal-II; H-1, Gal-V; H-1, Glc-I), 4.40 (1H, t, *J* = 6.5 Hz, H-5, Gal-III), 4.28 (1H, d, *J* = 2.7 Hz, H-4, Gal-III), 4.20 (1H, d, *J* = 3.0 Hz, H-4, GalNAc-IV), 4.11 – 4.04 (3H, m, H-2, GalNAc-IV; H-3, Gal-V, H-4, Gal-II), 4.03 – 3.53 (39H, m), 3.31 (1H, t, *J* = 8.5 Hz, H-2, Glc-I), 3.06 – 2.99 (2H, m, CH_2, pentyl), 2.80 – 2.70 (2H, m, H-3eq, Neu5Ac-VI; H-3eq, Neu5Ac-VII), 2.07 – 2.00 (9H, m, 3x NHAc), 1.79 (1H, t, *J* = 12.1 Hz, H-3ax, Neu5Ac-VI), 1.75 – 1.63 (5H, m, 2x CH_2, pentyl; H-3ax, Neu5Ac-VII), 1.51 – 1.43 (2H, m, CH_2, pentyl). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{59}H_{100}N_4O_{42}$, 1537.5812; found 1537.5874.

Microarray

Experimental procedures

The synthetic glycans (100 µM in sodium phosphate (250 mM), pH 8.5 buffer) were printed on activated glass slides (Nexterion Slide H, Schott Inc) by piezoelectric non-contact printing (sciFLEXARRAYER S3, Scienion Inc) with a drop volume of ~400 pL and 1 drop per spot at 50 % relative humidity. The compounds were printed as replicates of 6 with on each slide 24 subarrays (3x8). The slides were incubated overnight in a saturated NaCl chamber (providing a 75% relative humidity environment), after which the remaining activated esters were quenched with ethanolamine (50 mM) in TRIS (100 mM), pH 9.0. Slides were rinsed with DI water, dried by centrifugation, and stored in a desiccator at RT.

Sub-arrays were incubated with biotinylated lectins (*Maackia amurensis* leukagglutinin (MAL-II), Soybean agglutinin (SBA) and Wheat Germ agglutinin (WGA); from Vector Labs) at 10 µg/mL premixed with Streptavidin-AlexaFluor635 (5 µg/mL; ThermoFisher Scientific, S32364) in TSM binding buffer (20 mM Tris Cl, pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.05% Tween, 1% BSA) for 1 h followed by washing. Wash steps involved 4 successive washes with each 5 min soak time with 1) TSM wash buffer (20 mM Tris Cl, pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 0.05% Tween-20); 2) TSM buffer (20 mM Tris Cl, pH 7.4, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$); 3) deionized H_2O; and 4) deionized H_2O.

Biotin-conjugated ganglioside GM1 polyclonal antibody (2 µg/mL; Bioss, bs-2367R-Biotin) in TSM binding buffer was incubated for 1 h followed by washing as described above. Next the subarray was incubated with Streptavidin-AlexaFluor635 (5 µg/mL) for 1 h followed by washing.

Using the same buffers as above, recombinant human Siglec-7 comp (a gift from Dr. R.L. Schnaar, Johns Hopkins University School of Medicine, Baltimore, MD, USA) was assayed at 50 µg/mL premixed with 6x-His Tag monoclonal antibody-AlexaFluor647 (5 µg/mL; ThermoFisher Scientific MA1-135-A647) with an incubation for 2 h.

All incubation and wash steps were performed at RT. Washed arrays were dried by centrifugation and immediately scanned for fluorescence on a GenePix 4000 B microarray scanner (Molecular Devices) using a detection gain adjusted to avoid saturation of the signal. The data were processed with GenePix Pro 7 software and further analyzed using our home written Microsoft Excel macro. The lowest and highest value of the 6 replicates were excluded, after which the mean fluorescence intensities (corrected for mean background) and standard deviations (SD) were calculated (n=4). Data were fitted using Prism software (GraphPad Software, Inc). The lowest concentration required for good responsiveness in the optimum dynamic range was selected for all proteins examined.

Results and discussion printing controls

The printing of the synthetic compounds was validated by the plant lectins MAL II, SBA and WGA and a GM1 antibody.

MAL-II binds the terminal trisaccharide sequence Neu5Ac(α2-3)Gal(β1–4)GlcNAc/Glc.[45] Compounds **17** and **20** (Neu5Ac(α2-8)-Neu5Ac(α2-3)Gal(β1–4)Glc) have three terminal intact sugars, and as expected binds to MAL II. Compounds **5b** and **6b**, with the Neu5Ac(α2-3)-Gal-β1,3-GalNAc epitope at the terminal end, are not recognized by MAL II. Similarly, GT1b (**21**) with the same terminal epitope as **5b** and **6b** also did not show binding to MAL II.

SBA preferentially binds GalNAc, and also recognizes Gal residues although at much lower affinity.[46] Binding to SBA was observed for compounds with either a GalNAc or Gal at the terminal residue: **18** (GM2; with GalNAc at the terminal residue) and **4b** (Gb5; with Gal at terminal residue). As expected sialylated compounds **5b** and **6b** (sialylated Gb5; no Gal at the terminal residue) didn't show any binding. Also compound **19** (GM1a; with Gal at the terminal residue) didn't show any binding, due to the inhibition effect of Neu5Ac.

WGA preferentially binds GlcNAc moieties, and also interacts with some glycoproteins via terminal sialic acid residues. Indeed the terminal sialylated compound **21** (GT1b) and sialylated Gb5 (**5b** and **6b**) showed binding, while the non-terminal sialylated compounds **18** (GM2), **19** (GM1a) and **4b** (Gb5) did not bind. Compound **20** (GD3; with terminal α2,8-Neu5Ac-a2,3-Neu5Ac) also did not bind to WGA, apparently WGA does not recognize this sialylated epitope. As expected only GM1a (**19**) showed binding to the GM1 antibody.

References

1. (a) Schnaar, R. L.; Gerardy-Schahn, R.; Hildebrandt, H., *Physiological reviews* **2014**, *94* (2), 461-518; (b) Furukawa, K.; Ohmi, Y.; Tajima, O.; Ohkawa, Y.; Kondo, Y.; Shuting, J.; Hashimoto, N.; Furukawa, K., In *Gangliosides in Health and Disease* ed., Vol. 156; Schnaar, R. L.; Lopez, P. H. H., Eds.; Elsevier B.V., **2018**, pp 265-287.
2. Ho, M. Y.; Yu, A. L.; Yu, J., *Glycoconj. J.* **2017**, *34* (6), 765-777.
3. Liang, Y. J.; Kuo, H. H.; Lin, C. H.; Chen, Y. Y.; Yang, B. C.; Cheng, Y. Y.; Yu, A. L.; Khoo, K. H.; Yu, J., *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107* (52), 22564-22569.
4. (a) Chang, W. W.; Lee, C. H.; Lee, P.; Lin, J.; Hsu, C. W.; Hung, J. T.; Lin, J. J.; Yu, J. C.; Shao, L. E.; Yu, J.; Wong, C. H.; Yu, A. L., *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (33), 11667-11672; (b) Hung, T. C.; Lin, C. W.; Hsu, T. L.; Wu, C. Y.; Wong, C. H., *Journal of the American Chemical Society* **2013**, *135* (16), 5934-5937; (c) Sivasubramaniyan, K.; Harichandan, A.; Schilbach, K.; Mack, A. F.; Bedke, J.; Stenzl, A.; Kanz, L.; Niederfellner, G.; Buhring, H. J., *Glycobiology* **2015**, *25* (8), 902-917; (d) Cheung, S. K.; Chuang, P. K.; Huang, H. W.; Hwang-Verslues, W. W.; Cho, C. H.; Yang, W. B.; Shen, C. N.; Hsiao, M.; Hsu, T. L.; Chang, C. F.; Wong, C. H., *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113* (4), 960-965.
5. Olie, R. A.; Fenderson, B.; Daley, K.; Oosterhuis, J. W.; Murphy, J.; Looijenga, L. H., *British journal of cancer* **1996**, *74* (1), 133-140.
6. Ando, R.; Tokuda, N.; Yamamoto, T.; Ikeda, K.; Hashimoto, N.; Taguchi, R.; Fan, X.; Furukawa, K.; Niimura, Y.; Suzuki, A.; Goto, M.; Furukawa, K., *Glycoconj. J.* **2016**, *33* (2), 169-180.
7. Hakomori, S.; Zhang, Y., *Chem. Biol.* **1997**, *4*, 97-104.
8. (a) Tsai, Y. C.; Huang, J., R.; Cheng, J. Y.; Lin, J. J.; Hung, J. T.; Wu, Y. Y.; Yeh, K. T.; Yu, A. L., *J. Cancer Sc. Ther.* **2013**, *5* (7), 264-270; (b) Wondimu, A.; Liu, Y.; Su, Y.; Bobb, D.; Ma, J. S.; Chakrabarti, L.; Radoja, S.; Ladisch, S., *Cancer research* **2014**, *74* (19), 5449-5457.
9. Ono, M.; Hakomori, S., *Glycoconj. J.* **2004**, *20* (1), 71-78.
10. (a) Lassaletta, J. M.; Schmidt, R. R., *Tetrahedron Lett.* **1995**, *36* (24), 4209-4212; (b) Ishida, H.; Miyawaki, R.; Kiso, M.; Hasegawa, A., *J. Carbohydr. Chem.* **1996**, *15* (2), 163-182; (c) Hsu, C. H.; Chu, K. C.; Lin, Y. S.; Han, J. L.; Peng, Y. S.; Ren, C. T.; Wu, C. Y.; Wong, C. H., *Chemistry* **2010**, *16* (6), 1754-1760.
11. Danishefsky, S. J.; Shue, Y. K.; Chang, M. N.; Wong, C. H., *Accounts of chemical research* **2015**, *48* (3), 643-652.
12. Chen, X.; Varki, A., *ACS chemical biology* **2010**, *5* (2), 163-176.
13. (a) Yu, H.; Huang, S.; Chokhawala, H.; Sun, M.; Zheng, H.; Chen, X., *Angewandte Chemie* **2006**, *45* (24), 3938-3944; (b) Yu, H.; Li, Y.; Zeng, J.; Thon, V.; Nguyen, D. M.; Ly, T.; Kuang, H. Y.; Ngo, A.; Chen, X., *The Journal of organic chemistry* **2016**, *81* (22), 10809-10824.
14. Moremen, K. W.; Ramiah, A.; Stuart, M.; Steel, J.; Meng, L.; Forouhar, F.; Moniz, H. A.; Gahlay, G.; Gao, Z.; Chapla, D.; Wang, S.; Yang, J. Y.; Prabhakar, P. K.; Johnson, R.; Rosa, M. D.; Geisler, C.; Nairn, A. V.; Seetharaman, J.; Wu, S. C.; Tong, L.; Gilbert, H. J.; LaBaer, J.; Jarvis, D. L., *Nat Chem Biol* **2018**, *14* (2), 156-162.
15. (a) Saito, S.; Aoki, H.; Ito, A.; Ueno, S.; Wada, T.; Mitsuzuka, K.; Satoh, M.; Arai, Y.; Miyagi, T., *The Journal of biological chemistry* **2003**, *278* (29), 26474-26479; (b) Senda, M.; Ito, A.; Tsuchida, A.; Hagiwara, T.; Kaneda, T.; Nakamura, Y.; Kasama, K.; Kiso, M.; Yoshikawa, K.; Katagiri, Y.; Ono, Y.; Ogiso, M.; Urano, T.; Furukawa, K.; Oshima, S.; Furukawa, K., *The Biochemical journal* **2007**, *402* (3), 459-470; (c) Gupta, R.; Matta, K. L.; Neelamegham, S., *Biochem. Biophys. Res. Commun.* **2016**, *469* (3), 606-612.
16. Li, S.-P.; Hsiao, W.-C.; Yu, C.-C.; Chien, W.-T.; Lin, H.-J.; Huang, L.-D.; Lin, C.-H.; Wu, W.-L.; Wu, S.-H.; Lin, C.-C., *Adv. Synth. Catal.* **2014**, *356*, 3199-3213.
17. (a) Su, D. M.; Eguchi, H.; Yi, W.; Li, L.; Wang, P. G.; Xia, C., *Organic letters* **2008**, *10* (5), 1009-1012; (b) Tsai, T. I.; Lee, H. Y.; Chang, S. H.; Wang, C. H.; Tu, Y. C.; Lin, Y. C.; Hwang, D. R.; Wu, C. Y.; Wong, C. H., *Journal of the American Chemical Society* **2013**, *135* (39), 14831-14839.
18. Kumagai, D.; Miyazaki, M.; Nishimura, S., *Tetrahedron Lett.* **2001**, *42*, 1953-1956.
19. Wang, Z.; Gilbert, M.; Eguchi, H.; Yu, H.; Cheng, J.; Muthana, S.; Zhou, L.; Wang, P. G.; Chen, X.; Huang, X., *Adv. Synth. Catal.* **2008**, *350* (11-12), 1717-1728.
20. Unverzagt, C.; Kunz, H.; Paulson, J. C., *Journal of the American Chemical Society* **1990**, *112* (25), 9308-9309.

21. Crocker, P. R.; Paulson, J. C.; Varki, A., *Nat. Rev. Immunol.* **2007**, *7* (4), 255-266.
22. Hudak, J. E.; Canham, S. M.; Bertozzi, C. R., *Nature chemical biology* **2014**, *10* (1), 69-75.
23. Yamaji, T.; Teranishi, T.; Alphey, M. S.; Crocker, P. R.; Hashimoto, Y., *The Journal of biological chemistry* **2002**, *277* (8), 6324-6332.
24. Nicoll, G.; Avril, T.; Lock, K.; Furukawa, K.; Bovin, N.; Crocker, P. R., *European journal of immunology* **2003**, *33* (6), 1642-1648.
25. (a) Ito, A.; Handa, K.; Withers, D. A.; Satoh, M.; Hakomori, S., *Febs Lett* **2001**, *498* (1), 116-120; (b) Blixt, O.; Collins, B. E.; van den Nieuwenhof, I. M.; Crocker, P. R.; Paulson, J. C., *The Journal of biological chemistry* **2003**, *278* (33), 31007-31019.
26. Li, T.; Liu, L.; Wei, N.; Yang, J. Y.; Chapla, D. G.; Moremen, K. W.; Boons, G. J., *Nat. Chem.* **2019**, *11* (3), 229-236.
27. Li, P. J.; Huang, S. Y.; Chiang, P. Y.; Fan, C. Y.; Guo, L. J.; Wu, D. Y.; Angata, T.; Lin, C. C., *Angewandte Chemie* **2019**, In press.
28. Kawasaki, Y.; Ito, A.; Withers, D. A.; Taima, T.; Kakoi, N.; Saito, S.; Arai, Y., *Glycobiology* **2010**, *20* (11), 1373-1379.
29. Takashima, S.; Tsuji, S., *Trends Glycosc. Glycotechnol.* **2011**, *23* (132), 178-193.
30. Miyazaki, K.; Ohmori, K.; Izawa, M.; Koike, T.; Kumamoto, K.; Furukawa, K.; Ando, T.; Kiso, M.; Yamaji, T.; Hashimoto, Y.; Suzuki, A.; Yoshida, A.; Takeuchi, M.; Kannagi, R., *Cancer research* **2004**, *64* (13), 4498-4505.
31. Simons, K.; Toomre, D., *Nat Rev Mol Cell Bio* **2000**, *1* (1), 31-39.
32. Kawasaki, Y.; Ito, A.; Kakoi, N.; Shimada, S.; Itoh, J.; Mitsuzuka, K.; Arai, Y., *The Tohoku journal of experimental medicine* **2015**, *236* (1), 1-7.
33. Narimatsu, Y.; Joshi, H. J.; Nason, R.; Van Coillie, J.; Karlsson, R.; Sun, L.; Ye, Z.; Chen, Y. H.; Schjoldager, K. T.; Steentoft, C.; Furukawa, S.; Bensing, B. A.; Sullam, P. M.; Thompson, A. J.; Paulson, J. C.; Bull, C.; Adema, G. J.; Mandel, U.; Hansen, L.; Bennett, E. P.; Varki, A.; Vakhrushev, S. Y.; Yang, Z.; Clausen, H., *Molecular cell* **2019**, *75*, In press.
34. Vudhgiri, S.; Routhu, S. R.; Kumar, C. G.; Prasad, R. B. N.; Jala, R. C. R., *Med. Chem. Res.* **2018**, *27* (1), 285-307.
35. Boons, G. J.; Wang, Z., ed., Vol.; University of Georgia Research Foundation, Inc., **2018**.
36. Kuyama, H.; Nukada, T.; Nakahara, Y.; Ogawa, T., *Tetrahedron Lett.* **1993**, *34* (13), 2171-2174.
37. Noti, C.; de Paz, J. L.; Polito, L.; Seeberger, P. H., *Chemistry* **2006**, *12* (34), 8664-8686.
38. Imamura, A.; Ando, H.; Korogi, S.; Tanabe, G.; Muraoka, O.; Ishida, H.; Kiso, M., *Tetrahedron Lett.* **2003**, *44* (35), 6725-6728.
39. Meng, L.; Forouhar, F.; Thieker, D.; Gao, Z.; Ramiah, A.; Moniz, H.; Xiang, Y.; Seetharaman, J.; Milaninia, S.; Su, M.; Bridger, R.; Veillon, L.; Azadi, P.; Kornhaber, G.; Wells, L.; Montelione, G. T.; Woods, R. J.; Tong, L.; Moremen, K. W., *The Journal of biological chemistry* **2013**, *288* (48), 34680-34698.
40. Phan, J.; Zdanov, A.; Evdokimov, A. G.; Tropea, J. E.; Peters, H. K.; Kapust, R. B.; Li, M.; Wlodawer, A.; Waugh, D. S., *The Journal of biological chemistry* **2002**, *277* (52), 50564-50572.
41. Barb, A. W.; Meng, L.; Gao, Z.; Johnson, R. W.; Moremen, K. W.; Prestegard, J. H., *Biochemistry* **2012**, *51* (22), 4618-4626.
42. Vandersall-Nairn, A. S.; Merkle, R. K.; O'Brien, K.; Oeltmann, T. N.; Moremen, K. W., *Glycobiology* **1998**, *8* (12), 1183-1194.
43. Beckett, D.; Kovaleva, E.; Schatz, P. J., *Protein science : a publication of the Protein Society* **1999**, *8* (4), 921-929.
44. Pedelacq, J. D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S., *Nature biotechnology* **2006**, *24* (1), 79-88.
45. Goldstein, I. J.; Winter, H. C.; Poretz, R. D., In *New Comprehensive Biochemistry*ed., Vol. 29, Part B; Elsevier, **1997**, pp 403-474.
46. Rao, V. S. R.; Lam, K.; Qasba, P. K., *Journal of biomolecular structure & dynamics* **1998**, *15* (5), 853-860.

Chapter 3

Chemical and chemoenzymatic synthesis of globo-series glycosphingolipids

Abstract: Globo-series glycosphingolipids (GSLs) are stem cell biomarkers and play key roles in various diseases such as cancer and bacterial and viral infections. Recent reports indicate that the ceramide moiety of GSLs can modulate protein binding, which encouraged us to synthesize globo-series GSLs. Full chemical synthesis provided globally protected globopentaosylsphingosine (Gb5-Sph). However, Birch reduction to cleave the benzyl ethers was unsuccessful, possibly due to solubility issues. A chemoenzymatic approach was then pursued, in which Gb3-Sph was chemically synthesized and extended by the bifunctional bacterial enzyme LgtD from *Haemophilus inlfuenzae.* LgtD completely converted free reducing Gb3 and Gb4 to Gb4 and Gb5, respectively. Although partial conversion of the Gb3- and Gb4-glycolipids was observed, Gb4-Sph and Gb5-Sph could be obtained. Coupling of Gb3-Sph, Gb4-Sph and Gb5-Sph with palmitic acid provided Gb3-Cer, Gb4-Cer and Gb5-Cer.

Introduction

Glycosphingolipids (GSLs) are ceramides modified by glycans and are the major glycolipid component of mammalian cell membranes. Ceramides consist of a sphingoid base with a fatty acid at the C-2 amine. In mammals, the most common sphingoid base is sphingosine. The C-2 fatty acids can vary in their carbon chain length (14 to >30) and are occasionally unsaturated or bear α-hydroxyl groups. The biosynthesis of ceramide starts in the ER and glycan extensions are performed in the Golgi apparatus. Glycosylations are performed by various glycosyltransferases (GTs), which starts with the formation of lactosylceramide (Lac-Cer) and can then be extended into more complex glycan chains (Fig 1).[1,2,3] The first glycosylation of the globo-series GSLs is catalyzed by A4GalT, which adds α1,4-galactoside on Lac-Cer to provide Gb3-Cer (CD77, P^k).[4] Next, Gb3-Cer can be extended with acetylated β1,3-galactosamine by β3GalNAc-T1 to form Gb4-Cer.[5] The enzyme β3GalT5 adds β1,3-galactoside to Gb4-Cer to form Gb5-Cer.[6] Gb5-Cer can be decorated by α1,2-fucoside (Globo-H) or α2,3-neuraminic acid (MSGb5-Cer, SSEA-4) by FUT1 and FUT2 and ST3Gal2, respectively.[7,8] A second internal α2,6-neuraminic acid can be transferred by ST6GalNAc6 onto MSGb5-Cer to form DSGb5-Cer.[9] Upon completion, the newly synthesized GSLs are transported to the cell membrane.[1,2,3] GSLs mediate biological functions such as cell proliferation, apoptosis, adhesion and migration and the formation of endo- and exosomes.[1,2,3]

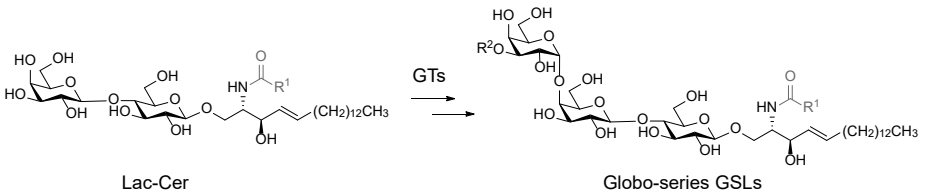


Figure 1. Lactosylceramide (Lac-Cer) is the starting point for a high variety of glycosphingolipids (GSLs). The globo-series start with the addition of α1,4-galactose to form Gb3-Cer. Black: lactosylsphingosine; Grey: fatty acid, where R^1 = a carbon chain varying from 14 to >30 and occasionally contains unsaturations or α-hydroxyl groups. R^2 = the glycosylation site to extend to more complex globo-series GSLs.

Globo-series GSLs have been found on kidney tissue, erythrocytes and are important biomarkers of embryonic and pluripotent stem cells.[10,11,12] Some of the globo GSLs are overexpressed by epithelial cancers, such breast, kidney and prostate cancers.[7,13,14]

GSLs can interact with components residing on the same cell (*cis* interactions) or other cells (*trans* interactions). Since GSLs often cluster in lipid rafts in the membrane, interactions take place in a multivalent manner, resulting in high avidity of binding.[1,15] Important protein interactions have been attributed to the globo-series GSLs. As an example, Gb4 promotes EGFR auto-phosphorylation, activating the downstream cascade and thereby promoting cell proliferation. Another example is the Fas (CD95) receptor, which binds Lac-Cer and Gb3, but not Gb4 or other gangliosides. Upon binding Gb3, ligand-bound Fas will undergo endocytosis and this results in a cell death signal. Furthermore, globo-series GSLs have been identified as interacting partners for viral and bacterial proteins. Gb3 binds gp120 on the human immunodeficiency virus (HIV) and verotoxins 1 and 2

(VT1, VT2), Gb4 to human and simian parvoviruses and Gb5 and MSGb5 to the human parvovirus.[16,17] More thoroughly studied gangliosides such as GD3 and GT1b are good binders of Siglec-7, an immune receptor mainly found on NK cells.[18,19,20,21] Among the globo-series glycosphingolipids disialylated DSGb5 has also been reported as a Siglec-7 ligand.[22,23]

Studying GSLs-protein interactions is challenging, since the isolation of GSLs usually provides heterogeneous mixtures and potential impurities. Besides, most *trans* interactions are thought to take place through the GSL glycan moieties. Therefore, most studies focus on the synthesis of the GSL glycan moieties.[24,25] We and others have synthesized globo-series oligosaccharides, by chemical synthesis[26,27], chemoenzymatic synthesis[28,29,30,] or enzymatic synthesis.[24,31]

Recent studies have attributed important binding interactions to the ceramide moiety of GSLs. One study demonstrated that milk-derived GD3 did bind Siglec-7, whereas GD3 expressed by DLD-1 cells did not bind this receptor. Structural differences were found in the ceramide moieties and turned out to be the reason for alterations in protein binding. These ceramide modifications were only small: phytosphingosine instead of sphingosine or an α-hydroxyl on the fatty acid.[32] Another study focused on how changes in lipid composition of Gb3 affects VT binding. In this study, shorter fatty acid chains (C12, C14) showed little binding, whereas intermediate and long fatty acid chains (C16-C24) showed stronger binding to VT. Also, unsaturated fatty acid moieties significantly increased VT binding.[33] These studies, suggest that ceramides can directly interact with proteins such as VT. The ceramide compositions can also affect raft formation with other components such as cholesterol, resulting in different presentations, which in turn may influence protein binding.[17]

Based on the above described observations, we envisaged that the presence of a ceramide of globo-series GSLs will affect protein binding. To illustrate the differences of the previously used[28] short (C5) aminopentyl linker (Fig. 2A), and the more complex ceramide moiety, structural analogues of Gb5 are presented in Fig. 2B. Chemically well-defined compounds are needed to be synthesized to investigate the importance of the ceramide moiety.

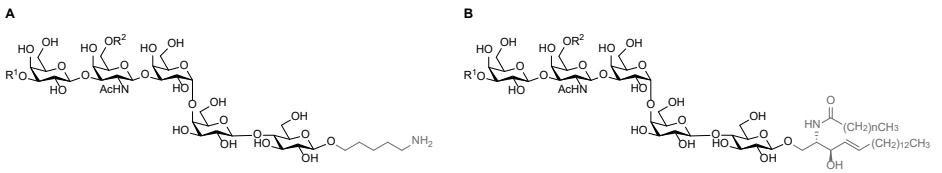


Figure 2. (**A**) Globo-series glycan moieties with an aminopentyl linker for microarray (Chapter 2). (**B**) Natural globo-series glycosphingolipids with natural ceramide, where the fatty acid chain length can vary between 14 to >30. $R^1 = R^2 =$ H: Gb5 | $R^1 = \alpha$2,3Neu5Ac, $R^2 =$ H: MSGb5 | $R^1 = \alpha$2,3Neu5Ac, $R^2 = \alpha$2,6Neu5Ac: DSGb5.

While chemical and enzymatic syntheses of the oligosaccharide moieties of globosides have been described, there are only a few reports dealing with the synthesis of the intact glycolipid of GSLs. Glycosphingolipid synthesis adds a number of challenges and different approaches can be considered.[34] Few reports deal with the chemical glycosylation of a sphingosine or ceramide acceptor with the fully protected saccharide donor obtained by chemical synthesis.[35,36] Since sphingosine or ceramide acceptors are of low reactivity, another strategy has been used where this key chemical glycosylation is performed in an earlier stage of GSL synthesis.[37] The latter chemical synthetic strategy is also called the "Glc-Cer cassette strategy" and was used to synthesize GSL mimic(s), such as GQ1b and the hybrid ganglioside X2.[37,38]

In addition to full chemical synthesis, chemoenzymatic approaches have been developed. The first approach is the enzymatic synthesis of an oligosaccharide moiety, which is protected by benzoyl esters and then converted into an imidate donor for chemical glycosylation with sphingosine.[39] A second strategy starts with the chemical synthesis of glycans having an anomeric α-fluoride, which can then be glycosylated with sphingosine or ceramide by the glycosidase mutant of EGC-II (EGC-E).[40] In this way, the less reactive acceptor is coupled more efficiently and without by-product formation. More recently, a chemoenzymatic strategy to synthesize GSLs has been reported, where initially lactosyl sphingosine is chemically synthesized and then extended by bacterial enzymes.[41]

In this chapter, the chemical and chemoenzymatic syntheses of various globo-series GSLs are described. Full chemical synthesis of protected Gb5-sphingosine (Gb5-Sph) was achieved, however benzyl ethers could not be cleaved by Birch reduction. Next to that, we explored a chemoenzymatic approach, starting from chemically synthesized Gb3-Sph. Extension of Gb3-Sph by the enzyme LgtD from *Haemophilus influenzae.* provided Gb4-Sph and Gb5-Sph consecutively. Lower conversion rates of LgtD were observed on the sphingosine analogues of Gb3, Gb4 and Gb5 compared to the free reducing glycans. With a number of optimizations, Gb3-, Gb4- and Gb5-Sph were successfully synthesized by this strategy. The full GSLs Gb3-Cer, Gb4-Cer and Gb5-Cer were obtained by coupling of the C-2 amine on sphingosine with palmitic acid.

Results and Discussion

Chemical synthesis of globopentaosylsphingosine

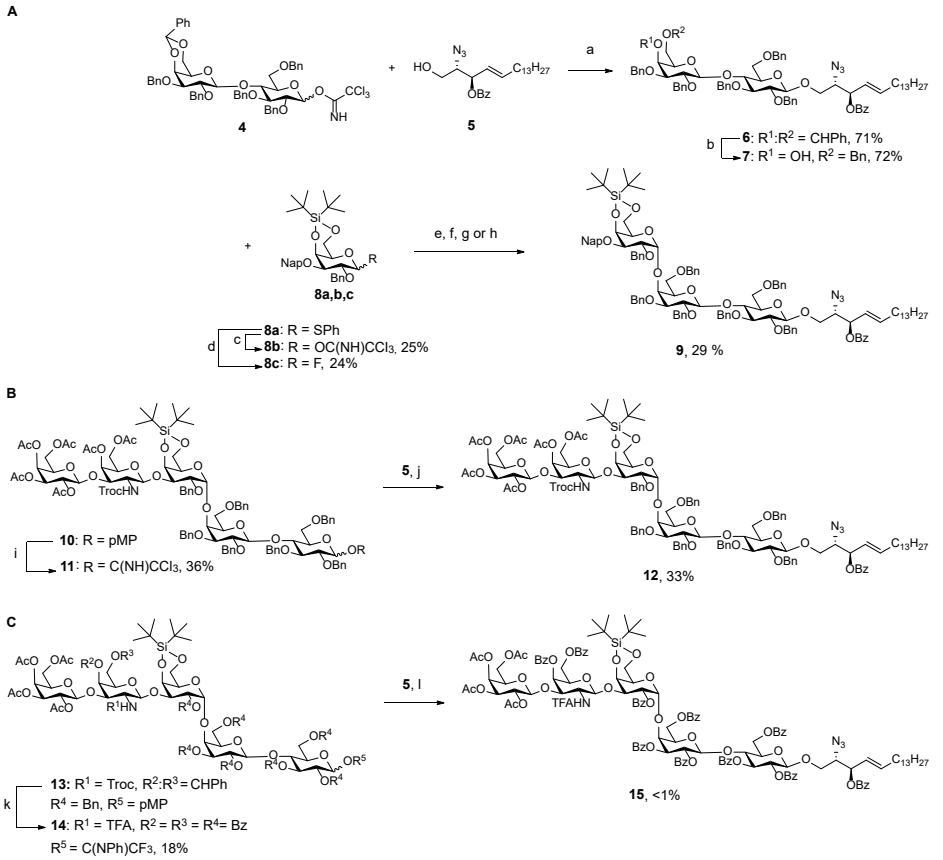
The chemical synthesis of the Gb5 pentasaccharide **10** is described in Chapter 2 and was envisaged as a good starting point for making the sphingolipid derivative.[28] Attachment of a ceramide to compound **10** requires adaption of the chemical synthesis. First, we searched for appropriate reaction conditions to form the β-glycosidic bond with sphingosine in the absence of C-2 neigbouring group participation.

Compound **4** was chosen as a model substrate for the optimization of the β-glycosylation in the absence of neighboring group participation (Scheme 1A). Sphingosine acceptor **5** was synthesized from phytosphingosine in 7 steps as described before.[41] Lactosyl donor **4** and sphingosine acceptor **5** were glycosylated in the presence of BF_3OEt_2 in DCM in the presence of 4 Å molecular sieves at -40 °C. These conditions provided lactosylsphingosine

6 as the pure β-anomer in 71% yield. Having solved the β-glycosylation without neighbouring group participation, we tried to extend lactosylsphingosine **6** with a α1,4-galactose to form globotriaosylsphingosine (Gb3-Sph). First the benzylidene acetal of **6** was selectively opened using Et_3SiH, TfOH at -70 °C to give **7** having C-4 hydroxyl.[28] Various glycosylation conditions of galactosyl donors (**8a-c**) with acceptor **7** were tested to provide the desired trisaccharide **9** (Scheme 1A). In each case low product yields and multiple by-products were obtained, which could be caused to side reactions of activating agents with the sphingosine alkene and low reactivity of the axial C-4 hydroxyl in the acceptor. Pure protected Gb3-Sph **9** could, however, be isolated in a 29% by preparative TLC.

