CHEMOENZYMATIC SYNTHESIS AND IMMUNOLOGICAL STUDIES OF XYLOSYLATED N-GLYCANS

Katarzyna Brzezicka

Glycotechnology Laboratory, CIC biomaGUNE

2016

This thesis has been carried out in the Glycotechnology Laboratory at The Centre for Cooperative Research in Biomaterials, CIC biomaGUNE, San Sebastián, País Vasco, Spain.

The research was financially supported by a fellowship obtained from the Spanish Ministry of Science and Innovation (BES-2012-052835) associated to the Microarray-Based Glycomics project (CTQ-2011-27874).

Resumen Tesis

Los carbohidratos junto a los ácidos nucleícos, las proteínas y los lípidos, son las macromoléculas fundamentales que constituyen todas las células. Antiguamente los carbohidratos eran considerados únicamente como moléculas proveedoras de energía (glucosa y glicógeno) o como elementos estructurales (celulosa, quitina). Hoy en día, se conoce que participan en numerosos procesos celulares y que codifican una gran cantidad de información, debido a su gran diversidad estructural. La superficie de las células está cubierta de una densa capa de carbohidratos que se denomina glicocálix. En esta capa los carbohidratos se encuentran formando glicoconjugados, como glicoproteínas, glicolípidos o proteoglicanos.

La expresión de macromoléculas glicosiladas es específica de cada tipo de célula o tejido y puede verse alterada por su estado de desarrollo o diferenciación. Los perfiles de glicosilación de las células también varían en diversas patologías. Por ello, patrones anormales de glicosilación pueden ser empleados como biomarcadores de enfermedades tan importantes como el cáncer o enfermedades autoinmunes. Adicionalmente, los carbohidratos juegan un papel muy importante en la activación de la respuesta inmune ya que los receptores de sistema inmunitario pueden ser activados por la presencia de carbohidratos exógenos en la superficie de patógenos. Por otro lado, muchos patógenos emplean la adhesión a los carbohidratos del huésped como el primer paso hacia la infección o incluso, expresan carbohidratos que mimetizan a los del huésped como estrategia para evadir al sistema inmunitario.

La glicosilación es una de las modificaciones postraduccionales más abundantes en las células eucariotas, se estima que alrededor del cincuenta por ciento de las proteínas se encuentran modificadas por carbohidratos en su superficie, cuya presencia modula de manera importante las propiedades como la solubilidad y la vida media. Los carbohidratos se unen a distintos aminoácidos mediante enlaces N- y O-glicosídicos. En concreto, las estructuras de

tipo N-glicano objeto de estudio de la presente Tesis doctoral se unen de manera covalente al nitrógeno de la amida de ciertos residuos de asparagina (Asn).

La composición del glicoma, repertorio de carbohidratos presentes en cada organismo, varía de manera muy significativa entre las diferentes especies. En el caso de los N-glicanos, a pesar de compartir una estructura de pentasacárido central, el número y la composición de las antenas así como la sustitución en la quitobiosa central varía de manera importante entre los distintos organismos. Algunas de estas estructuras han sido identificadas como inmunogénicas y se encuentran en plantas, invertebrados y en parásitos helmintos como *Schistosoma mansoni*. En la presente Tesis doctoral se pretende abordar la síntesis de N-glicanos que presentan epítopos supuestamente inmunogénicos y evaluar su posible función biológica, mediante el empleo de microarrays de carbohidratos y ensayos *in vitro* con células dendríticas.

En el primer capítulo de la Tesis se recoge la preparación de manera quimio-enzimática de 39 estructuras de tipo N-glicano que presentan todas ellas el epítopo β-1,2-xilosa. Para ello se aborda la síntesis orgánica de seis estructuras que posteriormente se modificarán empleando glicosiltransferasas recombinantes para instalar diferentes sustituyentes como pueden ser fucosas en la quitobiosa central (core α-1,6/core α-1,3) y diferentes elementos en las antenas como LDN (GalNAcβ-1,4-GlcNAc), Lewis X (Galβ-1,4(Fucα-1,3)GlcNAc)o LDNF (GalNAcβ-1,4(Fucα-1,3)GlcNAc).

El segundo capítulo de la Tesis se centra en la preparación de microarrays de carbohidratos para el estudio de interacciones carbohidrato-proteína. En las últimas décadas, los microarrays de carbohidratos han surgido como una herramienta altamente eficaz para el estudio de interacciones carbohidrato-proteína, así como para el estudio de enzimas que modifican carbohidratos. Las ventajas de esta metodología se centran en la evaluación simultánea de múltiples posibles interacciones (High-Throughput Screening) empleando cantidades mínimas de analitos. Las estructuras de tipo N-glicano previamente sintetizadas se inmovilizan de manera covalente sobre portas de vidrio y se incuban con disoluciones de las proteínas o anticuerpos a estudiar. La presencia de interacciones positivas se estudiará

empleando fluorescencia bien mediante el marcaje directo de la proteína o mediante la detección con un anticuerpo secundario marcado fluorescentemente.

El conjunto de proteínas que reconocen de manera específica a los carbohidratos sin modificarlos se denominan lectinas. Muchas de estas proteínas se aíslan de diferentes plantas y hongos, donde juegan un papel en la defensa frente a insectos y herbívoros, a través del reconocimiento de estructuras no presentes en los organismos que las producen. La especificidad hacia los carbohidratos de muchas de estas lectinas es conocida y hoy en día se han convertido en herramientas muy útiles en glicobiología. En la presente Tesis doctoral, se ha empleado un pequeño panel de lectinas para demostrar la funcionalidad de los microarrays de N-glicanos preparados. A pesar de que su especificidad se encuentra ampliamente documentada, se observó una importante influencia por la presencia de β-1,2-xilosa en los N-glicanos, que reduce el reconocimiento de *Concanavalina A* y completamente impide el reconocimiento de la aglutinina de *Galanthus nivalis*. Estos resultados demuestran el gran potencial de los microarrays de carbohidratos en el estudio y caracterización de interacciones carbohidrato-proteína.

Las lectinas de tipo C (C-type lectins) son una importante familia de receptores que se expresan en la superficie de las células dendríticas (CD) capaces de reconocer de manera específica a los carbohidratos que expresan diferentes patógenos en su superficie. Estas proteínas colaboran en la internalización de los antígenos por parte de las células dendríticas, que posteriormente los procesan y presentan los fragmentos como péptidos antigénicos a los linfocitos T. Una selección de lectinas de tipo C recombinantes de origen murino y humano se incubaron en el microarray de N-glicanos pudiendo observar el patrón de reconocimento de tres de ellas. En el caso particular de DC-SIGN, la presencia de la β-1,2-xilosa en los N-glicanos impide por completo el reconocimiento.

La esquistosomiasis es una enfermedad aguda y crónica causada por gusanos parásitos del género *Schistosoma* que según la Organización Mundial de la Salud, afecta a más de doscientos millones de personas. La principal fuente de contagio es el contacto con aguas

contaminadas, donde las personas se infectan cuando las larvas del parásito, liberadas por caracoles de agua dulce, penetran en la piel intacta. En el interior del organismo, las larvas se convierten en gusanos adultos, que viven en los vasos sanguíneos, donde las hembras depositan sus huevos. Algunos de esos huevos salen del organismo con las heces o la orina y continúan el ciclo vital del parásito infectado a caracoles de agua dulce. Otros quedan atrapados en los tejidos corporales, donde activan la respuesta inmune del huésped y producen un daño progresivo de los órganos. Durante su ciclo de vida, el parásito expresa un gran número de carbohidratos específicos en las distintas fases de su desarrollo y la estructura de muchos de ellos difiere marcadamente de los humanos. Está descrito que durante la infección se generan un gran número de anticuerpos anti-carbohidratos pero su papel en la protección del huésped o en el escape del parásito del sistema inmune no se conoce en detalle. En el tercer capítulo de la presente Tesis doctoral se emplearon microarrays de N-glicanos para caracterizar la especificidad de anticuerpos monoclonales generados a partir del suero de ratones infectados con *S. mansoni*. Posteriormente, se estudió la presencia de inmunoglobulinas G (IgG) y M (IgM) en el suero de pacientes infectados por *S. mansoni* que viven en una comunidad endémica en Butiaba, Uganda. Se observan marcadas diferencias en la respuesta de IgGs en el suero de pacientes de dos grupos de edad diferentes. La mayor respuesta en el suero de niños infectados, se observa hacia estructuras que contienen tanto β-1,2-xilosa como α-1,3-fucosa en las estructuras de N-glicano. Sin embargo, la respuesta de los adultos se centra sólo frente a N-glicanos sustituidos con β-1,2-xilosa, no reconociendo las estructuras que contienen α-1,3-fucosa. Parece indicar que el sistema inmune de los niños es capaz de reaccionar frente a un número mayor de epítopos inmunogénicos mientras que los adultos no muestran tal reacción probablemente por la exposición previa a dichos epítopos en infecciones pasadas.

El capítulo 4 de la Tesis Doctoral, se estudió la dirección de antígenos a células dendríticas murinas mediada por carbohidratos empleando glicoconjugados sintéticos. Basados en los resultados de unión del receptor DC-SIGN en el microarray de N-glicanos, donde se observaba que una pequeña modificación estructural como la presencia de β-1,2-xilosa en

los N-glicanos impedía la interacción con el receptor, se pretende estudiar la influencia de esta modificación en ensayos *in vitro*. Para ello se han sintetizado glicoconjugados de N-glicanos con la proteína antigénica modelo ovoalbúmina (OVA). El estudio de la diferente interacción de estos glicoconjugados con lectinas de tipo C se llevó a cabo mediante microarrays de proteínas, observándose una importante diferencia en el reconocimiento de dichos glicoconjugados por parte de la lectina de origen murino SignR3, proteína homologa al DC-SIGN humano. Adicionalmente, se han llevado a cabo ensayos de captación directa de los glicoconjugados en células dendríticas de origen murino, observándose una internalización más efectiva cuando el N-glicano no contiene el epítopo β-1,2-xilosa. Tras la internalización en las CDs, los antígenos son procesados y los distintos fragmentos peptídicos se presentan en el complejo principal de histocompatibilidad (MHC-II) a las células T, causando su activación y posterior diferenciación en células efectoras. Para cuantificar la distinta activación de las células T por parte de los glicoconjugados, se llevó a cabo un ensayo de co-cultivo de CDs y células T. Se cuantificó la expresión del marcador CD69 por citometría de flujo y la expresión de interleuquina 2 (IL-2) e interferon gamma (IFNγ) mediante ELISA (Enzyme-Linked ImmunoSorbent Assay), mostrando una mayor activación de las células T debido a la presencia del N-glicano que no contiene el epítopo β-1,2-xilosa, promoviendo la diferenciación hacia un fenotipo Th1.

Finalmente, se ha explorado la preparación de glicoconjugados de ovoalbúmina funcionalizados con nanoclusters de oro fluorescentes. Hoy en día los nanomateriales encuentran numerosas aplicaciones en campos como la biología y la medicina. Los nanoclusters de oro están formados por un número limitado de átomos de oro y poseen propiedades como alta solubilidad en agua, baja toxicidad y emisión de fluorescencia en el espectro del infrarrojo cercano, lo que posibilita su empleo en bioimagen tanto *in vitro* como *in vivo*. La preparación de nanoclusters de oro fluorescentes se ha llevado a cabo empleando los glicoconjugados previamente sintetizados como estabilizantes del núcleo de oro. La preparación de estos derivados se basa en la idea de combinar las propiedades antigénicas de la ovoalbúmina, la capacidad de dirección de antígenos del carbohidrato junto con las

propiedades de bioimagen del cluster de oro. Se ha llevado a cabo la completa caracterización de los nanoclusters de oro por diferentes técnicas físico/químicas y un ensayo inicial de captación directa de los nanoclusters en células dendríticas murinas.

Acknowledgments

I would like to say thank you to my supervisor, Niels, for giving me the opportunity to work in the field of Glycoscience and for trusting me enough to support my ideas and letting me follow the research directions I was interested in. Thank you to Manuel Martín Lomas for his guidance and good words over the years of my PhD and of course for the great facilities at CIC biomaGUNE.

I would also like to express my big gratitude to the collaborators who made this research work more complex and interesting. To Dr. Ron Hokke and Dr. Angela van Diepen from the Leiden Parasite Glycobiology Group for the inspiration and the work on the microarray-based screening of sera from *S. mansoni* infected children and adults. I would also like to say thank you to Prof. Bernd Lepenies for welcoming me in his former Glycoimmunology Group, at Max Planck Institute. For his time and precious comments regarding the work described in the Chapter 4, dealing with the DCs targeting. To all people from his group, especially Uwe who with big patience showed me how to perform immuno- assays and work with cells.

Thank you to all current and previous (Nerea, Carlos, Ana) members of Glycotechnology Lab, who not only helped me with the scientific part of my thesis but also made me feel welcome in San Sebastián, which with time became my second home. To Begoña who showed me and put with me my first glycosylation reaction and who helped me with synthetic but also all other possible problems. To Sonia, from the left side of my desk, who showed me how to express and work with enzymes and microarrays. But more importantly who was constantly pushing me and helping me with writing up of my thesis… For her motivation, guidance and finally for being my friend.

To my friend, Danijelka for her support in any possible situation and being there for me whenever I needed her. For waking up at 4 while being in Argentina, calming me down, laughter, tapping and snapping of cells… And to my Spanish Mother Marta for her smile, Mil and being always ready to help. And of course to my Family. Dziękuję Mamie, Tacie i Madziorowi za wsparcie przez te wszystkie lata, dzwonienie i podtrzymywanie mnie na duchu. Za Taty optymizm, Mamy pomoc i Madziora wsparcie i plotki. And of course to Javi, for his constant support and huge patience, his help whenever I need him and good ideas.

List of abbreviations

AAL Aleuria aurantia lectin

Ab Antibody

Ac Acetyl

APC Antigen presenting cells

AuCN Gold nanocluster

Bn Benzyl

Bz Benzoyl

BnAb Broadly neutralizing antibody

BSA Bovine serum albumin

BSL-II Bandeiraea simplicifolia lectin II

CLR C-type lectin receptor

ConA Concanavalin A

CRD Carbohydrate recognition domain

CTLD C-type lectin domains

DC Dendritic cell

DCM Dichloromethane

DC-SIGN Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin

DDQ 2,3-Dichloro-5,6-dicyjano-1,4-benzochinon

DMAP 4-Dimethylaminopyridine

DMF Dimetyloformamid

DMSO Dimethyl sulfoxide

DNA Deoxyribonucleic acid

DSS Disuccinimidyl suberate

EDTA Ethylenediaminetetraacetic acid

ELISA Enzyme-linked immunosorbent assay

ER Endoplasmic reticulum

EtOAc Ethyl acetate

Fab Fragment antigen binding

FACS Fluorescence-activated cell sorting

FCS Fetal calf serum

Fc Fragment of antibody

FucT Fucosyltransferase

GalT Galactosyltransferase

GalNAcT N-acetylgalactosaminetransferase

GBP Glycan binding protein

GNA Galanthus nivalis agglutinin

GNP Gold nanoparticle

GT Glycosyltransferases

HRP horseradish peroxidase

HSQC Heteronuclear Single Quantum Coherence

IL Interleukin

LC-MS Liquid chromatography-mass spectrometry

LPS Lipopolysaccharide

MACS Magnetic Assisted Cell Sorting

MALDI-TOF-MS Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

MCL Macrophage C-type lectin

MeOH Methanol

MHC Major histocompatibility complex

MS Molecular sieves

MW molecular weight

nAb natural antibody

Nap 2-Naphthylmethyl ether

NHS N-hydroxysuccinimide

NMR Nuclear Magnetic Resonance

NPhth Phthalamide

OVA Ovalbumin

PBS Phosphate buffer saline

PG Protecting group

RCA Ricin

TCA Trichloroacetimidate

TCR T cell receptor

TFA Trifluoro acetic acid

TfOH Trifluoromethanesulfonic acid/triflic acid

Th T helper cell

TLC Thin layer chromatography

TLR Toll like receptors

TMSOTf Trimethylsilyl trifluoromethanesulfonate

RT Room temperature

PRR Pattern recognition receptor

SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoreses

TNF-α Tumor necrosis factor-α

TLR Toll like receptor

Treg Regulatory T cells

UV Ultraviolet

WFL Wisteria floribunda lectin

Glycans and glycan fragments

Fuc Fucose

Gal Galactose

GalNAc N-acetylgalactosamine

Glc Glucose

GlcNAc N-acetylglucosamine

Lac Lactose

LDN LacdiNAc, GalNAcβ1-4GlcNAc

LDNF GalNAcβ1-4(Fucα1-3)GlcNAc

Lewis A Galβ1-3(Fucα1-4)GlcNAc

Lewis B (Fucα1-2)Galβ1-3(Fucα1-4)GlcNAc

Lewis X or Le^x Galβ1-4(Fucα1-3)GlcNAc

Lewis Y (Fucα1-2)Galβ1-4(Fucα1-3)GlcNAc

LN Galβ1-4GlcNAc

Man Mannose

Neu5Ac Sialic acid

Tn GalNAcα-Ser/Thr

Xyl Xylose

Amino acids

Asn (N) Asparagine

Asp (D) Aspartic acid

Gln (G) Glutamine

Glu (E) Glutamic acid

Pro (P) Proline

Ser (S) Serine

Thr (T) threonine

Table of Contents

GENERAL INTRODUCTION

General Introduction

Glycosylation is one of the most common post-translational modifications in eukaryotic cells. Glycoproteins together with other glycoconjugates decorate cell surface and form the so-called glycocalyx[1] (Figure 1, A-B). Glycosylation is species and tissue specific and changes during cell development and differentiation. Aberrant glycosylation profiles can also be found in many pathologies highlighting the potential of glycans as disease markers for example in cancer or autoimmune diseases.[2] Additionally, glycosylation can have an important role in cell-cell interactions, immune response and in pathogen recognition.[3,4]

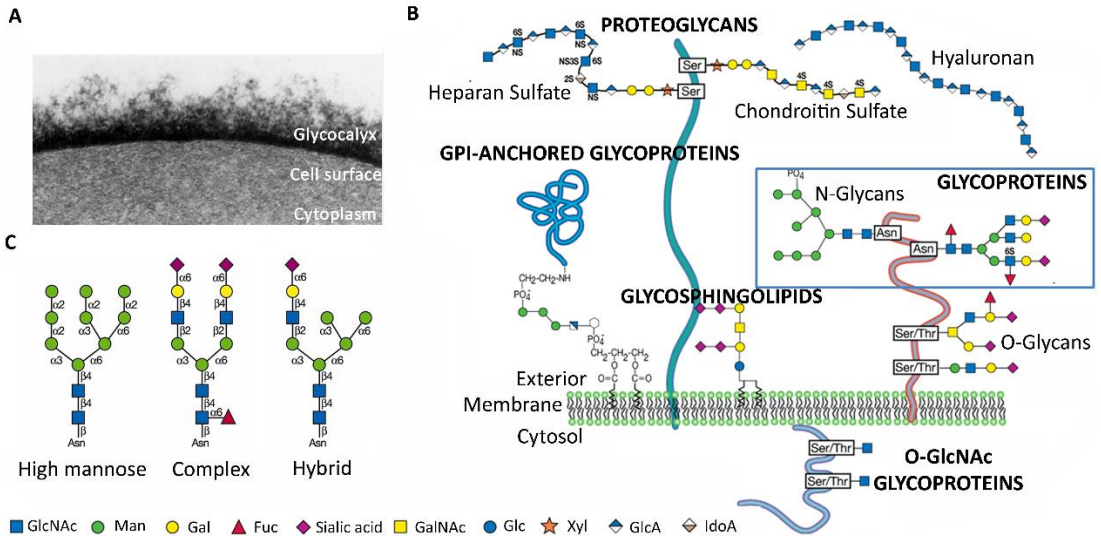


Figure 1. A. The erythrocyte glycocalyx visualized by electron microscopy (Taken from Voet and Voet, *Biochemistry*[5]); **B.** Glycans attached to the cell surface and inside the cell; **C.** Classification of N-glycans found in mammalian cells (B and C taken from *Essentials of Glycobiology,* Second Edition 2009[6]).

Around 50 % of mammalian proteins are glycosylated.[7] Carbohydrates can be attached to protein backbone either by N- or O-glycosidic bonds. **N-glycans** are covalently linked to the amide nitrogen of asparagine (Asn) residues present as part of the protein Asn-X-Ser/Thr

consensus sequence (where X is every amino acid except proline). O-glycans are attached to side chains of serine or threonine.[8] The presence of glycans in glycoproteins modulates their properties such as solubility and circulatory half-life and facilitates protein folding and trafficking. Mammalian N-glycans are generally composed of N-Acetyl-D-glucosamine (GlcNAc), D-mannose (Man), D-galactose (Gal), sialic acid (Neu5Ac), D-glucose (Glc) and L-fucose (Fuc). They share a common pentasaccharide core, Manα-1,6(Manα-1,3)Manβ-1,4GlcNAcβ-1,4GlcNAcβ-1,Asn- and can be classified into three types depending on the different substitution of pentasaccharide core: **high mannose** N-glycans in which only Man residues are attached to the core, **complex** N-glycans with GlcNAc substituted antennaes and **hybrid N-glycans**, a mixture of high-mannose and complex type N-glycans (Figure 1, C).

The biosynthetic pathway of N-glycans in eukaryotes is highly conserved and begins at the cytosolic part of the endoplasmic reticulum (ER) with the common precursor $Man_5GlcNAc_2$ transferred to a lipid dolichol phosphate (P-P-Dol).[8,9] $Man_5GlcNAc_2$-P-P-Dol is translocated across the ER membrane bilayer and is further extended in the ER lumen by glycosyltransferases (GTs). The intermediate $Gluc_3Man_9GlcNAc_2$-P-P-Dol is transferred onto Asn residues of the protein by the action of oligosaccharyltransferase (OST) (Figure 2). Before transport of a glycoprotein to the Golgi apparatus (GA) $Gluc_3Man_9GlcNAc_2$ is trimmed and N-glycans with eight or nine mannose residues are generated. At this point, quality control of misfolded protein takes place, once it is identified it is marked for further ER degradation. In the GA further glycan processing take place, glycan chains are differentiated with glycosyltransferases and glycosidases. Usually, in the early GA (*Cis*-Golgi) mannose moieties are trimmed while in the late GA (*Medial*-Golgi) the biosynthesis of complex and hybrid N-glycans is initiated. Addition of terminal GlcNAc moieties occurs which may result in branching of complex N-glycans. The final maturation of N-glycans take place in *trans*-Golgi where oligosaccharides can be further decorated with terminal galactose, N-acetyl galactosamine, sialic acid and core fucose moieties. In fact, there are over 250 glycosyltransferases present in the mammalian GA which catalyze transfer of activated sugars (glycan donor) to a glycan acceptor.[10] The high energy nucleotide glycan donors are

synthesized in the cytoplasm by nucleotide synthases and are imported to the Golgi lumen by nucleotide sugars transporters. Depending on the cell type or development stage, glycosylation is controlled by the transcription factors which regulate expression levels of glycosyltransferases, nucleotide sugars transporters and synthases creating overall huge heterogeneity of glycosylation.[9,10]

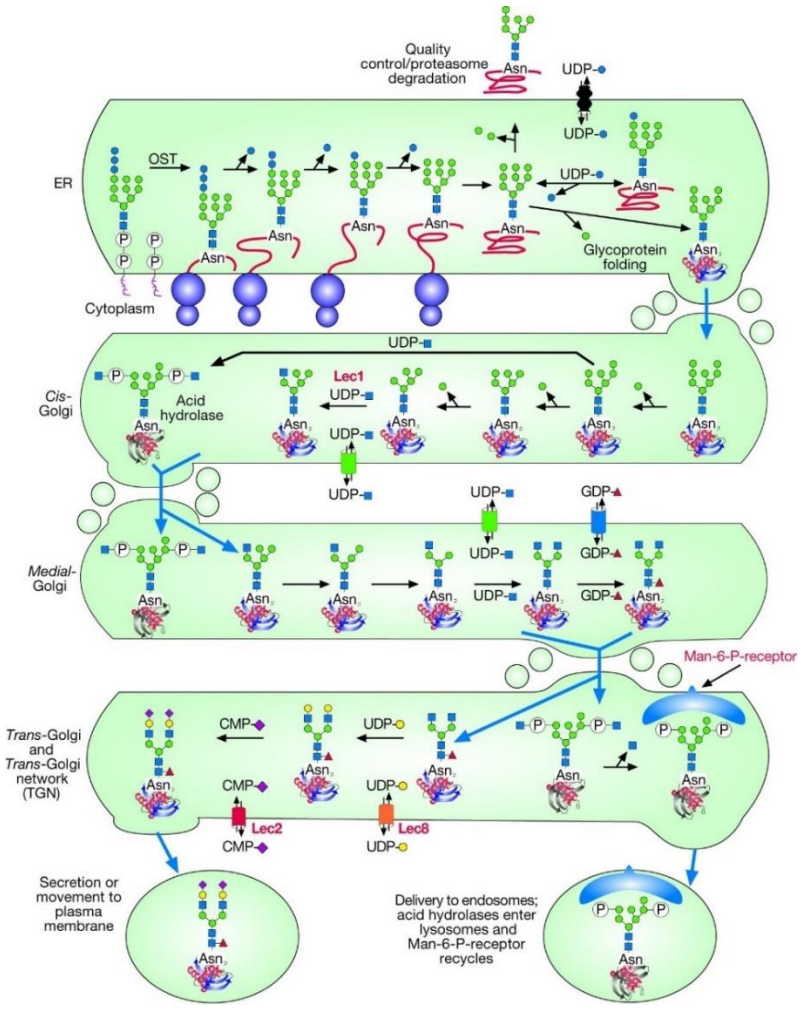


Figure 2. Processing and maturation of N-glycans (taken from *Essentials of Glycobiology,* Second Edition, 2009[8])

N-glycosylation differs largely between species (Figure 3). In yeast only high oligomannose N-glycans structures are formed. The N-glycan precursor which enters the *cis*-Golgi is not

trimmed but instead, additional mannose residues are added. Another feature that often distinguishes N-glycans of different organisms is the degree of fucosylation.[11] In mammals fucose can be linked α-1,2 to terminal Gal, α-1,3 or α-1,4 to GlcNAc of the antennae and core α-1,6 to the inner GlcNAc of the chitobiose core. The α-1,2-fucosyltransferase (FucT) activity is necessary for the biosynthesis of AB0 blood group antigens, while α-1,3 and α-1,4 FucT produce Lewis X (Galβ-1,4(Fucα-1,3)GlcNAc) related compounds. Core α-1,6-linked fucose of the inner GlcNAc on the other hand, is a typical mammalian feature, which is required for polysialylation of N-glycans[12] and which effects the properties of protein to which is linked, for example by blocking the hydrolytic activity of human asparaginase which cleaves the N-glycosidic linkage between protein and the N-glycan[13] or influencing the effector function of the IgG antibody.[14,15] The presence of core α-1,6-fucose substitution in the inner GlcNAc can be also found in the glycome of insects or invertebrates, but never in plants.[12]

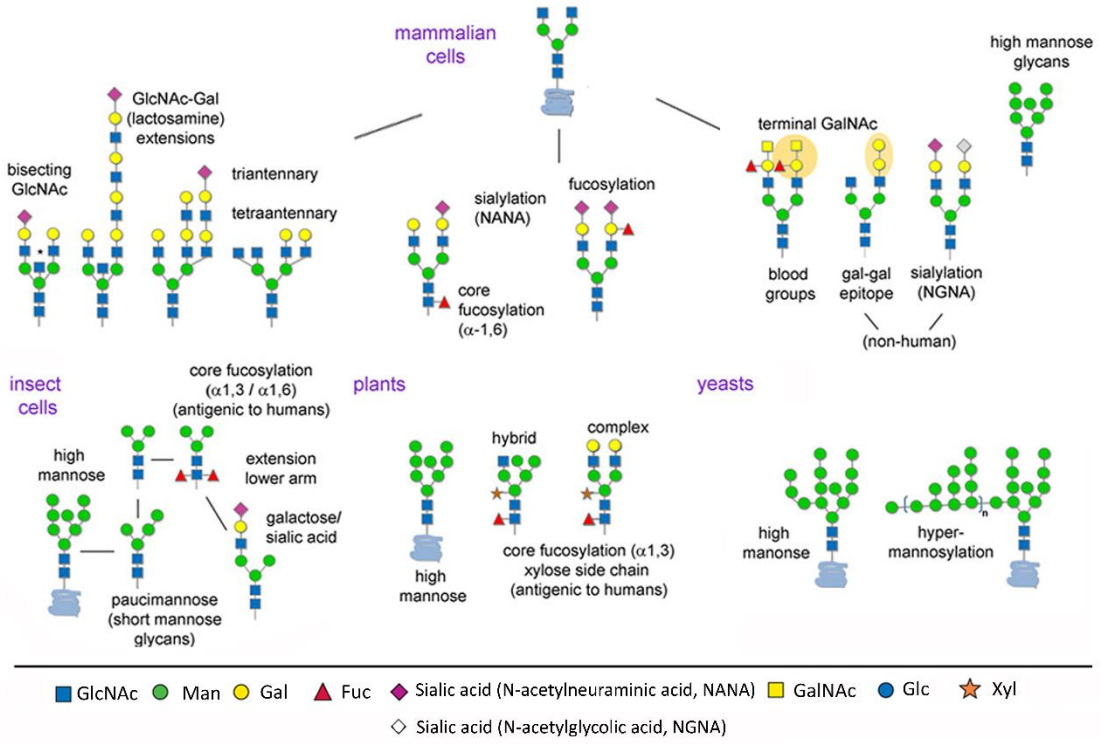


Figure 3. N-glycans commonly found in glycoproteins of different organisms (taken from New England BioLabs, https://www.neb.com)

Most of the plant-derived N-glycans are decorated with core α-1,3-fucose and/or core β-1,2-xylose. In mammals, both modifications are believed to be at least partially responsible for immunogenicity of plant-glycoproteins[16] and to be involved in the pathophysiology of several food allergens.[17] Core α-1,3-fucose and core β-1,2-xylose can be also found in the glycome of some invertebrates, such as insects and parasites and may have an important role in the induction of immune responses in parasites infected hosts.[17] These organisms express also smaller high mannose (paucimanosidic) N-glycans and bis-fucosylated structures presenting both core α-1,6-Fuc/α-1,3-Fuc modifications.

The complexity of glycans is tremendous. Nowadays it is understood that unlike RNA, DNA or proteins, glycosylation does not operate with templates as information about glycan structures is not genetically encoded. This is essential from the evolutionary point of view. In order to survive fast mutating pathogens, long-live hosts have to develop defense strategies, a feature known as Red Queen Effect- *it takes all the running you can to keep in the same place.*[1,18] It is well known that pathogens can attach to the host cells through certain glycans of the glycocalyx coating cell surface. Hence hosts may evade the pathogens invasion by changing of glycosylation patterns on their cell surface. This process can be achieved rapidly by the biosynthesis of various heterogenic glycostructures without DNA mutation. At the end this process results in huge diversity of glycans, what might be essential for the survival but this makes **Glycobiology** a challenging field of research.

In order to decode the information carried by a glycan structure, it is necessary to analyze and characterize the interactions with carbohydrate recognizing macromolecules. **Glycan binding proteins** (GBPs) and **carbohydrate-binding antibodies** play crucial roles in essential biological processes. Pathogens attach to host cells via GBPs which facilitate their further invasion.[19,20] On the other hand, GBPs expressed on host cells bind to non-self glycans, promote pathogen recognition and initiate activation of immune responses. The generation of anti-carbohydrate antibodies can be induced as a response to the glycan antigens presented on the pathogen, and can lead to their neutralization. The first HIV broadly neutralizing antibody (bnAbs) isolated from an asymptomatic HIV-positive patient is directed

towards high mannose type N-glycans presented with a high density on its viral envelope.[21,22,23] Interestingly, anti-carbohydrate antibodies are also found in the sera of all healthy individuals in the absence of the pathogens or immunization factors. So called natural antibodies (nAbs) are directed against structures found in normal human tissues and many of them are directed towards carbohydrate motifs. One of the first identified and well-known nAbs are directed against blood group antigens A and B.[24,25] Evaluation of glycan specificity of antibodies induced in bacterial, viral or parasitic infections became a crucial step towards development of carbohydrate-based vaccines.

The biggest challenge in the study of carbohydrate-protein interactions is the limited availability of glycan structures and their enormous complexity. Unlike proteins, carbohydrates cannot be cloned and their isolation from natural sources is very often laborious and time consuming, providing only low quantities of sugars and heterogeneous mixtures. Furthermore, the monovalent interaction between a carbohydrate moiety and a carbohydrate recognition domain (CRD) of GBPs is typically very weak. In nature to increase the binding affinity, proteins possess multiple CRDs or create assembles which results in multivalent complexes[26] whereas sugars are presented in multiple copies at the cell surface. To overcome these issues, new efficient and fast carbohydrate synthetic strategies have to be developed together with high throughput screening methods allowing for a simultaneous analysis of numerous carbohydrate-sugar interactions with minimization of the required amount of both carbohydrate and GBPs and the multivalent presentation of the former.

The development of **chemoenzymatic synthesis** towards complex carbohydrates has become the most promising method towards well defined and highly pure compounds.[27,28,29,30] In this strategy, conserved N-glycan structures are chemically synthesized and further diversify with recombinant glycosyltransferases (Figure 4, A). Use of glycan processing enzymes reduces amount of synthetic steps to minimum and allows for generation of glycan libraries containing highly defined glycans in relatively short times.

Together with the development of chemoenzymatic strategies, **glycan microarrays** have emerged as a high throughput screening method for carbohydrate-protein interactions

(Figure 4, B).[31,32,33,34] Based on the DNA microchips and subsequently protein array technology,[35] in 2002, the first glycan microarray was developed.[33] Typically, the sugar ligands are immobilized on the activated surface of glass slides and screened with solutions of GBPs. After incubation, the array is washed to remove any unbound material and binding events are measured and analyzed.

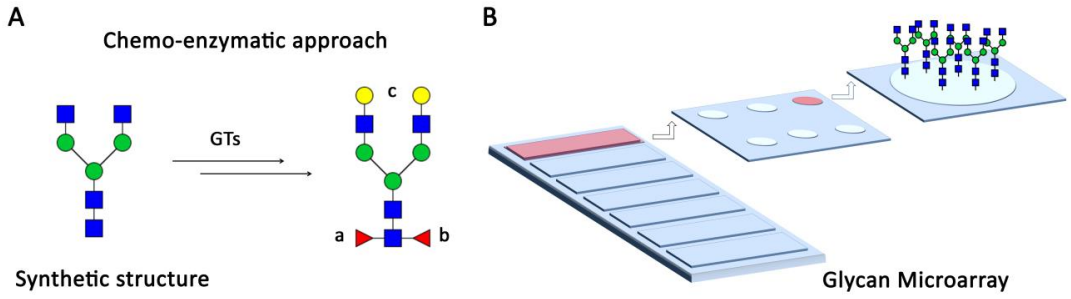


Figure 4. A. Chemoenzymatic synthesis of N-glycans, beginning with conserved, synthetic structure for its further differentiation with GTs; **B.** Schematic representation of glycan microarray.

The most common detection method used for the microarray screening is fluorescence read-out which involves the use of fluorescently labeled proteins or antibody directed against the analyte. The biggest advantage of the glycan microarray as a screening tool is the reduction of valuable material necessary to study carbohydrate-protein interactions and the capacity of simultaneous screening of hundreds glycans and multiple proteins.

Glycan microarrays have been successfully applied for the characterization of the sugar binding specificities of various GBPs, like **plant lectins** and **C-type lectins receptors** (CLR) and have provided important information about the role of glycans in different biological processes.

Plant lectins, are believed to play an important role in plant defense and act as storage proteins.[36,37] By the recognition of specific carbohydrates of predators, including plant eating insects or herbivores, plants lectins may assure the survival of tissues in which they are located. Today, many plant lectins are commercially available and they have become

important tools in glycobiology research.[38] Lectin affinity chromatography (Figure 5, A) is widely used for purifying glycoproteins while screening of glycan arrays with plant lectins is used as a functionality test to evaluate the correct immobilization of ligands on the array surface (Figure 5, B).

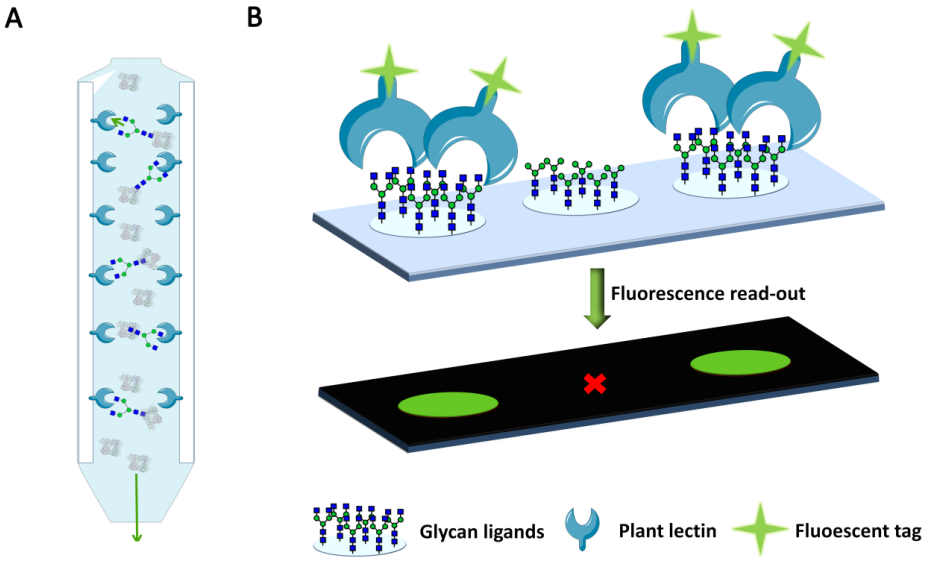


Figure 5. A. Schematic representation of lectin affinity chromatography; **B.** Schematic representation of glycan microarray functionality test applying fluorescently labeled plant lectin specific for terminal GlcNAc residues. Non-bound ligand shows complete lack of fluorescent.

C-type lectin receptors belong to the largest superfamily of animal lectins (over 1000 proteins).[39,40,41] CLRs contain one or more carbohydrate recognizing domain (CRDs) which mediates sugar binding, in most cases, in coordination with Ca^{2+}. CLRs can be found in either soluble or in transmembrane form and can be further classified depending on their carbohydrate specificity determined by certain amino acids present in the CRD. It has been reported that CLRs expressing EPN amino acid motif are specific to mannose-derived ligands while CLRs expressing QPD sequence are specific for galactose. Some of the transmembrane CLRs expressed on the cells of immune system play an important role, acting as pattern recognizing receptors (PRR). PRR on the antigen presenting cells (APC) can recognize self-

and non-self-carbohydrates of pathogens[42], capture and present them to lymphocytes. The most well-known and described APCs are dendritic cells (DCs) which express on their surface various C-type lectin receptors or even more extensively studied Toll like receptors (TLRs).[43] APCs act as a link between innate and adaptive immune system when the innate immune system alone cannot prevent from further pathogen invasion. Innate immune system is a first line of defense, which acts immediately in response to infection. Phagocytic cells of the innate immune system, like macrophages and neutrophils are constantly circulating in the organism ready to fight pathogens.[44] Adaptive immunity is normally silent and in order to be activated, foreign, antigenic substance has to be recognized and captured by APCs. The main family of adaptive immune cells are lymphocytes. B lymphocytes (which are also APC) produce antibodies and as a part of so called humoral immunity, defending against extracellular microbes. The whole microbes can be recognized by the B cells through their antigen receptors which are membrane-bound antibodies. Upon antigen recognition B lymphocytes are activated and differentiated into plasma cells producing and secreting antibodies directed against encountered pathogen. However, if the pathogen antigen is a protein, the effective antibody responses required second type of lymphocytes, called T cells. T cells are part of the cellular immunity and are mostly specialized to combat intracellular pathogens. T lymphocytes can be divided into helper T cells, (T_H, expressing $CD4^+$ surface protein), cytotoxic T lymphocytes (CTLs, $CD8^+$) and regulatory T calls (T_{reg}). $CD4^+$ Th cells help B cells to produce antibodies[45] and macrophages to kill ingested microbes, CTLs eliminate infected and tumor host cells whereas T_{reg} control immune responses.

Once DCs recognize the pathogen through C-type lectin receptors, the antigen is internalized and processed into smaller peptide fragments. Depending on the origin of the antigen, they are further presented on either major histocompatibility complex (MHC) molecule class-I or MHC-II which are recognized by $CD8^+$ T or $CD4^+$ T cells respectively. The processed peptide antigens (as well as glycopeptides) are presented by MHC-II and MHC-II/peptide complex are recognized by T cell receptors (TCRs) expressed by $CD4^+$ T cells. The activation of T cells requires not only recognition of MHC-II/peptide complex but also binding of co-stimulatory

molecules CD80 or CD86 of DCs by CD28 of T cells. Upon stimulation, $CD4^+$ T cells expand in number and begin to differentiate into effector T cells which secret soluble proteins called cytokines (Figure 6).

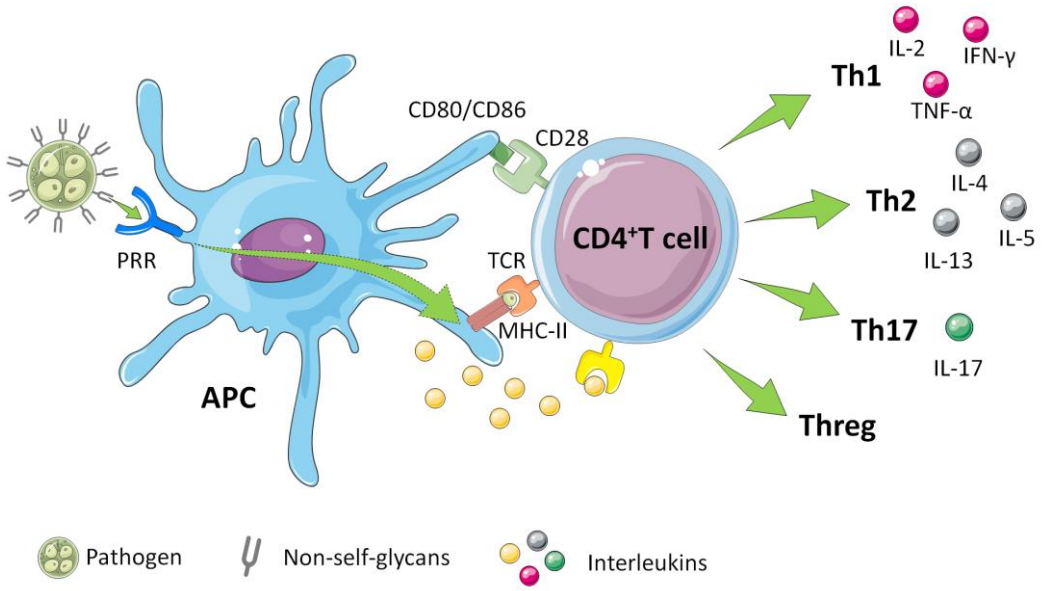


Figure 6. Schematic representation of DCs T cells interactions. Uptake, processing and presentation of antigen by DCs to $CD4^+$ T cells leads to it activation and differentiation into effector T cells.

Cytokines which mediate communication between leukocytes are called interleukins (IL). Depending on cytokines which are produced by $CD4^+$ T cells and activated DCs, T cells can further differentiate into Th1, Th2, Th17 cells or induced regulatory T cells (Treg)[46](Figure 6). Th1 cells are involved mostly in the cell-mediated immunity against viruses and intracellular bacteria and produce predominantly inflammatory cytokines, such as IL-2, interferon- γ (IFN-γ) and tumor necrosis factor-α (TNF-α). Unlike Th1, Th2 cells promote humoral immunity and are important for the defense against parasite infections but also play a role in allergic responses. Th2 cells secrete IL-4, IL-5 and IL-13. More recently, a third type of effector T helper cells was identified and named Th17. $CD4^+$ T cells differentiate into Th17 cells and further produce cytokine IL-17. Th17 elicit strong inflammatory response, play a role in the

host defense against extracellular bacteria and fungi but also can stimulate autoimmunity. Finally, Treg regulate Th1 and Th2 responses.

Finally, upon exposure to antigen the effector T cells proliferate and increase in number. Part of them are employed to clear the infection, however, the majority of these cells die. The remaining effector T cells become the memory cells and provide quick protection upon re-exposure to the same pathogen. Similarly, certain amount of antibody producing plasma cells migrate to the bone marrow where they can live for many years as memory cells. Half of our circulating IgGs are produced from these cells and reflect history of antigen exposure. The production of long live memory cells against pathogens is the main goal of effective vaccination.

In the last decades, the development of new vaccines preventing from bacterial infections emerged together with appearance antibiotic resistance. Carbohydrate based vaccines have a long history, already in 1923 Heidelberg and Avery identified a soluble polysaccharide from *Streptococcus pneumoniae*[47], which few years later after intradermal injection induced the development of anti-carbohydrate specific antibodies.[48]

Nowadays, the most successful carbohydrate-based vaccines are produced against bacterial infections and they are composed of the pathogen lipopolysaccharide or capsular polysaccharides coupled to protein carrier, which can elicit antibody responses in their host. Nowadays, the Food and Drug Administration of the United States (USFDA) has approved several polysaccharide-based vaccines preventing of infections caused by *Streptococcus pneumoniae*, *Haemophilus influenza* type b (Hib) and *Salmonella typhi.*[49]

The identification of the minimal bacterial carbohydrate antigen and their further conjugation to protein carriers is a crucial step in the design and construction of valuable synthetic carbohydrate based vaccines.[50] Carbohydrates alone are considered to activate B cells without involvement of T cells and thus do not stimulate effective immune protection. In conjugation with antigenic proteins, glycoproteins induce T-cell dependent immune responses which may provide long lasting protection and antibodies class switching.[51] The immunogenicity of glycoproteins is enhanced compared to polysaccharide as such. The first

large scale synthesis of conjugate vaccines composed of a fully synthetic capsular polysaccharide antigen coupled to tetanus toxoid (TT) was described and evaluated by Verez-Bencomo and colleagues and it was directed against *Haemophilus influenza.*[52]

The development of anti-parasitic vaccines is far more complicated than anti-bacterial vaccines. Parasites are biologically and genetically complex organisms, expressing thousands of different sugar structures, which change with the development stage of the parasite.[53]

The identification of glycan antigens for us in conjugate vaccine development is extremely challenging. Thus fundamental studies regarding the parasite glycome, identification of antigenic structures or the analysis of possible mechanisms of evading the host of immune system are needed.

In recent years, shotgun glycan microarrays composed of the whole glycome of parasites have become an attractive tool for the evaluation of possible roles of glycans during parasitic infection and identification of antigenic structures. Hokke and colleagues constructed natural shotgun microarrays which comprise N-glycans and lipid-glycans directly isolated from various life stages of *Schistosoma mansoni* parasites.[54,55] Extracts containing very complex sugars were isolated, subjected to detailed glycomic mass spectrometric analysis, purified and printed on the array. This natural microarray was further used for profiling anti-carbohydrate polyclonal antibodies from sera of *S. mansoni* infected patients. Natural glycan arrays are an excellent way to gain structural information about the specificity of glycan binding antibodies and GBP, especially when no prior knowledge about the potential ligands is available. One of the greatest benefit of the shotgun glycan microarrays is the presence of very complex structures which synthesis could be difficult and time consuming. However, because of difficult isolation, some of the fractions printed on the shotgun arrays are rather a mixture of different glycan structures than a single ligand. In this context, glycan arrays containing well defined and characterized glycans synthesized chemoenzymatically, could be an excellent complementary tool towards evaluation of the biological role of glycan structures in *S. mansoni* infections and to determine lead antigenic structures for the development of vaccines against *S. mansoni* (Figure 7).

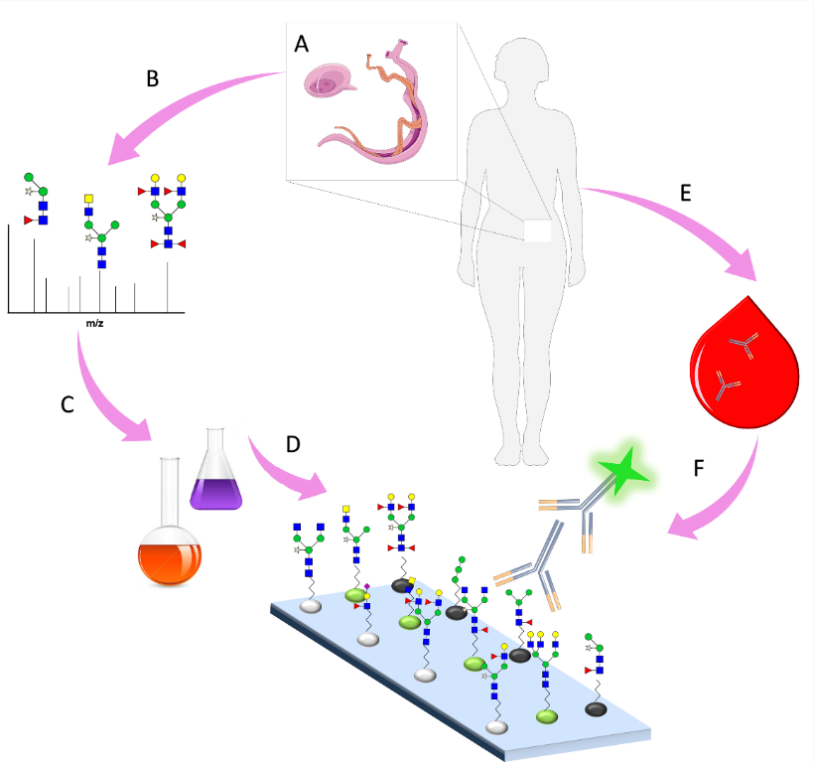


Figure 7. Use of synthetic glycan microarrays towards identification of antigenic glycans of *S. mansoni*; **A.** *S. mansoni* eggs and adult worms; **B.** Spectrometry analysis of parasite glycome; **C.** Chemoenzymatic synthesis of parasitic glycan structures; **D.** Design and preparation of glycan array containing previously synthesized compounds; **E.** Sera obtained from *S. mansoni* infected patients; **F.** Profiling of anti-carbohydrate polyclonal antibodies from sera of *S. mansoni* infected patients.

In recent years, beside the tremendous effort in the identification and synthesis of antigenic glycans towards design of carbohydrate-based vaccines, sugars have also received attention as molecules for DC targeting.[56] Targeting antigens to dendritic cells (DC) through CLRs is a promising strategy to enhance the efficacy of vaccines.[56] CLR targeting offers some advantages compared to antibody-mediated targeting[57], the most commonly employed strategy, such as, low immunogenicity, possibility to target several CLRs simultaneously and potentially a more favorable pharmacokinetic profile when compared to large protein drugs.

Objectives

The glycome differs largely between species. In particular, N-glycosylation in plants and invertebrates vary from the mammalian structures in terms of fucosylation degree and type, 2-O-xylosylation or the complete lack of sialylated structures which can be replaced by other terminal motifs. These differences have been found to have profound impact on the immune responses in mammals infected with parasites or being allergic to certain plants or insects.
S. mansoni is a parasitic worm which infects significant number of people in developing countries. *S. mansoni* expresses thousands of different sugar structures, which change with the developmental stage of the parasite. Some of these structures share common glycan elements with major plant allergens. The fundamental studies regarding the parasite glycome, the identification of antigenic structures or the analysis of host immune system evasion is a work in progress, which gather together parasitologists, chemists and immunologists to better understand *S. mansoni* infection, with the final goal of improved treatments or vaccines.

Drawing upon published antigenic carbohydrate structures identified by spectrometric analysis of the most infective stages of the parasite (cercariae and eggs), we will prepare chemoenzymatically library of non-mammalian, 2-O-xylosylated N-glycans. We aim to create a large set of defined N-glycans, with α-1,3 and α-1,6 core fucosylation and diverse terminal structural elements, including LN (Galβ-1,4-GlcNAc), LDN (GalNAcβ-1,4-GlcNAc), Lewis X (Galβ-1,4-(Fucα-1,3)GlcNAc), and LDNF (GalNAcβ-1,4-(Fucα-1,3)GlcNAc) and study their interactions with relevant glycan binding proteins (GBPs) and anti-carbohydrate antibodies induced in patients suffering from *S. mansoni* infection.

In **Chapter 1** the chemical synthesis of six conserve N-glycan structures will be presented, followed by their enzymatic diversification using a set of recombinant glycosyltransferases (Figure 8).

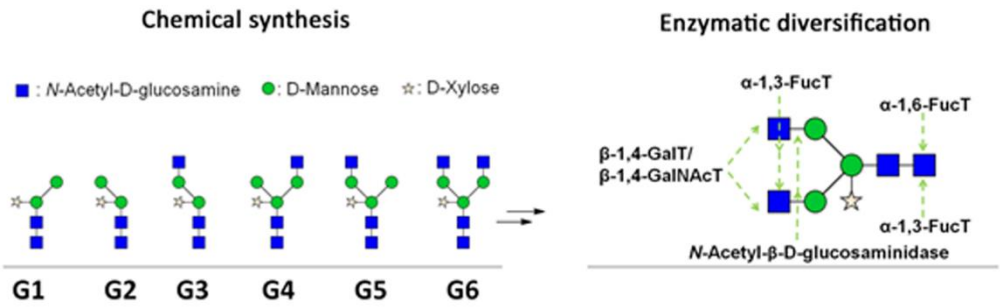


Figure 8. General strategy for the chemoenzymatic synthesis of xylosylated N-glycans; **A.** Target structures with a common xylosylated tetrasaccharide core (**G1**-**G6**); **B.** Exemplified enzymatic diversifications of core structures shown for the example of biantennary glycan **G6**.

In recent years, glycan microarrays have been successfully applied as a miniaturized, high throughput screening method for the characterization of sugar binding specificities of various GBPs, like plant lectins and C-type lectins receptors providing important information about the role of glycans in different biological processes.

In **Chapter 2,** the design and preparation of glycan microarrays which included the xylosylated N-glycans prepared in Chapter 1 will be described. Initially, the array functionality will be interrogated with plant lectins of known, general specificity, to be subsequently probed with set of human and murine CLRs. To better understand the role of parasite N-glycans in the interactions with the CLRs present on the surface of dendritic cells (DCs), we will pay special attention on the specificity of DC-SIGN, MGL-1 and Dectin-2 which were previously shown to internalize specific glycoproteins of parasites.

The majority of antibodies generated during the course of *S. mansoni* infection are directed against certain antigenic glycan elements. Their identification and characterization can be essential for the development of potential carbohydrate-based vaccines or to better understand the role of certain parasitic sugars during the infection, like the evasion of parasite from the host immune system.

In **Chapter 3**, glycan microarrays will be employed for the screening of antibodies raised against *S. mansoni*. First, the utility of glycan microarray towards the characterization of anti-carbohydrate antibodies will be tested using monoclonal antibodies produced against sera of infected mice. Second, the glycan microarray will be used to detect and characterize polyclonal antibodies from sera of *S. mansoni* infected patients from an endemic community. As control, sera from healthy individuals will be also screened and their natural carbohydrate antibodies will be characterized. Based on the glycan microarray screening of sera from patient infected with Schistosomiasis we aim to, first identify major antigenic N-glycans structures in a context of structures available on our synthetic glycan and secondly, propose their biological role in the host-parasite interactions.

In **Chapter 4** we want to explore the use of specific N-glycans as DCs targeting molecules. Based on the microarray-assisted screening of different C-type lectin receptors (CLRs) present on the surface of DCs we want to design an *in vitro* assay and study the correlation of microarray data with cellular trials.

References

(1) Homann, A.; Seibel, J. *Natural product reports* **2009**, *26*, 1555–1571.

(2) Tuccillo, F. M.; de Laurentiis, A.; Palmieri, C.; Fiume, G.; Bonelli, P.; Borrelli, A.; Tassone, P.; Scala, I.; Buonaguro, F. M.; Quinto, I. *BioMed research international* **2014**, *2014*:742831.

(3) Nagae, M.; Yamaguchi, Y. *International Journal of Molecular Sciences* **2012**, *13*, 8398–8429.

(4) Van Kooyk, Y.; Rabinovich, G. A. *Nature immunology* **2008**, *9*, 593–601.

(5) Voet, D.; Voet, J. G. *Biochemistry*, 4th ed, John Wiley & Sons: Hoboken, 2011.

(6) Varki, A.; Sharon, N. *Historical Background and Overview*.In *Essentials of glycobiology*, 2nd ed.; Cold Spring Harber Laboratory Press: New York, 2009,

(7) Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochimica* et *Biophysica Acta* **1999**, *1473*, 4–8.

(8) Stanley, P.; Schachter, H.; Taniguchi, N. *N-glycans.* In *Essentials of Glycobiology*, 2nd ed, Cold Spring Harber Laboratory Press: New York, 2009.

(9) Lannoo, N.; Van Damme, E. J. M. *Plant Science* **2015**, *239*, 67–83.

(10) Stanley, P. *Cold Spring Harb Perspect Biol* **2011**, *3*, a005199.

(11) Staudacher, E.; Altmann, F.; Wilson, I. B. H.; März, L. *Biochimica et Biophysica Acta -General Subjects* **1999**, *1473*, 216–236.

(12) Paschinger, K.; Staudacher, E.; Stemmer, U.; Fabini, G.; Wilson, I. B. *Glycobiology* **2005**, *15*, 463–474.

(13) Noronkoski, T.; Mononen, I. *Glycobiology* **1997**, *7*, 217–220.

(14) Ackerman, M. E.; Crispin, M.; Yu, X.; Baruah, K.; Boesch, A. W.; Harvey, D. J.; Dugast, A.-S.; Heizen, E. L.; Ercan, A.; Choi, I. *The Journal of clinical investigation* **2013**, *123*, 2183.

(15) Peipp, M.; van Bueren, J. J. L.; Schneider-Merck, T.; Bleeker, W. W.; Dechant, M.; Beyer, T.; Repp, R.; van Berkel, P. H.; Vink, T.; van de Winkel, J. G. *Blood* **2008**, *112*, 2390–2399.

(16) Van Ree, R.; Cabanes-Macheteau, M.; Akkerdaas, J.; Milazzo, J.-P.; Loutelier-Bourhis, C.; Rayon, C.; Villalba, M.; Koppelman, S.; Aalberse, R.; Rodriguez, R. *Journal of Biological Chemistry* **2000**, *275*, 11451–11458.

(17) Faveeuw, C.; Mallevaey, T.; Paschinger, K.; Wilson, I. B. H.; Fontaine, J.; Mollicone, R.; Oriol, R.; Altmann, F.; Lerouge, P.; Capron, M. *European journal of immunology* **2003**, *33*, 1271–281.

(18) Varki, A. *Cell* **2006**, *126*, 841–845.

(19) Guzman-Aranguez, A.; Argüeso, P. *The ocular surface* **2010**, *8*, 8–17.

(20) Geissner, A.; Anish, C.; Seeberger, P. H. *Current Opinion* in *Chemical Biology* **2014**, *18*, 38–45.

(21) Scanlan, C. N.; Pantophlet, R.; Wormald, M. R.; Saphire, E. O.; Stanfield, R.; Wilson, I. A.; Katinger, H.; Dwek, R. A.; Rudd, P. M.; Burton, D. R. *Journal of virology* **2002**, *76*, 7306–7321.

(22) Calarese, D. A.; Lee, H.-K.; Huang, C.-Y.; Best, M. D.; Astronomo, R. D.; Stanfield, R. L.; Katinger, H.; Burton, D. R.; Wong, C.-H.; Wilson, I. A. *Proceedings of the National Academy of Sciences* **2005**, *102*, 13372–13377.

(23) Burton, D. R.; Pyati, J.; Koduri, R.; Sharp, S. J.; Thornton, G. B.; Parren, P.; Sawyer, L.; Hendry, R. M.; Dunlop, N.; Nara, P. L. *Science* **1994**, *266*, 1024–1027.

(24) Obukhova, P.; Rieben, R.; Bovin, N. *Xenotransplantation* **2007**, *14*, 627–635.

(25) Huflejt, M. E.; Vuskovic, M.; Vasiliu, D.; Xu, H.; Obukhova, P.; Shilova, N.; Tuzikov, A.; Galanina, O.; Arun, B.; Lu, K.; Bovin, N. *Molecular Immunology* **2009**, *46*, 3037–3049.

(26) Collins, B. E.; Paulson, J. C. *Current opinion in chemical biology* **2004**, *8*, 617–625.

(27) Wang, Z.; Chinoy, Z. S.; Ambre, S. G.; Peng, W.; McBride, R.; de Vries, R. P.; Glushka, J.; Paulson, J. C.; Boons, G.-J. *Science* **2013**, *341*, 379–383.

(28) Blixt, O.; Razi, N. *Methods in enzymology* **2006**, *415*, 137–153.

(29) Unverzagt, C. *Carbohydrate research* **1997**, *305*, 423–431.

(30) Ichikawa, Y.; Lin, Y. C.; Dumas, D. P.; Shen, G. J.; Garcia-Junceda, E.; Williams, M. A.; Bayer, R.; Ketcham, C.; Walker, L. E. *Journal of the American Chemical Society* **1992**, *114*, 9283–9298.