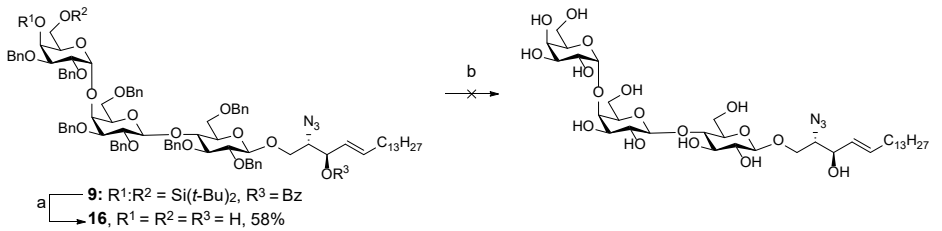
It was envisaged that glycosylations with Gb3-Sph **9** would result in poor yields and by-products. Therefore, we aimed for glycosylation of sphingosine acceptor **5** and pentasaccharide donor **11** (Scheme 1B). Protected pentasaccharide Gb5 **10** was synthesized as described before.[28] Donor **11** was prepared from compound **10** in two steps. First, the pMP protecting group was removed by multiple cycles of treatment with 1.3 eq cerium ammonium nitrate (CAN) to avoid by-product formation. The intermediate lactol was converted into trichloroacetimidate **11** by reaction with trichloroacetonitrile in the presence of 1,8-diazabicyclo(5.4.0)undec-7ene (DBU). Glycosylation of donor **11** with sphingosine acceptor **5** in the presence of BF_3OEt_2 in DCM at -50 °C resulted in the formation of **12** as a separable mixture of α/β anomers**.** The pure β-anomer of **12** (33%) was isolated after silica gel and SX1 size exclusion column chromatography.

Next, we explored another strategy based on C-2 neighbouring group participation to provide β-anomeric selectivity. A C-2 benzoyl ester was chosen as a participating group, since it is less prone to migration compared to an acetyl ester. Donor **14**, containing the C-2 benzoyl ester, was synthesized from pentasaccharide **13** in six steps. Thus, pentasaccharide **13** was treated with Zn and AcOH to remove *N*HTroc. The resulting free amine was by protected by TFA, using trifluoroacetic anhydride and Et_3N. The *N*HTroc was replaced by *N*HTFA to avoid dechlorination under subsequent reductive conditions. The benzyl ethers were cleaved by hydrogenation using $Pd(OH)_2/C$ in a mixture of MeOH/H_2O/HOAc. The resulting free hydroxyls were protected by benzoyl esters by treatment with benzoyl chloride and pyridine. Finally, pMP was removed as described above and *N*-phenyl trifluoroacetimidate was installed to obtain Gb5 donor **14**. Unfortunately, none of the tested glycosylation conditions of donor **14** with sphingosine accepter **5** provided the desired product. It was noted that 4 Å molecular sieves completely abolished glycosylation with sphingosine and only trace product **15** (<1%) was detected in the absence of molecular sieves (Scheme 1C). The lower reactivity of benzoylated donors over benzylated donors has been described before. This phenomenon is often described as “armed” for benzylated and “disarmed” for benzoylated glycosyl donors and is explained by the stronger electron-withdrawing properties of the ester groups over the ether groups.[42] The combination of both poor donor and poor acceptor reactivity are unfavorable for the chemical glycosylation.



Scheme 1. (**A**) Chemical glycosylation of benzylated lactose donor **4** and sphingosine accepter **5** to obtain β-linked lactosylsphingosine without neighbouring group participation. Lactosylsphingosine **6** was turned into acceptor **7** and glycosylated with donors **8a-c** in various conditions. After preparative TLC protected Gb3Sph **9** was isolated. (a) BF_3OEt_2, DCM, -40 °C to -20 °C, 4 Å mol sieves; (b) Et_3SiH, TfOH, DCM, -70 °C; (c) i. NBS, acetone/H_2O; ii. $NCCCl_3$, DBU, DCM, 0 °C; (d) i. NBS, acetone/H_2O, ii. DAST, DCM, -40 °C to -25 °C; (e) **8a**, NIS, TfOH, DCM, -35 °C; (f) **xa**, Ph_2SO, Tf_2O, TTBP, DCM, -70 °C; (g) **8b**, TMSOTf, DCM -60 °C; (h) **8c**, Sn(II)Cl_2, $AgClO_4$, DCM, -15 °C. (**B**) Chemical synthesis of protected Gb5-sphingosine (Gb5-Sph) **12**. Pentasaccharide donor **11** was synthesized from pentasaccharide **10** in two steps and glycosylated with sphingosine acceptor **5**. Reagents and conditions: (i) i. CAN, CH_3CN/H_2O, 0 °C, ii. $NCCCl_3$, DBU, DCM, 0 °C; (j) **5**, BF_3OEt_2, DCM, -50 °C, 4 Å mol sieves; (**C**) Replacement of protecting groups on compound **13** provided pentasaccharide donor **14**. Glycosylation with sphingosine acceptor **5** provided only traces of protected Gb5Sph **15**, in absence of molecular sieves. Reagents and conditions: (k) i. Zn, AcOH, ii. TFAA, Et_3N, DCM, 0 °C, iii. $Pd(OH)_2/C$, H_2, MeOH/H_2O/HOAc, iv. BzCl, pyridine, v. CAN, ACN/H_2O, 0 °C, vi. $ClC(NPh)CF_3$, Cs_2CO_3, 0 °C; (l) TfOH, DCM, 0 °C.

With protected Gb5-Sph **12** in hand (Scheme 1B), appropriate deprotection conditions were explored. Conventional benzyl ether cleavage by hydrogenation could not be used for compound **12**, since the sphingosine double bond will be reduced. Therefore, the benzyl ethers need to be cleaved by Birch reduction to keep the unsaturated sphingosine intact.[43] As a model substrate, Gb3-Sph **9** was converted into compound **16**. The benzoyl ester of compound **9** was removed under Zemplén conditions (*cat.* NaOMe in DCM/MeOH) and the crude intermediate was treated with HF·pyridine to cleave the silane and provide compound **16** (Scheme 2). Birch reduction on compound **16** proved unsuccessful, even when different addition orders of reagents were used or when heated to refluxing temperature (-30 °C). A possible explanation for the failure of the Birch reduction could be lack of the solubility of the substrate.



Scheme 2. Deprotections on compound **9** provided trisaccharide **16**. Birch reduction did not provide the desired deprotected Gb3Sph. Reagents and conditions: (a) i. NaOMe, DCM/MeOH, ii. HF·pyridine, pyridine; (b) Na, NH_3, THF/t-BuOH, -78 °C to -30 °C.

To obtain the desired globo-series glycosphingolipids we moved on to a chemoenzymatic strategy, which we expected to overcome some of the difficulties encountered in the chemical synthesis of Gb5-Sph.

Chemoenzymatic synthesis of Gb3-Cer, Gb4-Cer and Gb5-Cer.

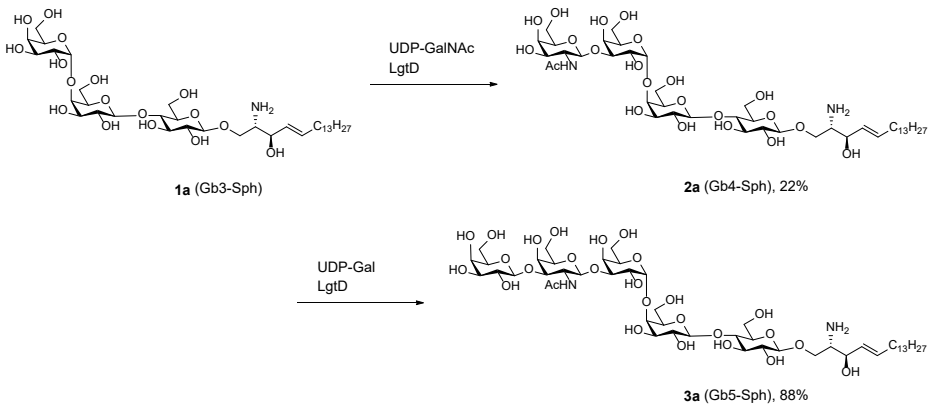
Enzymatic synthesis is usually performed by bacterial enzymes, since these are generally easier to express than mammalian enzymes and have broad substrate specificities. The bacterial enzymes that have been used in the synthesis of globo-series glycan moieties are: LgtC (Gb3), LgtD (Gb4, Gb5), CgtB (Gb5), FutC (Globo-H), JT-FAJ-16 (MSGb5), pmST1 (MSGb5), CST-I (MSGb5) and Psp2,6ST (DSGb5).[30,44,45,46] Two mammalian enzymes have also successfully been used for the synthesis of MSGb5 (ST3Gal1) and DSGb5 (ST6GalNAc5) by our group.[28]

Only one enzyme has been reported to synthesize Gb4 from Gb3; the bacterial enzyme LgtD from *Haemophilus influenzae*. For the second enzymatic extension, two enzymes can be used: CgtB and LgtD. LgtD is a bifunctional enzyme and it is able to catalyze both β1,3 GalNAc addition on Gb3 to form Gb4 and β1,3 Gal addition on Gb4 to form Gb5, however the efficiency of the latter step is lower.[45,46,47] CgtB has been used on a variety of substrates, including the oligosaccharide moiety of Gb4.[48,24]

Although bacterial enzymes LgtD and CgtB have been used for the synthesis of globo-series oligosaccharides, they have not been used for glycolipid synthesis. To ensure that all used enzymes were active, control reactions were performed on oligosaccharide substrates.

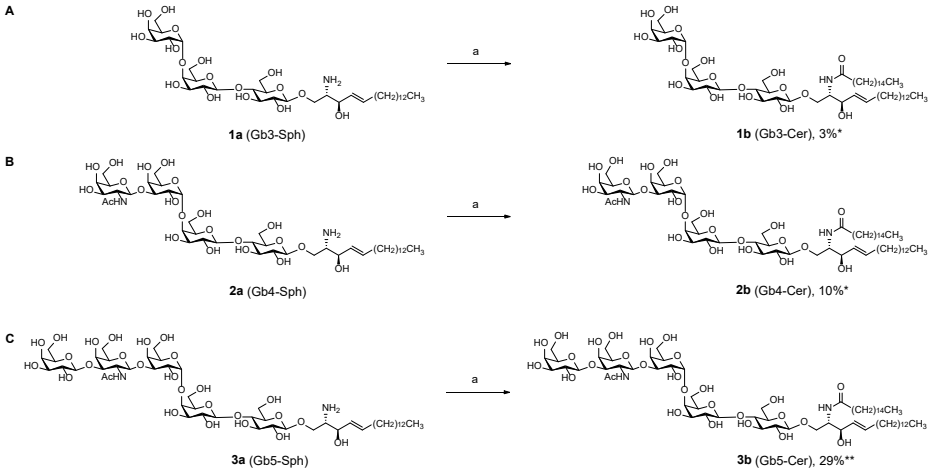
Our chemoenzymatic strategy started with the chemical synthesis of Gb3-Sph **1a** according to a reported procedure.[49] Gb3-Sph **1a** was mixed with UDP-GalNAc, $MgCl_2$ and LgtD in a Tris buffer at pH 7.7 (Scheme 3). Although full conversion was seen of the free oligosaccharide, only 30 % of product **2a** (Gb4-Sph) was formed (based on MALDI-TOF). Earlier reports had indicated that even small substituents at the reducing end such as a benzyl ether or glycolipid can reduce enzyme activity.[50,51,52] The use of a larger quantity of LgtD and UDP-GalNAc did not drive the reaction further. Therefore, the Gb3-Sph/Gb4-Sph **1a**/**2a** mixture was purified by C4 SPE cartridges. The obtained glycolipid mixture was resubmitted to fresh sugar nucleotide and enzyme, and in 5 runs all substrate was converted to Gb4-Sph **2a** (22%).

Gb4-Sph **2a** was mixed with UDP-Gal, $MgCl_2$ and CgtB in Tris buffer at pH 7.7. After 4 h, a mixture was obtained of Gb4-Sph (**2a**), Gb5-Sph (**3a**) and Gb5-Sph with a spectrum of 1-5 galactose additions (based on MALDI-TOF). Over-galactosylation has been decribed before for CgtB variants.[53,54] Conversion of **2a** by LgtD did not provide full conversion to **3a**, however, no over-galactosylation was observed. Therefore, LgtD was the better choice for the synthesis of Gb5-Sph **3a** (88%, 3 runs). Also, a highly improved isolated yield of **3a** over **2a** was obtained by changing SPE purifications from C18 to C4. Pure Gb3-Sph **1a**, Gb4-Sph **2a** and Gb5-Sph **3a** were characterized by 1 and 2D NMR spectroscopy and HRMS spectrometry.



Scheme 3. Enzymatic synthesis of Gb4-Sph **2a** and Gb5-Sph **3a** starting from chemically synthesized Gb3-Sph 1a. Reagents and conditions: substrate (10 mM), UDP-sugar (10-11 mM), Tris buffer pH 7.7 (100 mM), $MgCl_2$ (20 mM) and LgtD at 37 °C.

The successful synthesis of globo-series glycosylsphingosines **1a**, **2a** and **3a** encouraged us to synthesize the intact glycosylceramides. To this end, common amide coupling with a fatty acid was chosen. All coupling reactions were performed by shaking the respective glycosylsphingosine (**1a**-**3a**) with palmitic acid (C16:0) and EDC/HOBt in DMF (Scheme 4A-C). Full conversions were seen on TLC ($CHCl_3$/MeOH/H_2O) and by MALDI-TOF for all compounds (**1b-3b**). Again material was lost on the C18 cartridge for compounds **1b** and **2b**. Gb5-Cer, on the other hand, was purified by C4 SPE (MeOH/H_2O) purification and this resulted in a moderate yield (**3b**, 29%). All products (**1b-3b**) were characterized by 1D and 2D NMR and HRMS.



Scheme 4. (**A**) Chemical synthesis of Gb3Cer **1b** from Gb3Sph **1a**; (**B**) Gb4Cer **2b** from Gb4Sph **2a**; (**C**) and Gb5Cer **3b** from Gb5Sph **3a**. Reagents and conditions: (a) palmitic acid, EDC, HOBt, Et_3N, DMF. Isolated yields after C18* or C4** SPE purifications.

Conclusions

Several synthetic methods have been explored for the synthesis of globo-series glycosphingolipids. Gb5 was synthesized before and was used as a starting point for the synthesis of Gb5-Sph. Successful chemical glycosylation conditions were found to form β-lactosylsphingosine in the absence of neighbouring group participation. However, further glycosylations on lactosylsphingosine resulted in mixtures of products and lower yields, which might be due to side-reactions with the alkene moiety or poor acceptor reactivity. Glycosylation of a Gb5 trichloroacetimidate donor and sphingosine acceptor provided the desired protected Gb5-Sph as a separable α/β mixture in absence of neighbouring group participation. A pentasaccharide donor with a C-2 benzoyl ester was also prepared, in order to overcome the formation of an anomeric mixture. However, due to low reactivities of both the donor and acceptor only trace amounts of product were formed. Conventional benzyl ether cleavage by hydrogenation is not suitable in the presence of the unsaturated sphingosine. However, our attempts to cleave benzyl ethers of Gb3-Sph by Birch reduction failed, which might be due to solubility issues. Next, an chemoenzymatic approach was explored starting by the chemical synthesis

of Gb3-Sph. The bifunctional enzyme LgtD from *Haemophilus influenzae* was used to provide both Gb4-Sph and Gb5-Sph. Although the used enzymes showed higher activity on the free reducing glycans than on the glycolipids, cycles of resubmission converted all substrate to product. Coupling with palmitic acid provided glucosylceramides Gb3-Cer, Gb4-Cer and Gb5-Cer. We found that C4 SPE purifications using MeOH/H_2O as the eluent provided higher yields than earlier reported C18 SPE purifications with CH_3CN/H_2O. Future optimizations of the chemoenzymatic synthesis requires higher enzyme activities. We suggest that expression of engineered bacterial enzymes or the use of mammalian enzymes βGalNAc-T1 and β3GalT5 will provide better conversion of GSL substrates. Combined with slight optimizations in the purification methods sufficient material may be provided to study binding with immunological or cancer-associated proteins.

Experimental Section

Chemical synthesis

NMR nomenclature

The monosaccharides of glycosphingolipid Gb5-Cer have been labeled as shown in Fig. S1. Starting from the reducing end of the pentasaccharide core Gb5, these were labeled as Glc-I, Gal-II, Gal-III, GalNAc-IV, Gal-V. Globo-series glycans contain either sphingosine (Sph, R = H), or ceramide (Cer, R = $C_{16}H_{33}$).

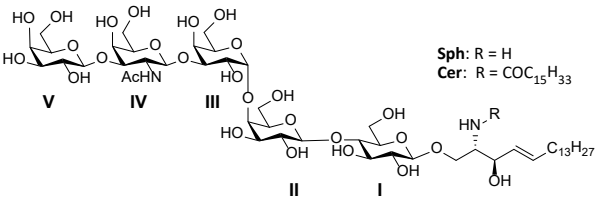
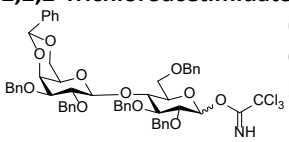
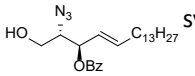


Figure S1. Monosaccharide labeling system for globo-series glycosphingolipids having either sphingosine (Sph, R = H) or ceramide (Cer, R = $COC_{16}H_{33}$).

Experimental procedures

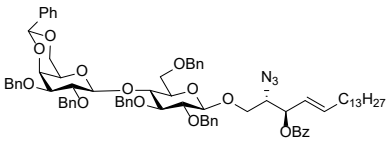


2,2,2-Trichloroacetimidate 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (4). Compound **4** was synthesized as described before (Chapter 2; compound 26).[28]



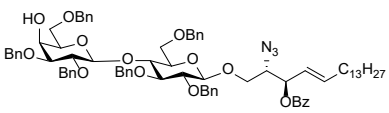
(2S,3R,4E)-2-Azido-3-O-benzoyloxy-octadec-4-ene-1-ol (5). Compound **5** was synthesized as described before.[41]

(2S,3R,4E)-2-azido-3-O-benzoyloxy-octadec-4-enyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (6). A mixture of



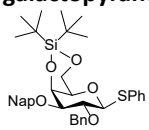
acceptor **5** (447 mg, 1.0 mmol), donor **4** (1.60 g, 1.6 mmol) and 4 Å molecular sieves was stirred in DCM for 3 h. The mixture was cooled to -50 °C and BF_3OEt_2 (38 µL, 0.3 mmol) was added. The mixture was allowed to warm to -20 °C over 2 h. The reaction mixture was quenched with Et_3N, filtered over Celite and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 7:3 v/v) as the eluent to afford compound **6** (962 mg, 71 %, only β anomer). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06 (2H, d, *J* = 7.3 Hz, H-Ar), 7.61 – 7.09 (33H, m, H-Ar), 5.96 – 5.80 (1H, m, CH=C*H*-CH_2), 5.71 – 5.61 (1H, m, C*H*OBz), 5.61 – 5.48 (1H, m, CHOBz-C*H*=CH), 5.45 (1H, s, C*H*-C_6H_5), 5.19 (1H, d, *J* = 10.6 Hz, CH*H*), 5.00 – 4.65 (7H, m), 4.52 (1H, d, *J* = 12.1 Hz, CH*H*), 4.46 (1H, d, *J* = 7.8 Hz, H-1, Gal-II), 4.39 (1H, d, *J* = 7.6 Hz, H-1, Glc-I), 4.30 (1H, d, *J* = 12.1 Hz, CH*H*), 4.19 (1H, d, *J* = 12.3 Hz, CH*H*), 4.06 – 3.54 (10H, m), 3.47 (1H, t, *J* = 8.3 Hz), 3.38 (2H, dd, *J* = 17.4, 7.4 Hz), 2.94 (1H, s), 2.02 (2H, d, *J* = 6.8 Hz, CH=CH-C*H*$_2$), 1.44 – 1.13 (22H, m, 11 x CH_2), 0.87 (3H, t, *J* = 6.6 Hz, CH_3, Sph). ESI HRMS (*m/z*): $[M + NH_4]^+$ calcd for C79H93N3O13; 1309.7047 found 1309.7098.

(2S,3R,4E)-2-azido-3-O-benzoyloxy-octadec-4-enyl 2,3-di-*O*-benzyl-6-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (7). Compound **6** (1.0

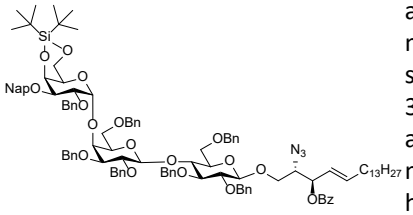


g, 0.77 mmol) was stirred in DCM (7 mL) with 4 Å molecular sieves for 2.5 h. The mixture was cooled to -70 °C and Et_3SiH (616 µL, 3.87 mmol) and TfOH (137 µL, 1.55 mmol) were added. After 1 h the same amounts of Et_3SiH and TfOH were added. The mixture was allowed to warm to -40 °C after which it was quenched by addition of Et_3N. The crude was filtered over Celite, concentrated *in vacuo* and purified by silica column chromatography using Hexane:EtOAc (1:0 to 7:3 v/v) as the eluent to afford compound **7** (719 mg, 72 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06 (2H, d, *J* = 7.3 Hz, H-Ar), 7.61 – 7.11 (32H, m, H-Ar), 5.92 – 5.83 (1H, m, CH=C*H*-CH_2), 5.66 (1H, dd, *J* = 8.0, 3.8 Hz, C*H*OBz), 5.55 (1H, dd, *J* = 15.4, 7.9 Hz, CHOBz-C*H*=CH), 4.99 (1H, d, *J* = 10.8 Hz, CH*H*), 4.88 (1H, d, *J* = 11.1 Hz, CH*H*), 4.80 – 4.62 (6H, m), 4.52 (1H, d, *J* = 12.2 Hz, CH*H*), 4.47 – 4.31 (5H, m, H-1, Gal-II; H-1, Glc-I), 4.04 – 3.90 (4H, m, H-4, Gal-II; C*H*N_3), 3.78 (1H, dd, *J* = 11.0, 4.2 Hz), 3.71 – 3.53 (5H, m, H-2, Gal-II), 3.49 – 3.29 (5H, m, H-2, Glc-I; H-3, Gal-II), 2.43 (1H, d, J = 2.4 Hz, O*H*), 2.04 – 1.97 (2H, m, CH=CH-CH_2, Sph), 1.44 – 1.17 (22H, m, 11 x CH_2, Sph), 0.87 (3H, t, J = 6.9 Hz, CH_3, Sph).

Phenyl 2-*O*-benzyl-3-*O*-(2-naphthyl)methyl-4,6-*O*-di-*tert*-butylsilanediyl-1-thio-β-D-galactopyranoside (8a) Compound **8a** was synthesized as described before (chapter 2: compound 11).[28]

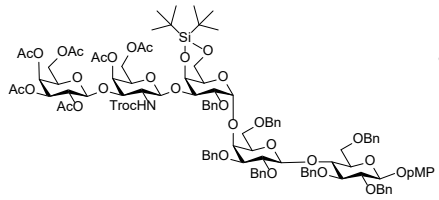


(2S,3R,4E)-2-azido-3-O-benzoyloxy-octadec-4-enyl 2-*O*-benzyl-3-*O*-(2-naphthyl) methyl-4,6-*O*-di-*tert*-butylsilanediyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (9). A mixture of



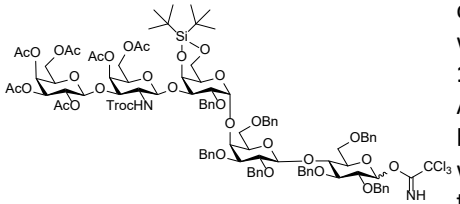
acceptor **7** (695 mg, 0.55 mmol), donor **8a** (483 mg, 0.75 mmol) and 4 Å molecular sieves was stirred in DCM for 1 h. The mixture was cooled to -35 °C and *N*-iodosuccinimide (242 mg, 1.07 mmol) and triflic acid (5 µL, 0.05 mmol) were added. The mixture was allowed to warm to -25 °C and after 1 h the reaction was quenched by addition of Et_3N. The mixture was filtered over Celite and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 15:1 v/v) as the eluent. Subsequent purification by preparative TLC using Toluene:EtOAc (95:7 v/v) as the eluent provided title compound **9** (286 mg, 29%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06 (2H, d, *J* = 7.4 Hz, H-Ar), 7.89 – 7.70 (5H, m, H-Ar), 7.68 (1H, s, H-Ar), 7.55 (2H, t, *J* = 7.6 Hz, H-Ar), 7.50 – 7.36 (7H, m, H-Ar), 7.35 – 7.06 (30H, m, H-Ar), 5.93 – 5.80 (1H, m, CH=C*H*-CH_2), 5.65 (1H, dd, *J* = 7.9, 3.7 Hz, C*H*OBz), 5.59 – 5.48 (1H, m, *J* = 15.8, 7.2 Hz, CHOBz-C*H*=CH), 5.07 (1H, d, *J* = 11.3 Hz, CH*H*), 4.97 (1H, d, *J* = 3.3 Hz, H-1, Gal-III), 4.86 (1H, d, *J* = 11.0 Hz, CH*H*), 4.82 – 4.63 (7H, m), 4.61 – 4.40 (6H, m, H-1, Gal-II; H-4, Gal-III), 4.37 (1H, d, *J* = 7.7 Hz, H-1, Glc-I), 4.35 – 4.23 (3H, m), 4.13 – 4.07 (1H, m), 4.07 – 3.87 (7H, m, H-4, Gal-II; H-2, Gal-III; C*H*N_3; H-4, Glc-I; H-3, Gal-III), 3.82 – 3.65 (4H, m, C*H*$_2$-CHN_3), 3.63 – 3.47 (4H, m, H-3, Glc-I; H-2, Gal-II), 3.46 – 3.40 (1H, m, H-2, Glc-I), 3.38 – 3.29 (2H, m), 3.26 (1H, dd, *J* = 10.0, 2.6 Hz, H-3, Gal-II), 2.01 (2H, dt, *J* = 14.0, 6.9 Hz, CH=CH-C*H*$_2$), 1.42 – 1.16 (22H, m, *J* = 26.4, 7.0 Hz, 11 x CH_2), 1.09 – 0.94 (18H, m, *J* = 27.7, 12.6 Hz, 2x *t*-Bu), 0.87 (3H, t, *J* = 6.7 Hz, CH_3, Sph). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.1 (C=O, Bz), 139.2, 138.8, 138.7 (CH=*C*H-CH_2), 138.6, 138.5, 138.3, 138.2, 136.7, 133.3, 133.1, 132.9, 130.0, 129.8, 128.6, 128.4, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.1, 128.0, 127.9, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 127.4, 127.4, 127.4, 127.3, 126.2, 126.0, 125.9, 125.9, 125.7, 125.7, 122.7 (CHOBz=*C*H=CH), 103.4 (C-1, Glc-I), 102.9 (C-1, Gal-II), 100.1 (C-1, Gal-III), 82.3 (C-3, Glc-I), 81.6 (C-2, Glc-I), 81.1 (C-3, Gal-II), 79.0 (C-2, Gal-II), 78.0 (C-3, Gal-III), 76.9, 75.29, 75.17, 75.0 (*C*HOBz), 74.88, 74.2 (C-2, Gal-III), 73.6, 73.4 (C-4, Gal-II), 73.32, 73.15, 73.09, 72.16, 71.2 (C-4, Gal-III), 70.57, 68.46, 68.32, 68.14, 67.60, 67.5 (C-4, Glc-I), 67.06, 64.0 (C$\underline{H}$$N_3$), 32.4 (CH=CH-C$\underline{H}$$_2$), 31.9, 29.7, 29.7, 29.7, 29.6, 29.4, 29.4, 29.2, 28.7, 27.7 (3x CH_3, *t*-Bu) 27.3 (3x CH_3, *t*-Bu), 23.3, 22.7, 20.7, 14.1 (CH_3, Sph). ESI HRMS (*m/z*): $[M + NH_4]^+$ calcd for $C_{111}H_{135}N_3O_{18}Si$; 1843.9848 found 1843.9883.

***Para*-methoxyphenyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-acetyl-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranosyl-(1→3)-2-*O*-benzyl-4,6-*O*-di-*tert*-butyl-silanediyl-α-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (10).**



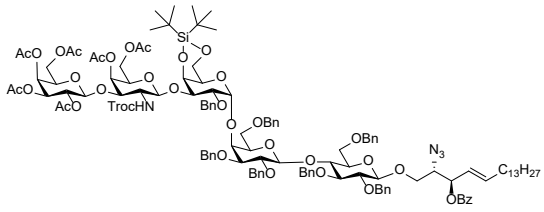
Compound **10** was synthesized as described before (Chapter 2; 16b).[28]

2,2,2-Trichloroacetimidate 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-acetyl-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranosyl-(1→3)-2-*O*-benzyl-4,6-*O*-di-*tert*-butyl-silanediyl-α-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (11). Oven-dried



ceric ammonium nitrate (63 mg, 0.122 mmol) was added to a stirring solution of compound **10** in CH_3CN (2.4 mL)/H_2O (0.4 mL) at 0 °C. After 45 min the mixture was quenched by addition of sat. aq. NaHCO3 and diluted with DCM. The layers were separated and the organic layer was washed with sat. aq. NaHCO3 (2 x), H2O, dried (Na2SO4), filtered and concentrated *in vacuo.* The obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 1:1 v/v) as the eluent to provide the product and starting material. The starting material was submitted again in the same manner as described before and this was repeated 3x. 2,2,2-Trichloroacetonitrile (34 µL, 0.343 mmol) and DBU (2 µL, 0.014 mmol) were added to the intermediate (139 mg, 0.069 mmol) in DCM (1.5 mL) with 4 Å molecular sieves at 0 °C. After 4.5 h the mixture was purified by silica column chromatography using Hexane:EtOAc (1:0 to 4:6 v/v) as the eluent to isolate compound **11**. Compound **11** was directly used in the next step.