(31) Paulson, J. C.; Blixt, O.; Collins, B. E. *Nature chemical biology* **2006**, *2*, 238–248.

(32) Park, S.; Lee, M.-R.; Shin, I. *Chemical Communications* **2008**, *37*, 4389–4399.

(33) Fukui, S.; Feizi, T.; Galustian, C.; Lawson, A. M.; Chai, W. *Nature biotechnology* **2002**, *20*, 1011–1017.

(34) Feizi, T.; Fazio, F.; Chai, W.; Wong, C.-H. *Current opinion in structural biology* **2003**, *13*, 637–645.

(35) Kulesh, D. A.; Clive, D. R.; Zarlenga, D. S.; Greene, J. J. *Proceedings of the National Academy of Sciences* **1987**, *84*, 8453–8457.

(36) Lannoo, N.; Van Damme, E. J. M. *Frontier in Plant Science* **2014**, *5*, 397.

(37) Peumans, W. J.; Van Damme, E. *Plant physiology* **1995**, *109*, 347.

(38) Peumans, W. J.; Damme, E. J. V. *Biotechnology and Genetic Engineering Reviews* **1998**, *15*, 199–228.

(39) Drickamer, K.; Taylor, M. E. *Current opinion in structural biology* **2015**, *34*, 26–34.

(40) Drickamer, K. *Current opinion in structural biology* **1999**, *9*, 585–590.

(41) Lepenies, B.; Lee, J.; Sonkaria, S. *Advanced drug delivery reviews* **2013**, *65*, 1271–1281.

(42) Cambi, A.; Koopman, M.; Figdor, C. G. *Cellular microbiology* **2005**, *7*, 481–488.

(43) Figdor, C. G.; van Kooyk, Y.; Adema, G. J. *Nature Reviews Immunology* **2002**, *2*, 77–84.

(44) Abbas, A. K.; Lichtman, A. H.; Pillai, S. *Cellular and Molecular Immunology*; Elsevier Health Sciences: Philadelphia, 2014.

(45) Crotty, S. *Nature Reviews Immunology* **2015**, *15*, 185–189.

(46) Geijtenbeek, T. B.; Gringhuis, S. I. *Nature Reviews Immunology* **2009**, *9*, 465–479.

(47) Heidelberger, M.; Avery, O. T. *The Journal of experimental medicine* **1923**, *38*, 73.

(48) Francis, T.; Tillett, W. S. *The Journal of experimental medicine* **1930**, *52*, 573–585.

(49) Oyelaran, O.; Gildersleeve, J. C. *Expert review of vaccines* **2007**, *6*, 957-969.

(50) Huang, Y.-L.; Wu, C.-Y. *Expert review of vaccines* **2010**, *9*, 1257-1274.

(51) Wolfert, M. A.; Boons, G.-J. *Nature chemical biology* **2013**, *9*, 776–784.

(52) Verez-Bencomo, V.; Fernández-Santana, V.; Hardy, E.; Toledo, M. E.; Rodríguez, M. C.; Heynngnezz, L.; Rodriguez, A.; Baly, A.; Herrera, L.; Izquierdo, M.; Villar, A.; Valdés, Y.; Cosme, K.; Deler, M. L.; Montane, M.; Garcia, E.; Ramos, A.; Aguilar, A.; Medina, E.; Toraño, G.; Sosa, I.; Hernandez, I.; Martínez, R.; Muzachio, A.; Carmenates, A.; Costa, L.; Cardoso, F.; Campa, C.; Diaz, M.; Roy, R. *Science* **2004**, *308*, 522–525.

(53) Hokke, C. H.; Deelder, A. M.; Hoffmann, K. F.; Wuhrer, M. *Experimental parasitology* **2007**, *117*, 275–283.

(54) De Boer, A. R.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Glycoconjugate journal* **2008**, *25*, 75–84.

(55) De Boer, A. R.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Analytical chemistry* **2007**, *79*, 8107–8113.

(56) Van Kooyk, Y.; Unger, W. W.; Fehres, C. M.; Kalay, H.; Garcia-Vallejo, J. J. *Molecular immunology* **2013**, *55*, 143–145.

(57) Tacken, P. J.; de Vries, I. J. M.; Gijzen, K.; Joosten, B.; Wu, D.; Rother, R. P.; Faas, S. J.; Punt, C. J.; Torensma, R.; Adema, G. J. *Blood* **2005**, *106*, 1278–1285.

CHAPTER 1

CHEMOENZYMATIC SYNTHESIS OF β2-XYLOSYLATED N-GLYCANS

1.1 Introduction

N-glycans have multiple roles in various biological processes thus their synthesis has been extensively studied over the years. However, due to polyhydroxyl nature of carbohydrates, the development of general and efficient chemical strategies of their synthesis is not trivial. The major challenge in the carbohydrate synthesis relies on the modification of a specific hydroxyl group in the presence of others and this task is usually solved by multiple protection/deprotection strategy. Moreover, the enormous complexity of oligosaccharides synthesis come from their similar chemical sequence: monosaccharide building blocks can adopt either α or β anomeric configuration (stereoisomers) and can be coupled by different linkages to the next monosaccharide unite (regioisomers). Improvement of synthetic methodology has been addressed by several research groups during the years, both in solution[1] and on solid phase.[2–4] One important contribution in the synthesis of N-glycans was made in mid-eighties by groups of Paulsen[5] and Ogawa[6–8] who used independent approaches to synthesized complex, N-glycans of glycoproteins.

Nevertheless, despite the big progress in synthetic carbohydrate chemistry, the preparation of complex N-glycans is still time consuming and often leads to low yields and high costs. As an alternative to purely organic synthesis, chemoenzymatic preparation of N-glycans have been developed in recent years.[9–12] It combines carbohydrate chemistry of a common scaffold with enzymatic glycosylation, taking advantage of the flexibility of the chemical approach and of the efficiency and selectivity of the modifications performed by enzymes. The chemoenzymatic strategy reduces the amount of necessary synthetic, structures to minimum and utilizes glycosyltransferases (GTs) or glycosidases to further diversify previously prepared structures. One of the first example of use glycosyltransferases towards synthesis of a simple glycan, N-acetyllactosamine was given by Whitesides and co-workers.[13] They designed the recycling system that generated UDP-galactose from inexpensive UDP-glucose by UDP-galactose epimerase. So created donor was used by galactosyltransferase (GalT) to elongate glucosamine acceptor whereas the hydrolyzed UDP

cofactor was further regenerated to UDP-glucose and UDP-galactose. The proposed enzymatic strategy overcame the biggest limitation of enzymatic glycosylation at that time, which was the high cost of sugar nucleotides. Few years later Paulson and co-workers described the efficient, fully enzymatic synthesis of sialyllactosamine,[14] based on the earlier study on the specificity of the sialyltransferase purified from bovine colostrum which transfers sialic acid to β-D-galactosides.[15] Finally, the chemoenzymatic approach was successfully applied for the generation of more complex N-glycans, in mid-nineties Unverzagt described the preparation of sialylated biantennary N-glycan with the help of two glycosyltransferases.[16] The terminal galactose and sialic acid residues were enzymatically introduced at the end onto a chemically prepared deprotected heptasaccharide with overall high yield of 87%.

In recent years, the use of glycosyltransferases for oligosaccharide synthesis has become a complementary approach to the chemical synthesis, and offers many advantages over traditional methods.[17] The use of enzymes results in the formation of defined products with complete regio- and stereoselectivity, without involvement of laborious protection/deprotection steps. Glycosyltransferases are known to be very specific towards both, glycosyl donor and acceptor but may also exhibit some flexibility, especially *in vitro*, when the substrates are externally provided. These features make glycosyltransferases very attractive tools for carbohydrate chemistry.[18,9] Beside glycosyltransferases, the glycosidic bond formation can be also achieved using glycosidase. *In vivo* these carbohydrate processing enzymes catalyze the hydrolysis of glycosidic linkage, however, they can also synthesize glycosides *in vitro*. Reaction with glycosidases gives the glycosyl-enzyme complex which can be intercepted either by water resulting in hydrolysed product or by a glycosyl acceptor in order to form new glycoside. The formation of new oligosaccharide *in vitro* can be achieved using appropriate glycosyl donors which are rapidly and irreversibly cleaved by the enzyme, for example p-nitrophenyl glycosides.[19]

In this Chapter, the chemoenzymatic synthesis of non-mammalian β-1,2-xylosylated N-glycans will be described. Some of these structures were identified in the glycome of

parasite helminths and trematodes such as *S. mansoni* or they are part of plant allergens. In the first part of this Chapter, we will focus on the chemical synthesis of six conserved core structures **G1**-**G6** (Figure 1, 1) with ready installed core β-1,2-xylosylation. Taking into account that in plants β-1,2-xylosylation has been suggested to occurs prior to other core modifications[20] the xylose containing N-glycans could be employed as starting point for further enzymatic diversifications. The introduction of β-1,2-xylose enzymatically will not be evaluated because it has been reported that the β-1,2-xylosyltransferase isolated from *Arabidopsis thaliana* (XYLT) always requires the presence of β-1,2-linked GlcNAc at the non-reducing terminus of the α-1,3-branched mannose residue (GlcNAcβ-1,2Manα-1,3Man) and does not tolerate α-1,3-linked fucose in the chitobiose core or β-1,4-linked terminal galactose residues of GlcNAcβ-1,2Manα-1,3Man branching.

In the second section of the Chapter, we will describe the expression of glycosyltransferases and enzymatic reactions performed on synthetic, deprotected structures (Figures 1, 2). We will introduce α-1,3-fucose or/and α-1,6-fucose core modifications using different fucosyltransferases. Our target structures are also decorated with terminal structural elements including LN (Galβ-1,4GlcNAc), LDN (LacdiNAc, GalNAcβ-1,4GlcNAc) or Lewis X (Galβ-1,4(Fucα-1,3)GlcNAc) motifs (Figure 1, 3). LN is the most common structural element of mammalian complex-type N-glycans, it decorates the terminal GlcNAc residues which can be additionally substituted with fucose, sialic acid and sulfate groups. LDN and Lewis X are an alternative type of nonreducing terminal glycan structures which are shared by both various parasites and their human host. [21,22] Both motifs seem to be highly abundant in helminths, including *S. mansoni* which express Lewis X, LDN and its fucosylated derivative LDNF within the different life stages.[23] In mammals, expression of Lewis X and LDN is only restricted to certain glycoproteins.[24] These both glycan moieties are commonly capped by sialic acid and are much lower abundant than LN element. Lewis X plays a role in the development of the nervous system of the vertebrates while the sialyl Lewis X mediates human sperm binding during fertilization.[25] Both structural elements are overexpressed on the surface of cancer cells. LDN-type glycans, on the other hand have been found on the N-

and O-glycans decorating certain hormons and other mammalian glycoproteins. Interestingly, it has been described that the humoral immune response of the host infected with *S.mansoni* is potentially directed towards LDN and Lewis X structural elements. Both motifs, are linked to the recognition of parasite by the host immune system[26] and might have a role in the modulation of immune responses.

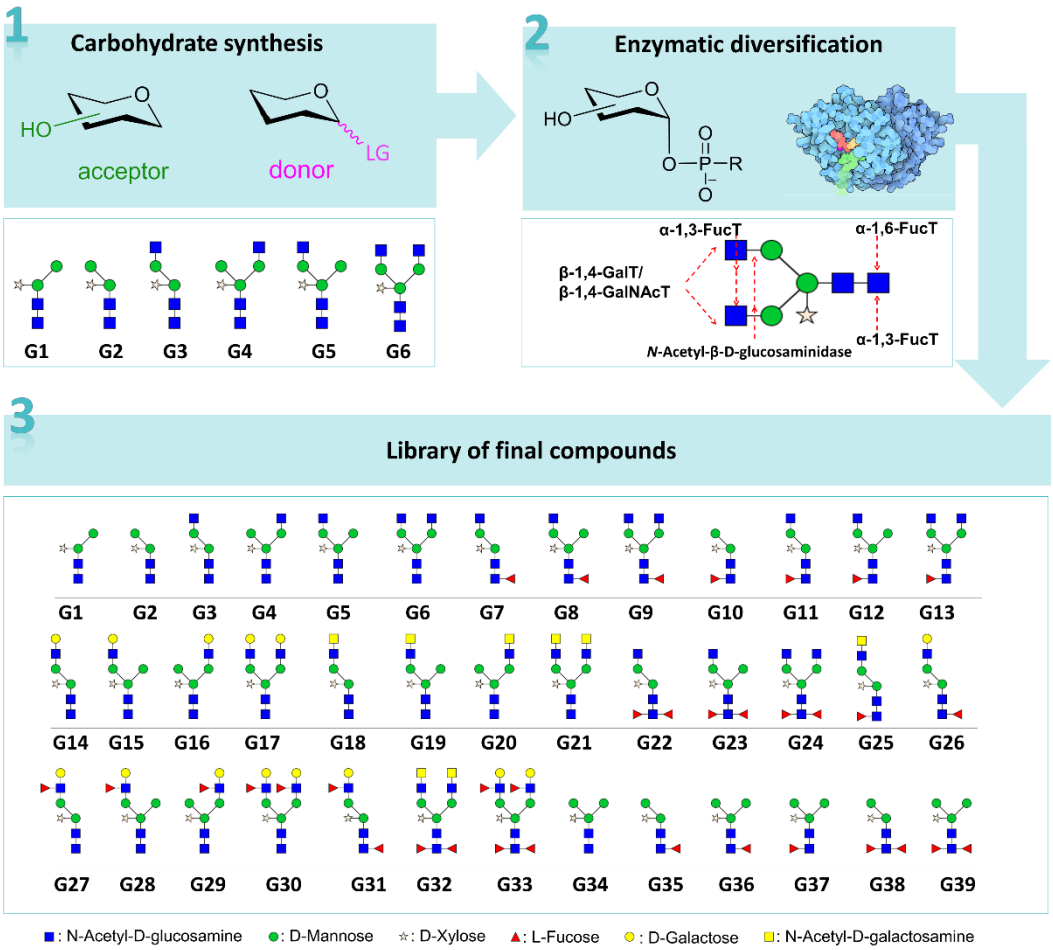
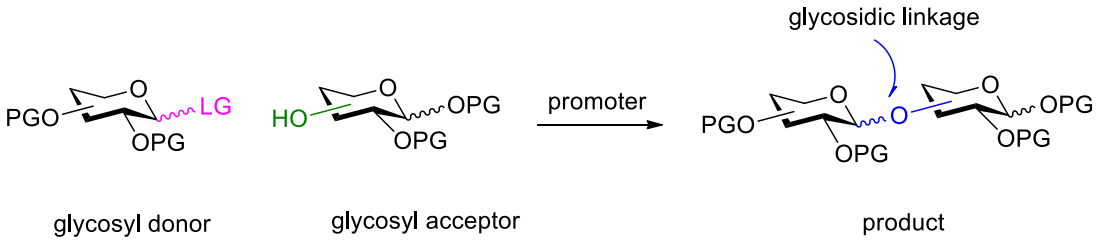


Figure 1. General strategy for the chemoenzymatic synthesis of xylosylated N-glycans; **1.** Target structures with a common xylosylated tetrasaccharide core prepared chemically; **2.** Exemplified enzymatic diversifications on the example of biantennary glycan G6; **3.** Library of final compounds obtained after chemoenzymatic synthesis.

1.1.1 Chemical synthesis

The chemical synthesis of carbohydrates requires sophisticated methodologies based on protecting group strategy and the modular assembly of carbohydrate building blocks. The formation of a glycosidic linkage between two carbohydrate units, known as glycosylation reaction, involves the coupling of activated OH-group of *glycosyl donor* with an OH-group of the *glycosyl acceptor*. Because carbohydrates contain more than one hydroxyl group, the differentiation between them is only possible by using protecting groups (PGs) masking all non-reacting hydroxyl functions except the one to be glycosylated (Scheme 1). One of the key aspects for the success in carbohydrate synthesis is a design of an orthogonal protecting group strategy. This means that the removal of a desired PG can be performed under conditions which do not affect the other set of protecting groups.

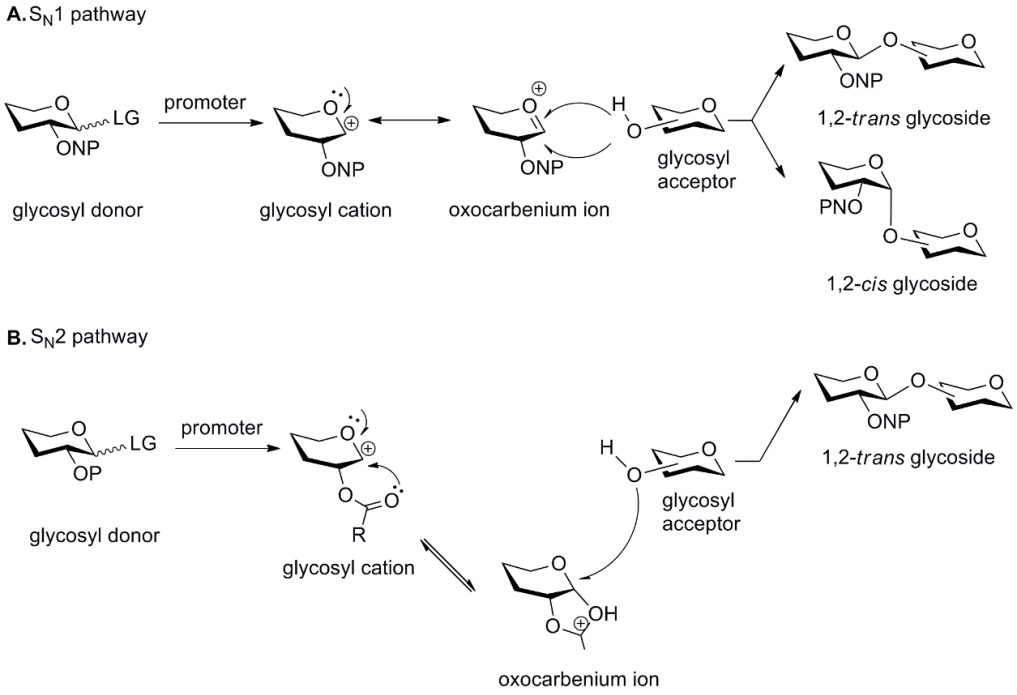


Scheme 1. Reaction of glycosylation between the glycosyl donor and acceptor catalyzed by the promoter.

In general, the protecting groups are divided into two types: permanent and temporary groups. Permanent PGs are used for the protection of functional groups which are not manipulated until the end of the synthesis; they mask the positions in which they are installed and prevent them from interfering in the synthetic steps. Permanent protecting groups are usually introduced in the first steps of synthesis and are removed at the end as for example benzyl ether. On the other hand, temporary protecting groups are much more easy and selective to remove without affecting permanent PGs and other temporary

groups. The nature of the protecting groups can have a profound effect on the overall reactivity of a carbohydrate building block and can also affect the stereochemical outcome of glycosylation reaction.

In general, glycosylation reaction involves the nucleophilic displacement of a leaving group installed at the anomeric carbon by a sugar acceptor following a S_N1 or S_N2 type pathway (Scheme 2). In most cases, the departure of this leaving group is assisted with an activator (promoter) and results on the formation of a glycosyl cation[27] that can be stabilized by resonance with the lone pairs of the adjacent oxygen forming an oxocarbenium ion.



Scheme 2. Glycosylation mechanism depending on the presence of **A.** non-participating group -ONP at C-2; **B.** participating group –OP and the different outcomes ,1,2-*trans or* 1,2-*cis* glycoside.

If the glycosyl donor carries a non-participating protecting group (Scheme 2, A) on the 2-hydroxyl group, this oxocarbenium ion could be attacked by the glycosyl acceptor from both top and bottom face of the ring following a S_N1 mechanism and allowing the possible

formation of both β and α anomers. One of the most successful ways to control the stereochemistry of the 1,2-*trans* glycosidic bond is the presence of a neighboring participating group, such as acetate or benzoate, on the 2-hydroxyl position of glycosyl donor (Scheme 2, B). In this case, the glycosyl cation is stabilized by the action of the adjacent carbonyl oxygen forming a cyclic oxonium ion, which is attacked by the nucleophile acceptor from the top face of the ring following a S_N2 mechanism. This results in the major formation of 1,2-*trans* glycosidic linkage.

Unlike 1,2-*trans* glycosides, the stereoselective formation of 1,2-*cis* glycosidic linkages, like β-mannoside requires an alternative synthetic strategy than the participating/non-participating group strategy and is a classical problem in the carbohydrates synthesis. In 1975 Lemieux reported so called *halide-assisted* or *in situ anomerization* strategy, which allows for the generation of 1,2-*cis* glycosides applying glycosyl bromide donor and the tetraethylammonium bromide as additive.[28]

The addition of halide salts results in the establishment of the quick equilibrium between the α and β bromide of the donor, without forcing of its departure, as the regular promoter does. The α-bromide is too stable to react with the acceptor, unlike the β-anomer which can be displaced by the nucleophile following the S_N2 pathway and resulting formation of the α-product. Once entire amount of the available β-bromide donor is turned into the product, new equilibrium between both anomers is established, more of the β-bromide is created and finally all the donor is used to form the 1,2-*cis* glycoside. Another typical approach of the β-mannosyl linkage formation was demonstrated by Stork and co-workers. They have described the intramolecular glycosylation (intramolecular aglycon delivery) performed on the temporarily tied glycosyl acceptor to the 2-OH group of glycosyl donor via a silicon ether tether.[29]

As mentioned before, the choice of protecting groups not only influence the stereoselectivity of the glycosylation as described above, but also can have an effect on the overall reactivity of a carbohydrate building block. Lemieux[30] and Fletcher[31,32] first linked

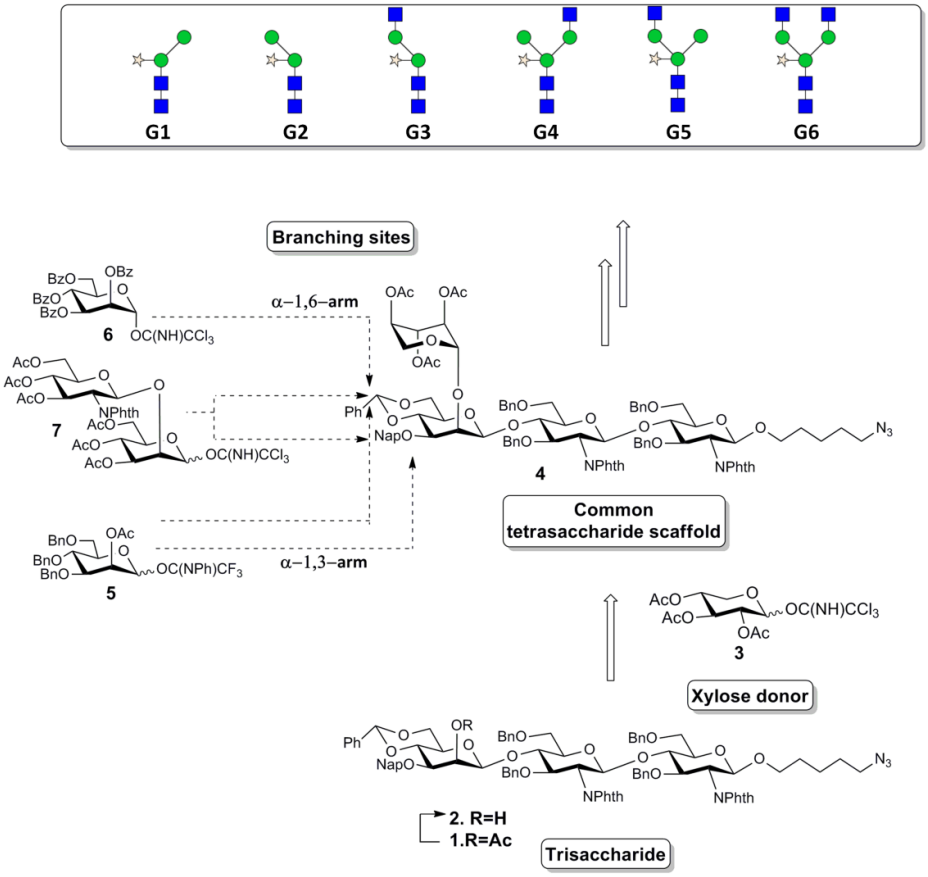
the reactivity and the selectivity of glycosyl halides and the armed/disarmed terminology was introduced. Carbohydrates substituted with strong electron-withdrawing groups like esters (-OCOR,) are called disarmed donors, they have a destabilizing effect on the oxocarbenium ion and they are less reactive that so called armed-donors. Armed-donors have as substituent electrodonating groups such as benzyl groups (-OBn) that reacts faster than disarmed donors.

During the years, development in carbohydrate synthesis was focused mostly on the design of new anomeric leaving groups and promoters. The first glycosylation was made by Michael and Fischer, who performed the reaction of anomeric position of glucose in the presence of methanol under acidic conditions.[33,34] Later on, Köenigs and Knorr[35] described the nucleophilic replacement of chlorine and bromine at the anomeric center using Ag_2CO_3 as a promoter. Subsequently Zemplen and Gerecs[36] improved the leaving-group ability by changing of promoter, into mercury (II) salt catalysts. For many decades halogen activated glycosyl donors and later hemiacetals and acetals were exclusively used as leaving groups in carbohydrate synthesis. Later, more elaborated leaving groups were introduced and nowadays, thioglycosides, O-acetimidates and fluorides are the most commonly sugar donors employed in complex carbohydrate chemistry.

1.1.2 Retrosynthetic analysis of target structures

Our synthetic strategy for the modular assembly of N-glycans is illustrated on Scheme 3, it is based on previously developed approach in our group[37] and utilizes a tetrasaccharide **4** as a starting point for further diversification into target structures (**G1**-**G6**). The synthesis of trisaccharide **1** has been described previously.[38] Trisaccharide **1** is protected by a set of orthogonal temporary protecting groups at OH-2, OH-3, OH-4 and OH-6, differentiated with acetate, 2-naphthyl methyl ether and benzylidene acetal protecting groups respectively. Moreover, the introduction of C5 azide containing linker at the anomeric position of the

first *N*-acetylglucosamine, gives us the possibility for conjugation of N-glycans on microarray surfaces and proteins.



Scheme 3. Retrosynthesis of **G1**-**G6** N-glycans.

The synthesis of N-glycans from the trisaccharidic core **1** starts with the selective deprotection of an acetate group at the axial position of terminal mannose, followed by glycosylation with xylose donor **3** to form the common tetrasaccharide scaffold. With tetrasaccharide **4** in hands, glycosylations on (**G1**, **G2**) or both (**G3**-**G6**) branching sites OH-3 and/or OH-6 will produce the desired compounds. In order to synthesize compounds **G2**-**G6**, tetrasaccharide **4** was treated with 2,3-dichloro-5,6-dicyano-*p*-benzoquinone (DDQ), to liberate OH-3. This glycosyl acceptor was subsequently substituted with donor **5** (final

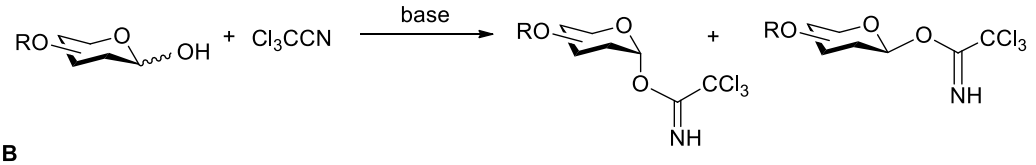
compound **G2**) or **7** (**G3**). In the next steps, the benzylidene acetal ring was cleaved and in regio- and stereoselective fashion the primary hydroxyl group OH-6 is glycosylated with donor **7** (**G4**) or **6** (**G6**). In order to synthesize compound **G1** we directly performed the reductive benzylidene acetal ring opening of tetrasaccharide **4** followed by the glycosylation with donor **5**.

1.1.3 Glycosyl imidates as donors in carbohydrate synthesis

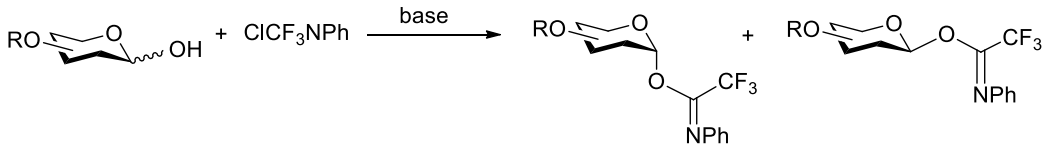
Synthesis of donors used during the generation of target structures (Scheme 3) was described previously by several groups[6,39,40] thus their detailed preparation and characterization will not be further discussed in this Chapter. However, because all of them are activated as O-imidates, we will shortly comment on the preparation and reactivity of trichloro- and trifluoroacetimidate donors.

In 1980 Schmidt and Michel introduced trichloroacetimidates (TCA) as donors[41] and they have fast become one of the most widely used glycosyl donors for the oligosaccharides synthesis.[42,43,44] Trichloroacetimidates are rapid and easy to prepare by treatment of hemiacetal with trichloroacetonitrile under basic conditions (Scheme 4, A).

A



B



Scheme 4. Preparation of *O*-imidates donors; **A.** Trichloroacetimidate; **B.** *N*-phenyl-trifluoroacetimidate.

Depending on the type of base used either the α-glycosyl (NaH/Cs_2CO_3) or β-glycosyl (K_2CO_3) trichloroacetimidates can be obtained, which while glycosylated under mild activation conditions can help to control the stereoselective outcome of the reaction following S_N2-like pathway. [42] The preparation of TCA under mild, basic conditions tolerates the use of almost all kind of protecting group patterns and glycosidic linkages in the donor molecule. Nevertheless, one of the biggest advantage of TCA over traditional glycosyl donors, like halides or thioglycosides, is the requirement of only catalytic amount of promoter for their activation, which can be achieve using Lewis acids, such as trimethylsilyl trifluoromethanesulfonate (TMSOTf) or boron trifluoride diethyletherate ($BF_3{\cdot}Et_2O$). Some years after the trichloroacetimidate synthesis, Schmidt reported the analogues, trifluoroacetimidates[45] and in 2010, Yu and Tao reviewed the use of N-phenyltrifluoroacetimidates (PTFAI) as donors in glycosylation reactions.[46] Similarly to TCA, PTFAI glycozyl donors are easy to prepare, starting from hemiacetals by treatment with N-phenyl trifluoroacetimidoyl chloride under mild, basic conditions. Compared to trichloroacetimidates, PTFAIs are reported to be less reactive and to generally require a higher amount of catalyst or higher temperatures for activation. The main advantage of PTFAIs over TCAs is the lower probability to undergo side reactions during glycosylation. When trichloroacetimidates donors are used, the trichloroacetamide liberated from the donor can react with the oxocarbenium ion and form a *N*-glycoside by-product. Particularly, this reaction takes place when the acceptor is low nucleophilic or sterically hindered. In the case of PTFAI, the lower N-basicity and the present of N-substitution help to avoid the formation of *N*-glycoside byproducts and thus the trifluoroacetimidates have become an alternative for a wide range of oligosaccharide synthesis, including preparation of β-mannosides.

1.2. Chemical synthesis of xylose containing N-glycans G1-G6

1.2.1. Assembly of β-1,2-xylosylated precursor 4

For the introduction of β-1,2-xylose moiety, we investigated the deprotection of the 2-O-acetyl group on the β-mannose moiety. Unexpectedly, standard Zemplén conditions employing sodium methoxide (NaOMe) treatment did not provide the deacetylated compound 2 and only unreacted starting material was recovered. The different trials in order to optimize this reaction are summarized in Table 1.

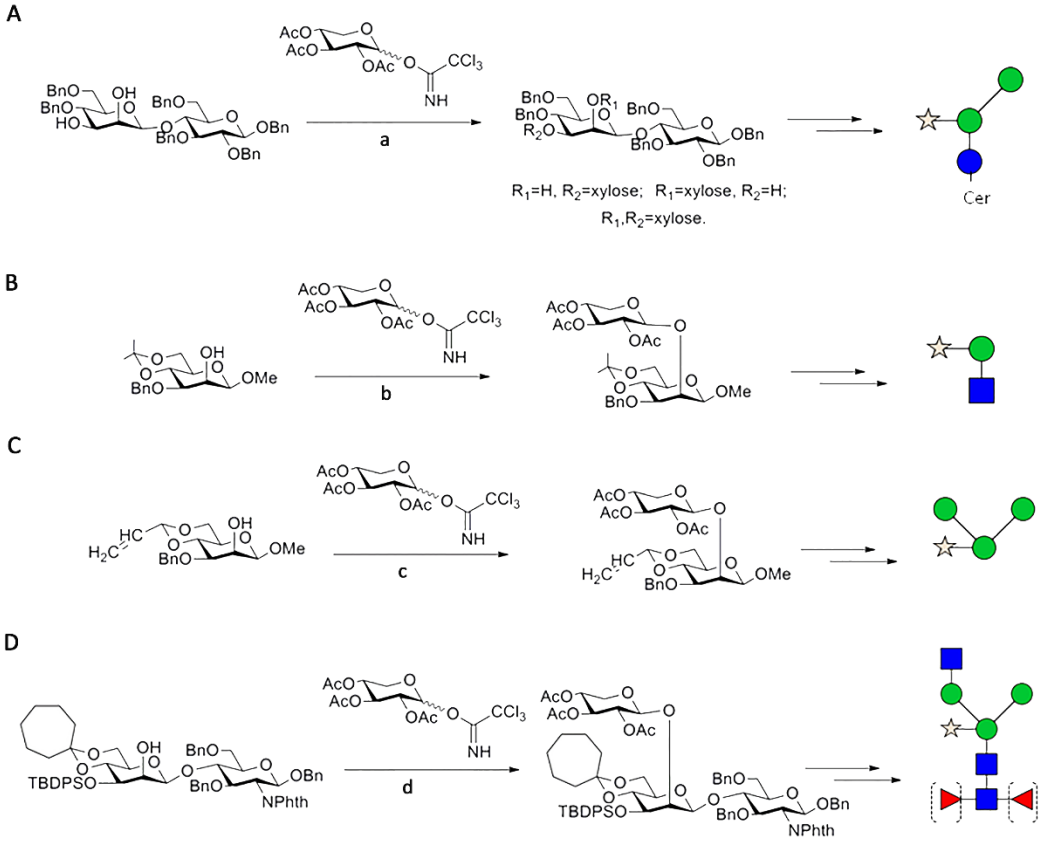
1 2

ENTRY	Base	Reagents and conditions	Comments
1	NaOMe	0.1 eq. NaOMe, MeOH:DCM (1:1)	No reaction
		1 eq NaOMe, MeOH:DCM (1:1)	SM+ product + byproduct
2	Bu_2SnO	1 eq. Bu_2SnO, DCM:MeOH (1:1)	No reaction
3	tBuOK	1 eq *t*BuOK, *t*BuOH:DCM (1:1)	No reaction
4	K_2CO_3	1eq K_2CO_3, MeOH:DCM (3:1), MW 90 °C (1 hour)	40% of product[a]
5	Cs_2CO_3	1 eq Cs_2CO_3, MeOH:DCM (3:1), MW 90 °C (1 hour)	57% of product, 17% BP, 26% SM[a]
6	KCN	3 eq. of KCN, MeOH, MW 80 °C (30 min)	50% of product[a]
7	$Mg(OMe)_2$	50 eq. $Mg(OMe)_2$ in DCM MW 80 °C (3x30 min)	82% of product[b]

Table 1. Different conditions of compound **1** 2-*O*-acetate deprotection of trisaccharide **1**. [a]Yield determine by UPLC. [b]Isolated yield.

Forcing the reaction condition by increasing the amount of sodium methoxide up to one equivalent resulted in the opening of the phthalimide groups (Table 1, entry 1). The opening of phthalimide groups under harsh basic conditions was confirmed by ultra-performance liquid chromatography-mass spectrometry (UPLC-MS). A new peak was detected corresponding to the m/z of the desired product plus 18 suggested addition of the water to the molecule. After several trials, using different bases, the 2-O-deacetylation was achieved by treatment of **1** with magnesium methoxide (Table 1, entry 7) under microwave heating at 80°C in three cycles of thirty minutes. The deacetylated compound **2** was obtained with a very good yield (86%) and no byproduct formation was observed. Next the determination of optimal conditions for the xylosylation on trisaccharide acceptor **2** was performed.

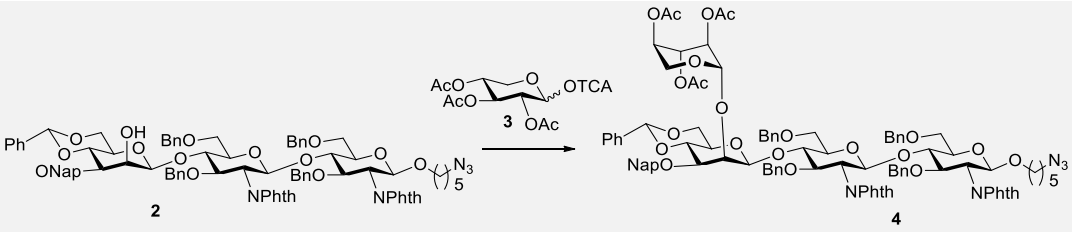
Only few examples describing the xylosylation of carbohydrates have been reported. One of the first was published in 1990 by Ogawa and coworkers when a series of synthetic mollusks glycosphingolipids were prepared[6] (Scheme 6, A). Few years later, Kerekgyarto *et al.*[47](Scheme 6, B) and van der Ven *et al.*[36] (Scheme 6, C) synthesized simple, protected tri- and tetrasaccharide N-glycan fragments. More recently Nakano *et al.* focused on the preparation of xylose containing N-glycans, common for the *S. mansoni* and *S. japonicum* parasites. The xylosylation was performed on the orthogonally protected disaccharide Manβ-1,4-GlcNAc (Scheme 6, D), showing a similar protecting group scheme as the trisaccharide acceptor **2** we were going to employ in our synthesis. The glycosylation was reported using a trichloroacetimidate xylose donor, 2 equivalents of TMSOTf to give the overall trisaccharide with a good yield of 67%.



Scheme 6. Previous examples of xylosylation reaction; **A.** Synthesis of glycosyl ceramides. Reagents and conditions: **a)** 0.3 eq. $BF_3{\cdot}Et_2O$, 1.6 eq. xylose donor, -20°C to room temp. mixture of three compounds (28%), (26%), (20%); **B.** Synthesis of xylosylated carbohydrates from glycoproteins of plant and animal origin. Reagents and conditions: **b)** 1.2 eq. TMSOTf, 1.2 eq. xylose donor, -40°C, yield 79%; **C.** Synthesis of xylosylated carbohydrates from N-glycoproteins Reagents and conditions: **c)** 0.25 eq. TMSOTf, 1.3 eq. xylose donor, -40°C, yield 89%; **D.** Synthesis of complex-type glycans derived from parasitic helminths Reagents and conditions: **d)** 2 eq. TMSOTf, 3 eq. xylose donor, -40°C, yield 67%.

Based on these previous results, we chose the peracetylated xylose trichloroacetimidate **3** and trimethylsilyl trifluoromethanesulfonate (TMSOTf) as promoter for the first trials of xylosylation. We investigated several reaction conditions (Table 2) and experienced that

this particular glycosylation requires strong acidic conditions in order to achieve the formation of the desired product and to avoid the formation of orthoester **8** which was the main obstacle of this reaction (Figure 2).



ENTRY	Donor [eq]	TMSOTf [eq]	Temperature	Concentration [M]	Product (yield)*
1	1.5	0.1	0 °C to RT	0.02	**4** (40%)
2	2.2	0.1	0 °C to RT	0.02	**4** (50%)
3	3	0.5	-40 °C	0.01	**8** (70%)
4	3	0.1	-40 °C to RT	0.07	**8** (85%)
5	3	0.4	-40 °C to RT	0.07	**4** (75%)

Table 2. Reaction conditions for the xylosylation of the trisaccharide **2** (*conversion determined by UPLC).

When xylosylation was performed at room temperature with 1.5 or 2 equivalents of donor and catalytic amount of promoter (Table 2, entries 1 and 2) the starting material was converted into tetrasaccharide **4** in 40-50% of yield, estimated on the basis of the UPLC chromatogram recorded after quenching of the reaction with triethylamine.

By lowering the reaction temperature and simultaneously increasing the amount of promoter and donor added (Table 2, entry 3), we observed the formation of the orthoester **8**, which sluggishly rearranges into acetylated or hydrolyzed starting material. The formation of orthoester **8,** was monitored by TLC and could be distinguished from the

reaction product **4** by a different retention time in UPLC-MS analysis (Figure 2). In order to avoid the formation of **8,** we have further increased the concentration of the reaction (0.07 M) and subsequently warmed it up to room temperature after addition of promoter. Unfortunately, again orthoester **8** was formed with very high yield (Table 2, entry 4). Finally, we kept previous conditions, but using an increased amount of promoter adding up to 0.4 equivalents (Table 2, entry 5). Under these conditions, the starting material was converted into the desired product with 75% of yield as determined by UPLC-MS. Unfortunately, the high amount of TMSOTf promoter favored the formation of the trimethylsilylated trisaccharide (see Figure 2B) byproduct, which complicated the isolation of the desired tetrasaccharide **4**. After the purification by flash column chromatography, the final compound was isolated in a 65% of yield.

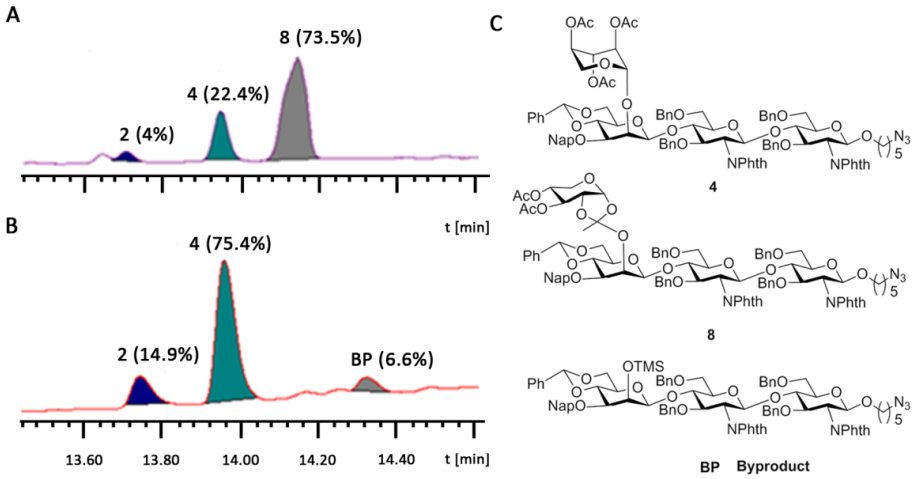


Figure 2. UPLC-chromatogram of the reaction crude after quenching with Et_3N under **A**. not optimized conditions with an orthoester **8** as a main product (Table 2, entry 3)**; B.** optimized conditions (Table 2, entry5).

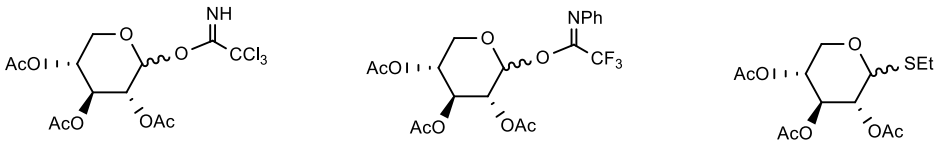
To avoid formation of trimethylsilyl substituted acceptor, we investigated alternative promoters such as trifluoromethanesulfonic acid (TfOH), copper (II)

trifluoromethanesulfonate $Cu(OTf)_2$ and ytterbium(III) trifluoromethanesulfonate $Yb(OTf)_3$ (Table 3).

Catalysts	Donor	Catalyst [eq.]	Temperature [°C]	4 [%]
TfOH	**3**	0.4	0 to r.t.	73
$Cu(OTf)_2$	**3**	2.5	0 to r.t.	55
$Yb(OTf)_3$	**3**	3	0 to r.t.	41
TfOH	**9**	0.4	0 to r.t.	50
NIS/TfOH	**10**	3.2/0.1	r.t.	60

Table 3. Different catalysts and donors used during xylosylation reaction.

We also investigated different donors such as thioglycoside **10** or *N*-phenyltrifluoroacetimidate **9** (Figure 3) but unfortunately, product formation was not improved and the xylose trichloroacetimidate was used in the reaction (Table 3).



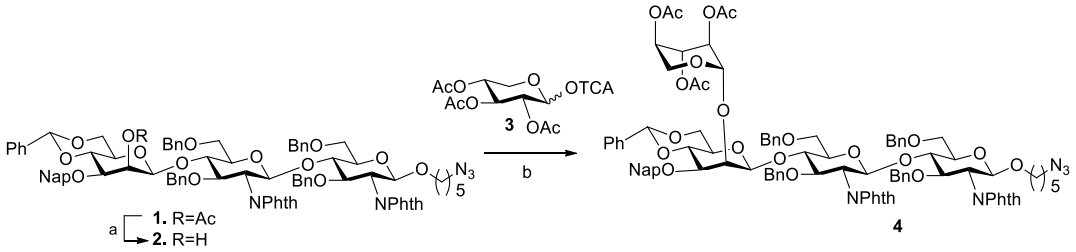
3 **9** **10**

Figure 3. Three donors tested in the xylosylation of trisaccharide.

Under previously optimized conditions (Table 2, entry 5) and using TfOH as promoter the xylosylated tetrasacchride **4** was obtained in 73% yield after purification.

Finally, we could establish the optimal conditions for the synthesis of tetrasaccharide **4** as summarized in Scheme 7. The removal of 2-OH acetate was achieved in high yield (86%) by using $Mg(OMe)_2$ under microwave irradiation. For the introduction β-1,2-xylose, we

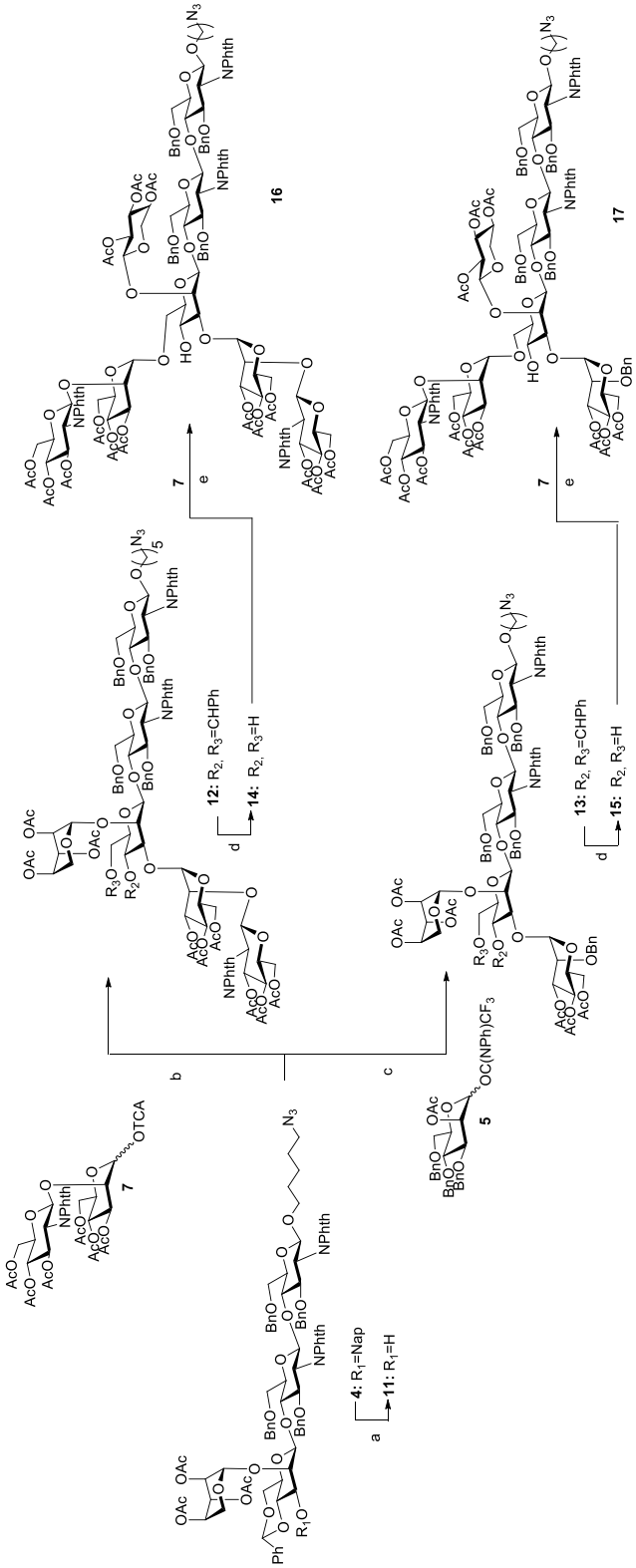
employed xylose trichloroacetimidate **3** activated with TfOH as acid catalyst to form the tetrasaccharide **4** in high yield (73%).



Scheme 7. Synthesis of tetrasaccharide **4**; Reagents and conditions: a) $Mg(OMe)_2$, DCM, MeOH, 80°C, MW, 86%; b) TfOH, DCM, 0°C to r.t, 73%.

1.2.2 Assembly of N-glycans. Glycosylation of 4 with glycosidic donors

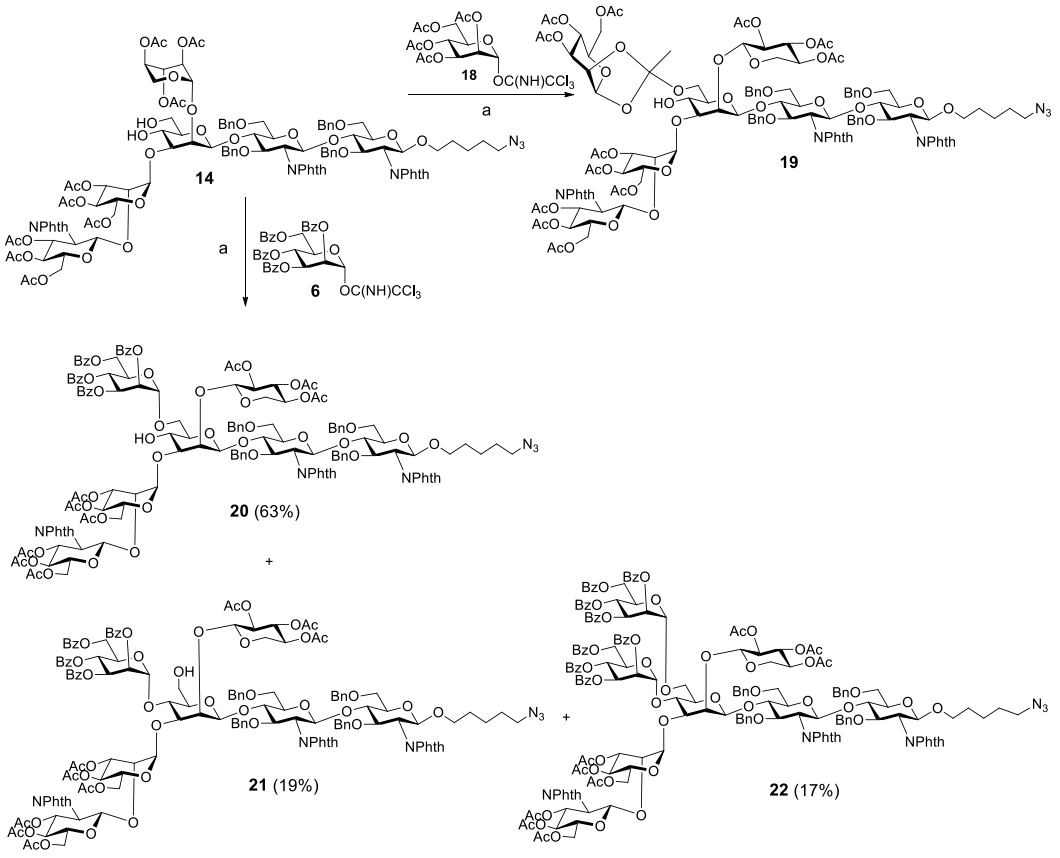
With the tetrasaccharide **4** in hands, we started to investigate glycosylation reactions with different glycosidic donors in one (**G1**, **G2**) or both (**G3**-**G6**) branching sites OH-3 and/or OH-6 to produce the final xylosylated N-glycans. In order to deprotect OH-3 position, tetrasaccharide **4** was treated with 2,3-dichloro-5,6-dicyano-*p*-benzoquinone (DDQ) to remove 2-naphthylmethyl ether (Nap) group[48,49,50] (Scheme 8). The desired tetrasaccharide acceptor **11** was obtained in modest yields between 50 to 60% while recovering part of the starting material. It was observed that increasing reaction time results in the loose of benzyl groups which hampered the isolation of the final product. Therefore, the optimal conditions were established by quenching the reaction before complete deprotection of Nap group to both maintain a reasonable yield and maximize the recovery of unmodified starting material. The tetrasaccharide acceptor **11** was subsequently glycosylated with disaccharide **7** or mannose donor **5** under TMSOTf activation to produce both hexasaccharide **12** (82%) and pentasaccharide **13** (94%) in excellent yields (Scheme 8). Removal of benzylidene acetal protecting group to liberate OH-4 and OH-6 was achieved by acid hydrolysis catalyzed by BF_3OEt_2 in the presence of ethanethiol to afford the corresponding diol acceptors **14** (79%) and **15** (86%) in good yields.



Scheme 8 Chemical synthesis of xylosylated glycans **12, 13, 16** and **17.** Reagents and conditions: **a)** DDQ, DCM:MeOH (4:1) 60%; **b)** TMSOTf, DCM, 4Å MS, -40ºC, 82%; **c)** TMSOTf, DCM, 4Å MS, -20 ºC, 94%; **d)** $BF_3{\cdot}Et_2O$, EtSH, DCM, **14** (79%) and **15** (86%); e) TMSOTf, DCM, 4Å MS, -45ºC, **16** (73%) **17** (73%).

The glycosyl acceptors **14** and **15** were reacted with disaccharide donor **7** under TMSOTf activation to produce octasaccharide **16** (73%) and heptasaccharide **17** (73%) with complete stereo- and regio-selectivity in good yields (Scheme 8, e). A strict control of low temperature as well as high dilution of reaction mixture (250 mL/mmol) helped preventing the formation of regioisomers.[39]

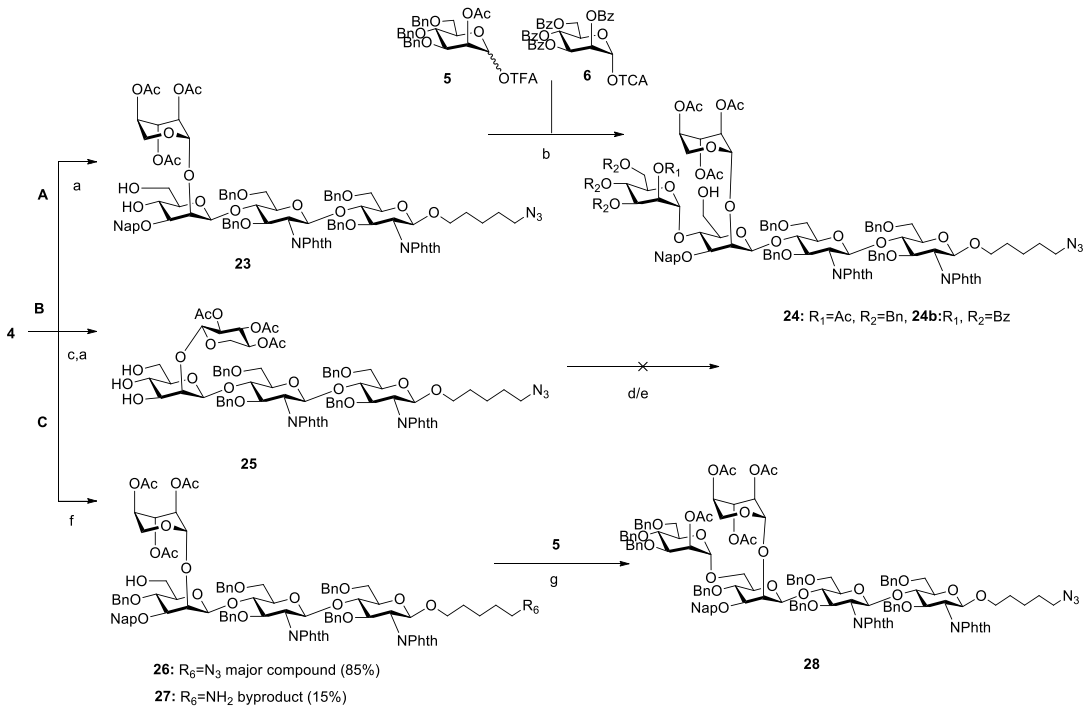
The initial attempt for the preparation of hybrid N-glycan **G5** employed a similar strategy as for the biantennary compound **16**, starting with acceptor **14** and coupling with peracetylated mannose trichloroacetimidate donor **18** (Scheme 9).



Scheme 9. Synthesis of the heptasaccharide **G5**; Reagents and conditions; a) TMSOTf, DCM, 4Å MS, -30°C. Yields determined by UPLC before chromatographic purification of desired compound **20**.

Unfortunately, during this reaction an unusually stable orthoester **19** was formed as major reaction product, which after rearrangement under acidic conditions afforded a complex mixture of both OH-4 and OH-6 substituted regioisomers and unreacted starting material **14** (ratio 1:1:1, as determined by UPLC). In order to avoid the orthoester formation, we synthesized the mannose donor **6**, fully substituted with benzyl ester groups that are bulkier than acetates. Glycosylation with acceptor **14** was conducted under activation with TMSOTf and even though we could observe the presence of a regioisomer OH-4 **21** and a bis-substituted by-product **22** (Scheme 9, b) we could isolate the desired heptasaccharide **20** in a moderate isolated yield (50%).

Finally, we attempted the synthesis of the truncated N-glycan **G1**. We assumed that glycosylation of acceptor **23** with armed mannose donor **5** would preferable lead to substitution of the less sterically hindered and more reactive primary OH-6 position providing the desired pentasaccharide in a regioselective manner (Scheme 10, route A, donor **5**). Unfortunately, detailed analysis of HSQC and comparison of ^{13}C and ^{1}H chemical shift values at C-4 and C-6 attachment sites between starting material and product revealed formation mainly of the undesired regioisomer OH-4 **24** (Figure 4). We observed a downfield shift for the C-4 ^{13}C resonance from 66 ppm in **23** to 71 ppm for compound **24.** Likewise, a downfield shift was observed for the H-4 ^{1}H resonance from 3.7 ppm to 4.1 ppm in **24**[51] while the C-6 ^{13}C resonance remained unaltered between starting material **23** and product **24**. Both were in the range of 60-65 ppm, which is a characteristic value for unsubstituted derivatives (Figure 4). A similar result was obtained when disarmed donor **6** was employed (Scheme 10, route A, **24b**).



Scheme 10. Synthesis of compound **28**. Reaction and conditions: a) $BF_3{\cdot}Et_2O$, EtSH in DCM; b) TMSOTf, 4Å MS in DCM; c) DDQ, DCM:MeOH (4:1); d) butane-2,3-dione, $(MeO)_3CH$, BF_3Et_2O or CSA, MeOH; e) 2,2,3,3-tetramethoxybutane, $(MeO)_3CH$, CSA, MeOH; f) $BH_3{\cdot}THF$, n-Bu_2BOTf, 0°C, 2h; g) 15% TMSOTf, 4Å MS in DCM, -20°C.