(2S,3R,4E)-2-azido-3-O-benzoyloxy-octadec-4-enyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-acetyl-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranosyl-(1→3)-2-*O*-benzyl-4,6-*O*-di-*tert*-butyl-silanediyl-α-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-galactopyranosyl-(1→4)-2,3,6-tri-*O*-benzyl-β-D-glucopyranoside (12). A mixture of acceptor **5** (15 mg, 0.035 mmol) and



donor **11** (100 mg, 0.046 mmol) was co-evaporated with toluene (3 x) and dried under high vacuum overnight. The dried compound mixture was stirred with 4 Å molecular sieves in DCM for 1.5 h. The mixture was cooled to -50 °C and BF_3OEt_2 (0.4 µL in 100µL DCM, 0.0035 mmol) was added. After 3 h the reaction was quenched by addition of Et_3N. The mixture was filtered over Celite and concentrated *in vacuo*. The obtained residue was purified by silica gel chromatography using Toluene:EtOAc (1:0 to 8:2 v/v) as the eluent. Subsequent purification by Biogel SX-1 using Toluene:Acetone (1:1 v/v) as the eluent provided the β-anomer of the title compound. ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.05 (2H, d, *J* = 7.6 Hz, H-Ar), 7.58 – 7.11 (38H, m, H-Ar), 5.89 – 5.81 (1H, m, CH=C*H*-CH_2), 5.62 (1H, dd, *J* = 7.8, 4.1 Hz, C*H*OBz), 5.53 (1H, dd, *J* = 15.3, 7.9 Hz, CHOBz-C*H*=CH), 5.36 (1H, d, *J* = 3.1 Hz, H-4, Gal-V), 5.21 (1H, d, *J* = 3.0 Hz, H-4, Gal-IV), 5.07 – 4.98 (2H, m, H-2, Gal-V), 4.94 (1H, d, *J* = 11.7 Hz, CH*H*), 4.89 – 4.77 (4H, m, H-3, Gal-V; H-1, Gal-III), 4.75 – 4.64 (4H, m), 4.59 (1H, s, H-4, Gal-III), 4.56 – 4.42 (5H, m, H-1, Gal-II; H-1, Gal-V), 4.41 – 4.11 (9H, m, H-1, Glc-I, H-1, GalNAc-IV), 4.10 – 3.71 (15H, m, H-4, Glc-I; C*H*N_3; H-4, Gal-II; H-2, Gal-III; H-5, Gal-V; H-3, Gal-III), 3.68 – 3.46 (7H, m, H-2, Gal-II; H-3, GalNac-IV; H-3, Glc-I), 3.41 (1H, t, *J* = 8.4 Hz, H-2, Glc-I), 3.33 – 3.26 (3H, m, H-3, Gal-II), 2.16 (3H, s, CH_3, OAc), 2.10

– 2.06 (6H, m, 2 x CH_3, OAc), 2.04 – 1.95 (11H, m, 3 x CH_3, OAc; CH=CH-CH_2), 1.38 – 1.18 (22H, m, 11 x CH_2), 0.97 (9H, s, 3 x CH_3, *t*-Bu), 0.93 (9H, s, 3 x CH_3, *t*-Bu), 0.88 (3H, t, *J* = 7.0 Hz, CH_3, Sph). ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.4 (C=O, Ac), 170.4 (C=O, Ac), 170.3 (C=O, Ac), 170.1 (C=O, Ac), 170.0 (C=O, Ac), 169.4 (C=O, Ac), 165.1 (C=O, Bz), 153.8 (C=O, Troc), 139.8, 138.9 (CH=CH-CH_2), 138.5, 138.5, 138.3, 138.2, 138.0, 138.0, 133.2, 130.0, 129.8, 129.7, 128.6, 128.5, 128.4, 128.4, 128.4, 128.3, 128.3, 128.3, 127.8, 127.8, 127.8, 127.6, 127.6, 127.5, 127.3, 122.7 (CHOBz-CH=CH), 103.4 (C-1, Glc-I), 102.6 (C-1, Gal-II), 102.4 (C-1, GalNAc-IV), 100.9 (C-1 Gal-V), 100.0 (C-1, Gal-III), 95.2 (CCl_3, Troc), 81.5 (C-2, Glc-I), 81.1 (C-3, Gal-II), 81.0 (C-3, Gal-III), 80.7 (C-3, Gal-I), 79.5 (C-3, GalNAc-IV), 76.2 (C-4, Gal-II), 75.2, 75.0, 75.0 (C-2, Gal-II), 74.9 (CHOBz), 74.2, 74.0 (C-4, Gal-III), 73.7, 73.5, 73.1, 73.0 (C-2, Gal-III), 72.6 (C-4, Glc-I), 72.0, 70.8 (C-5, Gal-V), 70.60, 70.5 (C-3, Gal-V), 68.6 (C-2, Gal-V), 68.4, 68.3 (C-4, GalNAc-IV), 68.1, 67.5, 67.2, 67.1, 66.8 (C-4, Gal-V), 64.0 (CHN_3), 62.3, 61.2, 54.0 (C-2, GalNAc-IV), 32.4 (CH=CH=CH_2), 31.9, 29.7, 29.6, 29.4, 29.4, 29.2, 28.7, 27.5 (3x CH_3, *t*-Bu), 27.3 (3x CH_3, *t*-Bu), 23.3, 22.7, 20.8 (CH_3, OAc), 20.8 (CH_3, OAc), 20.8 (CH_3, OAc), 20.8 (CH_3, OAc), 20.7 (CH_3, OAc), 20.6 (CH_3, OAc), 14.2 (CH_3, Sph). ESI HRMS (*m/z*): $[M + NH_4]^+$ calcd for $C_{127}H_{161}Cl_3N_4O_{35}Si$; 2453.0114 found 2453.0146.

3

Chemoenzymatic synthesis

Expression of LgtD from *Haemophilus influenzae*

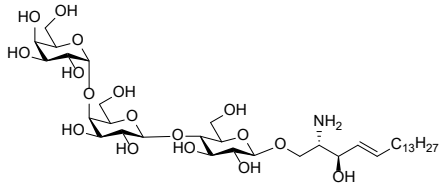
Materials. The full-length codon-optimized gene encoding for HiLgtD (ENA: AAC23227) was synthesized by Genscript and subcloned into pET15b+ (NdeI, BamHI). Bl21(DE3) (C2527H) cells were purchased from New England Biolabs (NEB). Ampicillin (sodium salt) (14417) was ordered through Cayman Chemicals. Imidazole (56750), TRIS-HCl (T5941), NaCl (S9888), lysozyme (62971) and triton X-100 (T8787) came from Sigma-Aldrich. 2xYT medium (BP97432) and 2xYT Agar (BP2466) were ordered from Fisher BioReagents. Isopropyl-beta-D-thiogalactopyranoside (IPTG) (R0393) as well as GelCode Blue Stain Reagent (24592) was ordered from Thermo Scientific. The Ni-NTA resin (17-5318-01) and the Superdex 200 Increase 10/300 GL column (28990944) were ordered from GE Healthcare. Sartorius provided Vivaspin 6, 10000 MWCO (VS0602) spinfilters. For SDS-PAGE, SurePAGE, Bis-Tris, 10x8, 4-12%, 12 wells (M00653) from GenScript were used. Laemmli sample buffer 2x (161-0737) and Precision Plus Protein Dual Color Standards (161-0374) were ordered from Bio-Rad Laboratories. A standard buffer consisting of 50 mM TRIS-HCl and 250 mM NaCl pH 8 was made. Imidazole was added to the standard buffer at concentrations of 20 mM, 50 mM or 250 mM to make wash 1, wash 2 and elution buffer (re-adjusted to pH 8 if needed) respectively. To wash 1, lysozyme (1mg/mL) and triton X-100 (0.1%) was added to make the lysis buffer.

Protocol. BL21(DE3) cells were transformed with each of the individual vectors, a colony was picked from the 2xYT agar plate with ampicillin (100 µg/mL) and expanded to a cell culture volume of 500 mL ampicillin (100 µg/mL) containing 2xYT medium at 37°C. The cells were induced at OD_{600} = 0.6 with IPTG (final concentration was 1 mM) and cultured overnight at 20°C. Then, cells were pelleted at 3000 *xg*, resuspended in lysis buffer (5% of culture volume), incubated at 37°C for 1 h and sonicated for 30' on ice. Clear supernatant, by removal of cell debris by centrifugation at 10000 *xg*, was loaded onto a gravity flow 4 mL Ni-NTA column in case of HiLgtD. Once washed with 10 column volumes

(CV) of wash 1, 10 CV of wash 2, HiLgtD was eluted in 3 CV of elution buffer. The eluate was concentrated using a Vivaspin 6 filter and further purified on a PBS equilibrated Superdex 200 Increase 10/300 GL column attached to a Shimadzu Nexera system using a flow of 0.45 ml/min collecting fractions each minute. Fractions containing HiLgtD were pooled, concentrated, aliquoted and stored at -20°C. Further characterization of the enzyme was done by SurePage 4-12% gel, using Laemmli sample buffer and the Precision Plus Protein marker as reference.

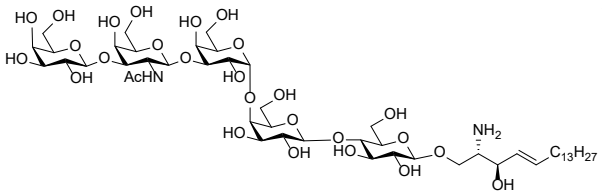
Experimental procedures

(2S,3R,4E)-2-amino-3-hydroxy-octadec-4-enyl α-D-galactopyranosyl-(1->4)-β-D-galacto-pyranosyl-(1→4)-β-D-glucopyranoside (1a). Compound **1a** (Gb3Sph) was synthesized as described before.[49]



(2S,3R,4E)-2-amino-3-hydroxy-octadec-4-enyl 2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (2a). LgtD (100 µL, 0.5 mg) was added to a mixture of compound **1a** (5.0 mg, 10 mM final concentration) in H_2O with UDP-GalNAc (3.9 mg, 10 mM), $MgCl_2$ (20 mM), Tris buffer (100 mM, pH = 7.7). The mixture was shaken at 37 °C overnight and monitored by Maldi-TOF. Purification by bench top C4 column using MeOH:H_2O (1:0 to 0:1 v/v) as the eluent provided a mixture of Gb3Sph/Gb4Sph. The glycolipid mixture was resubmitted as described above. This was repeated until no more starting material was seen after C4 purification. (compound **2a** (Gb4Sph), 1.4 mg, 22 %). ^{1}H NMR (600 MHz, MeOD) δ 5.82 – 5.72 (1H, m, CH=C*H*-CH_2), 5.55 – 5.44 (1H, m, CHOH-C*H*=CH), 4.94 (1H, d, *J* = 3.9 Hz, H-1, Gal-III), 4.62 (1H, d, *J* = 8.0 Hz, H-1, GalNAc-IV), 4.44 – 4.38 (1H, m, H-1, Gal-II), 4.33 – 4.29 (1H, m, H-1, Glc-I), 4.27 (1H, t, *J* = 6.2 Hz, H-5, Gal-III), 4.17 (1H, d, *J* = 2.2 Hz, H-4, Gal-III), 4.03 (1H, t, *J* = 6.7 Hz, CH=CH-C*H*OH), 4.00 – 3.48 (22H, m), 3.44 – 3.38 (1H, m), 3.36 – 3.25 (1H, m, H-2, Glc-I), 3.01 – 2.94 (1H, m, *H*C-NH_2, Sph), 2.13 – 2.05 (2H, m, CH=CH-C*H*$_2$), 2.01 (3H, s, CH_3, NHAc), 1.47 – 1.24 (22H, m, 11 x CH_2, Sph), 0.90 (3H, t, *J* = 7.0 Hz, CH_3, Sph). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{44}H_{80}N_2O_{22}$; 988.5203 found 989.5283.

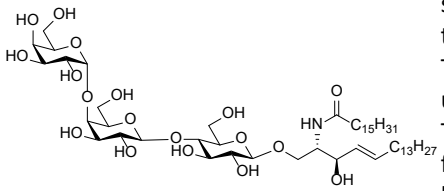
(2S,3R,4E)-2-amino-3-hydroxy-octadec-4-enyl β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (3a). LgtD (50 µL, 0.25 mg) was added to a mixture of compound **2a** (0.9 mg, 10 mM final concentration) in H_2O with UDP-Gal (0.6 mg, 11 mM), $MgCl_2$ (20 mM), Tris buffer (100 mM, pH = 7.7). The mixture was shaken at 37 °C overnight and monitored by Maldi-TOF. Purification by bench top C4 column using MeOH:H_2O (1:0 to 0:1 v/v) as the eluent provided a mixture of Gb4Sph/Gb5Sph. The glycolipid mixture was resubmitted as described above. This was repeated until no more starting material was seen after C4 purification. (compound **3a** (Gb5Sph), 0.92 mg, 88 %). ^{1}H NMR (600 MHz, MeOD) δ 5.83 – 5.73 (1H, m, CH=*CH*-CH_2), 5.49 (1H, dd, *J* = 15.1, 8.0 Hz, CHOH-C*H*=CH), 4.93 (1H, d, J = 4.0 Hz, H-1, Gal-III), 4.71 (1H, d, *J* = 8.4 Hz, H-1, GalNAc-IV), 4.41 (1H, s, H-1, Gal-II), 4.34 (1H, d, *J* = 7.8 Hz, H-1, Gal-V), 4.32 (1H, d, *J* = 7.7 Hz, H-1, Glc-I), 4.30 – 4.23 (1H, m), 4.17 (1H, s), 4.12 – 4.02 (2H, m), 3.99 (1H, s), 3.96 – 3.16 (29H, m), 2.10 (2H, dd, *J* = 14.1, 6.8 Hz, CH=CH-C*H*$_2$), 1.98 (3H, s, CH_3, NHAc), 1.47 – 1.19 (22H, m, 11 x CH_2), 0.90 (3H, t, *J* = 6.9 Hz, CH_3, Sph). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{50}H_{90}N_2O_{27}$; 1150.5731 found 1151.5823.



General procedure for ceramide synthesis

Glycosylsphingosine (1 eq, 10mM), was dissolved in DMF and palmitic acid (3 eq), EDC·HCl (3 eq), Et_3N (5-10% v/v) and HOBt (3 eq) were added from freshly prepared stock solutions in DMF. The reaction mixture was shaken in an 1.5mL eppendorf tube at room temperature overnight. The reaction was monitored by MALDI-TOF and after overnight shaking, H_2O was added and the mixture was lyophilized. The crude residue was purified on a pre-equilibrated C4 or C18 cartridge.

(2S,3R,4E)-3-hydroxy-2-(hexadecanamido)octadec-4-enyl (α-D-galactopyranosyl)-(1→4)-(β-D-galactopyranosyl)-(1→4)-β-D-glucopyranoside (1b). Compound **1b** was synthesized from compound **1a** according to the general procedure for ceramide synthesis. The crude was purified by a 1mL C18 cartridge using H_2O:CH_3CN (1:0 to 2:8 v/v) as the eluent. The organic solvent was removed under N_2 flow and the residual water was removed by lyophilization. (0.03mg, 3%) ^{1}H NMR (600 MHz, MeOD) δ 5.72 – 5.63 (1H, m, CH=C*H*-CH_2), 5.44 (1H, dd, *J* = 15.2, 7.7 Hz, CHOH-C*H*=CH), 4.94 (1H, d, *J* = 3.7 Hz, H-1, Gal-III), 4.40 (1H, d, *J* = 6.7 Hz, H-1, Gal-II), 4.29 (1H, d, *J* = 7.8 Hz, H-1, Glc-I), 4.25 (1H, t, *J* = 6.6 Hz), 4.17 (1H, dd, *J* = 10.1, 4.6 Hz), 4.05 (1H, t, *J* = 8.2 Hz, HOC*H*-CH=CH), 3.97 – 3.22 (22H, m), 2.24 – 2.11 (2H, m, *J* = 16.4, 14.4, 8.2 Hz, CH_2, Cer), 2.06 – 1.97 (2H, m, CH_2, Cer), 1.66 – 1.50 (2H, m, CH_2, Cer), 1.42 – 1.21 (46, m, *J* = 13.8 Hz, 23 x CH_2, Cer), 0.89 (6H, t, *J* = 6.9 Hz, 2x CH_3, Cer). ESI HRMS (*m/z*): $[M + Na]^+$ calcd for $C_{52}H_{97}NO_{18}$; 1046.6598 found 1046.6606.



(2S,3R,4E)-3-hydroxy-2-(hexadecanamido)octadec-4-enyl 2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (2b). Compound **2b** was synthesized from compound **2a** according to the general procedure for ceramide synthesis. The crude was purified by a 1mL C18 cartridge using H_2O: CH_3CN (1:0 to 0:1 v/v) as the eluent. The organic solvent was removed under N_2 flow and the residual water was removed by lyophilization. (0.1 mg, 10%) ^{1}H NMR (600 MHz, MeOD) δ 5.72 – 5.63 (1H, m, CH=C*H*-CH_2), 5.44 (1H, dd, *J* = 15.3, 7.9 Hz, HOCH-C*H*=CH), 4.93 (1H, d, *J* = 3.9 Hz, H-1, Gal-III), 4.61 (1H, d, *J* = 9.0 Hz, H-1, GalNAc-IV), 4.43 – 4.38 (1H, m, H-1, Gal-II), 4.29 (1H, d, *J* = 7.8 Hz, H-1, Glc-I), 4.28 – 4.23 (1H, m), 4.16 (2H, dd, *J* = 9.3, 5.1 Hz), 4.08 – 4.04 (1H, m, HOC*H*-CH=CH), 3.99 – 3.23 (24H, m), 2.22 – 2.13 (2H, m, CH_2, Cer), 2.05 – 1.97 (5H, m, *J* = 16.6 Hz, CH_2, Cer; CH_3, NHAc), 1.63 – 1.53 (2H, m, CH_2, Cer), 1.43 – 1.20 (46H, m, 23 x CH_2, Cer), 0.89 (6H, t, *J* = 6.7 Hz, 2 x CH_3, Cer). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{60}H_{110}N_2O_{23}$; 1227.7572 found 1227.7580.

(2S,3R,4E)-3-hydroxy-2-(hexadecanamido)octadec-4-enyl β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→3)-α-D-galactopyranosyl-(1→4)-β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside (**3b**). Compound **3b** was synthesized from compound **3a** according to the general procedure for ceramide synthesis. The crude was purified by a 1mL C4 cartridge using H_2O:MeOH (1:0 to 0:1 v/v) as the eluent. The organic solvent was removed under N_2 flow and the residual water was removed by lyophilization. (0.2 mg, 29%) ^{1}H NMR (600 MHz, MeOD) δ 5.68 (1H, dd, *J* = 14.4, 7.2 Hz, CH=C*H*-CH_2), 5.44 (1H, dd, *J* = 14.7, 7.4 Hz, HOCH-C*H*=CH), 4.93 (1H, d, *J* = 3.3 Hz, H-1, Gal-III), 4.70 (1H, d, *J* = 8.5 Hz, H-1, GalNAc-IV), 4.42 – 4.38 (1H, m, H-1, Gal-II), 4.33 (1H, d, *J* = 7.4 Hz, H-1, Gal-V), 4.29 (1H, d, *J* = 7.6 Hz, H-1, Glc-I), 4.27 – 4.22 (1H, m), 4.19 – 3.24 (33H, m), 2.24 – 2.17 (2H, m, CH_2, Cer), 2.06 – 1.95 (2H, m, CH_2, Cer), 1.91 (3H, s, CH_3, NHAc), 1.66 – 1.55 (2H, m, CH_2, Cer), 1.43 – 1.21 (46H, m, 23 x CH_2, Cer), 0.89 (6H, t, *J* = 7.0 Hz, 2 x CH_3, Cer). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{66}H_{120}N_2O_{28}$; 1389.8100 found 1389.8138.

References

1. Schnaar, R. L.; Kinoshita, T., Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), **2015-2017**.
2. Hannun, Y. A.; Obeid, L. M., *Nat Rev Mol Cell Biol* **2018**, *19* (3), 175-191.
3. D'Angelo, G.; Capasso, S.; Sticco, L.; Russo, D., *FEBS J* **2013**, *280* (24), 6338-53.
4. Keusch, J. J.; Manzella, S. M.; Nyame, K. A.; Cummings, R. D.; Baenziger, J. U., *J Biol Chem* **2000**, *275* (33), 25315-21.
5. Okajima, T.; Nakamura, Y.; Uchikawa, M.; Haslam, D. B.; Numata, S. I.; Furukawa, K.; Urano, T.; Furukawa, K., *J Biol Chem* **2000**, *275* (51), 40498-503.
6. Zhou, D.; Henion, T. R.; Jungalwala, F. B.; Berger, E. G.; Hennet, T., *J Biol Chem* **2000**, *275*(30), 22631-4.
7. Chang, W. W.; Lee, C. H.; Lee, P.; Lin, J.; Hsu, C. W.; Hung, J. T.; Lin, J. J.; Yu, J. C.; Shao, L. E.; Yu, J.; Wong, C. H.; Yu, A. L., *Proc Natl Acad Sci U S A.* **2008**, *105* (33), 11667-11672.
8. Saito, S.; Aoki, H.; Ito, A.; Ueno, S.; Wada, T.; Mitsuzuka, K.; Satoh, M.; Arai, Y.; Miyagi, T., *J Biol Chem* **2003**, *278* (29), 26474-9.
9. Senda, M.; Ito, A.; Tsuchida, A.; Hagiwara, T.; Kaneda, T.; Nakamura, Y.; Kasama, K.; Kiso, M.; Yoshikawa, K.; Katagiri, Y.; Ono, Y.; Ogiso, M.; Urano, T.; Furukawa, K.; Oshima, S.; Furukawa, K., *Biochem J* **2007**, *402* (3), 459-70.
10. Morace, I.; Pilz, R.; Federico, G.; Jennemann, R.; Krunic, D.; Nordstrom, V.; von Gerichten, J.; Marsching, C.; Schiessl, I. M.; Muthing, J.; Wunder, C.; Johannes, L.; Sandhoff, R.; Grone, H. J., *Kidney Int* **2019**, *96* (2), 327-341.
11. Brimble, S. N.; Sherrer, E. S.; Uhl, E. W.; Wang, E.; Kelly, S.; Merrill, A. H., Jr.; Robins, A. J.; Schulz, T. C., *Stem Cells* **2007**, *25* (1), 54-62.
12. Levery, S. B.; Salyan, M. E.; Steele, S. J.; Kannagi, R.; Dasgupta, S.; Chien, J. L.; Hogan, E. L.; van Halbeek, H.; Hakomori, S., *Arch Biochem Biophys.* **1994**, *312* (1), 125-134.
13. Yu, A. L.; Hung, J. T.; Ho, M. Y.; Yu, J., *Stem Cells Dev* **2016**, *25* (20), 1532-1548.
14. Shimada, S.; Ito, A.; Kawasaki, Y.; Kakoi, N.; Taima, T.; Mitsuzuka, K.; Watanabe, M.; Saito, S.; Arai, Y., *Prostate Cancer Prostatic Dis* **2014**, *17* (2), 199-205.
15. Sonnino, S.; Chiricozzi, E.; Grassi, S.; Mauri, L.; Prioni, S.; Prinetti, A., *Prog Mol Biol Transl Sci* **2018**, *156*, 83-120.
16. Russo, D.; Parashuraman, S.; D'Angelo, G., *Int J Mol Sci* **2016**, *17* (10).
17. Mahfoud, R.; Manis, A.; Binnington, B.; Ackerley, C.; Lingwood, C. A., *J Biol Chem* **2010**, *285* (46), 36049-59.
18. Ledeen, R. W.; Kopitz, J.; Abad-Rodriguez, J.; Gabius, H. J., *Prog Mol Biol Transl Sci* **2018**, *156*, 289-324.
19. O'Reilly, M. K.; Paulson, J. C., In *Glycomics*ed., Vol., **2010**, pp 343-363.
20. Rapoport, E.; Mikhalyov, I.; Zhang, J.; Crocker, P.; Bovin, N., *Bioorg Med Chem Lett.* **2003**, *13* (4), 675-678.
21. Lopez, P. H.; Schnaar, R. L., *Curr Opin Struct Biol* **2009**, *19* (5), 549-57.
22. Kawasaki, Y.; Ito, A.; Withers, D. A.; Taima, T.; Kakoi, N.; Saito, S.; Arai, Y., *Glycobiology* **2010**, *20* (11), 1373-1379.
23. Ito, A.; Handa, K.; Withers, D.; Satoh, M.; Hakomori, S., *FEBS Letters* **2001**, *498*, 116-120.
24. Yu, H.; Li, Y.; Zeng, J.; Thon, V.; Nguyen, D. M.; Ly, T.; Kuang, H. Y.; Ngo, A.; Chen, X., *J Org Chem* **2016**, *81* (22), 10809-10824.
25. Cheng, C. W.; Wu, C. Y.; Hsu, W. L.; Wong, C. H., *Biochemistry* **2019**.
26. Jeon, I.; Iyer, K.; Danishefsky, S. J., *J Org Chem* **2009**, *74* (21), 8452-5.
27. Zhu, T.; Boons, G.-J., *Angew Chem Int Ed* **1999**, *38* (23), 3495-3497.
28. t Hart, I. M. E.; Li, T.; Wolfert, M. A.; Wang, S.; Moremen, K. W.; Boons, G. J., *Org Biomol Chem* **2019**, *17* (31), 7304-7308.
29. Wang, Z.; Gilbert, M.; Eguchi, H.; Yu, H.; Cheng, J.; Muthana, S.; Zhou, L.; Wang, P. G.; Chen, X.; Huang, X., *Adv Synth Catal* **2008**, *350* (11-12), 1717-1728.
30. Li, P. J.; Huang, S. Y.; Chiang, P. Y.; Fan, C. Y.; Guo, L. J.; Wu, D. Y.; Angata, T.; Lin, C. C., *Angew Chem Int Ed Engl* **2019**, *58* (33), 11273-11278.
31. Danishefsky, S. J.; Shue, Y. K.; Chang, M. N.; Wong, C. H., *Acc Chem Res* **2015**, *48* (3), 643-52.
32. Hashimoto, N.; Ito, S.; Tsuchida, A.; Bhuiyan, R. H.; Okajima, T.; Yamamoto, A.; Furukawa, K.; Ohmi,

Y.; Furukawa, K., *J Biol Chem* **2019**, *294* (28), 10833-10845.
33. Kiarash, A.; Boyd, B.; Lingwood, C. A., *J. Biol. Chem.* **1994**, *269* (15), 11138-11346.
34. Hunter, C. D.; Guo, T.; Daskhan, G.; Richards, M. R.; Cairo, C. W., *Chem Rev* **2018**, *118* (17), 8188-8241.
35. Lassaletta, J. M.; Carlsson, K.; Garegg, P. J.; Schmidt, R. R., *J Org Chem* **1996**, *61* (20), 6873-6880.
36. Gege, C.; Schumacher, G.; Rothe, U.; Schmidt, R. R.; Bendas, G., *Carbohydr Res* **2008**, *343* (14), 2361-8.
37. Nakashima, S.; Ando, H.; Saito, R.; Tamai, H.; Ishida, H.; Kiso, M., *Chem Asian J* **2012**, *7* (5), 1041-51.
38. Imamura, A.; Kiso, M., *Methods Mol Biol* **2018**, *1804*, 293-310.
39. Liu, Y.; Wen, L.; Li, L.; Gadi, M. R.; Guan, W.; Huang, K.; Xiao, Z.; Wei, M.; Ma, C.; Zhang, Q.; Yu, H.; Chen, X.; Wang, P. G.; Fang, J., *European J Org Chem* **2016**, *2016* (25), 4315-4320.
40. Rich, J. R.; Withers, S. G., *Angew Chem Int Ed Engl* **2012**, *51* (34), 8640-3.
41. Santra, A.; Li, Y.; Yu, H.; Slack, T. J.; Wang, P. G.; Chen, X., *Chem Commun (Camb)* **2017**, *53* (59), 8280-8283.
42. Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraser-Reid, B., *J Am Chem Soc* **1988**, *110* (16), 5583-5584.
43. Franchini, L.; Matto, P.; Ronchetti, F.; Panza, L.; Barbieri, L.; Costantino, V.; Mangoni, A.; Cavallari, M.; Mori, L.; De Libero, G., *Bioorg Med Chem* **2007**, *15* (16), 5529-36.
44. Johnson, K. F., *Glycoconjugate Journal* **1999**, *16* (2), 141-146.
45. Randriantsoa, M.; Drouillard, S.; Breton, C.; Samain, E., *FEBS Lett* **2007**, *581* (14), 2652-6.
46. Tsai, T. I.; Lee, H. Y.; Chang, S. H.; Wang, C. H.; Tu, Y. C.; Lin, Y. C.; Hwang, D. R.; Wu, C. Y.; Wong, C. H., *J Am Chem Soc* **2013**, *135* (39), 14831-9.
47. Shao, J.; Zhang, J.; Kowal, P.; Lu, Y.; Wang, P. G., *Biochem Biophys Res Commun* **2002**, *295* (1), 1-8.
48. Gilbert, M.; Brisson, J. R.; Karwaski, M. F.; Michniewicz, J.; Cunningham, A. M.; Wu, Y.; Young, N. M.; Wakarchuk, W. W., *J Biol Chem* **2000**, 3896-3906.
49. Gold, H.; Boot, R. G.; Aerts, J. M. F. G.; Overkleeft, H. S.; Codée, J. D. C.; van der Marel, G. A., *European Journal of Organic Chemistry* **2011**, *2011* (9), 1652-1663.
50. Zhou, G.; Liu, X.; Su, D.; Li, L.; Xiao, M.; Wang, P. G., *Bioorg Med Chem Lett* **2011**, *21* (1), 311-4.
51. Lingwood, C. A., *Glycoconjugate Journal* **1996**, *13* (4), 495-503.
52. Yu, H.; Santra, A.; Li, Y.; McArthur, J. B.; Ghosh, T.; Yang, X.; Wang, P. G.; Chen, X., *Org Biomol Chem* **2018**, *16* (22), 4076-4080.
53. Bernatchez, S.; Gilbert, M.; Blanchard, M. C.; Karwaski, M. F.; Li, J.; Defrees, S.; Wakarchuk, W. W., *Glycobiology* **2007**, *17* (12), 1333-43.
54. Malekan, H.; Fung, G.; Thon, V.; Khedri, Z.; Yu, H.; Qu, J.; Li, Y.; Ding, L.; Lam, K. S.; Chen, X., *Bioorg Med Chem* **2013**, *21* (16), 4778-85.

Chapter 4

Chemical synthesis of an *O*-glycan analog of sulfoglycolipid SM1a: SM1-core3

A potential glycan antigen for cancer-specific monoclonal antibody HAE3

Abstract: Sulfoglycolipids (SGLs) are important cancer biomarkers, making them interesting synthetic targets. A microarray study revealed SGL SM1a as an antigen for anti-epiglycanin antibody (AE3 or HAE3). It was surprising that a glycolipid antigen for antibody HAE3 was identified, since HAE3 is known to bind mucin glycoproteins. *O*-glycan epitopes similar to SM1a have been found on mucin glycoproteins, having a core3 (GlcNAcβ1,3GalNAcα-Ser/Thr) glycan residue in common. Therefore, a hybrid structure of SM1a and the core3 disaccharide was proposed as a potential ligand for HAE3, here named as SM1-core3 (Galβ1,3GalNAcβ1,4Galβ1,3GlcNAcβ1,3GalNAcα). We report here the chemical synthesis of the protected SM1-core3 glycan having an α-linked aminopentyl group at the reducing end for immobilization. A block synthetic approach (2 + 3) was chosen and the respective disaccharide donor and trisaccharide acceptor were synthesized from five monosaccharide building blocks. After deprotection and sulfation of SM1-core 3, this sulfoglycolipid analog will be studied for its HAE3 binding properties in a glycan microarray

Introduction

Cancer cells often express aberrant glycans, such as truncated structures, or having altered fucosylation or increased sialylation.[1] Over the years, many tumor associated carbohydrate antigens (TACAs) have been discovered.[2] Glycoprotein TACAs, such as (sialyl-)Thomson-nouveau (sTn), (sialyl-)Thomsen-Friedenreich (sT) and (sialyl-)Lewis (sLe) antigens[3,4] have received considerable attention. Furthermore, glycosphingolipid and sulfoglycolipid antigens have been found to be key players in various cancers. Examples include ganglio-series GM3, SM3, GM2, GD3, GD2 and SB1a and globo-series Gb3, Gb4, Gb5, Globo-H, MSGb5 and DSGb5.[5,6,7] Antibodies against these carbohydrate antigens are attractive candidates for cancer cell detection and vaccine development.[2,8,9]

Epiglycanin is a mucin-type glycoprotein that was first isolated from mouse mammary carcinoma cells.[10,11] The anti-epiglycanin antibody (AE3, later named HAE3) was raised against epiglycanin. HAE3 cross-reacts with human epithelial carcinoma-associated antigen (HCA) in serum samples of patients with epithelial carcinomas such as; lung, bladder, prostate, esophagus, ovarian and breast cancer.[12,13] However, little is known about the specific glycoantigen that HAE3 recognizes. A carbohydrate microarray study revealed that HAE3 binds similar structures as the lectin PNA (specific to Galβ1,3GalNAc), but did not show any binding to blood group or Lewis antigens. Surprisingly, the sulfated glycolipid SM1a showed strong binding to HAE3.[14]

SM1a is the sulfoglycolipid (SGL) analog of ganglio-series glycosphingolipid GM1a (Fig 1A). GM1a has a sialic acid (α2,3-)linked to the inner galactose residue, while SM1a has a (3-*O*-)sulfate group on the C-3 of this galactose. SM1a was first isolated from rat and green monkey in kidney cells and is proposed to be a precursor of SB1a (Fig 1A).[15,16] Little is known about the role of SM1a and its analogs in healthy or cancerous cells. Nevertheless, other SGLs have been associated with different types of cancer. In breast cancer, upregulated 3-*O*-sulfotransferase activity was detected towards core 3 *O*-glycans, lactosamine and globo-sequences.[17] Furthermore, SM3 (sulfated analog of GM3) and disulfated SB1a have been reported to be upregulated in liver carcinoma.[18,19]

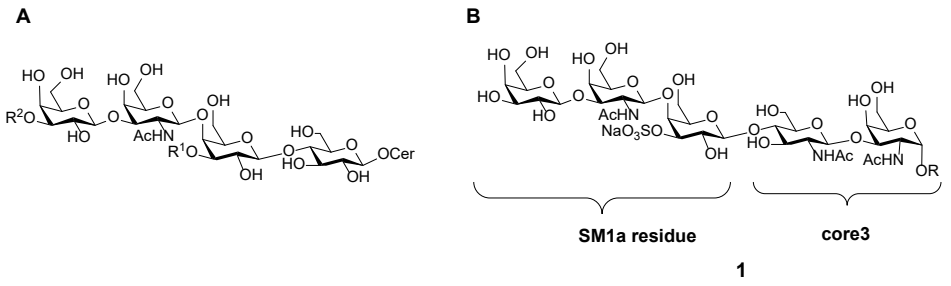


Figure 1. (**A**) **SM1a**: $R^1 = SO_3^-$, R^2 = H, SB1a, **SB1a**: $R^1 = R^2$: SO_3^- and **GM1a**: R^1 = α2,3Neu5Ac, R^2 = H | (**B**) Proposed **SM1-core3** glycan[14] (R = Ser or Thr). Target of this chapter: **1**, R = $(CH_2)_5NH_2$.

The HAE3 antibody binds mucin-type *O*-glycans, which are mainly found in the gut. In an effort to find sulfated oligosaccharide sequences on *O*-glycans in the colon that are similar to SM1a, Sd(a)/Cad-antigen-like epitopes (GalNAcβ1,4(NeuAcα2,3)Gal) were found, which all contain a core3 *O*-glycan sequence (GlcNAcβ1,3GalNAcαSer/Thr).[20] From these findings a hybrid structure of SM1a and a core3 disaccharide was proposed as a potential ligand for HAE3, here named SM1-core3 (Figure 1B).[14]

SM1-core 3 has not been identified or isolated from mucin or other sources. Isolation of intact sulfoglycolipids and its analogs is tedious due to the presence of both charged and lipophilic residues. Furthermore, the isolation process often results in heterogenous mixtures. Therefore, a synthetic approach is required to obtain SM1-core 3. Chemical synthesis of SM1a with two different sphingosine compositions has been described.[21,22] However, no sulfoglycolipid *O*-glycan hybrids have been synthesized. A principle difference between *O*-glycans and SGLs are their characteristic reducing glycan moieties. The reducing moiety of *O*-glycans consists of GalNAc α-linked to serine (Ser) or threonine (Thr) of peptides, whereas complex glycolipids contain lactosyl residues β-linked to ceramide.[23,24] Therefore, a different chemical synthetic strategy is required to synthesize a SGL O-glycan hybrid compared to SGL synthesis. First of all, the core3 reducing moiety require chemical glycosylation between GlcNAc and GalNAc, whereas SGLs can be synthesized from commercially available lactose. Furthermore, terminal 1,2-*cis* (in this case α-) anomeric selectivity is required to mimic the natural *O*-glycan linkage. Chemical synthesis of 1,2-*cis* linkages is more challenging than 1,2-*trans* anomers (β in the case of lactosylceramide), where neighboring group participation can be used to direct the stereoselective outcome.[25]

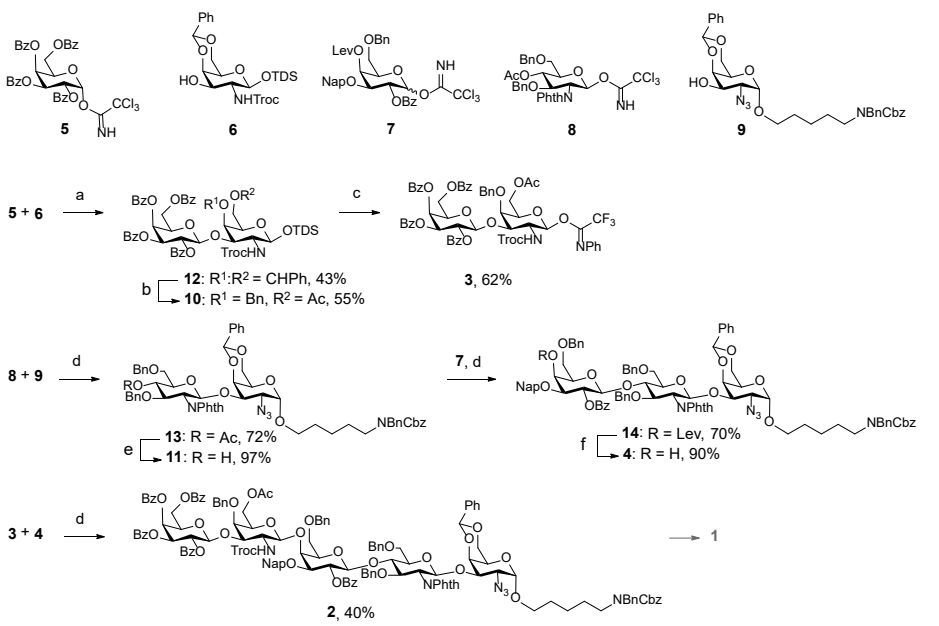
Although HAE3 is known to bind tumor glycan markers, further elucidation on the specificity to various glycan structures is required. Here we report the first chemical synthesis of a protected form of the SM1-core 3 glycan; an potential ligand for monoclonal antibody HAE3. Global deprotection and 3-*O*-sulfation will provide the desired synthetic target. An α-anomeric aminopentyl spacer has been incorporated at the reducing end to enable facile immobilization for microarray binding studies. The results of such a binding study will provide insight in the binding specificity of HAE3. Furthermore, it will encourage the search for more relevant sulfoglycolipid (anolog)s as possible cancer-associated glycan epitopes. Highly defined cancer-associated antigens will provide more insight in cancer cell mechanisms and serve as handles for cancer specific therapies.

Results and Discussion

Chemical synthesis

The chemical synthesis of protected pentasaccharide **2** was designed in such way that the position of the 3-*O*-sulfate is protected with an *O*-2-naphthylmethylether (Nap) as an orthogonally cleavable group (Scheme 1). All glycosidic linkages are in the β-configuration, except the reducing aminopentyl linker, resembling the α-linkage (to Ser/Thr) found in *O*-glycans. Appropriate C-2 ester protecting groups (*O*Bz, *N*Phth, *N*HTroc) were chosen to install β-anomeric linkages by neighboring group participation during glycosylation. A block synthesis approach was used, in which coupling of disaccharide **3**

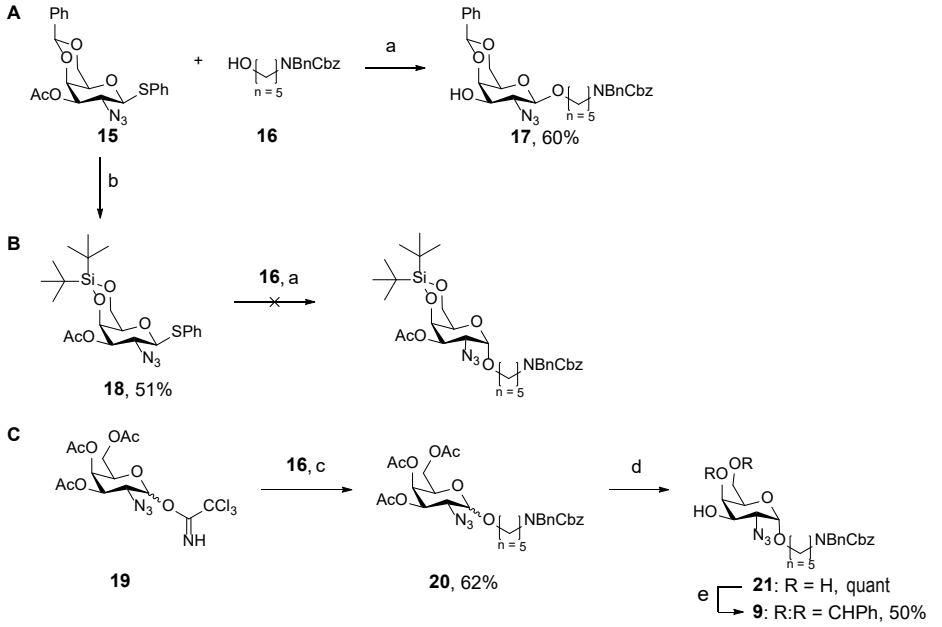
and trisaccharide **4** provides pentasaccharide **2**. Disaccharide **3** is similar to the Galβ1,3-GalNAc disaccharides synthesized before (Chapter 2 of this Thesis). Trisaccharide **4** (Galβ1,4GlcNAcβ1,3-GalNAc-α-$(CH)_2NH_2$) was synthesized from three monosaccharide building blocks and an aminopentyl linker. Monosaccharide building blocks **5** - **9** were synthesized from the respective free reducing monosaccharides.



Scheme 1. Chemical synthesis of SM1-core3. Reagents and conditions: (a) TMSOTf, CH_2Cl_2, 4 Å MS, -40 °C; (b) i. Et_3SiH, Cl_2PhB, CH_2Cl_2, 4 Å MS, -68 °C, ii. Ac_2O, pyridine; (c) i. HF·pyridine, pyridine, ii. $CF_3C(NPh)Cl$, Cs_2CO_3, CH_2Cl_2, 0 °C; (d) TMSOTf, CH_2Cl_2, 4 Å MS, -70 °C; (e) NaOMe, CH_2Cl_2/MeOH; (f) NH_2NH_2HOAc, CH_2Cl_2/MeOH, 50 °C.

Benzylidene protected $GalN_3$ donor **15** was chosen to form the α-linkage with the protected aminopentanol linker **16**[26] (Scheme 2A). Surprisingly, only the β-anomer of **17** was formed, even though the C-2 ester was expected to provide neighboring group participation. This was confirmed by decoupled HSQC NMR spectroscopy ($^1J_{CH}$ = 162.0 Hz). The high β-anomeric selectivity can be attributed to a combination of high reactivity of the nucleophile **16** and stabilization of the intermediate anomeric triflate in presence of a C-2 azide. The C-2 azide has an electron withdrawing effect, thereby stabilizing a covalent α-triflate intermediate. A strong nucleophile can replace the α-triflate, providing the β-product.[27] To incorporate an α-directing protecting group, the benzylidene acetal of donor **15** was replaced in two steps for a di-*tert*-butylsilylene group (**18**, Scheme 2B). However, this donor proved to be highly unreactive after activation by NIS/TfOH and no product was formed. Also, no donor hydrolysis was observed and intact donor **18** was recovered from the reaction mixture. Finally, glycosylation of trichloroactimidate **19**[28] and protected aminopentanol acceptor **16** in the presence of thiophene provided an

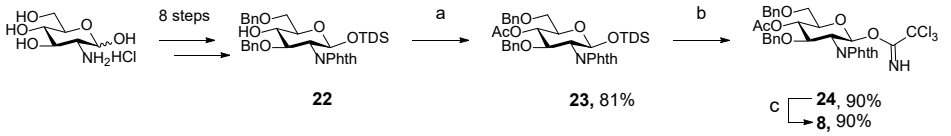
inseparable α/β mixture of compound **20** (Scheme 2C). After deacetylation under Zemplén conditions (*cat.* NaOMe in MeOH) compound **21** was obtained. Protection of the C-4 and C-6 hydroxyls of compound **21** by a benzylidene acetal provided **9** as a separable mixture of α/β anomers. The desired α-anomer **9** was isolated by silica column chromatography.



Scheme 2. Chemical glycosylations of $GalN_3$ donors **15**, **18** and **19** with acceptor **16** were performed to obtain building block **9** as the α-anomer. (**A**) Donor **15** solely provided the undesired β-anomer. (**B**) Replacement of the benzylidene acetal by an α-directing di-*tert*-butyl-silane protecting group, provided donor **18**. Chemical glycosylation of donor **18** with acceptor **16** did not provide any product. (**C**) Glycosylation of acetylated donor **19** with acceptor **16** provided an α/β mixture (1/1) of compound **20**. Deacetylation of **20** provided compound **21** and subsequent protection of the C-4 and C-6 hydroxyl by a benzylidene acetal provided a separable mixture of anomers. Purification by silica column chromatography provided compound **9**. Reagents and conditions: (a) NIS, TfOH, -20 °C, 4 Å MS; (b) i. HF·Pyr, pyridine, ii. $(t\text{-Bu})_2Si(OTf)_2$, pyridine, DMF, - 40 °C; (c) TMSOTf, thiophene, Et_2O, 0 °C, 4 Å MS; (d) NaOMe, MeOH; (e) $PhCH(OMe)_2$, $pTsOH\cdot H_2O$, CH_3CN.

Protected glucosamine **22** was synthesized from glucosamine·HCl in eight steps as described before.[29] Starting from $GlcNH_2\cdot HCl$, phthalimide was chosen as a protecting group for the amine, since it enables neighboring group participation during glycosylation and sufficiently stabile towards bases such as sodium hydride required for benzylation of hydroxyls. The subsequent acetylation of all hydroxyls, anomeric deacetylation by hydrazine acetate and protection with *tert*-hexyl-*di*-methyl-silyl (TDS) group were performed as described (40%, 4 steps). Then, all acetyl esters were cleaved and the 4- and 6- hydroxyls were protected as benzylidene acetal, after which the C-3 hydroxyl was protected with a benzyl ether (70%, 3 steps). The benzylidene acetal was opened in a regioselective manner by Et_3SiH and TfOH in DCM at -70 °C to expose the C-4 hydroxyl (**22**, 52%).

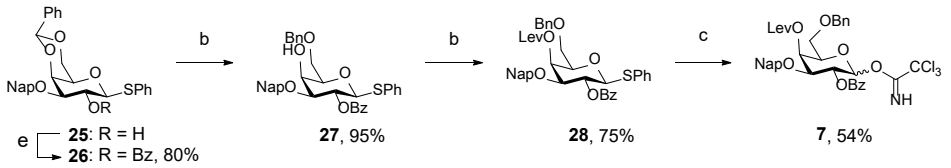
Compound **22** was acetylated by treatment with acetic anhydride in pyridine to provide compound **23** (81%). The anomeric TDS group was cleaved by HF·Pyridine to provide lactol **24**, which was converted into trichloroacetimidate **8** by reaction with trichloroacetonitrile in the presence of Cs_2CO_3 (81%, 2 steps: Scheme 3).



Scheme 3. Chemical synthesis of GlcNAc-II trichloroacetimidate donor **8**. Reagents and conditions: (a) i. Ac_2O, Pyridine, ii. HF·Pyridine, Pyridine; (b) i. $NCCCl_3$, DBU, CH_2Cl_2, 0 °C; (c) $ClC(NPh)CF_3$, Cs_2CO_3, CH_2Cl_2, 0 °C to r.t., 4 Å MS.

Chemical glycosylation of donor **8** and acceptor **9** provided disaccharide **13** in a high yield (72%). The acetyl esters was easily removed under Zemplén conditions (*cat*. NaOMe, MeOH) to form disaccharide acceptor **11** (Scheme S1).

Galactose donor **7** contains flexible protecting groups to furnish neighbouring group participation (OBz), acceptor formation (OLev) and a sulfation site (ONap) (Scheme 4). Starting from compound **25**[30], the C-2 hydroxyl was protected by a benzoyl ester (**26**, 80%). Next, selective opening of the benzylidene by TFA and Et_3SiH in DCM at 0 °C, smoothly provided compound **27**. Protection of the C-4 hydroxyl of compound **27** by treatment with levulinic acid in the presence of EDC/DMAP/Et_3N resulted in 50% conversion to **28**. Silica column purification provided **28** and starting material **27**, which was submitted for Lev protection once more, providing a total yield of 75% of compound **28**. The axial hydroxyl of galactose is known to have a low reactivity, especially when bulky electrophiles are used.[31]



Scheme 4. Chemical synthesis of Gal-III donor **7**. Reagents and conditions: (a) BzCl, pyridine; (b) i. TFA, CH_2Cl_2, 0 °C, ii. LevOH, EDC, DMAP, Et_3N; (c) i. TFA, NIS, H_2O ii. DBU, CH_2Cl_2, 0 °C.

Initially, thioglycoside donor **28** was used for a glycosylation with acceptor **11** employing NIS/TfOH as the promotor system at 0 °C or -30 °C. However, in both cases an inseparable anomeric mixture of **14** was formed. Therefore, donor **7** was synthesized from compound **28** by subsequent hydrolysis of the anomeric thiol ether by NIS/TFA/H_2O and formation of the trichloroacetimidate (Scheme 4). Glycosylation of donor **7** with acceptor **11** at -70 °C provided **14** as product, which was confirmed to be the β-anomer by 1D and 2D NMR spectroscopy (Scheme S2).

Trisaccharide acceptor **4** was obtained after removal of the Lev-ester by hydrazine acetate. It was noted that heating (50 °C) was required to drive the reaction to completion. Complete conversion to product **4** was required, since a mixture of compounds **14** and **4** was inseparable by silica column chromatography.

Disaccharide **12** was synthesized from donor **5** and acceptor **6** in a similar fashion as in Chapter 2 of this Thesis.[32] Earlier studies showed that the benzylidene acetal on a Galβ1,3GalNAc disaccharide donor provided low reactivity and poor anomeric selectivity.[32] Also, replacement of the benzylidene for 4- and 6-O acetyl esters decreased donor reactivity drastically, providing only traces of product (data not shown). Therefore, the benzylidene acetal was selectively opened by treatment with Cl_2BPh and Et_3SiH to obtain a C-4 benzyl ether and C-6 hydroxyl. The exposed C-6 hydroxyl was protected by an acetyl ester providing intermediate disaccharide **10** (55%, 2 steps). The anomeric silyl protecting group (TDS) was cleaved by treatment with HF·pyridine and the intermediate lactol was converted into the *N*-phenyl trifluoroacetimidate, providing donor **3** (62%, 2 steps; Scheme S3). Glycosylation of disaccharide donor **3** and trisaccharide acceptor **4** successfully provided pentasaccharide **2** (40%).

Deprotections and sulfation of compound **2** will provide target compound **1**; a sulfoglycolipid-*O*-glycan hybrid. SM1a-core 3 will be screened in a microarray binding study as the potential cancer specific antigen HAE3.

Conclusions

New cancer-associated antigens have been discovered among the sulfoglycolipids, but little is known about their mechanisms of action. Sulfoglycolipid SM1a was detected as a ligand for the cancer specific antibody HAE3 in a glycan microarray screening. It is known that HAE3 binds mucin-type *O*-glycans and a search for epitopes similar to SM1a provided a set of core 3 *O*-glycans. Therefore, a SM1-core3 hybrid was proposed as a ligand for HAE3. Here we report the first total synthesis of a protected SM1-core3 glycan with a reducing α-aminopentyl linker for immobilization. After global deprotection and sulfation, a microarray study will reveal the HAE3 binding properties to SM1-core 3. A binding interaction will support the presence of SM1-core 3 and similar sulfated *O*-glycans in the mucin glycoproteins. This will encourage the search for sulfated mucin glycans as a new class of cancer biomarkers. Furthermore, the optimized synthetic strategy will be useful for future synthesis of sulfoglycolipids and its analogs.

Experimental section

Chemical synthesis

NMR nomenclature

The monosaccharides of glycan SM1-core3 have been labeled as shown in **Figure S1.** Starting from the reducing end these were labeled GalNAc-I, GlcNAc-II, Gal-III, GalNAc-IV and Gal-V.

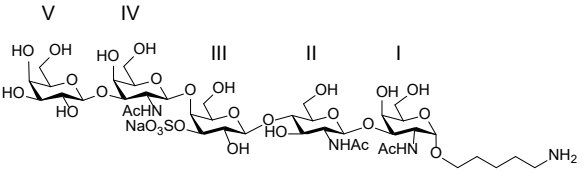
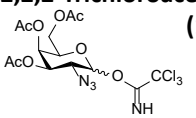


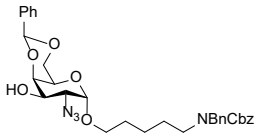
Figure S1. Monosaccharide labeling system for SM1-core3.

2,2,2-Trichloroacetimidate 3,4,6-*O*-acetyl-*O*-2-azido-2-deoxy-α/β-D-galactopyranoside (15). Compound **15** was synthesized as described before.[28]



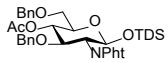
***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentan-1-ol (16).** Compound **16** was synthesized according to previous procedures.[26]

***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl *O*-2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside (9).** A mixture of acceptor **16** (1.93 g, 5.9 mmol), donor **15** (1.87 g, 3.9 mmol), thiophene (6.3 mL, 78.7 mmol) and 4 Å molecular sieves was stirred in Et_2O (40 mL) for 20 min. The mixture was cooled in an ice bath and TMSOTf (71 µL, 0.4 mmol) was added to the mixture. The reaction was monitored by TLC (hexanes: EtOAc 6.5 : 3.5 v/v) and more TMSOTf (2 x 140 µL) was added over time to drive the reaction to completion. After 2 h, Et_3N was added, the mixture was filtered, and the filtrate concentrated *in vacuo* and separated between DCM and sat. aq. $NaHCO_3$. The organic layer was washed with brine, dried (Na_2SO_4), filtered and concentrated *in vacuo*. The resulting residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 24 : 8 v/v) as the eluent affording the product as an α/β mixture (1.57 g, 62 %). The obtained α/β mixture of compound **20** was dissolved in MeOH (25 mL) and freshly prepared NaOMe (2.5 mL) was added. After 1 h H^+ Amberilte resin was added, the mixture was filtered and concentrated *in vacuo*. Compound **21** was used in the next reaction without further purification. Benzaldehyde dimethyl acetal (444 µL, 2.96 mmol) and pTsOH·H_2O (43 mg, 0.23 mmol) were added to a solution of compound **21** (1.17 g, 2.27 mmol) in CH_3CN (25 mL). After 1 h the mixture was quenched with Et_3N, concentrated *in vacuo* and the obtained residue was purified by silica column chromatography using toluene:EtOAc (1:0 to 8:2 v/v) as the eluent to obtain compound **9** (α-anomer: 688 mg, 50 %; β-anomer: 490 mg, 36 %). The NMRs of the α-anomer are depicted below. ^1H NMR (600 MHz, $CDCl_3$) δ 7.52 – 7.46 (2H, m, H-Ar), 7.42 – 7.10 (13H, m, H-Ar), 5.57 (1H, s, C*H*-C_6H_5), 5.18 (2H, d, *J* = 20.1 Hz, CH_2), 4.96 (1H, d, *J* = 13.9 Hz, H-1), 4.50 (2H, d, *J* = 12.0 Hz, CH_2), 4.31 – 4.22 (2H, m, H-4; H-6a), 4.21 – 4.11 (1H, m, H-3), 4.11 – 4.02 (1H, m, H-6b), 3.78 – 3.58 (2H, m, H-5; C*H*H, pentyl), 3.53 (1H, d, *J* = 8.8 Hz, H-2), 3.49 – 3.34 (1H, m, CH*H*, pentyl), 3.34 – 3.13 (2H, m, CH_2, pentyl), 2.42 (1H, d, *J* = 10.8 Hz, O*H*), 1.68 – 1.46 (4H, m, 2x CH_2, pentyl), 1.41 – 1.24 (2H, m, 2x CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 137.3 (C=O, Cbz), 129.4, 128.6, 128.5, 128. 3, 127.9, 127.8, 127.2, 126.2, 101.3 (*C*H-C_6H_5), 98.7 (C-1), 75.5 (C-4), 69.3 (C-6), 68.4 (CH_2, pentyl), 67.4 (C-3), 67.2 (CH_2), 62.8 (C-5), 60.7 (C-2), 50.3 (CH_2), 47.1 (CH_2, pentyl), 29.1 (CH_2, pentyl), 27.4 (CH_2, pentyl), 23.4



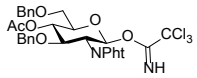
(CH_2, pentyl). ESI HRMS (*m/z*): $[M + Na]^+$ calcd for $C_{33}H_{38}N_4O_7$; 625.2638 found 625.2615.

Dimethylthexylsilyl 4-*O*-acetyl-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (23). Compound **22** was synthesized according to a reported procedure.[29]

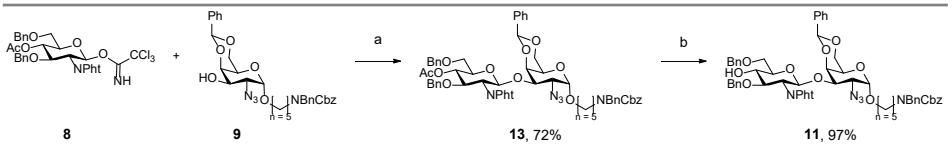


Compound **22** (3.48 g, 5.5 mmol) was dissolved in pyridine (10 mL) and Ac_2O (1.04 mL) was added to the stirring mixture. The reaction mixture was stirred for 1 h and concentrated *in vacuo*. The obtained oil was dissolved in EtOAc, washed with sat. aq. $NaHCO_3$ and brine, dried (Na_2SO_4), filtered, concentrated *in vacuo.* The obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 43 : 7) as the eluent to afford compound **23** (3.00 g, 81 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.88 – 7.66 (4H, m, H-Ar), 7.39 – 7.24 (5H, m, H-Ar), 7.04 – 6.89 (5H, m, H-Ar), 5.37 (1H, d, *J* = 8.0 Hz, H-1), 5.12 (1H, t, *J* = 9.5 Hz, H-4), 4.59 (1H, d, *J* = 12.0 Hz, CH*H*, Bn), 4.55 (s, 2H, CH_2), 4.45 (1H dd, H-3), 4.35 (1H, d, *J* = 12.0 Hz, CH*H*, Bn), 4.21 (1H, dd, *J* = 10.9, 8.1 Hz, H-2), 3.81 – 3.71 (1H, m, H-5), 3.65 – 3.56 (2H, m, H-6), 1.95 (3H, s, OAc), 1.42 – 1.30 (1H, m, C*H*, TDS), 0.66 – 0.54 (12H, m, 4x CH_3, TDS), 0.12 (3H, s, CH_3-Si), -0.03 (3H, s, CH_3-Si). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.7 (C, OAc), 138.1 (C, Phth), 137.8 (C, Phth), 133.8, 128.3, 128.1, 127.8, 127.7, 127.6, 127.4, 123.1 (CH, Phth), 93.3 (C-1), 76.7 (C-3), 73.5 (CH_2, Bn), 73.5 (CH_2, Bn), 73.5 (C-5), 72.6 (C-4), 70.0 (C-6), 57.5 (C-2), 33.8 (*C*H, TDS), 24.5 (C, TDS), 20.9 (CH_3, OAc) , 19.8 (CH_3, TDS), 19.7 (CH_3, TDS), 18.3 (CH_3, TDS), 18.2 (CH_3, TDS) ,-1.8 (CH_3-Si) , -3.9 (CH_3-Si). ESI HRMS (*m/z*): $[M + Na]^+$ calcd for $C_{38}H_{47}NO_8Si$; 696.2969 found 696.2930

2,2,2-Trichloroacetimidate 4-*O*-acetyl-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (8). HF·pyridine (70% HF, 3.0 mL) was added to a stirring solution of



compound **23** (3.0 g, 4.5 mmol) in pyridine (25 mL) in a plastic round bottom flask. After overnight stirring, sat. aq. $NaHCO_3$ was added to the mixture and stirring was contintued for another 30 min. The reaction mixture was diluted with EtOAc, washed with sat. aq. $NaHCO_3$, brine and dried (Na_2SO_4), filtered, and the filtrate concentrated *in vacuo* to give a residue that was co-evaporated with toluene (3x) and DCM (1x). Crystallization from DCM/Hexane afforded compound **24** (2.14 g, 90 %). 2,2,2-Trichloroacetonitrile (471 µL, 4.7 mmol) and DBU (28 µL, 0.19 mmol) were added to a stirring solution of compound **24** (500 mg, 0.94 mmol) in DCM (10 mL) at 0 °C. After 1 h the reaction mixture was concentrated *in vacuo* and the obtained residue was directly purified by silica column chromatography using Hexane(1% Et_3N):EtOAc (1:0 to 3:1 v/v) as the eluent to afford compound **8** (572 mg, 90 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.58 (1H, s, N*H*), 7.76 – 7.62 (4H, m, H-Ar), 7.38 – 7.30 (5H, m, H-Ar), 7.06 – 6.85 (5H, m, H-Ar), 6.43 (1H, d, *J* = 8.5 Hz, H-1), 5.25 (1H, dd, *J* = 10.0, 8.6 Hz, H-4), 4.63 (1H, d, *J* = 12.1 Hz, CH*H*, Bn), 4.60 – 4.50 (4H, m, H-3; CH_2, Bn; H-2), 4.36 (1H, d, *J* = 12.1 Hz, CH*H*, Bn), 3.95 (1H, ddd, *J* = 10.0, 5.0, 3.6 Hz, H-5), 3.70 – 3.61 (2H, m, H-6), 1.94 (3H, s, OAc). ^{13}C NMR (151 MHz, $CDCl_3$) δ 169.5 (C, OAc), 160.8 (C=NH), 137.8 (C, Phth), 137.6 (C, Phth), 134.0, 131.4, 128.4, 128.1, 128.0, 127.8, 127.7, 127.5, 123.4, 93.9 (C-1), 90.3 (CCl_3), 76.7 (C-3), 74.5 (C-5), 74.0 (CH_2, Bn), 73.5 (CH_2, Bn), 71.9 (C-4), 69.0 (C-6), 54.4 (C-2), 20.9 (CH_3, OAc). Open anomer: ESI HRMS (*m/z*): $[M + Na]^+$ calcd for $C_{30}H_{29}NO_8$; 554.1791 found 554.1818.



Scheme S1. Chemical glycosylation of GlcNAc-II donor **8** and GalNAc-I acceptor **9** and formation of disaccharide acceptor **11**. Reagents and conditions: (a) TMSOTf, CH_2Cl_2, -70 °C; (b) NaOMe, CH_2Cl_2/MeOH.

***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 4-*O*-acetyl-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside (13).** A mixture of acceptor **9** (188 mg, 0.31 mmol) and donor **8** (316 mg, 0.47 mmol) and 4 Å molecular sieves was stirred in DCM (3.6 mL) for 2 h. The reaction mixture was cooled to -70 °C and a solution of TMSOTf (11 µL, 0.06 mmol) in DCM (100 µL) was added. After 40 min the reaction was quenched with Et_3N, filtered over a pad of Celite and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 8.5:1.5 v/v) as the eluent to afford compound **13** (251 mg, 72 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.80 – 7.57 (4H, m, Ar-H), 7.47 (2H, d, *J* = 7.6 Hz, Ar-H), 7.42 – 7.21 (18H, m, Ar-H), 6.95 (5H, ddd, *J* = 25.0, 21.7, 7.3 Hz, Ar-H), 5.43 (1H, d, *J* = 8.4 Hz, H-1, GlcNAc-II), 5.39 (1H, s, C*H*-C_6H_5), 5.17 (2H, d, *J* = 11.8 Hz, CH_2), 5.09 (1H, t, *J* = 9.4 Hz, H-4, GlcNac-II), 4.79 (1H, d, *J* = 14.8 Hz, H-1, GalNAc-I), 4.60 (1H, d, *J* = 12.0 Hz, C*H*H, Bn), 4.55 – 4.31 (8H, m, C*H*H, Bn; CH_2; H-3, GlcNAc-II; CH*H*, Bn; H-4, GalNAc-I; H-2, GlcNAc-II; CH*H*, Bn), 4.04 (1H, d, *J* = 11.6 Hz, H-6a, GalNAc-I), 3.95 (1H, d, *J* = 7.8 Hz, H-3, GalNAc-I), 3.84 (1H, dd, *J* = 12.1, 5.0 Hz, H-5, GlcNAc-II), 3.67 – 3.11 (9H, m, H-6, GlcNAc-II; H-6b, GalNAc-I; H-2, GalNAc-I; C*H*H, pentyl; H-5, GalNAc-I; CH*H*, pentyl; CH_2, pentyl), 1.99 (3H, s, CH_3, OAc), 1.49 (4H, d, *J* = 33.8 Hz, 2x CH_2, pentyl), 1.23 (2H, d, *J* = 37.0 Hz, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 169.8 (C, OAc), 138.0 (C=O, Cbz), 137.8 (C, NPhth), 137.6 (C, NPhth), 133.7, 131.7, 129.1, 128.7, 128.6, 128.5, 128.3, 128.2, 128.1, 128.0, 127.9, 127.9, 127.8, 127.5, 127.2, 126.1, 123.3, 100.4 (*C*H-C_6H_5), 99.8 (C-1, GlcNAc-II), 98.4 (C-1, GalNAc-I), 77.0 (C-3, GlcNAc-II), 76.3 (C-3, GalNAc-I), 75.3 (C-4, GalNAc-I), 73.9 (CH_2, Bn), 73.6 (CH_2, Bn), 73.3 (C-5, GlcNAc-II), 72.3 (C-4, GlcNAc-II), 70.3 (C-6, GlcNAc-II), 68.9 (C-6, GalNAc-I), 68.2 (CH_2, pentyl), 67.2 (CH_2), 63.0 (C-5, GalNAc-I), 58.2 (C-2, GalNAc-I), 55.4 (C-2, GlcNAc-II), 50.3 (CH_2), 47.1 (CH_2, pentyl), 46.11 (CH_2, pentyl), 29.0 (CH_2, pentyl), 27.9 (CH_2, pentyl), 27.4 (CH_2, pentyl), 23.2 (CH_2, pentyl), 20.9 (*C*H_3, OAc). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{63}H_{65}N_5O_{14}$; 1138.4426 found 1138.4455.