Fairbanks and coworkers observed similar phenomena on the set of a 4,6-diol disaccharide acceptors that were glycosylated with a different glycosidic donors.[52] They demonstrated that the regioselectivity of the glycosylation reaction is strongly dependent on the functionalization of the neighboring OH-3 group of the acceptor and that the OH-4 regioisomer is preferably formed when OH-3 is protected with benzyl ether and when donor possesses an ester participating group at C-2. Similar to our findings, they also demonstrated the preferential formation of OH-6 regioisomer when the OH-3 position had been previously glycosylated. Later our group have reported further experimental observation regarding the role of the glycosyl donor on the regiochemical outcome of the glycosylation reaction performed on the 4,6-diol system of OH-3 glycosylated acceptor.[53]

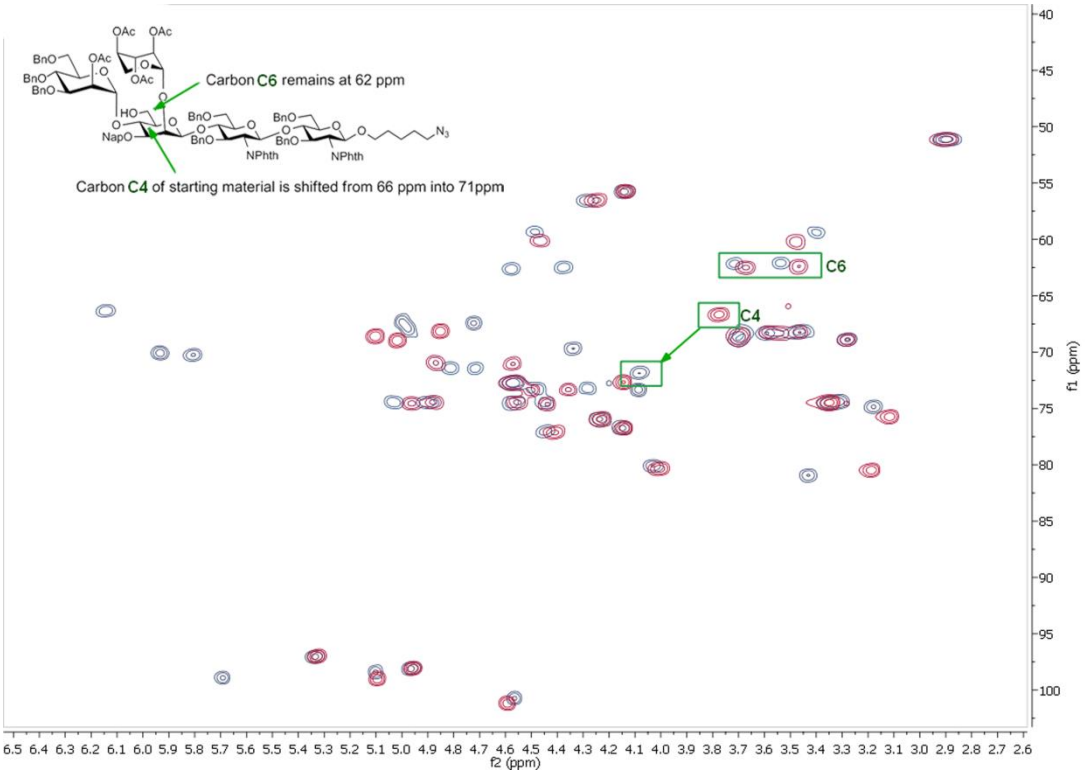


Figure 4. HSQC spectra comparison of starting material **23** (red) and product **24** (blue).

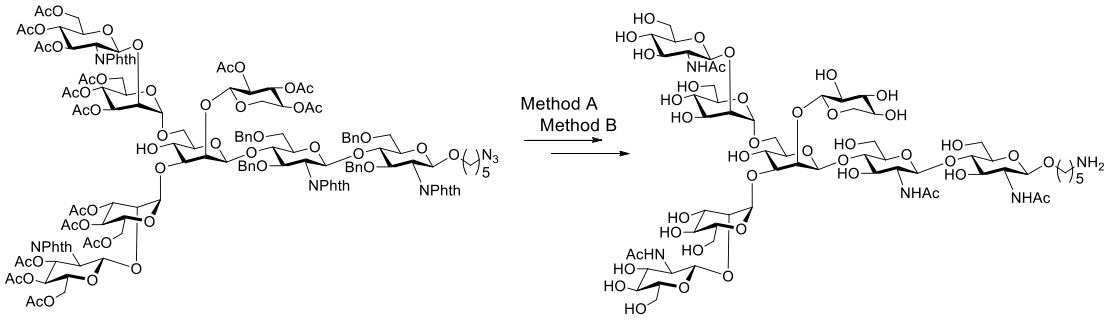
In the next attempt (Scheme 10, route B) we removed Nap and benzylidene acetal protecting groups from tetrasaccharide **4** and we tried the selective protection of the two equatorial hydroxyl groups OH-3 and OH-4 of compound **25** *via* a butane-2,3-diacetal. Even though these reaction conditions has been applicable in the synthesis of several 6-substituted sugars[54,55], in our case we the conversion of triol into protected sugar was not achieved.

Finally, we could selectively perform the reductive ring opening of the benzylidene acetal in tetrasaccharide **4** employing borane tetrahydrofuran complex (BH_3·THF) to yield acceptor with primary alcohol OH-6 **26** (Scheme 10, route C). When TMSOTf or $Cu(OTf)_2$ were used as a catalyst in this reduction, conversion into desired product was low (20-30%). In these reactions, different byproducts lacking acetates or with a reduced azide group were formed, as monitored by UPLC-MS. Changing the catalyst into *n*-Bu_2BOTf, provided free OH-

6 acceptor **26** with excellent selectivity[56,57] and good isolated yield (65%). Nevertheless, the formation of a byproduct **27** with the azido group reduced to a primary amine could not be completely abolished (15% as detected by UPLC-MS). After purification, tetrasaccharide acceptor **26** was reacted with mannose donor **5** under TMSOTf activation to provide desired pentasaccharide **28** with a moderate yield (52%).

1.2.3 Final deprotection of N-glycans

The N-glycan precursors **12**, **13**, **16, 17, 20** and **28** were deprotected to yield the target N-glycans **G1-G6** either in a three or four step sequence (Scheme 11, Table 4).



Scheme 11. Two ways of deprotection of final compounds on the example of biantennary compound **G6**; Method **A.** Three steps deprotection. Reagents and conditions; **a:** $NH_2CH_2CH_2NH_2$:*n*BuOH, MW (120ºC, 3x30'); **b)** Ac_2O, pyridine.; **c)** Na, NH_3(liquid), -78ºC; Method **B.** Four steps deprotection Reagents and conditions: **a, b)**, **d)** NaOMe, MeOH; **e)** H-cube®, 10% Pd/C CatCart®, MeOH, 0.1% TFA.

In both cases, we initially performed the aminolysis of phtalimide protecting groups by treatment with ethanolamine in *n*BuOH under heating, followed by acetylation of free amino and hydroxyl groups generated after aminolysis with acetic anhydride in pyridine. In the three steps deprotection, we have subsequently removed O-acetates, O-benzyl groups

and reduced the azido group of the linker to a primary amine in a single step applying Birch reaction using metallic sodium in liquid ammonia (Scheme 11, method A).

For other glycans the dissolving metal conditions, led to the formation of a side product with a mass loss 18Da compared to the target compounds, presumably formed by the elimination of an acetate group. This side reaction was avoided by deprotection of O-acetate groups by methanolysis using sodium methoxide, followed by the removal of remaining benzyl ether and the reduction of azido group by palladium catalyzed hydrogenation (Scheme 11, method B). Following this procedure, the fully deprotected target glycans **G1**, **G3, G5**, and **G6**, were obtained in good yields (Table 4).

Entry	N-glycan	Precursor	Method	Total yield [%]
1	**G1**	**28**	B	49
2	**G2**	**13**	A	69
3	**G3**	**12**	B	65
4	**G4**	**17**	A	56
5	**G5**	**20**	B	58
6	**G6**	**16**	B	69

Table 4. Deprotection of glycans followed by two strategies, A and B.

The final compounds were easily purified using reverse phase C18 cartridges. With **G1-G6** in hand we focused on the preparation of a library of core-xylosylated compounds by enzymatic diversification employing recombinant glycosyltransferases.

1.3. Conformational changes of xylose moiety in synthetic N-glycans

During the synthesis of xylose donor **3** and tetrasaccharide **4**, we have encountered an atypical conformational behavior of D-xylose compared to other sugars. In this section we will discuss the conformational changes of the xylose moiety observed during the synthesis of β-1,2-xylosylated N-glycans in detail.

D-xylose is an aldopentaose analogue of the aldohexose D-glucose that lack the hydroxymethyl group at position C5. In solution, D-xylose exists predominantly in the $^4C_1$ chair pyranose form as an α- or β- anomer in a ratio of 35 to 65% respectively.[58] Interestingly, the derivatives of β-xylopyranose can also adopt $^1C_4$ chair conformation (Figure 5) depending on the nature of the substituents.[59]

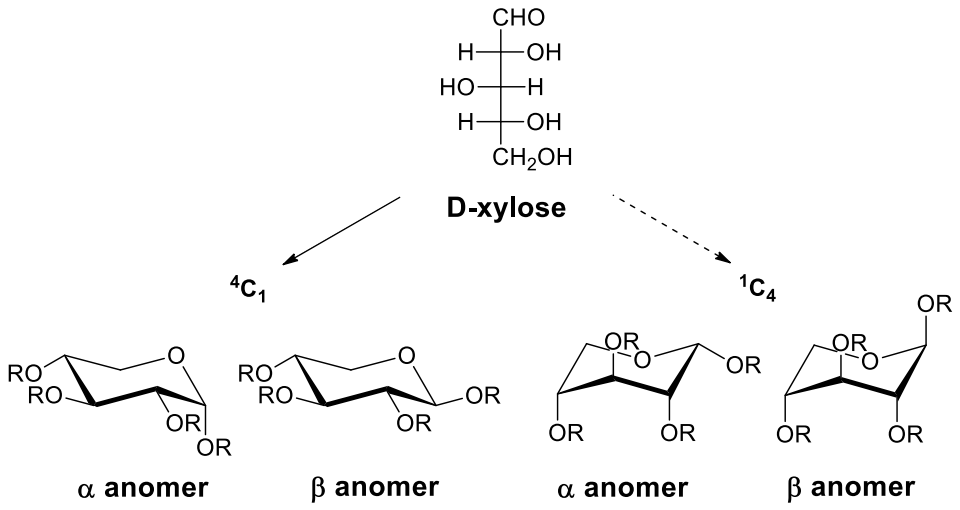


Figure 5. Open and cyclic form of D-xylose, with $^4C_1$ and $^1C_4$ chair conformations present as an α- or β- anomers.

We have therefore closely investigated the structure of trichloroacetimidate xylose donor **3** by Nuclear Magnetic Resonance (NMR), in order to establish which anomer (α- or β-) and in which conformation ($^4C_1$ or $^1C_4$) was favored.

The α- and β- anomers of xylose have in principle distinguishable NMR spectra, vicinal proton–proton coupling constants (J_{H1-H2}) in NMR depend on the dihedral angle between them and this is described by the Karplus equation (Figure 6).[60] It states that vicinal proton-proton coupling constant are maximal for protons with 0° and 180° dihedral angles and that coupling will be minimal for protons that are 90° from each other. In general, coupling constants will be larger for the protons in axial-axial orientation, followed by an axial-equatorial, equatorial-axial and with lowest couplings constants values for equatorial-equatorial protons positions (Figure 6).

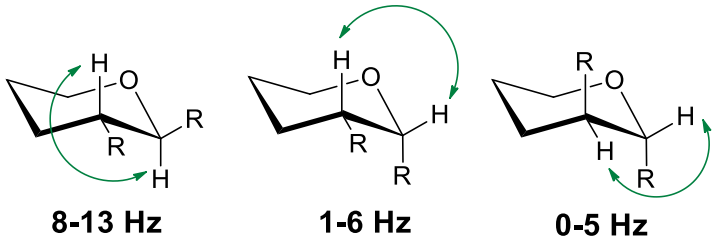


Figure 6. Schematic of coupling constants in nuclear magnetic resonance following Karplus equation.

In xylose pyranose ring, the anomeric configuration is assigned based on the basic ^{1}H NMR spectra with coupling constant, values between $^{3}J_{H1-H2}$= 7–9 Hz are typical for the diaxial protons coupling and $^{3}J_{H1-H2}$= 2–4 Hz indicative of the equatorial–axial coupling. Beside the three bond proton-proton H1-H2 coupling constant, the anomeric configuration can be also determined based on the single-bond ^{1}H–^{13}C coupling constant.[61] This information can be easily obtained from an undecoupled ^{1}H–^{13}C heteronuclear single quantum correlation spectrum (HSQC), where the value of coupling constants $^{1}J_{C1-H1}$ above or equal to 170Hz indicates an equatorial proton at C-1, while $^{1}J_{C1-H1}$ 160 Hz indicates an axial orientation of the anomeric proton.

The evaluation of ^{1}H-NMR spectrum of 2,3,4-tri-*O*-acetyl-xylopyranosyl trichloroacetimidate **3** revealed the presence of two species (**A** and **B**) with coupling constants of $^{3}J_{H1-H2}$ = 3.6 and $^{3}J_{H1-H2}$ = 4.4 Hz for anomeric protons (**A**H-1, **B**H-1), respectively (Figure 7). Taking into

account the values of $^{3}J_{H4\text{-}H5'}$ and $^{3}J_{H4\text{-}H5''}$ we assumed, that xylose donor **B** adopted the standard $^{4}C_{1}$ chair conformation, as α anomer. This finding was confirmed by the HSQC experiment coupling constants $^{1}J_{H1\text{-}C1}$ value 174 Hz, which is indicative of axial glycosides.

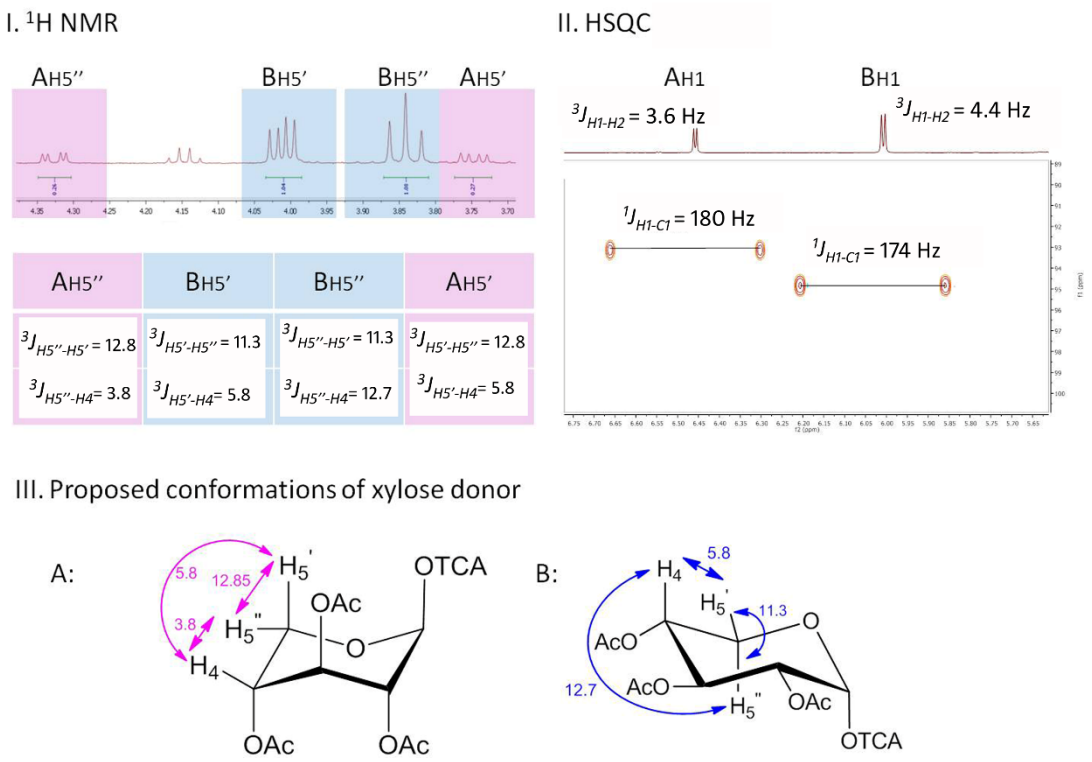


Figure 7. NMR studies of the conformations of 2,3,4-tri-*O*-acetyl-xylopyranosyl trichloroacetimidate **3**, showing two species **A** and **B**; **I.** Fragment of the ^{1}H NMR spectrum with peaks corresponding to an axial and equatorial H5 protons of compounds **A** and **B**; **II.** undecoupled 1H-13C HSQC spectrum showing coupling constants between H1 at the anomeric carbon of **A** and **B**; **III.** Ilustration of β 1C4 (A) and α 4C1 (B) conformation of xylose donor **3**.

On the other hand, xylose donor **A** revealed smaller values of coupling constants than donor **B** in the entire molecule, with a value of $^{3}J_{H4\text{-}H5''}$=3.8 Hz, which could be explained by the equatorial-equatorial orientation of vicinal protons H-4 and H-5'' present in the inverted chair conformation $^{1}C_{4}$. Moreover, coupling constant $^{3}J_{H1\text{-}H2}$ =3.6 Hz is inconsistent

with the β-anomer of the $^4C_1$ D-xylose conformation but is in a good agreement with the β anomer of a $^1C_4$ chair conformation, where vicinal protons H-1 and H-2 have equatorial-equatorial orientation. The observation can be also supported by the high coupling constant $^1J_{H1-C1}$=180 Hz which is characteristic again for axial glycosides. The ratio between both conformations α and β was 4:1.

These findings were in agreement with data reported previously in the literature. Durette and Horton investigated the conformational equilibria of peracetylated and perbenzoylated D-xylose halides.[62] They demonstrated that α anomer of both D-xylose chloride and bromide exist mainly in the $^4C_1$ form while the β anomer exist in the inverted chair conformation $^1C_4$. Later, Kerekgyarto[47] and Crich[63] reported the conformational changes of xylose moiety, which always produce deviations in *J* values between the protons on the xylose skeleton. These conformational changes were explained by a possible compensation of sterically unfavored axial oriented *O*-acetyl groups destabilizing effect by the favorable dipolar interactions in the anomeric effect.[64]

After having determined the possible conformations of xylose donor **3**, we investigated the preferential conformation of the xylose moiety on the tetrasaccharide **4** (Figure 8). We expected the glycosylation of xylose trichloroacetimidate **3** with trisaccharide **2** to produce a single β-1,2-xylose glycoside due to the O-2 acetyl neighboring group participation in donor **3**. In the analysis of HSQC experiment of tetrasaccharide **4**, we could observe presence of anomeric carbon C1 at δ 99.4 ppm, with a coupling constant of $^1J_{C1-H1}$ =171 Hz characteristic for axial glycosides. Due to the overlapping of signals in ^{1}H NMR spectra of tetrasaccharide **4**, additional NMR experiments were required to determine all coupling constants between the different protons in the xylose moiety. We employed monodimensional-TOCSY (Total Correlation Spectroscopy) experiments, in which anomeric proton of xylose is irradiated to produce a ^{1}H-NMR spectra of the xylose unit (Figure 8). With this experiment, we were able to measure the coupling constants between H-1 and H-2 ($^3J_{H1-H2}$=3.6 Hz) as well as between H-4 and H-5 ($^3J_{H4-H5'}$=3.68 Hz and $^3J_{H4-H5''}$=4.1 Hz) which are consistent with H-4 and H-5 protons adopting a equatorial-equatorial conformation.

Based on these findings, we could assign a $^{1}C_{4}$ chair conformation for the β-1,2-xylose moiety.

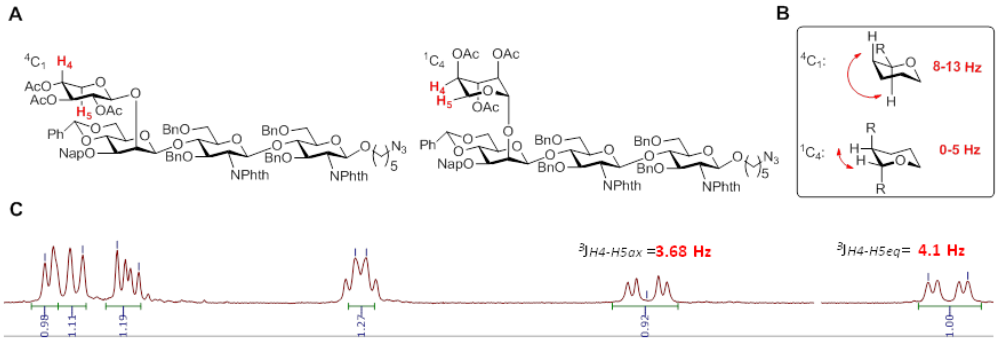


Figure 8. Analysis of the xylose conformation in the tetrasaccharide **4; A.** Possible β-conformations of xylose in tetrasaccharide; **B.** Values of $^{3}J_{H5\text{-}H4}$ coupling constants predicted based on Karplus equation; **C.** ^{1}H-^{1}H TOCSY (Total Correlated Spectroscopy) spectrum of tetrasaccharide **4** showing J values in agreement with a 1C4 conformation of β-xylose.

During the synthesis of N-glycans **G1-G6**, we observed protecting group dependent conformational changes for the β-D-xylose residue in various intermediates, most of which have been assigned and are listed in the Table 5. Generally, removal of the benzylidene acetal on the β-1,4-core mannose moiety of protected N-glycans results in the flipping of the xylose residue into the $^{4}C_{1}$ configuration (**15**) persisted after glycosylation of position OH-6 (**17**). Surprisingly, the reductive opening of the benzylidene acetal and the presence of benzyl ethers in the position OH-4 (**20**), showed no influence on xylose conformation maintained the conformation adopted in tetrasaccharide (**4**), even after glycosylation of position OH-6 (**28**). Importantly, after final deprotection of N-glycans, β-D-xylose residue always adopts a "standard" $^{4}C_{1}$ configuration, with $^{3}J_{H1\text{-}H2}$ values between 7.4-7.6 Hz.

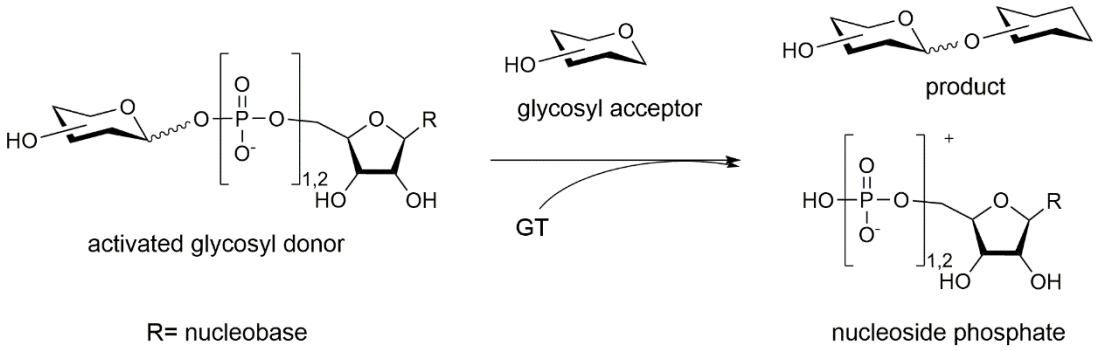
Protected compounds (NMR data, $CDCl_3$)										
Compound		$^3J_{H1-H2}$	$^3J_{H2-H3}$	$^3J_{H3-H4}$	$^3J_{H4-H5a}$	$^3J_{H4-H5b}$	$^3J_{H5a-H5b}$	$^1J_{C1-H1}$	δC1	Conformation
3	**A**	4.4	5.9	5.9	3.8	5.8	12.86	180	93.1	$^1C_4$
	B	3.6	9.9	9.9	5.9	12.7	11.3	174	94.8	$^4C_1$
4		3.63	5.45	-	3.7	4.1	13	171	99.4	$^1C_4$
11		6.7	8.4	8.4	-	-	-	168	100.9	$^4C_1$
12		-	-	-	4.2	-	12.4	171	99.6	$^1C_4$
13		4.4	6.6	6.6	3.9	5.7	12.6	171	99.7	$^1C_4$
15		6.6	8.4	8.4	5.1	8.5	11.8	168	99.5	$^4C_1$
26		3.35	5.06	5.06	-	-	-	171	98.9	$^1C_4$

Table 5. Coupling constants (*J*, Hz) between the skeleton protons of the xylose residue and the $^1J_{C1H1}$ values in the protected and deprotected synthesized compounds.

1.4 Glycosyltransferases. Introduction

In the second part of this Chapter the expression and use of recombinant glycosyltransferases for the enzymatic diversification of the synthetic compounds **G1**-**G6** will be described.

Glycosyltransferases (GTs) (EC 2.4.1.-) are the family of carbohydrate-processing enzymes which catalyze glycosidic bond formation by transferring of a sugar moiety from an activated donor to an acceptor substrate (Scheme 13).



Scheme 13. Representation of the glycosyltransferases mechanism of action on the example of nucleotide monosphosphate NMP/ nucleotide diphosphate NDP activated glycosyl donor.

The majority of glycosyl donors are activated as a nucleoside mono- or di- phosphates (Leloir donors), however, the mono phosphates (non-Leloir donor) substituted donors can be also found (Figure 9). Non-activated sugars in form of sucrose or starch are also utilized by some GTs, however their action is limited to a use of a glucose or fructose donors.

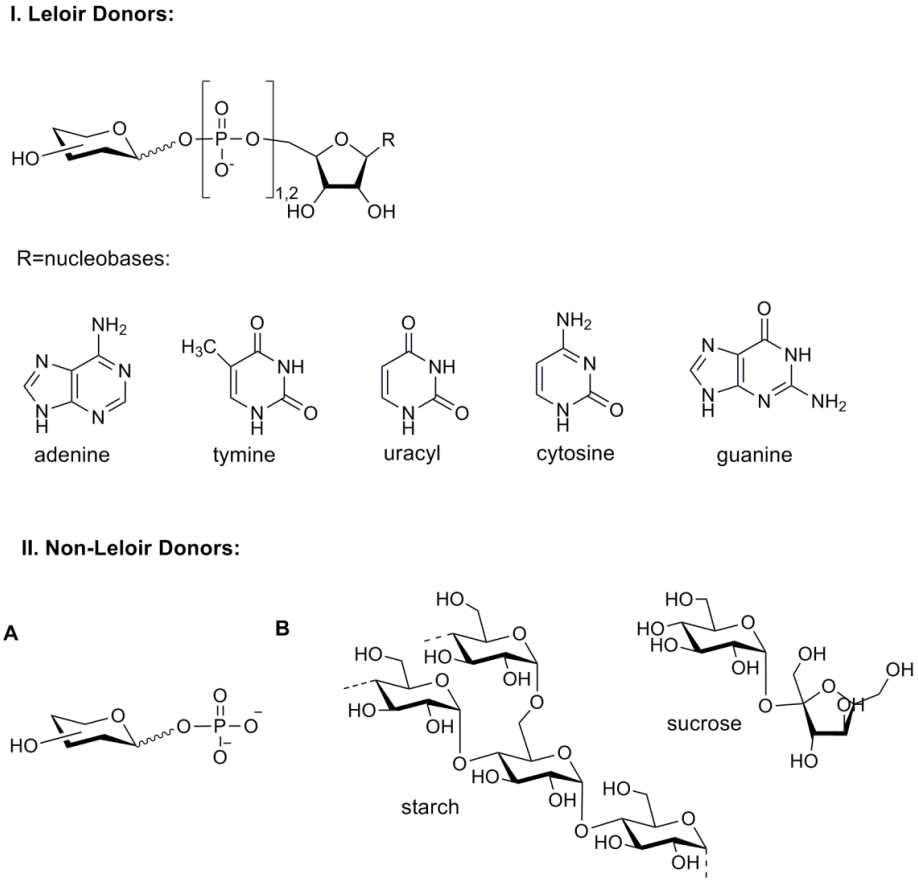


Figure 9. Different donors utilized by GTs divided into two categories: I. Leloir donors II. non-Leloir donors **A.** activated with phosphate **B.** non-activated sucrose and starch.

The majority of mammalian GTs found in Golgi apparatus and endoplasmic reticulum belongs to so called GT-A superfamily (Figure 10), with the catalytic domain being able to bind both, nucleotide sugar donor and acceptor, generally in a divalent metal-ion-dependent matter.[65,66] Mn^{2+} or Mg^{2+} ions are coordinated in the enzyme active site and further facilitate the departure of leaving group by stabilization of negative charge of phosphate group of sugar donor. Unlike GT-A enzymes, GT-B proteins possess two distinct domains and do not required metal ion for their correct action.[67] More recently a third superfamily of glycosyltransferases, GT-C was identified and the first full length crystal

structure of bacterial oligosaccharyltransferase PglB from *Campylobacter lari* has been resolved.[68] GT-C utilizes lipid phosphate-linked sugars as sugar donors.

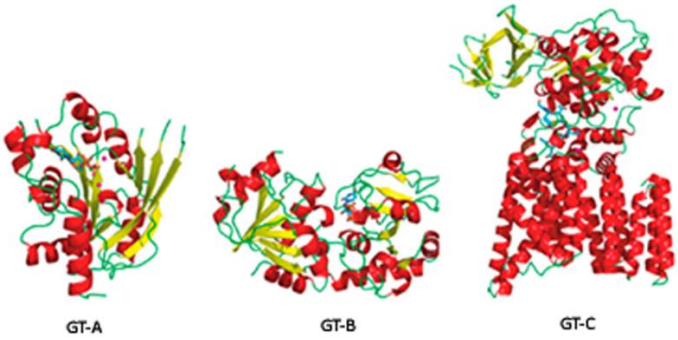


Figure 10. Representative structures of glycosyltransferases type GT-A, GT-B and GT-C. (Gloster *et al.* 2014[65])

Depending on the stereochemical outcome of glycosyltransferase action, their mechanism can be further classified as an inverting (α-->β) or retaining (α-->α) (Figure 11).

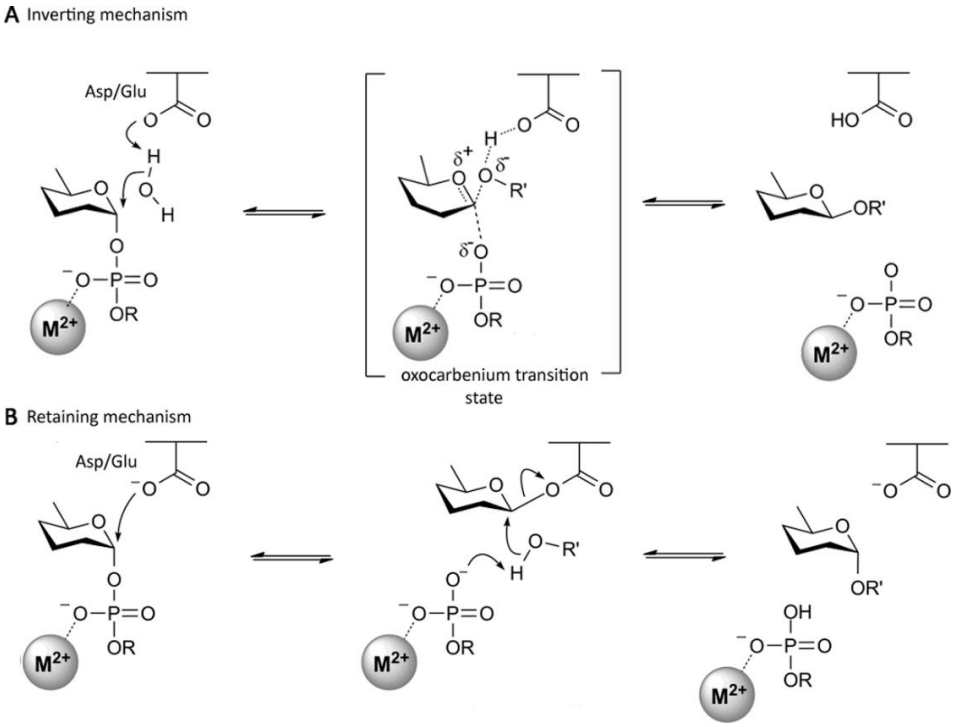


Figure 11. Proposed mechanism of GTs action; **A.** Inverting and **B.** Retaining mechanism.

The inversion of configuration at C1 of glycosyl donor is usually explained as a direct displacement S_N2-like mechanism (Figure 11, A). The carboxylate residue from aspartic acid (Asp) or glutamic acid (Glu) in the enzyme catalytic site deprotonates the glycosyl acceptor, which subsequently attacks the C1 atom of the glycosyl donor (S_N2). An oxocarbenium like transition state is formed which rearrange to the glycoside bond with an inversion of configuration at C1. The divalent metal ion facilitates a leaving group departure.

The retaining mechanism is far less understood, however the double-displacement theory has been proposed (Figure 11, B). In this mechanism, after the direct S_N2 nucleophile attack of carboxylic residue on the C1 of glycosyl donor the β linked enzyme-donor intermediate is formed followed by reaction with the glycosyl acceptor with a second inversion of stereochemistry.[69]

1.4.1 Expression of recombinant glycosyltransferases

The glycosyltransferases are powerful tools for the synthesis of oligosaccharides. However, it is difficult to obtain them in a sufficient amount by direct isolation from tissues, fluids or natural extracts. Recent progress in molecular biology allows for the recombinant expression of proteins in host organisms such as bacteria, yeast, insect cells or mammalian cells.[70] This methodology can provide pure enzymes in large quantities that find application in the enzymatic preparation of glycoconjugates. In our N-glycan synthesis, we have employed GTs expressed in bacterial cultures (*E. coli*) and in yeast cultures (*P. pastoris*). In the next section, the advantages and limitations of both expressions will be briefly discussed.

1.4.1.1 E. coli expression system

Escherichia coli host organism is usually the first choice when expressing heterologous proteins, due to its simplicity and easy manipulation. *E. coli* culture is inexpensive and can be performed in standard laboratories. Moreover, proteins are produced very rapidly and in large quantities. There are many commercial expression vectors available with different N- and C-terminal tags and several strains optimized for special applications. Nevertheless, the biggest drawback of using *E. coli* as expression system is the inability to form disulfide bonds and post-translational modifications. Therefore, many eukaryotic proteins do not fold properly in *E. coli* and form insoluble aggregates, called inclusion bodies. Inclusion bodies are biologically inactive and need to be further processed to recover their biological function. The protein folding from inclusion bodies is an empirical process that need to be optimized for each individual protein.[71–77] The process involves several steps such as the isolation of inclusion bodies from the bacterial pellet, solubilization in the appropriate chaotrophic reagent, optimization of refolding buffer and finally the purification/concentration of the desired protein. Folding buffers usually contain a large number of components, including oxidized/reduced thiol pairs, such as glutathione (GSH/GSSG) or cysteamine/cystamine. Both oxidizing/reducing agents help the formation of disulfide bonds, which are crucial in the process of refolding. Several additives can be present in the folding buffer, like L-arginine, detergents, low molecular concentration of chaotrophic reagents such as urea or sugars. Nevertheless, there are some advantages of protein expressed as inclusion bodies production, including their lower degradability, resistance to proteolytic attack or very high level of expression.

1.4.1.2 Pichia pastoris expression system

Pichia pastoris is frequently used as an expression system for the production of recombinant proteins. Yeasts are easy to cultivate, they grow on inexpensive media without

any special requirement and they are considered to be simpler to handle compared to other host such as insect and mammalian cell cultures.

P. pastoris is a single-cell eukaryotic microorganism, capable to perform several posttranslational modifications typical of eukaryotic organisms including folding, disulfide bond formation and glycosylation of proteins. Thus, proteins produced in the system, either intracellular or secreted into medium are directly biological active, and beside purification, do not required further manipulations. *P. pastoris* is a methylotrophic yeast capable to utilize methanol as only carbon source. It has a specialized organelle, called peroxisome, where the methanol is oxidized to form formaldehyde and hydrogen peroxide by a set of specialized enzymes called alcohol oxidases (AOX). There are two genes encoding AOX enzymes in *P. pastoris*, *AOX1* and *AOX2*, however it is the *AOX1* which is the principal responsible for alcohol oxidase activity and whose expression is strictly regulated and induced by the presence of methanol.

In order to express a recombinant protein in *P. pastoris*, the foreign cDNA sequence encoding the target protein has to be introduced into the yeast genome, in a process called transformation. There are several expression vectors specially designed for *P. pastoris*, sharing variety of common characteristics, including the presence of a foreign gene expression cassette. The expression cassette is composed of *P. pastoris* AOX1 promoter gene, followed by the unique restriction sites for the insertion of the recombinant protein sequence and transcriptional termination sequence. The expression vectors might additionally carry drug resistance markers, which allow for the enrichment of strains with multiple copies of foreign gene. Once the foreign gene encoding the recombinant protein is introduced into the genome of *P. pastoris*, the yeasts are cultivated and protein expression is simply induced by addition of methanol to the growing media. In general, proteins in *P. pastoris* are expressed extracellularly and the protein is secreted to the expression media being in many cases the major component of the media facilitating subsequent isolation and/or purification.

1.5 Expression of GTs used for diversification of synthetic N-glycans

Recombinant glycosyltransferases which were used for the diversification of the target N-glycans (**G1**-**G6**) were produced in *Pichia pastoris* and *E. coli* expression systems as summarized in Table 6.

Enzyme	Expression system
Bovine milk β-1,4 GalT[86,87]	*Escherichia coli*
C. elegans β-1,4-GalNAcT[82]	*Pichia pastoris*
C. elegans α-1,3-FucT (CeFUT6)[90,84]	*Pichia pastoris*
C. elegans α-1,3-FucT core type (CeFUT1)[90,84]	*Pichia pastoris*
C. elegans α-1,6-FucT core type (CeFUT8)[79]	*Pichia pastoris*
A. thaliana α-1,3-FucT core type (AtFucTA)[83]	*Pichia pastoris*

Table 6. Glycosyltransferases (GTs) expressed and used during the chemo-enzymatic preparation of N-glycans structures.

1.5.1 Expression of glycosyltransferases in *P. pastoris*

The target N-glycan library in this study differs significantly from the mammalian set of N-glycan structures. The synthetic compounds **G1**-**G6** contain β-1,2-linked xylose which is typically found in invertebrates and plants but never in mammalian organisms. We want to further diversify the synthetic structures and introduce other modifications which will resembles antigenic structures presented in some pathogens, like *S. mansoni* or in plant allergens.[78,79] We therefore focused on the preparation of multiple core α-fucose decorated N-glycans and structures with terminal LDN, LDNF or Lewis X-motifs. We centered our attention in glycosyltransferases available for recombinant expression that come from two

model organisms, such as the nematode *Caenorhabditis elegans* and the plant *Arabidopsis thaliana*.

C. elegans express a high number of fucose containing N-glycans, and it is able to accommodate up to three fucose residues in the chitobiose core of N-glycans.[80,81] Moreover, the β-1,4-galactosaminyltransferase activity involved in the LDN synthesis as well as the α-1,3-fucosyltransferase able to prepare Lewis X like structures have been identified.[82] On the other hand, *A. thaliana* express core α-1,3-fucose[83] and β-1,2-xylose modifications on its N-glycans. Our group has participated in the characterization of some of these enzymes using glycan microarrays[84] and therefore, expression of these and some other enzymes has been set up in our laboratory.

In this thesis *P. pastoris* strains encoding four different fucosyltransferases: AtFucTA, CeFUT1, CeFUT6, CeFUT8 and one *N*-acetylgalactosaminetransferase GalNAcT were used. The gene encoding AtFucTA was originally identified and isolated from the *Arabidopsis thaliana* plant.[83,85] The enzyme possesses core α-1,3-fucosyltransferase activity and is involved in the synthesis of N-glycans specific to plants and invertebrates which have at least one non-reducing terminal N-acetylglucosamine residue on the 3-arm (Figure 12, A-I). CeFUT1, CeFUT6 and CeFUT8 were identified in *C. elegans*. CeFUT1 exhibit core α-1,3-fucosyltransferase activity similar to AtFucTA, however its substrate requirement differs.[79] Unlike the plant enzyme, the nematode CeFUT1 accepts paucimannose N-glycans lacking terminal GlcNAc residues (Figure 12, A-II). Nematode CeFUT6 exhibits dual activity. It has been described to synthesize Lewis X type motifs *in vitro* but the *C. elegans* glycome does not contain these type of structures and the *in vivo* activity of the enzyme has been established recently, CeFUT6 is capable to introduce α-1,3-fucose in the distal GlcNAc of truncated N-glycans lacking 6-arm (Figure 12, C).[84] CeFUT8 is a homologue of mammalian core α-1,6-fucosyltransferase, with the typical substrate specificity for this type of enzyme, requiring the presence of a non-reducing terminal N-acetylglucosamine residue (Figure 12, B).[80] The *N*-acetylgalactosaminetransferase from *C. elegans* decorates terminal, non-

reducing GlcNAc moieties with β-1,4-GalNAc residues creating LDN epitopes (Figure 12, D).[82]

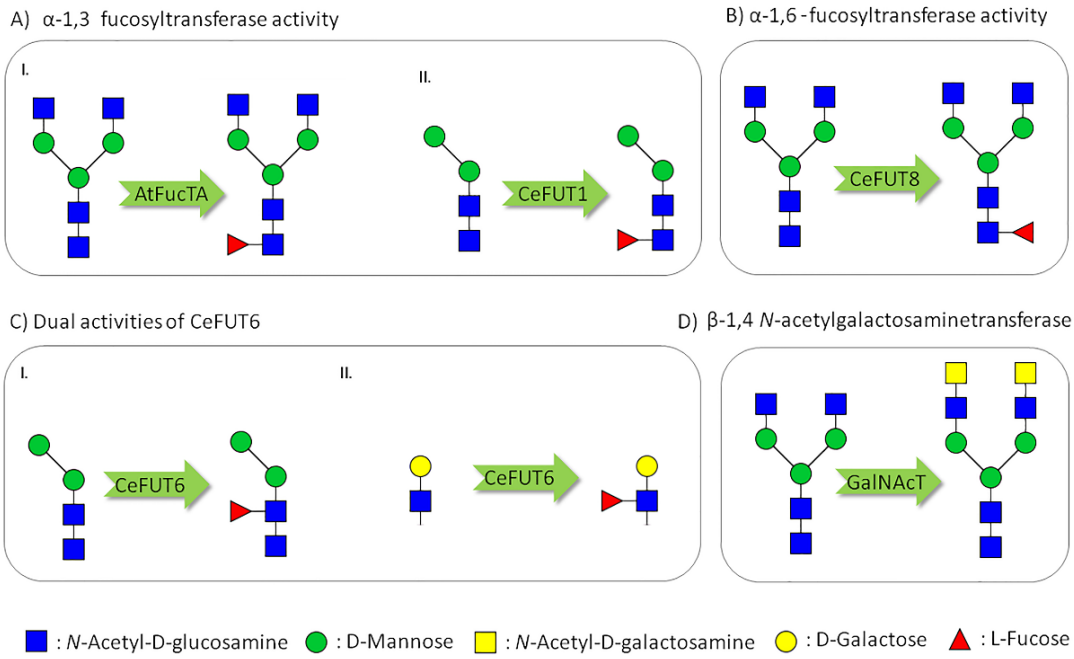


Figure 12. Schematic representation of recombinant glycosyltransferases activities and their substrate specificity; **A-I**. α-1,3-fucosyltransferase activity of AtFucTA; **A-II.** α-1,3-fucosyltransferase activity of CeFUT1 towards paucimannose N-glycans; **B.** α-1,6-fucosyltransferase activity of CeFUT8; **C-I.** α-1,3-fucosyltransferase activity of CeFUT6 towards distal GlcNAc of truncated N-glycans; **C-II.** Lewis X type α-1,3-fucosyltransferase activity of CeFUT6; **D.** β-1,4-*N*-acetylgalactosaminetransferase activity.

The protein expression in *P. pastoris* is summarized in Figure 13, the frozen glycerol stocks of recombinant *P. pastoris* were inoculated into agar plates containing Zeocin™ as selection marker. The agar plates were incubated at 30°C until yeast colonies appeared. A single colony was selected and inoculated in a pre-culture medium containing glycerol as carbon source. The generation of the biomass was monitored by measuring the absorbance at 600 nm (OD_{600nm}=1 corresponds to $3x10^7$ cells/mL).

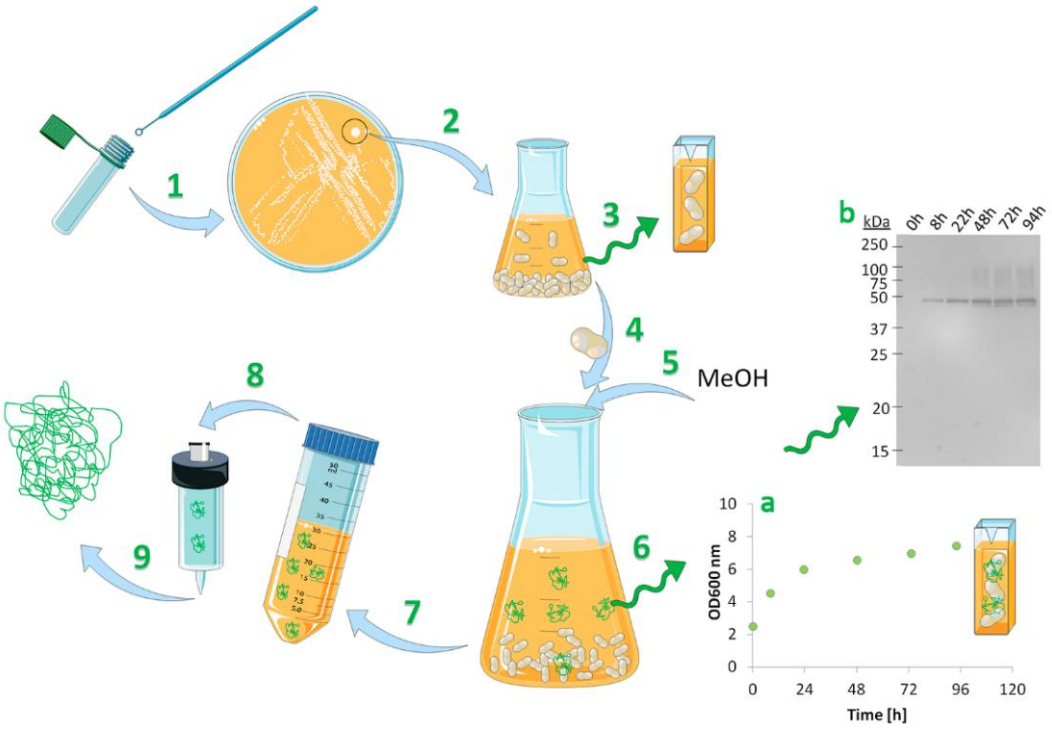


Figure 13. Schematic representation of the recombinant proteins expression in *P.pastoris* system; **1.** Frozen glycerol stock is inoculated on the agar plate; **2.** A single colony is selected and introduced to the pre-culture medium; **3.** The growth of the yeast is controlled by the OD_{600} measurement; **4.** Cells are seeded in medium for protein expression; **5.** The protein expression is induced by the methanol (1%) addition; **6.** The protein expression is monitored by the gel electrophoresis or Western Blot and the cell growth by OD measurements: **a.** absorbance at 600 nm at different time points after CeFUT1 protein induction by MeOH addition, **b.** image of CeFUT1 protein expression taken at different time points monitored by Western blot detected with Anti-His tag antibody; **7.** Collection of the soluble proteins; **8.** Purification of the His-tagged proteins on the NTA-Ni cartridge; **9.** Purified recombinant protein.

After the initial step for biomass generation, cells were collected by centrifugation and seeded to OD_{600nm}=1 in medium for protein expression, in this case containing methanol as sole carbon source. The induction of protein was maintained by the addition of methanol (1%) every 24 hours and protein expression in the medium supernatant was monitored by

gel electrophoresis and/or by Western Blot. In our case, the optimal time for yeast harvesting was 72 hours after protein induction. At this point, the *P. pastoris* culture reaches the steady state of growth and the amount of protein expressed reaches a plateau as shown by Western Blot analysis (Figure 13). After 72h of protein expression the culture is centrifuged and the protein presented in the medium was purified by immobilized metal affinity chromatography (IMAC) or used directly after concentration by ultracentrifugation.

All expressed fucosyltransferases contain a hexahistidine tag (6xHis tag) on their N-terminus and they were purified using Ni-NTA (nitrilotriacetic acid) columns (Figure 14, A).

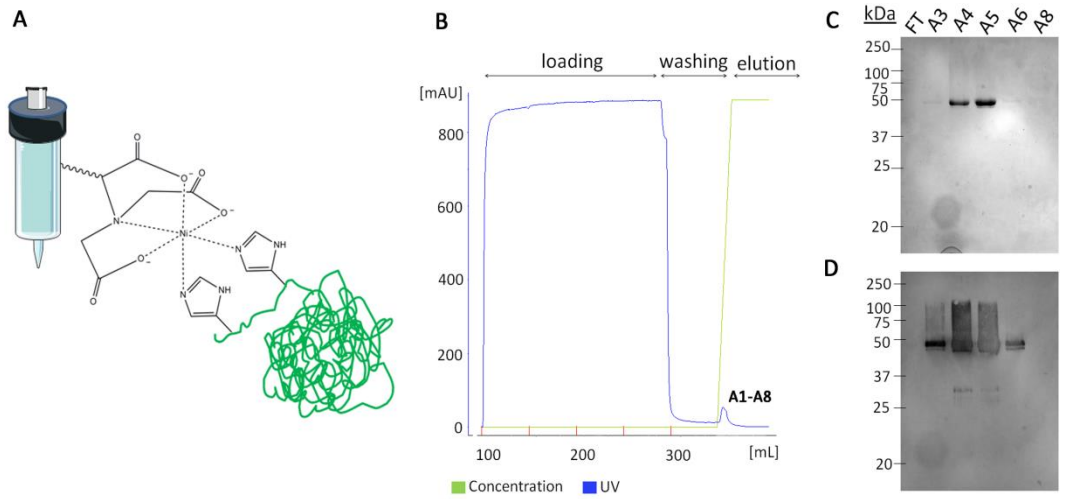


Figure 14. Purification of the recombinant proteins using immobilized metal affinity chromatography (IMAC); **A.** Ni-NTA loaded column and the Ni-His complex; **B.** Chromatogram showing loading, washing and elution steps of the recombinant protein purification; **C.** Collected fractions monitored by gel electrophoresis; **D.** Collected fractions monitored by the Western Blot with the anti-His antibody.

His-tagged recombinant proteins expressed high affinity towards nickel and forming stable complexes which could be trapped on the column. At the end, each Ni^{2+} coordinates with one nitrilotriacetic acid and two histidines from the recombinant protein. The protein is

eluted from the column by using an imidazole gradient which competes for binding to Ni-NTA and causes protein elution from the column (Figure 14, B). Protein fractions were analyzed by gel electrophoresis and/or by Western Blot (Figure 14, C-D) and dialyzed to remove imidazole and/or concentrated by ultracentrifugation. Purified recombinant fucosyltransferases were stored at -4ºC and were highly stable.

1.5.2 Expression of bovine milk β1,4- galactosyltransferase in *E. coli*

Bovine milk galactosyltransferase (β1,4-GalT) is a well-studied enzyme which catalytic domain has been successfully cloned and expressed in *E. coli*, as previously described.[86,87] This enzyme catalyzes the transfer of galactose from UDP-galactose to terminal N-acetylglucosamine residues and it is used for elongation of oligosaccharide chains (Figure 15, A).

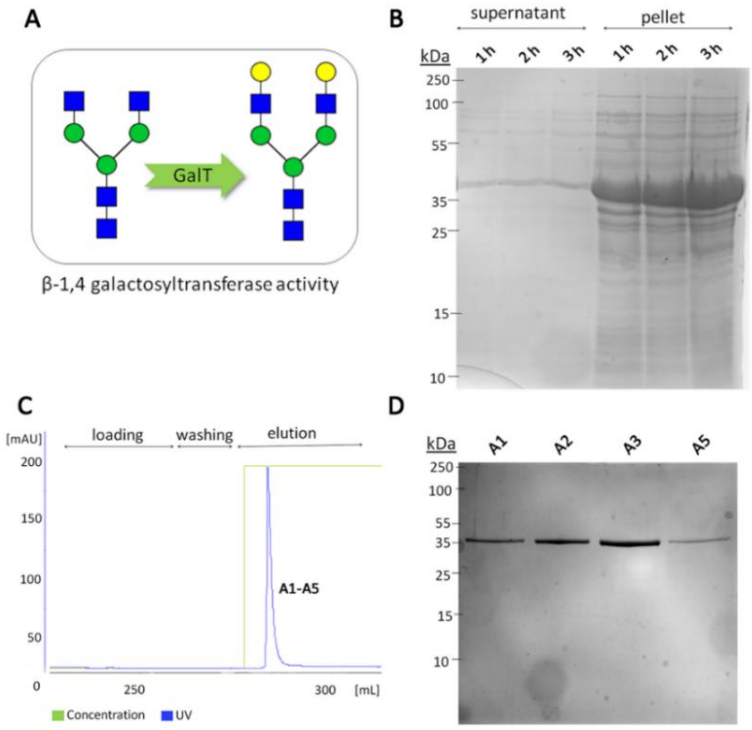


Figure 15. β-1,4-GalT; **A.** Specificity of enzyme; **B.** SDS-PAGE analysis of protein expression present in the soluble fraction and in pellet after 1h, 2h and 3h; **C.** Purification of GalT on the Ni-NTA column; **D.** SDS-PAGE of collected fractions A1-A5.

E. coli cell line containing the catalytic domain of β-1,4-GalT gene was seeded in Luria Bertani (LB) agar plates containing the appropriate antibiotic and incubated at 37°C overnight. A single colony was inoculated in liquid LB and incubated overnight. The initial culture was diluted and bacterial growth was maintained at 37ºC until cells reached the OD_{600nm} value of 0.6-0.7. At this point, protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) and monitored at different time points by gel electrophoresis (Figure 15, B).

Protein expression was analyzed both in the cell lysate supernatant and in the bacterial pellet. An over expressed protein of around 37 kDa was found in the bacterial pellet expressed as inclusion bodies. The washed inclusion bodies were folded as previously described in the literature and purified by NiNTA affinity chromatography to yield the functional β1,4-GalT.[86]

1.6 Diversification of G1-G6 N-glycans with recombinant glycosyltransferases

Having fully characterized N-glycans **G1**-**G6** (Chapter 1 Section 1.2) and purified, active recombinant glycosyltransferases (Chapter 1 Section 1.5) we have further performed the enzymatic diversification of the synthetic structures in solution. The key aspect of this strategy is based on previous knowledge of substrate specificity of enzymes used during the study, mentioned in the Sections 1.5.1 and 1.5.2 (Figure 12, Figure 15). Taking into account the previously described substrate specificity of glycosyltransferases and the 'order' of their action *in vivo*, GTs can be easily translated into the *in vitro* situation, and be further used as a synthetic tool for the substrate diversification.

Glycosyltransferases are very specific both for the sugar-nucleotide glycosyl donor and the acceptor sugar. For example, in mammals up to seven fucosyltransferases act on the disaccharide substrate Galα-1,3/4-GlcNAc but their specificities differ. A common feature of the family of enzymes transferring fucose onto the GlcNAc residue is that they do not tolerate substitution on position-6 of the galactose residue, indicating that the corresponding hydroxyl group is involved in substrate binding to the active site.

In our study, xylose containing synthetic N-glycans **G1**-**G6** have been used as a starting point for their further diversification by the action of recombinant glycosyltransferases. First the fucosyltransferases were applied. Based on the microarray-assisted probing of CeFUT8 and AtFucTA performed in our laboratory[88] and the earlier literature[80] it was shown, that core α-1,6-fucosylation occurs prior to α-1,3-fucosylation and that the α-1,3-fucosylated compounds are no longer substrate for the CeFUT8. Fucosylated or bisfucosylated compounds can be treated with GalT or GalNAcT resulting in the installation of terminal β-1,4-linked LN and LDN elements on the non-reducing GlcNAc of glycans antennae. When the terminal β-1,4-Gal and β-1,4-GalNAc are introduced prior the core fucosylation, the subsequent installation of α-1,6-fucose moieties in the reducing GlcNAc is impeded. At the end, LN or LDN decorated glycans can be further diversify with CeFUT6, which introduces α-1,3-fucose to the GlcNAc of antennae, creating corresponding Lewis X

compounds. As mentioned in the Section 1.5.1 of this Chapter, besides the Lewis X type fucosyltransferase activity of CeFUT6, the enzyme can also catalyze the fucosylation of the distal GlcNAc of truncated N-glycans, both paucimannose and with non-reducing GlcNAc on the 3-arm (See Figure 17, **G41** and **G48**).[84] In both cases, only glycans without further substituents in O-6 of the central β-mannose are tolerated. Here, as an additional constraint, we found that the presence of a core xylose impedes the fucosylation with CeFUT6 on glycans which are otherwise good substrates for the enzyme.[89] This behavior was demonstrated in solution followed by MALDI-TOF MS and on glycan arrays, prepared by the printing the synthetic structures **G1**-**G6** and the corresponding N-glycans lacking β-1,2-xylose available in the lab on NHS-activated glass slides (Figure 16). A further description of the microarray technique will be given in Chapter 2, here it is only important to mention that glycan arrays can be a suitable and fast tool for the screening of enzyme specificity.

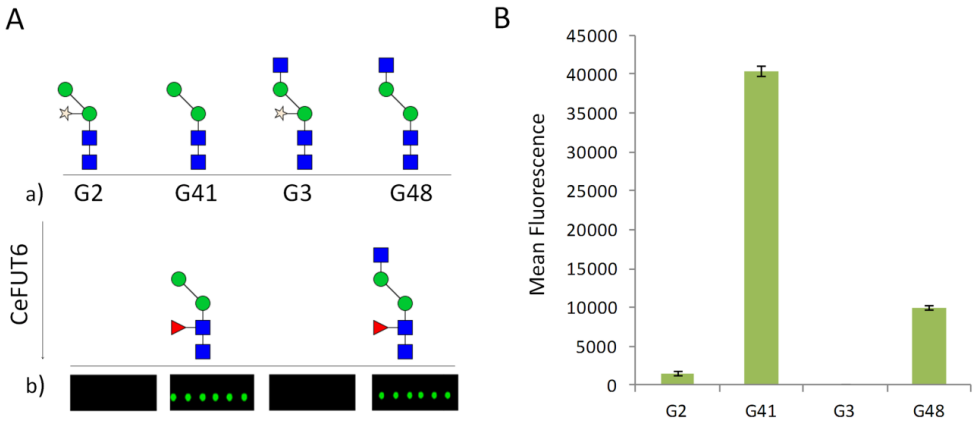


Figure 16. Screening for CeFUT6 substrates on the glycan array; **A.** Enzymatic on-chip elongation with CeFUT6 (a); the N-glycan microarray image after incubation with fucose-recognizing AAL-555 lectin (b); **B.** Mean fluorescence intensities of the CeFUT6-incubated (fucosylated) array after incubation with AAL-555 lectin.

The structures immobilized on the microarray surface were incubated with *C. elegans* α-1,3 fucosyltransferase CeFUT6 in the presence of GDP-fucose. The presence of enzymatically

extended structures was verified by incubation with fluorescently tagged fucose specific *Aleuria aurantia* lectin (Figure 16). As a result of the CeFUT6 fucosyltransferase action on the xylosylated compounds **G1**-**G6**, only Lewis X or LDNF element can be introduced, never α-1,3 fucose of the innermost GlcNAc of chitobiose core.

In general, the synthesis of target compounds involves up to 4 subsequent elongations with recombinant glycosyltransferases starting with core fucosylation/s (α-1,6 prior α-1,3; CeFUT8, AtFucTA, CeFUT1), followed by terminal β-1,4 GalT or β-1,4 GalNAcT modifications and the final introduction of Lewis X type element.

In addition, the paucimannose compounds **G35-G39** were accessible by incubation of several enzymatically diversified N-glycans or biantennary compound **G6** (see Figure 17, orange arrows) with a commercially available β-*N*-acetyl-glucosaminidase from *Canavalia ensiformis*, which cleaves glycosidic bonds at terminal GlcNAc moieties of N-glycan antennae. Enzymatic reaction could be easily followed by MALDI-TOF mass spectrometry and reaction was continued until a single product was formed. After each enzymatic elongation, the reaction product was purified on the reverse phase C18 cartridges and fully characterized by NMR and mass spectrometry.

At this point, it is necessary to point out the great effectiveness of enzymatic preparation of carbohydrates. Starting from six synthetic N-glycan structures and by the combine action of 6 recombinant glycosyltransferases and 1 commercially available glycosyl hydrolase, we could achieve the formation of 33 new fully characterized N-glycan structures. The enzymatic reactions proceeded with complete stereo- and regioselectivity leading to the formation of desired products in high yields.

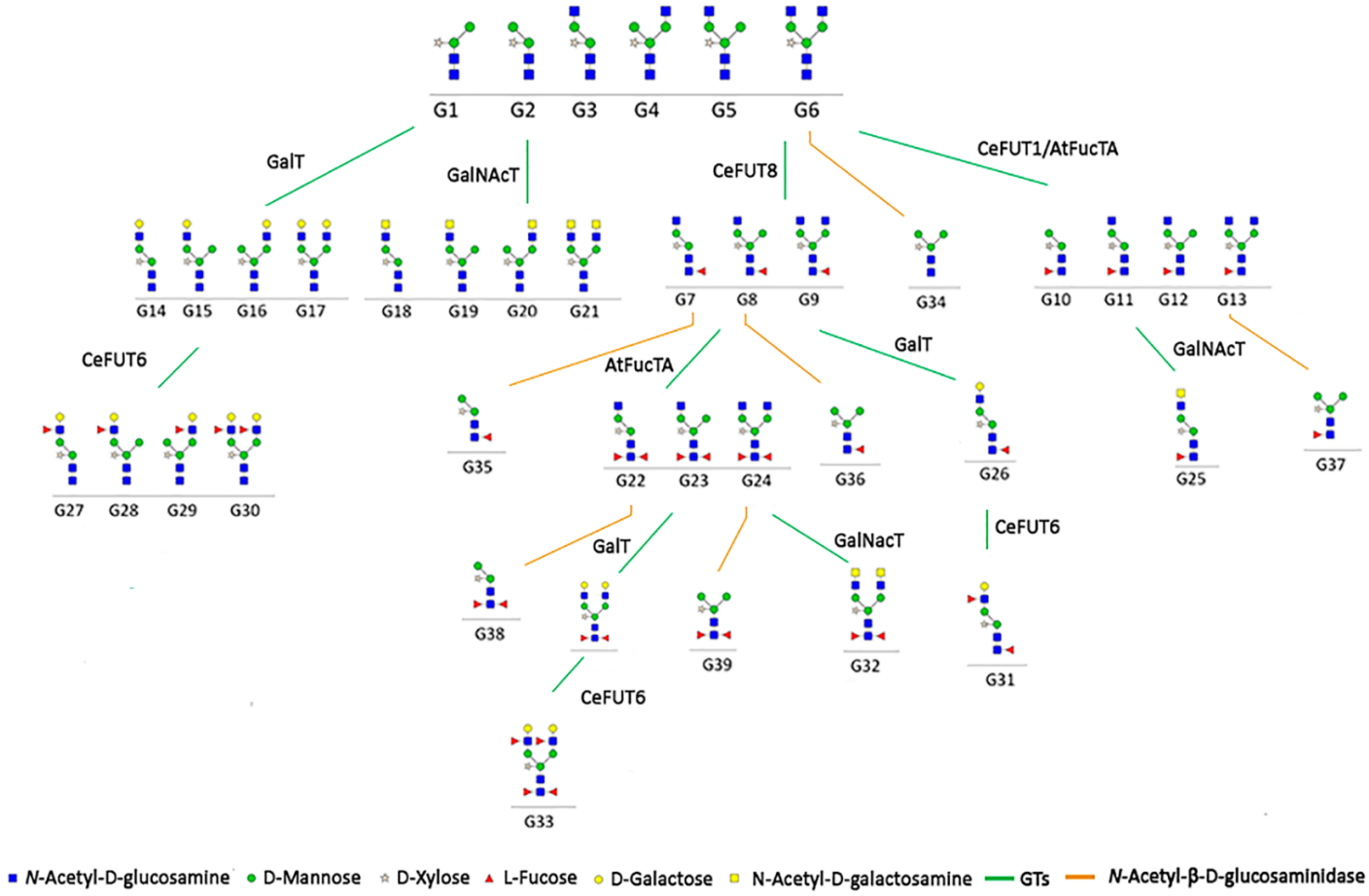


Figure 17. Enzymatic diversification of synthetic N-glycans (G1-G6) using glycosyltransferases GTs (green arrows) and *N*-acetyl-β-D-glucosaminidase (orange arrows)

1.7 Summary

In this chapter, the chemoenzymatic synthesis of 39 β-1,2-xylosylated N-glycans carrying core and terminal modifications, including α-1,3- or/and α-1,6-core fucosylation, LN, LDN and Lewis X elements (Figure 18) was described. In the first part, we have focused on the modular assembly of carbohydrates leading to the six synthetic structures **G1**-**G6**. Later we have described the expression of recombinant glycosyltransferases available in our laboratory and finally, their use as complementary tool to the carbohydrate synthesis towards diversification of synthetic structures (compounds **G7**-**G39**).

Some of the chemoenzymatically prepared structures are of biological relevance in the field of helminth infections as well as in food allergies. Therefore, in the next chapters we will describe the preparation of glycan arrays including above mentioned compounds and glycan array-based screenings of C-type lectin receptors which might be involved in the pathogen recognition, as well as monoclonal and polyclonal antibodies of mice and humans suffering from *S. mansoni* infections.

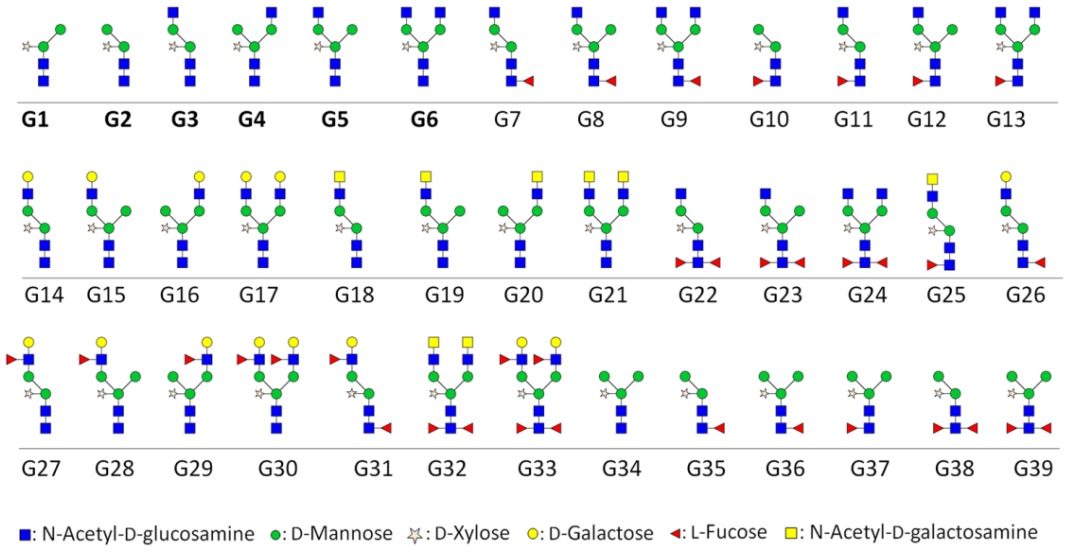


Figure 18. List of the total number of xylosylated N-glycans prepared chemoenzymatically.

1.8 Experimental Part

General Methods

Materials: Chemicals were purchased from Sigma-Aldrich or Acros Organics and were used without further purification. All reaction solvents were dried over activated 4Å or 3Å molecular sieves. Thin layer chromatography was carried out using Merck aluminium sheets silica gel 60 F_{254} and visualized by UV irradiation (254 nm) or by staining with vanillin solution The recombinant enzymes α-1,3-fucosyltransferases (AtFucTA) from *Arabidopsis Thaliana,* α-1,6-fucosyltransferase, (CeFUT8), core α-1,3-fucosyltransferases (CeFUT6/CeFUT1) and β-1,4-N-acetylgalactosaminyltransferase from *C. elegans,* bovine milk β-1,4-galactosyltransferase were expressed as described previously.[90,79,82,83,84,85,86,87] β-N-acetylglucosaminidase from *Canavalia ensiformis* (Jack bean) was purchased from Sigma Aldrich. Uridine 5'-diphosphogalactose disodium salt (UDP-Gal), uridine 5'-diphospho-N-acetylgalactosamine disodium salt (UDP-GalNAc) and guanosine 5'-diphospho-β-L-fucose sodium salt (GDP-Fuc) were purchased from Carbosynth (San Diego, USA). *Aleuria aurantia* lectin (AAL) lectin was purchased from VectorLabs and was labeled with Hilyte Plus™ 555 protein labeling kits from AnaSpec, Freemont, USA.

Microwave irradiation was performed on Biotage Initiator monomode oven, (Biotage AB, Uppsala, Sweden). Hydrogenation reactions were performed in continuous-flow hydrogenation reactor H-Cube® and CatCarts® 10% Pd/C from ThalesNano Nanotechnology Inc. (Budapest, Hungary). Purifications of compounds were performed on: size-exclusion chromatography using LP2 gel from Biorad (Hercules, USA), SampliQ high performance graphitized carbon cartridges (1 mL) from Agilent Technologies (Santa Clara, CA, USA), C18 Sep-Pak Cartridges (1 mL) from Waters (Milford, Massachusetts, USA), flash chromatography using Merck 62Å 230-400 mesh silica gel or on a Biotage SP4 automated flash chromatography system, (Biotage AB, Uppsala Sweden) employing prepacked silica cartridges. All aqueous solutions were prepared from nanopure water produced with a

Diamond UV water purification system (Branstead International, IA, Madrid, Spain). Pooled glycan containing fractions were lyophilized on an ALPHA-2-4 LSC freeze-dryer from Christ, Osterode, Germany. All compounds in organic solvents were concentrated using rotary evaporator.

^{1}H and ^{13}C spectra were acquired on Bruker 500 MHz spectrometer or on Bruker BioSpin GmbH 800 MHz. 13C-HSQC spectra recorded with Bruker BioSpin GmbH 800 MHz use a folded window in the F1(13C) dimension (with a sweepwidth value of 40 ppm). Chemical shifts were reported in ppm (δ) and referenced to the residual signal of the solvent used (MeOD 4.87 ppm; $CDCl_3$ 7.26ppm; D_2O 4.79 ppm). Plitting patterns are designated as s, singlet; d, doublet; t, triplet; m, multiplet. Coupling constants (*J*) are reported in Hz. High resolution mass spectra were acquired on a Waters LCT Premier XE instrument, (Waters, Manchester, UK) equipped with a standard ESI source by direct injection. The instrument was operated with a capillary voltage of 1.0 kV and a cone voltage of 200 V. Cone and desolvation gas flow were set to 50 and 600 L/h, respectively; source and desolvation temperatures were 100 ºC. MALDI-TOF mass analyses were performed on an Ultraflextreme III time-of-flight mass spectrometer equipped with a pulsed N_2 laser (337 nm) and controlled by FlexControl 3.3 software (Bruker Daltonics). The acquisitions were carried out in positive ion reflectron mode at a laser frequency of 500 Hz. Microarrays were printed employing a robotic non-contact piezoelectric SciFLEXARRAYER spotter S11 (Scienion, Berlin, Germany). Fluorescence measurements were performed in an Agilent G2565BA microarray scanner system (Agilent Technologies, Santa Clara, USA) with 10 µm resolution, using 2 lasers (532 nm, 633 nm). Quantification of fluorescence was achieved by ProScanArray® Express software from Perkin Elmer, Shelton, USA.

Chemical synthesis of N-glycans G1-G6:

5-azidopentyl 4,6-*O*-benzylidene-3-O-(2-naphthylmethyl)-β-D-mannopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-gluco-pyranoside (2): To a solution of trisaccharide **1** (200 mg, 132.9 µmol) in CH_2Cl_2 (3 mL) 12 mL of $Mg(OMe)_2$ were added. The reaction mixture was heated to 80 ºC by microwave irradiation (4 cycles of 30 minutes). The suspension was filtered through a pad of Celite® and the filtrate was neutralized with 10% acetic acid. The organic layer was washed with saturated $NaHCO_3$, water, saturated NaCl, dried over $MgSO_4$ and concentrated. The crude product was purified by flash column chromatography (SiO_2, Hexane:EtOAc, 3:1 to 2:1) to afford the title compound as a white foam (167 mg, 86%); **$[\alpha]^{20}_{589}$**= +14.2 (c, 0.5, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.87–7.66 (m, 12H, Ar), 7.48–7.46 (m, 5H, Ar), 7.39–7.29(m, 8H, Ar), 7.24–7.14 (m, 5H, Ar), 7.02-7.00 (m, 2H, Ar), 6.97-6.88 (m, 5H, Ar), 6.78-6.74 (m, 3H, Ar), 5.51 (s, 1H, OC*H*Ph), 5.27 (d, *J*= 8.7 Hz, 1H, H-1Aβ), 4.93–4.90 (m, 2H, H-1Bβ, C*H*$_2$Ph), 4.87–4.82 (m, 3H, C*H*$_2$Ph), 4.65 (s, 1H, H-1Cβ), 4.54-4.35 (m, 7H, H-3A, C*H*$_2$Ph), 4.24-4.20 (m, 2H, H-2A, H-4B), 4.14-4.09 (m, 4H, H-2B, H-6aC, H-3B, H-4A), 4.05-4.00 (m, 2H, H-2C, H-4C), 3.69-3.62 (m, 2H, H-6aA, O*CH*H$(CH_2)_4N_3$), 3.58-3.48 (m, 4H, H-6bA, H-6aB, H-6bC, H-3C), 3.41 (dd, J_{6a-6b}= 11.1 Hz, J_{6b-5}= 3.7 Hz, 1H, H-6bB), 3.34-3.22 (m, 3H, H-5A, H-5B, O*CH*H$(CH_2)_4N_3$), 3.14-3.09 (m, 1H, H-5C), 2.91-2.81 (m, 2H,-C*H*$_2N_3$), 1.40-1.28 (m, 4H, $OCH_2CH_2CH_2CH_2CH_2N_3$), 1.10-1.02 (m, 2H, $O(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 138.7-123.3, 101.7, 100.8, 98.2, 97.1, 79.1, 78.3, 75.8, 74.8, 74.6, 74.5, 74.5, 73.4, 72.7, 72.6, 69.9, 69.0, 68.6, 68.3, 67.9, 66.8, 56.7, 55.8, 51.2, 29.8, 28.4, 23.1; **HRMS (MALDI-Tof) *m/z*** $[M+Na]^+$ calculated for $C_{85}H_{83}N_5O_{18}$ 1484.5625, found 1484.5649.

5-azidopentyl (2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl)-(1→2)-(4,6-*O*-benzylidene-3-*O*-(2-naphthylmethyl)-β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (4): A mixture of **2** (172 mg, 117.6 µmol) and **3** (125 mg, 297.1 µmol) was dried under high

vacuum overnight. The mixture was dissolved in anhydrous CH_2Cl_2 (1.1 mL) and freshly activated molecular sieves (3Å) were added. After stirring for 40 minutes at room temperature, the mixture was cooled to 0°C and 5% solution of trifluoromethanesulfonic acid (50 µL, 0.28 eq) in CH_2Cl_2 was slowly added. The mixture was allowed to warm up to room temperature and stirred for 2h. The reaction mixture was then quenched with Et_3N, filtered through a pad of Celite® and concentrated. The crude product was purified by automated flash column chromatography (toluene:acetone 10:1) to afford the title compound (128 mg, 68%) as a white foam; **$[\alpha]^{20}_{589}$**= -4 (c, 0.5, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.87-7.78 (m, 4H, Ar), 7.74-7.63 (m, 7H, Ar), 7.51-7.43 (m, 5H, Ar), 7.40-7.28 (m, 9H, Ar), 7.23-7.16 (m, 2H, Ar), 7.15-7.09 (m, 3H, Ar), 7.06-7.02 (m, 2H, Ar), 6.99-6.93 (m, 4H, Ar), 6.91-6.86 (m, 1H, Ar), 6.80-6.73 (m, 3H, Ar), 5.46 (s, 1H, OC*H*Ph), 5.30 (d, *J* = 8.4 Hz, 1H, H-1Aβ), 5.10-5.07 (m, 2H, H-1Xβ, H-3X), 5.04 (dd, J=3.47Hz, J=5.4Hz, 1H, H-2X), 4.94-4.92 (m, 1H, H-1Bβ), 4.90-4.82 (m, 5H, H-4X, OCH_2Ph), 4.60 (dd, $J_{5xa\text{-}5xb}$=13Hz, $J_{5xa\text{-}4}$=3.4Hz, 1H, H-5Xa) 4.58-4.48 (m, 4H, OCH_2Ph, H-1Cβ), 4.43 (d, *J*= 12.1 Hz, 1H, OCH_2Ph), 4.38-4.32 (m, 2H, H3A ,OCH_2Ph), 4.26-4.19 (m, 3H, OCH_2Ph, H-4B, H-2A), 4.14-4.06 (m, 4H, H-2C, H-3B, H-2B, H-6_aC), 4.03-3.99 (m, 1H, H-4A), 3.89 (t, *J*= 9.5 Hz, 1H, H-4C), 3.70-3.62 (m, 2H, O*CH*H$(CH_2)_4N_3$, H-6aB), 3.55 (d, *J*= 10.6 Hz, 1H, H-6bB), 3.52-3.35 (m, 5H, H-5Xb, H-6_bC, H-3C, H-6aA, H-6bA), 3.34-3.28 (m, 2H, H-5A, H-5B), 3.27-3.21 (m, 1H, O*CH*H$(CH_2)_4N_3$), 3.12-3.05 (m, 1H,H-5C), 2.93-2.80 (m, 2H, -CH_2N_3), 2.10 (s, 3H, CH_3CO), 2.02, 1.99 (2 s, 2x3H, CH_3CO), 1.41-1.28 (m, 4H, OCH$_2$CH_2CH$_2$CH_2CH$_2$N$_3$), 1.12-1.01 (m, 2H, O$(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 170.1, 169.5, 169.1, 167.7, 138.8-131.8, 129.9-123.3, 102.0, 101.7, 99.4, 98.2, 97.2, 80.6, 78.4, 75.9, 76.0, 75.8, 75.2, 74.8, 74.6, 74.5, 73.5, 73.4, 72.8, 72.1, 69.1, 69.0, 68.7, 68.5, 68.4, 68.3, 68.3, 67.4 60.3, 56.7, 55.9, 51.2, 28.8, 28.4, 23.1, 21.1- 20.8; **HRMS (ESI) *m/z*** $[M+NH_4]^+$ calculated for $C_{96}H_{97}N_5O_{25}NH_4$ 1737.6811, found 1737.6862.

5-azidopentyl (2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl)-(1→2)-(4,6-*O*-benzylidene-β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (11): Tetrasaccharide **4** (110 mg, 63.9 µmol) was dissolved in a mixture of CH_2Cl_2:MeOH (0.6 ml, 4:1) and treated with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) (35 mg, 154 µmol). After 90 min. of stirring at room temperature, the reaction was quenched with saturated $NaHCO_3$ and extracted with EtOAc (3x). The combined organic layers were washed with saturated $NaHCO_3$, water, saturated NaCl, dried over $MgSO_4$ and concentrated. The crude product was purified by flash column chromatography (SiO_2, hexane:EtOAc 4:2 to 2:2) to afford the title compound (60 mg, 60%) as a white solid**; $[\alpha]^{20}_{589}$**= -8.6^o , (c, 0.5, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.85-7.28 (m, 23H, Ar), 7.04-6.78 (m, 10H, Ar), 5.41 (s, 1H, OC*H*Ph), 5.28 (d, *J*= 8.1 Hz, 1H, H-1Aβ), 5.24 (t, 1H, *J*=8.4 Hz, H-3X), 5.02-4.95 (m, 2H, H-2X, H-4X), 4.94-4.92 (m, 1H, H-1Bβ), 4.86 (d, 1H, *J*=12.8 Hz, OCH*H*Ph), 4.81– 4.78 (m, 2H, $J_{1\text{-}2}$=6.7 Hz, *J*=11.9 Hz, OCH*H*Ph, H-1Xβ), 4.66 (s, 1H, H-1Cβ), 4.61-4.44 (m, 5H, 3x OCH_2Ph), 4.32 (d, 1H, *J*=12.2Hz, OCH*H*Ph), 4.29-4.20 (m, 4H, H-4A, H-3B, H-5$_a$X, H-2A), 4.15-4.11 (m, 2H, H-2B, H-4B), 4.05-4.00 (m, 2H, H-6$_a$C, H-3A), 3.91 (d, *J*= 2.7 Hz, 1H, H-2C), 3.70-3.65 (m, 2H, H-6aA, OCH*H*$(CH_2)_4N_3$), 3.59-3.47 (m, 4H, H-6aB, H-6bA, H-4C, H-3C), 3.38-3.35 (m, 3H, H-5X, H-6$_b$C, H-6bB), 3.34-3.28 (m, 2H, H-5A, H-5B), 3.27-3.21 (m, 1H, OCH*H*$(CH_2)_4N_3$), 3.12-3.05 (m, 1H, H-5C), 2.93-2.80 (m, 2H, -CH_2N_3), 2.10, 2.02, 1.99 (3 s, 3x3H, CH_3CO), 1.41-1.28 (m, 4H,$OCH_2CH_2CH_2CH_2CH_2N_3$), 1.12-1.01 (m, 2H, $O(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 170.0, 169.9, 169.4, 138.8, 138.6, 137.7, 137.3, 134.0, 131.5, 129.2-126.5, 123.3, 102.2, 102.0, 101.0, 98.3, 97.2, 80.9, 79.0, 78.9, 76.1, 74.7, 74.6, 74.5, 73.7, 72.7, 71.1, 70.8, 70.4, 69.0, 68.6, 68.5, 68.3, 68.1, 67.5, 62.1, 56.6, 55.9, 51.2, 29.8, 29.8, 29.7, 28.8, 28.4, 23.1, 21.1, 20.9, 20.9; **HRMS (ESI) *m/z*** $[M+NH_4]^+$ calculated for $C_{96}H_{97}N_5O_{25}$ 1737.6811, found 1737.6862.

5-azidopentyl (3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→2)-(3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl)-(1→3)-[2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-(4,6-*O*-benzylidene-β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (12): A mixture of **11** (55 mg, 34.79 µmol) and **7** (50 mg, 57.6 µmol) was dried under high vacuum overnight, then dissolved in anhydrous CH_2Cl_2 (1.1 mL) and freshly activated ground molecular sieves (3Å) were added. After 1h of stirring at room temperature, the mixture was cooled to -40°C and TMSOTf (1 µL, 5.52 µmol) was added. After 1h the reaction mixture was quenched with Et_3N, filtered through a pad of Celite® and concentrated. The crude product was purified by flash column chromatography (hexane: EtOAc, 3:2 to 2:3) to afford the title compound as a white foam (63 mg, 82%); **$[\alpha]^{20}_{589}$**= -21 (c, 0.5, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.86-7.53 (m, 12H, Ar), 7.77-7.29 (m, 15H, Ar), 7.02-6.79 (m, 10H, Ar), 5.49 (dd, $J_{3\text{-}2}$= 10.9 Hz, $J_{3\text{-}4}$= 9.3 Hz, 1H, H-3G), 5.36 (s, 1H, OC*H*Ph), 5.31 (d, $J_{1\text{-}2}$= 8.4 Hz, 1H, H-1Aβ), 5.09-5.06 (m, 2H, H-3M, H-3X), 5.01 (t, *J* = 10.01, *J* =11.1 Hz, 1H, H-4M), 4.98-4.95 (m, 2H, H-4G, H-2X), 4.95-4.87(m, 5H, H-1Bβ, H-1Mα, H-1Xβ, OCH*H*Ph, H-1Gβ), 4.85-4.80 (m, 2H, H-4X, OCH*H*Ph), 4.66 (d, J=11.8 Hz,1H, OCH_2Ph), 4.57-4.48 (m, 3H, OCH_2Ph), 4.43 (s, 1H, H-1Cβ), 4.41-4.36 (m, 2H, H-5_aX, H-3A), 4.31-4.24 (m, 3H, OCH_2Ph, H-2G), 4.24-4.18 (m, 2H, H-2A, H-4B), 4.14-4.09 (m, 2H, H-2B, H-3B), 4.07-4.02 (m, 3H, H-6_aC, H-2M, H-4A), 3.95 (dd, $J_{6a\text{-}6b}$= 12.4 Hz, $J_{6a\text{-}5}$= 3.7 Hz, 1H, H-6_aG), 3.90-3.87 (m, 1H, H-5M), 3.79 (d, *J*= 3.4 Hz, 1H, H-2C), 3.72-3.63 (m, 6H, H-6_bG, OCH*H*$(CH_2)_4N_3$, H-6_aA, H4C, H-$6_{a,b}$M), 3.59 (dd, $J_{5a\text{-}5b}$= 12.4 Hz, $J_{5b\text{-}4}$= 4.2 Hz, 1H, H-5_bX), 3.55-3.50 (m, 2H, H-6_bA, H-6_aB), 3.45-3.41 (m, 2H, H-3C, H-6_bB), 3.36-3.29 (m, 3H, H-5A, H-5B, H-6_bC), 3.27– 3.22 (m, 1H, OCH*H*$(CH_2)_4N_3$), 2.94– 2.90 (m, 1H, H-5C), 2.89-2.81 (m, 2H, -*C*H_2N_3), 2.11-2.10 (m, 1H, H-5G), 2.07, 2.05, 2.03, 1.98, 1.97, 1.94, 1.87, 1.86 (9 s, 9x3H, C*H*$_3$CO), 1.38-1.29 (m, 4H, $OCH_2CH_2CH_2CH_2CH_2N_3$), 1.10-1.03 (m, 2H, $O(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 170.7, 170.5, 170.4, 170.3, 169.8, 169.5, 169.3, 169.0, 138.9, 138.8, 138.6, 137.9, 137.6, 134.35, 133.8, 131.5, 130.3, 129.1, 128.9, 128.7, 128.7, 128.6, 128.3, 128.2, 128.0, 127.8, 127.5, 127.5, 127.1, 127.0, 123.7, 123.3, 102.5, 101.3, 99.6, 98.4, 98.2, 97.2, 95.8, 79.8,

78.4, 76.6, 76.1, 74.6, 74.6, 73.6, 72.8, 72.7, 71.2, 70.5, 69.2, 69.0, 68.8, 68.7, 68.5, 68.3, 67.9, 66.7, 66.2, 65.5, 63.2, 61.1, 60.3, 56.6, 55.8, 54.2, 51.2, 29.8, 28.8, 28.4, 23.1, 21.1, 21.0, 20.9, 20.8, 20.8, 20.7, 20.6; **HRMS (MALDI-Tof) *m/z*** $[M+Na]^+$ calculated for $C_{117}H_{124}N_6O_{42}$ 2307.7644, found 2307.7554.

5-azidopentyl [2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-[2-*O*-acetyl-3,4,6-tri-*O*-benzyl-α-D-mannopyranosyl-(1→3)]-(4,6-O-benzylidene- β-D-mannopyranosyl)-(1→4)-(3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (13): A mixture of **11** (43mg, 27.2 µmol) and **5** (25mg, 37.6 µmol) was dried under high vacuum overnight. The mixture was dissolved in anhydrous CH_2Cl_2 (0.6 mL) and freshly activated molecular sieves (3Å) were added. After stirring for 40 minutes at room temperature, the mixture was cooled to -20°C and 0.1 eq of trimethylsilyl trifluoromethanesulfonate TMSOTf (10% in CH_2Cl_2, 4.5 µL, 2.75 µmol) was slowly added. After stirring for 45 min the suspension was quenched with Et_3N, filtered through a pad of Celite® and concentrated. The crude product was purified by flash column chromatography (SiO_2; hexane:EtOAc, 4:2 to 3:2) to afford the pentasaccharide as a white solid (52 mg, 94%); **$[\alpha]^{20}_{589}$**= -5 (c, 0.5, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.91 -7.66 (m, 12H, Phth), 7.44-7.16 (m, 30H, Ar), 7.18-6.76 (m, 10H, Ar), 5.60-5.55 (m, 1H, H-2M), 5.42 (s, 1H, OC*H*Ph), 5.28 (d, 1H, $J_{1\text{-}2}$= 8.3 Hz, H-1Aβ), 5.24-5.20 (m, 1H, H-1Mα), 5.09 (t, $J_{3,2/4}$ = 6.6 Hz, 1H, H-3X), 5.03-4.97 (m, $J_{2\text{-}3}$ = 6.6Hz, $J_{2\text{-}1}$= 4.4 Hz, 1H, H-2X), 4.95 (d, 1H, $J_{1\text{-}2}$ = 4.4 Hz, H-1Xβ), 4.93-4.88 (m, 2H, H-1B, CH_2O), 4.84 (dd, *J*= 12.5 Hz, J=5.3 Hz, 2H, OCH_2Ph), 4.80-4.74 (q, 1H, $J_{4\text{-}3}$= 6.6 Hz, $J_{4\text{-}5b}$= 5.7 Hz, H-4X), 4.72 (d, *J*= 11.3 Hz, 1H, OCH*H*Ph), 4.63 (d, 2H, OCH*H*Ph, H-1Cβ), 4.59-4.44 (m, 8H, 4xOCH_2Ph), 4.41-4.30 (m, 3H, OCH_2Ph, H-3A), 4.26 (dd, $J_{5a\text{-}5b}$= 12.6 Hz, $J_{5a\text{-}4}$= 3.9 Hz, 1H, H-5Xa,) 4.23-4.19 (m, 2H, H-2A, H-4B), 4.11-4.09 (m, 2H, H-2B, H-3B), 4.05-3.96 (m, 5H, H-6_aC, H-2C,H-4A, H-3C, H-3M), 3.87-3.72 (m, 4H, H-4M, H-4C,H-5M, H-6_aM), 3.69-3.63 (m, 3H, H-6_bM, H-6_aA, OCH*H*$(CH_2)_4N_3$), 3.54 (d, *J*= 10.9 Hz, 1H, H-6_bA), 3.47-3.39 (m, 2H, H-6abB), 3.38-3.27 (m, 4H, H-6_bC,H-5bX, H-5A, H-5B), 3.27-3.21 (m, 1H, OCH*H*$(CH_2)_4N_3$), 3.05-2.99 (m, $J_{5\text{-}6}$= 9.5 Hz, $J_{5\text{-}4}$= 4.7 Hz, 1H, H-5C), 2.92-2.82 (m, 2H,-

C*H*$_2$N$_3$), 2.11, 2.03, 2.01, 1.93(4 s, 4x3H, C*H*$_3$CO), 1.37-1.27 (m, 4H, OCH$_2$*CH$_2$* CH$_2$*CH$_2$* CH$_2$N$_3$), 1.10-1.03 (m, 2H, OCH$_2$CH$_2$ *CH$_2$*CH$_2$ CH$_2$N$_3$); **^{13}C NMR (126 MHz, CDCl$_3$) δ** 170.1, 169.7, 169.6, 169.2, 168.6, 167.7, 138.8, 138.7, 138.5, 138.3, 138.1, 137.8, 137.2, 133.7, 131.8, 131.5, 128.9-127.0, 126.0, 123.8, 123.3, 102.1, 101.2, 99.7, 99.1, 98.2, 97.1, 80.6, 78.2, 77.8, 76.7, 76.0, 75.2, 74.7, 74.6, 74.6, 74.5, 74.5, 74.3, 73.6, 73.5, 73.4, 72.7, 72.1, 71.7, 69.6, 69.1, 69.0, 68.5, 68.4, 68.3, 68.3, 68.1, 67.2, 60.7, 56.6, 55.8, 51.2, 28.8, 28.4, 23.1, 21.2-21.0; **HRMS (MALDI-Tof) *m/z*** [M+Na]$^+$ calculated for C$_{114}$H$_{119}$N$_5$O$_{31}$ 2076.7781, found 2076.7880.

5-azidopentyl (3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→2)-(3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl)-(1→3)-[2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-(β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (14): Hexasaccharide **12** (69 mg, 30.18 µmol) was dissolved in dry CH$_2$Cl$_2$ (0.32 mL) and EtSH (13.5 µL) and a 10% solution of BF$_3$·Et$_2$O in CH$_2$Cl$_2$ (7.4 µL, 0.2 eq) were added. After 2h of stirring at room temperature, the reaction mixture was quenched with Et$_3$N and concentrated. The crude product was purified by flash column chromatography (SiO$_2$, hexane: EtOAc, 2:1 to 1:2) to afford the title compound as a white solid (52 mg, 79%); **[α]$^{20}_{589}$**= -3.5 , (c, 0.5, CHCl$_3$); **^{1}H NMR (500 MHz, CDCl$_3$) δ** 7.86-7.55 (m, 12H, Ar), 7.38-7.23 (m, 10H, Ar), 7.06-6.78 (m, 10H, Ar), 5.78 (dd, $J_{3\text{-}2}$= 10.75 Hz, $J_{3\text{-}4}$= 9.21 Hz, 1H, H-3G), 5.42 (d, $J_{1\text{-}2}$= 8.4 Hz, 1H, H-1Gβ), 5.27 (d, $J_{1\text{-}2}$= 8.5 Hz, 1H, H-1Aβ), 5.19-5.09 (m, 3H, H-4G, H-3X, H-4M), 5.05-5.02 (m, 2H, H-2X, H-3M), 4.95 (d, 1H, $J_{1\text{-}2}$= 1.4Hz, H-1Mα), 4.93-4.84 (m, 5H, H-1Bβ, H-4X, H-1Xβ, OC*H$_2$*Ph), 4.59-4.48 (m, 5H, 2xOC*H$_2$*Ph, H-1Cβ), 4.43-4.37 (m, 4H, OC*H$_2$*Ph, H-2G, H-3A), 4.32 (dd, 1H, $J_{6a\text{-}6b}$= 12.4Hz, $J_{6a\text{-}5}$= 4.8Hz, H-6$_a$G), 4.28 (t, 1H, $J_{2\text{-}4}$= 3.6Hz, $J_{2\text{-}3}$= 2.4 Hz, H-2M), 4.24-4.17 (m, 3H, H-2A, H-5$_a$X, H-4B), 4.14-4.08 (m, 3H, H-2B, H-3B, H-6$_b$G), 3.99-3.97 (m, 1H, H-4A), 3.95-3.90 (m, 2H, H-5M, H-5G), 3.84 (d, 1H, *J*= 2.9Hz, H-2C), 3.78 (dd, 1H, $J_{6a\text{-}6b}$= 12.1Hz, $J_{6a\text{-}5}$= 2.6 Hz , H-6$_a$M), 3.73-3.64 (m, 4H, H-6$_b$M, H-4C, H-6$_a$A, OCH*H*(CH$_2$)$_4$N$_3$), 3.61 (dd, 1H, $J_{6a\text{-}6b}$= 12.0 Hz, $J_{6a\text{-}5}$= 2.7 Hz, H-6$_a$C), 3.54-3.46 (m, 3H, H-6$_b$A, H-6$_a$B, H-6$_b$C), 3.43-3.39 (m, 2H, H-6$_b$B, H-5$_b$X), 3.33– 3.29 (m, 2H, H-5A, H-5B), 3.26– 3.19 (m,

2H, OCH*H*$(CH_2)_4N_3$, H-3C), 2.96-2.93 (m, 1H, H-5C), 2.92-2.81 (m, 2H, -*CH*$_2N_3$), 2.11, 2.07, 2.04, 2.02, 1.94, 1.91, 1.86 (9 s, 9x3H, C*H*$_3$CO), 1.35-1.29 (m, 4H, OCH_2*CH*$_2CH_2$*CH*$_2CH_2N_3$), 1.09-1.04 (m, 2H, $O(CH_2)_2$*CH*$_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 170.9, 170.6, 170.5, 170.3, 169.8, 169.6, 169.5, 138.8, 138.5, 137.7, 134.4, 134.1, 133.8, 131.8, 131.5, 128.8-127.0, 123.7, 123.3, 101.2, 99.8, 98.2, 97.8, 97.1, 96.9, 80.1, 79.9, 76.5, 76.1, 75.9, 74.7, 74.6, 74.6, 74.3, 74.2, 73.5, 72.7, 71.9, 70.7, 70.3, 69.9, 69.9, 69.1, 69.0, 68.7, 68.6, 68.3, 67.7, 66.3, 66.1, 63.1, 62.3, 62.2, 61.1, 56.6, 55.8, 54.6, 51.2, 29.8, 28.8, 28.4, 23.1, 21.1-20.6**; HRMS (MALDI-Tof) *m/z*** [M+ Na]$^+$ calculated for $C_{110}H_{120}N_6O_{42}$ 2219.7331, found 2219.7351.

5-azidopentyl [2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-[2-*O*-acetyl-3,4,6-tri-*O*-benzyl-α-D-mannopyranosyl-(1→3)]-(β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (15): Pentasaccharide **13** (51 mg, 24.81 µmol) was dissolved in dry CH_2Cl_2 (0.24 mL) and EtSH (10 µL, 5.7 eq) and a 10% solution of $BF_3.Et_2O$ in CH_2Cl_2 (8.5 µL, 0.27 eq) were added. After 3h of stirring at room temperature the reaction mixture was quenched with Et_3N and concentrated. The crude product was purified by flash column chromatography (SiO_2, hexane: EtOAc, 2:1 to 1:1) to afford the title compound as a white solid (42 mg, 86%); **$[\alpha]^{20}_{589}$**= +0.8 ,(c, 0.35, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.84 -7.64 (m, 6H, Ar), 7.39-7.24 (m, 25H, Ar), 7.19-7.17 (m, 2H, Ar), 7.04-7.02 (m, 2H, Ar), 6.99-6.97 (m, 2H, Ar), 6.92-6.86 (m, 3H, Ar), 6.78-6.767 (m, 2H, Ar), 5.57 (dd, 1H, $J_{2\text{-}3}$= 3.2 Hz, $J_{2\text{-}1}$= 1.8 Hz, H-2M), 5.43 (d, 1H, $J_{1\text{-}2}$= 1.8 Hz, H-1Mα), 5.26 (d, $J_{1\text{-}2}$= 8.3 Hz, 1H, H-1Aβ), 5.21 (t, 1H, $J_{3\text{-}4/2}$= 8.4 Hz, H-3X), 5.04 (dt, 1H, $J_{4\text{-}3}$ = 8.4 Hz, $J_{4\text{-}5a}$ = 5.1 Hz, H-4X), 4.95 (d, 1H, $J_{1\text{-}2}$= 6.6 Hz, H-1Xβ), 4.94-4.91 (m, 1H, H-1Bβ), 4.90-4.77 (m, 5H, H-2X, 2x OC*H*$_2$Ph), 4.69 (d, 1H, $J_{1\text{-}2}$= 8.4 Hz, H-1Bβ), 4.56-4.41 (m, 10H, 5xOC*H*$_2$Ph), 4.28 (dd, *J*= 10.6 Hz, *J*= 8.6 Hz, 1H, H-3A), 4.22-4.10 (m, 7H, H-5X$_a$, H-5M, H-2A, H-2C, H-4B, H-3B, H-2B), 4.0-3.94 (m, 2H, H-3M, H-4A), 3.85 (t, 1H, $J_{4\text{-}3/5}$= 9.7 Hz, H-4C), 3.74 (dd, 1H, $J_{6a\text{-}6b}$= 10.3 Hz, $J_{6a\text{-}5}$= 1.7 Hz, H-6M$_a$), 3.68-3.57 (m, 5H, H-6C$_a$, H-4M, OCH*H*$(CH_2)_4N_3$, H-6M$_b$, H-6Ba), 3.53-3.46 (m, 3H, H-6A$_{ab}$, H-6C$_b$), 3.44 (dd, 1H, $J_{3\text{-}}$

$_4$= 9.7 Hz, J_{3-2}=3.1 Hz, H-3C), 3.40 (dd, 1H, J_{6a-6b}= 10.9 Hz, J_{6-5}= 4.1 Hz, H-6Bb), 3.31-3.22 (m, 4H, H-5A, H-5B, H-5X$_b$, OCH*H*$(CH_2)_4N_3$), 3.08-3.04 (m, 1H, H-5C), 2.91-2.81 (m, 2H, -*CH*$_2N_3$), 2.17, 2.07, 2.01, 1.79 (4 s, 4x3H, C*H*$_3$CO), 1.37-1.24 (m, 4H, OCH$_2$*CH*$_2$ CH$_2$*CH*$_2$ CH$_2N_3$), 1.19-1.03 (m, 2H, OCH$_2$CH$_2$*CH*$_2$CH$_2$CH$_2N_3$); **^{13}C NMR (126 MHz, CDCl$_3$) δ** 170.4, 170.0, 169.8, 169.6, 138.7, 138.6, 138.5, 138.3, 138.1, 137.7, 134.1, 133.9, 133.8, 131.8, 131.5, 128.7-127.0, 123.7, 123.3, 101.9, 99.5, 98.2, 97.1, 96.9, 80.5, 79.7, 77.9, 76.1, 76.1, 75.3, 75.1, 74.8, 74.7, 74.7, 74.6, 74.6, 73.6, 73.5, 72.7, 71.8, 71.4, 71.0, 70.8, 69.7, 69.1, 69.0, 68.6, 68.3, 67.8, 62.9, 62.1, 56.6, 55.8, 51.2, 29.8, 28.8, 28.4, 23.1, 21.2, 21.0, 20.8, 20.7; **HRMS (MALDI-Tof) *m/z*** [M+ Na]$^+$ calculated for $C_{107}H_{115}N_5O_{31}$ 1988.7468, found 1988.7417.

5-azidopentyl (3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→2)-(3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl)-(1→3)-[2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-[(3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→2)-(3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl)-(1→6)]-(β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (16): A mixture of **14** (28 mg, 12.73 µmol) and **7** (14.3 mg, 16.5 µmol) was dried under high vacuum overnight. The mixture was dissolved in anhydrous CH_2Cl_2 (3 mL, 250 mL/mmol) in the presence of freshly activated ground 3Å MS. After 1h at room temperature, the mixture was cooled to -45^{o}C and TMSOTf (3.4 µL, 0.15 eq, 1.9 µmol) was slowly added. After 2h the reaction mixture was quenched with Et_3N, filtered through a pad of Celite and concentrated. The crude product was purified by flash chromatography (SiO_2, hexane: acetone, 3:2) to afford the title compound as a wihte solid (25 mg, 70%); **$[\alpha]^{20}_{589}$**= +8.2 , (c, 0.5, $CHCl_3$); **^{1}H NMR (500 MHz, CDCl$_3$) δ** 5.76-5.69 (m, 2H, H-3G, H-3G'), 5.45 (d, 1H, J_{1-2} = 8.5 Hz, H-1Gβ), 5.29 (d, 1H, J_{1-2}= 8.8 Hz, H-1G'β), 5.25 (d, 1H, J_{1-2}= 8.4 Hz, H-1Aβ), 5.20-5.11 (m, 5H, H-4M, H-4G, H-4M', H-4G', H-3X), 5.07 (dd, 1H, J_{3-4}= 10.1 Hz, J_{3-2}=3.3 Hz, H-3M), 4.96-4.86 (m, 6H, H-3M',H-2X, H-1Bβ, H-1Mα, H-1Xβ, H-4X), 4.84 (d, 1H, *J* = 12.8 Hz, OCH*H*Ph), 4.72 (d, 1H, *J* = 12.5 Hz, OCH*H*Ph), 4.58 (d, 1H, *J*= 12.0 Hz, OC*H*$_2$Ph), 4.52-4.43 (m, 4H, OC*H*$_2$Ph, H-1Cβ), 4.41-4.30 (m, 7H, H-1M'α, H-2G, H-2G', OC*H*$_2$Ph, H-2C, H-6Ga) 4.27-

4.21 (m, 3H, H-6G'a, H-5Xa, H-3A), 4.17-4.08 (m, 6H, H-4B, H-2M', H-3B, H-2A, H-2B, H-6Gb), 3.97-3.93 (m, 4H, H-5G, H-5Xb, H-4A, H-5M), 3.83-3.71 (m, 5H, H-6Mab, H-6M'ab, H-2C), 3.68-3.55 (m, 6H, H-5G', H-5M', H-6Cab, H-4C, OCH*H*$(CH_2)_4N_3$), 3.51-3.36 (m, 4H, H-6Aab, H-6Ba, H-5Xb), 3.29-3.21 (m, 4H, H-5A, H-5B, H-6Bb, OCH*H*$(CH_2)_4N_3$), 3.15 (dd, 1H, J_{3-4}= 9.6 Hz, J_{3-2}= 3.0 Hz, H-3C), 3.04-3.00 (m, 1H, H-5C), 2.91-2.81 (m, 2H, -CH_2N_3), 2.08 , 2.06 (2 s, 2x3H, CH_3CO), 2.04-2.03 (3 s, 3x3H, CH_3CO), 2.00-1.96 (m, 21H, CH_3CO), 1.84 (s, 6H, CH_3CO), 1.81 (s, 3H, CH_3CO), 1.37– 1.27 (m, 4H, $OCH_2CH_2CH_2CH_2CH_2N_3$), 1.09-1.02 (m, 2H, $O(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 171.0-168.4, 138.8, 138.6, 137.8, 134.4, 133.8, 131.5, 128.8-127.0, 123.7, 123.3, 101.2, 99.6, 99.1, 98.2, 98.0, 97.4, 97.2, 79.9, 79.6, 76.0, 75.4, 74.7-74.4, 73.6, 73.5, 72.7, 71.8, 71.7, 70.8, 70.2, 70.1, 69.9, 69.9, 69.7, 69.4, 69.0, 68.9, 68.7, 68.2, 67.7, 66.1, 65.6, 62.9, 62.4, 61.9, 61.7, 61.1, 56.7, 55.8, 54.7, 54.6, 51.2, 29.8, 28.8, 28.4, 23.1, 20.8-20.7; **HRMS (MALDI-Tof) *m/z*** $[M+Na]^+$ calculated for $C_{142}H_{155}N_7O_{59}$ 2924.9236, found 2924.9144.

5-azidopentyl (3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→2)-(3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl)-(1→6)-[2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-[2-*O*-acetyl-3,4,6-tri-*O*-benzyl-α-D-mannopyranosyl)-(1→3)]-(β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (17): A mixture of **15** (40 mg, 20.33 µmol) and **7** (21 mg, 24.19 µmol) was dried under high vacuum overnight. The mixture was dissolved in anhydrous CH_2Cl_2 (5 mL, 250 mL/mmol) in the presence of freshly activated ground 3Å MS. After stirring for 1h, the reaction mixture was cooled to -45 ^oC and TMSOTf (7 µL, 0.2 eq, 4.06 µmol) was slowly added. After 2h of stirring the suspension was quenched with Et_3N, filtered through a pad of Celite® and concentrated. The crude product was purified by flash chromatography (SiO_2, hexane: acetone, 3:2) to afford the tittle compound as a white solid (40 mg, 73%); **$[\alpha]^{20}_{589}$**= -0.6 (c, 0.35, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.79 -7.55 (m, 12H, Ar), 7.38-7.27 (m, 24H, Ar), 7.16-7.15 (m, 2H, Ar), 7.04-7.03 (m, 2H, Ar), 6.94-6.93 (m, 2H, Ar), 6.81-6.74(m, 5H, Ar), δ 5.66 (t, 1H, *J* = 9.9

Hz, H-3G), 5.59 (s, 1H, H-2M), 5.36 (s, 1H, H-1Mα), 5.24 (t, 2H, *J*= 7.9 Hz, H-1Aβ, H-1Gβ), 5.19 (t, *J*= 8.4 Hz, H-3X), 5.14 (t, 1H, *J*= 9.9 Hz, H-4M), 5.08 (t, 1H, *J*= 9.7 Hz, H-4G), 5.02-4.99 (m, 1H, H-4X), 4.98-4.93 (m, 2H, H-3M', H-1Xβ), 4.91-4.83 (m, 4H, H-2X, H-1Bβ, OC*H*$_2$Ph), 4.78 (d, 1H, *J*= 11.3 Hz, OCH*H*Ph), 4.67 (d, 1H, *J*= 12.6 Hz, OCH*H*Ph), 4.59-4.54 (m, 5H, H-C1β, 2xOC*H*$_2$Ph), 4.51-4.33 (m, 8H, H-1M'α, H-2G, 3xOC*H*$_2$Ph), 4.23-4.06 (m, 10H, H-5Xa, H-6aG, H-3A, H-5M, H-2A, H-2B, H-2M', H-2C, H-3B, H-4B), 4.00 (dd, 1H, $J_{3\text{-}4}$= 9.9Hz, $J_{3\text{-}2}$= 3.4 Hz, H-3M), 3.95 (t, 1H, *J*=9.0 Hz, H-4A), 3.87 (d, 1H, *J*= 12.0 Hz, H-6Gb), 3.80 (dd, 1H, $J_{6a\text{-}6b}$= 12.2 Hz, $J_{6a\text{-}5}$= 4.3 Hz, H-6M'a), 3.72-3.61 (m, 8H, H-4C, H-4M, H-5M', H-6M'b, OCH*H*(CH$_2$)$_4$N$_3$, H-6Mab, H-6Aa), 3.56-3.35 (m, 7H, H-5G, H-3C, H-6Ab, H-6Bab, H-6Cab), 3.30-3.22 (m, 4H, H-5A, H-5B, H-5Xb, OCH*H*(CH$_2$)$_4$N$_3$), 3.13-3.10 (m, 1H, H-5C), 2.91-2.80 (m, 2H, -*CH*$_2$N$_3$), 2.16, 2.05, 2.02, 2.00, 1.99, 1.98, 1.93, 1.92, 1.82, 1.71 (10 s, 10x3H, C*H*$_3$CO), 1.41-1.29 (m, 4H, OCH$_2$*CH*$_2$CH$_2$*CH*$_2$CH$_2$N$_3$), 1.107-1.03 (m, 2H, OCH$_2$CH$_2$*CH*$_2$CH$_2$CH$_2$N$_3$); **^{13}C NMR (126 MHz, CDCl$_3$) δ** 170.9-167.4, 138.8, 138.5, 138.1, 138.0, 134.4, 134.0, 133.7, 131.9, 131.7, 131.5, 128.7-127.0, 123.6, 123.3, 101.8, 99.6, 98.2, 97.8, 97.7, 97.2, 80.7, 79.5, 77.8, 76.1, 75.1, 75.1, 74.8, 74.7, 74.6, 74.4, 73.5, 73.4, 72.7, 71.8, 71.7, 71.6, 71.0, 70.7, 70.0, 69.4, 69.1, 69.0, 68.9, 68.7, 68.6, 68.4, 68.3, 68.0, 66.8, 65.6, 62.5, 62.0, 61.7, 56.5, 55.8, 54.5, 51.2, 29.8, 28.8, 28.4, 23.1, 21.2-20.6; **HRMS (ESI) *m/z*** [M+Na]$^+$ calculated for $C_{139}H_{150}N_6O_{48}$ 2693.9373, found 2693.9343.

5-azidopentyl (3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→2)-(3,4,6-tri-*O*-acetyl-α-D-mannopyranosyl)-(1→3)-[2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-[(2,3,4,6-tetra-*O*-benzoyl-α-D-mannopyranosyl)-(1→6)]-(β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (20): A mixture of **14** (21 mg, 9.55 µmol) and **6** (8.5 mg, 11.47 µmol) was dried under high vacuum overnight. The mixture was dissolved in anhydrous CH_2Cl_2 (2.2 mL, 250 mL/mmol) and in the presence of freshly activated ground 3Å MS. After 1h at room temperature, the mixture was cooled to -45 ºC and TMSOTf (6.5 µL, 0.4 eq, 3.82 µmol) was added. After 5 hours stirring at -45 ºC, the

reaction mixture was quenched with Et_3N, filtered through a pad of Celite® and concentrated. The crude product was purified by preparative HPLC (acetonitrile:H_2O; initial condition 85:15; 2 min 90:10, 35min 99:1, 36min 85:15) to afford the tittle compound as a white solid (12.4 mg, 47%); **$[\alpha]^{20}_{589}$**= -8.8 (c, 0.52, $CHCl_3$)**; ^{1}H NMR (500 MHz, $CDCl_3$) δ** 8.12–8.11 (m, 2H, Ar), 7.99-7.97(m, 2H, Ar), 7.91-7.89 (m, 2H ,Ar), 7.83-7.71 (m, 5H, Ar), 7.66-7.61(m, 5H, Ar), 7.57-7.47 (m, 6H, Ar), 7.43-7.39 (m, 3H, Ar), 7.35-7.28 (m, 12H, Ar), 7.24-7.21 (m, 4H, Ar), 7.05 (d, *J*= 7.9 Hz, 2H, Ar), 6.98-6.96 (m, 2H, Ar), 6.83 (t, J=7.92Hz, 2H, Ar), 6.78-6.76 (m, 3H, Ar), 6.74-6.71 (m, 1H, Ar), 6.09 (t, *J*= 10.0 Hz, H-4M'), 5.83-5.79 (m, 2H, H-3M', H-3G), 5.66-5.64 (m, 1H, H-2M'), 5.49 (d, $J_{2\text{-}1/3}$= 8.5 Hz, 1H, H-1Gβ), 5.27 (d, $J_{1\text{-}2}$= 8.4 Hz, 1H, H-1Aβ), 5.21-5.08 (m, 4H, H-3X, H-4M, H-4G, H-3M), 5.04-5.02 (m, *3*H, H-2X, H-1M, H-1M'), 4.95-4.90 (m, 4H, H-1Xβ, H-1Bβ, H-4X, OCH*H*Ph), 4.87 (d, *J*= 12.9 Hz, 1H, OCH*H*Ph), 4.80-4.77 (m, 1H, H-6M'a), 4.61-4.33 (m, 12H, 6xOCH_2Ph, H-1Cβ, H-5M', H-2G, H-2M, H-3A, H-6M'b), 4.25-4.20 (m, 3H, H-2A, H-6Ga, H-5Xa), 4.18-4.09 (m, 3H, H-4B, H-3B, H-2B), 4.04 (t, *J*= 9.2 Hz, 1H, H-4A), 3.99- 3.96 (m, 2H, H-5M, H-6Gb), 3.91-3.84 (m, 3H, H-6Ca, H-2C, H-5G), 3.80-3.64 (m, 6H, H-6Cb, H-4C, H-6Mab, H-6Aa, OCH*H*$(CH_2)_4N_3$), 3.56-3.47 (m, 2H, H6Ab, H6Ba), 3.42-3.33 (m, 2H, H6Bb, H5Xb), 3.35-3.22 (m, 4H, OCH*H*$(CH_2)_4N_3$, H-3C, H-5A, H-5B), 3.18-3.15 (m, 1H, H-5C) 2.90-2.81 (m, 2H, CH_2N_3), 2.06, 2.04, 2.03, 2.015, 2.01, 1.93, 1.91, 1.85 (9 s, 3x3H, CH_3CO), 1.41-1.29 (m, 4H, $OCH_2CH_2CH_2CH_2CH_2N_3$), 1.06-1.05 (m, 2H, $OCH_2CH_2CH_2CH_2CH_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 170.9, 170.6, 170.5, 170.3, 169.9, 169.8, 169.6, 169.5, 169.2, 166.5, 165.6, 165.5, 165.2, 138.8, 138.7, 138.5, 137.8, 134.3, 133.7, 133.4, 133.4, 133.2, 131.6, 130.0, 129.9 129.9, 129.4, 129.3, 129.2, 128.9, 128.8, 128.6, 128.6, 128.5, 128.3, 128.3, 128.2, 128.1, 128.0, 128.0, 128.0, 127.8, 127.5, 127.4, 127.1, 127.0, 101.9, 100.1, 98.8, 98.3, 98.2, 97.2, 97.0, 80.5, 79.3, 76.1, 75.9, 74.8, 74.7, 74.6, 74.5, 74.2, 73.4, 72.7, 71.8, 70.9, 70.7, 70.5, 70.4, 70.3, 70.2, 69.9, 69.4, 69.0, 69.0, 68.7, 68.3, 67.9, 67.8, 66.7, 66.2, 63.1, 62.4, 62.0, 61.4, 56.7, 55.8, 54.7, 51.2, 32.0, 28.8, 28.4, 23.1, 20.9-20.6; **HRMS (ESI) *m/z*** $[M+ Na]^+$ calculated for $C_{144}H_{146}N_6O_{51}Na$ 1410.4399, found 1410.4358.

5-azidopentyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)-4-*O*-benzyl-3-*O*-(2-naphthylmethyl)- β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (26): A solution of BH_3·THF (310 µL, 310 µmol, 1.0 M in THF) was added to the tetrasaccharide **4** (50 mg, 29.0 µmol) and the reaction mixture cooled to 0°C. *n*-Bu_2BOTf (92 µL of 1.0 M solution in Et_2O, 92 µmol) were then added, and the reaction mixture was stirred at 0°C for 2 h, quenched with Et_3N/MeOH and concentrated *in vacuo*. The crude mixture was coevaporated with MeOH (2 times) and purified by flash column chromatography (SiO_2, toluene:EtOAc, 3:1) to afford the title compound as a white solid (32 mg, 65%)**; $[\alpha]^{20}_{589}$**= + 9.6 (c, 0.52, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.85-7.77 (m, 5H, Ar), 7.73-7.61 (m, 7H, Ar), 7.49-7.45 (m, 3H, Ar), 7.36-7.27 (m, 11H, Ar), 7.14-7.10 (m, 3H, Ar), 7.10-7.07(m, 2H, Ar), 7.02-6.91 (m, 6H, Ar), 6.77-6.75 (m, 3H, Ar), 5.29 (d, $J_{1\text{-}2}$ = 8.3 Hz, 1H, H-1Aβ), 5.09 (d, $J_{1\text{-}2}$ = 3.35 Hz, 1H, H-1Xβ), 5.07 (t, $J_{3\text{-}2/1}$ = 5.06 Hz, 1H, H-3X), 5.01 (q, 1H, $J_{2\text{-}1}$= 3.40 Hz, $J_{2\text{-}3}$=4.82Hz, H-2X), 4.98 (d, *J*= 12.5 Hz, 1H, OCH*H*Ph), 4.94-4.92 (m, 1H, H-1Bβ), 4.90 (d, *J*= 11.0 Hz, 1H, OCH*H*Ph), 4.85-4.82 (m, 3H, OCH_2Ph, H-4X), 4.68 (d, *J*= 11.9 Hz, 1H, OCH*H*Ph), 4.58-4.49 (m, 6H, H-5Xa, H-1Cβ, 2xOCH_2Ph), 4.42 (d, *J*= 12.5 Hz, 2H, OCH_2Ph), 4.38-4.36 (m, 1H, H-3A), 4.28 (d, *J*= 12.1 Hz, 1H, OCH*H*Ph), 4.24-4.17 (m, 2H, H-2A, H-4B), 4.11-4.07 (m, 3H, H-2B, H-3B, H-2C), 3.96 (t, 1H, *J*= 9.3 Hz, H-4A), 3.75-3.54 (m, 5H, H-4C, H-6Ca, H-6Aa, H-6Ba, O*CH*H(CH2)4N3), 3.49-3.38 (m, 5H, H-5Xb, H-6Ab, H-6Bb, H-6Cb, H-3C), 3.33-3.22 (m, 3H, H-5A, H-5B, O*CH*H(CH2)4N3), 3.12-3.08 (m, 1H, H-5C), 2.92-2.82 (m, 2H, -CH_2N_3), 2.11, 2.04, 1.99 (3 s, 3x3H, CH_3CO), 1.43-1.31 (m, 4H, $OCH_2CH_2CH_2CH_2CH_2N_3$), 1.12-1.01 (m, 2H, $O(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$) δ** 170.1, 169.6, 169.2, 138.9, 138.8, 138.6, 138.3, 137.8, 135.5, 133.7, 133.4, 133.1, 128.6, 128.5, 128.4, 128.3, 128.3, 128.1, 128.0, 127.9, 127.8, 127.7, 127.6, 127.6, 127.1, 127.0, 126.7, 126.6, 126.4, 126.2, 125.9, 123.3, 101.2, 98.9, 98.2, 97.2, 81.1, 80.3, 76.0, 75.7, 75.2, 74.6, 74.6, 74.4, 74.0, 73.8, 73.4, 72.8, 71.7, 70.7, 69.0, 68.8, 68.4, 68.3, 68.1, 62.1, 59.9, 56.7, 55.8, 51.2, 29.8, 28.8, 28.4, 23.1, 21.1, 21.1, 20.9; **HRMS (ESI) *m/z*** [M+ Na]$^+$ calculated for $C_{96}H_{99}N_5O_{25}Na$ 1744.6521, found 1744.6563.

5-azidopentyl 2-*O*-acetyl-3,4,6-tri-*O*-benzyl-α-D-mannopyranosyl-(1→6)-[2,3,4-tri-*O*-acetyl-β-D-xylopyranosyl-(1→2)]-4-*O*-benzyl-3-*O*-(2-naphthylmethyl)- β-D-mannopyranosyl)-(1→4)-(3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(1→4)-3,6-di-*O*-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (28): A mixture of **26** (24 mg, 13.93 µmol) and **5** (13.4 mg, 20.19 µmol) was dried under high vacuum overnight. The mixture was dissolved in anhydrous CH_2Cl_2 (0.4 mL) and freshly activated, ground molecular sieves (3Å) were added. After 30 minutes of stirring at room temperature, the mixture was cooled to -20^{o}C and TMSOTf (3.6 µL, 0.15 eq, 2.08 µmol) was slowly added. After 1h stirring at -20 ºC, the reaction was quenched with Et_3N, filtered through a pad of Celite® and concentrated. The crude product was purified by preparative thin layer chromatography (TLC) on Silica gel 60 F_{254}, Merck KGaA, Germany (hexane: EtOAc 3:1 to 3:2) to afford product in 52% yield (16 mg); **$[\alpha]^{20}_{589}$**= +2.1 (c, 0.45, $CHCl_3$); **^{1}H NMR (500 MHz, $CDCl_3$) δ** 7.83-7.76 (m, 4H, Ar), 7.71-7.62 (m, 7H, Ar), 7.48– 7.44 (m, 3H, Ar), 7.33-7.29 (m, 7H, Ar), 7.24-7.15 (m, 20H, Ar), 7.17-7.11 (m, 2H, Ar), 7.07-7.04 (m, 2H, Ar), 7.04-6.98 (m, 2H, Ar), 6.93-6.92 (m, 2H, Ar), 6.77-6.68 (m, 6H, Ar), 5.38-5.37 (dd, $J_{2\text{-}3}$= 3.3 Hz, $J_{2\text{-}1}$= 1.8 Hz, 1H, H-2M), 5.26-5.22 (m, 2H, H-3X, H-1Aβ), 5.10-5.06 (m, 2H, H-1X, H-2X), 4.98-4.89 (m, 4H, H-4X, H-1Bβ, OC*H*$_2$Ph), 4.82-4.75 (m, 4H, H1M-α, OC*H*$_2$Ph), 4.62-4.26 (m, 16H, H-1Cβ, OC*H*$_2$Ph, H-5Xa, H-3A), 4.21-4.15 (m, 3H, H-2A, H-2C, H-4B), 4.09-4.06 (m, 2H, H-2B, H-3B), 4.01 (t, $J_{4\text{-}3/5}$=9.1 Hz, 1H, H-4A), 3.91 (dd, $J_{3\text{-}4}$= 9.6 Hz, $J_{3\text{-}2}$= 3.3 Hz, 1H, H-3M), 3.77 (t, $J_{4\text{-}3/5}$= 9.6 Hz, 1H, H-4M), 3.73-3.63 (m, 5H, H-4C, H-6Ca, H-6Aa, H-6Ba, O*CH*H(CH2)4N3), 3.60-3.54 (m, 2H, H-5M, H-6Cb), 3.53-3.49 (m, 2H, H-6Ab, H-6Bb), 3.45-3.37 (m, 4H, H-5Xb, H-3C, H-6Mab), 3.33-3-21 (m, 4H, H-5A, H-5B, H-5C, O*CH*H(CH2)4N3), 2.91-2.81 (m, 2H, -C*H*$_2$N$_3$), 2.10, 1.98, 1.97, 1.97, (4 s, 4x3H, C*H*$_3$CO), 1.38-1.29 (m, 4H, $OCH_2CH_2CH_2CH_2CH_2N_3$), 1.09–1.03 (m, 2H, $O(CH_2)_2CH_2(CH_2)_2N_3$); **^{13}C NMR (126 MHz, $CDCl_3$)** δ 170.4-169.1, 138.7-133.1, 128.4-123.2, 102.7, 99.8, 98.0, 97.4, 96.9, 81.3, 78.1, 76.5, 76.0, 75.0-74.1, 73.3, 72.9, 72.7, 71.6, 71.3, 70.4, 69.0- 68.3, 66.7, 56.7, 55.8, 51.2, 28.8, 28.4, 23.1, 21.0, 20.9; **HRMS (ESI) *m/z*** $[M+H+Na]^{2+}$ calculated for $C_{125}H_{129}N_5O_{31}$ 1120.9227, found 1120.9224.