***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside (11).** Freshly prepared NaOMe (2.5 mL) was added to a stirring solution of compound **13** (243 mg, 0.22 mmol) in DCM/MeOH (2.5/5 mL). After 1.5 h, Amberlite H^+ resin was added and the reaction mixture was filtered and concentrated *in vacuo* to afford compound **11** (226 mg, 97 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.84 – 7.55 (4H, m, H-Ar), 7.49 – 7.42 (2H, m,

H-Ar), 7.40 – 7.13 (18H, m, H-Ar), 7.10 – 7.05 (2H, m, H-Ar), 7.04 – 6.94 (3H, m, H-Ar), 5.47 – 5.38 (2H, m, H-1, GlcNAc-II; C*H*-C_6H_5), 5.17 (2H, d, *J* = 13.3 Hz, CH_2), 4.81 (1H, d, *J* = 13.2 Hz, H-1, GalNAc-I), 4.68 (1H, d, *J* = 12.1 Hz, CH*H*), 4.60 (1H, d, *J* = 11.6 Hz, CH*H*), 4.57 – 4.43 (4H, m, CH*H*; CH*H*; CH_2), 4.38 (1H, d, *J* = 1.8 Hz, H-4, GalNAc-I), 4.33 – 4.26 (2H, m H-3, GlcNAc-II; H-2, GlcNAc-II), 4.10 (1H, d, *J* = 10.7 Hz, H-6a, GalNAc-I), 3.95 (1H, d, *J* = 9.8 Hz, H-3, GalNAc-I), 3.87 – 3.69 (5H, m, H-6, GlcNAc-II; H-6b, GalNAc-I; H-5, GlcNAc-II; H-4, GlcNAc-II), 3.61 – 3.44 (3H, m, H-2, GalNAc-I; CH*H,* pentyl; H-5, GalNAc-I), 3.39 – 3.10 (3H, m, CH*H*, pentyl, CH_2, pentyl), 2.61 (1H, d, *J* = 10.4 Hz, O*H*), 1.62 – 1.39 (4H, m, 2x CH_2, pentyl), 1.32 – 1.12 (2H, m, *J* = 36.9, 11.8 Hz, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 138.0 (C=O, Cbz), 137.7 (C, NPhth), 133.7, 131.8, 128.6, 128.5, 128.3, 128.0, 127.9, 127.9, 127.8, 127.6, 126.1, 123.2, 100.4 (*C*H-C_6H_5) , 99.8 (C-1, GlcNAc-II), 98.4 (C-1, GalNAc-I), 79.0 (C-3, GlcNAc-II), 75.8 (C-3, GalNAc-I), 75.5 (C-4, GalNAc-I), 74.3 (CH_2) , 73.8 (CH_2), 73.5 (C-5, GlcNAc-II), 71.0 (C-6, GlcNAc-II), 69.0 (C-6, GalNAc-I), 68.2 (CH_2, pentyl), 67.1 (CH_2), 62.9 (C-5, GalNAc-I), 58.3 (C-2, GalNAc-I), 55.3 (C-2, GlcNAc-II), 29.0 (CH_2 , pentyl) 23.3 (CH_2, pentyl). ESI HRMS (m/z): $[M + Na]^+$ calcd for $C_{61}H_{63}N_5O_{13}$; 1096.4320 found 1096.4250.

Phenyl 2-*O*-benzoyl-3-*O*-(2-naphthyl)methyl-4,6-*O*-benzylidene-1-thio-β-D-galactopyranoside (26). Compound **25** was synthesized as described before.[30] Benzoyl chloride (0.72 mL, 6.2 mmol) was added to a stirring solution of compound **25** (2.08 g, 4.1 mmol) in pyridine (10 mL). After 1 h the mixture was concentrated *in vacuo*, separated between EtOAc/H_2O, dried (Na_2SO_4), filtered and the filtrate concentrated *in vacuo*. The obtained residue was dissolved in a small amount of DCM and crystallized by adding hexane. The resulting solids were filtered, collected and dried in high *vacuo* (1.99 g, 80 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.03 (2H, d, *J* = 7.7 Hz, H-Ar), 7.74 (1H, d, *J* = 7.8 Hz, H-Ar), 7.65 – 7.55 (4H, m, H-Ar), 7.53 – 7.34 (9H, m, H-Ar), 7.32 – 7.20 (6H, m, H-Ar), 5.59 (1H, t, *J* = 9.7 Hz, H-2), 5.49 (1H, s, C*H*-C_6H_5), 4.82 – 4.76 (2H, m, CH*H*, Nap; H-1), 4.72 (1H, d, *J* = 13.0 Hz, CH*H*, Nap), 4.38 (1H, d, *J* = 12.1 Hz, H-6a), 4.28 (1H, s, H-4), 4.02 (1H, d, *J* = 12.1 Hz, H-6b), 3.81 (1H, d, *J* = 9.6 Hz, H-3), 3.49 (1H, s, H-5). ^{13}C NMR (151 MHz, $CDCl_3$) δ 133.8, 133.0, 129.9, 129.1, 128.7, 128.5, 128.4, 128.2, 128.0, 127.8, 127.7, 126.7, 126.5, 126.1, 126.0, 125.7, 101.3 (*C*H-C_6H_5), 85.4 (C-1), 78.1 (C-3), 73.2 (C-4), 71.1 (CH_2), 70.1 (C-5), 69.3 (C-6), 69.0 (C-2). ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{37}H_{32}O_6S$; 622.2263 found 622.2273.

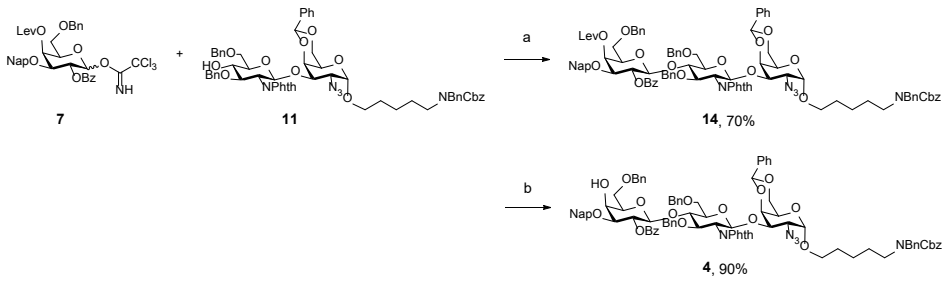
Phenyl 2-*O*-benzoyl-6-*O*-benzyl-3-*O*-(2-naphthyl)methyl-1-thio-β-D-galactopyranoside (27). Compound **26** (1.17 g, 1.9 mmol) was stirred with 4 Å molecular sieves in DCM (20 mL) for 1 h. The mixture was cooled by an ice bath and Et_3SiH (2.5 mL, 15.5 mmol) and TFA (0.89 mL, 11.6 mmol) were added. The progress of the reaction was monitored by TLC (Toluene:EtOAc 9:1 v/v) and more reagents (2x 1.25 mL Et_3SiH and 2 x 0.44 mL TFA) were added over time (after 1 and 2 h) to drive the reaction to completion. After stirring overnight, H_2O (4 mL) was added and the mixture was filtered. The filtrate was diluted with DCM, washed with sat. aq. $NaHCO_3$, brine and the organic layer was dried (Na_2SO_4), filtered and the filtrate concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 17:3 v/v) as the eluent to afford compound **27** (1.12 g, 95 %). 1H NMR (600 MHz, CDCl3) δ 7.97 (2H, d, J = 7.2 Hz, H-Ar), 7.71 (1H, d, J = 7.7 Hz, H-Ar), 7.61 – 7.55 (3H, m, H-Ar), 7.52 (1H, d, J = 8.4 Hz, H-Ar), 7.48 – 7.12 (15H, m, H-Ar), 5.52 (1H, t, J = 9.7 Hz, H-2),

4.83 (1H, d, J = 12.5 Hz, CH*H*), 4.72 (1H, d, J = 10.1 Hz, H-1), 4.65 (1H, d, J = 12.5 Hz, CH*H*), 4.60 (2H, s, CH_2), 4.22 (1H, d, J = 2.8 Hz, H-4), 3.89 – 3.78 (2H, m, H-6a,b), 3.73 – 3.66 (2H, m, H-3; H-5). ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.4 (C=O), 138.0, 134.4, 133.3, 133.2, 133.0, 132.2, 129.9, 128.8, 128.5, 128.4, 128.4, 127.9, 127.8, 127.7, 127.7, 126.9, 126.2, 126.1, 125.8, 86.7 (C-1), 79.1 (C-3), 77.5 (C-5), 73.8 (CH_2), 71.3 (CH_2), 69.6 (C-2), 69.3 (C-6), 66.3 (C-4). ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{37}H_{34}O_6S$; 624.2430 found 624.2420.

Phenyl 2-*O*-benzoyl-6-*O*-benzyl-4-*O*-levulenoyl-3-*O*-(2-naphthyl)methyl-1-thio-β-D-galactopyranoside (28). A mixture of levulinic acid (0.21 mL, 2.0 mmol), EDCHCl (391 mg, 2.0 mmol), NEt_3 (0.28 mL, 2.0 mmol) and DMAP (30 mg, 0.24 mmol) was stirred in DCM (3 mL) for 1 h. This mixture was added to compound **27** (495 mg, 0.82 mmol) in DCM (2 mL). The reaction mixture was stirred overnight and a fresh LevOH/EDC mix was made and after stirring for an additional 1 h, added to the reaction mixture. This was repeated once more and then the mixture was washed with H_2O, brine, extracted with DCM and the organic layer was dried (Na_2SO_4), filtered and and the filtrate concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:Acetone (1:0 to 100:10 v/v) as the eluent to afford compound **28** (294 mg, 51 %) and starting material **27** (257 mg). The starting material was submitted for Lev protection as described above to obtain more compound **28** (140 mg, 47 %). (Total overall yield: 434 mg, 75 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.95 (2H, d, *J* = 7.9 Hz, H-Ar), 7.70 (1H, d, *J* = 8.0 Hz, H-Ar), 7.64 – 7.53 (2H, m, *J* = 19.5, 12.1 Hz, H-Ar), 7.51 (1H, d, *J* = 8.0 Hz, H-Ar), 7.47 – 7.12 (16H, m, H-Ar), 5.73 (1H, d, *J* = 3.0 Hz, H-4), 5.42 (1H, t, *J* = 9.8 Hz, H-2), 4.79 – 4.72 (2H, m, CH*H*; H-1), 4.57 – 4.51 (3H, m, CH*H*; CH_2), 3.82 (1H, t, *J* = 6.3 Hz, H-5), 3.73 – 3.68 (2H, m, H-6a; H-3), 3.61 (1H, dd, *J* = 9.5, 6.9 Hz, H-6b), 2.85 – 2.56 (4H, m, 2x CH_2, Lev), 2.17 (3H, s, CH_3, Lev). ^{13}C NMR (151 MHz, $CDCl_3$) δ 206.4 (H_3C-C=O, Lev), 172.1 (O-C=O, Lev), 165.3 (C=O, Bz), 134.6, 133.2, 132.4, 130.0, 128.8, 128.5, 128.4, 128.2, 127.8, 127.6, 126.9, 126.0, 125.9, 86.8 (C-1), 77.1 (C-3), 76.3 (C-5), 73.8 (CH_2), 70.8 (CH_2), 69.5 (C-2), 68.2 (C-6), 66.4 (C-4), 38.1 (CH_2, Lev), 29.9 (CH_3, Lev), 28.1 (CH_2, Lev). ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{42}H_{40}O_8S$; 722.2788 found 722.2799.

2,2,2-Trichloroacetimidate 2-*O*-benzoyl-6-*O*-benzyl-4-*O*-levulenoyl-3-*O*-(2-naphthyl) methyl-α/β-D-galactopyranoside (7). *N*-iodosuccinimide (255 mg, 1.13 mmol) and TFA (87 µL, 1.13 mmol) were added to a cooled (0 °C) mixture of compound **28** (400 mg, 0.57 mmol) in DCM (10 mL). After 2 h the mixture was washed with sat. aq. $NaHCO_3$, extracted with DCM, washed with 5 % NaS_2O_3, dried (Na_2SO_4), filtered and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Hexane:EtOAc (1:0 to 1:1 v/v) as the eluent to afford the free reducing intermediate (236 mg, 68 %). 2,2,2-trichloroacetonitrile (192 µL, 1.92 mmol) and DBU (11 µL, 0.077 mmol) were added to this intermediate (235 mg, 0.38 mmol) in DCM (10 mL) at 0 °C. After 1 h the mixture was concentrated *in vacuo* and the residue was purified by silica column chromatography using Hexane (1% Et_3N):EtOAc (1:0 to 1:1 v/v) as the eluent to afford compound **7** as an α/β mixture (230 mg, 79 %). ^{1}H NMR (α/β mixture 2/1) (600 MHz, $CDCl_3$) δ 8.56 (0.5H, s, N*H*, β), 8.40 (1H, s, N*H*, α), 7.90 – 7.83 (3H, m, H-Ar), 7.79 – 7.27 (24H, m, H-Ar), 7.20 (0.5H, dd, *J* = 8.4, 1.6 Hz, H-Ar), 6.63 (1H, d, *J* = 3.5 Hz, H-1, α), 5.90 – 5.86 (1H, m, H-4, α), 5.83 (0.5H, d, *J* = 8.4 Hz, H-1, β), 5.79 (0.5H, dd, *J* = 3.3, 1.0 Hz, H-4, β), 5.66 (0.5H,

dd, *J* = 10.0, 8.4 Hz, H-2, β), 5.56 (1H, dd, *J* = 9.1, 4.5 Hz, H-2, α), 4.87 – 4.79 (1.5H, m, *J* = 18.7, 12.6 Hz, CH*H* α, β), 4.66 – 4.57 (1.5H, m, CH*H* α, β), 4.56 – 4.53 (3H, m, CH_2, α, β), 4.39 – 4.35 (1H, m, H-5, α), 4.22 (1H, dd, *J* = 9.2, 4.6 Hz, H-3, α), 3.99 – 3.94 (0.5H , m, H-5, β), 3.77 (0.5H, dd, *J* = 9.5, 4.8 Hz, H-3, β), 3.72 – 3.57 (3H, m, H-6 α, β), 2.92 – 2.60 (6H, m, 2x CH_2, Lev, α, β), 2.18 (s, 1.5H, CH_3, Lev, β), 2.16 (s, 3H, CH_3, Lev, α). ^{13}C NMR (151 MHz, $CDCl_3$) δ 206.3 (H_3C-C=O, Lev), 172.0 (O-C=O, Lev), 165.6 (C=O, Bz), 160.5 (C=NH), 137.7, 134.7, 133.0, 129.9, 128.5, 128.4, 128.3, 128.3, 128.2, 128.2, 127.9, 127.6, 127.2, 126.3, 126.1, 96.4 (C-1, β), 94.2 (C-1, α), 75.9 (C-3, β), 73.8 (CH_2), 73.5 (C-5, β), 72.1 (C-3, α), 71.3 (CH_2, α) , 71.0 (CH_2, β), 70.7 (C-5, α), 69.9 (C-2, β), 69.1 (C-2, α), 67.6 (C-6, α), 67.3 (C-6, β), 67.0 (C-4, α), 66.0 (C-4, β), 38.1 (CH_2, Lev), 29.8 (CH_3, Lev), 28.1 (CH_2, Lev).

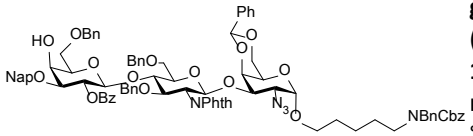


Scheme S2. Chemical glycosylation of galactose donor **7** and disaccharide acceptor **11** to form trisaccharide **14**. Lev removal on compound **14** formed the trisaccharide acceptor **4**. Reagents and conditions: (a) TMSOTf, CH_2Cl_2, 4 Å MS, -70 °C; (b) NH_2NH_2.HOAc, CH_2Cl_2/MeOH, 50 °C.

4

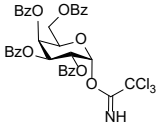
***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2-*O*-benzoyl-6-*O*-benzyl-4-*O*-levulenoyl-3-*O*-(2-naphthyl)methyl-β-D-galactopyranoside-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside (14)**. A mixture of acceptor **11** (234 mg, 0.22 mmol) and donor **7** (231 mg, 0.31 mmol) and 4 Å molecular sieves was stirred in DCM (3 mL) for 3 h. The reaction mixture was cooled to -70°C and a solution of TMSOTf (4 µL, 0.022 mmol) in DCM (100 µL) was added. After 1 h Et_3N was added and the reaction mixture was concentrated *in vacuo*. The obtained residue was purified by SX1 size exclusion chromatography (Toluene:Acetone 1:1 v/v), followed by silica column chromatography using Toluene:EtOAc (1:0 to 4:1 v/v) as the eluent to afford compound **14** (256 mg, 70 %). 1H NMR (600 MHz, $CDCl_{3)}$ δ 7.90 (2H, d, J = 7.4 Hz, H-Ar), 7.77 – 6.84 (44H, m, H-Ar), 5.69 (1H, d, J = 3.1 Hz, H-4, Gal-III), 5.40 (1H, dd, J = 9.8, 8.1 Hz, H-2, Gal-III), 5.32 – 5.26 (2H, m, C*H*-C_6H_5; H-1, GlcNAc-II), 5.15 (2H, d, J = 7.6 Hz, CH_2), 4.88 (1H, d, CH*H*), 4.80 – 4.71 (2H, m, CH*H*; H-1, GalNAc-I), 4.64 (1H, d, J = 8.0 Hz, H-1, Gal-III), 4.56 – 4.20 (m, 10H, CH*H*; CH_2; CH*H*; CH_2; CH*H*; H-3, GlcNAc-II; H-4, GalNAc-I; H-2 GlcNAc-II), 4.08 – 3.97 (2H, m, J = 9.3 Hz, CH*H*; H-6a, GalNAc-I), 3.91 (1H, s, H-4, GlcNAc-II), 3.84 (1H, d, J = 9.4 Hz, H-3, GalNAc-I), 3.67 – 3.08 (14H, m, H-5, Gal-III; H-6b, GalNAc-I; H-3, Gal-III; H-6, GlcNAc-II; H-2, GalNAc-I; H-5, GlcNAc-II; H-6, Gal-III; H-5, GalNAc-I; CH_2, pentyl, CH_2, pentyl), 2.84 – 2.58 (4H, m, 2x CH_2, Lev), 2.09 (3H, s, CH_3, Lev), 1.45 (4H, d, J = 33.3 Hz, 2x CH_2, pentyl), 1.18 (2H, d, J = 36.3 Hz, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 206.1

(H_3C-C=O, Lev), 172.0 (O-C=O, Lev), 165.0 (C=O, Bz), 138.6, 137.9, 137.8, 137.7, 134.8, 133.6, 133.3, 133.0, 133.0, 131.8, 129.9, 129.7, 128.5, 128.5, 128.3, 128.1, 128.0, 127.9, 127.8, 127.7, 127.6, 127.0, 126.9, 126.0, 123.2, 100.9 (C-1, Gal-III), 100.3 (CH-C_6H_5), 99.6 (C-1, GalNAc-I), 98.4 (C-1, GlcNAc-II), 78.3 (C-4, GlcNAc-II), 77.2 (C-3, GlcNAc-II), 76.2 (C-3, Gal-III), 75.8 (C-3, GalNAc-I), 75.3 (C-4, GalNAc-I), 74.6 (CH_2), 74.3 (C-5, GlcNAc-II), 73.7 (CH_2), 73.4 (CH_2), 72.2 (C-5, Gal-III), 71.8 (C-2, Gal-III), 70.7 (CH_2), 68.9 (C-6, GalNAc-I), 68.8 (C-6, GlcNAc-II), 68.1 (CH_2, pentyl), 67.5 (C-6, Gal-III) 67.1 (CH_2), 66.0 (C-4, Gal-III), 62.8 (C-5, GalNAc-I), 58.1 (C-2, GalNAc-I), 55.6 (C-2, GlcNAc-II), 50.6 (CH_2), 47.0 (CH_2, pentyl) , 38.0 (CH_2, Lev) , 29.7 (CH_3, Lev), 28.0 (CH_2, pentyl) , 27.4 (CH_2, pentyl), 23.2 (CH_2, pentyl). ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{97}H_{97}N_5O_{21}$; 1685.7020 found 1685.7068.

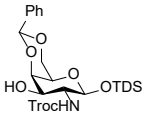
***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2-*O*-benzoyl-6-*O*-benzyl-3-*O*-(2-naphthyl)methyl-β-D-galactopyranoside-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside (4).** 

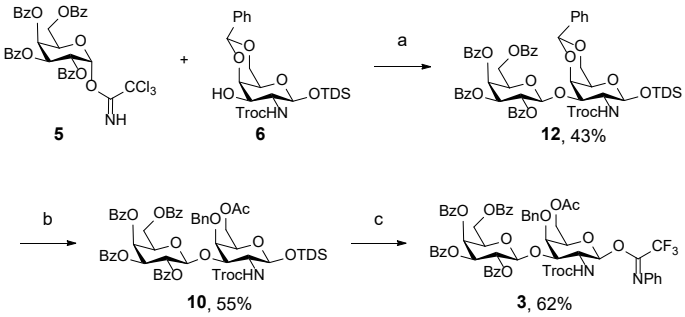
Hydrazine acetate (21 mg, 0.22 mg) was added to compound **14** (250 mg, 0.15 mmol) in DCM/MeOH (2/1 mL). The reaction mixture was heated to 50 °C and stirred overnight. After full conversion of the starting material was observed by MALDI-TOF, the mixture was concentrated *in vacuo.* The obtained residue was purified by silica column chromatography using Toluene:Acetone (1:0 to 9:1 v/v) as the eluent to afford compound **4** (212 mg, 90 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.92 (2H, d, *J* = 7.7 Hz, H-Ar), 7.73 (1H, d, *J* = 7.9 Hz, H-Ar), 7.69 – 6.91 (40H, m, H-Ar), 6.88 – 6.78 (3H, m, H-Ar), 5.54 – 5.50 (1H, m, H-2, Gal-III), 5.31 (1H, s, C*H*-C_6H_5), 5.28 (1H, d, *J* = 8.2 Hz, H-1, GlcNAc-II), 5.15 (2H, d, *J* = 10.2 Hz, CH_2), 4.90 – 4.81 (2H, m, CH*H*; CH*H*), 4.74 (1H, d, *J* = 15.0 Hz, H-1, GalNAc-I), 4.65 – 4.57 (2H, m, H-1, Gal-III; CH*H*), 4.51 – 4.41 (5H, m, CH_2; CH*H*; CH_2), 4.35 (1H, t, *J* = 13.2 Hz, CH*H*), 4.31 – 4.21 (3H, m, H-3, GlcNAc-II; H-4, GalNAc-I; H-2, GlcNAc-II), 4.19 (1H, d, *J* = 2.8 Hz, H-4, Gal-III), 4.07 – 3.98 (2H, m, CH*H*; H-6a, GalNAc-I), 3.92 (1H, d, *J* = 7.7 Hz, H-4, GlcNAc-II), 3.84 (1H, s, H-3, GalNAc-I), 3.73 (1H, dd, *J* = 9.6, 6.7 Hz, H-6a, Gal-III), 3.63 – 3.43 (m, 9H, H-6, GlcNAc-II; H-6b, GalNAc-I; H-5, Gal-III; H-3, Gal-III; H-2, GalNAc-I; H-6b, Gal-III; H-5, GlcNAc-II; CH*H*, pentyl), 3.38 (1H, d, *J* = 26.2 Hz, H-5, GalNAc-I), 3.31 – 3.10 (3H, m, CH*H*, pentyl CH_2, pentyl), 1.52 – 1.33 (4H, m, 2x CH_2, pentyl), 1.30 – 1.08 (2H, m, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.2 (C=O, Bz), 138.6, 138.0, 138.0, 137.7, 137.3, 134.6, 133.5, 133.3, 133.1, 131.8, 129.9, 129.8, 129.1, 128.6, 128.5, 128.5, 128.3, 128.0, 128.0, 127.9, 127.8, 127.8, 127.7, 127.7, 127.7, 127.2, 127.0, 126.9, 126.3, 126.1, 126.1, 125.8, 123.2, 100.9 (C-1, Gal-III), 100.3 (*C*H-C_6H_5), 99.6 (C-1, GlcNAc-II), 98.4 (C-1, GalNAc-I), 78.4 (C-4, GlcNAc-II), 78.1 (C-5, Gal-III), 77.3 (C-3, GlcNAc-II), 75.9 (C-3, GalNAc-I), 75.4 (C-4, GalNAc-I), 74.8 (CH_2), 74.4 (C-5, GlcNAc-II), 73.7 (CH_2), 73.4 (CH_2), 73.3 (C-3, Gal-III), 71.8 (C-2, Gal-III), 70.9 (CH_2), 69.0 (C-6, GalNAc-I), 69.0 (C-6, GlcNAc-II), 68.7 (C-6, Gal-III), 68.2 (CH_2, pentyl), 67.2 (CH_2), 65.6 (C-4, Gal-III), 62.9 (C-5, GalNAc-I), 58.2 (C-2, GalNAc-I), 55.7 (C-2, GlcNAc-II), 50.6 (CH_2), 46.1 (CH_2, pentyl), 29.0 (CH_2, pentyl), 27.4 (CH_2, pentyl), 23.2 (CH_2, pentyl). ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{91}H_{91}N_5O_{19}$; 1587.6652 found 1587.6690.

2,2,2-Trichloroacetimidate 2,3,4,6-tetra-*O*-benzoyl-α-D-galactopyranoside (5). Compound **5** was synthesized as described before.[33]



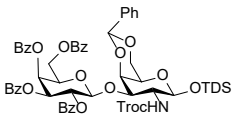
Dimethylthexylsilyl 4,6-*O*-benzylidene-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranoside (6). Compound **6** was synthesized as described before.[32]





Scheme S3. Chemical glycosylation of galactose donor **5** and galactosamine acceptor **6** to form disaccharide **12**. Selective opening of the benzylidene acetal and subsequent acetylation provided disaccharide **10**. Removal of anomeric protecting group (TDS) and formation of the acetimidate provided disaccharide donor **3**. Reagents and conditions: (a) TMSOTf, CH_2Cl_2, 4 Å MS, -40 °C; (b) i. i) Cl_2BPh, Et_3SiH, CH_2Cl_2, 4 Å MS, -68 °C, ii. Ac_2O, Pyridine; (c) i. HF·Pyridine, Pyr, ii. ClC(NPh)CF_3, Cs_2CO_3, CH_2Cl_2.

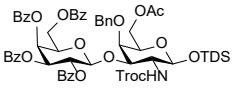
Dimethylthexylsilyl 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl-(1→3)-4,6-*O*-benzylidene-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranoside (12).



A mixture of acceptor **6** (300 mg, 0.51 mmol), donor **5** (570 mg, 0.77 mmol) and 4 Å molecular sieves was stirred in DCM (5 mL) for 2.5 h. The mixture was cooled to -40 °C and TMSOTf (18 µL, 0.10 mmol) was added. After 10 min Et_3N/MeOH was added and the mixture was concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 100:4 v/v) as the eluent to afford compound **12** (256 mg, 43 %)**.** ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.11 – 7.12 (m, 25H, H-Ar), 5.96 (1H, s, H-4, Gal-V), 5.82 (1H, t, *J* = 9.1 Hz, H-2, Gal-V), 5.54 (1H, d, *J* = 10.3 Hz, H-3, Gal-V), 5.49 – 5.43 (1H, m, C*H*-C_6H_5), 5.11 – 5.00 (2H, m, H-1, Gal-V; H-1; GalNAc-IV), 4.75 (1H, dd, *J* = 11.0, 7.3 Hz, H-6a, Gal-V), 4.60 – 4.52 (2H, m, H-3, GalNAc-IV; CH*H*, Troc), 4.42 – 4.30 (3H, m, H-6b, Gal-V; H-5, Gal-V; H-4, GalNAc-IV), 4.11 (1H, d, *J* = 12.1 Hz, H-6a, GalNAc-IV), 3.99 (1H, d, *J* = 12.0 Hz, CH*H*, Troc), 3.74 (1H, d, *J* = 12.0 Hz, H-6b, GalNAc-IV), 3.42 – 3.33 (1H m, H-2, GalNAc-IV), 3.20 (1H, s, H-5, GalNAc-IV), 1.63 – 1.55 (1H, m, C*H*-TDS), 0.85 – 0.76 (12H, m, *J* = 19.8, 6.0 Hz, 4x CH_3, TDS), 0.14 (3H, s, CH_3-Si), 0.08 (3H, s, CH_3-Si). ^{13}C NMR (151 MHz, $CDCl_3$) δ 166.0 (C=O, Bz), 165.6 (C=O, Bz), 165.6 (C=O, Bz), 165.0 (C=O, Bz), 153.7 (C=O, Troc), 133.6 (C, Bz),

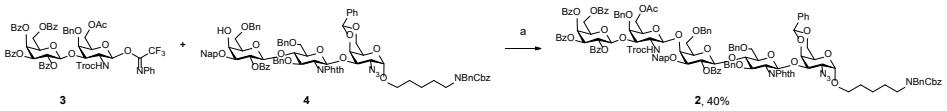
133.5 (C, Bz), 133.4 (C, Bz), 133.3 (C, Bz), 130.1, 129.8, 129.8, 129.5, 129.1, 129.0, 128.8, 128.7, 128.6, 128.3, 128.2, 125.3, 102.2 (C-1, Gal-V), 100.8 (*C*H-C_6H_5), 95.4 (CCl_3, Troc), 94.3 (C-1, GalNAc-IV), 75.8 (C-4, GalNAc-IV), 73.8 (C-3, GalNAc-IV), 71.9 (C-3, Gal-V), 71.5 (C-5, Gal-V), 69.9 (C-2, Gal-V), 69.2 (C-6, GalNAc-IV), 68.2 (C-4, Gal-V), 66.4 (C-5, GalNAc-IV), 62.5 (C-6, Gal-V), 56.1 (C-2, GalNAc-IV), 34.1 (*CH*, TDS) 24.8 (C, TDS), 20.2 (CH_3, TDS) 20.1 (CH_3, TDS), 18.6 (CH_3, TDS), 18.5 (CH_3, TDS), -1.8 (CH_3-Si) , -3.0 (CH_3-Si). ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{58}H_{62}Cl_3NO_{16}Si$; 1179.3247 found 1179.3288.

Dimethylthexylsilyl 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl-(1→3)-6-*O*-acetyl-4-*O*-benzyl-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranoside (10).



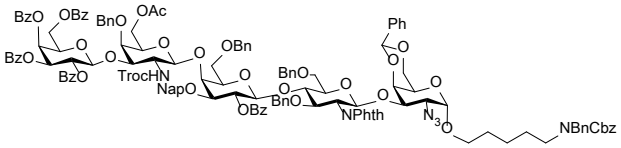
Compound **12** (144 mg, 0.12 mmol) was stirred with 4 Å molecular sieves in DCM (4 mL) for 2.5 h. The mixture was cooled to -68 °C and Et_3SiH (80 µL, 0.50 mmol) and Cl_2PhB (50 µL, 0.38 mmol) were added. After 15 min Et_3N was added and the reaction mixture was concentrated *in vacuo.* The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 100:15 v/v) as the eluent to afford the desired compound**.** The obtained intermediate (85 mg, 0.073 mmol) was dissolved in pyridine (2 mL) and Ac_2O (1 mL) was added. After 2.5 h the reaction mixture was concentrated *in vacuo* and co-evaporated with Toluene. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 10:2 v/v) as the eluent to afford compound **10** (82 mg, 55 % over 2 steps)**.** ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.06 – 7.92 (6H, m, H-Ar), 7.78 (2H, d, *J* = 7.7 Hz, H-Ar), 7.59 – 7.14 (17H, m, H-Ar), 6.00 (1H, d, *J* = 3.2 Hz, H-4, Gal-V), 5.93 (1H, dd, *J* = 10.4, 8.0 Hz, H-2, Gal-V), 5.57 (1H, dd, *J* = 10.5, 3.0 Hz, H-3, Gal-V), 5.14 (1H, d, *J* = 11.2 Hz, CH*H*), 5.04 (1H, d, *J* = 7.9 Hz, H-1, Gal-V), 4.94 (1H, d, *J* = 6.7 Hz, N*H*), 4.92 (1H, d, *J* = 8.0 Hz, H-1, GalNAc-IV), 4.78 – 4.69 (2H, m, CH*H*; H-6a, Gal-V), 4.59 (1H, dd, *J* = 11.0, 2.3 Hz, H-3, GalNAc-IV), 4.55 – 4.42 (3H, m, CH_2; H-6b, Gal-V), 4.38 – 4.32 (1H, m, H-5, Gal-V), 4.18 (1H, dd, *J* = 11.1, 7.3 Hz, H-6a, GalNAc-IV), 3.98 (1H, s, H-4, GalNAc-IV), 3.89 (1H, dd, *J* = 11.1, 5.2 Hz, H-6b, GalNAc-IV), 3.47 (1H, t, *J* = 6.2 Hz, H-5, GalNAc-IV), 3.33 – 3.22 (1H, m, *J* = 17.7, 8.3 Hz, H-2, GalNAc-IV), 1.91 (3H, s, CH_3, OAc), 1.59 – 1.50 (1H, m, C*H*, TDS), 0.83 – 0.72 (12H, m, 4 x CH_3, TDS), 0.08 (3H, s, CH_3-Si), 0.05 (3H, s, CH_3-Si). ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.6 (C=O, Ac), 166.1 (C=O, Bz), 165.7 (C=O, Bz), 165.5 (C=O, Bz), 165.3 (C=O, Bz), 153.9 (C=O, Troc), 138.4 (C, Bn), 133.8 (C, Bz), 133.6 (C, Bz), 133.5 (C, Bz), 133.5 (C, Bz), 129.9, 129.9, 129.4, 129.2, 128.8, 128.8, 128.7, 128.5, 128.5, 128.4, 128.0, 102.4 (C-1, Gal-V), 95.5 (CCl_3, Troc), 94.4 (C-1, GalNAc-IV), 77.7 (C-3, GalNAc-IV), 75.3 (C-4, GalNAc-IV), 74.9 (CH_2), 74.3 (CH_2), 72.1 (C-5, GalNAc-IV), 71.6 (H-3, Gal-V), 71.6 (H-5, Gal-V), 70.2 (C-2, Gal-V), 68.3 (C-4, Gal-V), 63.3 (C-6, GalNAc-IV), 62.2 (C-6, Gal-V), 57.1 (C-2, GalNAc-IV), 34.1 (*C*H, TDS), 24.9 (C, TDS), 20.8 (CH_3, Ac) , 20.2 (CH_3, TDS), 20.1 (CH_3, TDS), 18.7 (CH_3, TDS), 18.6 (CH_3, TDS), -1.8 (CH_3-Si), -3.4 (CH_3-Si) . ESI HRMS (m/z): $[M + NH]^+$ calcd for $C_{60}H_{66}Cl_3NO_{17}Si$; 2123.3509 found 2123.3547.

(*N*-phenyl)-2,2,2-trifluoroacetimidate 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl-(1→3)-6-*O*-acetyl-4-*O*-benzyl-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranoside (3). HF·Pyridine (70% HF, 80 µL) was added to a stirring solution of compound **10** (80 mg, 0.066 mmol) in pyridine (1 mL) in a plastic round bottom flask. After overnight stirring the mixture was diluted with DCM and quenched by addition of sat. aq. $NaHCO_3$. The organic phase was washed with sat. aq. $NaHCO_3$, the aq. layers were extracted with DCM and the combined organic phases were dried (Na_2SO_4), filtered, concentrated *in vacuo* and co-evaporated with toluene. The obtained residue was dissolved in DCM (1 mL) and Cs_2CO_3 (86 mg, 0.27 mmol) and CF_3NPhimidate (159 L, 0.99 mmol) were added. After 3 h the mixture was concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc (1:0 to 10:2 v/v) as the eluent to afford compound **3** (51 mg, 62 % over 2 steps)**.** ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.01 – 7.93 (6H, m, H-Ar), 7.76 (2H, d, *J* = 7.8 Hz, H-Ar), 7.58 – 7.14 (19H, m, H-Ar), 7.09 (1H, t, *J* = 7.5 Hz, H-Ar), 6.74 (2H, d, *J* = 7.4 Hz, H-Ar), 6.04 (1H, d, *J* = 2.9 Hz, H-4, Gal-V), 5.99 – 5.90 (1H, m, H-2, Gal-V), 5.68 (1H, d, *J* = 9.9 Hz, H-3, Gal-V), 5.26 – 5.13 (2H, m, CH*H*; H-1, Gal-V), 4.93 (1H, s, N*H*), 4.80 – 4.74 (1H, m, H-6a, Gal-V), 4.70 (1H, d, *J* = 10.8 Hz, CH*H*), 4.54 – 4.43 (4H, m, H-2, GalNAc-IV; H-6b, Gal-V; CH*H*; H-5, Gal-V), 4.21 – 4.13 (2H, m, H-3, GalNAc-IV; H-6a, GalNAc-IV), 4.09 – 4.04 (1H, m, H-4, GalNAc-IV), 4.01 (1H, dd, *J* = 11.1, 4.7 Hz, H-6b, GalNAc-IV), 3.93 (1H, s, H-5, GalNAc-IV), 3.40 (1H, d, *J* = 12.1 Hz, CH*H*), 1.97 (3H, s, CH_3, OAc). ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.4 (C=O, Ac), 165.9 (C=O, Bz), 165.7 (C=O, Bz), 165.6 (C=O, Bz), 165.4 (C=O, Bz), 153.7 (C=O, Troc), 143.1 (C, NPh), 137.9 (C, Bn), 133.8 (C, Bz), 133.5 (C, Bz), 133.5 (C, Bz), 133.4 (C, Bz), 129.8, 129.8, 129.6, 129.1, 128.9, 128.7, 128.6, 128.5, 128.4, 128.0, 125.3, 119.2, 102.4 (C-1, Gal-V), 95.4 (CCl_3, Troc), 77.5 (C-4, GalNAc-IV), 75.2 (C-3, GalNAc-IV), 75.1 (CH_2), 73.9 (CH_2), 71.9 (C-5, Gal-V), 71.4 (C-3, Gal-V), 70.9 (C-5, GalNAc-IV), 69.9 (C-2, Gal-V), 68.1 (C-4, Gal-V), 63.0 (C-6, GalNAc-IV), 62.1 (C-6, Gal-V), 50.1 (C-2, GalNAc-IV), 20.7 (CH_3, Ac).



Scheme S4. Chemical glycosylation of disaccharide donor **3** with trisaccharide acceptor **4** provided pentasaccharide **2**. Reagents and conditions: (a) TMSOTf, CH_2Cl_2, 4 Å MS, -74 °C.

5-aminopentyl 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl-(1→3)-4-*O*-acetyl-6-*O*-benzyl-2-[(2,2,2-trichloroethoxy)carbonylamino]-β-D-galactopyranosyl-(1→4)-2-*O*-benzoyl-6-*O*-benzyl-3-*O*-(2-naphthyl)methyl-β-D-glucopyranosyl-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→3)-2-azido-4,6-*O*-benzylidene-2-deoxy-α-D-galactopyranoside (2). A mixture of acceptor **4** (54 mg, 0.034 mmol), donor **3** (51 mg, 0.041 mmol) and 4 Å molecular sieves was stirred in DCM (2 mL) for 2 h. The mixture was cooled to -74 °C and TMSOTf (1.24 µL, 0.0069 mmol) in DCM (200 µL) was added. The reaction was monitored by MALDI-TOF-MS and after 30 min Et_3N was added. The mixture was allowed to warm to room temperature, filtered over a pad of Celite and concentrated *in vacuo*. The obtained residue was purified by SX1 size exclusion chromatography using Toluene:Acetone (1:1 v/v) as the eluent to afford compound **2** (34 mg, 40 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.04 – 7.91 (8H, m, H-Ar), 7.81 – 7.09 (56H, m, H-Ar), 7.05 (2H, d, *J* = 5.9 Hz, H-Ar), 6.89 (2H, d, *J* = 7.5 Hz, H-Ar), 6.80 (2H, t, *J* = 7.5 Hz, H-Ar), 6.68 (1H, t, *J* = 7.3 Hz, H-Ar), 5.98 (1H, d, *J* = 3.0 Hz, H-4, Gal-V), 5.92 (1H, dd, *J* = 10.2, 8.3 Hz, H-2, Gal-V), 5.42 (1H, dd, *J* = 10.5, 3.1 Hz, H-3, Gal-V), 5.40 – 5.36 (1H, m, H-2, Gal-III), 5.31 – 5.22 (3H, m, C*H*-C_6H_5; H-1, GlcNAc-II; N*H*Troc), 5.15 (3H, d, *J* = 10.7 Hz, CH_2; C*H*H), 5.08 (1H, d, *J* = 8.2 Hz, H-1, GalNAc-IV), 4.95 (1H, d, *J* = 8.0 Hz, H-1, Gal-V), 4.81 – 4.59 (8H, m, CH*H*; H-6a, Gal-V; H-1, GalNAc-I; CH*H*; CH*H*; CH*H*; H-3, GalNAc-IV; H-1, Gal-III), 4.54 – 4.38 (4H, m, *J* = 21.3, 17.9, 9.4 Hz, CH*H*; CH_2; H-6b, Gal-V), 4.37 – 4.16 (9H, m, CH*H*; H-3, GlcNAc-II; H-4, GalNAc-I; CH*H*; CH_2; H-2, GlcNAc-II; H-5, Gal-V; H-4, Gal-III), 4.14 – 4.06 (2H, H-6, GalNAc-IV), 4.06 – 3.91 (4H, m, CH*H*; CH*H*; H-6a, GalNAc-I; H-4, GalNAc-IV), 3.89 – 3.76 (2H, m, H-4, GlcNAc-II; H-3, GalNAc-I), 3.60 – 3.09 (16H, m, H-3, Gal-III; H-6b, GalNAc-I; H-5, Gal-III; H-6a, Gal-III; H-5, GalNAc-IV; CH*H*, pentyl; H-6, GlcNAc-II; H-5, GlcNAc-II; H-5, GalNAc-I; H-2, GalNAc-IV; H-6b, Gal-III; CH*H*, pentyl; CH_2, pentyl), 1.87 (3H, s, CH_3, OAc), 1.46 (4H, d, *J* = 34.4 Hz, 2x CH_2, pentyl), 1.24 – 1.10 (2H, m, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 170.5 (C=O, Ac), 166.0 (C=O, Bz), 165.5 (C=O, Bz), 165.4 (C=O, Bz), 165.4 (C=O, Bz), 164.6 (C=O, Cbz), 154.3 (C=O, Troc), 138.4 , 138.3, 138.3, 138.0, 137.9, 137.9, 137.7, 134.8, 133.6, 133.6, 133.4, 133.3, 133.3, 133.1, 131.8, 129.9, 129.8, 129.8, 129.4, 129.3, 129.1, 129.1, 129.0, 128.7, 128.7, 128.6, 128.5, 128.4, 128.3, 128.3, 128.0, 127.9, 127.9, 127.7, 127.7, 127.5, 127.4, 127.3, 127.2, 126.9, 126.3, 126.2, 126.0, 125.9, 125. 3, 123.2, 102.0 (C-1, Gal-V), 101.0 (C-1, Gal-III), 100.3 (*C*H-C_6H_5), 99.7 (C-1, GlcNAc-II), 99.2 (C-1, GalNAc-IV), 98.4 (C-1, GalNAc-I), 96.0 (CCl_3, Troc), 79.7 (C-3, Gal-III), 78.2 (C-3, GalNAc-I), 77.5 (C-3, GlcNAc-II), 76.9 (C-3, GalNAc-IV), 75.9 (C-4, GlcNAc-II), 75.3 (C-4, GalNAc-I), 75.2 (C-4, GalNAc-IV), 75.0 (CH_2), 74.6 (CH_2), 74.4 (C-5, GalNAc-IV), 74.0 (CH_2), 73.9 (C-5, GlcNAc-II), 73.5 (CH_2), 73.4 (CH_2), 71.9 (C-5, Gal-III), 71.9 (C-2, Gal-III), 71.8 (CH_2), 71.5 (C-3, Gal-V), 71.3 (C-5, Gal-V), 71.1 (C-4, Gal-III), 69.7 (C-2, Gal-V), 69.1 (C-6, GlcNAc-II), 68.9 (C-6, GalNAc-I), 68.8 (C-6, Gal-III), 68.2 (C-4, Gal-V), 68.1 (CH_2, pentyl), 67.2 (CH_2), 62.9 (C-5, GalNAc-I), 62.8 (C-6, GalNAc-IV), 61.8 (C-6, Gal-V), 58.1 (C-2, GalNAc-I), 55.6 (C-2, GlcNAc-II), 55.2 (C-2, GalNAc-IV), 50.5 (CH_2) 46.1 (CH_2, pentyl), 29.0 (CH_2, pentyl), 27.4 (CH_2, pentyl) , 23.2 (CH_2, pentyl), 20.8 (CH_3, OAc). ESI HRMS (m/z): $[M + NH_4]^+$ calcd for $C_{144}H_{137}Cl_3N_6O_{35}$; 2632.8529 found 2632.8522.



References

1. Ghazarian, H.; Idoni, B.; Oppenheimer, S. B., Acta Histochem 2011, 113 (3), 236-47.
2. Jin, K. T.; Lan, H. R.; Chen, X. Y.; Wang, S. B.; Ying, X. J.; Lin, Y.; Mou, X. Z., Biotechnol Lett 2019, 41 (6-7), 641-650.
3. Chandler, K.; Goldman, R., Mol Cell Proteomics 2013, 12 (4), 836-45.
4. Fu, C.; Zhao, H.; Wang, Y.; Cai, H.; Xiao, Y.; Zeng, Y.; Chen, H., HLA 2016, 88 (6), 275-286.
5. Yu, A. L.; Hung, J. T.; Ho, M. Y.; Yu, J., Stem Cells Dev 2016, 25 (20), 1532-1548.
6. Chuang, P. K.; Hsiao, M.; Hsu, T. L.; Chang, C. F.; Wu, C. Y.; Chen, B. R.; Huang, H. W.; Liao, K. S.; Chen, C. C.; Chen, C. L.; Yang, S. M.; Kuo, C. W.; Chen, P.; Chiu, P. T.; Chen, I. J.; Lai, J. S.; Yu, C. T.; Wong, C. H., Proc Natl Acad Sci U S A 2019, 116 (9), 3518-3523.
7. Zhu, T.; Xu, L.; Xu, X.; Wang, Z.; Zhu, J.; Xie, Q.; Zhang, B.; Wang, Y.; Ju, L.; He, Y.; Ye, X.; Zhou, D.; Li, Y., Carbohydr Res 2015, 402, 189-99.
8. Okuda, T.; Shimizu, K.; Hasaba, S.; Date, M., Sci Rep 2019, 9 (1), 18803.
9. O'Cearbhaill, R. E.; Ragupathi, G.; Zhu, J.; Wan, Q.; Mironov, S.; Yang, G.; Spassova, M. K.; Iasonos, A.; Kravetz, S.; Ouerfelli, O.; Spriggs, D. R.; Danishefsky, S. J.; Sabbatini, P. J., Cancers (Basel) 2016, 8 (4).
10. Codington, J. F.; Haavik, S.; Nikrui, N.; Kuter, I.; Vassileva, C.; Zhang, H.; Matson, S.; Chen, X.; Wu, Z., Cancer 2002, 94 (3), 803-13.
11. Thingstad, T.; Vos, H. L.; Hilkens, J., Biochem. J. 2001, 353, 33-40.
12. Wang, D.; Tang, J.; Liu, S.; Huang, J., J Immunol Res 2015, 2015, 510810.
13. Haavik, S.; Nilsen, M.; Thingstad, T.; Barsett, H.; Renouf, D. V.; Hounsell, E. F.; Codington, J. F., Glycoconj. J. 1999, 16, 229-236.
14. Palma, A. S.; Liu, Y.; Childs, R. A.; Herbert, C.; Wang, D.; Chai, W.; Feizi, T., Biochem Biophys Res Commun 2011, 408 (4), 548-52.
15. Tadano-Aritomi, K.; Kubo, H.; Ireland, P.; Okuda, M.; Kasama, T.; Handa, S.; Ishizuka, I., Carbohydrate Research 1995, 273 (1), 41-52.
16. Niimura, Y.; Ishizuka, I., Glycobiology 2006, 16 (8), 729-35.
17. Chandrasekaran, E. V.; Xue, J.; Neelamegham, S.; Matta, K. L., Carbohydr Res 2006, 341 (8), 983-94.
18. Hiraiwa, N.; Iida, N.; Ischizuka, I.; Itai, S.; Shigeta, K.; Kannagi, R.; Fukuda, Y.; Imura, H., Cancer Res. 1988, 48, 6769-6774.
19. Hiraiwa, N.; Fukuda, Y.; Imura, H.; Tadano-Aritomi, K.; Nagai, K.-I.; Ischizuka, I.; Kannagi, R., Cancer Res. 1990, 50, 2917-2928.
20. Capon, C.; Maes, E.; Michalski, J. C.; Leffler, H.; Kim, Y. S., Biochem. J. 2001, 358, 657-664.
21. Zhang, P.; Wang, K.; Zhang, J.; Li, C.; Guan, H., European Journal of Organic Chemistry 2015, 2015 (3), 570-583.
22. Hirose, H.; Tamai, H.; Gao, C.; Imamura, A.; Ando, H.; Ishida, H.; Feizi, T.; Kiso, M., Org Biomol Chem 2015, 13 (45), 11105-17.
23. Brockhausen, I.; Stanley, P., Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 10. 2017.
24. Schnaar, R. L.; Kinoshita, T., Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2015-2017.
25. Boltje, T. J.; Buskas, T.; Boons, G. J., Nat Chem 2009, 1 (8), 611-22.
26. Noti, C.; de Paz, J. L.; Polito, L.; Seeberger, P. H., Chemistry 2006, 12 (34), 8664-86.
27. van der Vorm, S.; van Hengst, J. M. A.; Bakker, M.; Overkleeft, H. S.; van der Marel, G. A.; Codee, J. D. C., Angew Chem Int Ed Engl 2018, 57 (27), 8240-8244.
28. Fusari, M.; Fallarini, S.; Lombardi, G.; Lay, L., Bioorg Med Chem 2015, 23 (23), 7439-47.
29. Wu, X.; Schmidt, R. R., J. Org. Chem. 2004, 69 (6), 1853-1857.
30. Crich, D.; Vinogradova, O., J. . Org. Chem. 2007, 72 (17), 6513-6520.
31. Dimakos, V.; Taylor, M. S., Chem Rev 2018, 118 (23), 11457-11517.
32. t Hart, I. M. E.; Li, T.; Wolfert, M. A.; Wang, S.; Moremen, K. W.; Boons, G. J., Org Biomol Chem 2019, 17 (31), 7304-7308.
33. Rio, S.; Beau, J.-M.; Jacquinet, J.-C., Carbohydr Res 1991, 219, 71-90.

Chapter 5

Chemoenzymatic synthesis of *C. jejuni* heptose-ganglioside mimics to study antibody binding patterns in Guillain-Barré Syndrome

Abstract: Guillian-Barré syndrome (GBS) is an autoimmune disease mainly affecting peripheral neurons, resulting in muscle weakness. One third of GBS cases is triggered by a *Campylobacter jejuni* infection. The lipooligosaccharides (LOS) in the outer membrane of *C. jejuni* can contain ganglioside mimics, which are thought to trigger cross-reactive antibodies for human endogenous gangliosides. The aim of this study is to assess binding of truncated LOS structures and endogenous gangliosides to GBS serum antibodies. The smallest difference between LOS and natural gangliosides is the monosaccharide L-glycero-D-manno-heptose (Hep). Therefore, heptosyl-ganglioside mimics where synthesized by a chemoenzymatic approach, starting with the chemical synthesis of a Galβ1,3Hep disaccharide. Further enzymatic extensions by PmST1, CgtA and CgtB provided Hep-GM3, Hep-GM2 and Hep-GM1. All structures were equipped with an aminopentyl linker for microarray immobilization. Glycan microarray studies will reveal the effect of microbial-specific heptose on antiganglioside antibody binding.

Introduction

Autoimmune diseases are triggered by autoantibodies. These antibodies can be naturally present or are induced after infection by viruses, bacteria or parasites.[1,2] Antiganglioside antibodies are involved in acute and chronic neuropathies, such as Alzheimer's disease, Parkinson's disease and amyotrophic lateral sclerosis (ALS).[3]

Guillain-Barré Syndrome (GBS) is an acute peripheral neuropathic disease, in which antiganglioside antibodies are mainly formed after infections.[4] GBS is the most common paralytic disease in the world, affects 0.6-4/100,000 person each year and symptoms vary from weakness and tingling sensations up to complete muscular failure.[5,6,7]

Approximately one quarter to a third of GBS cases is preceded by *Campylobacter jejuni* infection.[5,6] These bacteria bear glycans in their core lipooligosaccharides (LOS) that are structurally similar to gangliosides present in human nervous tissue. In some cases, *C. jejuni* elicits antibodies that cross-react with endogenous gangliosides resulting in an autoimmune disease.[2,4,7]

C. jejuni is a gram-negative bacterium and is the main cause for gastroenteritis. The full genome has been sequenced in 2000. High variability was detected in the *C. jejuni* genome in sequences responsible for glycosyltransferases and surface structures such as its LOS.[8] Gram-negative bacteria have a double membrane and the outer membrane contains lipopolysaccharides (LPS). The outer membrane of *C. jejuni* contains a short form of LPS, LOS, which lacks a repeating oligosaccharide polymer.[9] LOS consists of lipid A and an oligosaccharide core, which can be further divided into an inner core and outer core. Various LOS structures of *C. jejuni* have been described which may be beneficial for adaptation or avoiding host defenses. The inner core contains unusual monosaccharides such as 3-deoxy-α-D-manno-2-octulopyranosonic acid (Kdo) and L-glycero-D-manno-heptose (Hep) and is highly conserved among *C. jejuni* serotypes (Figure 1A).[10,11,12] The outer core varies between bacterial strains and often contains sialylated structures mimicking gangliosides (Figure 1).[11,12] LOS ganglioside mimics of *C. jejuni* are biosynthesized by the enzymes Cst-II (α2,3- and α2,8-sialyltransferase), CgtA (β1,4-N-acetylgalactosamine-transferase) and CgtB (β1,3-galactosyltransferase).[13]

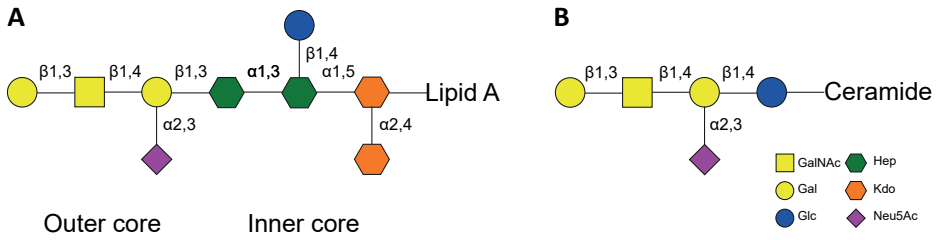


Figure 1. (**A**) The lipooligosaccharide (LOS) core structure from *C. jejuni* mimicking GM1. (**B**) Human GM1 ganglioside.

Isolation, separation and characterization of LOS glycans from *C. jejuni* has been performed by various techniques.[14,15,16] The presence of *C. jejuni* molecular mimics in GBS is mainly supported by the detection of antiganglioside antibodies in patient serum.

Various antiganglioside antibodies have been identified in serum samples of GBS patients, mainly directed to the gangliosides GM1, GD3, GD1a, GD1b, GT1a, GT1b and GQ1b and rarely to GM2. Antibodies can be directed to one single ganglioside, but also to mixtures of gangliosides.[3,11,17] As an example, antibodies towards GD1a/GD1b or GQ1b/GM1 complexes have been described and seem to be associated with a severe form of GBS.[18]

GM1, GM2 and GM3 are major monosialylgangliosides present in the peripheral nervous system.[19] In addition, the gangliosides GM1 and GD1a are highly enriched at the nodes of Ranvier, which is an important location for the generation and continuance of action potentials.[20] Antibodies isolated from LOS-immunized rabbits and a GBS patient where shown to cross-react with ganglioside epitopes at the nodes of Ranvier.[21] Furthermore, the GM1 ganglioside is mainly expressed in motor neurons.[22] This explains why GM1 antibodies are mainly detected in a motor subtype of GBS and in in multifocal motor neuropathy (MMN), although no *C. jejuni* infections have been reported preceding MMN.[23]

The minimal difference of the LOS ganglioside mimics and the natural gangliosides is the monosaccharide L-*glycero*-D-manno-heptose (Hep, Figure 1). Heptose is strictly found in microbes and is present in LPS of almost all Gram-negative bacteria.[24]

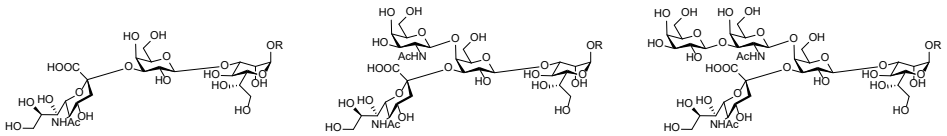


Figure 2. Chemical structures of truncated *C. jejuni* LOS ganglioside mimics GM3 (**2**), GM2 (**3**) and GM1 (**4**): synthetic targets for this study. R = $(CH)_5NH_2$ linker for immobilization.

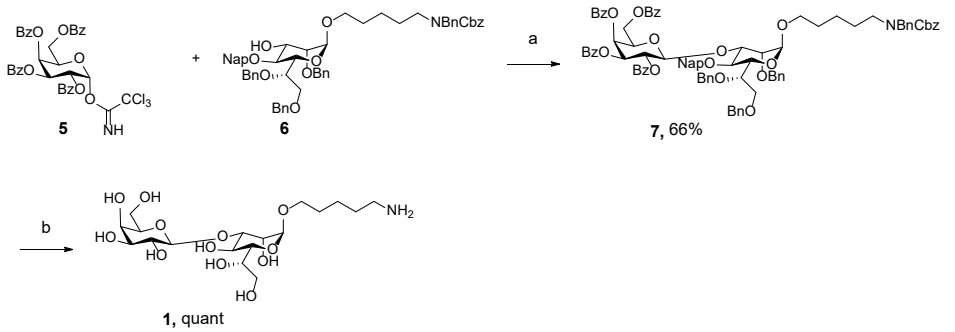
To study the importance of the heptoside for eliciting antibodies cross reactive with ganglioside, we aimed to synthesize Hep-gangiloside mimics **2**-**4** (Figure 2). An aminopentyl linker at the reducing end was incorporated for immobilization on NHS-activated microarray slides. Binding of serum antiganglioside antibodies of GBS patients to gangliosides and the bacterial heptose-analogs will be compared.

Earlier synthetic efforts have focused on full chemical assembly of LOS oligosaccharides[25], however bacterial enzymes that can synthesis the ganglio-series oligosaccharides GM3, GM2 and GM1 are readily available.[26] Therefore, we aimed for a chemoenzymatic approach, starting by chemically synthesizing Gal-β1,3-Hep followed by enzymatic extension to obtain the desired ganglioside mimics. To our knowledge, these truncated LOS structures have not been synthesized before. The heptose-gangliosides will make it possible to gain insight in how GBS serum antibodies bind to gangliosides and its mimics.

Results and Discussion

Chemical synthesis of galactose-heptose disaccharide

Heptosyl acceptor **6**, equipped with an α-anomeric aminopentyl linker, was chemically synthesized starting from D-mannose according to a previously reported procedure.[27,28] Galactostyl donor **5** was synthesized from D-galactose as described before (Chapter 4).[29] Heptosyl acceptor **6** was glycosylated with galactosyl donor **5** in the presence of TMSOTf to give disaccharide **7** in a yield of 66% after purification by silica gel column chromatography (Scheme 1). Disaccharide **7** was debenzoylated, followed by catalytic hydrogenation over $Pd(OH)_2/C$ and H_2 in a mixture of $MeOH/H_2O/HOAc$ to remove the benzyl, Nap and Cbz protecting groups. Disaccharide **1** was purified by Bio-Gel P-2 size exclusion chromatography. After careful characterization by HRMS and 1D and 2D NMR analysis disacchardie **1** was used for further enzymatic extensions.



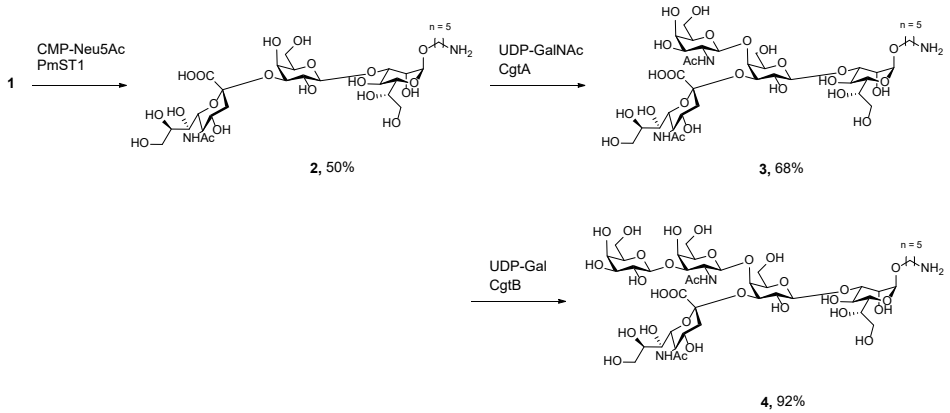
Scheme 1. Chemical synthesis of disaccharide **7** from galactosyl donor **5** and heptosyl acceptor **6**. Deprotections provided the disaccharide Galβ1,3-Hep **1**. Reagents and conditions: (a) TMSOTf, DCM, 4 Å MS, -10 °C to 0 °C; (b) i. NaOMe, $MeOH/CH_2Cl_2$; ii. $Pd(OH)_2/C$, H_2, $MeOH/H_2O/HOAc$.

Enzymatic synthesis of ganglioside mimics of GM3, GM2 and GM1a

Enzymatic carbohydrate synthesis is often performed by bacterial enzymes, since these are generally easier to express and have broader substrate specificities compared to similar mammalian enzymes. Various bacterial enzymes have been used to synthesize ganglio-series glycans including GM3, GM2 and GM1.[26,30] GM3 has been synthesized by the α2,3-sialyltransferases: PmST1 (M144D), PmST3 and Cst-II.[26,30,31] GM2 has been synthesized from GM1 by the β1,4-GalNAc transferase CgtA. A study dealing with examing substrate specificities of CgtA has indicated a requirement of a sialic acid residue, which preferentially should be α2,3- linked.[26] GM1 is synthesized from GM2 by the β1,3-Gal transferase CgtB. Variants of CgtB showed different preferences when tested on a series of substrates (e.g. GM2-, GalNAcα-, GalNAcβ-) linked to an 6-(fluorescein-5-carbaxamido)hexanoic acid succinimidyl ester.[32] Furthermore, we and others found that CgtB is able to transfer more than one galactosyl residue if an excess of UDP-galactose is used (Chapter 3).[30]

Disaccharide **1** was extended with α2,3-linked Neu5Ac moiety to give trisaccharide **2** (Hep-GM3) using PmST1 and CMP-Neu5Ac (Scheme 2). Trisaccharide **2** was isolated in an average yield (50 %), although LC-MS analysis indicated full conversion. The average yield might have been caused by presence of salts in starting material **1** or

losses during P2 Bio-Gel purification. The β1,4 linkage between GalNAc and lactose is catalyzed by the bacterial enzyme CgtA. However, initial attempts on UDP-GalNAc extension by commercially available CgtA did not show any activity on disaccharide **2** or 2,3-Neu5Aclactose (data not shown). In-house expressed and purified CgtA proved inactive as well. CgtA was expressed again and the crude lysate was directly used to successfully convert trisaccharide **2** to tetrasaccharide **3**. Still, CgtA lysate was quite labile and lost its activity after overnight storage at 4 °C. Immediate storage at -20 °C was required to maintain enzyme activity over a longer period of time.[30] Hep-GM2 (**3**) was successfully extended by a β1,3-Gal moiety by treatment with CgtB in the presence of UDP-Gal to provide pentasaccharide **4** (Hep-GM1). Each enzymatic conversion was monitored by LC-MS, and products were purified by Bio-Gel P-2 size exclusion chromatography and the structural identified was confirmed by NMR spectroscopy (1D and 2D) and HRMS.



Scheme 2. Enzymatic synthesis of LOS ganglioside mimics: Hep-GM3 (**2**), Hep-GM2 (**3**) and Hep-GM1 (**4**). Reagents and conditions: substrate (10 mM), sugar-nucleotide (11 mM), Tris buffer pH 7.7 (100 mM), $MgCl_2$ (20 mM) and enzyme (pmST1, CgtA or CgtB) were shaken at 37 °C.