General procedure A: Deprotection of compounds G2 and G4

A mixture of the glycans (1 eq.) and ethylenediamine:*n*BuOH (1:4; 42 mL/mmol) was heated at 120 ºC under microwave irradiation (3 cycles x 30 minutes). The reaction mixture was concentrated and coevaporated with toluene (3x) and EtOH (3x). The crude product was dissolved in anhydrous pyridine (60 µL/µmol), cooled to 0 ºC and Ac_2O (28 µL/µmol) was added dropwise. The reaction was warmed to room temperature and stirred overnight. The mixture was concentrated and coevaporated with toluene (3 times). The crude product was purified by flash column chromatography (SiO_2, EtOAc:MeOH, 99:1). Sodium (25 mg) was added to liquid ammonia (15 mL) at -78 ºC and the resulting blue colored solution was stirred for 5 minutes. A solution of the protected glycan in anhydrous THF (0.11 mL/µmol) was added dropwise and the resulting mixture was stirred at -78 ºC until the blue color disappeared (1 h). The reaction was quenched by adding solid NH_4Cl until neutral pH and evaporated to dryness. The crude was purified by graphitized carbon cartridge and the deprotected N-glycans were eluted using H_2O:MeOH (1:1).

General procedure B: Deprotection of compounds G1, G3, G5 and G6.

A mixture of the protected glycan (1 eq.) and ethylenediamine:*n*BuOH (1:4; 42 mL/mmol) was heated at 120 ºC under microwave irradiation (3 cycles x 30 minutes). The reaction mixture was concentrated and coevaporated with toluene (3 times) and EtOH (3 times). The crude product was dissolved in anhydrous pyridine (60 µL/µmol), cooled to 0 ºC and Ac_2O (28 µL/µmol) was added dropwise. The reaction was warmed to rt and stirred overnight. The mixture was concentrated and co-evaporated with toluene (3 times) to remove pyridine. The crude product was purified by size exclusion chromatography (CH_2Cl_2:MeOH, 1:2). The compound was dissolved in MeOH and 0.5 M NaOMe in MeOH was added (0.2 eq. per each acetate group). After 2 h of stirring at room temperature, the reaction mixture was quenched with Amberlite® IR120 (H), filtered and concentrated. The crude product was

dissolved in MeOH containing 0.1% of trifluoroacetic acid (TFA) and hydrogenated with H-Cube® reactor on 10% Pd/C cartridge (full H_2 mode; 0.5 mL/min; 50 ºC). The solution was quenched with $NaHCO_3$, the crude was purified by graphitized carbon cartridge and the deprotected N-glycans were eluted using H_2O:MeOH (1:1).

5-Aminopentyl [α-D-mannopyranosyl]-(1→6),[β-D-xylopyranosyl]-(1→2)-β-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (G1): Following the general procedure B, compound **28** (24.5 mg, 11.17 µmol) was deprotected to give **G1** as a white solid (5.3 mg, 49 % over 4 steps); **^{1}H NMR (500 MHz, D_2O) δ** 4.93 (s, 1H), 4.87 (s, 1H), 4.62 (d, *J*= 7.5 Hz, 1H), 4.50 (d, *J*= 7.5 Hz, 1H), 4.46 (d, *J*= 7.5 Hz, 1H), 4.28 (m, 1H), 3.99– 3.59 (m, 33H), 3.54-3.37 (m, 4H), 3.27 (t, *J*= 11.1Hz, 1H), 2.94 (t, *J*= 7.5 Hz, 2H, C*H*$_2$NH$_2$), 2.09, 2.04 (2 s, 2x3H, -NH*Ac*), 1.66-1.58 (m, 4H, CH_2 linker), 1.43-1.37 (m, 2H, CH_2 linker); **^{13}C NMR (126 MHz, D_2O) δ** 174.7, 174.4, 104.2, 101.3, 101.0, 100.7, 99.7, 79.7, 79.3, 78.0, 75.3, 74.5, 74.4, 73.2, 72.7, 72.4, 72.1, 70.4, 70.1, 69.8, 69.2, 67.0, 66.7, 65.8, 65.0, 60.9, 60.1, 55.0, 39.5, 28.1, 27.0, 22.2, 22.1; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{38}H_{67}N_3O_{25}H$ 966.4142, found 966.4161.

5-Aminopentyl [α-D-mannopyranosyl]-(1→3),[β-D-xylopyranosyl]-(1→2)-β-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (G2): Following the general procedure B, compound **13** (31 mg, 15.08 µmol) was deprotected to give **G2** as a white solid (7.43 mg, 50 % over 3 steps); **^{1}H NMR (500 MHz, D_2O) δ** 5.14 (s, 1H), 4.86 (s, 1H), 4.60 (d, *J*= 7.6 Hz, 1H), 4.50 (d, *J*= 7.8 Hz, 1H), 4.47(d, *J*= 7.6 Hz, 1H), 4.26 (d, *J*= 3.0 Hz, 1H), 4.06– 4.04 (m, 1H), 4.03-3.99 (m, 2H), 3.93-3.57 (m, 27H), 3.52-3.43 (m, 3H), 3.38 (t, *J*= 8.6 Hz, 1H), 3.25 (t, *J*= 11.3 Hz, 1H), 2.93 (t, *J*= 7.2 Hz, 2H), 2.08, 2.04 (2 s, 2x3H, -NH*Ac*), 1.67-1.56 (m, 4H, CH_2 linker), 1.42-1.36 (m, 2H CH_2 linker); **^{13}C NMR (126 MHz, D_2O, peaks assigned from HSQC) δ** 104.8, 102.2, 101.3, 101.0, 100.1, 79.3, 79.0, 78.7, 77.3, 76.2, 74.5, 74.5, 73.5, 73.2, 72.4, 72.0, 70.3, 70.3, 70.1,

70.0, 69.2, 66.8, 66.3, 64.9, 64.9, 60.3, 60.1, 60.1, 55.1, 55.0, 39.5, 28.2, 27.2, 22.3, 22.2, 22.2; **HRMS (ESI) *m/z*** $[M+Na]^+$ calculated for $C_{38}H_{67}N_3O_{25}Na$ 988.3956, found 988.3972.

5-Aminopentyl [2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl]-(1→3),[β-D-xylopyranosyl]-(1→2)- β**-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-**β**-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-**β**-D-glucopyranoside (G3):** Following the general procedure B, compound **12** (36 mg, 15.7 µmol) was deprotected to give **G3** as a white solid (12.5 mg, 65% over 4 steps); **^{1}H NMR (500 MHz, D_2O) δ** 5.15 (s, 1H, H1M), 4.86 (s, 1H, H1M'), 4.60 (d, *J*= 7.7 Hz, 1H), 4.53-4.49 (m, 2H), 4.45 (d, *J*= 7.6 Hz, 1H, H1X), 4.25 (d, *J*= 2.8 Hz, 1H, H2M), 4.16 (m, 1H, H2M'), 4.04-4.00 (m, 1H), 3.93-3.85 (m, 10H), 3.80– 3.42 (m, 28H), 3.42-3.39 (m, 1H), 3.27-3.23 (m, 1H), 2.92 (t, *J*= 7.6 Hz, 2H,-C*H*$_2$$NH_2$), 2.07, 2.06, 2.04 (3 s, 3x3H, -NH*Ac*), 1.66-1.57 (m, 4H, CH_2 linker), 1.42-1.37 (m, 2H, CH_2 linker); **^{13}C NMR (126 MHz, D_2O) δ** 174.7, 174.6, 174.4, 104.9, 101.3, 101.0, 100.1, 99.7, 99.2, 79.3, 79.1, 78.7, 77.1, 76.6, 76.3, 75.8, 75.3, 74.5, 73.6, 73.3, 73.2, 72.4, 71.9, 70.1, 69.9, 69.4, 69.3, 67.2, 66.5, 64.9, 61.7, 60.6, 60.2, 60.1, 55.3, 55.1, 55.0, 39.5, 28.1, 27.3, 22.3, 22.1, 22.1, 22.1; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{46}H_{80}N_4O_{30}H$ 1169.493, found 1169.4907.

5-Aminopentyl [α-D-mannopyranosyl]-(1→3), [2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl]- (1→6), [β-D-xylopyranosyl]-(1→2)-β**-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-**β**-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-**β**-D-glucopyranoside (G4):** Following the general procedure A compound **17** (40 mg, 15.0 µmol) was deprotected to give **G4** as a white solid (11.9 mg, 56 %); **^{1}H NMR (500 MHz, D_2O) δ** 5.13 (s, 1H), 4.91 (s, 1H), 4.88 (s, 1H), 4.60 (d, *J*= 7.8 Hz, 1H), 4.56 (d, *J*= 8.4 Hz, 1H), 4.50 (m, *J*= 7.8 Hz, 1H), 4.45 (d, *J*= 7.5 Hz, 1H), 4.27 (d, *J*= 2.8 Hz, 1H), 4.12-4.11 (m, 1H), 4.06-4.04 (m, 1H), 4.02-3.98 (m, 2H), 3.94-3.36 (m, 43H), 3.25 (t, *J*= 11.2 Hz, 1H), 2.92 (t, *J*= 7.5 Hz, 2H, C*H*$_2$$NH_2$), 2.09, 2.06, 2.04 (3 s, 3x3H, -NH*Ac*), 1.65-1.58 (m, 4H, CH_2 linker), 1.40-1.37 (m, 2H, CH_2 linker); **^{13}C NMR (126 MHz, D_2O) δ** 174.8, 174.6, 174.4, 105.1, 102.2, 101.3, 101.0, 100.6, 99.6, 96.9, 79.5, 79.3, 77.3, 76.2, 75.8, 75.3, 74.5, 74.4, 74.3, 73.5, 73.3, 72.8, 72.4, 72.0, 70.3, 70.1, 69.9, 69.9, 69.7, 69.5, 69.3, 67.3, 66.7, 66.4, 65.4, 64.9, 61.6, 61.2, 60.6,

60.1, 60.0, 55.3, 55.0, 39.5, 28.2, 28.1, 27.4, 27.2, 22.3, 22.2, 22.1; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{52}H_{90}N_4O_{35}H$ 1331.5458, found 1331.5494.

5-Aminopentyl [2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl]-(1→3),[α-D-mannopyranosyl]-(1→6),[β-D-xylopyranosyl]-(1→2)-β-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (G5): Following the general procedure B, compound **20** (24.8 mg, 8.9 µmol) was deprotected to give **G5** as a white solid (58% over 4 steps); **^{1}H NMR (500 MHz, D_2O) δ** 5.15 (s, 1H), 4.93 (d, 1H, *J*= 1.7 Hz), 4.88 (s, 1H), 4.61 (d, *J*= 8.0 Hz, 1H), 4.53 (d, 1H, *J*= 8.3 Hz), 4.50 (d, *J*= 8.2Hz, 1H), 4.45 (d, *J*= 7.6 Hz, 1H), 4.27 (d, *J*= 3.2 Hz, 1H), 4.20-4.15 (dd, *J*= 3.2, 1.4 Hz, 1H), 4.05-3.37 (m, 45H), 3.26 (t, *J*= 11.1 Hz, 1H), 2.99 (t, *J*= 7.8 Hz, 1H, linker), 2.09, 2.06, 2.04 (3 s, 3x3H, -NH*Ac*), 1.71-1.65 (m, 2H, linker), 1.64-1.58 (m, 2H, linker), 1.43-1.39 (m, 2H, linker); **^{13}C NMR (126 MHz, D_2O) δ** 105.1, 101.3, 101.0, 100.4, 99.7, 99.5, 99.1, 79.5, 79.3, 79.2, 77.0, 76.5, 75.6, 74.5, 74.2, 73.7, 73.3, 73.2, 72.7, 72.4, 71.8, 70.3, 70.2, 70.1, 69.8, 69.7, 69.7, 69.4, 69.2, 67.2, 66.6, 66.5, 65.2, 65.2, 64.9, 64.8, 61.7, 60.6, 60.6, 60.1, 54.9, 55.1, 39.4, 28.1, 26.5, 22.3, 22.1; **HRMS (ESI)** *m/z* $[M+H]^+$ calculated for $C_{52}H_{90}N_4O_{35}H$ 1331.5458, found 1331.5502.

5-Aminopentyl di[2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl]-(1→3),(1→6)-[β-D-xylopyranosyl]-(1→2)-β-D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (G6): Following the general procedure B, compound **16** (95 mg, 54.2 µmol) was deprotected to give **G6** as a white solid (50.2 mg, 60 % over 4 steps);**^{1}H NMR (500 MHz, D_2O) δ** 5.13 (s, 1H), 4.90 (s, 1H), 4.86 (s, 1H), 4.59 (d, *J*= 7.7 Hz, 1H), 4.57 (d, *J*= 8.4 Hz, 1H), 4.52 (d, *J*= 8.3 Hz, 1H), 4.51 (d, *J*= 7.6 Hz, 1H), 4.45(d, *J*= 7.7 Hz, 1H), 4.26 (d, *J*= 2.6 Hz 1H), 4.15-4.12 (m, 1H), 4.12-4.07 (m, 1H), 4.05 –3.35 (m, 51H), 3.24 (t, *J*= 11.2 Hz, 1H), 2.91 (t, *J*= 7.8 Hz, 2H), 2.07, 2.05, 2.04, 2.02 (4 s, 4x3H, -NH*Ac*), 1.65-1.55 (m, 4H, linker), 1.40-1.34 (m, 2H, linker); **^{13}C NMR (126 MHz, D_2O) δ** 174.8, 174.7, 174.5, 105.1, 101.4, 101.0, 100.5, 99.7, 99.6, 99.2, 96.9, 79.6, 79.4, 76.6, 76.3, 75.8, 75.4, 74.5, 74.4, 73.7, 73.4, 73.3, 72.9, 72.5, 72.1, 70.2,

69.9, 69.6, 69.5, 69.3, 67.3, 67.2, 66.5, 65.0, 61.8, 61.6, 60.7, 60.1, 55.4, 55.1, 39.6, 28.1, 27.3, 22.4, 22.2, 22.2; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{60}H_{103}N_5O_{40}H$ 1534.6252, found 1534.6322.

Screening of CeFUT6 specificity on the glycan array (more detailed description of glycan array preparation will be given in Chapter 2)

50 µM solutions of N-glycans G1-G6 (and corresponding non-xylosylated structures) were prepared from stock solutions (1 mM in water) by dilution with sodium phosphate buffer (300 mM, pH 8.4, 0.005% Tween 20). 40 µl of each glycans solutions were pippeted into a 384 well plate (Scienion, Berlin, Germany) that was used as a source plate for printing glycans onto NHS functionalized glass slides (each spot: 1.25 nL, 5 drops of 250 pL). N-glycans were printed at a spot to spot distance of 350 µm (x and y axies). Ligands were printed in 6 replicates (3 different N-glycans/row). After printing, the slides were placed in a 75 % humidity chamber (saturated NaCl solution) at 25 ºC for 18 hours. The unreacted NHS groups were quenched by placing the slides in a 50 mM solution of ethanolamine in sodium borate buffer 50 mM, pH 9.0, for 1h. The standard washing of the slides was performed with PBST (PBS solution containing 0.5% Tween 20), PBS and water. The slides were dried in a slide spinner.

The array was used to predict possible enzymatic elongations of xylosylated compounds (G1-G6) and to compare enzyme specificities with corresponding, non-xylosylated structures. A solution (100 µL) containing 10 µL of CeFUT6 fucosyltransferase, GDP-Fuc (2 mM), $MnCl_2$ (20 mM) in MES buffer (80 mM, pH 6.5) was added to each well and the slide was incubated at room temperature for 24 hours. The slide was then washed with PBS and water. To test efficient fucosylation one well was incubated with *Aleuria aurantia* lectin (AAL-555, 20 µg/mL) in PBS, 5mM $CaCl_2$, 5 mM $MgCl_2$ containing 0.1% Tween-20, for 1h at

room temperature. The slide was washed with PBST, PBS and water, dried under a stream of Ar and analyzed.

Enzymatic synthesis of N-glycans

Expression of bovine milk β-1,4-galactosyltransferase in *E. coli*.

A single colony was inoculated in 10 mL of LB medium (1 % NaCl, 1 % peptone and 0.5 % yeast extract) containing kanamycin (50 μg/mL) and incubated at 37 ºC/250 rpm overnight. This culture was used to inoculate 400 mL of LB medium containing kanamycin (50 μg/mL). Cells were grown at 37ºC/250rpm until they reached OD_{600nm} value of 0.6-0.7. Protein expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) during 3 hours at 37ºC/250rpm. The culture was harvested by centrifugation (30000 rpm, 4ºC, 10 minutes) and washed with Tris 50 mM, 150 mM NaCl, pH=8.5 twice. The pellet was resuspended in 10 mL of lysis buffer (Tris 50 mM, 150 mM NaCl, 1% Triton X-100, 10 mM EDTA, 1 mM phenylmethylsulfonyl fluoride PMSF pH=8.5) and lysis was performed by sonication followed by centrifugation. The treatment with lysis buffer was performed twice and the inclusion bodies were washed once with water. The inclusion bodies were solubilized in 5M guanidine·HCl until OD_{280nm} value of 1.9 to 2.0 (1 mg/mL). The protein solution was filtered through 0.22 μm filter and diluted ten times (100 μg/mL) in folding buffer containing Tris·HCl 50 mM pH=8.5, 10.56 mM NaCl, 0.44 mM KCl, 2.2 mM $MgCl_2$, 2.2 mM $CaCl_2$, 0.5 M guanidine·HCl, 8 mM cysteamine, 4 mM cystamine, 0.055 % PEG-4000, 0.55 M L-arginine. The protein was allowed to renature for 48 h at 4 ºC and then dialyzed against Tris 25 mM, 150 mM NaCl, pH 8.5 (24 h, 4 changes of buffer). The precipitated protein was removed by centrifugation at 5000 rpm, 10 mM imidazole was added to the supernatant and purified by HisTrap column (5 mL) controlled by an ÄKTA™ protein purifier. The column was washed with 5 CV of binding buffer and the protein is eluted with Tris 25 mM, 300 mM NaCl, 500 mM imidazole, pH =8.5. Elution fractions were analyzed by SDS-

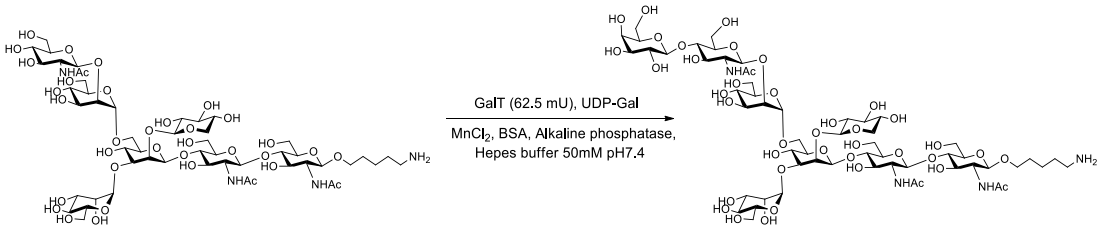
PAGE**.** SDS-PAGE was performed in 12 % acrylamide gels that were stained with Coomassie Blue G-250. Protein containing fractions were pooled and dialyzed against Tris 25 mM, 150 mM NaCl, pH 7.5**.**

Expression of glycosyltransferases in *P. pastoris*:

Recombinant strains of *Pichia pastoris* were inoculated in 10 mL of MGYC medium (1% (w/v) yeast extract, 2% (w/v) peptone, 1% (w/v) casamino acid, 1.34% (w/v) yeast nitrogen base, 1% (v/v) glycerol, 0.00004% (w/v) biotin). The cultures were grown at 30 ºC with shaking at 250 rpm in baffled flasks for 24 h. The culture was centrifuged at 500g for 10 minutes at 4 ºC and re-suspended in 100 mL of MMYC (composition as for MGYC except that MeOH 1% (v/v) substitutes glycerol). The protein expression was carried out at 16 ºC for 72 hours, adding every 24 hours additional MeOH (to 1% (v/v)). After expression phenylmethylsulfonyl fluoride (PMSF, final concentration 0.1 mg/mL) was added to the culture medium and centrifuged at 1500x *g* for 10 minutes. The supernatant was concentrated in Vivaspin® 20 devices. *C. elegans* β-1,4-N-acetylgalactosaminyltransferase was employed in enzymatic reactions without further purification. The fucosyltransferases were further purified by affinity chromatography. The crude extracts were dialyzed against Tris 50 mM, 300 mM NaCl, 7 mM imidazole pH=8.5 and were loaded on nickel Histrap™ HP 1 mL affinity columns controlled by an ÄKTA™ protein purifier. The column was washed with 5 CV of binding buffer and the protein is eluted with a gradient from 0% to 100% of elution buffer (Tris 25 mM, 300 mM NaCl, 500 mM imidazole, pH =8.5). Elution fractions were analyzed by SDS-PAGE/Western Blot**.** SDS - PAGE was performed in 12 % acrylamide gels that were stained with Coomassie Blue G - 250. For Western blotting, the polyacrylamide gel was rinsed with blotting solution (50 mM Tris, 40 mM glycine, 0.036% (w/v) SDS and 20 % (v/v) MeOH) for 5 minutes and placed into the transfer sandwich soaked with blotting solution and proteins were transferred by semidry blotting to a nitrocellulose membrane. After the protein transfer, the membrane was rinsed with TBS buffer (10 mM Tris, 150 mM

NaCl, pH 7.4) for 5 minutes. The membrane was then blocked with TBST (10 mM Tris, 150 mM NaCl, 0.05% (v/v) Tween 20, pH 7.4) containing 0.5% (w/v) of bovine serum albumin (BSA) during 1 hour. The membrane was incubated with anti-HIS monoclonal antibody raised in mouse (General Electrics) 1:3000 in TBST containing 0.5% (w/v) of BSA. After one hour of incubation, the membrane was washed three times with TBST and incubated with the secondary antibody anti - Mouse IgG alkaline phosphatase conjugate (Sigma Aldrich), 1:10000 in TBST containing 0.5% (w/v) BSA for one hour. After washing as above mentioned the membrane was immersed in staining solution FAST™BCIP/NBT (Sigma Aldrich). The reaction was stopped with 5% acetic acid. Protein fractions were pooled, concentrated with Vivaspin® 20 devices and dialyzed against Tris·HCl 25 mM, 150 mM NaCl, pH 7.5.

General procedure C: Reaction with bovine milk β-1,4-galactosyltransferase

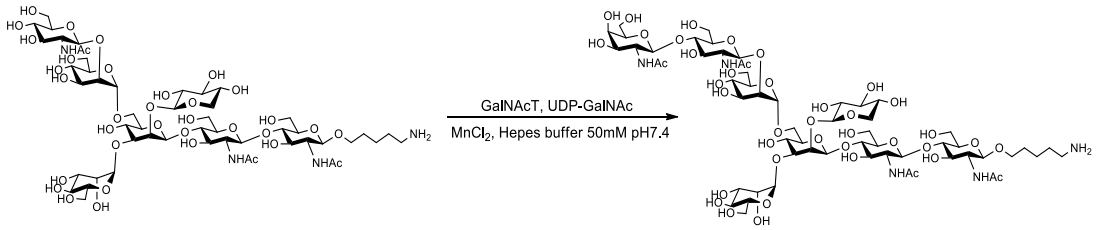


A solution (500 µL/mg) of glycan (1eq.), UDP-Gal (Uridine 5'-diphosphogalactose disodium salt) (1.2 eq.), bovine serum albumin BSA (0.5 mg), bovine milk β-1,4-galactosyltransferase, alkaline phosphatase and 10 mM $MnCl_2$ in HEPES buffer (50 mM, pH=7.4) was incubated at 37 ºC overnight. The resulting mixture was heated at 95ºC for 5 minutes to precipitate the enzyme. After centrifugation, the supernatant was purified on a C18 cartridge and the product was eluted using a mixture of ACN: aqueous 0.1% TFA (1:1). Fractions were analyzed by MALDI-TOF and the compound containing fractions were pooled and freeze dried to obtain the title compound.

General procedure D: Reaction with fucosyltransferases.

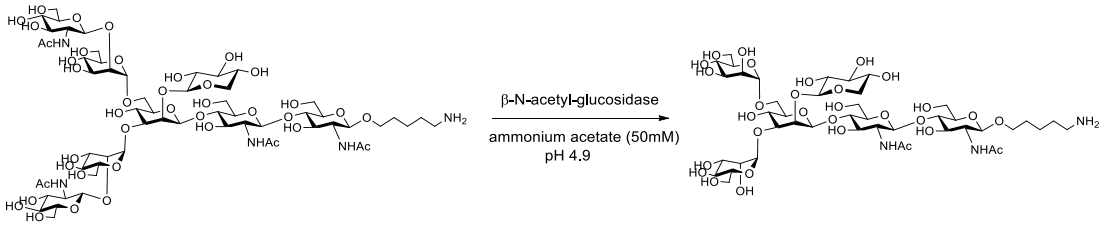
A solution (500 µL/mg) of glycan (1 eq.), GDP-Fuc (guanosine 5′-diphospho-β-L-fucose sodium salt) (1.2eq.), CeFUT6, AtFucTA, CeFUT1 or CeFUT8 and 20 mM $MnCl_2$ in MES buffer (40 mM, pH=6.5) was incubated at room temperature overnight. The resulting mixture was heated at 95ºC for 5 minutes to precipitate the enzyme. After centrifugation the supernatant was purified on a C18 cartridge and the product was eluted using a mixture of ACN: aqueous 0.1% TFA(1:1). Fractions were analyzed by MALDI-TOF and the compound containing fractions were pooled and freeze dried.

General procedure E: Reaction with β-1,4-N-acetylgalactosaminyltransferase



A solution (500 µL/mg) of glycan (1 eq.), UDP-GalNAc (Uridine 5′-diphospho-N-acetylgalactosamine disodium salt) (1.2 eq.), concentrated *P. pastoris* extract containing β-1,4-GalNAcT and 10 mM $MnCl_2$ in HEPES buffer (40 mM, pH=7.4) was incubated at room temperature for 72h. The resulting mixture was heated at 95ºC for 5 minutes to precipitate the enzyme**.** After centrifugation the supernatant was purified on a C18 cartridge and the product was eluted using a mixture of ACN: aqueous 0.1% TFA(1:1). Fractions were analyzed by MALDI-TOF and the compound containing fractions were pooled and freeze dried.

General procedure F: Reaction with Jack bean β-N-acetylglucosaminidase.



A solution (500 µL/mg) of compound glycan (1 eq.), Jack bean β-N-acetylglucosaminidase in ammonium acetate buffer (50 mM, pH 4.9) was incubated at room temperature for 96 h. The resulting mixture was heated at 95ºC for 5 minutes to precipitate the enzyme. After centrifugation the supernatant was purified on a C18 cartridge and the product was eluted using a mixture of ACN: aqueous 0.1% TFA (1:1).. Fractions were analyzed by MALDI-TOF and the compound containing fractions were pooled and freeze dried.

Heptasaccharide G7: Following the general procedure D, compound **G7** (1.35 mg, 98.5%) was prepared from **G3** (1.22 mg, 1.04 µmol) using CeFUT8 (30 µL, 1.1 mg/mL) and GDP-Fuc (792 µg, 1.25 µmol). **^{1}H NMR (500 MHz, D_2O) δ** 5.15 (s, 1H, H-$1_{\alpha Man}$), 4.91 (d, *J*= 4.0 Hz, 1H, H-1_{Fuc}), 4.86 (s, 1H, H-$1_{\beta Man}$), 4.68 (d, *J*= 7.6 Hz, 1H), 4.51 (t, *J*= 8.4 Hz, 2H), 4.46 (d, *J*= 7.6 Hz, 1H), 4.26-4.25 (m, 1H, H-2_{Man}), 4.17-4.16 (m, 1H, H-2_{Man}), 4.16-4.12 (m, 1H, H-5_{Fuc}), 4.02 (dd, *J*= 11.5, 5.4 Hz, 1H), 3.96-3.44 (m, 38H), 3.40-3.35 (m, 1H), 3.26 (t, *J*= 11.1 Hz, 1H), 2.99 (t, *J*= 7.6 Hz, 2H, linker), 2.08, 2.06, 2.04 (3 s, 3x3H, NH*Ac*), 1.69-1.65 (m, 2H, linker), 1.63-1.58 (m, 2H, linker), 1.44-1.38 (m, 2H, linker), 1.24 (d, *J*= 6.7 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (126 MHz, D_2O) δ**,105.0, 101.2, 101.0, 100.1, 99.7, 99.4, 99.2, 79.2, 78.7, 78.6, 77.17, 76.6, 76.3, 75.7, 73.6, 73.4, 73.3, 73.1, 72.3, 72.0, 71.8, 70.0, 69.8, 69.6, 69.5, 69.5, 69.4, 69.3, 68.1, 68.0, 67.2, 66.8, 66.5, 66.5, 64.9, 64.9, 61.7, 60.3, 60.3, 55.2, 55.1, 39.3, 28.1, 26.4, 22.3(x3), 22.1, 15.4; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{52}H_{90}N_4O_{34}H$ 1315.5509, found 1315.5499.

Octasaccharide G8: Following the general procedure D, compound **G8** (1.24 mg, 99%) was prepared from **G5** (1.03 mg, 0.76 µmol) using core α-1,6-fucosyltransferase CeFUT8 (30 µL,

1.1 mg/ml), and GDP-Fuc (589 µg, 0.93 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.15 (s, 1H, H-$1_{\alpha Man}$), 4.93 (s, 1H, H-$1_{\alpha Man}$), 4.92 (d, *J*= 4.0 Hz, 1H, H-1_{Fuc}), 4.88 (s, 1H), 4.68 (d, *J*= 8.1 Hz, 1H), 4.53 (d, *J*= 8.1 Hz, 1H), 4.51 (d, *J*= 8.1 Hz, 1H), 4.45 (d, *J*= 7.6 Hz, 1H), 4.27-4.26 (m, 1H, H-2_{Man}), 4.18-4.16 (m, 1H, H-2_{Man}), 4.15-4.12 (m, 1H, H-5_{Fuc}), 4.04-3.51 (m, 45H), 3.40-3.36 (m, 1H), 3.26 (t, *J*= 11.3 Hz, 1H), 2.99 (t, *J*= 7.5 Hz, 2H, linker), 2.10, 2.06, 2.04 (3 s, 3x3H, NH*Ac*), 1.71-1.65 (m, 2H, linker), 1.63-1.58 (m, 2H, linker), 1.43-1.38 (m, 2H, linker), 1.24 (d, *J*= 6.7 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (126 MHz, D_2O) δ** 105.2, 101.1, 101.0, 100.6, 99.7, 99.5, 99.1, 79.6, 79.6, 79.4, 78.6, 76.9, 76.5, 75.6, 74.4, 74.3, 73.6, 73.3, 73.3, 72.8, 72.6, 72.1, 71.8, 70.2, 70.0, 69.9, 69.8, 69.6, 69.5, 69.3, 68.1, 68.0, 67.2, 66.8, 66.7, 66.6, 66.5, 65.4, 65.3, 64.9, 64.9, 61.8, 60.7, 60.6, 55.20, 54.3, 54.8, 39.3, 28.1, 26.4, 22.3, 22.1, 22.1 (x2), 15.3; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{58}H_{100}N_4O_{39}H$ 1477.6037, found 1477.6073.

Nonasaccharide G9: Following the general procedure D, compound **G9** (0.82 mg, 75%) was prepared from **G6** (1 mg, 0.651 µmol) using core α-1,6-fucosyltransferase CeFUT8 (30 µL, 1.1 mg/mL) and GDP-Fuc (589 µg, 0.93 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s, 1H, H-$1_{\alpha Man}$), 4.92 (s, 2H, H-$1_{\alpha Man,}$ H-$1_{\alpha Fuc}$), 4.89 (s, 1H, H-$1_{\beta Man}$), 4.68 (d, *J*= 7.9 Hz, 1H), 4.57 (d, *J*= 8.3 Hz, 1H), 4.53 (d, *J*= 8.3 Hz, 1H), 4.51 (d, *J*= 8.0 Hz, 1H), 4.45 (d, *J*= 7.6 Hz, 1H), 4.27-4.25 (m, 1H, H-2_{Man}), 4.18-4.16 (m, 1H, H-2_{Man}), 4.14-4.12 (m, 2H), 4.04-4.00 (m, 2H), 3.94-3.37 (m, 49H), 3.26 (t, *J*= 11.1 Hz, 1H), 3.00 (t, *J*= 7.7 Hz, 2H, linker), 2.10, 2.07, 2.06, 2.04 (4 s, 4x3H, NH*Ac*), 1.69-1.66 (m, 2H, linker), 1.62-1.59 (m, 2H, linker), 1.42-1.38 (m, 2H, linker), 1.24 (d, *J*= 6.6 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (126 MHz, D_2O) δ** 105.2, 101.2, 101.1, 100.6, 99.7, 99.3, 99.2, 99.2, 99.1, 81.8, 79.6, 79.6, 79.4, 79.2, 77.1, 76.6, 76.3, 76.2, 75.7, 74.3, 73.9, 73.3, 73.3, 73.1, 73.0, 72.2, 71.8, 71.70, 71.1, 70.2, 69.8,69.5, 69.4, 68.2, 68.0, 67.1, 66.8, 66.5, 66.3, 66.3, 65.5, 65.1, 64.9, 64.2, 61.7, 61.7, 60.1, 60.6, 60.3, 55.2, 55.1, 54.9, 53.5, 39.3, 28.1, 26.5, 22.2, 22.2, 22.1, 15.4; **HRMS (ESI) *m/z*** $[M+Na]^+$ calculated for $C_{66}H_{113}N_5O_{44}Na$ 1702.6651, found 1702.6577.

Hexasaccharide G10: Following the general procedure D, compound **G10** (720 µg, 99%) was prepared from **G2** (548 µg, 0.57 µmol) using core α-1,3-fucosyltransferase CeFUT1 (26 µL,

0.9 mg/mL), and GDP-Fuc (430 µg, 0.68 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.15 (s, 1H, H-1_{Fuc}), 5.14 (s, 1H, H-$1_{\alpha Man}$), 4.84 (s, 1H, H-$1_{\beta Man}$), 4.74-4.70 (m, 1H, H-5_{Fuc}), 4.55 (d, *J*= 7.9 Hz, 1H), 4.49 (t, *J*= 7.6 Hz, 2H), 4.26 (d, *J*= 3.2 Hz, 1H, H-2_{Man}), 4.06-3.39 (m, 29H), 3.26 (t, *J*= 11.2 Hz, 1H), 2.95 (t, *J*= 7.7 Hz, 2H, linker), 2.06, 2.03 (2 s, 2x3H, NH*Ac*), 1.68-1.57 (m, 4H, linker), 1.41-1.37 (m, 2H, linker), 1.28 (d, *J*= 6.7 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (126 MHz, D_2O)** δ 104.9, 102.2, 100.9, 100.4, 100.1, 98.4, 79.9, 78.9, 77.4, 76.2, 75.3, 75.3, 74.7, 73.9, 73.5, 73.3, 72.2, 72.0, 70.3, 70.3, 70.2, 69.9, 69.4, 69.1, 67.7, 66.8, 66.7, 66.3, 64.9, 64.7, 61.2, 61.0, 60.1, 55.7, 55.2, 39.4, 28.1, 26.9(2x), 22.3, 22.2, 15.5; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{44}H_{77}N_3O_{29}$ H 1112.4716 found 1112.4758.

Heptasaccharide G11: Following the general procedure D, compound **G11** (1.2 mg, 85 %) was prepared from **G3** (1.27 mg, 1.08 µmol) using AtFucTA (50 µL, 0.55 mg/mL) and GDP-Fuc (0.8 mg, 1.27 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s, 1H, H-1_{Fuc}), 5.14 (d, *J* = 4.2 Hz, 1H, H-$1_{\alpha Man}$), 4.72 (m, 1H, H-5_{Fuc}), 4.85 (s, 1H, H-$1_{\beta Man}$), 4.55 (d, *J*= 7.91 Hz, 1H), 4.53 (d, *J*= 8.4 Hz, 1H), 4.50 (d, *J*= 7.91 Hz, 1H), 4.48 (d, *J*= 7.7 Hz, 1H), 4.25 (d, *J*= 2.27 Hz, 1H, H-2_{Man}), 4.19-4.17 (m, 1H, H-2_{Man}), 4.03 (dd, *J*= 11.7, 5.5 Hz, 1H), 3.99 (dd, *J*= 10.5, 3.26 Hz, 1H), 3.94-3.43 (m, 34H), 3.38 (dd, *J*= 9.8, 7.9 Hz, 1H), 3.27 (t, *J*= 11.1 Hz, 1H), 2.91 (t, *J*= 7.6 Hz, 2H, linker), 2.06, 2.05, 2.03 (3 s, 3x3H, NH*Ac)*, 1.66-1.57 (m, 4H, linker), 1.42-1.37 (m, 2H, linker), 1.28 (d, *J*= 6.9 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (126 MHz, D_2O) δ** 104.9, 101.0, 100.4, 100.1, 99.7, 99.1, 98.3, 79.9, 79.7, 79.1, 77.0, 76.6, 76.1, 75.9 75.4, 74.7, 73.3, 73.7, 73.2, 72.2, 71.9, 70.2, 70.2, 69.8, 69.4, 69.3, 69.2, 67.8, 67.2, 66.8, 66.1, 64.9, 64.9, 61.8, 60.9, 60.5, 60.3, 60.2, 55.6, 55.3, 55.3, 39.7, 28.1, 27.3, 22.2 (3x), 22.1, 15.5; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{52}H_{90}N_4O_{34}Na$ 1337.5329, found 1337.5347.

Octasaccharide G12: Following the general procedure D, compound **G12** (470 µg, 84%) was prepared from **G5** (560 µg, 0.38 µmol) using core α-1,3-fucosyltransferase AtFucTA (25 µL, 0.25 mg/mL), and GDP-Fuc (291 µg, 0.46 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.15 (s, 1H, H-$1_{\alpha Man}$), 5.13 (d, *J*= 3.9 Hz, 1H, H-1_{Fuc}), 4.91 (d, *J*= 1.3 Hz, 1H, H1- $H_{\alpha Man}$), 4.85 (s, 1H, H1 $_{\beta Man}$), 4.71-4.69 (m, 1H, H-5_{Fuc}), 4.54 (d, *J*= 8.3 Hz, 1H, H1- H_{GlcNAc}), 4.52 (d, *J*= 8.7 Hz, 1H, H1-

H_{GlcNAc}), 4.49 (d, *J*= 8.3 Hz, 1H, H1- H_{GlcNAc}), 4.45 (d, *J*= 7.4 Hz, 1H, H1- H_{xyl}), 4.25 (d, *J*= 3.4 Hz, 1H, H2- H_{Man}), 4.16-4.15 (m, 1H, H2- H_{Man}), 4.02 (dd, *J*= 11.6, 5.3 Hz, 1H), 3.98-3.95 (m, 2H), 3.93-3.83 (m, 12H), 3.81-3.80 (m,1H), 3.80-3.72 (m, 9H), 3.68-3.43 (m, 17H), 3.37 (dd, *J*= 9.4, 7.8 Hz, 1H), 3.26 (t, *J*= 11 Hz, 1H), 2.98 (t, *J*= 7.8 Hz, 2H, linker), 2.05, 2.05, 2.02 (3 s, 3x3H, NHAc), 1.68-1.64 (m, 2H, linker), 1.60-1.57 (m, 2H, linker), 1.41-1.37 (m, 2H, linker), 1.27 (d, *J*= 6.6 Hz, 3H, C*H*$_3$ Fuc); **^{13}C NMR (201 MHz, D2O, peaks from HSQC experiment) δ** 107.8, 103.6, 103.2, 103.0, 102.3, 102.3, 101.7, 100.9, 83.3, 81.8, 79.6, 79.2, 78.4, 78.0, 77.1, 77.1, 76.9, 76.5, 76.4, 75.9, 75.8, 75.4, 74.9, 74.7, 74.4, 73.1, 72.5, 72.4, 72.0, 71.9, 71.8, 69.8, 69.4, 69.3, 69.1, 68.0, 67.8, 67.8, 64.4, 64.4, 63.5, 63.5, 63.0, 62.6, 62.4, 58.2, 58.1, 57.7, 41.9, 30.7, 29.0, 24.8, 24.8, 24.8, 18.1; **HRMS (ESI) *m/z*** $[M+Na+H]^{2+}$ calculated for $C_{58}H_{101}N_4O_{39}Na$ 851.8362, found 851.8347.

Nonasaccharide G13: Following the general procedure D, compound **G13** (620 µg, 99%) was prepared from **G6** (570 µg, 0.37 µmol) using AtFucTA (25 µL, 0.25 mg/mL), and GDP-Fuc (267 µg, 0.42 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s, 1H, H-$1_{\alpha Man}$), 5.15 (d, *J*= 3.7 Hz, 1H, H-1_{Fuc}), 4.91 (s, 1H, H-$1_{\alpha Man}$), 4.87 (s, 1H, H-$1_{\beta Man}$),4.75 (m, 1H, H-5_{Fuc}), 4.55 (m, 3H), 4.50 (d, *J*= 8.0 Hz, 1H), 4.46 (d, *J*= 7.6 Hz, 1H), 4.27-4.26 (m, 1H, H-$2_{Man\beta}$), 4.17-4.16 (m, 1H, H-$2_{Man\alpha}$), 4.12-4.11 (m, 1H, H-$2_{Man\alpha}$), 4.05-3.43 (m, 47H), 3.40 (t, *J*= 8.8 Hz 1H), 3.27 (t, *J*= 11.5 Hz, 1H), 3.00 (t, *J*= 7.8 Hz, 2H, linker), 2.07, 2.06, 2.03 (3 s, 4x3H, NH*Ac*), 1.71- 1.58 (m, 4H, linker), 1.43-1.37 (m, 2H, linker), 1.28 (d, *J*= 6.5 Hz, 3H, CH_{3Fuc}); **^{13}C NMR(201 MHz, D_2O) δ** 104.7, 100.5, 100.1, 99.9, 99.3, 99.1, 98.6, 97.8, 96.5, 80.2, 78.8, 76.6,76.0, 75.7, 75.3, 74.9, 74.8, 74.1, 74.1, 73.8, 73.3, 72.8, 72.8, 72.4, 71.8, 71.5, 71.2, 69.7, 69.7, 69.4, 69.0, 68.9, 68.8, 68.7, 68.0, 67.3, 66.7, 66.2, 65.9, 64.9, 64.5, 64.4, 61.2, 61.2, 61.1, 60.5, 60.5, 60.1, 60.1, 60.1, 59.5, 59.4, 55.1, 54.8, 54.5, 38.9, 27.6, 26.0, 21.8, 21.7, 21.6, 15.0; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{66}H_{113}N_5O_{44}Na$ 1702.6651, found 1702.6577.

Heptasaccharide G14: Following the general procedure C, compound **G14** (630 µg, 99%) was prepared from **G3** (0.5 mg, 0.427 µmol) using 5mU of bovine milk β-1,4-galactosyltransferase and UDP-Gal (390 µg, 0.64 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s,

1H, H-$1_{\alpha Man}$), 4.86 (s, 1H, H-$1_{\beta\text{-}Man}$), 4.61 (d, *J*= 7.7 Hz, 1H), 4.55 (d, *J*= 7.6 Hz, 1H, H-$1_{\beta Gal}$), 4.51 (d, *J*= 8.0 Hz, 1H), 4.48 (d, *J*= 7.8 Hz, 1H), 4.46 (d, *J*= 7.7 Hz, 1H), 4.26 (d, *J*= 2.8 Hz, 1H, H-2_{Man}), 4.18-4.17 (m, 1H, H-2_{Man}), 4.04-3.98 (m, 2H), 3.94-3.43 (m, 37H), 3.40-3.36 (m, 1H) 3.26 (t, *J*= 11.4 Hz, 1H), 2.95 (t, *J*= 7.4 Hz, 2H, linker), 2.08, 2.06, 2.04 (3 s, 3x3H,-NH*Ac*), 1.70-1.56 (m, 4H, linker)1.44-1.37 (m, 2H, linker); **^{13}C NMR (126 MHz, D_2O) δ** 105.0, 102.9, 101.3, 101.0, 100.2, 99.6, 99.2, 79.1, 78.5, 77.2, 76.5, 76.4, 75.3, 75.3, 74.7, 74.6, 74.5, 73.7, 73.3, 72.4, 71.9, 71.0, 70.2, 70.2, 70.1, 69.4, 68.5, 67.2, 66.5, 65.0, 64.9, 61.8, 61.8, 60.9, 60.3, 60.3, 60.2, 59.9, 59.9,58.3, 55.1, 54.9, 39.4, 28.0, 26.8, 22.2, 22.1; **HRMS (ESI) *m/z*** $[M+Na]^+$ calculated for $C_{52}H_{90}N_4O_{35}Na$ 1353.5278, found 1353.5264.

Octasaccharide G15: Following the general procedure C, compound **G15** (640 µg, 95%) was prepared from **G5** (600 µg, 0.45 µmol) applying using 94 mU of bovine milk β-1,4-galactosyltransferase and, UDP-Gal (590 µg, 0.98 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s, 1H, H-$1_{\alpha Man}$), 4.93 (s, 1H, H-$1_{\alpha Man}$), 4.88 (s, 1H, H-$1_{\beta Man}$), 4.61 (d, *J*= 7.7 Hz, 1H), 4.55 (d, *J*= 7.7 Hz, 1H), 4.50 (d, *J*= 8.0 Hz, 1H), 4.48 (d, *J*= 8.0 Hz, 1H), 4.45 (d, *J*= 7.7 Hz, 1H), 4.27 (d, *J*= 2.8 Hz, 1H, H-2_{Man}), 4.18-4.17 (m, 1H, H-2_{Man}), 4.04-3.43 (m, 46H), 3.26 (t, *J*= 11.3 Hz, 1H), 2.96 (t, 2H, *J*= 7.5 Hz, linker), 2.09, 2.06, 2.04 (3 s, 3x3H, -NH*Ac*), 1.67-1.59 (m, 4H, linker), 1.43-1.39 (m, 2H, linker); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 107.8, 105.5, 104.0, 103.7, 103.2, 102.3, 102.2, 101.8, 82.1, 82.0, 81.9, 81.1, 79.7, 79.1, 78.0, 77.9, 77.1, 77.1, 77.0, 76.3, 75.3, 75.0, 74.5, 73.8, 73.1, 72.8, 72.4, 71.9, 71.2, 69.8, 69.4, 69.3, 68.0, 67.6, 67.6, 64.4, 64.4, 63.6, 63.5, 62.6, 62.6, 62.6, 60.8, 57.6, 57.6, 57.5, 42.1, 30.7, 29.5, 24.9, 24.8, 24.7, 24.7; **HRMS (ESI) *m/z*** $[M+Na+H]^{2+}$ calculated for $C_{58}H_{101}N_4O_{40}Na$ 758.2939 found 758.2911.

Octasaccharide G16: Following the general procedure C, compound G16 (0.65 mg, 99 %) was prepared from **G4** (582 µg, 0.437 µmol) using 62.5 mU of bovine milk β-1,4-galactosyltransferase and UDP-Gal (306 µg, 0.502 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.14 (s, 1H, H-$1_{\alpha Man}$), 4.92 (s, 1H, H-$1_{\alpha Man}$), 4.89 (s, 1H, H-$1_{\beta Man}$), 4.61 (d, *J*= 8.4 Hz, 1H), 4.59 (d, *J*= 7.90 Hz, 1H), 4.51 (d *J*= 7.9 Hz, 1H), 4.49 (d, *J*= 7.9 Hz, 1H), 4.46 (d, *J*= 7.6 Hz, 1H), 4.29-4.27

(m, 1H, H-2_{Man}), 4.13-4.11 (m, 1H, H-2_{Man}), 4.06-4.05 (m, 1H, H-2_{Man}), 4.02-3.50 (m, 45H), 3.45 (t, *J*= 9.5 Hz, 1H), 3.41-3.36 (m,1H), 3.26 (t, *J*= 11.1 Hz, 1H), 3.00 (t, *J*= 7.6 Hz, 2H, linker), 2.09, 2.06, 2.04 (3 s, 3x3H, NH*Ac*), 1.71 – 1.65 (m, 2H, linker), 1.63-1.58 (m, 2H, linker), 1.44-1.38 (m, 2H, linker); **^{13}C NMR (126 MHz, D_2O) δ** 105.1, 102.9, 102.2, 101.4, 101.0, 100.6, 99.4, 96.9, 79.7, 79.4, 79.2, 78.5, 78.3, 77.3, 76.2, 75.3, 74.5, 74.4, 73.6, 73.3, 73.1, 72.9, 72.6, 72.3, 72.0, 70.9, 70.2, 70.09, 69.8, 69.9, 69.7, 69.6, 69.5, 69.2, 68.5, 66.6, 66.3, 65.5, 65.1, 65.0, 61.5, 61.4, 61.1, 60.1, 60.0, 60.0, 55.0, 54.9, 39.3, 28.0, 26.3, 22.2, 22.1, 22.1; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{58}H_{101}N_4O_{40}$ 1493.5987, found 1493.6013.

Decasaccharide G17: Following the general procedure C, compound **G17** (700 µg, 99%) was prepared from **G4** (470 µg, 0.306 µmol) using 88 mU of bovine milk β-1,4-galactosyltransferase and, UDP-Gal (550 µg, 0.91 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.14 (s, 1H, H-$1_{\alpha Man}$), 4.91 (s, 1H, H-$1_{\alpha Man}$), 4.87 (s, 1H, H-$1_{\beta Man}$), 4.59 (dd, *J*= 8.6, 7.9 Hz, 2H), 4.54 (d, *J*= 7.9 Hz, 1H), 4.49 (d, *J*= 7.9 Hz, 1H), 4.47 (dd, *J*= 7.8, 4.7 Hz, 2H), 4.44 (d, *J*= 7.7 Hz, 1H), 4.26-4.25 (d, *J*= 3.2 Hz, 1H, H-2_{Man}), 4.16-4.15 (m, 1H, H-2_{Man}), 4.11-4.10 (m, 1H, H-2_{Man}), 4.02-3.97 (m, 4H), 3.93-3.48 (m, 51H), 3.45-3.42 (m, *J*= 9.2, 2.3 Hz, 1H), 3.38-3.36 (m, 1H), 3.25 (t, *J*= 11.3 Hz, 1H), 2.98 (t, 2H, *J*= 7.7 Hz, linker), 2.08, 2.05, 2.04, 2.03 (4 s, 4x3H, NH*Ac*), 1.69-1.65 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.40-1.38 (m, 2H, linker); **^{13}C NMR(201 MHz, D2O) δ** 107.7, 105.6, 105.6, 104.0, 103.7, 103.2, 102.2, 102.1, 101.7, 99.7, 82.1, 82.05, 81.9, 81.1, 79.7, 79.1, 78.9, 78.0, 77.9, 77.9, 77.4, 77.4, 77.1, 77.0, 76.3, 76.2, 75.9, 75.5, 75.1, 74.7, 74.6, 74.4, 73.6, 73.5, 72.8, 72.7, 72.2, 72.1, 71.9, 71.9, 71.4, 71.2, 69.8, 69.1, 68.0, 67.9, 67.6, 67.6, 64.3, 64.2, 63.6, 62.9, 62.7, 62.6, 62.6, 62.6, 62.6, 62.5, 57.6, 57.6, 57.5, 41.9, 30.7, 29.0, 24.9, 24.9, 24.8, 24.7; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{72}H_{123}N_5O_{50}Na$ 1880.7128, found 1880.7042.

Heptasaccharide G18: Following the general procedure E, compound **G18** (680 µg, 99%) was prepared from **G3** (590 µg, 0.50 µmol) using concentrated *P. pastoris* extract containing β-1,4-GalNAcT (40 µL) and UDP-GalNAc (390 µg, 0.60 µmol); **^{1}H NMR (500 MHz,**

D_2O) δ 5.15 (s, 1H, H-$1_{\alpha Man}$), 4.86 (s, 1H, H-$1_{\beta Man}$), 4.61 (d, *J*= 7.5 Hz, 1H, H-$1_{\beta GlcNAc}$), 4.54-4.50 (m, 3H), 4.46 (d, *J*= 7.5 Hz, 1H, H-$1_{\beta Xyl}$), 4.26-4.24 (m, 1H, H-2_{Man}), 4.18-4.16 (m, 1H, H-2_{Man}), 4.05-3.99 (dd, *J*= 11.4, 5.5 Hz, 1H), 3.96-3.43 (m, 38H), 3.40-3.36 (m, 1H), 3.26 (t, *J* = 11.1 Hz, 1H), 2.98 (t, *J*= 7.7 Hz, 2H, linker), 2.08, 2.06, 2.04 (3 s, 3x3H, NH*Ac*), 1.71-1.65 (m, 2H, linker), 1.62-1.58 (m, 2H, linker), 1.44-1.39 (m, 2H, linker); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 107.6, 104.4, 104.0, 103.7, 102.7, 102.2, 101.8, 81.9, 81.8, 81.7, 81.3, 79.8, 79.1, 78.9, 78.0, 77.9, 77.2, 77.1, 76.3, 75.9, 75.0, 74.6, 74.4, 73.3, 72.8, 72.8, 72.1, 72.0, 70.2, 69.8, 69.1, 67.6, 64.4, 64.4, 63.6, 63.0, 62.7, 62.7, 62.6, 57.8, 57.6, 57.2, 55.1, 41.9, 30.7, 29.0, 24.8, 24.7, 24.7, 24.7**; HRMS (ESI)** ***m/z*** $[M+H]^+$ calculated for $C_{54}H_{94}N_5O_{35}$ 1372.5724, found 1372.5740.

Octasaccharide G19: Following the general procedure E, compound **G19** (500 µg, 99%) was prepared from **G4** (450 µg, 0.34 µmol), using concentrated *P. pastoris* extract containing β-1,4-GalNAcT (75 µL) and UDP-GalNAc (270 µg, 0.41 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.14 (s, 1H, H-$1_{\alpha Man}$), 4.92 (s, 1H, H-$1_{\alpha Man}$), 4.87 (s, 1H, H-$1_{\beta Man}$), 4.59 (d, *J*= 8.3 Hz, 1H), 4.52-4.50 (m, 2H), 4.49 (d, *J*= 7.9 Hz, 1H), 4.44 (d, *J*= 7.6 Hz, 1H), 4.25 (d, *J*= 3.3 Hz, 1H, H-2_{Man}), 4.14 (m, 1H, H-2_{Man}), 4.01 (dd, *J*= 11.7, 5.6 Hz, 1H), 3.99-3.98 (m, 1H), 3.97-3.54 (m, 39H), 3.52-3.49 (m, 3H), 3.43 (t, *J*= 9.2 Hz, 1H), 3.37 (dd, *J*= 9.3, 7.6 Hz, 1H), 3.25 (t, *J*= 11.1 Hz, 1H), 2.98 (t, *J*= 7.7 Hz, 2H, linker), 2.08, 2.07, 2.04, 2.03 (4 s, 4x3H, NH*Ac*), 1.69-1.64 (m, 2H, linker), 1.62-1.57 (m, 2H, linker), 1.40-1.37 (m, 2H); **^{13}C NMR (201 MHz, D2O, peaks from HSQC experiment) δ** 105.0, 101.6, 101.2, 101.0, 100.4, 99.5, 99.4, 98.8, 82.2, 81.9, 81.9, 81.8, 79.7, 79.1, 77.9, 77.9, 77.1, 77.0, 77.0, 76.3, 75.9, 75.3, 75.1, 74.6, 74.4, 73.3, 73.1, 72.7, 72.7, 72.4, 72.2, 72.0, 71.9, 70.2, 69.8, 69.3, 68.0, 64.4, 64.4, 63.5, 63.5, 62.9, 62.7, 62.7, 57.6, 57.6, 57.2, 55.1**,** 41.9, 30.7, 29.0, 24.9, 24.8, 24.8, 24.8, 24.7; **HRMS (ESI)** ***m/z*** $[M+H]^+$ calculated for $C_{60}H_{104}N_5O_{40}$ 1534.6252, found 1534.6299.

Octasaccharide G20: Following the general procedure E, compound G20 (700 µg, 99 %) was prepared from **G4** (610 µg, 0.46 µmol) using concentrated *P. pastoris* extract containing β-1,4-GalNAcT (120 µL) and UDP-GalNAc (752 µg, 1.15 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.14

(s, 1H, H-1$_{\alpha Man}$), 4.91 (s, 1H, H-1$_{\alpha Man}$), 4.89 (s, 1H, H-1$_{\beta Man}$), 4.61 (d, *J*= 7.9 Hz, 1H), 4.58-4.55 (d, *J*= 7.8 Hz, 1H), 4.52 (d, *J* = 8.3 Hz, 1H), 4.51 (d, *J*= 7.8 Hz, 1H), 4.46 (d, *J*= 7.6 Hz, 1H), 4.28 (d, *J*= 3.0 Hz, 1H, H-2$_{Man}$), 4.12-4.10 (m, 1H, H-2$_{Man}$), 4.05 (dd, *J*= 3.2, 1.4 Hz, 1H, H-2$_{Man}$), 4.03-3.58 (m, 40H), 3.53-3.49 (m, 3H), 3.45 (t, *J* = 9.26 Hz, 1H), 3.40-3.37 (m, 1H), 3.26 (t, *J*= 11.1 Hz, 1H), 3.00 (t, *J*= 7.7 Hz, 2H, linker), 2.10, 2.09, 2.05, 2.04 (4 s, 4x3H, NH*Ac*), 1.71-1.65 (m, 2H, linker), 1.62-1.58 (m, 2H, linker), 1.44-1.38 (m, 2H, linker); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ:**107.6, 104.8, 104.4, 104.0, 103.7, 103.3, 102.1, 99.5, 82.3, 82.2, 81.9, 79.9, 78.8, 78.0, 78.0, 77.1, 77.0, 76.2, 75.9, 75.5, 75.0, 74.7, 74.3, 73.3, 72.9, 72.8, 72.7, 72.6, 72.2, 72.2, 71.9, 70.2, 69.9, 69.4, 69.0, 68.0, 67.6, 67.5, 64.2, 63.8, 63.7, 62.8, 62.7, 62.7, 57.6, 57.6, 57.3, 55.1, 41.9, 30.7, 29.0, 24.9, 24.9, 24.8, 24.8, 24.7; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{60}H_{104}N_5O_{40}$ 1534.6252, found 1534.6201.

Decasaccharide G21: Following the general procedure E, compound **G21** (700 µg, 99%) was prepared from **G6** (570 µg, 0.37 µmol), using concentrated *P. pastoris* extract containing β-1,4-GalNAcT (150 µL) and UDP-GalNAc (960 µg, 1.48 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.15 (s, 1H, H-1$_{\alpha Man}$), 4.91 (s, 1H, H-1$_{\alpha Man}$), 4.88 (s, 1H, H-1$_{\beta Man}$), 4.61 (d, *J*= 8.0 Hz, 1H), 4.58-4.50 (m, 5H), 4.45 (d, *J*= 8.1 Hz, 1H), 4.28-4.26 (m, 1H, H-2$_{Man}$), 4.16-4.14 (m, 1H, H-2$_{Man}$), 4.01-3.42 (m, 56H), 3.38 (s, 2H), 3.28-3.23 (m, 1H), 2.97 (t, *J*= 7.4 Hz, 2H), 2.10-2.07 (m, 9H, NH*Ac*), 2.06-2.04 (m, 9H, NH*Ac*), 1.69-1.63 (m, 2H, linker), 1.62-1.57 (m, 2H, linker), 1.42-1.38 (m, 2H, linker); **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{76}H_{129}N_7O_{50}Na$ 1962.7659, found 1962.7672.

Octasaccharide G22: Following the general procedure D, compound **G22** (650 µg, 99%) was prepared from **G7** (470 µg, 0.36 µmol) using AtFucTA (25µL, 0.25 mg/mL), and GDP-Fuc (267 µg, 0.42 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s, 1H, H-1$_{\alpha Man}$), 5.14 (d, *J*= 4.1 Hz, 1H, H-$_{1Fuc\ \alpha1,3}$), 4.94 (d, *J*= 4.1 Hz, 1H, H-$_{1Fuc\ \alpha1,6}$), 4.84 (s, 1H, H-1$_{\beta Man}$), 4.74-4.72 (m, 1H, H-5$_{Fuc\ \alpha1,3)}$, 4.69 (d, *J*= 7.9 Hz, 1H), 4.52 (d, *J*= 8.5 Hz, 1H), 4.49 (t, *J*= 8.9Hz, 2H), 4.25 (d, *J*= 3.17 Hz, 1H, H-2$_{Man}$), 4.18-4.16 (m, 2H, H-2$_{Man}$,H-5$_{Fuc\ \alpha1,6}$), 4.03-3.46 (m, 39H), 3.40-3.35 (m, 1H), 3.27 (t, *J*= 11.3 Hz, 1H), 2.99 (t, *J*= 7.7 Hz, 2H, linker), 2.06, 2.05, 2.03 (3 s, 3x3H, NH*Ac*), 1.71-1.65

(m, 2H, linker), 1.62-1.57(m, 2H, linker), 1.41-1.36 (m, 2H, linker), 1.28 (d, *J*= 6.8 Hz, 3H, C$H_{3\text{Fuc }\alpha1,3}$), 1.24 (d, *J*= 6.7 Hz, 3H, C$H_{3\text{Fuc }\alpha1,6}$); **^{13}C NMR (201 MHz, D2O) δ** 107.6, 103.6, 102.7, 102.7, 102.4, 101.8, 101.6, 100.9, 82.7, 81.8, 79.7, 79.2, 78.8, 78.5, 77.9, 77.7, 77.3, 76.6, 76.3, 76.0,75.9, 75.9, 74.8, 74.4, 72.7, 72.5, 72.1, 72.0, 71.9, 71.8, 70.7, 70.4, 69.8, 69.4, 69.4, 69.1, 67.6, 67.5, 65.3, 64.4, 64.3, 63.2, 63.0, 58.3, 57.9, 57.9, 42.0, 30.7, 29.2, 24.8, 24.8, 24.8, 24.8, 18.1, 18.0; **HRMS (ESI)** *m/z* $[M+Na]^+$ calculated for $C_{58}H_{100}N_4O_{38}Na$ 1483.5908 found 1483.5907.

Nonasaccharide G23: Following the general procedure D, compound **G23** (340 µg, 91.5%) was prepared from **G8** (340 µg, 0.23 µmol) using AtFucTA (25 µL, 0.25 mg/mL) and GDP-Fuc (228 µg, 0.36 µmol); **^{1}H NMR (500 MHz, D_2O) δ** 5.16 (s, 1H, H-$1_{\alpha Man}$), 5.14 (d, *J*= 3.9 Hz, 1H, H-$_{1\text{Fuc }\alpha1,3}$), 4.94 (d, *J*= 3.9 Hz, 1H, H-$_{1\text{Fuc }\alpha1,6}$), 4.9 (d, *J*= 1.6 Hz, 1H, H-$1_{\alpha Man}$), 4.87 (s, 1H, H-$1_{\beta Man}$), 4.74-4.72 (m, 1H, H-$5_{\text{Fuc }\alpha1,3}$), 4.70 (d, *J*= 7.9 Hz, 1H), 4.53 (d, *J* = 8.6 Hz, 1H), 4.50 (d, *J* = 8.2 Hz, 1H), 4.48 (d, *J*= 7.8 Hz, 1H), 4.27 (d, *J*= 3.1 Hz, 1H, H-2_{Man}), 4.18-4.15 (m, 2H, H-2_{Man}, H-$5_{\text{Fuc }\alpha1,6}$), 4.05-3.45 (m, 42H), 3.41-3.36 (m, 1H), 3.28 (t, *J*= 11.3 Hz, 1H), 3.00 (t, *J*= 7.5 Hz, 2H, linker), 2.07, 2.06 ,2.03 (3 s, 3x3H, NH*Ac*), 1.71-1.64 (m, 2H, linker), 1.63-1.58 (m, 2H, linker), 1.43-1.38 (m, 2H, linker), 1.29 (d, *J*= 6.5 Hz, 3H, C$H_{3\text{Fuc }\alpha1,3}$), 1.24 (d, *J*= 6.7 Hz, 3H, C$H_{3\text{Fuc}\alpha1,6}$); **^{13}C NMR (201 MHz, D2O) δ** 107.8, 103.6, 103.2, 102.6, 102.4, 102.3, 101.8, 101.6, 100.9, 83.6, 82.05, 79.6, 79.2, 78.4, 77.9, 77.3, 77.1, 77.0, 76.6, 76.3, 76.0, 75.9, 75.8, 75.3, 75.00, 74.5, 73.1, 72.7, 72.4, 72.3, 72.1, 72.0, 71.9, 71.8, 70.7, 70.4, 70.0, 69.4, 69.4, 69.3, 69.2, 68.0, 68.1, 67.6, 67.5, 64.4, 64.3, 63.5, 63.2, 58.3, 57.9, 57.7, 42.0, 30.7, 29.3, 24.9, 24.8, 24.8, 24.8, 18.1, 18.0; **HRMS (ESI)** ***m/z*** $[M+Na]^+$ calculated for $C_{64}H_{110}N_4O_{43}Na$ 1645.6436 found 1645.6415.

Decasaccharide G24: Following the general procedure D, compound **G24** (1.64 mg, 93 %) was prepared from **G9** (1.62 mg, 0.96 µmol) using AtFucTA (30 µL, 0.25 mg/mL), and GDP-Fuc (728 µg, 1.15 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.15 (s, 1H, H-$1_{\alpha Man}$), 5.13 (d, *J*= 3.6 Hz, 1H, H-$1_{\text{Fuc}\alpha1\text{-}3}$), 4.94 (d, *J*= 3.6 Hz, 1H, H-$1_{\text{Fuc}\alpha1\text{-}6}$), 4.90 (s, 1H, H-$1_{\alpha Man}$), 4.86 (s, 1H, H-$1_{\beta Man}$), 4.68 (d, *J*= 8.3 Hz, 1H), 4.56 (d, *J*= 8.6 Hz, 1H), 4.51 (d, *J*= 8.6 Hz, 1H), 4.49-4.48 (m, 1H), 4.46

(d, *J*= 7.6 Hz, 1H), 4.26 (d, *J*= 3.2 Hz, 1H, H-2_{Man}), 4.16-4.15 (m, 2H, H-$2_{Man,}$ H-$5_{Fucα1-6}$), 4.11-4.10 (m, 1H), 4.03-3.42 (m, 51H), 3.38 (dt, *J*= 9.8, 1.8 Hz, 1H), 3.27 (t, *J*= 11.3 Hz, 1H), 2.97-2.94 (m, 2H, linker), 2.06 (s, 3H, NH*Ac*), 2.05 (m, 6H, NH*Ac*), 2.02 (s, 3H, NH*Ac*), 1.67-1.63 (m, 2H, linker), 1.60-1.57 (m, 2H, linker), 1.40-1.37 (m, 2H, linker), 1.27 (d, *J*= 6.6 Hz, 3H, $CH_{3Fucα1-3}$), 1.23 (d, *J*= 6.5 Hz, 3H, $CH_{3Fucα1-6}$); **^{13}C NMR(201 MHz, D2O) δ** 106.6, 102.4, 102.0, 101.4, 101.3, 100.9, 100.5, 100.4, 99.7, 98.5, 82.3, 80.8, 78.4, 77.9, 77.5, 77.2, 76.6, 76.0, 75.9, 75.8, 75.3, 75.0, 74.8, 74.7, 74.7, 74.2, 73.8, 73.3, 73.1, 71.4, 71.3, 70.9, 70.8, 70.6, 70.5, 69.9, 69.5, 69.1, 68.6, 68.2, 68.1, 67.8, 67.4, 67.4, 66.9, 66.4, 66.3, 63.1, 63.1, 62.9, 62.5, 62.4, 62.0, 62.0, 61.7, 57.0, 56.7, 56.4, 40.8, 29.5, 28.1, 23.7, 23.6, 23.6, 23.6, 23.5, 16.9, 16.8; **HRMS (ESI) *m/z*** $[M+Na+H]^{2+}$ calculated for $C_{72}H_{123}N_5O_{48}$ 924.8651, found 924.8627.

Octasaccharide G25: Following the general procedure E, compound **G25** (0.37 mg, 99%) was prepared from **G11** (0.32 mg, 0.24 µmol) using concentrated *P. pastoris* extract containing β-1,4-GalNacT (50 µL) and UDP-GalNAc (189 µg, 0.291 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.14-5.13 (m, 2H, H-$1_{αMan,}$ H-1_{Fuc}), 4.83 (s, 1H, H-$1_{βMan}$), 4.71-4.69 (m, 1H, H-5_{Fuc}), 4.54 (d, 1H, *J*=8.5 Hz), 4.52-4.50 (m, 2H), 4.48 (d, *J*= 8.2 Hz, 1H), 4.45 (d, *J*= 7.8 Hz, 1H), 4.25-4.23 (m, 1H, H-2_{Man}), 4.15-4.14 (m, 1H, H-2_{Man}), 4.00 (dd, *J*= 11.4, 5.6 Hz, 1H), 3.97 (dd, *J*= 10.5, 3.4 Hz, 1H), 3.94-3.43 (m, 40H), 3.37 (dd, *J*= 10.0, 7.8 Hz, 1H), 3.25 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.9 Hz, 2H, linker), 2.07, 2.04, 2.04, 2.02 (4 s, 4x3H, NH*Ac*), 1.69-1.65 (m, 2H, linker), 1.60-1.57 (m, 2H, linker), 1.41-1.38 (m, 2H, linker), 1.26 (d, *J*= 6.6 Hz, 3H, CH_{3Fuc}); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 107.6, 104.4, 103.7, 103.1, 102.7, 102.2, 101.7, 100.9, 82.5, 81.8, 81.7, 79.7, 79.0, 78.9, 78.0, 78.0, 77.9, 77.1, 76.4, 76.4, 75.9, 74.8, 74.7, 74.4, 73.3, 72.8, 72.7, 72.1, 72.0, 71.9, 70.4, 70.2, 69.8, 69.4, 69.1, 67.6, 67.5, 64.4, 64.4, 63.6, 63.6, 63.5, 62.9, 62.7, 62.7, 62.7, 58.2, 57.9, 57.2, 55.2, 30.7, 29.0, 24.9, 24.8, 24.8, 24.7, 18.1; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for 1540.6122 $C_{60}H_{103}N_5O_{39}$, found 1540.6124.

Octasaccharide G26: Following the general procedure C, compound **G26** (830 µg, 89%) was prepared from **G7** (830 µg, 0.63 µmol) using 87 mU of bovine milk β-1,4-galactosyltransferase and UDP-Gal (464 µg, 0.76 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.14 (s, 1H, H-$1_{\alpha Man}$), 4.91 (s, 1H), 4.85 (s, 1H), 4.66 (d, *J*= 8.3 Hz, 1H), 4.54 (d, *J*= 7.6 Hz, 1H), 4.50 (d, *J*= 8.6 Hz, 1H), 4.45 (dd, *J*= 7.9, 7.7 Hz, 2H), 4.24 (s, 1H, H-2_{Man}), 4.16 (s, 1H, H-2_{Man}), 4.14-4.12 (m, 1H, H-5_{Fuc}), 4.02-3.97 (m, 2H), 3.92-3.44 (m, 40H), 3.37 (t, *J*= 8.5 Hz, 1H), 3.24 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.9 Hz, 2H), 2.08, 2.05, 2.03 (3 s, 3x3H, NH*Ac*), 1.68-1.65 (m, 2H, linker), 1.62-1.58 (m, 2H, linker), 1.41-1.37 (m, 2H, linker), 1.23 (d, *J*= 6.6 Hz, 3H, CH_{3Fuc}); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 106.4, 104.3, 102.5, 102.5, 101.5, 101.0, 100.8, 100.5, 80.5, 79.8, 78.5, 77.9, 77.7, 76.7, 76.6, 76.1, 75.9, 75.0, 74.9, 74.7, 73.3, 73.1, 72.9, 72.3, 71.4, 70.9, 70.8, 70.7, 69.9, 69.5, 68.6, 68.2, 67.9, 67.8, 66.3, 66.3, 63.2, 63.1, 62.4, 61.6, 61.4, 61.3, 61.3, 61.3, 56.6, 56.5, 56.4, 40.7, 29.5, 27.81, 23.7, 23.6, 23.5, 23.5, 16.8; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{58}H_{100}N_4O_{39}Na$ 1499.5857, found 1499.5786.

Octasaccharide G27: Following the general procedure D, compound **G27** (380 µg, 99%) was prepared from **G14** (350 µg, 0.26 µmol) using CeFUT6 (20 µL, 1.07 mg/mL) and GDP-Fuc (202 µg, 0.32 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13 (s, 1H, H-$1_{\alpha Man}$), 5.13 (d, *J*= 4.1 Hz, 1H, H-1_{Fuc}), 4.85(s, 1H, H-$1_{\beta Man}$), 4.84-4.82 (m, 1H, H-5_{Fuc}), 4.59 (d, *J*= 8.1 Hz, 1H), 4.55-4.53 (m, 1H), 4.49 (d, *J*= 7.7 Hz, 1H), 4.44 (d, *J*= 8.2 Hz, 2H), 4.24 (d, *J*= 2.8 Hz, 1H, H-2_{Man}), 4.16-4.15 (m, 1H, H-2_{Man}), 4.03-3.99 (m, 2H), 3.96-3.84 (m, 14H), 3.80-3.56 (m, 21H), 3.52-3.46 (m, 4H), 3.43 (t, *J*= 9.3 Hz, 1H), 3.37 (dd, *J*= 9.5, 7.7 Hz, 1H), 3.24 (t, *J*= 11.1 Hz, 1H), 2.98 (t, *J* = 7.7 Hz, 2H, linker), 2.06, 2.04, 2.03 (3 s, 3x3H, NH*Ac*), 1.69-1.64 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.42-1.37 (m, 2H, linker), 1.17 (d, *J*= 6.6 Hz, 3H, CH_{3Fuc}); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 106.4, 103.2, 102.8, 102.4, 101.3, 100.8, 100.5, 99.93, 80.7, 80.5, 80.0, 78.5, 77.8, 77.7, 76.3, 75.9, 75.8, 75.1, 74.7, 74.6, 73.8, 73.3, 73.0, 72.5, 71.5, 70.8, 70.7, 70.6, 69.8, 69.1, 68.6, 68.1, 67.8, 66.4, 66.3, 63.2, 63.1, 62.9, 61.7, 61.6, 61.5, 61.4, 61.3, 61.1, 57.1, 56.6, 56.4, 40.7, 29.4, 27.8, 23.8, 23.6, 23.5, 23.4, 16.7; **HRMS (ESI) *m/z*** [M+H] calculated for $C_{58}H_{101}N_4O_{39}$ 1477.6037, found 1477.6034.

Nonasaccharide G28: Following the general procedure D, compound **G28** (380 µg, 99%) was prepared from **G15** (325 µg, 0.22 µmol) using CeFUT6 (15 µL, 1.07 mg/mL) and GDP-Fuc (164 µg, 0.26 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13 (s, 1H, H-$1_{\alpha Man}$), 5.12(d, *J* = 3.9 Hz, 1H, H-1_{Fuc}), 4.92 (s, 1H, H-$1_{\alpha Man}$), 4.87 (s, 1H, H-$1_{\beta Man}$), 4.60 (d, *J*= 8.0 Hz, 1H), 4.56-4.54 (m, 1H), 4.5 (d, *J* = 7.8 Hz, 1H), 4.45-4.43 (m, 2H), 4.26 (d, *J*= 2.7 Hz, 1H, H-2_{Man}), 4.16-4.15 (m, 1H, H-2_{Man}), 4.13-4.09 (m, 1H), 4.02-3.48 (m, 48H), 3.44 (t, *J*= 9.3 Hz, 1H), 3.37 (dd, *J*= 9.6, 8.0 Hz, 1H), 3.25 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.8 Hz, 2H, linker), 2.08, 2.04, 2.03 (3 s, 3x3H, NH*Ac*), 1.68-1.65 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.41-1.38 (m, 2H, linker), 1.18 (d, *J*= 6.5 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (201 MHz, D2O, peaks from HSQC experiment) δ** 107.8, 104.4, 104.0, 103.7, 103.2, 102.3, 101.9, 101.7, 101.2, 82.1, 82.0, 81.9, 79.7, 79.0, 77.9, 77.9, 77.6, 77.2, 77.1, 77.0, 76.3, 75.9, 75.9, 75.3, 75.0, 75.0, 74.6, 74.6, 74.4, 73.7, 73.4, 72.7, 72.7, 72.4, 72.0, 71.9, 71.7, 70.9, 70.3, 69.8, 69.5, 69.3, 69.2, 68.2, 67.6, 67.5, 64.8, 64.1, 63.5, 62.9, 63.0, 62.3, 57.6, 57.6, 30.7, 29.1, 25.0, 24.8, 24.8, 24.7, 17.9; **HRMS (ESI)** ***m/z*** $[M+Na+H]^{2+}$ calculated for $C_{64}H_{110}N_4O_{44}$ 831.3229, found 831.3253.

Nonasaccharide G29: Following the general procedure D, compound **G29** (260 µg, 74%) was prepared from **G16** (320 µg, 0.22 µmol) using CeFUT6 (15 µL, 1.07 mg/mL) and GDP-Fuc (164 µg, 0.26 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.14-5.12 (m, 2H, H-$1_{\alpha Man,}$ H-$1_{Fuc\alpha\text{-}1,3}$), 4.90 (s, 1H, H-$1_{\alpha Man}$), 4.88 (s, 1H, H-$1_{\beta Man}$), 4.59-4.58 (m, 2H), 4.49 (d, *J*= 7.62 Hz, 1H), 4.44 (dd, *J*= 7.91, 2.8 Hz, 2H), 4.27 (d, *J*= 2.8 Hz, 1H, H-2_{Man}), 4.11-4.10 (m, 1H), 4.04-3.54 (m, 45H), 3.51-3.47 (m, 3H), 3.44 (t, *J*= 9.4 Hz, 1H), 3.39-3.36 (m, 1H), 3.25 (t, *J*= 11.1 Hz, 1H), 2.98 (t, *J*= 7.7 Hz, 2H, linker), 2.08, 2.04, 2.03 (3 s, 3x3H, NH*Ac*), 1.68-1.63 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.40-1.38 (m, 2H, linker), 1.17 (d, *J*= 6.6 Hz, 3H, $CH_{3Fuc\alpha\text{-}1,3}$); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 107.7, 104.8, 104.4, 104.0, 103.7 ,103.3, 101.8, 101.2, 99.4, 82.2, 81.9, 81.9, 79.9, 78.7, 77.9, 77.9, 77.6, 77.3, 77.1, 77.0, 76.2, 75.9, 75.9, 75.5, 75.1, 75.1, 74.7, 74.6, 74.3, 73.7, 72.9, 72.7, 72.6, 72.2, 71.8, 71.1, 71.0, 70.3, 69.9, 69.4, 69.3, 69.1, 67.5, 64.1, 63.8, 63.0, 62.7, 62.7, 62.3, 62.1, 58.3, 57.6, 57.6, 41.9, 30.7, 28.9, 25.0, 24.8, 24.8, 22.6, 17.9; **HRMS (ESI)** ***m/z*** $[M+Na+H]^{2+}$ calculated for $C_{64}H_{111}N_4O_{44}Na$ 831.3229, found 831.3199.

Dodecasaccharide G30: Following the general procedure D, compound **G30** (400 µg, 99%) was prepared from **G17** (360 µg, 0.19 µmol) using CeFUT6 (15 µL, 1.07 mg/mL) and GDP-Fuc (291 µg, 0.46 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13-5.11 (m, 3H, H-$1_{\alpha Man,}$ H-$1_{Fuc\alpha\text{-}1,3,}$ H-$1_{Fuc\alpha\text{-}1,3}$), 4.89 (s, 1H, H-$1_{\alpha Man}$), 4.87(s, 1H, H-$1_{\beta Man}$), 4.84 (m, 1H H-$5_{Fuc\alpha\text{-}1,3}$), 4.60-4.58 (m, 2H), 4.55-4.53 (m, 1H), 4.48 (d, *J*= 8.1 Hz, 1H), 4.45-4.43 (m, 3H), 4.25 (d, *J*= 3.1 Hz, 1H, H-2_{Man}), 4.16-4.15 (m, 1H, H-2_{Man}), 4.11-4.10 (m, 1H, H-2_{Man}), 4.02-3.55 (m, 57H), 3.50-3.47 (m, 5H), 3.44-3.42 (m, 1H), 3.39-3.36 (m, 1H), 3.24 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.7 Hz, 2H, linker), 2.08, 2.04, 2.04, 2.03 (4 s, 4x3H, NH*Ac*), 1.68-1.64 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.41-1.38 (m, 2H, linker), 1.17 (m, 6H, $CH_{3Fuc\alpha\text{-}1,3,}$ $CH_{3Fuc\alpha\text{-}1,3}$); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 108.0**,** 104.5**,** 104.4**,** 103.9**,** 103.7**,** 103.2**,** 101.9**,** 101.8**,** 101.7**,** 101.2**,** 101.2**,** 99.5, 82.2, 82.0, 81.9, 79.9, 79.0, 78.7, 77.9, 77.9, 77.6, 77.2, 77.1, 76.9, 76.3, 75.9, 75.9, 75.5, 75.1, 75.0, 75.0, 74.7, 74.5, 74.4, 73.7, 73.7, 72.7, 72.2, 72.2, 72.2, 72.1, 71.9, 71.8, 71.7, 71.0, 70.3, 70.3, 69.8, 69.3, 69.3, 69.1, 68.3, 68.1, 67.6, 67.6, 64.4, 64.3, 64.2, 64.1, 63.0, 62.8, 62.7, 62.7, 62.6, 62.4, 62.3, 58.3, 58.3, 57.6, 57.6, 41.9, 30.7, 29.0, 25.0, 25.0, 24.8, 24.8, 24.7, 17.9, 17.9; **HRMS (ESI) *m/z*** $[M+K]^+$ calculated for $C_{84}H_{143}N_5O_{57}K$ 2172.8286, found 2172.8364.

Nonasaccharide G31: Following the general procedure D, compound **G31** (420 µg, 79%) was prepared from **G26** (480 µg, 0.32 µmol) using CeFUT6 (30 µL, 1.07 mg/mL) and GDP-Fuc (297 µg, 0.47 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13 (s, 1H, H-$1_{\alpha Man}$), 5.12 (d, *J* = 4.1 Hz, 1H, H-$1_{Fuc\alpha\text{-}1,3}$), 4.90 (d, *J* = 3.9 Hz, 1H, H-$1_{Fuc\alpha\text{-}1,6}$), 4.85 (s, 1H, H-$1_{\beta Man}$), 4.83-4.81 (m, 1H, H-$5_{Fuc\alpha\text{-}1,3}$), 4.65 (d, *J*= 8.0 Hz, 1H), 4.55-4.53 (m, 1H), 4.50 (d, *J*= 8.5, 1H), 4.44 (dd, *J*= 7.4, 3.9 Hz, 2H), 4.24 (d, *J*= 2.94 Hz, 1H, H-$2_{Man\beta}$), 4.16-4.15 (m, 1H, H-$2_{Man\alpha}$), 4.12 (q, *J*= 6.7 Hz, 1H, H-$5_{Fuc\alpha\text{-}1,6}$), 4.02-3.99 (m, 2H), 3.96-3.56 (m, 39H), 3.52-3.48 (m, 3H), 3.45 (t, *J*= 9.2 Hz, 1H), 3.38-3.36 (m, 1H), 3.25 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.7 Hz, 2H, linker), 2.08, 2.04, 2.03 (3 s, 3x3H, NH*Ac*), 1.69-1.65 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.41-1.37 (m, 2H, linker), 1.22 (d, *J*= 6.6 Hz, 3H, $CH_{3Fuc\alpha\text{-}1,3}$), 1.17 (d, *J*= 6.5 Hz, 3H, $CH_{3Fuc\alpha\text{-}1,6}$); **^{13}C NMR(201 MHz, D2O) δ** 106.4, 103.2, 102.5, 102.4, 101.5, 100.8, 100.7, 100.4, 99.9, 80.6, 80.1, 78.5, 77.8, 77.7, 76.7, 76.6, 76.0, 75.9, 75.1, 74.9, 74.7, 74.6, 73.8, 73.4, 73.3, 73.1, 72.5, 71.4, 71.4, 70.8,

70.8, 70.6, 69.7, 69.5, 69.1, 68.6, 68.2, 68.1, 67.9, 67.8, 66.3, 66.3, 63.1, 62.9, 61.6, 61.6, 61.5, 61.1, 61.0, 57.1, 56.5, 56.4, 40.6, 29.4, 27.8, 23.7, 23.6, 23.5, 23.5, 16.7, 16.7; **HRMS (ESI) *m/z*** $[M+Na]^+$ calculated for $C_{64}H_{110}N_4O_{43}Na$ 1645.6436, found 1645.6499.

Dodecasaccharide G32: Following the general procedure E, compound **G32** (620 µg, 63%) was prepared from **G24** (900 µg, 0.49 µmol) using concentrated *P. pastoris* extract containing β-1,4-GalNAcT (210 µL) and UDP-GalNAc (320 µg, 1.97 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.15-5.13 (m, 2H, H-$1_{\alpha Man}$, H-$1_{Fuc\alpha 1\text{-}3}$), 4.95, (m, 1H, H-$1_{Fuc\alpha 1\text{-}6}$), 4.90 (m, 1H, H-$1_{\alpha Man}$), 4.86 (s, 1H, H-$1_{\beta Man}$), 4.69 (m, 1H), 4.57-4.45 (m, 6H), 4.26-4.25 (m, 1H, H-2_{Man}), 4.16-4.15 (m, 2H, H-$2_{Man,}$ H-$5_{Fuc\alpha 1\text{-}6}$), 4.10-4.09 (m, 2H), 4.03-3.34 (m, 69H), 2.98 (m, 2H, linker), 2.08, 2.07, 2.05, 2.02 (4 s, 4x3H, NH*Ac*), 1.69-1.64 (m, 2H, linker), 1.62-1.55 (m, 2H, linker), 1.42-1.36 (m, 2H, linker), 1.28 (d, *J*= 6.7 Hz, 3H, $CH_{3Fuc\alpha 1\text{-}3}$), 1.24 (d, *J*= 6.6 Hz, 3H, $CH_{3Fuc\alpha 1\text{-}6}$); **HRMS (ESI) *m/z*** $[M+Na]^+$ calculated for $C_{88}H_{149}N_7O_{58}$ 2254.8817, found 2254.885.

Tetradecasaccharide G33: Following the general procedures C and D, compound **G24** (1.2 mg, 0.67 mmol) was first galactosylated employing bovine milk β-1,4-GalT (35 µL) and UDP-Gal. The galactosylated intermediate (714 µg, 0.331 µmol) was purified and fucosylated using CeFUT6 (75 µL, 1.07 mg/mL) and GDP-Fuc (633 µg, 1.0 µmol) to produce compound **G33** (800 µg, 99%); **^{1}H NMR (500 MHz, D_2O)** δ 5.16-5.13 (m, 4H, H-$1_{\alpha Man,}$ H-$1_{Fuc\alpha\text{-}1,3,}$ H-$1_{Fuc\alpha\text{-}1,3,}$ H-$1_{Fuc\alpha\text{-}1,3}$), 4.95 (d, *J* = 3.61 Hz, 1H), 4.91 (s, 1H, H-$1_{\alpha Man,}$), 4.87-4.83 (m, 3H, H-$1_{\beta Man,}$ 2xH-$5_{Fuc\alpha 1,3/LewisX,}$), 4.73-4.69 (m, 1H, H-$5_{Fuc\alpha 1,3}$), 4.69 (d, *J*= 8.0 Hz, 1H), 4.60 (d, *J*= 8.4 Hz, 1H), 4.56 (d, *J*= 7.6 Hz, 1H), 4.50 (d, *J*=8.4 Hz, 1H), 4.48-4.45 (m, 3H), 4.27-4.26 (m, 1H, H-2_{Man}), 4.17-4.16 (m, 1H, H-$2_{Man,}$), 4.11-4.10 (m, 1H, H-$2_{Man,}$), 4.04 –3.47 (m, 70H), 3.42-3.37 (m, 1H), 3.28 (t, *J* = 10.8 Hz, 1H), 3.01 (t, *J*= 7.7 Hz, 2H, linker), 2.08, 2.05, 2.03 (3 s, 3x3H, NH*Ac*), 1.70-1.57 (m, 4H, linker), 1.43-1.38 (m, 2H,linker), 1.29 (d, *J* = 6.54 Hz, 3H, $CH_{3Fuc\alpha 1,3}$), 1.24 (d, *J*= 6.5 Hz, 3H, $CH_{3Fuc\alpha 1,6}$), 1.17 (m, 6H, 2x$CH_{3Fuc/LewisX}$); **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{96}H_{164}N_5O_{66}$ 1232.9759, found 1232.9803.

Hexasaccharide G34: Following the general procedure F, compound **G34** (670 µg, 99 %) was prepared from **G6** (924 µg, 0.60 µmol) using 1.4 U of Jack bean β-N-acetylglucosaminidase;

^{1}H NMR (800 MHz, D_2O) δ 5.12 (d, *J*= 1.6 Hz, 1H, H-$1_{\alpha Man}$), 4.91 (d, *J*= 1.8 Hz, 1H, H-$1_{\alpha Man}$), 4.87 (s, 1H, H-$1_{\beta Man}$), 4.60 (d, *J*= 8.2 Hz, 1H, H-1_{GlcNAc}), 4.50 (d, *J*= 8.0 Hz, 1H, H-1_{GlcNAc}), 4.45 (d, *J*= 7.7, 1H, H-1_{Xyl}), 4.27 (d, *J*= 3.1 Hz, 1H, H-$2_{Man\beta}$), 4.04 (dd, *J*= 3.38 Hz, *J* =1.6 Hz, 1H, H-$2_{Man\alpha}$), 4.01 (dd, *J*= 11.62 Hz, *J* =5.4 Hz, 1H), 4.00-3.94 (m, 3H), 3.92-3.58 (m, 27H), 3.51-3.49 (m, 1H), 3.44 (td, *J*= 9.2Hz, *J* =2.3 Hz, 1H), 3.39-3.37 (m, 1H), 3.25 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.9 Hz, 2H, linker), 2.08, 2.03 (2 s, 2x3H, NH*Ac*), 1.69-1.65 (m, 2H, linker), 1.61-1.57 (m, 2H, linker), 1.41-1.38 (m, 2H, linker); **^{13}C NMR(201 MHz, D2O) δ** 107.7, 104.8, 103.9, 103.7, 103.2, 102.2, 82.2, 81.9, 81.9, 79.9, 77.9, 77.2, 77.1, 76.1, 75.9, 75.2, 75.0, 74.7, 73.1, 72.6, 72.4, 72.2, 71.9, 69.1, 68.0, 67.6, 67.5, 63.9, 63.8, 63.5, 63.5, 62.8, 62.7, 62.7, 57.6, 57.6, 41.9, 30.7, 28.9, 24.8, 24.7, 24.7; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{44}H_{77}N_3O_{30}Na$ 1150.4484, found 1150.4433.

Hexasaccharide G35: Following the general procedure F, compound **G35** (1 mg, 90%) was prepared from **G7** (1.3 mg, 0.99 µmol) using 0.5 U of jack bean β-N-acetylglucosaminidase; **^{1}H NMR (500 MHz, D_2O) δ** 5.14 (s, 1H, H-$1_{\alpha Man}$), 4.92 (d, *J*= 3.9 Hz, 1H, H-$1_{\alpha Fuc}$), 4.87 (s, 1H, H-$1_{\beta Man}$), 4.67 (d, *J*= 7.7 Hz, 1H), 4.51 (d, *J*= 7.7 Hz, 1H), 4.47 (d, *J*= 7.9 Hz, 1H), 4.25 (d, *J*= 3.1 Hz, 1H, H-2_{Man}), 4.14 (dd, *J*= 13.4 Hz, *J*= 6.22 Hz, 1H, H-5_{Fuc}), 4.06 (dd, *J*= 3.3, 1.5 Hz, 1H, H-2_{Man}), 4.04-3.99 (m, 2H), 3.94-3.58 (m, 20H), 3.51-3.45 (m, 2H), 3.41-3.36 (m, 1H), 3.26 (t, *J*= 11.4 Hz, 1H), 3.00 (t, *J*= 7.6 Hz, 2H, linker), 2.09, 2.05 (2 s, 2x3H, NH*Ac*), 1.71-1.67 (m, 2H, linker), 1.64-1.58 (m, 2H, linker), 1.44-1.39 (m, 2H, linker), 1.24 (d, *J*= 6.7 Hz, 3H, CH_{3Fuc}); **^{13}C NMR (126 MHz, D_2O) δ** 104.9, 102.2, 101.2, 101.1, 100.1, *99.4*, 79.1, 78.7, 77.3, 75.4, 74.5, 73.5, 73.3, 72.4, 72.3, 72.0, 71.8, 71.7, 70.3, 70.0, 69.9, 69.6, 69.1, 68.1, 66.8, 66.7, 66.5, 66.43, 64.90, 64.87, 61.7, 60.2, 60.2, 55.1, 55.1, 39.3, 28.1, 26.4, 22.2, 22.1, 22.1, 15.3; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{45}H_{77}N_3O_{29}$ 1112.4716, found 1112.4758**.**

Heptasaccharide G37: Following the general procedure D, compound **G37** (430 µg, 99%) was prepared from **G34** (340 µg, 0.30 µmol) using CeFUT1 (20 µL, 0.9 mg/mL), and GDP-Fuc (272 µg, 0.43 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13 (d, *J*= 3.7 Hz, 1H, H-1_{Fuc}), 5.12 (s, 1H, H-

$1_{\alpha Man}$), 4.91 (s, 1H, H-$1_{\alpha Man}$), 4.85 (s, 1H, H-$1_{\beta Man}$), 4.72-4.70 (m, 1H, H-5_{Fuc}), 4.54 (d, *J*= 8.3 Hz, 1H), 4.49-4.46 (m, 2H), 4.26 (d, *J*= 3.1 Hz, 1H, H-$2_{Man\beta}$), 4.04-4.03 (m, 1H, H-$2_{Man\alpha}$), 4.03-3.44 (m, 36H), 3.39-3.37 (m, 1H), 3.26 (t, *J*= 11.0 Hz, 1H), 2.98 (t, *J*= 7.8 Hz, 2H, linker), 2.05, 2.02 (2 s, 2x3H, NH*Ac*), 1.69-1.65 (m, 2H, linker), 1.59-1.57 (m, 2H, linker), 1.41-1.38 (m, 2H, linker), 1.27 (d, *J* = 6.7 Hz, 3H, CH_{3Fuc}); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 106.4, 103.7, 102.4, 101.9, 101.8, 101.1, 99.7, 82.1, 80.6, 78.7, 76.7, 76.7, 75.9, 75.9, 75.7, 75.2, 74.9, 74.7, 74.0, 73.7, 73.4, 73.1, 71.7, 71.7, 71.3, 71.2, 70.7, 70.5, 69.1, 68.2, 68.1, 67.9, 66.6, 66.4, 62.6, 62.4, 62.3, 62.3, 61.7, 57.0, 56.5, 40.7, 29.5, 27.8, 23.6, 23.5, 23.6, 16.8; **HRMS (MALDI-tof) *m/z*** $[M+Na]^+$ calculated for $C_{50}H_{87}N_3O_{34}$ 1296.5063, found 1296.512.

Heptasaccharide G38: Following the general procedure D, compound **G38** (570 µg, 99%) was prepared from **G35** (460 µg, 0.41 µmol) using CeFUT1 (20 µL, 0.9 mg/mL) and GDP-Fuc (310 µg, 0.49 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13 (d, *J*= 3.8 Hz, 1H, H-$1_{Fuc\alpha1,3}$), 5.12 (s, 1H, H-$1_{\alpha Man}$), 4.93 (d, *J*= 3.8 Hz, 1H, H-$1_{Fuc\alpha1,6}$), 4.91 (s, 1H, H-$1_{\alpha Man}$), 4.85 (s, 1H. H-$1_{\beta Man}$), 4.73-4.71 (m, 1H, H-$5_{Fuc\alpha1,3}$), 4.69 (d, *J*= 8.6 Hz, 1H), 4.48 (d, *J* = 8.9 Hz, 1H), 4.47 (d, *J* = 7.6 Hz, 1H), 4.26 (d, *J*= 3.1 Hz, 1H, H-$2_{Man\beta}$), 4.15 (q, *J*= 6.8 Hz, 1H, H-$5_{Fuc\alpha1,6}$), 4.04 (dd, *J*= 3.4, 1.4 Hz, 1H, H-$2_{Man\alpha}$), 3.99-3.55 (m, 37H), 3.49 (t, *J*= 9.2 Hz, 1H), 3.46 (t, *J*= 8.9 Hz, 1H), 3.38 (dd, *J* = 9.4, 7.5 Hz, 1H), 3.27 (t, *J*= 11.3 Hz, 1H), 2.98 (t, *J*= 7.9 Hz, 2H, linker), 2.06, 2.02 (2 s, 2x3H, NH*Ac*), 1.68-1.64 (m, 2H, linker), 1.60-1.57 (m, 2H, linker), 1.40-1.36 (m, 2H, linker), 1.28 (d, *J*= 6.6 Hz, 3H, $CH_{3Fuc\alpha1,3}$), 1.23 (d, *J*= 6.6 Hz, 3H, $CH_{3Fuc\alpha1,6}$); **^{13}C NMR(201 MHz, D2O, peaks from HSQC experiment) δ** 104.5, 101.9, 100.6, 99.6, 99.6, 98.0, 97.9, 79.6, 78.7, 76.8, 75.8, 74.8, 74.3, 74.2, 73.6, 73.0, 72.9, 71.5, 71.3, 69.8, 69.6, 69.5, 69.5, 69.1, 68.7, 67.6, 66.5, 66.4, 65.7, 65.3, 64.9, 64.4, 64.4, 60.7, 59.9, 59.9, 58.7, 56.7, 55.3, 55.0, 38.9, 27.7, 26.1, 21.8, 21.7, 21.7, 15.1, 14.9; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{50}H_{88}N_3O_{33}$ 1258.5295, found 1258.5330.

Octasaccharide G39: Following the general procedure D, compound **G39** (360 µg, 99%) was prepared from **G36** (320 µg, 0.26 µmol) using CeFUT1 (20 µL, 0.9 mg/mL) and GDP-Fuc (190

µg, 0.31 µmol); **^{1}H NMR (800 MHz, D_2O) δ** 5.13 (d, *J* = 3.8 Hz, 1H, H-$1_{Fucα\text{-}1,3}$), 5.12 (s, 1H, H-$1_{αMan}$), 4.93 (d, *J* = 3.8 Hz, 1H, H-$1_{Fucα1,6}$), 4.91 (s, 1H, H-$1_{αMan}$), 4.85 (s, 1H. H-$1_{βMan}$), 4.73-4.71 (m, 1H, H-$5_{Fuc1,3}$), 4.69 (d, *J*= 8.6 Hz, 1H), 4.48 (d, *J*= 8.9 Hz, 1H), 4.47 (d, *J*= 7.6 Hz, 1H), 4.26 (d, *J*= 3.1 Hz, 1H, H-$2_{Manβ}$), 4.15 (q, *J*= 6.8 Hz, 1H, H-$5_{Fucα1,6}$), 4.04 (dd, *J*= 3.4, 1.4 Hz, 1H, H-$2_{Manα}$), 3.99-3.55(m, 37H), 3.49 (t, *J*= 9.2 Hz, 1H), 3.46 (t, *J* = 8.9 Hz, 1H), 3.38 (dd, *J*= 9.4, 7.5 Hz, 1H), 3.27 (t, *J*= 11.3 Hz, 1H), 2.98 (t, *J*= 7.9 Hz, 2H, linker), 2.06, 2.02 (2 s, 2x3H, NH*Ac*), 1.68-1.64 (m, 2H, linker), 1.60-1.57 (m, 2H, linker), 1.40-1.36 (m, 2H, linker), 1.28 (d, *J*= 6.6 Hz, 3H, $CH_{3Fucα1,3}$), 1.23 (d, *J*= 6.6 Hz, 3H, $CH_{3Fucα1,6}$); **^{13}C NMR(201 MHz, D_2O) δ:** 106.5, 103.6, 102.4, 102.0, 101.4, 101.1, 100.34, 99.71, 82.3, 80.8, 78.6, 76.6, 76.0, 75.9, 75.8, 75.4, 74.9, 74.8, 74.7, 74.1, 73.7, 73.3, 71.8, 71.3, 71.3, 71.2, 70.9, 70.9, 70.6, 70.5, 69.4, 68.2, 68.2, 68.1, 67.8, 67.5, 66.8, 66.8, 66.3, 62.6, 62.6, 62.4, 62.3, 62.3, 56.9, 56.4, 40.7, 29.5, 27.8, 23.6, 23.6, 23.5, 16.9, 16.7; **HRMS (ESI) *m/z*** $[M+H]^+$ calculated for $C_{56}H_{98}N_3O_{38}$ 1420.5823, found 1420.5786.

References

(1) Walczak, M. A.; Danishefsky, S. J. *Journal of the American Chemical Society* **2012**, *134*, 16430–16433.

(2) Rademann, J.; Geyer, A.; Schmidt, R. R. *Angewandte Chemie International Edition* **1998**, *37*, 1241–1245.

(3) Seeberger, P. H. *Carbohydrate research* **2008**, *343*, 1889–1896.

(4) Plante, O. J.; Palmacci, E. R.; Seeberger, P. H. *Science* **2001**, *291*, 1523–1527.

(5) Paulsen, H. *Chemical Society Reviews* **1984**, *13*, 15–45.

(6) Mori, M.; Ito, Y.; Ogawa, T. *Carbohydrate research* **1990**, *195*, 199–224.

(7) Nakahara, Y.; Shibayama, S.; Nakahara, Y.; Ogawa, T. *Carbohydrate research* **1996**, *280*, 67–84.

(8) Ogawa, T.; Sugimoto, M.; Kitajima, T.; Sadozai, K. K.; Nukada, T. *Tetrahedron letters* **1986**, *27*, 5739–5742.

(9) Wang, Z.; Chinoy, Z. S.; Ambre, S. G.; Peng, W.; McBride, R.; de Vries, R. P.; Glushka, J.; Paulson, J. C.; Boons, G.-J. *Science* **2013**, *341*, 379–383.

(10) Blixt, O.; Razi, N. *Methods in enzymology* **2006**, *415*, 137–153.

(11) Ichikawa, Y.; Lin, Y. C.; Dumas, D. P.; Shen, G. J.; Garcia-Junceda, E.; Williams, M. A.; Bayer, R.; Ketcham, C.; Walker, L. E. *Journal of the American Chemical Society* **1992**, *114*, 9283–9298.

(12) Unverzagt, C. *Carbohydrate research* **1997**, *305*, 423–431.

(13) Wong, C. H.; Haynie, S. L.; Whitesides, G. M. *The Journal of Organic Chemistry* **1982**, *47*, 5416–5418.

(14) Unverzagt, C.; Kunz, H.; Paulson, J. C. *Journal of the American Chemical Society* **1990**, *112*, 9308–9309.

(15) Paulson, J. C.; Beranek, W. E.; Hill, R. L. *Journal of Biological Chemistry* **1977**, *252*, 2356–2362.

(16) Unverzagt, C. *Angewandte Chemie International Edition in English* **1996**, *35*, 2350–2353.

(17) Koeller, K. M.; Wong, C.-H. *Chemical Reviews* **2000**, *100*, 4465–4494.

(18) Unverzagt, C. *Tetrahedron letters* **1997**, *38*, 5627–5630.

(19) Weijers, C. A.; Franssen, M. C.; Visser, G. M. *Biotechnology advances* **2008**, *26*, 436–456.

(20) Kajiura, H.; Okamoto, T.; Misaki, R.; Matsuura, Y.; Fujiyama, K. *Journal of Bioscience and Bioengineering* **2011**, *113*, 1347-4421.

(21) Kawar, Z. S.; Haslam, S. M.; Morris, H. R.; Dell, A.; Cummings, R. D. *Journal of Biological Chemistry* **2005**, *280*, 12810–12819.

(22) Wuhrer, M.; Koeleman, C. A. M.; Deelder, A. M.; Hokke, C. H. *FEBS Journal* **2006**, *273*, 347–361.

(23) Hokke, C. H.; Deelder, A. M.; Hoffmann, K. F.; Wuhrer, M. *Experimental parasitology* **2007**, *117*, 275–283.

(24) Van Die, I.; Cummings, R. D. *Glycobiology* **2010**, *20*, 2-12.

(25) Li, L.; Liu, Y.; Ma, C.; Qu, J.; Calderon, A. D.; Wu, B.; Wei, N.; Wang, X.; Guo, Y.; Xiao, Z.; others. *Chemical Science* **2015**, *6*, 5652–5661.

(26) Van den Berg, T. K.; Honing, H.; Franke, N.; van Remoortere, A.; Schiphorst, W. E. C. M.; Liu, F.-T.; Deelder, A. M.; Cummings, R. D.; Hokke, C. H.; van Die, I. *The Journal of Immunology* **2004**, *173*, 1902–1907.

(27) Martin, A.; Arda, A.; Désiré, J.; Martin-Mingot, A.; Probst, N.; Sinaÿ, P.; Jiménez-Barbero, J.; Thibaudeau, S.; Blériot, Y. *Nature Chemistry* **2015**, *8*, 186-191.

(28) Lemieux, R.; Hendriks, K.; Stick, R.; James, K. *Journal of the American Chemical Society* **1975**, *97*, 4056–4062.

(29) Stork, G.; Kim, G. *Journal of the American Chemical Society* **1992**, *114*, 1087–1088.

(30) Lemieux, R. *Advances in carbohydrate chemistry* **1953**, *9*, 1–57.

(31) Ness, R. K.; Fletcher Jr, H. G. *Journal of the American Chemical Society* **1956**, *78*, 4710–4714.

(32) Ness, R. K.; Fletcher Jr, H. G.; Hudson, C. *Journal of the American Chemical Society* **1951**, *73*, 959–963.

(33) Fischer, E. *Chemische Berichte* **1893**, *26*, 2400–2412.

(34) Michael, A. *American Chemical Journal* **1879**, *1*, 305–312.

(35) Koenigs, W.; Knorr, E. *Chemische Berichte* **1901**, *34*, 957-981.

(36) Zemplen, G.; Gerecs, A. *Chemische Berichte* **1930**, *63B*, 2720–2729.

(37) Serna, S.; Etxebarria, J.; Ruiz, N.; Martin-Lomas, M.; Reichardt, N. C. *Chemistry-A European Journal* **2010**, *16*, 13163–13175.

(38) Serna, S.; Kardak, B.; Reichardt, N. C.; Martin-Lomas, M. *Tetrahedron: Asymmetry* **2009**, *20*, 851–856.

(39) Unverzagt, C.; Eller, S.; Mezzato, S.; Schuberth, R. *Chemistry-A European Journal* **2008**, *14*, 1304–1311.

(40) Paulsen, H.; Helpap, B. *Carbohydrate Research* **1991**, *216*, 289–313.

(41) Schmidt, R. R.; Michel, J. *Angewandte Chemie International Edition in English* **1980**, *19*, 731–732.

(42) Schmidt, R. R. *Angewandte Chemie International Edition in English* **1986**, *25*, 212–235.

(43) Schmidt, R. R.; Castro-Palomino, J. C.; Retz, O. *Pure and applied chemistry* **1999**, *71*, 729–744.

(44) Schmidt, R. R.; Kinzy, W. *Advances in* Carbohydrate Chemistry *and* Biochemistry **1994**, *50*, 21–123.

(45) Schmidt, R. R.; Michel, J.; Roos, M. *Liebigs Annalen der Chemie* **1984**, *1984*, 1343–1357.

(46) Yu, B.; Sun, J. *Chemical Communications* **2010**, *46*, 4668–4679.

(47) Kerekgyarto, J.; van der Ven, J. G. M.; Kamerling, J. P.; Liptak, A.; Vliegenthart, J. F. G. *Carbohydrate research* **1993**, *238*, 135–145.

(48) Wright, J. A.; Yu, J.; Spencer, J. B. *Tetrahedron Letters* **2001**, *42*, 4033–4036.

(49) Xia, J.; Abbas, S. A.; Locke, R. D.; Piskorz, C. F.; Alderfer, J. L.; Matta, K. L. *Tetrahedron Letters* **2000**, *41*, 169–173.

(50) Nakano, J.; Ohta, H.; Ito, Y. *Bioorganic & medicinal chemistry letters* **2006**, *16*, 928–933.

(51) Ruiz, N.; Ferreira, S. S.; Padro, D.; Reichardt, N.-C.; Martín-Lomas, M. *Carbohydrate Research* **2011**, *346*, 1581–1591.

(52) Rising, T. W.; Heidecke, C. D.; Fairbanks, A. J. *Synlett* **2007**, *2007*, 1421–1425.

(53) Ruiz, N.; Ferreira, S. S.; Padro, D.; Reichardt, N.-C.; Martin-Lomas, M. *Carbohydrate research* **2011**, *346*, 1581–1591.

(54) Ley, S. V.; Baeschlin, D. K.; Dixon, D. J.; Foster, A. C.; Ince, S. J.; Priepke, H. W.; Reynolds, D. J.; others. *Chemical Reviews-Columbus* **2001**, *101*, 53–80.

(55) Hense, A.; Ley, S. V.; Osborn, H. M. I.; Owen, D. R.; Poisson, J. F.; Warriner, S. L.; Wesson, K. E. *Journal of the Chemical Society, Perkin Transactions 1* **1997**, 2023–2032.

(56) Shie, C.-R.; Tzeng, Z.-H.; Wang, C.-C.; Hung, S.-C. *Journal of the Chinese Chemical Society* **2009**, *56*, 510.

(57) Jiang, L.; Chan, T.-H. *Tetrahedron letters* **1998**, *39*, 355–358.

(58) Peña, I.; Mata, S.; Martin, A.; Cabezas, C.; Daly, A. M.; Alonso, J. L. *Physical* Chemistry *Chemical Physics* **2013**, *15*, 18243–18248.

(59) Lichtenthaler, F. W.; Lindner, H. J. *Carbohydrate Research* **1990**, *200*, 91–99.

(60) Bubb, W. A. *Concepts in Magnetic Resonance Part A* **2003**, *19*, 1–19.

(61) Bock, K.; Lundt, I.; Pedersen, C. *Tetrahedron Letters* **1973**, *14*, 1037–1040.

(62) Durette, P.; Horton, D. *Carbohydrate Research* **1971**, *18*, 403–418.

(63) Crich, D.; Dai, Z. *Tetrahedron* **1999**, *55*, 1569–1580.

(64) Van der Ven, J. G.; Wijkmans, J. C.; Kamerling, J. P.; Vliegenthart, J. F. *Carbohydrate research* **1994**, *253*, 121–139.

(65) Gloster, T. M. *Current opinion in structural biology* **2014**, *28*, 131–141.

(66) Rini, J.; Esko, J.; Varki, A. *Glycosyltransferases and glycan-processing enzymes.* In *Essentials of glycobiology*, 2nd ed.; Cold Spring Harber Laboratory Press: New York, 2009.

(67) Albesa-Jové, D.; Giganti, D.; Jackson, M.; Alzari, P. M.; Guerin, M. E. *Glycobiology* **2014**, *24*, 108–124.

(68) Lizak, C.; Gerber, S.; Numao, S.; Aebi, M.; Locher, K. P. *Nature* **2011**, *474*, 350–355.

(69) Albesa-Jové, D.; Mendoza, F.; Rodrigo-Unzueta, A.; Gomollón-Bel, F.; Cifuente, J. O.; Urresti, S.; Comino, N.; Gómez, H.; Romero-Garcia, J.; Lluch, J. M.; others. *Angewandte Chemie* **2015**, *127*, 10036–10040.

(70) Fukuda, M.; Bierhuizen, M. F.; Nakayama, J. *Glycobiology* **1996**, *6*, 683–689.

(71) Fischer, B.; Sumner, I.; Goodenough, P. *Biotechnology and bioengineering* **1993**, *41*, 3–13.

(72) Shin, H.-C. *Biotechnology and Bioprocess Engineering* **2001**, *6*, 237–243.

(73) De Bernardez Clark, E. *Current Opinion in Biotechnology* **1998**, *9*, 157–163.

(74) Vallejo, L. F.; Rinas, U. *Microbial Cell Factories* **2004**, *3*: 11.

(75) Lilie, H.; Schwarz, E.; Rudolph, R. *Current opinion in biotechnology* **1998**, *9*, 497–501.

(76) Phan, J.; Yamout, N.; Schmidberger, J.; Bottomley, S. P.; Buckle, A. M. *Refolding your protein with a little help from REFOLD.* In *Protein Folding, Misfolding, and Disease*; Springer, 2011; pp. 45–57.

(77) Singh, S. M.; Panda, A. K. *Journal of bioscience and bioengineering* **2005**, *99*, 303–310.

(78) Faveeuw, C.; Mallevaey, T.; Paschinger, K.; Wilson, I. B. H.; Fontaine, J.; Mollicone, R.; Oriol, R.; Altmann, F.; Lerouge, P.; Capron, M.; others. *European journal of immunology* **2003**, *33*, 1271–1281.

(79) Both, P.; Sobczak, L.; Breton, C.; Hann, S.; Nöbauer, K.; Paschinger, K.; Kozmon, S.; Mucha, J.; Wilson, I. B. *Glycobiology* **2011**, *21*, 1401–1415.

(80) Paschinger, K.; Staudacher, E.; Stemmer, U.; Fabini, G.; Wilson, I. B. *Glycobiology* **2005**, *15*, 463–474.

(81) Oriol, R.; Mollicone, R.; Cailleau, A.; Balanzino, L.; Breton, C. *Glycobiology* **1999**, *9*, 323–334.

(82) Kawar, Z. S.; Van Die, I.; Cummings, R. D. *Journal of Biological Chemistry* **2002**, *277*, 34924–34932.

(83) Wilson, I. B.; Rendi'c, D.; Freilinger, A.; Dumi'c, J.; Altmann, F.; Mucha, J.; Müller, S.; Hauser, M.-T. *Biochimica et Biophysica Acta (BBA)-General Subjects* **2001**, *1527*, 88–96.

(84) Yan, S.; Serna, S.; Reichardt, N.-C.; Paschinger, K.; Wilson, I. B. *Journal of Biological Chemistry* **2013**, *288*, 21015–21028.

(85) Bencúrová, M.; Rendi'c, D.; Fabini, G.; Kopecky, E.-M.; Altmann, F.; Wilson, I. B. *Biochimie* **2003**, *85*, 413–422.

(86) Boeggeman, E. E.; Ramakrishnan, B.; Qasba, P. K. *Protein expression and purification* **2003**, *30*, 219–229.

(87) Ramakrishnan, B.; Qasba, P. K. *Journal of Biological Chemistry* **2002**, *277*, 20833–20839.

(88) Serna, S.; Yan, S.; Martin-Lomas, M.; Wilson, I. B.; Reichardt, N.-C. *Journal of the American Chemical Society* **2011**, *133*, 16495–16502.

(89) Brzezicka, K.; Echeverria, B.; Serna, S.; van Diepen, A.; Hokke, C. H.; Reichardt, N.-C. *ACS Chemical Biology* **2015**, *10*, 1290–1302.

(90) Nguyen, K.; van Die, I.; Grundahl, K. M.; Kawar, Z. S.; Cummings, R. D. *Glycobiology* **2007**, *17*, 586–599.

CHAPTER 2

GLYCAN MICROARRAYS

2.1 Introduction

2.1.1 Miniaturization of assay-glycan microarrays

Cell surface glycans and glycan binding proteins (GBPs) have multiple functions and play important roles in numerous biological events such as cell proliferation, cell-cell recognition and the immune response.[1] Pathogenic GBPs mediate attachment to glycans in host cells and facilitate its subsequent invasion[2,3] whereas the host GBPs presented in immune cells promote pathogen recognition, immune activation and clearance of pathogens.

In order to 'decode' the information carried by each glycan and to study their possible role in the organism, the interactions with GBPs have been extensively studied over past last years. Relatively weak affinities of proteins to a single carbohydrate ligand can be enforced by the multivalent interactions between multiple carbohydrate recognizing domains of protein and multiple glycans [4] or/and adjustment of their spatial orientations. To this end, the glycan multivalent presentation may be achieved by ligand clustering on proteins, dendrimers, nanoparticles or array surfaces. Taking into account the huge diversity of carbohydrate structures as well as the enormous complexity of the protein-carbohydrate interactions, the development of the miniaturized, high throughput analytical screening methods such as microarrays emerged in recent years.[5,6,7]

The concept of the miniaturized assay in the form of *microspots*, was successfully applied for the first time around 15 years ago to study gene expression and to detect nucleic acids in the form of DNA chips (Figure 1).[8] The simplicity of the DNA microarrays development was based on the ability to synthesize DNA chains in a controlled and fast manner. This technology was later extended to protein and glycan microarrays (Figure 1),[9] which development raised together with the improvement of their ligands synthesis techniques. As mentioned in the Chapter 1, nowadays, the glycan libraries can be effectively prepared by chemoenzymatic strategies whereas some of the carbohydrates can be also purified

from natural sources.[10,11]

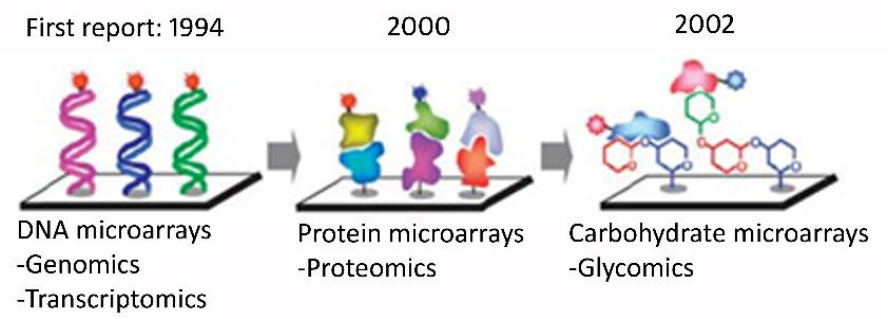


Figure 1. Microarray-based technologies for applications to genomics, transcriptomics, proteomics and glycomics (Shin *et al. Chem.Commun.* 2008[5])

The miniaturization of glycan-binding assays into a microarray format permits to reduce amount of both, carbohydrates and analyte employed in the screening. It also allows for the simultaneous analysis of hundreds ligands and multiple glycan binding proteins in a single assay. Arrays are created by immobilization of sugars on the solid support. Individual glycans are spotted in extremely small amount of femtomoles to create a microspot of an average diameter of 100-200 microns. Approximately 100 µL of protein solution (concentration usually vary between 1-10 µg·mL) can be further apply to analyse interactions with over 100 glycans at this same time. Binding of molecules to glycans is usually visualized by fluorescence of directly labelled protein or by fluorescently tagged antibody directed against the analyte.

In this Chapter, we will describe the design and preparation of glycan microarrays which will contain xylosylated N-glycans synthesized in Chapter 1 and show their utility towards screening of glycan binding proteins, such as plant lectins and C-type lectin receptors from mammals.