Saccharides **1** - **4** will be immobilized on a NHS-activated microarray slide, together with the oligosaccharide moieties of the human gangliosides. Binding studies will be performed with various serum samples derived from GBS patients. This study will elucidate the effect of the heptoside on antiganglioside antibody binding.

Conclusions

Molecular mimicry by the lipooligosaccharides of *C. jejuni* has been identified as a trigger for Guillain-Barré Syndrome. However, exact immune responses to these ganglioside mimics are not well understood. Our aim is to study the effect of bacterial saccharide components, such as heptose, on eliciting antiganglioside antibodies. For this purpose, we designed a chemoenzymatic strategy to obtain heptosyl containing analogs of the oligosaccharide moiety of the gangliosides GM3, GM2 and GM1. The disaccharide, Hepβ1,3Galα$(CH_2)_5NH_2$, was successfully obtained by chemical glycosylation of a galactosyl donor with a properly protected heptosyl acceptor. The deprotected disaccharide was then extended by the bacterial enzymes pmST1, CgtA and CgtB to

provide Hep-GM3, Hep-GM2 and Hep-GM1. The four heptosyl-ganglioside mimics will be immobilized on NHS-activated microarray slides, along with the corresponding ganglioside oligosaccharides. Serum antibodies from GBS patients will be screened for binding towards these gangliosides and mimics to provide insight in the effect of heptoside to trigger cross reactive antibodies during *C. jejuni* infection and triggering GBS.provide Hep-GM3, Hep-GM2 and Hep-GM1. The four heptosyl-ganglioside mimics will be immobilized on NHS-activated microarray slides, along with the corresponding ganglioside oligosaccharides. Serum antibodies from GBS patients will be screened for binding towards these gangliosides and mimics to provide insight in the effect of heptoside to trigger cross reactive antibodies during *C. jejuni* infection and triggering GBS.

Experimental section

Chemical synthesis

NMR nomenclature

The monosaccharides of *C. jejuni* core heptose-ganglioside mimics have been labeled as shown in **Figure S1.** Starting from the reducing end these were labeled Hep-I, Gal-II, Neu5Ac-III, GalNAc-IV and Gal-V.

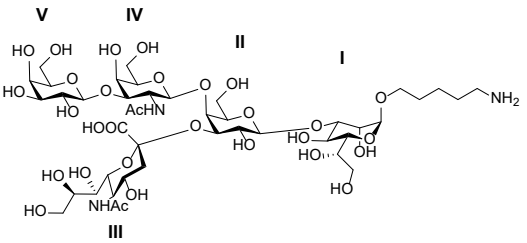
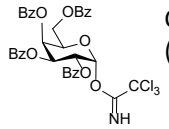
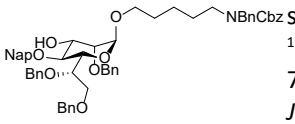


Figure S1. Monosaccharide labeling system for heptose-ganglioside mimics.

Experimental procedures

2,2,2-Trichloroacetimidate 2,3,4,6-tetra-*O*-benzoyl-α-D-galactopyranoside (5). 

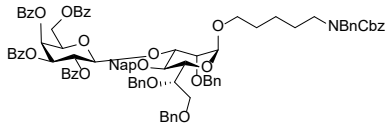
Compound **5** was synthesized from galactose as described before (Chapter 4: 5).[29]

***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2,6,7-tetra-*O*-benzyl-4-*O*-(2-naphthyl) methyl-L-*glycero*-α-D-*manno*heptopyranoside (6).** Compound **6** was synthesized in 

similar fashion as reported for α-D-mannosopyranoside.[27,28] ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.82 – 7.70 (3H, m, H-Ar), 7.48 – 7.03 (29H, m, H-Ar), 5.16 (2H, d, *J* = 20.7 Hz, CH_2), 5.02 (1H, d, *J* = 11.3 Hz, CH*H*), 4.94 (1H, d, *J* = 10.8 Hz, H-1), 4.83 (1H, d, *J* = 11.9 Hz, CH*H*), 4.79 (1H, d, *J* = 10.9 Hz, CH*H*), 4.56 – 4.41 (6H, m, 3 x CH_2), 4.16 (1H, t, *J* = 6.1 Hz, H-6), 4.06 – 3.99 (1H, m, H-3), 3.92 (1H, t, *J* = 9.4 Hz, H-4), 3.81 (1H, s, CH*H*,

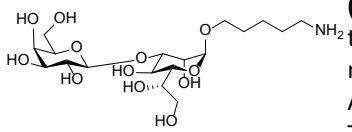
H-7a), 3.78 – 3.65 (3H, m, H-5; CH*H*, H-7b; H-2), 3.61 – 3.49 (1H, m, CH*H*, pentyl), 3.32 – 3.07 (3H, m, CH*H* + CH_2, pentyl), 2.37 (2H, d, *J* = 9.8 Hz, O*H*), 1.51 – 1.35 (4H, m, 2 x CH_2, pentyl), 1.23 – 1.05 (2H, m, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 138.7, 138.1, 137.9, 137.8, 136.2, 133.3, 132.9, 128.6, 128.4, 128.3, 128.1, 128.0, 127.9, 127.7, 127.7, 127.6, 127.6, 126.3, 125.9, 125.8, 96.6 (C-1), 78.5 (C-2), 76.3 (C-4), 75.1 (C-6), 74.5, 73.4, 72.8, 72.7, 72.3 (C-3), 70.6 (C-5), 70.3, 70.1, 67.3, 67.2, 50.5, 50.2, 47.1, 46.1, 29.2, 28.0, 27.6, 23.4. ESI HRMS (*m/z*): $[M + NH_4]^+$ calcd for $C_{59}H_{63}NO_9$; 947.4841 found 947.4884.

***N*-(Benzyl)-benzyloxycarbonyl-5-aminopentyl 2,3,4,6-tetra-*O*-benzoyl-β-D-galactopyranosyl-(1→3)-2,6,7-tetra-*O*-benzyl-4-*O*-(2-naphthyl)methyl-L-*glycero*-α-D-*manno*heptopyranoside (7).** A mixture of acceptor **6** (83 mg, 0.089 mmol), donor



5 (99 mg, 0.134 mmol) and 4 Å molecular sieves was stirred in DCM (1.5 mL) for 1 h. The reaction was cooled to -10 °C and TMSOTf (1.6 µL, 0.1 mmol in 100 µL DCM) was added. After 1 h the reaction was quenched with Et_3N (at 0 °C), filtered over Celite and concentrated *in vacuo*. The obtained residue was purified by silica column chromatography using Toluene:EtOAc as the eluent (1:0 to 10:1 v/v) and SX1 chromatography using Toluene:Acetone as the eluent (1:1 v/v) to afford compound **7** (85 mg, 66 %). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.00 – 7.10 (50H, m, H-Ar), 7.06 – 6.98 (2H, m, H-Ar), 5.94 (1H, s, H-4, Gal-II), 5.93 – 5.87 (1H, m, H-2, Gal-II), 5.60 (1H, dd, *J* = 10.3, 3.3 Hz, H-3, Gal-II), 5.33 (1H, d, *J* = 11.0 Hz, CH*H*), 5.23 – 5.12 (2H, m, *J* = 23.4 Hz, CH_2), 5.08 (1H, s, H-1, Gal-II), 4.81 – 4.71 (2H, m, H-1, Hep-I; CH*H*), 4.53 – 4.18 (13H, m, 4 x CH_2; H-6, Gal-II; H-3, Hep-I; H-4, Hep-I; H-5, Gal-II), 4.16 – 4.09 (1H, m, H-6, Hep-I), 3.81 – 3.73 (2H, m, H-5, Hep-I; H-7a, Hep-I), 3.71 (2H, s, H-2, Hep-I; H-7b, Hep-I), 3.49 (1H, s, CH*H*), 3.16 (3H, d, *J* = 40.6 Hz, CH_2, CH*H*, pentyl), 1.50 – 1.29 (4H, m, 2 x CH_2, pentyl), 1.10 (2H, dd, *J* = 37.4, 8.1 Hz, CH_2, pentyl). ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.9 (C=O, Bz), 165.6 (C=O, Bz), 165.5 (C=O, Bz), 165.1 (C=O, Bz), 138.8, 138.2, 133.5, 133.3, 133.2, 132.8, 129.9, 129.8, 129.8, 129.7, 129.4, 129.2, 129.1, 129.0, 128.8, 128.6, 128.6, 128.4, 128.4, 128.3, 128.3, 128.3, 128.0, 128.0, 127.9, 127.7, 127.6, 127.5, 127.3, 127.2, 125.9, 125.6, 125.3, 99.9 (C-1, Gal-II), 97.6 (C-1, Hep-I), 80.1 (C-3, Hep-I), 76.0 (C-2, Hep-I), 75.0 (C-6, Hep-I), 74.4, 73.4 (C-4, Hep-I), 72.82, 72.72, 72.1 (C-3, Gal-II), 71.6 (C-5, Hep-I), 71.2 (C-5, Gal-II), 70.4 (C-2, Gal-II), 68.0 (C-4, Gal-II), 67.2, 67.2, 61.6, 50.5, 50.2, 47.1, 46.1, 29.0, 27.9, 27.6, 23.4, 23.3, 21.5. ESI HRMS (*m/z*): $[M + NH_4]^+$ calcd for $C_{93}H_{89}NO_{18}$; 1525.6418 found 1525.6489

5-Aminopentyl β-D-Galactopyranosyl-(1→3)-L-*glycero*-α-D-*manno*heptopyranoside



(1). Freshly prepared NaOMe in MeOH (1 mL) was added to a solution of disaccharide **7** in DCM/MeOH (1 mL/2 mL). The mixture was stirred overnight, neutralized with Amberlite H^+ resin, filtered and concentrated *in vacuo*. The obtained intermediate was dissolved in MeOH/H_2O/HOAc (2.5/0.5/0.5 mL), followed by addition of $Pd(OH)_2$/C (60 mg) and the reaction mixture was stirred overnight under a hydrogen atmosphere. After that time, it was filtered over Celite, concentrated *in vacuo* and the residue purified by BioGel P-2 column chromatography using BuOH:H_2O (5:95 v/v) as the eluent to give the desired product **1**

(25 mg, quant). ^{1}H NMR (600 MHz, D_2O) δ 4.86 (1H, s, H-1, Hep-I), 4.47 (1H, d, *J* = 7.7 Hz, H-1, Gal-II), 4.07 (1H, s, H-2, Hep-I), 4.01 (1H, t, *J* = 6.5 Hz, H-6, Hep-I), 3.98 – 3.90 (2H, m, H-3, Hep-I; H-4, Hep-I), 3.88 (1H, s, H-4, Gal-II), 3.78 – 3.61 (7H, m, CH_2, pentyl; H-6, Gal-II; H-5, Gal-II; H-6a, Gal-II; H-3, Gal-II), 3.60 – 3.53 (2H, m, H-5, Hep-I; H-2, Gal-II), 3.52 – 3.46 (1H, m, H-6b, Gal-II), 2.94 (2H, t, *J* = 7.6 Hz, CH_2, pentyl), 1.69 – 1.55 (4H, m, 2 x CH_2, pentyl), 1.46 – 1.33 (2H, m, CH_2, pentyl). ^{13}C NMR (151 MHz, D_2O) δ 100.9 (C-1, Gal-II), 99.3 (C-1, Hep-I), 78.6 (C-3, Hep-I), 75.3 (C-5, Gal-II), 72.6 (C-3, Gal-II), 71.1 (C-5, Hep-I), 70.7 (C-2, Gal-II), 68.7 (C-6, Hep-I), 68.6 (C-4, Gal-II), 67.7 (C-2, Hep-I), 67.3, 64.4 (C-4, Hep-I), 62.6, 61.1, 39.3, 28.0, 26.5, 22.5. ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{18}H_{35}NO_{12}$; 458.2232 found 458.2238

Enzymatic synthesis

Expression of enzymes CgtA and CgtB from *Campylobacter jejuni*

Materials: The full-length codon-optimized gene encoding for CjCgtB (ENA: CAL35256.1) was synthesized by Genscript and subcloned into pET15b+ (NdeI, BamHI). The same applied to CjCgtA (ENA: AAL05993.1), except the first 30 base pairs (first 10 amino acids) were omitted. Bl21(DE3) (C2527H) cells were purchased from New England Biolabs (NEB). Ampicillin (sodium salt) (14417) was ordered through Cayman Chemicals. Imidazole (56750), TRIS-HCl (T5941), NaCl (S9888), lysozyme (62971) and triton X-100 (T8787) came from Sigma-Aldrich. 2xYT medium (BP97432) and 2xYT Agar (BP2466) were ordered from Fisher BioReagents. Isopropyl-beta-D-thiogalactopyranoside (IPTG) (R0393) as well as GelCode Blue Stain Reagent (24592) was ordered from Thermo Scientific. The Ni-NTA resin (17-5318-01) and the Superdex 200 Increase 10/300 GL column (28990944) were ordered from GE Healthcare. Sartorius provided Vivaspin 6, 10000 MWCO (VS0602) spinfilters. For SDS-PAGE, SurePAGE, Bis-Tris, 10x8, 4-12%, 12 wells (M00653) from GenScript were used. Laemmli sample buffer 2x (161-0737) and Precision Plus Protein Dual Color Standards (161-0374) were ordered from Bio-Rad Laboratories. A standard buffer consisting of 50 mM TRIS-HCl and 250 mM NaCl pH 8 was made. Imidazole was added to the standard buffer at concentrations of 20 mM, 50 mM or 250 mM to make wash 1, wash 2 and elution buffer (re-adjusted to pH 8 if needed) respectively. To wash 1, lysozyme (1mg/mL) and triton X-100 (0.1%) was added to make the lysis buffer.

Protocol: BL21(DE3) cells were transformed with each of the individual vectors, a colony was picked from the 2xYT agar plate with ampicillin (100 µg/mL) and expanded to a cell culture volume of 500 mL ampicillin (100 µg/mL) containing 2xYT medium at 37°C. The cells were induced at OD_{600} = 0.6 with IPTG (final concentration was 1 mM) and cultured overnight at 20°C. Then, cells were pelleted at 3000 *xg*, resuspended in lysis buffer (5% of culture volume), incubated at 37°C for 1 h and sonicated for 30' on ice. As His-tag purification of CjCgtA or CjCgtB inherently resulted in loss of activity, clarified lysate was used directly or stored at -20°C for later use. Further characterization of the enzymes was done by SurePage 4-12% gel, using Laemmli sample buffer and the Precision Plus Protein marker as reference.

Experimental procedures

5-Aminopentyl αNeu5Ac-(2→3)-β-D-Galactopyranosyl-(1→3)-L-*glycero*-α-D-*manno*heptopyranoside (2). PmST1 M144D (75 µL of a 2 mg/mL solution) was added to compound **1** (14 mg, 30.6 µmol, 10 mM final concentration) in H_2O with CMP-Neu5Ac (11 mM), $MgCl_2$ (20 mM) and Tris buffer (100 mM, pH = 7.7). After 4 h the mixture was lyophilized overnight. The residue was redissolved in H_2O and purified by Bio-Gel P2 size exclusion column chromatography to provide compound **2** (11.5 mg, 50 %). ^{1}H NMR (600 MHz, D_2O) δ 4.88 (1H, s, H-1, Hep-I), 4.58 (1H, d, *J* = 7.8 Hz, H-1, Gal-II), 4.13 – 4.08 (2H, m, H-2, Hep-I), 4.06 – 3.47 (22H, m), 3.00 – 2.94 (2H, m, CH_2, pentyl), 2.75 (1H, dd, *J* = 12.4, 4.6 Hz, H-3eq, Neu5Ac-III), 2.01 (3H, s, CH_3, NHAc), 1.78 (1H, t, *J* = 12.1 Hz, H-3ax, Neu5Ac-III), 1.71 – 1.56 (4H, m, 2 x CH_2, pentyl), 1.49 – 1.37 (2H, m, CH_2, pentyl). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{29}H_{52}N_2O_{20}$; 749.3186 found 749.3199

5-Aminopentyl 2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→4)-[αNeu5Ac-(2→3)]-β-D-Galactopyranosyl-(1→3)-L-*glycero*-α-D-*manno*heptopyranoside (3). CgtA (fresh lysate: 250 µL) was added to compound **2** (4.1 mg, 5.5 µmol, 10 mM final concentration) in H_2O with UDP-GalNAc (11 mM), $MgCl_2$ (20 mM) and Tris buffer (100 mM, pH = 7.7). After 19 h the mixture was lyophilized overnight. The residue was redissolved in H_2O and purified by Bio-Gel P2 size exclusion chromatography to provide compound **3** (3.5 mg, 68 %). ^{1}H NMR (600 MHz, D_2O) δ 4.88 (1H, s, H-1, Hep-I), 4.71 (1H, d, *J* = 8.5 Hz, H-1, GalNAc-IV), 4.57 (1H, d, *J* = 7.9 Hz, H-1, Gal-II), 4.15 (1H, d, *J* = 9.9 Hz, H-3, Gal-II), 4.10 (2H, s, H-4, Gal-II; H-2, Hep-I), 4.04 (1H, t, *J* = 6.5 Hz, H-6, Hep-I), 3.98 – 3.93 (2H, m, H-4, Hep-I; H-3, Hep-I), 3.92 – 3.56 (19H, m), 3.54 – 3.45 (2H, m, CHH, pentyl; H-6, Neu5Ac-III), 3.41 (1H, t, *J* = 8.8 Hz, H-2, Gal-II), 2.98 (2H, t, *J* = 7.6 Hz, CH_2, pentyl), 2.66 (1H, dd, *J* = 12.6, 4.4 Hz, H-3eq, Neu5Ac-III), 2.01 (6H, s, 2 x CH_3, NHAc), 1.90 (1H, t, *J* = 12.0 Hz, H-3ax, Neu5Ac-III), 1.72 – 1.57 (4H, m, 2 x CH_2, pentyl), 1.50 – 1.36 (2H, m, 2 x CH_2, pentyl). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{37}H_{65}N_3O_{25}$; 952.3980 found 952.3962.

5-Aminopentyl β-D-Galactopyranosyl-(13)-2-acetamido-2-deoxy-β-D-galacto-pyranosyl-(1→4)-[αNeu5Ac-(2→3)]-β-D-Galactopyranosyl-(1→3)-L-*glycero*-α-D-*manno*heptopyranoside (4). CgtB (lysate: 50 µL) was added to compound **3** (1.2 mg, 1.3 µmol, 10 mM final concentration) in H_2O with UDP-Gal (11 mM), $MgCl_2$ (20 mM) and Tris buffer (100 mM, pH = 7.7). After 20 h the mixture was lyophilized overnight. The residue was redissolved in H_2O and purified by Bio-Gel P2 size exclusion chromatography to provide compound **4** (1.3 mg, 92 %).^{1}H NMR (600 MHz, D_2O) δ 4.91 (1H, s, H-1, Hep-I), 4.79 – 4.76 (1H, m, H-1, GalNAc-IV), 4.60 (1H,

d, *J* = 7.9 Hz, H-1, Gal-II), 4.55 (1H, d, *J* = 7.8 Hz, H-1, Gal-V), 4.21 – 4.16 (2H, m, H-3, Gal-II; H-4, GalNAc-IV), 4.15 – 4.10 (2H, m, H-4, Gal-II; H-2, Hep-I), 4.09 – 4.02 (2H, m, H-6, Hep-I; H-2, GalNAc-IV), 4.01 – 3.96 (2H, m, H-3, Hep-I; H-4, Hep-I), 3.93 (1H, d, *J* = 3.2 Hz, H-4, Gal-V), 3.90 – 3.49 (24H, m), 3.47 – 3.41 (1H, m, H-2, Gal-II), 3.01 (2H, t, *J* = 7.6 Hz, CH_2, pentyl), 2.68 (1H, dd, *J* = 12.7, 4.5 Hz, H-3eq, Neu5Ac-III), 2.04 (3H, s, CH_3, NHAc), 2.03 (3H, s, CH_3, NHAc), 1.94 (1H, t, *J* = 12.0 Hz, H-3ax, Neu5Ac-III), 1.74 – 1.61 (4H, m, 2 x CH_2, pentyl), 1.52 – 1.39 (2H, m, CH_2, pentyl). ESI HRMS (*m/z*): $[M + H]^+$ calcd for $C_{43}H_{75}N_3O_{30}$; 1114.4508 found 1114.4531.

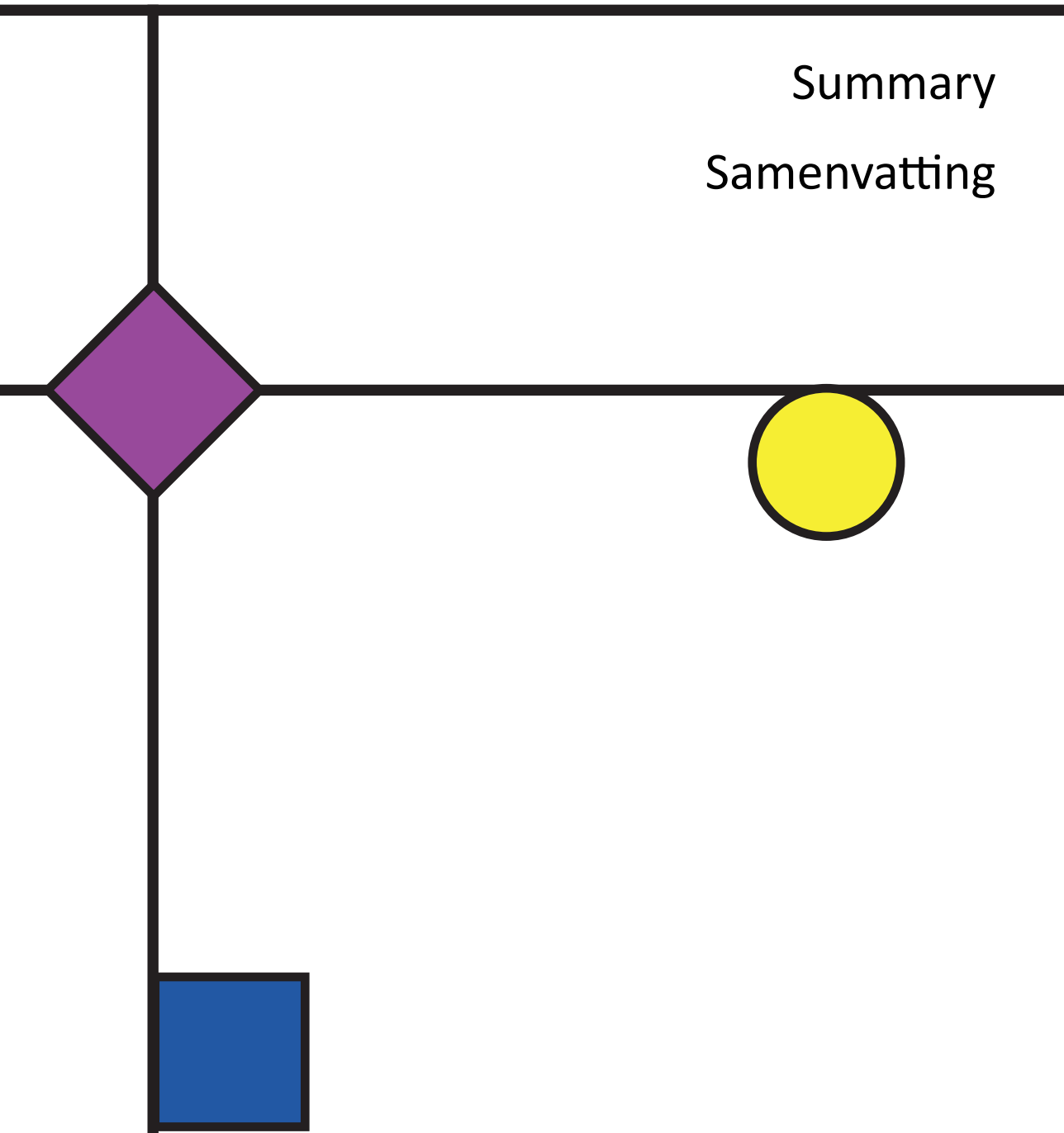
References

1. Goodfellow, J. A.; Willison, H. J., Prog Mol Biol Transl Sci 2018, 156, 355-382.
2. Rivera-Correa, J.; Rodriguez, A., Trends Immunol 2018, 39 (7), 515-522.
3. Wanleenuwat, P.; Iwanowski, P.; Kozubski, W., J Neurol Sci 2020, 408, 116576.
4. van den Berg, B.; Walgaard, C.; Drenthen, J.; Fokke, C.; Jacobs, B. C.; van Doorn, P. A., Nat Rev Neurol 2014, 10 (8), 469-82.
5. Israeli, E.; Agmon-Levin, N.; Blank, M.; Chapman, J.; Shoenfeld, Y., Clin Rev Allergy Immunol 2012, 42 (2), 121-30.
6. Hughes, R. A. C.; Cornblath, D. R., The Lancet 2005, 366 (9497), 1653-1666.
7. Schnaar, R. L.; Kinoshita, T., Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2015-2017.
8. Parkhill, J.; Wren, B. W.; Mungall, K.; Ketley, J. M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R. M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A. V.; Moule, S.; Pallen, M. J.; Penn, C. W.; Quail, M. A.; Rajandream, M. A.; Rutherford, K. M.; van Vliet, A. H.; Whitehead, S.; Barrell, B. G., Nature 2000, 403 (6770), 665-668.
9. Burnham, P. M.; Hendrixson, D. R., Nat Rev Microbiol 2018, 16 (9), 551-565.
10. Karlyshev, A. V.; Ketley, J. M.; Wren, B. W., FEMS Microbiol Rev 2005, 29 (2), 377-90.
11. Moran, A. P.; Prendergast, M. M.; Appelmelk, B. J., FEMS Immunol Med Microbiol. 1996, 16 (2), 105-115.
12. Kanipes, M. I.; Tan, X.; Akelaitis, A.; Li, J.; Rockabrand, D.; Guerry, P.; Monteiro, M. A., J Bacteriol 2008, 190 (5), 1568-74.
13. Gilbert, M.; Brisson, J.-R.; Karwaski, M.-F.; Michniewicz, J.; Cunningham, A.-M.; Wu, Y.; N.M., Y.; Wakarchuk, W. W., J. Biol. Chem. 2000, 275, 3896-3906.
14. Kilar, A.; Dornyei, A.; Kocsis, B., Mass Spectrom Rev 2013, 32 (2), 90-117.
15. Houliston, R. S.; Vinogradov, E.; Dzieciatkowska, M.; Li, J.; St Michael, F.; Karwaski, M. F.; Brochu, D.; Jarrell, H. C.; Parker, C. T.; Yuki, N.; Mandrell, R. E.; Gilbert, M., J Biol Chem 2011, 286 (14), 12361-70.
16. Semchenko, E. A.; Day, C. J.; Wilson, J. C.; Grice, I. D.; Moran, A. P.; Korolik, V., BMC Microbiol. 2010, 10, 305.
17. Jasti, A. K.; Selmi, C.; Sarmiento-Monroy, J. C.; Vega, D. A.; Anaya, J. M.; Gershwin, M. E., Expert Rev Clin Immunol 2016, 12 (11), 1175-1189.
18. Kaida, K.; Kusunoki, S., Future Lipidology 2008, 3 (4), 425-434.
19. Willison, H. J.; Yuki, N., Brain 2002, 125, 2591-2625.
20. Susuki, K.; Yuki, N.; Schafer, D. P.; Hirata, K.; Zhang, G.; Funakoshi, K.; Rasband, M. N., Exp Neurol 2012, 233 (1), 534-42.
21. Moran, A. P.; Annuk, H.; Prendergast, M. M., J Neuroimmunol 2005, 165 (1-2), 179-85.
22. Kaida, K.; Ariga, T.; Yu, R. K., Glycobiology 2009, 19 (7), 676-92.
23. Yeh, W. Z.; Dyck, P. J.; van den Berg, L. H.; Kiernan, M. C.; Taylor, B. V., J Neurol Neurosurg Psychiatry 2020, 91 (2), 140-148.
24. Gaudet, R. G.; Gray-Owen, S. D., PLoS Pathog 2016, 12 (9), e1005807.
25. Yoshida, F.; Yoshinaka, H.; Tanaka, H.; Hanashima, S.; Yamaguchi, Y.; Ishihara, M.; Saburomaru, M.; Kato, Y.; Saito, R.; Ando, H.; Kiso, M.; Imamura, A.; Ishida, H., Chemistry – A European Journal 2018, 25 (3), 796-805.
26. Blixt, O.; Vasiliu, D.; Allin, K.; Jacobsen, N.; Warnock, D.; Razi, N.; Paulson, J. C.; Bernatchez, S.; Gilbert, M.; Wakarchuk, W., Carbohydr Res 2005, 340 (12), 1963-72.
27. Boons, G.; Overhand, M.; van der Marel, G.; van Boom, J., Angew. Chem. Int. Ed. Engl. 1989, 28, 1504-1506.
28. Boons, G. J. P. H.; Steyger, R.; Overhand, M.; van der Marel, G. A.; van Boom, J. H., Journal of Carbohydrate Chemistry 1991, 10 (6), 995-1007.
29. Rio, S.; Beau, J.-M.; Jacquinet, J.-C., Carb. Res. 1991, 219, 71-90.
30. Yu, H.; Li, Y.; Zeng, J.; Thon, V.; Nguyen, D. M.; Ly, T.; Kuang, H. Y.; Ngo, A.; Chen, X., J Org Chem 2016, 81 (22), 10809-10824.
31. Yu, H.; Santra, A.; Li, Y.; McArthur, J. B.; Ghosh, T.; Yang, X.; Wang, P. G.; Chen, X., Org Biomol Chem 2018, 16 (22), 4076-4080.
32. Bernatchez, S.; Gilbert, M.; Blanchard, M. C.; Karwaski, M. F.; Li, J.; Defrees, S.; Wakarchuk, W. W., Glycobiology 2007, 17 (12), 1333-43.

Chapter 6

Summary

Samenvatting



Chapter 1

The chemical structures and biological roles of glycosphingolipids (GSLs) and sulfoglycolipids (SGLs) are introduced in this chapter. Chemical and chemoenzymatic synthetic strategies to obtain GSL oligosaccharides and glycolipids are discussed.

GSLs are composed of a glycan part, covering hundreds of different oligosaccharides, and a lipid part: ceramide. Due to their amphiphilic nature, they reside in the cell membrane, where the glycan moiety faces the extracellular milieu and the lipid is anchored in the membrane. GSLs can interact with both extracellular components (*trans* interactions) and components in the same cell membrane (*cis* interactions). The biosynthesis of GSLs starts with the formation of glucosylceramide (Glc-Cer) in the ER and early Golgi apparatus. Next, Glc-Cer is extended to Lac-Cer: the common precursor for both the globo- and ganglio-series of GSLs, which cover the structures discussed in this thesis. Various glycosyltransferases are involved in the biosynthesis of globo- and ganglio-series glycolipids. GSLs are involved in various processes in the cell, such as cell adhesion and transmembrane signaling. Furthermore, GSLs are involved in various diseases, which are mainly neurological (Tay Sachs, Sandhoff, Alzheimer's, Parkinson's, MS) or immune related (verotoxins, HIV). Other diseases involving GSLs are discussed more in depth: cancer and the auto-immune disease Guillain-Barré Syndrome (GBS). A common feature in carcinogenesis is the upregulation of sialylated structures. Sialylated GSLs are known as gangliosides and play important roles in different types of cancer. A lesser investigated ganglioside, DSGb5, has been found in renal cell carcinomas. DSGb5 was reported to bind Siglec-7: an immune inhibitory receptor mainly found on NK-cells. If cancer cells bind receptors such as Siglec-7, they are able to evade an immune response. This makes DSGb5 an interesting target for cancer detection or treatment. Cancer specific antibodies have been elicited towards various glycans. The anti-epiglycanin antibody HAE3, is known to bind mucin-type glycans and is selective binds epithelial cancer cells. Surprisingly, HAE3 strongly bound to the sulfoglycolipid SM1a in a glycan-microarray. Other SGLs, such as GM3 and SB1a, have been associated with liver cancer, however little is known of the exact role of SM1a. Another important disease involving gangliosides, is the auto-immune disease Guillain-Barré syndrome GBS. In GBS, the peripheral nervous system is damaged and antiganglioside antibodies seem to be a leading cause. The formation of antiganglioside antibodies seems to be triggered by the gram-negative bacterium *Campylobacter jejuni*. Ganglioside mimics can be found on lipooligosaccharides (LOS) in the outer membrane of *C. jejuni* and seem to trigger the formation of cross-reactive antiganglioside antibodies. In GBS serum samples, antibodies towards gangliosides such as GM1a have been detected. These gangliosides are mainly present in the peripheral nervous system, explaining certain symptoms of the disease. To study the biological roles and protein interactions of GSLs more closely *ex vivo*, pure homogeneous GSLs are required. Isolation of GSLs is a tedious job, especially for negatively charged compounds such as gangliosides and SGLs. Furthermore, isolation often results in heterogeneous mixtures. Therefore, chemical or enzymatic synthesis is the best option to obtain these GSLs. Different strategies have been designed for the synthesis of GSL glycolipids. Each approach has its advantages and limitations. The most resent approaches are the Glc-Cer cassette approach, to solve the problem of the low reactivity of a chemical glycosylation between an oligosaccharide donor and ceramide acceptor. Another promising approach is the chemoenzymatic strategy, where chemically synthesized lactosylceramide (Lac-

Cer) is extended by bacterial glycosyltransferases (GTs). Stereo- and regioselectivity are the key factors in carbohydrate synthesis. In chemical synthesis, regioselectivity is usually obtained by careful choosing orthogonal protecting groups. This way, the acceptor can be designed in such a way that only one free hydroxyl can react as the nucleophile. A stereoselective outcome is more challenging to obtain in chemical glycosylations. The general approach to form 1,2-*trans* linkages (β in the case of Glc or Gal) is by neighboring group participation of a C-2 ester. Forming a 1,2-*cis* linkage (α in the case of Glc and Gal) is more challenging, although some effective ways have been designed where the β face is blocked. One way is the use of a chiral auxiliary between C-1 and C-2, which forms a *trans*-decalin intermediate, or a di-*tert*-butane-silane group on the C-4 and C-6 of galactose which creates a steric bulk. Enzymatic synthesis provides both high stereo- and regioselectivity, although not all enzymes have been identified or expressed and sugar nucleotide donors are often expensive, thereby limiting the scale. Enzymes from both mammalian and bacterial sources have been used to synthesize oligosaccharides. Mammalian enzymes are usually more selective, but harder to express. Bacterial enzymes are easier to express, but also show less selectivity, which can be an advantage or disadvantage. Bacterial enzymes have been identified, expressed and used for the synthesis of both globo- and the ganglio-series GSLs.