2.1.2 Surface chemistry and printing of glycan microarrays

The immobilization of glycans on the microarray surface can be performed either by non-covalent interactions or covalent coupling.[7,9] For example, the non-covalent approach utilizes nitrocellulose as a support and the ligands are immobilized through hydrophobic forces. Wang *et al*. first demonstrated how large molecules including glycoproteins and polisaccharides can be directly printed on nitrocellulose-coated slide without any prior derivatization.[12] Later on, Feizi and colleagues established a new system in which smaller molecules, oligosaccharides, were linked to lipids to form neoglycolipids (NGLs) which could be further immobilized on nitrocellulose-coated glass slides.[7,13] The other example of non-covalent attachment methods utilizes functionalized aluminum oxide or indium tin oxide (ITO) slides which surface is coated with a hydrophobic layer. Taking advantage from the transparency and conductivity of these slides, both platforms have been mostly exploited to monitor enzymes activity and specificity by on-chip mass spectrometry combined with fluorescence read-out.[14] Alternatively, glycans can be functionalized with oligonucleotide tags and immobilized on a surface by hybridization to a printed complementary DNA sequence. Likewise, biotinylated glycans can be attached to streptavidin covered glass slides via non-covalent but very strong interactions.[15]

The covalent immobilization of compounds on the array requires use of slides with functionalized surface (Figure 2). Free reducing glycans (Figure 2, A) can be directly printed on hydrazide or aminooxy-derivatized glass slides based on the hydrazone oxime formation, however, the resulting glycosidic linkage may be either cyclic or acyclic with beta or alfa configuration (Figure 2, A2-3).[16] Additionally, this conjugation technique requires relatively high immobilization concentrations (mM range) for the efficient construction of microarrays.[17]

Beside the aldehyde function of reducing sugars, most glycans lack selective, reactive functional group and therefore require chemical modifications to be subsequently immobilized on the surface-activated glass slides. For this purposes, a large variety of

coupling chemistry, including for example the thiol-maleimide, Diels-Alder reaction on gold surface or click chemistry between an alkyne functionalized surface with azide modified glycan have been developed and nowadays, many commercial slides are available.[18] However, the most frequently used microarray slides are *N*-hydroxysuccinimide (NHS)-ester and/or epoxy- activated surfaces, reactive for free amino groups (Figure 2, B).[6,19]

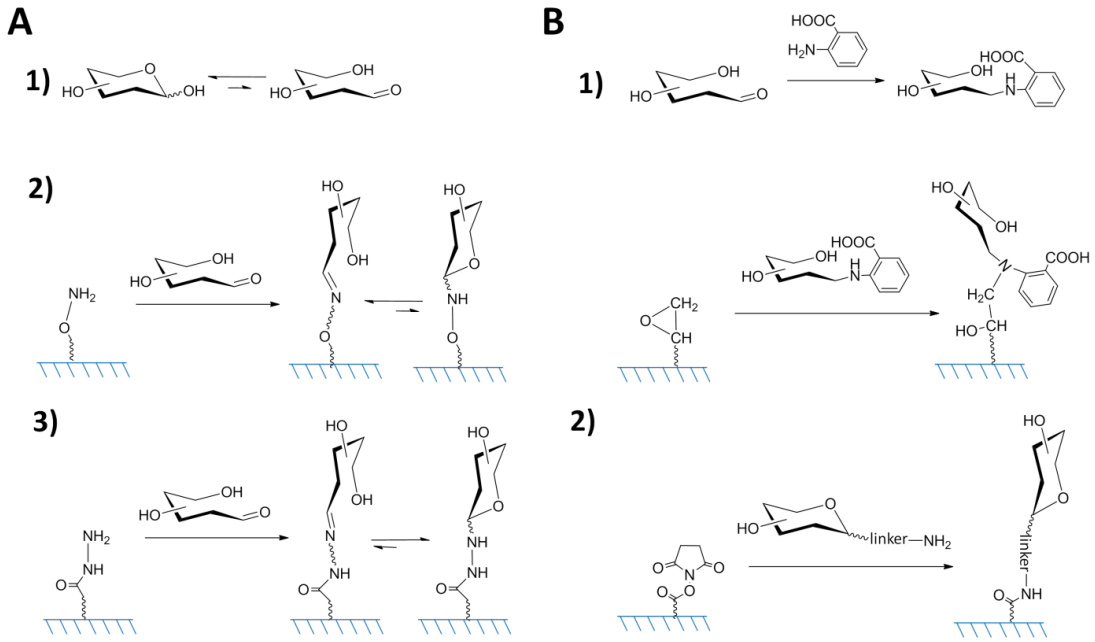


Figure 2. Covalent immobilization of sugars into the microarray surface; **A.** direct immobilization of free reducing carbohydrates using: **A2.** aminooxy activated glass slides and **A3.** hydrazide activated glass slides; **B.** immobilization of carbohydrates modified at their reducing end with: **B1.** 2AA tag and its further immobilization of the epoxy glass slide and **B2.** aminolinker and its immobilization on the NHS-activated slide.

NHS-slides are reacting very efficiently with primary amines and the immobilization leads to the formation of a very stable amide bond (Figure 2, B2). These arrays are widely used for chemically amine-modified glycans as well as for glycoproteins or proteins. Epoxy slides bind to secondary and aromatic amines with the formation of tertiary or secondary amines (Figure 2, B1) and are most widely used for the preparation of so called natural glycan

microarrays. Natural arrays contain ligands isolated from the natural sources usually tagged with either 2-aminobenzamide (2AB) or 2-aminobenzoic acid (2AA) which enable their purification by the HPLC.[20] So isolated 2AB or 2AA-tagged glycans are subsequently spotted on the epoxy activated glass slides to create the natural glycan microarray.

The microarray manufacturing always involves transfer of the glycans from a microplate to the solid surface, such as glass slide. In general, once the chemistry of immobilization is chosen, ligands can be either printed in a contact or non-contact manner.[9] During contact printing a set of steel pins (up to 64) is dipped into the ligand solutions and a small quantity of the sample is transferred to the surface by direct contact with it. Amount of solution delivered into the slide depends strongly on the time that pin is in contact with its surface. Non-contact printing is more elaborate and precise. It is accomplished with a Piezo-electric printer ensuring consistent drop size using controlled electric signals and thus resulting in more homogenous morphology and size of spots compared to contact-printing. The major disadvantage of the Piezo-electronic printers, however, is a limited number of tips (4 or 8) which results in longer printing times compared to contact printing methodology. The piezoelectric effect was discovered in 1880 by two brothers, Jacques and Pierre Currie, when they were in their early twenties. They describe it as electricity resulting from pressure. Later on this phenomenon was used for a construction of Piezo-electric printers containing a glass capillary (nozzle) surrounded by a piezoelectric quartz crystal, which flexes when an electric current is applied. The deformation provides pressure on the nozzle filled with sample and forces the drop out of it. The remaining ligand solution can be return back to the source plate once its printing is accomplished.

2.1.3 Detection strategies

The most commonly used strategies for the detection of binding events of proteins or other biological samples to the glycan array are fluorescence-based methods. They are highly

sensitive, allowing for the detection of femtomoles of ligands and very convenient to read-out with a high resolution microarray scanner operating with a resolution of 2-10 μm.[5] In this procedure, carbohydrate arrays are incubated with fluorescently labeled protein (Figure 3, A) or secondary reagents (usually antibody) that specifically recognize primary protein (Figure 3, B). The fluorescence intensity of individual spot is proportional to the amount of bound protein and is determined using a microarray scanner and quantified with specially designed softwares (for example ProScanArray Express from Perkin Elmer). Quantitation methods are used to construct a pixel by pixel map that indicates the property of each pixel in the image corresponding to spots and their background. The adaptive circle quantification method fits all spots in the image with circles which diameters are estimated separately for each spot on the array. So obtained data represents the intensity of each spot with the substracted background fluorescence for that spot. The histograms can be subsequently represented as a mean or median fluorescence intensity of spots with local background subtraction.

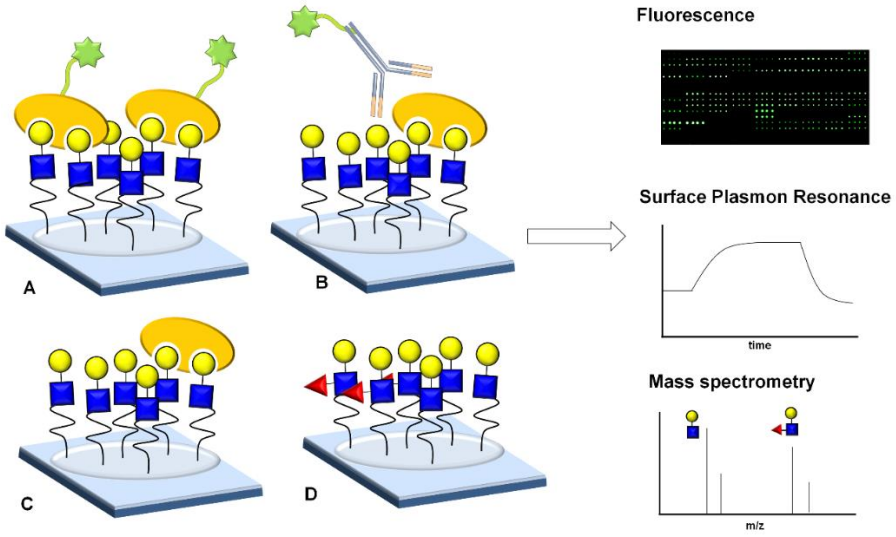


Figure 3. Detection strategies used for the carbohydrate-protein interactions quantification; **A.** and **B.** fluorescence-based methods applying directly labelled proteins or secondary antibody directed against analysed protein; **C.** SPR-based methods; **D.** Mass spectrometry usually used for the detection of enzymatic reactions performed on the array.

There are also alternative label-free detection methods for the array-based screenings, including surface plasmon resonance (SPR, Figure 3, C) and mass spectrometry (MS, Figure 3D). In the SPR imagining technology the real-time data on the affinity, specificity and kinetics of protein interactions can be measured. The approach is based on an optical detection system that monitors real-time changes in protein mass on the sensor surface. Binding events result in changes in protein mass that, in turn, alter surface plasmon resonance, an electro-magnetic phenomenon that dampens the intensity of light reflected off of the surface of the sensor chip at a specific angle. The change of resonance angle resulting from analyte-ligand interaction is measured in resonance units (RU) and plotted in sensorgrams as a function of time.[21] Although in this method the protein does not have to be labelled (Figure 3,C) it requires a gold surface, which is suitable to control the surface plasmon phenomena and the modification of ligands with SH-linker to be afterwards directly conjugated onto gold chip *via* Au-S linkage. Mass spectrometry, on the other hand is usually applied to follow the modifications of sugars printed on the array but also can be used for the identification of protein bound to the certain ligand.[7,14] For example, MALDI-TOF MS is perfectly suitable to assign the specificity of carbohydrate processing enzymes, like glycosyltransferases or glycosidases directly on the array.[14,22,23] The difference in mass spectrum acquired after the reactions and compared to the starting material indicates the outcome of reaction (Figure 3, D).

2.2 Design and construction of glycan microarrays

All glycans prepared in our laboratory are modified with a C5-amino linker for immobilization. We have chosen commercially available (NHS)-activated glass slides as the most convenient microarray support for our purposes. Next, regarding number of all glycans available in our lab we established the right array design and finally, we spotted carbohydrates with Piezo-electric robotic non-contact printer (Figure 4).

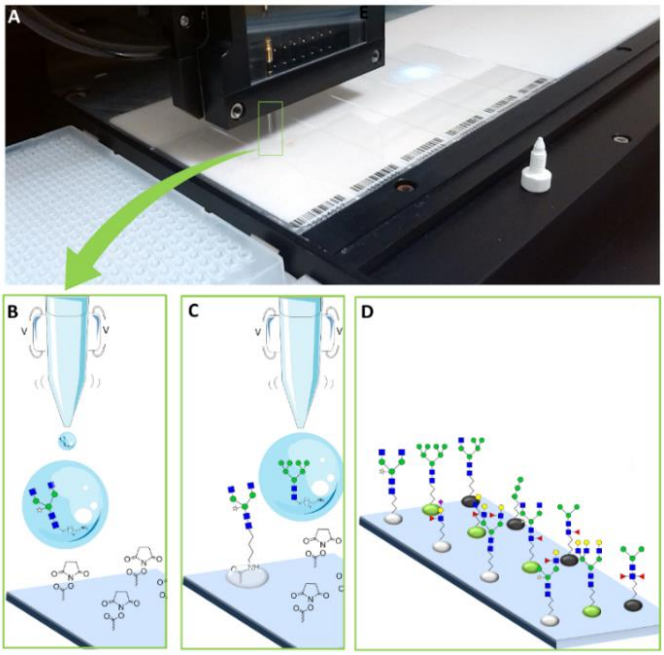


Figure 4. Preparation of the N-glycans microarrays; **A.** Piezo-electric printing of carbohydrates microarray; **B-C.** Schematic representation of the immobilization of glycans onto the NHS-activated glass slide; **D.** Final glycan array.

We have built a library of over hundred glycan structures, including 39 xylosylated N-glycans (**G1**-**G39**) prepared in Chapter 1 of this thesis[24], combined with diverse carbohydrates synthesized earlier in our laboratory and commercially available O-glycans (Figure 5).

The glycan library of 126 structures was printed in a 7x1 microarray format (Figure 6) which

is consisted of seven identical subarrays (Figure 6, A-B) representing copies of same N-glycan library.

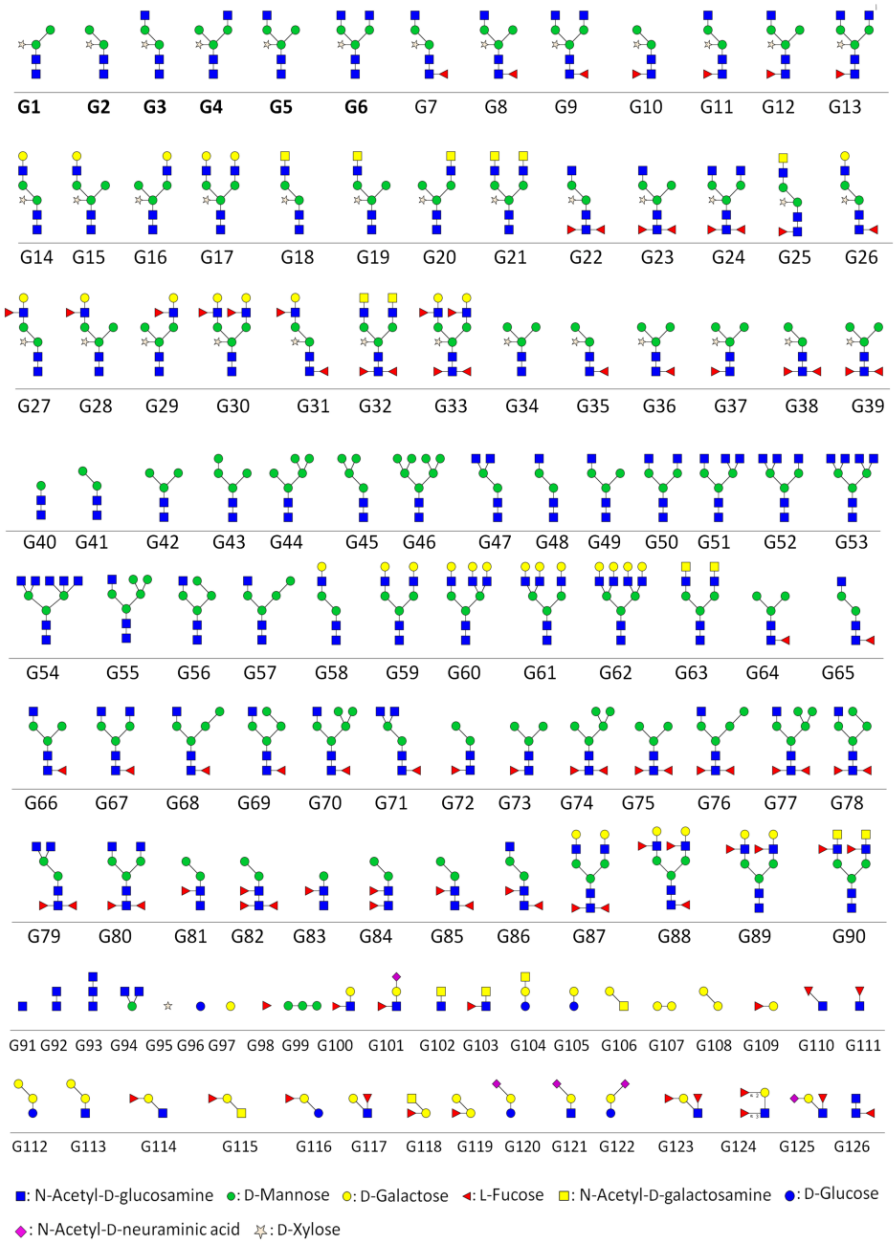


Figure 5. Total number of glycans printed on the glycan array; **G1**-**G39** xylosylated N-glycans prepared in Chapter 1; **G40**-**G103** structures prepared in our laboratory; **G104**-**G126** commercial glycans.

Each subarray is designed to fit maximum of 648 spots with 36 spots per row and 18 spots per column (Figure 6, C). We have printed all glycans in 4 replicates, what resulted in 9 glycans per row of each subarray (Figure 6, C) and a total of 3528 spots on every slide. The average spot diameter was 195 µm, with a pitch (distance between spots) of 400 µm in both, x and y axes (Figure 6, D).

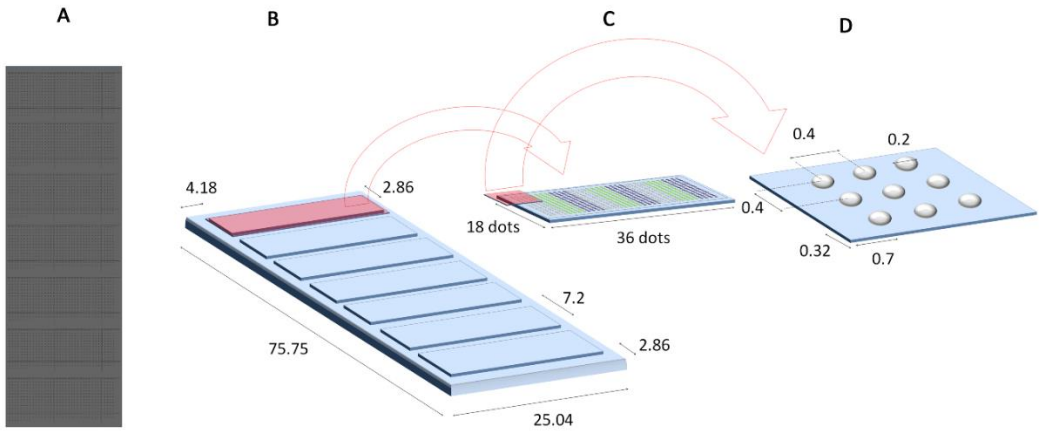


Figure 6. Design of glycan microarray used during our study; **A.** Real picture of the array taken directly after glycans printing; **B.** NHS-activated glass slide consistent of seven identical subarrays; **C.** subarray with 18x36 dots printed on its surface; **D.** distances between the spots and their average diameter.

Glycan spotting solutions were prepared from 1mM stock solutions by dilution with sodium phosphate buffer (pH 8.4, 300 mM, 0.005% Tween 20) to a final spotting concentration of 50 µM. The concentration of printing glycan solutions had been established in our laboratory and was supported by the data found in the literature.[6] This concentration result in a spot saturation. Glycans were immobilized by the robotic printing by spotting of 5 drops of glycan solution per every spot on the array. Taking into account that each drop has an average volume of 250 pL (1.25 nL/spot) the final amount of each glycan used for the printing of one spot is about 0.62 fmol/spot. The microarray strategy is an extremely

powerful tool, which allow for a miniaturization of the binding assays and use of very small quantities of glycans and simultaneous screening of several samples on one slide. To demonstrate the big advantage of the microspot-strategy over usual techniques (like ELISA), a calculation was performed. With our glycan microarray, we use 1.25 nL/spot of each glycan. Taking into account that each sugar is printed in 4 replicates on each subarray and that there are 7 copies of subarrays on the slide, it gives total volume of 35 nL per array, equal less than 1 µL per 25 arrays. Usually, amounts of synthesized compound vary between 0.5-1 mg, which allows for its good characterization, by both NMR and high resolution mass spectrometry. In order to prepare 50 µM printing solutions of each glycan (average mass of 1200 g/mol) starting with 0.5 mg of compound, we would potentially prepare 7 mL of printing solution. As 1 µL would be used to print 25 microarrays, above volume would allow us for a screening of roughly 175 000 arrays.

This theoretical calculation shows, how many binding studies can be done, regarding amount of synthesized glycan. Additionally, screening of different carbohydrates binding proteins can be done simultaneously on one slide reducing the amount of the required sample and of time needed for the experiment. Using multiple fluorescent probes, we can also measure different interaction patterns in one subarray format, what can be done for example for an antibody isotype class differentiation.

2.3 Screening of carbohydrate binding proteins using glycan arrays

In this section of the Chapter the microarray-based studies of carbohydrate-protein interactions will be described.

2.3.1 Carbohydrate binding proteins. Introduction

Lectins are a family of proteins capable of binding to carbohydrate moieties. They form a heterogeneous family of mono- and multivalent proteins which are widely distributed in different species, including microorganisms, plants and animals.[25] Lectins have diverse functions and play a crucial role in both physiological and pathological processes, like fertilization, pathogen recognition and endocytosis or immune defence. Animal lectins, called galectins for example, modulate the turnover of glycoproteins[26] whereas selectins promote the migration of leukocytes to the site of injury in the inflammation process.[27] Through the interactions with leukocytes surface glycoproteins, selectins expressed by the inflamed endothelium of the blood vessel wall, ensure the contact between the endothelium and circulating in the blood stream leukocytes.

Nowadays, lectins with well characterized specificity are commercially available and have found many applications in biochemistry, like in glycan profiling or in glycoprotein purification and characterization.[28] The two groups of lectins described and analyzed in this Chapter are plant and fungal lectins and animal C-type lectins receptors (CLRs).

The first plant lectin was discovered by Stillmark in 1888, he described the hemagglutinin activity of the castor beans extracts and this activity was assigned to ricin. Since then, several hundreds of proteins have been isolated from plants and fungi and characterised regarding their carbohydrate specificity. Nowadays, the term-lectin, is used for proteins which possess at least one non-catalytic domain which reversible bind to mono- and/or oligosaccharides. The physiological roles of plant lectins have been extensively studied over the years. It is believed, that they might have a dual functions: act as storage proteins and

have a role in the plant defence against various plant eating organisms.[29–31] This theory was supported by the observation, that some of plant lectins possess sugar specificity towards carbohydrates present outside the plant kingdom. They are extremely stable in unfavorable conditions like harsh pH, heat or insect/animal proteases and they are associated with the survival of plant tissues. Additionally, the high cytotoxicity of some of these lectins (ricin) seems to assure good protection against predators like animals (herbivores), insects or some fungi.

Other group of carbohydrate binding proteins are C-type lectins receptors which are present in animal organisms. They create group of over 1000 proteins, in either soluble or membrane bound form (Figure 7). C-type lectins are characterized by presence of at least one carbohydrate recognizing domain (CRD) which is a conserved structural motif containing two protein loops stabilized by two disulfide bridges.[25,32] Because the second loop is more flexible than the first one, it is usually the place where ligands are bound.

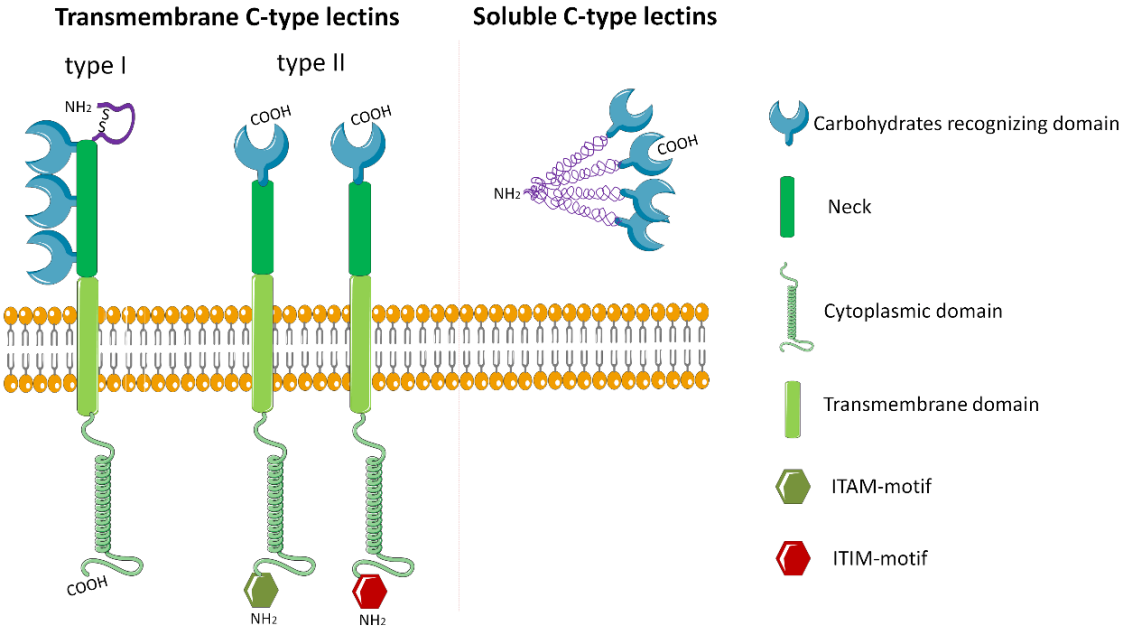


Figure 7. C-type lectins receptors belonging to transmembrane (type I and type II) and soluble subfamilies of proteins.

Furthermore, the carbohydrate specificity of CLRs depends on the conserved motifs present in the CRD, either EPN that could bind to mannose, N-acetylglucosamine, fucose and glucose or QPD that could bind to galactose and N-acetylgalactosamine moieties. C-type lectins are usually associated with calcium mediated binding of exo- and endogenous carbohydrates, however, less often they can also bind to sugars in calcium independent manner or recognize other ligands, including lipids.

Membrane bound C-type lectin receptors can be further divided into two groups, type I (e.g. Mannose receptor, DEC-205) or type II transmembrane CLRs (e.g. DC-SIGN, MGL, DCIR).[33] Proteins of the first group have their N-terminus tail pointing outside the cytoplasm, they contain an amino-terminal cystein-rich region (S-S) and between eight to ten CRDs. Type II C-type lectins on the other hand are characterized by the single CRD facing outside the cell with the N-terminus tail placed inside the cell.

C-type lectins play important function in the immune system by acting as both an adhesion protein and pattern recognizing receptor (PRRs).[25,33–35] The adhesion protein ensures the cell-cell contact (selectins) while the PRR are involved in pathogen recognition, uptake and further presentation (mannose receptor). Pattern recognizing receptors bind pathogen-associated molecular patterns (PAMPs) and are able to further modulate the immune response. Depending on the intracellular internalization motifs present in the cytoplasmic domains of C-type lectins, ITAM-immunoreceptor tyrosine-based activation motif or ITIM-immunoreceptor tyrosine-based inhibition motif, C-type lectins can have potential immune-suppression or immune-activation functions.[36] Because of these properties, C-type lectins have become important target for carbohydrate based vaccine candidates and the identification of their natural ligands emerged in recent years. One of a good example whene glycan arrays were utilized for the identification of the exact glycan binding motifs of C-type lectins has been given by Skerra and coworkers.[37] They have characterized two homologues C-type lectins, DC-SIGN and Langerin, express on the surface of different dendritic cells. Both proteins, were found to play a role in HIV-1 infection, with DC-SIGN facilitating virus survival and transmission, whereas Langerin promoting HIV-1 uptake and

subsequent degradation. Even though, both C-type lectins share a common tripeptide sequence (EPN) and thus bind towards mannose and fucose containing structures, microarray analysis revealed more specific binding epitope of each protein. DC-SIGN was found to bind internal mannose residues of high mannose structures and fucose-containing blood type antigens, while Langerin was identified to recognize more preferentially terminal mannose moieties, N-acetylglucosamine and some of the blood type structures. This information may aid for the future design of HIV-1 vaccines candidates that specifically affect the virus interactions with DC-SIGN receptor while not compromising the protective function of Langerin.

2.3.2 Characterisation of plant lectins

To demonstrate the functionality of printed glycan structures on the microarray, we incubated different subarrays in parallel with fluorescently tagged lectins (Table 1) including *Galanthus nivalis, Wisteria floribunda, Aleuria Aurantia*, *Bandeiraea* simplicifolia, *Canavalia ensiformis* and *Ricinus communis* agglutinin, expressing different glycan specificities resumed in Table 1.[38]

Entry	Lectin	Acronym-Label	Specificity
1	*Galanthus nivalis* agglutinin	GNA-555	High mannose-type N-glycans, preference to α-1,3 mannose
2	*Concanavalin A*	ConA-647	High mannose-type N-glycans
3	*Aleuria aurantia* lectin	AAL-555	Fucose
4	*Bandeiraea simplicifolia* lectin II	BSLII-647	Terminal non-reducing α- or β-linked N-acetyl-D-glucosamine
5	*Wisteria floribunda* lectin	WFL-555	N-acetyl-D-galactosamine
6	*Ricinus communis* agglutinin-*120*	RCA-555	β–galactose residues

Table 1. Lectins probed on the glycan arrays (column two) with the given acronym and fluorescent label (column three) as well as its previously described specificity[38] (column four).

Lectins were fluorescently labelled on primary amino groups of lysines and the N-terminal with commercial Alexa Fluor® 555 NHS ester (sulfonated xanthene) and Alexa Fluor® 647 NHS ester (sulfonated castasteron). Solutions of lectins in sodium carbonate buffer (pH=9.0) were incubated in the dark for 1 hour with the corresponding fluorescent dye NHS ester. Excess of dye was removed by dialysis or by ultracentrifugation. Subsequently lectins dilutions in PBS containing 5 mM $CaCl_2$, 5 mM $MgCl_2$ and 0.01% Tween-20 were incubated on glycan microarrays in the dark at room temperature for 1 hour. The slides were washed with PBS containing Tween-20 followed by water and scanned in a microarray scanner. The fluorescence bound to the surface was quantified and represented as histograms.

Figures 8-10 show the histograms corresponding to incubation of lectins conducted on the glycan arrays. As a result, we could observe the carbohydrate specificity previously described for the lectins employed but also identify new features regarding for example influence of xylose modification on the recognition of glycans by the mannose binding lectins. Even though the general glycan specificity of analysed plant lectins is already known, in some cases it is based on simple inhibition assays performed with mono- or disaccharides and does not provide additional information about the specificity in context of larger, more complex glycans. Having on our arrays various, well-defined structures of plant, parasitic but also mammalian origins, the microarray-based screenings could not only confirm functionality of the array but also reveal more biological relevant ligands of analysed lectins and allow us to study the effect of differentiated epitope presentation on glycan-protein interactions.

The glycan array probed with *Ricinus communis* agglutinin-120 (RCA-555) and with *Wisteria floribunda* lectin (WFL-555) (Figure 8), showed binding towards terminal N-acetyl lactosamine and towards terminated N-glycans with terminal N-acetyl-D-galactosamine respectively (Figure 8). The analysis of fluorescence further revealed, that the multivalent presentation of galactose moiety on the antennae of N-glycans, have an influence on the binding affinities of tetrameric RCA lectin towards mono-, di-, tri- and tetrantenary compounds. The latter two N-glycans were giving the highest fluorescence intensity values

indicating strengthen of the protein binding towards N-glycans presenting multiple galactose moieties. In this context, the multiantennary N-glycans can be seen as miniature analogues of dendrimers that present multiple copies of a terminal motif in a very define space.[39,40]

The WFL, on the other hand was found to bind to the LDN glycan element, in a context of larger N-glycans but also to a simple disaccharide directly conjugated to the array surface. No significant difference of the LDN presentation on the affinity to the lectins was observed.

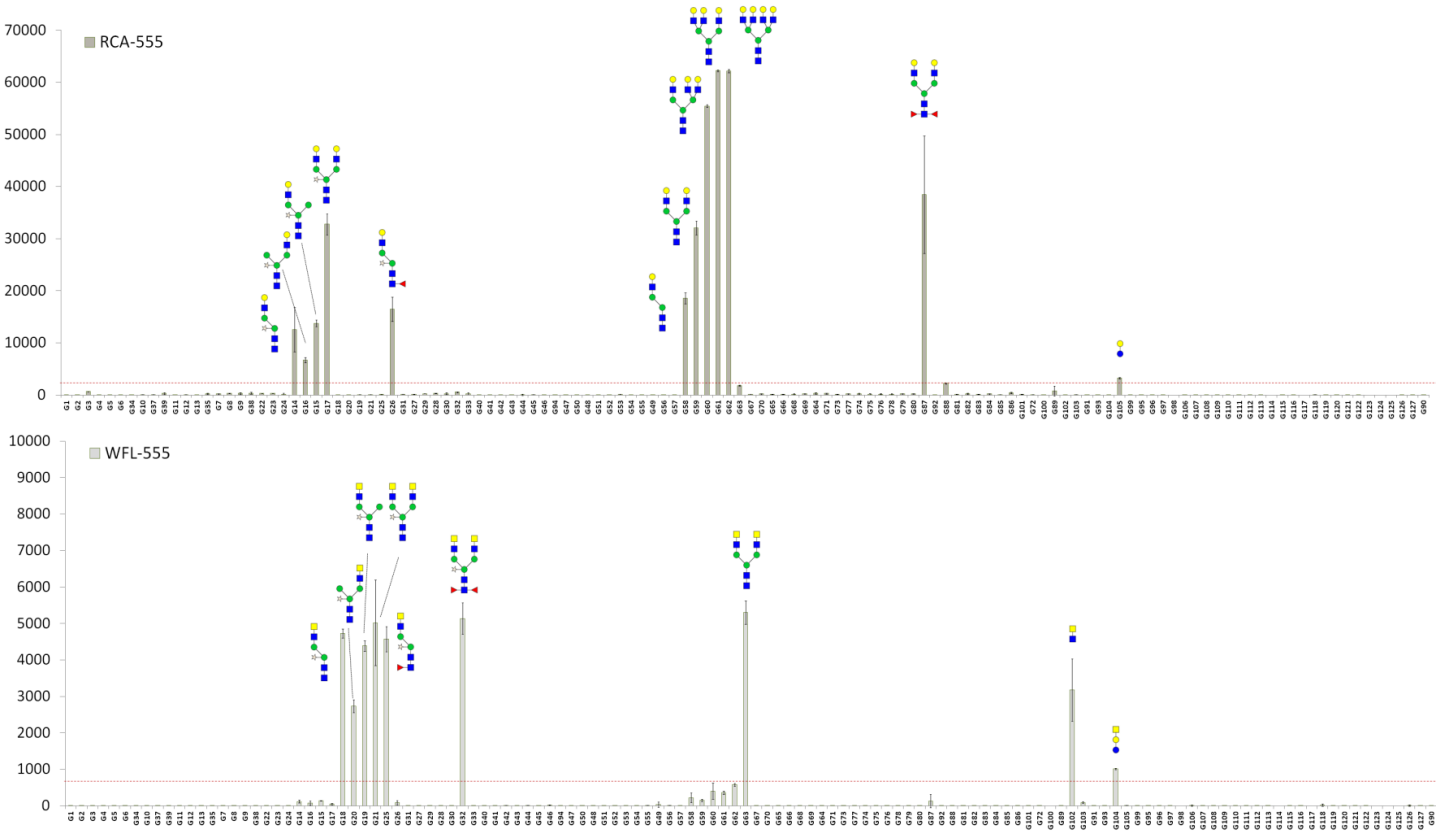


Figure 8. Fluorescence intensities for N-acetyl lactosamine recognizing plant lectin **RCA-555** (upper) and β-linked galactose binding **WFL-555** (lower) on glycan array.

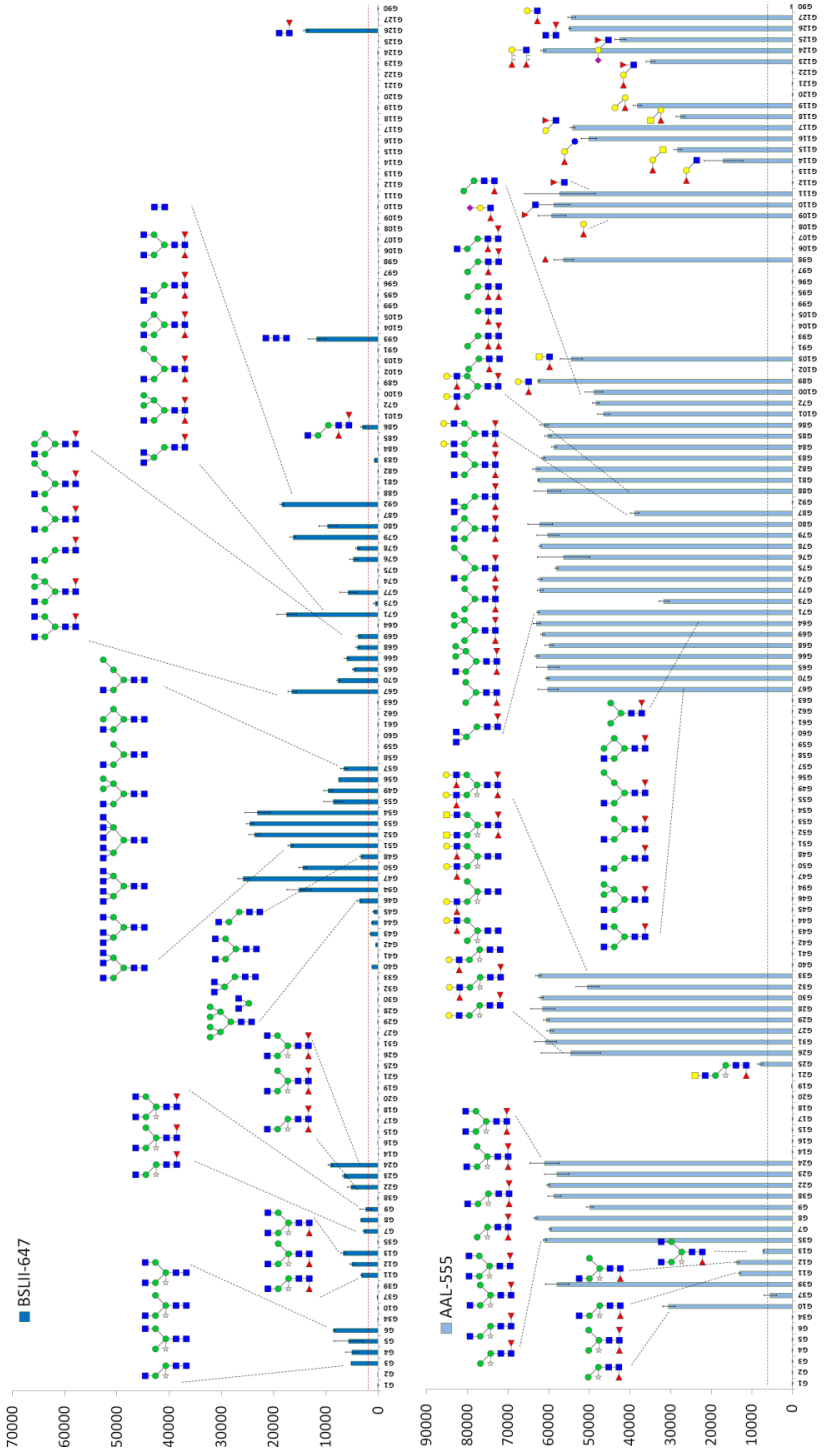


Figure 9. Glycan array probed with **BSLII-647** (upper) recognizing terminal, non-reducing α- or β-linked GlcNAc and **AAL-555** (lower) fucose specific lectin.

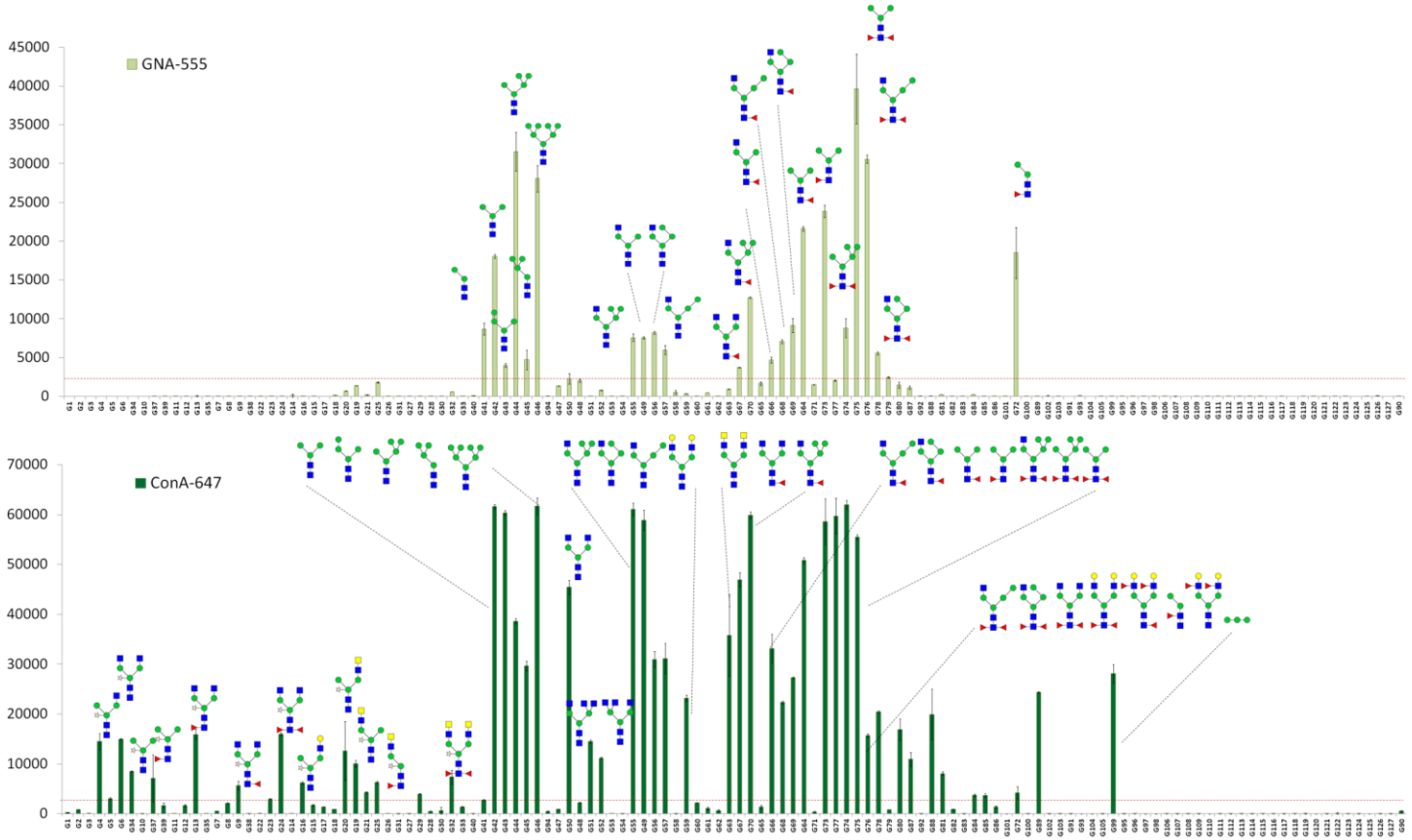


Figure 10. Fluorescence intensities for binding of mannose recognizing plant lectins: **GNA-555** (upper) and **ConA-647** (lower) on glycan array.

Further screening of two other lectins, BSLII-647 and AAL-555 revealed that specificity of both proteins follows previously described trends, with *Bandeiraea simplicifolia* recognizing the terminal N-acetyl-D-glucosamine moieties and *Aleuria aurantia* lectin L-fucose decorated glycans (Figure 9). A more detailed analysis of the carbohydrate binding profile of BSL-II showed the strongest affinity towards complex tri- and tetraantennary N-glycans with terminal GlcNAc moieties and slightly lower interactions with the biantennary compounds. Hybrid N-glycans, with only one GlcNAc substituted antennae were still bound, however, to the lower extent. It is also worth mentioning about the strong interactions of BSL-II with chitobiose core, what can be easily correlated with the role of *Bandeiraea simplicifolia* lectin in nature. This plant protein defends against certain insects by binding to the chitin or the glycoproteins of the epithelial cells of the incest.[41]

Next, following the binding profile of AAL, we could observe specificity towards compounds containing fucose, regardless the position (α-1,3 or/and α-1,6) or presentation (core fucose of Lewis X-type of structures). Additionally, a simple fucose was also a good ligand.

Binding of RCA-555, WFL-555, BSL-II-647 and AAL-555 to glycans was not affected by the presence of core xylosylation. The incubation of glycan microarray with mannose recognizing lectins, such as *Concanavalin A* (ConA-647) and *Galanthus nivalis* agglutinin (GNA-555) followed the general specificity of both proteins (Figure 10). However, an interesting influence of the presence of core β-1,2-xylose on the interaction was observed. In the case of *Concanavalin A,* the presence of core β-1,2-xylose strongly reduced the binding when compared to the non-xylosylated counterparts (Figure 11, A,B). We could also observe that core fucosylation also reduced binding to ConA but to a lower extent (see ligand pairs **G66**/**G49** and **G42**/**G75**) and not in all cases (**G50**/**G67**). In the case of *Galanthus nivalis* agglutinin the presence of β-1,2-xylosylation completely abolished binding (Figure 11, A, C). ConA is known to have broad specificity and bind high mannose as well as complex-type N-glycans.[42] Microarray-assisted study showed, that it can recognize even in the context of a biantennary N-glycan with two bulky terminal LDN or Lewis X groups (Figure 11, A), however, the analyzed fluorescence intensities are approximately half of

those found for high mannose N-glycans.[9]

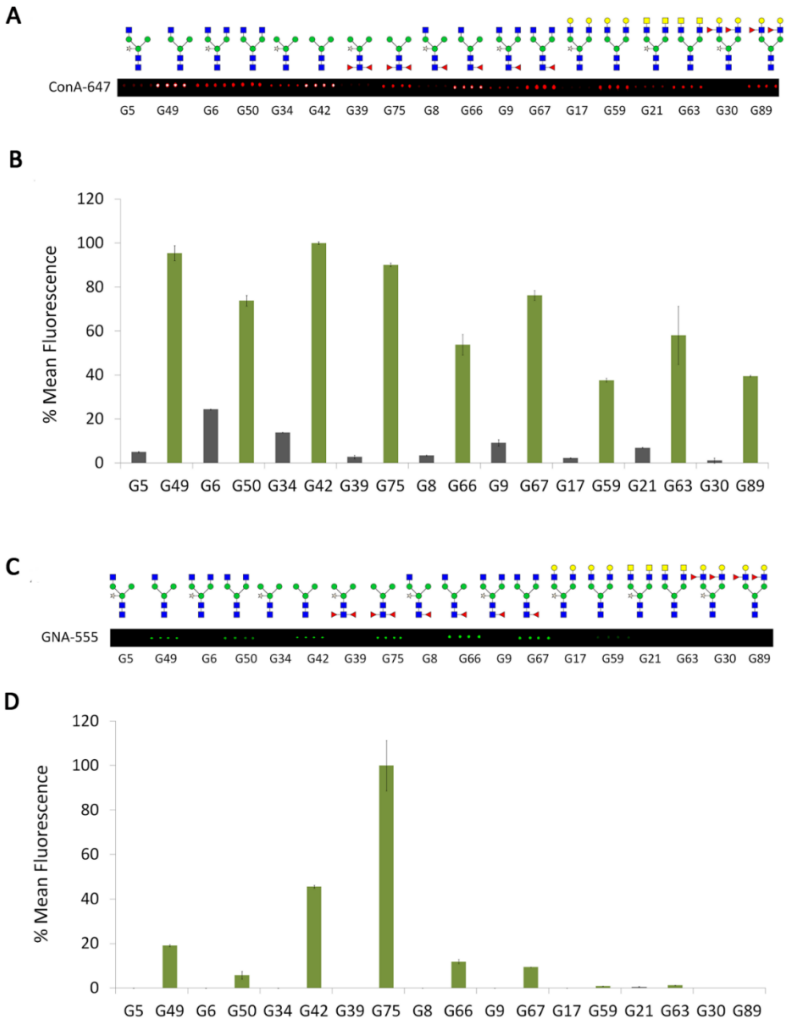


Figure 11. Glycan array incubated with mannose recognizing fluorescently labeled lectins ConA-647 and GNA-555; **A,C.** Selected ligands in the interaction with lectins. **B.** Decrease of binding for xylosylated (grey bars) and non-xylosylated structures (green bars) in the interaction with ConA-647; **D.** Lack of recognition of xylosylated structures compared to the corresponding non-xylosylated N-glycans (green bars) after incubation with fluorescently labeled GNA-555.

GNA on the other hand, binds to terminal α-1,3 and α-1,6 linked mannose residues and unlike ConA tolerates further extension of the β-mannose with GlcNAc but not with LN or

LDN residues. As these lectins could be involved in plant defense against predators[31,43], this could be presumably achieved by attaching to insect or vertebrate type glycans lacking the core xylose motif representative of plant glycans. The strongly reduced binding affinity of ConA and GNA towards core-xylosylated plant glycans could be explained as a tolerance mechanism to avoid interference with self proteins.

Overall, the incubation of printed glycan microarrays with plant lectins confirmed the high specificity of these proteins towards certain carbohydrates and showed the potential application for multiple analyte characterization. Moreover, taking advantage of the unique structures having on our array, we could further detailed the specificity of analyzed plant lectins in the context of larger N-glycans. It was demonstrated, that microarray-assisted binding studies with well characterized, chemo-enzymatically prepared ligands allow for a very precise insight into the protein specificity and the structural requirements of glycans for effective binding.

2.3.3 Screening of C-type lectin receptors

After the characterization of plant lectins, we screened a selection of human and murine C-type lectins receptors. Glycan binding patterns for seven recombinant murine Fc fusion CLRs were studied with the glycan array: DC immunoactivating receptor (DCAR), DC immunoreceptor (DCIR), Clec-2, Clec-9A, Macrophage inducible C-type lectin (Mincle), Macrophage-galactose-type lectin (MGL-1) and the murine DC-SIGN homologue SignR3. CLR-Fc fusion proteins were produced in a mammalian expression system in the laboratory of Prof. Bernd Lepenies. They were expressed as fusion proteins containing the extracellular part of CLR and the Fc part of human IgG1 antibody.[44] We additionally included in the screening three commercially available human Fc fusion proteins: Dectin-2, DC-SIGN and MGL-1. The previously described specificity and some general information about the function of these CLRs is listed in Table 2.

Lectin	Function/Characteristic	Ligands	Array binding study
Clec-2	-Platelet activation receptor; -ITAM motif[47,48]	Exo/endogenous ligands, rhodocytin (snake venom), podoplanin, sulfated sugars[45]	-No binding
Mincle	-Pathogen recognition/internalization[49,50]; -ITAM motif	Glycolipids (e.g. TDM= trehalose-6,6'-dimycolate); recognition of pathogenic fungi and mycobacterium	-No binding
Clec-9A	-Binds damaged or dead cells via exposed actin; -Activation receptor[48] (ITAM motif)	Unknown	-No binding
DCAR	-Unknown physiological role [51];-May be activating receptor (ITAM motif)	Unknown	-No binding
DCIR	-Internalization of HIV-1(gp-120) and HCV(E2)[52]; -Immune suppression[53] (ITIM motif)	Unknown	-No binding
mMGL-1	-Internalization/antigen presentation; -Suggested role in a tumor immunity	Highly specific for Lewis X and Lewis A; no binding towards Lewis B or Lewis Y[46]	- Lewis X and Lewis A
SignR3	-Pathogen recognition/presentation; -Homologue of human DC-SIGN[54- 56]	Similar carbohydrate specificity to DC-SIGN (fucose, high mannose glycans)	-No binding
Dectin-2	-Impairment of UV-induced tolerance[51]; -Response to allergens; -ITAM motif	α-mannans; *M. tuberculosis*; house dust mite; fungi	-No binding
DC-SIGN	-Pathogen recognition/presentation[57–60]; -DCs-T cells interactions	Fucose and high-mannose; ICAM-3; blood group B antigens; Le^b>Le^y>Le^a>Le^x, LDNF	-Confirmed known spec. -Xylosylated biantennary compounds (**G6**, **G9**, **G24**, **G17**, **G21**) do not bind
MGL-1	-Pattern recognition receptor; -Tumor immunity, -Recognition of helminthes[61]	α and β-linked terminal GalNAc moieties; Tn antigen (GalNAcα-Ser/Thr)	-GalNAc terminated N-glycans (LDN and LDNF)

Table 2. C-type lectins Fc fusion proteins screened on the glycan array.

We have tried two different detection modes for C-type lectins on the glycan array, the so-called sandwich assay and a single step incubation. In the first strategy, C-type lectins were diluted with TSMT binding buffer pH=7.5 (20 mM Tris·HCl, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 1% BSA and 0.005% Tween-20) to a final concentration of 10 µg/mL and incubated on the glycan array overnight at 4 °C. Unbound proteins were washed off and the arrays were subsequently incubated with anti-human IgG1 (Fc specific)-Cy3 antibody for 1h. The slides were washed and scanned in a microarray scanner. In the one step incubation, fusion C-type lectins were premixed with anti-human IgG1-Cy3 antibody prior overnight incubation at 4 °C on the glycan array. In this strategy, one washing step is avoided but more importantly complexing of the protein with antibody could result in its dimerization and thus strengthen avidity. Results of both screenings were similar, although slightly lower fluorescence intensities where observed applying the single step incubation. Therefore, results presented in this Chapter are obtained by the sandwich incubation of C-type lectins on the glycan array.

First, we determined carbohydrate recognition profiles of two closely related lectins: human macrophage galactose-type lectin MGL-1 and its murine homologue, mMGL-1 (Figure 12). Both proteins belong to family II transmembrane lectins and are expressed on the surface of immature macrophages and dendritic cells where are involved in pathogen and tumor antigens recognition.[46] Humans and murine MGL-1 carbohydrate recognizing domains predict binding towards galactose-related structures; however their exact carbohydrate specificity differs.[62] It has been described that human MGL-1 recognizes Tn-antigens (GalNAcα-Ser/Thr) which are expressed by adenocarcinoma cells, but they are also part of the parasite *Schistosoma mansoni* egg glycome.[61] On the other hand, murine MGL-1 binds to Lewis X related structures, found in a high abundance in tumor cells (for example colon carcinoma) and in helminths organisms. In our glycan microarray screening, the known specificities of both human and murine MGL-1 were confirmed. The murine lectin bound to Le^X decorated glycans, as well as tri- and tetra-antennary compounds with terminal galactose.[40] Taking advantage from the collection of multiantennary N-glycans on our

synthetic array, it was interesting to see the impact of a multivalent presentation of galactose on the binding strength to the lectin. Murine MGL-1 bound higher towards tetra-antennary compound compared to triantennary[40] and no binding for galactosylated biantennary N-glycan was detected. Human MGL-1 was giving binding towards GalNAc terminated structures, both in LDN and LDNF epitopes. Also simple di- and trisacharides were recognized by the lectin and we could not observe any effect of epitope presentation in the context of larger N-glycans on the binding strength to the MGL-1. There was also no significant influence of core xylose or fucose on the interactions with protein.

Next, we studied carbohydrate specificity of DC-SIGN to core-modified structures (Figure 13, 14)[24]. This C-type lectin signaling receptor expressed by human dendritic cells can bind and internalizes several antigens some of which may be mannose- or fucose-containing glycans.[60] DC-SIGN targeting is an efficient strategy of pathogens like human immunodeficiency virus (HIV), Hepatitis virus, Ebola virus or *Mycobacterium tuberculosis* to evade the immune system.[24,63,64] DC-SIGN is also one of the receptors for recognition and uptake of *Schistosoma mansoni* antigens via Le^X, LDNF and mannosylated and fucosylated N-glycan motifs.[60,65,66]

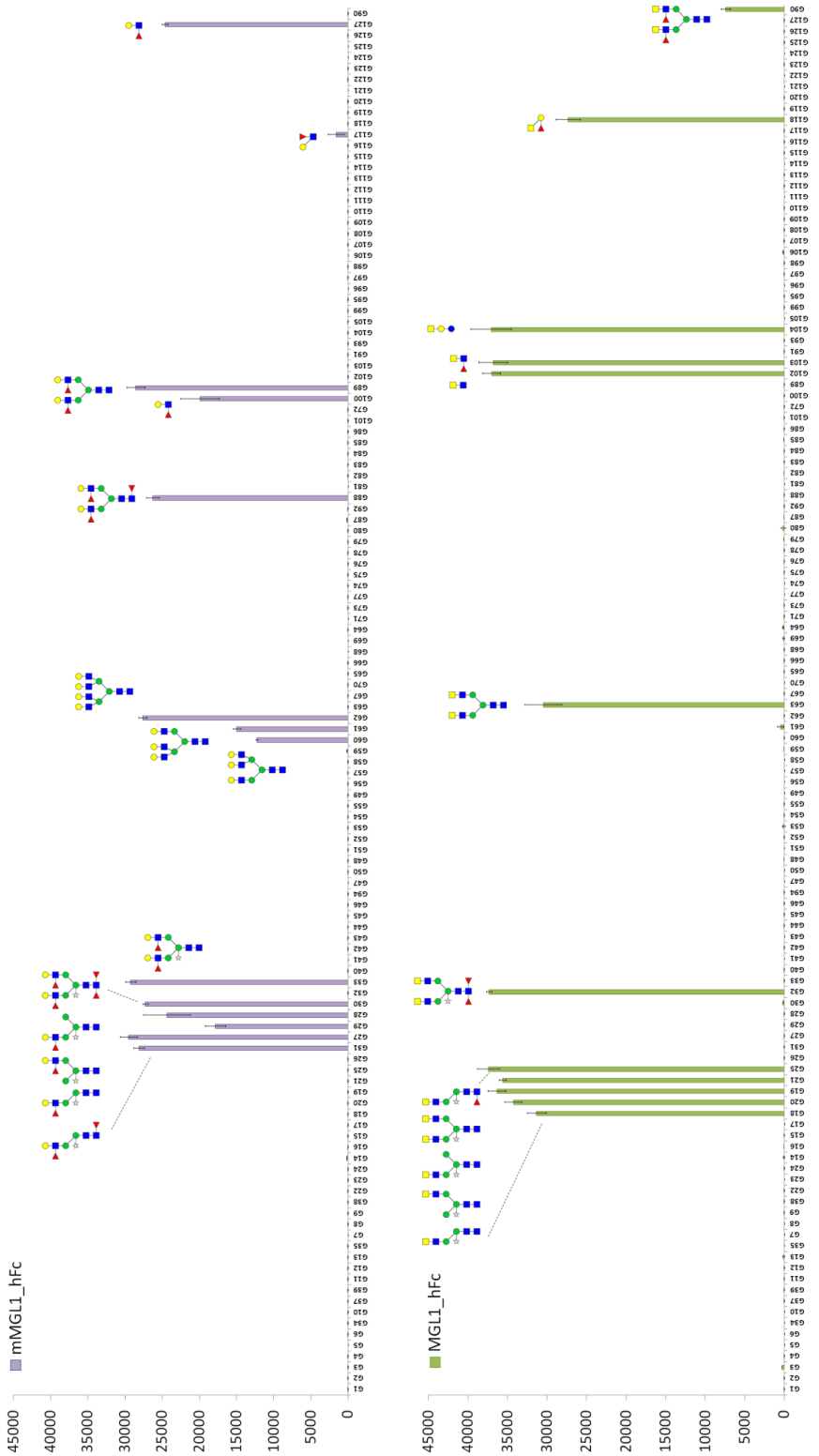


Figure 12. Fluorescence intensity after incubation with mMGL1 and MGL1 (10µg/mL) followed incubation with secondary labeled anti-human antibody. Each bar in the histogram shows the average of fluorescence from 4 spots and the standard deviation as error bar.

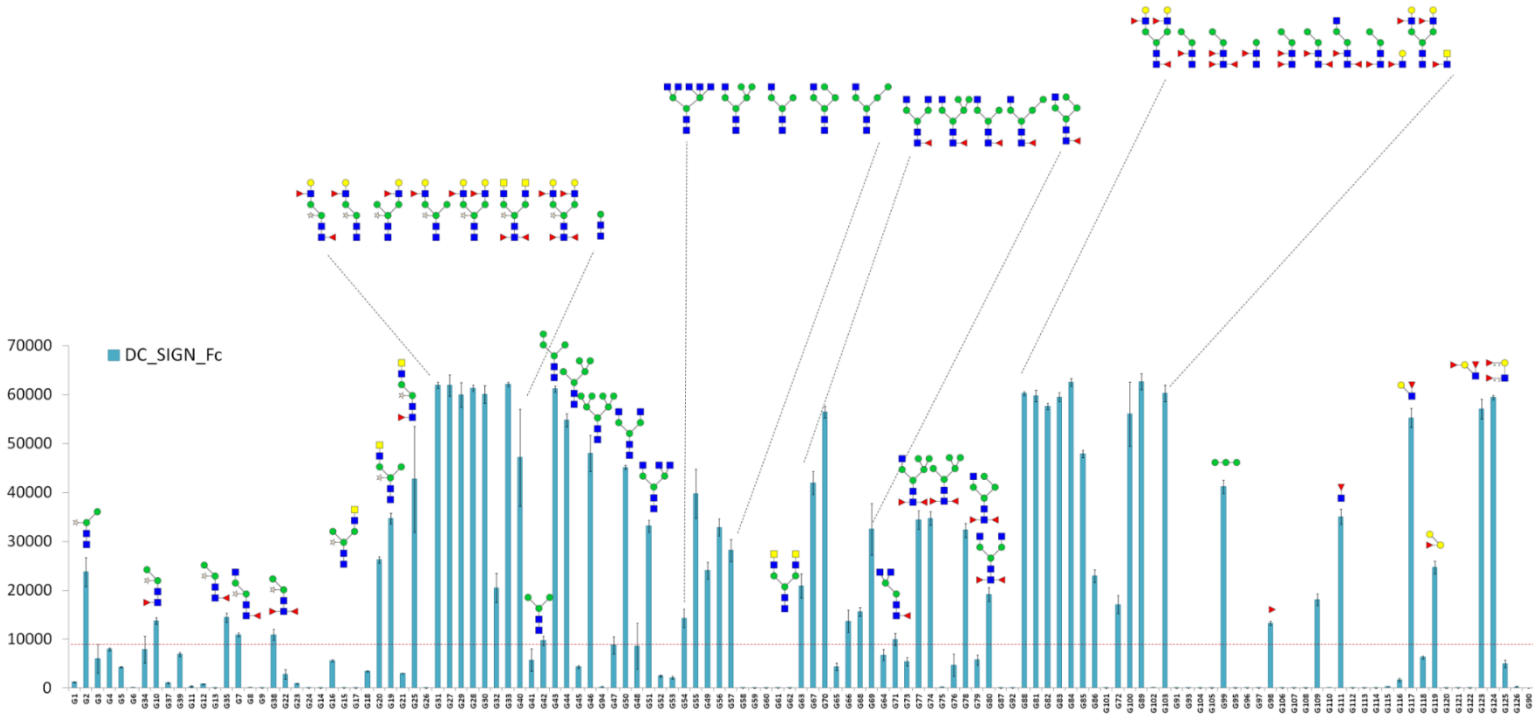


Figure 13. Fluorescence intensity after incubation with DC-SIGN (10µg/mL) followed incubation with secondary labeled anti-human antibody. Each bar in the histogram shows the average of fluorescence from 4 spots and the standard deviation as error bar.