Chapter 2

The globo-series ganglioside, disialosyl globopentaosylceramide (DSGb5), is often expressed in renal cell carcinoma. DSGb5 is proposed to bind Siglec-7; an immunoinhibitory receptor mainly found on NK-cells. Therefore, DSGb5 would aid cancer cells in escaping immune detection and immune-mediated clearance. We designed a chemoenzymatic strategy to synthesize DSGb5 and study its binding interaction with Siglec-7. The core pentasaccharide Gb5 was chemically synthesized by a 2 + 3 block approach. Four saccharide building blocks were protected to enable proper stereo- and regioselective outcomes. The α1,4-linkage between galactose and lactose was obtained by the use of a bulky di-*tert*-butyl-silane group on the galactose donor, solely providing α-anomeric product. Other chemical glycosylations provided β-linkages through neighboring group participation of C-2 esters. Chemical glycosylation of the Galβ1,3GalNAc disaccharide donor with the Galα1,4Lac (Gb3) trisaccharide acceptor provided anomeric mixtures when the donor contained a benzylidene acetal. Replacement of the benzylidene acetal by C-4 and C-6 acetyls provided pure β-anomeric selectivity. Gb5 was globally deprotected and mammalian sialyltransferases ST3Gal1 and ST6GalNAc5 installed the terminal α2,3- and internal α2,6-sialic acids. Glycan microarray binding studies indicated proper printing of Gb5, MSGb5 and DSGb5, by the plant lectins Soybean Agglutinin (SBA: binds to Gal or GalNAc residues) and Wheat Germ Agglutinin (WGA: binds to GlcNAc or sialic acid residues). It was found that Siglec-7 does not recognize DSGb5, and preferentially binds Neu5Acα(2,8)Neu5Ac containing glycans.

Chapter 3

In the previous chapter the focus was on the synthesis of the globo-series oligosaccharides. Recent literature studies on the importance of the lipid moiety of GSLs on protein binding and cellular organization encouraged us to synthesize the full glycolipids of globo-series GSLs. Incorporation of a lipid moiety introduces new challenges, such as the introduction of new functional groups, the lower reactivity of lipids for chemical glycosylation and altered solubility properties. A first fully chemical synthetic strategy was designed and resulted in the fully protected form of globopentaosylsphingosine (Gb5-Sph). The removal of the benzyl ether protecting groups required a Birch reduction, to avoid reduction of the unsaturated sphingosine. However, removal of the benzyl ether protecting groups proved unsuccessful on a model glycolipid, which might be due to solubility issues. Next, a chemoenzymatic approach was chosen, starting from chemically synthesized globotriaosylsphingosine (Gb3-Sph). A bifunctional bacterial enzyme, LgtD from *Haemophilus Influenzae*, was used to extend the glycan chain of Gb3-Sph to obtain Gb4-Sph. Gb4-Sph was then converted to Gb5-Sph by the same enzyme. Although a lower conversion was seen on the sphingosine-linked glycans compared to the free reducing oligosaccharides, both substrates could be converted to their respective products. To obtain full conversion, multiple enzymatic reactions were required. Gb3-Sph, Gb4-Sph and Gb5-Sph were coupled with palmitic acid, under amide coupling conditions (HOBt/EDC) and provided globotriaosylceramide (Gb3-Cer), Gb4-Cer and Gb5-Cer. We propose that engineered bacterial enzymes or mammalian enzymes β3GalNAc-T1 (for Gb4) and β3GalT5 (for Gb5) will show higher toleration of the reducing lipid residues. This way, the lower enzymatic conversions of GSL substrates can be overcome.

Chapter 4

Sulfoglycolipids (SGLs) are in many ways similar to gangliosides, since they both contain a negatively charged glycan residue and a reducing-end ceramide. Also, SGLs are often sulfated on the same position as the sialic acid in a ganglioside, and a good example is found in the SGL SM1a and ganglioside GM1a. SGLs such as SM3 and SB1a have been associated with cancer, however little is known of SM1a. A microarray study from Ten Feizi *et al*. revealed that SM1a strongly binds to the cancer-specific anti-epiglycanin antibody HAE3. A surprise finding, since HAE3 is known to bind mucin-type *O*-glycans. *O*-glycan epitopes similar to SM1a were found on mucin glycoproteins, all having a core3 (GlcNAcβ1,3GalNAcα-Ser/Thr) glycan residue. Therefore, a hybrid glycan of SM1a and core3 was proposed, here named SM1-core3 (Galβ1,3GalNAcβ1,4Galβ1,3GlcNAcβ1,3GalNAcα). SM1-core3 is the synthetic target of this chapter and the protected pentasaccharide was successfully obtained through chemical synthesis. First, all monosaccharide building blocks were synthesized with the appropriate orthogonal protecting groups and C-2 esters to selectively form all β-anomeric linkages. The reducing *N*-acetylgalactosamine was equipped with an α-anomeric aminopentyl linker, to resemble the stereochemistry in *O*-glycans and to provide a handle for immobilization. Various $GalN_3$ donors were synthesized and tested to form the 1,2-*cis* α-linkage with an aminopentanol acceptor. The 3,4,6-tri-*O*-acetylated $GalN_3$ trichloroacetimidoyl donor provided the best outcome (~1/1 α/β anomeric mixture). After hydrolysis of the acetyls and installation of a benzylidene acetal on C-4 and C-6, the anomeric mixture could be purified by silica column chromatography. All

remaining glycosylations were performed in DCM in the presence of TMSOTf and 4 Å MS at -40 °C to -70 °C, providing pure β-anomeric selectivity and yields between 40% - 72%. After deprotection and sulfation of SM1a-core3, this sulfoglycolipid analog will be studied for its HAE3 binding properties in a glycan microarray.

Chapter 5

Gangliosides and mimics thereof, are not just associated with cancer, but also with auto-immune diseases. Guillian-Barré syndrome (GBS) is such an autoimmune disease, where the peripheral neurons are mainly affected, resulting in muscle weakness to various degrees. One third of GBS cases is triggered by *Campylobacter jejuni* infections. *C. jejuni* carries ganglioside mimics on its membrane glycan lipooligosaccharides (LOS) which are thought to elicit cross-reactive antibodies towards human endogenous gangliosides. Antibodies towards various the ganglio-series gangliosides, such as GM1a, GM1b and GD1a, have been detected in GBS serum. These gangliosides are mainly present in the peripheral nervous system, partly explaining the symptoms in this disease. In this chapter we described the synthesis of ganglioside mimics to study binding patterns of these serum antiganglioside antibodies. The smallest difference of bacterial ganglioside mimics and endogenous ganglioside is the monosaccharide L-*glycero*-D-manno-heptose (Hep). Heptose is solely found in gram-negative bacteria and possibly plays a role in triggering antiganglioside antibody formation. Therefore, we designed a chemoenzymatic approach to synthesize heptose-ganglioside mimics. Our synthetic approach started with the chemical glycosylation of a galactosyl donor with a heptose acceptor, to form the desired Galβ1,3Hep disaccharide. After global deprotection, the disaccharide was subsequently extended by bacterial enzymes PmST1 to form Hep-GM3, CgtA to form Hep-GM2 and CgtB to form Hep-GM1. It was observed that the enzyme CgtA quickly lost activity and the lysate should be directly used after expression. All structures were equipped with an aminopentyl linker for microarray immobilization. The synthesized ganglioside mimics and the ganglioside oligosaccharides will be studied for their serum antiganglioside antibody binding properties in a glycan microarray.

Hoofdstuk 1

De chemische structuur en biologische rol van glycosphingolipiden (GSLs) en sulfoglycolipiden (SGLs) worden geïntroduceerd in dit hoofdstuk. Verschillende strategieën voor de chemische en enzymatische synthese van GSL oligosachariden en glycolipiden worden bediscussieerd. GSLs zijn opgebouwd uit een sacharide deel, met honderden mogelijke structuren, en een lipide deel: ceramide. Vanwege de amfifiliciteit verblijven GSLs voornamelijk in het celmembraan, met de sacharide in het extracellulaire milieu en de ceramide verankerd in het membraan. Interacties van GSLs kunnen plaatsvinden met componenten buiten de cel (*trans* interacties) en componenten in hetzelfde celmembraan (*cis* interacties). De biosynthese van GSLs begint in het ER en Golgi en start met de vorming van glucosylceramide (Glc-Cer). Enzymen verlengen Glc-Cer naar lactosylceramide (Lac-Cer): de uitgangsstof voor zowel de globo- als ganglio-serie GSLs, welke de structuren van deze thesis bevatten. Verschillende glycosyltransferases vormen de globo- en ganglio-serie glycolipiden. The resulterende GSLs zijn betrokken bij cel processen, zoals cel adhesie en signaal transductie. Daarnaast zijn GSLs betrokken bij verschillende ziektes, voornamelijk neurologische ziektes (Tay Sachs, Sandhoff, Alzheimer, Parkinson, MS) en immuun gerelateerde ziektes of infecties (verotoxines, HIV). Twee ziektes waar GSLs een belangrijke rol in spelen worden in meer detail besproken: kanker en de auto-immuun ziekte Guillain-Barré syndroom (GBS). Een veelvoorkomend fenomeen in carcinogenese is een toename in structuren met siaalzuren. Gangliosiden zijn gesialileerde GSLs en vormen daarmee relevante structuren voor het detecteren en mogelijk behandelen van kanker. DSGb5 is een minder bestudeerd ganglioside, en is gedetecteerd in nierkanker. Daarnaast wordt er gesuggereerd dat DSGb5 bindt aan Siglec-7, een immuun remmend eiwit dat voornamelijk te vinden is op NK cellen. Wanneer kankercellen binden aan eiwitten zoals Siglec-7, kunnen ze een immuunreactie ontwijken. Verder zijn er ook antilichamen gevonden voor kanker-specifieke oligosachariden. Een van deze antilichamen is HAE3, welke bindt aan epitheelkanker cellen. SM1a werd gedetecteerd als een sterke bindend ligand voor HAE3. Andere SGLs, zoals GM3 en SB1a, zijn gevonden in leverkanker, maar er is nog maar weinig bekend van de rol van SM1a. In de auto-immuunziekte GBS wordt het perifere zenuwstelsel aangetast. Het lijkt erop dat anti-ganglioside antilichamen een mogelijke oorzaak hiervan zijn. Deze antilichamen kunnen gevormd worden na infectie met de gram-negatieve bacterie *Campylobacter jejuni*. Ganglioside-achtige structuren kunnen aanwezig zijn in de lipooligosachariden in het buitenste membraan van *C. jejuni* en lijken de vorming van anti-ganglioside antilichamen te veroorzaken. Antilichamen voor ganliosides zoals GM1a zijn gevonden in serum monsters van GBS patienten. Over het algemeen richten de antilichamen zich op gangliosiden die zich voornamelijk in het perifere zenuwstelsel bevinden. Om de biologisch rol en eiwit interacties van GSLs nader te bestuderen buiten de cel, zijn zuivere en homogene GSLs nodig. Isolatie van GSLs uit biologische monsters is lastig, zeker wanneer deze negatief geladen groepen bevatten zoals in gangliosiden en SGLs. Daarnaast resulteert isolatie van GSLs vaak in heterogene mengsels. Vanwege de lastige isolatie van GSLs bieden chemische en enzymatische synthese strategieën een uitkomst. Er bestaan een aantal strategieën voor de synthese van de GSL glycolipiden, met ieder hun eigen voor- en nadelen. Eén van de methodes is de Glc-Cer cassette benadering, waardoor de lage reactiviteit van een oligosacharide donor en ceramide acceptor in een chemische glycosylering verholpen

wordt. Een andere veelbelovende methode is een chemo-enzymatische strategie waar chemisch gesynthetiseerd lactosylceramide (Lac-Cer) verder wordt uitgebouwd door bacteriële glycosyltransferases. Stereo- en regioselectiviteit zijn belangrijke factoren in de suikerchemie. Regioselectiviteit wordt over het algemeen verkregen door het gebruik van orthogonale beschermgroepen. Op deze manier kan er één hydroxylgroep beschikbaar zijn als nucleofiel in de acceptor. Stereoselectiviteit sturen is uitdagender in chemische glycosyleringen. Over het algemeen worden 1,2-*trans* verbindingen (β in het geval van Glc of Gal) gecontroleerd gesynthetiseerd door naburige groep participatie door een C-2 ester. De vorming van 1,2-*cis* verbindingen (α in het geval van Glc of Gal) is moeilijker, alhoewel er een aantal manieren zijn gevonden om de β-kant te blokkeren. Eén van deze benaderingen gaat uit van een chirale groep, gelinkt tussen C-1 en C-2 door een *trans*-decalin systeem. Een andere benadering maakt gebruik van sterische bulk van een di-*tert*-butyl-silaan groep op de C-4 en C-6 hydroxylgroepen van galactose. Naast chemische synthese kunnen oligosachariden ook door enzymen gesynthetiseerd worden. Over het algemeen zijn enzymen stereo- en regioselectief, maar kunnen het expressievolume en de activiteit van het enzym en de dure suiker nucleotiden zorgen voor een kleinere schaal en lagere opbrengsten. Enzymen uit zoogdieren zijn vaak selectiever, maar moeilijk tot expressie te brengen. Bacteriële enzymen zijn makkelijker tot expressie te brengen, maar zijn ook minder selectief, wat zowel een voor- als nadelig gevolg kan hebben. Verschillende bacteriële enzymen zijn gevonden en tot expressie gebracht, die gebruikt kunnen worden voor de synthese van globo- en ganglio-serie van GSLs.

Hoofdstuk 2

Een ganglioside uit de globo-serie, disialosyl globopentaosylceramide (DSGb5), wordt vaak gedetecteerd in nierkankercellen. Er is eens sterk vermoeden dat DSGb5 Siglec-7 bindt: een immuunonderdrukkend eiwit, vooral gevonden op NK-cellen. Op deze manier kan DSGb5 ervoor zorgen dat kanker cellen ontkomen aan detectie en opruiming door immuuncellen. Wij hebben een chemo-enzymatische strategie ontworpen om DSGb5 te synthetiseren, zodat we de bindingseigenschappen met Siglec-7 nader kunnen bestuderen. Het pentasacharide Gb5 vormt de basis van DSGb5 en werd gesynthetiseerd door de chemische glycosylering van een di- en trisacharide. Allereerst werden vier bouwstenen gesynthetiseerd met de benodigde beschermgroepen om de juiste stereo- en regioselectiviteit te kunnen faciliteren. De α-configuratie in Galα1,4Lactose werd gevormd door een galactose donor met een di-*tert*-butyl-silane groep, welke zorgt voor een sterische bulk en daarmee vorming van het β-anomeer verhinderd. Alle overige glycosyleringen werden gefaciliteerd door C-2 esters, om zo β-configuratie te verkrijgen. De chemische glycosylering van het Galβ1,3GalNAc disacharide met het Galα1,4Lac trisacharide resulteerde in een mengsel van anomeren, wanneer het disacharide een benzylideen acetaal bevatte. Toen het benzylideen acetaal vervangen werd door C-4 en C-6 acetyl groepen, werd alleen het gewenste β-product gevormd. Na volledige ontscherming van Gb5 werden twee enzymen, ST3Gal1 en ST6GalNAc6, gebruikt om zowel het terminale α2,3- en de interne α2,6-siaalzuren te koppelen. De plantlectines SBA (herkent Gal of GalNAc) en WGA (herkent GlcNAc of siaalzuren) toonden aan dat Gb5, MSGb5 en DSGb5 correct waren geprint op de micro-array slide. Verder werd er geconstateerd dat Siglec-7 niet aan DSGb5 bindt, maar wel aan structuren die α2,3-α2,8-gelinkte siaalzuren bevatten.

Hoofdstuk 3

In het vorige hoofdstuk lag de focus op de synthese van de oligosachariden van de globo-serie GSLs. Nu willen we de volledige glycolipiden uit de globo-serie synthethiseren, omdat er gesuggereerd wordt dat het lipide een belangrijke rol speelt in het binden van eiwitten en de cellulaire organisatie. Het includeren van een lipide zorgt voor nieuwe uitdagingen: er worden nieuwe functionele groepen geïntroduceerd, de gebruikte lipiden hebben een lage reactiviteit in chemische glycosyleringen en de oplosbaarheid van de stoffen zal veranderen. Een volledig chemische synthetische strategie resulteerde in de beschermde vorm van Gb5-sphingosine (Gb5-Sph). Het verwijderen van de benzyl ethers vereist een Birch reductie, om te voorkomen dat ook het onverzadigde sphingosine wordt gereduceerd. De Birch reductie werkte niet op het modelsubstraat Gb3-Sph, wat mogelijk komt door problemen met de oplosbaarheid. Vervolgens kozen we een chemo-enzymatische strategie, beginnende met de chemische synthese van Gb3-Sph. Een bacterieel enzym met dubbele functie, LgtD uit *Haemophilus Influenzae,* was gekozen voor de synthese van Gb4-Sph en Gb5-Sph. Hoewel er minder omzetting werd gezien op de glycolipide substraten vergeleken met de respectievelijke oligosacharides, konden zowel Gb4-Sph als Gb5-Sph worden verkregen. Om een volledige omzetting van de substraten naar producten te krijgen, moesten er meerdere reactiecycli plaatsvinden. De substraten Gb3-Sph, Gb4-Sph en Gb5-Sph werden vervolgens gekoppeld met palmitinezuur door condities gebruikt voor amide koppeling (EDC, HOBt) en resulteerde in de Gb3-Cer, Gb4-Cer en Gb4-Cer. Voor toekomstige optimalisatie stellen wij het gebruik van aangepaste bacteriële enzymen voor. Een andere optie is het gebruik van enzymen verkregen uit zoogdieren, zoals β3GalNAc-T1 (voor Gb4) en β3GalT5 (voor Gb5). Deze enzymen zullen mogelijk meer tolerantie voor aanwezige lipiden tonen en zorgen voor hogere omzettingen van substraat in product.

Hoofdstuk 4

Sulfoglycolipiden (SGLs) hebben veel overeenkomsten met gangliosiden, omdat ze beide negatief geladen oligosachariden en een ceramide bevatten. Daarnaast is vaak dezelfde positie in SGLs gesulfateerd waar normaal de siaalzuren in gangliosiden zich bevinden; een goed voorbeeld hiervan zijn SGL SM1a en ganglioside GM1a. SGLs zoals SM3 en SB1a zijn geassocieerd met verschillende vormen van kanker, maar er is weinig bekend over SM1a. SM1a was gevonden als een sterke binder voor het anti-epiglycanine antilichaam HAE3 in een microarray. Het was verrassend om een SGL te vinden als ligand voor HAE3, aangezien HAE3 normaalgesproken *O*-glycanen van mucine bindt. Vergelijkbare structuren als SM1a waren gevonden in *O*-glycanen van mucine en alle structuren bevatten het core3 disacharide (GlcNAcβ1,3GalNAcα-Ser/Thr). Uit deze bevindingen werd een hybride structuur van SM1a en core3 voorgesteld (Galβ1,3GalNAcβ1,4Galβ1,3GlcNAcβ1,3GalNAcα). Dit oligosacharide wordt in dit hoofdstuk SM1-core3 genoemd en vormt tevens het synthetische doel. De volledig beschermde vorm van SM1-core3 is verkregen door chemische synthese. Om te beginnen werden alle monosachariden zo beschermd dat ze de juiste stereo- en regioselectiviteit kunnen vormen. Voor alle β-gelinkte sachariden werd er gebruik gemaakt van C-2 esters om de gewenste stereoselectiviteit te faciliteren. Er bevind zich één α-configuratie in het molecuul, namelijk tussen galactosamine en de aminopentyl linker. De α-configuratie

was gekozen om zo dicht mogelijk bij het natuurlijk gelinkte *O*-glycaan te blijven en de linker was gekozen om het oligosaccharide te kunnen immobiliseren in een latere fase. Verschillende $GalN_3$ donoren werden getest om de gewenste 1,2-*cis* (α)-configuratie te vormen. Het beste resultaat werd verkregen door een 3,4,6-tri-*O*-geacetyleerd trichloroacetimidaat donor (~1/1 α/β mengsel). Na verwijdering van de acetylgroepen en de installatie van een benzylideen acetaal op C-4 en C-6, konden de α- en β-anomeren gescheiden worden door silica kolom chromatografie. Alle overige chemische glycosyleringen vonden plaats in DCM in de aanwezigheid van TMSOTf en 4 Å MS bij een temperatuur van 40 °C to -70 °C. Deze condities resulteerden in alle gevallen in het pure β-anomeer en opbrengsten varieerden van 40% tot 72%. Na de ontschermingen en sulfatering van SM1a-core3 zal dit sulfoglycolipide analoog worden bestudeerd voor zijn bindingseigenschappen voor het HAE3 antilichaam in een microarray.

Hoofdstuk 5

Gangliosiden en ganglioside-achtige structuren zijn niet alleen veel geassocieerd met kanker, maar ook met auto-immuun ziekten. Guillian-Barré syndroom (GBS) is zo'n autoimmuun ziekte, waar de perifere zenuwen worden aangetast, wat kan resulteren in verzwakte spieren of zelfs verlamming. Een derde van de gevallen van GBS wordt voorafgegaan door een infectie met de bacterie *Campylobacter jejuni. C. jejuni* is een gramnegatieve bacterie en heeft lipooligosacchariden (LOS) in het buitenste membraan. Dit LOS kan ganglioside-achtige structuren bevatten en zorgen mogelijk voor antilichamen die ook gangliosiden in het menselijk lichaam kunnen aantasten. Verschillende anti-ganglioside antilichamen zijn gevonden in serum van GBS patiënten, waaronder GM1a, GM1b en GD1a. Deze ganglio-serie gangliosiden kunnen hoofdzakelijk gevonden worden in het perifere zenuwstelsel, wat de symptomen van de ziekte zou kunnen verklaren. In dit hoofdstuk beschrijven wij de synthese van ganglioside-achtige structuren om op zoek te gaan naar een bindingspatroon van anti-ganglioside antilichamen in GBS serum. Het kleinste verschil tussen de ganglioside-achtige structuur van *C. jejuni* en de menselijke gangliosiden is de monosacharide L-*glycero*-D-manno-heptose (Hep). Heptose wordt alleen gevonden in gramnegatieve bacteriën en kan mogelijk een rol spelen in de vorming van de anti-ganglioside antilichamen. Om deze reden hebben wij een chemo-enzymatische strategie bedacht om heptose-gangliosiden te synthetiseren. Allereerst werd het Galβ1,3Hep disacharide verkregen door chemische glycosylering van een galactose donor en heptose acceptor. Vervolgens is dit disacharide ontdaan van zijn beschermgroepen en werden bacteriële enzymen gebruikt om de ganglioside-achtige structuren te maken. PmST1 werd gebruikt om Hep-GM3 te maken, CgtA voor Hep-GM2 en CgtB voor Hep-GM1. Het enzym CgtA was opvallend onstabiel en het cellysaat moest direct na expressie gebruikt worden om voldoende activiteit te hebben voor de omzetten van Hep-GM3 naar Hep-GM2. Alle gesynthetiseerde structuren bevatten een aminopentyl linker voor immobilisatie op een microarray. Naast de gesynthetiseerde structuren zullen ook de ganglioside oligosachariden geprint worden, om zo een patroon te kunnen ontdekken in de binding van anti-ganglioside antilichamen in GBS serum.

Appendices

Curriculum Vitae

List of Publications

Acknowledgements / Dankwoord

Curriculum Vitae

Ingrid 't Hart was born on the 10th of April 1992 in Krimpen a/d IJssel, The Netherlands. She graduated from the Comenius College in Capelle a/d IJssel in 2010 and started the bachelor Bio-Pharmaceutical Sciences (BPS) in Leiden the same year. In 2013, she completed her bachelor and continued with the BPS master in Leiden, specializing in organic synthesis. Her first masters internship was in the department of Medicinal Chemistry at the Leiden Academic Centre for Drug Research (LACDR) in Leiden under supervision of Jaco van Veldhoven. In 2014, she successfully completed her internship and contributed as co-author on two publications. In 2015 she moved to Dortmund, Germany to start her second internship at the Max Planck Institute of Molecular Physiology under supervision of Pablo Martin-Gago. She completed her master thesis "Covalent inhibition of PDEδ by sulfonyl fluoride derivatives: a possible treatment for K-Ras dependent carcinomas" and obtained her Master of Science degree *cum laude*. Throughout her studies she has actively participated in committees from the study association L.P.S.V. „Aesculapius". In 2015, she started her PhD research in the group of Chemical Biology & Drug Discovery (CBDD) at the Utrecht Institute for Pharmaceutical Sciences (UIPS) in Utrecht. Under the supervision of Geert-Jan Boons, she worked on four projects all focused on the chemical or chemoenzymatic synthesis of glycosphingolipids and the results are described in this thesis. In May 2020 she started her job as a scientific advisor at the National Institute for Public Health and the Environment (RIVM) in Bilthoven.

List of publications

Chemoenzymatic synthesis of the oligosaccharide moiety of the tumor-associated antigen disialosyl globopentaosylceramide. Ingrid M. E. ’t Hart, Tiehai Li, Margreet A. Wolfert, Shuo Wang, Kelley W. Moremen and Geert-Jan Boons. *Org. Biomol. Chem.* **2019**, 17: 7304-7308

Other publications

Structure−Affinity Relationships (SARs) and Structure−Kinetics Relationships (SKRs) of Kv11.1 Blockers, Zhiyi Yu, Jacobus P. D. van Veldhoven, Julien Louvel, Ingrid M. E. ’t Hart, Martin B. Rook, Marcel A. G. van der Heyden, Laura H. Heitman, and Adriaan P. IJzerman. *J. Med. Chem.,* **2015**, 58: 5916-5929

Synthesis and biological evaluation of negative allosteric modulators of the Kv11.1 (hERG) channel. Zhiyi Yu, Jacobus P.D. van Veldhoven, Ingrid M.E. ˈt Hart, Adrian H. Kopf, Laura H. Heitman, Adriaan P. IJzerman. *Eur. J. Med. Chem*., **2015**, 106: 50-59

Poster presentations

Eurocarb, 1-4 July 2019, Leiden

CHAINS 3-4 Dec 2018, Veldhoven

Eurocarb 2-6 July 2017, Barcelona

Bijvoet Symposium 25-26 April 2016, Soesterberg

Acknowledgements / Dankwoord

First of all, I would like to thank **Geert-Jan** for his guidance and supervision during my PhD research. Before I started in Utrecht, I had no background in sugar chemistry, but in the four years of my PhD I have learned a lot, both practically and theoretically. I would also like to thank **Margreet**; especially for the microarray studies. **Roland**, thanks for all the helpful questions and suggestions during group meetings and the nice collaboration during teaching Organic Chemistry 3. Thanks to **Tom** (**Wennekes**) for all work-related discussions and fun times. Special thanks for all your help in the job hunt, this proved to be very fruitful!

I would like to thank **Johan**, **Hans** and **Justyna** for their help with NMR measurements and **Javier** and **John** for their help with MS measurements on various machines and for providing all HRMS data. I want to thank **Gerlof** for the expression of various enzymes; some of the enzymes proved to be quite tricky! I want to thank **Dowson** for the precious heptose building block that he synthesized for me.

Furthermore, I would like to thank **Nathaniel** and his **group** for all the nice times here in Utrecht. Our Nemo fish has reminded me of your group every day. Dear **Charlotte, Tom, Emma** and **Nicola**, we had a lot of fun organizing all sorts of activities, such as the lab retreat, borrels and lasertag matches. I am glad we are still staying in touch and have our regular dinner in Utrecht or Leiden, King's day celebrations or the occasional 90's or 00's party. All the best in the beautiful city of Leiden!

I would like to thank my labmates **Apoorva** and **Ivan**, who helped me a lot during my first year in the lab in Utrecht. I am impressed by the way you all learned to speak Dutch in such a short time, from such different languages! Thanks to **Mehman** for the collaboration of the sulfoglycolipid project and discussions at our fumehoods. Also thanks to (captain) **Jack**, always ready for questions and the occasional joke in the lab. **Enrico** and **Arwin**, you have been my next-door labmates, always ready for a question on a chemical reaction or a (really) bad joke. I also want to thank my student, **Caitlin**: you were a highly motivated and bright student and made the supervision very easy for me!

Thanks to **Hanna, Pieter, Tim, Nino**, **Frederik, Cindy**, **Helena, Seino** for the all the support, discussions and lunch breaks. We have discussed all sorts of things, from chemicals and toxicity to moving (and renovation a house), movies (such as "It"), recipes and the Dutch and German language.

Rosanne, I am happy that you want to be my paranimf. Already since you started your internship we could get along very well and a had lot of fun moments. I am happy that you came back to the department for your PhD and I wish you all the best in your research projects! I am also glad that we could discuss anything through the years and I definitely hope that we will stay in touch! (nog ééntje dan?)

Yvette (YMCA), I am happy that you want to be my paranimf! From the start of both our PhDs we could get along very well. And even though we are working on quite different topics (biology vs chemistry!), we had some very nice in-depth discussions. You were

always ready for a chat or a joke or a game of Mario Kart on the switch with a snack from the vegetable drawer. Even though you moved to your villa in Overloon, I hope we will stay in touch. Good luck with the last part of your PhD and the job hunt!

I would also like to thank anyone of **CBDD** or **CCRC** (Athens, Georgia) that I haven't mentioned by name and wish you all the best!

Lieve **Cindy, Richard, Marianne, Deborah, Lisanne en Bastiaan**. Wij kennen elkaar al van de 4e of 5e klas van de middelbare school en ik vind het enorm leuk dat we nog steeds zo goed contact hebben. Ook al begrijpen we elkaars werk niet altijd, toch is het fijn om af en toe iets te kunnen delen hierover. Dank dat jullie er altijd voor mij geweest zijn, ook als het contact even niet van mijn kant kwam. Ik hoop dat we nog een hoop mooie en leuke dingen mogen delen samen.

Lieve **pa** en **ma**, **Martin**, **Peter** en **Judith**. Jullie hebben me altijd gesteund en in mij geloofd. Ik ben blij dat ik opgegroeid ben in zo'n liefdevolle familie en alle mogelijkheden heb gekregen om me te ontwikkelen, zowel persoonlijk als in opleidingen. Ik ben trots op jullie allemaal en zijn ook allemaal een voorbeeld voor me geweest (en zijn dat nog steeds)!

Lieve **Mark**, wat ben ik blij dat ik jou ken en met jou mocht trouwen tijdens mijn PhD. Je bent er altijd voor mij geweest en daar ben ik heel dankbaar voor. Ik hoop er zo ook te kunnen zijn voor jou in de laatste fase van je eigen PhD. Nog even en we mogen ons Dr. & Dr. Bakker noemen!