Incubation of DC-SIGN in our glycan microarray confirmed the known specificity towards high-mannose, LDNF, blood type antigens and perhaps less known, complex glycans with terminal GlcNAc on the 3-arm. Best binders in terms of fluorescence signal intensity were Le^X presenting N-glycans (**G27**-**31**, **G33**, **G88**, **G89**) (Figure 13).

Interestingly, at this point we could observe that the presence of core xylose in N-glycan structures completely abolished binding to the GlcNAc terminal (3-arm) oligosaccharides including biantennary and hybrid N-glycans (Figure 14). The binding pattern is similar to that of other mannose binding proteins observed previously, ConA and GNA. The presence of xylose residues seems to distort the glycan conformation in such a way that the lectin can no longer make favorable contacts with the binding epitope. No significant effect of xylose on binding to the Lewis X presenting N-glycans structures was observed (**G30**, **G33**), the conformation and accessibility of the Lewis X branches is clearly not compromised by the additional core xylose modification.[67] Xylosylated glycans which do not bind DC-SIGN but conserve binding to other C-type lectins could become lead structures for glycan mediated specific targeting of antigens to C-type lectin receptors on dendritic cells in immune therapy.

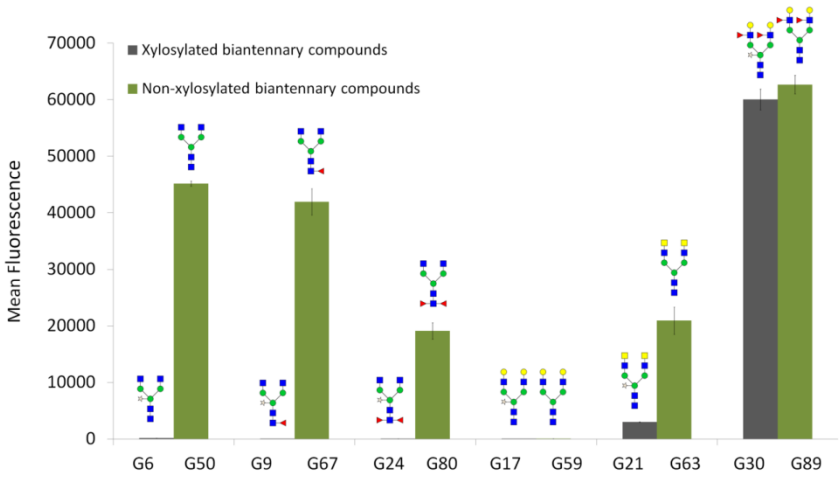


Figure 14. Interactions of DC-SIGN with xylosylated (grey histograms) and non-xylosylated glycan series of bi-antennary compounds (green histograms).

The other C-type lectins tested on the glycan microarray did not give binding towards any structure presented. Several reasons could be argued to explain this behavior, one of the most obvious is that we might simply lack the natural ligands of the proteins or that the glycan fragments available on the array have to be presented in the context of a larger structure, like protein or lipid. The other important issue could be the presentation and the accessibility of the C-type lectins CRD which might not be sufficient for effective binding of glycans. Further explanation will be given in the Chapter 4, where the protein G-coated C-type lectin microarray was successfully prepared and probed with different fluorescently labeled glycoconjugates.

2.4 Summary

Including the newly synthesized 39 core-xylosylated N-glycans which preparation was presented in the Chapter 1, a large microarray containing non-xylosylated glycan structures and smaller glycan fragments was designed and prepared for the array-based analysis of glycan binding specificities of series of plant, fungi and animal lectins. In total 126 carbohydrates were printed on NHS-activated glass slide using a Piezo-electric robotic non contact printer. We have shown the first functionality screenings of glycan array with plant and fungi lectins, specifically recognizing certain glycan elements. Already here, we could observe a great potential of our synthetic array towards identification of the ligands of carbohydrate binding proteins in the context of larger N-glycan structure and have an insight into the finer specificity of analyzed lectin. We discovered the influence of the core xylose on the interaction of N-glycans with mannose binding lectins, C*oncanavalin A* and *Galanthus nivalis*. The presence of core xylose reduced the interactions with ConA and completely abolished binding towards GNA. Taking into account that the β-1,2-xylose is a common modification of plant N-glycans, above mentioned impact of xylose moiety on the lectins specificity could be rationalized by a tolerance mechanism avoiding interactions between plant lectins and self-glycans.

Further, we have performed a screening of murine and human C-type lectins receptors, paying a special attention on the proteins which may be involved in the host-pathogen interactions during curse of *S. mansoni* infection. As a result, we could observe binding of human MGL-1 receptor to the LDN and LDNF glycan elements and the Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin (DC-SIGN) to set of Lewis X terminated glycan structures. Our results confirmed the previous reports describing MGL as a pattern recognizing receptor for *S. mansoni* binding to the soluble egg antigens (SEA) expressing LDN and LDNF glycoconjugates and the DC-SIGN binding the Lewis X type of glycans that have been found on the glycoconjugates from all life cycle stages of *S. mansoni*. Interestingly, as it had place during the GNA characterisation, we noted the differences in the binding of DC-SIGN receptor towards xylosylated and non-xylosylated biantennary N-glycans. Similar to the plant lectin, the presence of core xylose completely abolished binding towards DC-SIGN studied on the glycan microarray. Inspired by the influence of the xylose moiety on the recognition of biantennary compound by the DC-SIGN receptor, we have further design an *in vitro* assay employing OVA glycoconjugates containing above mentioned sugars (**G6** and **G50**) to compare their abilities towards targeting of murine dendritic cells. The study will be described in details in the Chapter 4.

2.5 Experimental Part

General Methods

Materials: *Canavalia ensiformis* lectin (ConA), Wheat germ agglutinin (WGA), *Ricinus communis* agglutini -120 (RCA-120), *Aleuria aurantia* lectin (AAL), *Pisum sativum* agglutinin (PSA), *Galantus nivalis* agglutinin (GNA), *Griffonia* (*Bandeiraea) simplicifolia* lectin *II* (BSL-II) and *Wisteria floribunda* lectin (WFL) were purchased from VectorLabs. Lectins were labeled with Hilyte Plus^{TM} 647 and Hilyte Plus^{TM} 555 protein labeling kits from AnaSpec, Freemont, USA. Recombinant human DC-SIGN, MGL-1 and Dectin-2 Fc chimeras were purchased from

Sino Biological Inc., Beijing, P. R. China. Murine-hFc C-type lectins were provided by Prof. Dr. Bernd Lepenies.[44] Anti-human IgG (Fc specific)-Cy3, developed in goat, was purchased from Sigma-Aldrich. Glycan arrays were compartmentalized with 8 Well ProPlate Module 7 x 16 mm Well from Electron Microscopy Sciences. NHS activated glass slides, Nexterion® H were purchased from Schott AG, Mainz, Germany. All aqueous solutions were prepared from nanopure water produced with a Diamond UV water purification system (Branstead International, IA, Madrid, Spain). Microarrays were printed employing a robotic non contact piezoelectric SciFLEXARRAYER spotter S11 (Scienion, Berlin, Germany). Arrays were imaged on an Agilent G2565BA fluorescence scanner system (Agilent Technologies, Santa Clara, USA) at 10 µm resolution, using 2 lasers (532 nm, 633 nm). Quantification of fluorescence was achieved by ProScanArray® Express software from Perkin Elmer, Shelton, USA. A quantification method using an adaptive circle with minimum diameter of 50 µm and maximum diameter of 200 µm was employed. Each histogram represents the average of mean fluorescence from four replicates with local background subtraction.

Immobilization of the ligands and test for biofunctionality of the printed ligands

Ligand solutions were prepared from stock solutions (1 mM in water) by dilution with sodium phosphate buffer (300 mM, pH 8.4, 0.005% Tween 20) to a final concentration of 50 µM. 40 µL of each glycans spotting solution was placed into a 384 well source plate (Scienion, Berlin, Germany) which was stored at -20°C and reused if necessary. These solutions (1.25 nL, 5 drops of 250 pL) were spatially arrayed onto NHS functionalized glass slides (Nexterion® H, Schott AG, Mainz, Germany) with a distance of 400 µm. 126 glycans were spotted in 4 replicates (9 different N-glycans/row), establishing the complete 36x18 array which was printed in 7 copies onto each slide. After printing, the slides were placed in a 75 % humidity chamber (saturated NaCl solution) at 25°C for 18 hours. The remaining NHS groups were quenched by placing the slides in a 50 mM solution of ethanolamine in sodium borate buffer 50 mM, pH 9.0, for 1h. The standard washing of the slides was performed

with PBST (PBS solution containing 0.5% Tween 20), PBS and water. The slides were dried in a slide spinner.

Interaction with plant lectins

Each subarray was incubated with a solution of fluorescently labelled lectin. Solutions of *Galanthus nivalis* agglutinin (GNA-555, 22 µg/mL), *ConcanavalinA* (ConA-647, 22 µg/mL), *Aleuria aurantia* lectin (AAL-555, 20 µg/mL), *Bandeiraea simplicifolia* lectin (BSL-II-647, 24 µg/mL) (Figure X), *Wisteria floribunda* lectin (WFL-555, 10 µg/mL) and *Ricinus communis* agglutinin (RCA-120-555, 1 µg/mL), were prepared in PBS, 2 mM $CaCl_2$, 2 mM $MgCl_2$ containing 0.01% Tween-20. For incubations, 200 µL of each lectin solution was applied to each subarray and incubated in the dark for one hour at room temperature. The slides were washed under standard conditions, dried and the fluorescence was analyzed with a microarray scanner.

Glycan binding analysis of C-type lectins

mMGL-1, MGL-1 and DC-SIGN solutions were prepared from stock solutions (0.25 mg/mL in PBS) by dilution with TSMT binding buffer (20 mM Tris·HCl, 150 mM NaCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 1% BSA and 0.005% Tween-20 pH=7.5) to a final concentration of 10 µg/mL whereas the Dectin-2 to the final concentration of 10, 20, 75 and 100 µg/mL. The murine C-type lectins Clec-2, Mincle, Clec-9A, DCAR, DCIR, SignR3 were diluted with TSMT binding buffer to a final concentration of 10 and 20 µg/mL. Lectins solutions (200 µL) were incubated overnight at 4 °C with gentle shaking. Solutions containing protein were removed and each subarray was washed with TSMT binding buffer, followed by 1h incubation with anti-human IgG (Fc specific)-Cy3 in TSM binding buffer (1:1000). The washing of slides was performed with TSM_T (TSM containing 0.005% Tween-20), TSM and water. Slides were dried in a slide spinner. The fluorescence was analyzed with a microarray scanner.

References

(1) Maverakis, E.; Kim, K.; Shimoda, M.; Gershwin, M. E.; Patel, F.; Wilken, R.; Raychaudhuri, S.; Ruhaak, L. R.; Lebrilla, C. B. *Journal of autoimmunity* **2015**, *57*, 1–13.

(2) Guzman-Aranguez, A.; Argüeso, P. *The ocular surface* **2010**, *8*, 8–17.

(3) Geissner, A.; Anish, C.; Seeberger, P. H. *Current opinion in chemical biology* **2014**, *18*, 38–45.

(4) Collins, B. E.; Paulson, J. C. *Current opinion in chemical biology* **2004**, *8*, 617–625.

(5) Park, S.; Lee, M.-R.; Shin, I. *Chemical Communications* **2008**, *37*, 4389–4399.

(6) Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; others. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 17033–17038.

(7) Feizi, T.; Fazio, F.; Chai, W.; Wong, C.-H. *Current opinion in structural biology* **2003**, *13*, 637–645.

(8) Kulesh, D. A.; Clive, D. R.; Zarlenga, D. S.; Greene, J. J. *Proceedings of the National Academy of Sciences of the United States of America* **1987**, *84*, 8453–8457.

(9) Heimburg-Molinaro, J.; Song, X.; Smith, D. F.; Cummings, R. D. *Current protocols in protein science* **2011**, *12*, 1–33.

(10) Kajihara, Y.; Suzuki, Y.; Sasaki, K.; Juneja, L. R. *Methods in enzymology* **2003**, *362*, 44–64.

(11) Novotny, M. V.; Alley Jr, W. R. *Isolation and Purification of Glycans from Natural Sources for Positive Identification*. In *The Proceedings of the Beilstein Glyco-Bioinformatics Symposium* 2013, 133–148.

(12) Wang, D.; Liu, S.; Trummer, B. J.; Deng, C.; Wang, A. *Nature biotechnology* **2002**, *20*, 275–281.

(13) Fukui, S.; Feizi, T.; Galustian, C.; Lawson, A. M.; Chai, W. *Nature biotechnology* **2002**, *20*, 1011–1017.

(14) Beloqui, A.; Calvo, J.; Serna, S.; Yan, S.; Wilson, I.BH.; Martin-Lomas, M.; Reichardt, N.C. *Angewandte Chemie International Edition* **2013**, *52*, 7477--7481.

(15) Reichardt, N.C.; Serna, S.; Echevarria, J. *Glycoarrays: An Invaluable Tool for Glycomics*. In *Carbohydrate Chemistry: State of the Art and Challenges for Drug Development: An Overview on Structure, Biological Roles, Synthetic Methods and Application as Therapeutics;* World

Scientific, 2015; 147-172.

(16) Oyelaran, O.; Gildersleeve, J. C. *Current Opinion in Chemical Biology* **2009**, *13*, 406–413.

(17) Park, S.; Lee, M.-R.; Shin, I. *Bioconjugate chemistry* **2009**, *20*, 155–162.

(18) Zhou, X.; Zhou, J. *Biosensors and Bioelectronics* **2006**, *21*, 1451–1458.

(19) Serna, S.; Etxebarria, J.; Ruiz, N.; Martin-Lomas, M.; Reichardt, N. C. *Chemistry-A European Journal* **2010**, *16*, 13163–13175.

(20) De Boer, A. R.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Analytical chemistry* **2007**, *79*, 8107–8113.

(21) Zhang, X.; Oglesbee, M. *Biological procedures online* **2003**, *5*, 170–181.

(22) Min, D.-H.; Yeo, W.-S.; Mrksich, M. *Analytical chemistry* **2004**, *76*, 3923–3929.

(23) Northen, T. R.; Lee, J.-C.; Hoang, L.; Raymond, J.; Hwang, D.-R.; Yannone, S. M.; Wong, C.-H.; Siuzdak, G. *Proceedings of the National Academy of Sciences* **2008**, *105*, 3678–3683.

(24) Brzezicka, K.; Echeverria, B.; Serna, S.; van Diepen, A.; Hokke, C. H.; Reichardt, N.-C. *ACS Chemical Biology* **2015**, *10*, 1290–1302.

(25) Vázquez-Mendoza, A.; Carrero, J. C.; Rodriguez-Sosa, M. *BioMed research international* **2013**, *2013*, 1–11.

(26) Marth, J. D.; Grewal, P. K. *Nature Reviews Immunology* **2008**, *8*, 874–887.

(27) Cohen, M. *Biomolecules* **2015**, *5*, 2056–2072.

(28) Peumans, W. J.; Damme, E. J. V. *Biotechnology and Genetic Engineering Reviews* **1998**, *15*, 199–228.

(29) Lannoo, N.; Van Damme, E. J. M. *Frontier in Plant Science* **2014**, *5*, 397.

(30) Howe, G. A.; Schaller, A. *Direct defenses in plants and their induction by wounding and insect herbivores*. In *Induced plant resistance to herbivory*; Springer, 2008; 7–29.

(31) Peumans, W. J.; Van Damme, E. *Plant physiology* **1995**, *109*, 347-352.

(32) Lepenies, B.; Lee, J.; Sonkaria, S. *Advanced drug delivery reviews* **2013**, *65*, 1271–1281.

(33) Cambi, A.; Koopman, M.; Figdor, C. G. *Cellular microbiology* **2005**, *7*, 481–488.

(34) Figdor, C. G.; van Kooyk, Y.; Adema, G. J. *Nature Reviews Immunology* **2002**, *2*, 77–84.

(35) Van Kooyk, Y.; Rabinovich, G. A. *Nature immunology* **2008**, *9*, 593–601.

(36) Geijtenbeek, T. B.; Gringhuis, S. I. *Nature Reviews Immunology* **2009**, *9*, 465–479.

(37) Holla, A.; Skerra, A. *Protein Engineering Design and Selection* **2011**, *24*, 659–669.

(38) Manimala, J. C.; Roach, T. A.; Li, Z.; Gildersleeve, J. C. *Angewandte Chemie* **2006**, *118*, 3689–3692.

(39) Garcia-Vallejo, J. J.; Unger, W. W.; Kalay, H.; van Kooyk, Y. *Oncoimmunology* **2013**, *2*, 230401-230403.

(40) Eriksson, M.; Serna, S.; Maglinao, M.; Schlegel, M. K.; Seeberger, P. H.; Reichardt, N.-C.; Lepenies, B. *Chembiochem* **2014**, *15*, 844–851.

(41) Zhu-Salzman, K.; Shade, R. E.; Koiwa, H.; Salzman, R. A.; Narasimhan, M.; Bressan, R. A.; Hasegawa, P. M.; Murdock, L. L. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 15123–15128.

(42) Porter, A.; Yue, T.; Heeringa, L.; Day, S.; Suh, E.; Haab, B. B. *Glycobiology* **2010**, *20*, 369–380.

(43) Ramos, M. V.; Grangeiro, T. B.; Cavada, B. S.; Shepherd, I.; Lopes, R. O. de M.; Sampaio, A. H. *Brazilian Archives of Biology and Technology* **2000**, *43*, 1–11.

(44) Maglinao, M.; Eriksson, M.; Schlegel, M. K.; Zimmermann, S.; Johannssen, T.; Götze, S.; Seeberger, P. H.; Lepenies, B. *Journal of Controlled Release* **2014**, *175*, 36–42.

(45) Hsu, T.-L.; Cheng, S.-C.; Yang, W.-B.; Chin, S.-W.; Chen, B.-H.; Huang, M.-T.; Hsieh, S.-L.; Wong, C.-H. *Journal of Biological Chemistry* **2009**, *284*, 34479–34489.

(46) Singh, S. K.; Streng-Ouwehand, I.; Litjens, M.; Weelij, D. R.; Garcia-Vallejo, J. J.; van Vliet, S. J.; Saeland, E.; van Kooyk, Y. *Molecular immunology* **2009**, *46*, 1240–1249.

(47) Suzuki-Inoue, K.; Inoue, O.; Ozaki, Y. *Journal of Thrombosis and Haemostasis* **2011**, *9*, 44–55.

(48) Huysamen, C.; Brown, G. D. *FEMS microbiology letters* **2009**, *290*, 121–128.

(49) Furukawa, A.; Kamishikiryo, J.; Mori, D.; Toyonaga, K.; Okabe, Y.; Toji, A.; Kanda, R.; Miyake, Y.; Ose, T.; Yamasaki, S.; others. *Proceedings of the National Academy of Sciences of the United States of Americ* **2013**, *110*, 17438–17443.

(50) Feinberg, H.; Jégouzo, S. A.; Rowntree, T. J.; Guan, Y.; Brash, M. A.; Taylor, M. E.; Weis, W. I.; Drickamer, K. *Journal of Biological Chemistry* **2013**, *288*, 28457–28465.

(51) Kerscher, B.; Willment, J. A.; Brown, G. D. *International immunology* **2013**, *25*, 271–277.

(52) Lambert, A. A.; Gilbert, C.; Richard, M.; Beaulieu, A. D.; Tremblay, M. J. *Blood* **2008**, *112*, 1299–1307.

(53) Bloem, K.; Vuist, I. M.; van der Plas, A.-J.; Knippels, L. M.; Garssen, J.; Garc'\ia-Vallejo, J. J.; van Vliet, S. J.; van Kooyk, Y. *PloS one* **2013**, *8*, e66266.

(54) Galustian, C.; Park, C. G.; Chai, W.; Kiso, M.; Bruening, S. A.; Kang, Y.-S.; Steinman, R. M.; Feizi, T. *International immunology* **2004**, *16*, 853–866.

(55) Takahara, K.; Yashima, Y.; Omatsu, Y.; Yoshida, H.; Kimura, Y.; Kang, Y.-S.; Steinman, R. M.; Park, C. G.; Inaba, K. *International immunology* **2004**, *16*, 819–829.

(56) Eriksson, M.; Johannssen, T.; von Smolinski, D.; Gruber, A. D.; Seeberger, P. H.; Lepenies, B. *Frontiers in immunology* **2013**, *4,* 196.

(57) Zhang, F.; Ren, S.; Zuo, Y. *International reviews of immunology* **2014**, *33*, 54–66.

(58) Garcia-Vallejo, J. J.; van Kooyk, Y. *Trends in immunology* **2013**, *34*, 482–486.

(59) Kooyk, Y. van; Engering, A.; Lekkerkerker, A. N.; Ludwig, I. S.; Geijtenbeek, T. B. *Current opinion in immunology* **2004**, *16*, 488–493.

(60) Van Liempt, E.; Bank, C.; Mehta, P.; Kawar, Z. S.; Geyer, R.; Alvarez, R. A.; Cummings, R. D.; Kooyk, Y. van; van Die, I.; others. *FEBS letters* **2006**, *580*, 6123–6131.

(61) Van Vliet, S. J.; van Liempt, E.; Saeland, E.; Aarnoudse, C. A.; Appelmelk, B.; Irimura, T.; Geijtenbeek, T. B.; Blixt, O.; Alvarez, R.; van Die, I.; others. *International immunology* **2005**, *17*, 661–669.

(62) Drickamer, K. *Current opinion in structural biology* **1999**, *9*, 585–590.

(63) Tanne, A.; Ma, B.; Boudou, F.; Tailleux, L.; Botella, H.; Badell, E.; Levillain, F.; Taylor, M. E.; Drickamer, K.; Nigou, J.; others. *The Journal of experimental medicine* **2009**, *206*, 2205–2220.

(64) Gringhuis, S. I.; den Dunnen, J.; Litjens, M.; van der Vlist, M.; Geijtenbeek, T. B. *Nature immunology* **2009**, *10*, 1081–1088.

(65) Meyer, S.; Van Liempt, E.; Imberty, A.; Van Kooyk, Y.; Geyer, H.; Geyer, R.; Van Die, I. *Journal of Biological Chemistry* **2005**, *280*, 37349–37359.

(66) Van Liempt, E.; Imberty, A.; Bank, C. M.; Van Vliet, S. J.; Van Kooyk, Y.; Geijtenbeek, T. B.; Van Die, I. *Journal of Biological Chemistry* **2004**, *279*, 33161–33167.

(67) Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I. *Science* **2001**, *294*, 2163–2166.

CHAPTER 3

GLYCAN MICROARRAY PROFILING AND ANTI-CARBOHYDRATE ANTIBODIES

3.1 Introduction

3.1.1 Glycan arrays as a tool for antibody profiling

Carbohydrates found on pathogens are often significantly different from the mammalian set of glycans and may induce important immune responses. In recent years, carbohydrate antigen arrays have become an excellent tool for high-throughput screening of anti-glycan antibodies and have been applied for the carbohydrate-based vaccines development.[1–4] In general, glycan array can be used in the discovery and characterization of antigens but also to evaluate the immune response to vaccine candidates.

Arrays of microbial carbohydrates have been used for the identification of bacterial antigens and for diagnosis by analyzing serum samples from infected patients. For example, Wang and coworkers were able to identify an antigenic sugar structure from gram positive bacteria *Bacillus anthracis*, using a glycan array composed of carbohydrates from the major surface glycoprotein (Bcla) of bacterial spores.[5] In the other study, glycan arrays have been successfully applied for the characterization of carbohydrate specificity of broadly neutralizing antibodies (bnAbs) in HIV infection. The binding profile of the HIV-1 2G12 bnAb isolated from an asymptomatic HIV-positive patients[6,7,8,9] was analyzed by screening of a glycan array composed of oligomannose structures found on HIV viral coat glycoprotein (gp120). The study revealed that the antibody was capable to bind high mannose type N-glycans bearing Manα-1,2-Man residues.[6] Further investigation defined the minimal epitope required for antibody to bind and identified the structure with highest affinity towards 2G12 bnAb, which was Man_4 oligomannoside.[7] In a more recent study, Wong and colleagues compiled several bi-, tri- and tetra-antennary sialylated complex oligosaccharides to a focused glycan array which was subsequently probed with the HIV-1 bnAbs, PG9 and PG16. The authors identified natural and un-natural glycan ligands of PG16 on their array, which were then conjugated to the diphtheria toxin and analysed in a dose

dependant manner with the antibodies as a first step towards a novel HIV vaccine candidate.[10]

Beside the bacterial or viral infections, identification and characterization of glycan antigens can be also important for the development of new treatment or diagnostics towards cancer. Cancer cells express numbers of tumor-associated carbohydrate antigens, such as sialyl Lewis A ((Neu5Acα1-3)Galβ1-3(Fucα1-4)GlcNAc), Tn (GalNAcα-Ser/Thr) , Lewis Y ((Fucα1-2)Galβ1-4(Fucα1-3)GlcNAc) or GloboH, which may be used as a biomarker of disease but also can become a target for the treatment development.[11] GloboH is a glycolipid hexasaccharide expressed on several tumors including breast, colon, lung, prostate and ovary and thus this glycan with conjugation to antigenic protein has become a good starting point for anti-cancer vaccine development.[1] Screening of anti-GloboH monoclonal antibodies and polyclonal antibodies from cancer patient sera on glycan microarray defined the minimal epitope of GloboH required for antibody recognition[12]. A tetrasaccharide was found to bind anti-GloboH antibodies with same affinity as larger hexasaccharide and because of reduced synthetic steps became a better candidate for the development of potential anti-cancer vaccine.

There are also several examples where glycan arrays have been used for the study of biological functions of glycans during the course of parasite infections. Hokke and coworkers profiled the monoclonal and polyclonal antibodies from *Schistosoma mansoni* infected mice and humans employing natural shotgun glycan microarrays.[13,14] Schistosomiasis is a chronic and potentially deadly parasitic disease, which can be treated with praziquantel (PZQ)[15], however the drug does not prevent from re-infection and thus the need of vaccine development emerged in recent years. It is known that the host immune response induced against parasite is directed towards its certain carbohydrates.[16] Their identification and further characterization can be performed on the shotgun array comprises N-glycans and lipid-glycan structures directly isolated from the four life stages of *S. mansoni*: larvae (cercariae), male adult worm, female adult worm and eggs. The microarray-based studies allow for the detection of serum antibodies towards the whole

natural glycome of parasite, including many unique glycan structures still not available synthetically. Screenings performed on the natural glycan array containing glycans or the glycan mixtures from the whole glycome of the parasite, is a good method to identify antigenic elements and to narrow the number of potentially important structures when no previous data is available. However, because some of the ligands printed on the shotgun array contains mixture of different glycan elements and structures, in some cases it may be difficult to speculate which specific responses occur to which glycan element. [14]

In this chapter we will describe use of our focused, synthetic glycan array as a complementary tool towards studying the biological role of glycan in *S. mansoni* infections. This work was performed in collaboration with Dr. C.H. Hokke from the Leiden Parasite Glycobiology Group (Leiden University Medical Center, the Netherlands) and brings together results of monoclonal antibody incubation with the screening of serum samples from *S. mansoni* infected patients.

3.1.2 Glycome of *Schistosoma mansoni*

Schistosomiasis also known as bilharzia, is a tropical disease caused by trematode worms, which infect over 200 millions of people worldwide according to the WHO organization (Figure 1, A). It is particularly abundant in the developing countries with inadequate sanitation where parasite is transmitted to the humans in a contact with excreta contaminated water.[17,18] There are three main disease-causing species, *S. haematobium*, *S. japonicum* and *S. mansoni*, with the latter one being the interest of this Chapter. Human *S. mansoni* has a very complex life-cycle (Figure 1, B) involving a definitive host (mammalian) where sexual worm reproduction takes place and a freshwater snail where the parasite multiplies in an asexual manner. The infection is initiated by cercaria which are released by the snail into the water and which can penetrate through even intact skin of human host. During the skin penetration, cercariae loses its tail and transforms into schistosomula which migrates to the portal blood system and maturate into male and female adult worms.

Adult worms form male-female pairs and migrate to the blood vessels in the lower intestine. Adult females deposit eggs in the bloodstream (up to 200 eggs per day), many of which embolize in smaller blood vessels or capillaries of diverse organs, including urinary bladder or intestine. From there, eggs can be transported to the liver or are excreted in feces or urine.

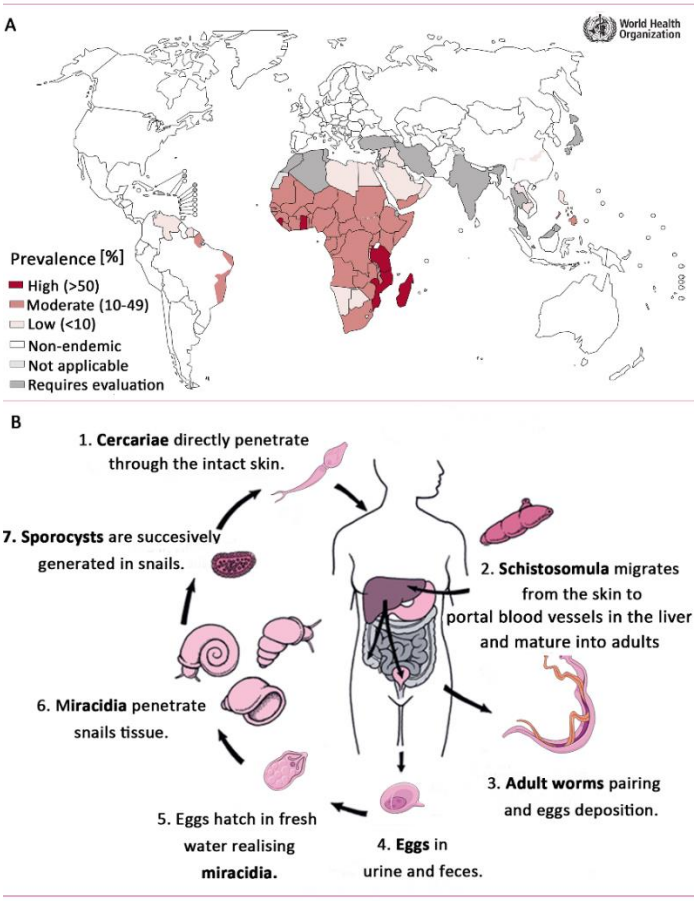


Figure 1. *Schistosoma mansoni*; **A.** Worldwide distribution of parasite (2012) showing the national prevalence [%] (data source: World Health Organization WHO); **B.** Life cycle of *S. mansoni*.

Eggs from feces or urine hatch in fresh water and release miracidia which subsequently penetrate snail tissues and successively generate the free-swimming cercariae to complete the developmental cycle of parasite.[19]

There are two immunological phases of schistosomiasis infection, the acute and chronic phase.[20] The acute phase starts when the cercariae penetrates the host skin and lasts until the deposition of eggs by the adult female (usually sixth week after infection). It is dominated by the Th1 pro-inflammatory immune response. The chronic phase of schistosomiasis is characterized by the spontaneous down modulation of the initial inflammatory response and switch from Th1 into Th2 immune response. It is also characterized by high levels of IgE antibodies. In this state, *Schistosoma* may resist in the host for years and even after PZQ drug treatment, reinfection usually occurs. The progress of the disease depends on the individual and in some cases, it might lead to life-threatening inflammatory and obstructive conditions.[18] It has been demonstrated that infected individuals living in endemic areas acquire certain level of immunity towards the parasite. However, it takes many years of exposure, several reinfections and drug treatments for the development of a protective immune response. Possibly, when humans are treated with PZQ the parasite dies, it falls apart and exposes antigens which further induce more effective immune responses in the host. It may subsequently lead to maturation of immune system and in some cases produce immunity against the parasite.[13,20]

During its life cycle, *S. mansoni* express a high number of glycans as a part of its glycoproteins and glycolipids. The detailed glycomic spectrometric analysis of the parasite extracts at different developmental stages revealed that *S. mansoni* express very complex and stage-specific sugars (Figure 2) and some of them, markedly differs from the mammalian ones (Figure 2).[21,22,23] In a typical experiment N-glycans are released from the total protein extracts by treatment with PNGase-F to cleave off all asparagine-linked complex, hybrid, or high mannose oligosaccharides unless their core contains an α-1,3-fucose. Fractions are subsequently incubated with trypsin and then treated with PNGase-A to cleave remaining α-1,3-fucose containing glycans. Released glycans are purified, labeled and fractionated by hydrophobic interaction liquid chromatography with fluorescent detection.[21] Collected fractions are subsequently analyzed by MALDI-TOF MS.

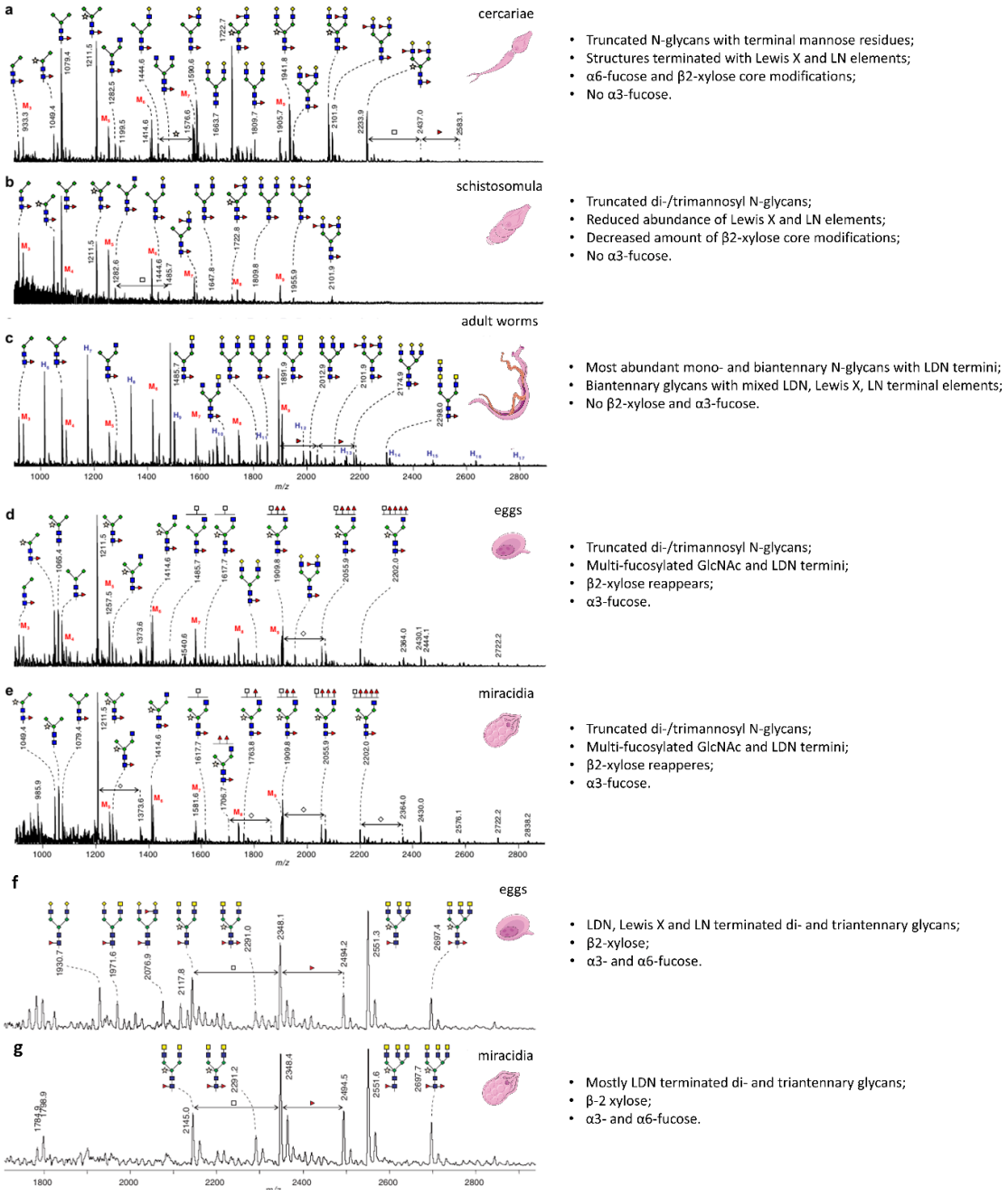


Figure 2. Profiles of PNGase F (**a-e**) and PNGase A (**f-g**) released N-glycans from different *S. mansoni* life-cycle stages and measured by MALDI-TOF-MS in the positive ion-reflector mode (taken from Hokke *at al*.[21])

It was shown that the majority of cercarial N-glycans are modified with β-1,2-xylose and core α-1,6-fucose while no core α-1,3-fucose is observed. Additionally, several truncated mannose terminated N-glycans appears and sugars presenting Lewis X antigen (Figure 2, a). Schistosomula express lower amounts of terminal Lewis X sugars and core xylosylated compounds compared to cercariae but the presence of truncated di- or trimannosidic N-glycans is more pronounced (Figure 2, b).

In adult worms, β-1,2-xylose completely disappears and Lewis X elements become less abundant compared to schistosomula N-glycans. N-glycosylation of adult parasites is mostly characterized by the presence of mono- and biantennary glycans with LDN termini and core α-1,6-fucosylation (Figure 2, c).

In eggs, the N-glycan profile changes again and β-1,2-xylosylated, α-1,3-fucosylated and difucosylated (core α-1,3/α-1,6) structures can be now detected (Figure 2, d,f). The truncated N-glycans with terminal mannose residues reappears at high abundance; the antenna of remaining complex N-glycans are decorated with LDN termini, including Lewis X and LDNF moieties. Another characteristic of egg glycans is the multi-fucosylation of GlcNAc antenna and the presence of unusual multifucosylated motifes like, Fuc(α1-2)Fuc(α1-3)GlcNAc motif or Fuc(α1-2)Fuc(α1-3)GalNAc(β1-4)[Fuc(α1-2)Fuc(α1-3)]GlcNAc (DF-LDN-DF). N-glycans from miracidia show comparable N-glycan profile to eggs. (Figure 2, e,g).

Even though *S. mansoni* expresses multiple and often unique glycans, their role during the parasite infection in the host is still not completely understood. It is well known that high levels of anti-carbohydrates antibodies are generated during the course of infection, however, they might be involved either in host protection or act like a smoke screen that could be beneficial for the parasite rather than for the host.[20,24] The smoke screen theory refers to the subverting of the immune system away from epitopes that could provoke protective immunity by the high antibody responses induced against other carbohydrate elements.[20] Generally, during the course of *S. mansoni* infection, high levels of IgE and IgA antibody responses are associated with increased resistance to re-infection. Two IgG subclasses, IgG1 and IgG3 are correlated with protection while IgG2 can have dual functions

and act as protective or blocking antibody. IgG4 on the other hand can block protective IgE functions and high IgG4/IgE ratio is associated with susceptibility to re-infection similar to high levels of IgM. [20]

In this Chapter, we will evaluate functionality of our glycan arrays towards profiling of anti-glycan antibodies using anti-glycan monoclonal antibodies (mAbs) of *S. mansoni* infected mice and compare acquire results of specificities in a context of complex N-glycans printed on our array with previously published studies. Later we will focus on the characterization of polyclonal antibodies from the sera of infected individuals from schistosomiasis endemic areas. We will study the age-related anti-carbohydrate IgG and IgM responses in sera of infected children and adults and compare our results with the data obtained from the shotgun arrays. Finally, we will try to evaluate potential role of sugars during the course of infection.

3.2. Glycan array assisted characterization of *S. mansoni* specific monoclonal antibodies

To test the functionality of our glycan array towards characterization of anti-carbohydrate antibodies but also to characterize the specificity of antibodies in the context of glycans present on the array, we first probed a group of 17 monoclonal antibodies (mAbs) generated from serum of *S. mansoni* infected mice.[25]

Monoclonal antibodies are single type antibodies secreted by immortal hybridoma cells that have been created by fusion of antigen activated B-cells isolated from the spleen of immunized mice with a non-secreting myeloid cell line (Figure 3).[26] Hybridomas can grow in culture media indefinitely and secrete antibodies of desired specificity what makes them especially attractive for research. Nowadays, monoclonal antibodies have found multiple applications in the diagnosis and treatment of many diseases, like rheumatoid arthritis (Rituxan) or certain cancers (Cetuximab).[27]

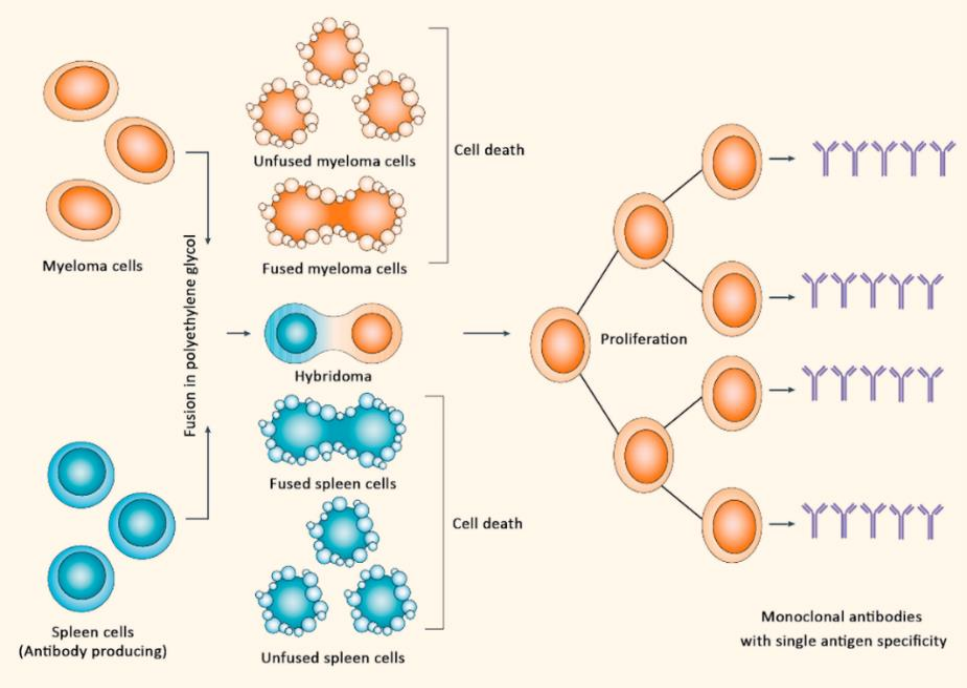


Figure 3. Production of monoclonal antibodies (taken from Sefik S. Alkan *at al*.[26])

The available library of mAbs specific for schistosome antigens has included two immunoglobulin isotypes, immunoglobulin M (IgM) and G (IgG). Pentameric-IgM is the largest antibody with five complement-binding sites and plays a major role in the intravascular neutralization of for example certain viruses. IgM is produced immediately after exposure to the antigen, disappears few weeks later and thus usually indicates acute or current infections. IgG on the other hand is the most abundant immunoglobulin isotype, smaller than IgM and thus able to penetrate into human tissues. It is also the most important antibody of the humoral immune response, which is associated with the acquisition of long term immunity, induction of opsonization or antibody dependent cell cytotoxicity (ADCC).[28,29]

The specificities of several IgG and IgM mAbs have been previously studied based on the screening against a small panel of synthetic glycoconjugates[25,30] or applying shotgun arrays[31] and were generally divided into four groups, based on the recognition either to

Lewis X, LDN, LDN-DF or both LDNF and LDNDF motifs. Here we will to study the specificities of mAbs in the context of structures presented on our synthetic array, taking a closer insight into the structural requirements of antibodies to bind.

We have analyzed the supernatants of hybridoma culture media containing mAbs. First, the media was diluted in PBS to be subsequently incubated on the glycan array at room temperature during one hour. Slides were washed and subsequently incubated with anti-mouse IgG-647 and anti-mouse IgM-555 (1:1000 dilution) for one hour in the dark. Arrays were washed from unbound antibodies, dried and fluorescence was analyzed in a microarray scanner.

The majority of antibodies bound only to the tri- or tetrasaccharide glycan fragments, Lewis X, LDN and LDNF but with subtle differences regarding their presentation (Figure 4-5). Five of the mAbs, including 291-2G3, 128-4F9, 114-4D12, 291-4D10 and 99-163[25,30] were specific for Lewis X motif, three of which (291-2G3, 128-4F9, 114-4D12) were found to bind exclusively the monovalent trisaccharide, showing no recognition of Lewis X terminated N-glycans (Figure 4, A-C). 291-4D10 IgM recognized trisaccharide Lewis X in its monovalent form and within the context of the oligosascharides only on the arm 6 of N-glycans (Figure 4, D). In contrast, 99-1G3 IgM recognized Lewis X motif exclusively when presented on the arm-3 of truncated N-glycans lacking α-1,6-mannose on the pentasaccharide core. (Figure 4, E).

MAbs 114-4E8 IgG3[32] and 290-2E6IgM had been previously characterized as anti-LDNF antibodies. We could observe binding of both antibodies against LDNF (Figure 5 A, B). However, due to the lack of N-glycans presenting this epitope on the antennae, it was not possible to determine if the presentation of LDNF in the context of larger oligosaccharide would improve binding. MAbs 259-2A1 IgG3[25] and 114-3A5 IgG3 had been previously described as LDN binding antibodies. We could observe that both mAbs bound the simple disaccharide fragment (Figure 5 C, D). While 259-2A1 IgG3 exclusively recognized simple LDN structure, 114-3A5 IgG3 also showed binding towards N-glycans containing terminal

LDN on truncated N-glycans with LDN on arm-3, on hybrid N-glycans with LDN on arm-3 and on biantennary LDN terminated compounds.

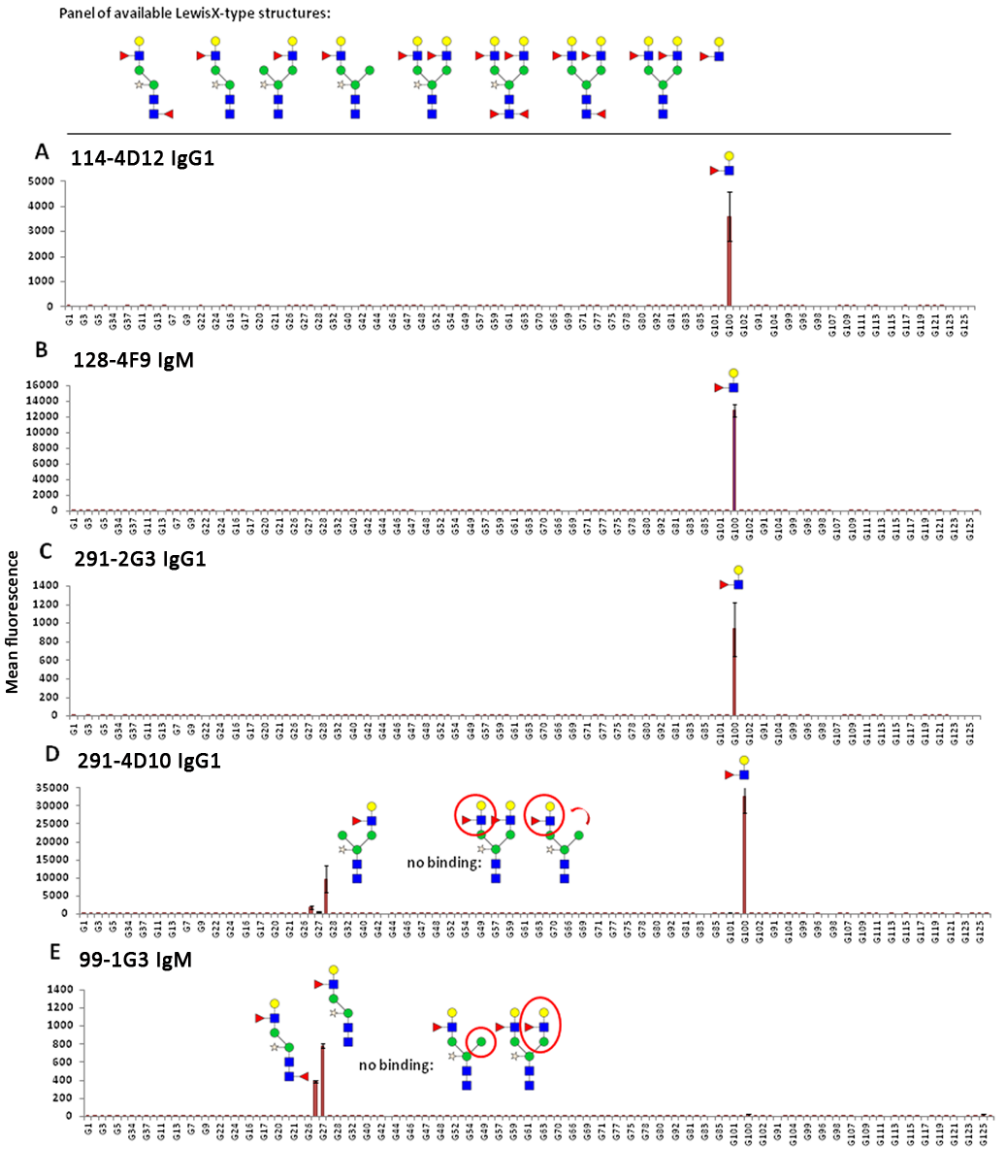


Figure 4. Glycan array screening of five monoclonal antibodies, showing carbohydrate specificity towards Lewis X. **A-E.** Fluorescence intensities after incubation with mAbs followed by anti-IgG-647/anti-IgM-555. Each histogram represents the average RFU values for four replicates (mean±SD).

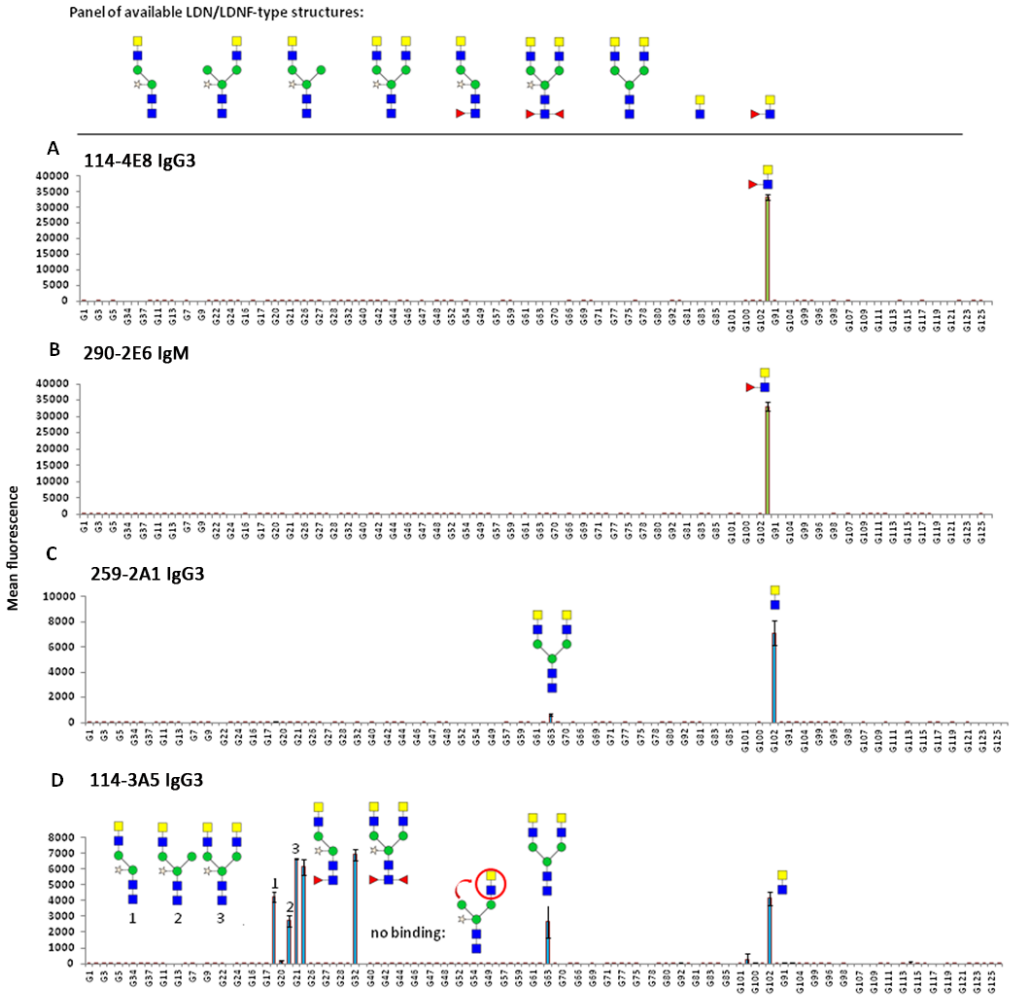


Figure 5. Glycan array screening of four monoclonal antibodies, showing carbohydrate specificity towards LDN or LDNF. **A-D**. Fluorescence intensities after incubation with mAbs followed by anti-IgG-647/anti-IgM-555. Each histogram represents the average RFU values for four replicates (mean±SD).

Lewis X, as well as LDN motif either with (LDNF) or without fucose are major antigens responsible for anti-carbohydrate immune response upon *S. mansoni* infection, although both motifs are also found to a minor extent in mammalian glycome. The consequences of the cross-reactivity for schistosomal induced antibodies with host structures are still not fully understood but could be part of a strategy to evade recognition by molecular

mimickry.[33] Additionally, we could observe that the binding of several mAbs depended not only on the presence of particular recognition element but also on its structural context. These results are supported by the fact that *S. mansoni* expresses different structural forms of Lewis X or LDNF motifs during its development stages and which result in induction of context-specific antibodies against structurally related but not identical presentations of these elements. For example, in *Schistosoma* glycome Lewis X can be present as a monomeric but also oligo- or polymeric structures, it can decorate glycoproteins both through N-linkage or O-linkage and it could also decorate glycolipids.[20] Monoclonal antibodies which are specific exclusively against simple trisaccharide Lewis X might suggest the specificity towards structures expressed for example in O-glycans or glycolipids. LDN and LDNF-based structures are present in all life stages of *S. mansoni* including cercarial O-glycocalyx, the surface of schistosomula, adult worms and the egg-shell and sub-shell areas.[23] Fucosylated and non-fucosylated LN and LDN motifs can be part of parasite N- or O-glycans or glycolipids.

Finally, we have characterized the specificity of 100-4G11 IgM monoclonal antibody[34] (Figure 6) which has been described to cross-react with *S. mansoni* and allergens present in plant and insect glycoproteins. It has been demonstrated that 100-4G11 IgM is able to recognize horseradish peroxidase (HRP), a plant glycoprotein and phospholipase A2 (PLA2) from honeybee venom. The most abundant N-glycans which are shared by both glycoproteins and parasite glycans are classified as paucimannose N-glycans modified with core α-1,3-Fuc/α-1,6-Fuc, but in contrast to HRP (Figure 6B, **G37**), PLA2 does not present β-1,2-xylosylation (Figure 6B, **G73**, **G75**).

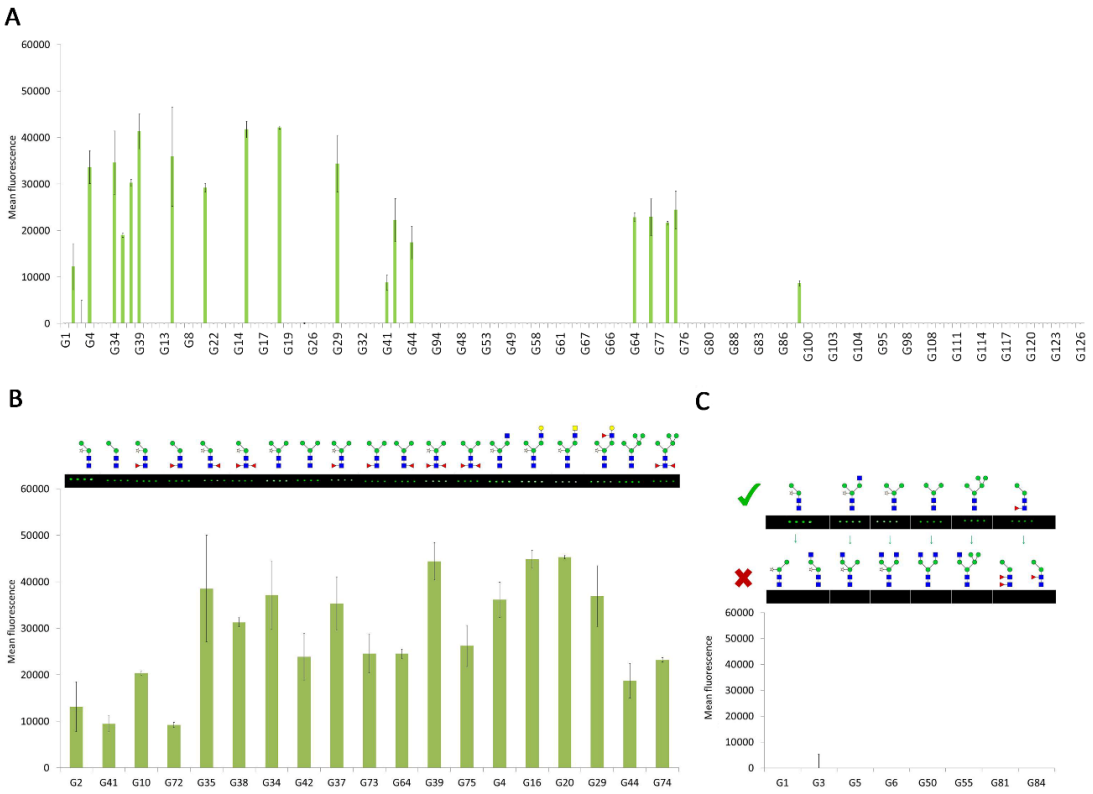


Figure 6. Anti-carbohydrate specificity of 100-4G11 IgM monoclonal antibody; **A.** Fluorescence intensities after incubation of glycan array with mAb 100-4G11/anti-IGM-555. Each histogram represents the average RFU values for four replicates (mean±SD). **B.** Fluorescence images and quantification of glycans recognized by 100-4G11 IgM; **C.** Comparison between structurally similar N-glycans which are recognized (green tick) and not (red cross) by 100-4G11 IgM.

We have screening the anti-carbohydrate specificity of 100-4G11 IgM mAb on our glycan array. Our results strongly suggested that 100-4G11 IgM mAb recognizes paucimannose N-glycans as reported before[34], but also binds to other structurally similar N-glycans carrying additional core modifications (Figure 6, B). Recognition of certain hybrid compounds (Figure 6B; **G4**, **G16**, **G20**, **G29**) was comparable to paucimannose mannose structures. On the other hand, analogues isomeric compounds **G5**, **G55** (Figure 6C) and biantennary complex type N-glycans **G6**, **G50** were not recognized by the mAb. Those data demonstrate that 100-

4G11 antibody requires a free terminal α-1,3-linked mannose, but it can also tolerate different substitutions on the 6-arm. Based on glycan array results, the minimum epitope required for antibody binding consist of Manα-1,3-Manβ-1,4-GlcNAc-β-1,4-GlcNAc epitope and α1-6-linked mannose is not required as suggested in earlier study[34] (Figure 6B, compounds **G2**, **G41**, **G10**, **G72**, **G35**, **G38**). The presence of a core α-1,3-Fuc and α-1,6-Fuc linked to the proximal core GlcNAc as well as β-1,2-xylose does not affect the binding. The presence of core α-1,3-linked fucose in the distal GlcNAc of the chitobiose core completely abolishes binding of mAb (Figure 6C; **G81**, **G84**). In Table 1, a summary of anti-carbohydrate mAbs specificities established during our glycan array study and previous reported data is compiled.

Entry	MAb	Isotype	Reported ligand	Glycan array
1	291-2G3[25]	IgG1	Lewis X	Simple Le^X
2	291-4D10[25]	IgM	Lewis X	Simple Le^X, Le^X (arm-6)
3	128-4F9[25]	IgM	Lewis X	Simple Le^X
4	99-1G3[25]	IgM	Lewis X	Le^X (arm-3)
5	114-4D12[30]	IgG1	DF-LDN-DF	Simple Le^X
6	114-4E8[32]	IgG3	LDNF	Simple LDNF
7	290-2E6[25]	IgM	LDN-DF+ LDNF	Simple LDNF
8	114-3A5	IgG3	Unknown?	Simple LDN, LDN (arm-3), G0
9	259-2A1[25]	IgG3	LDN	Simple LDN
10	100-4G11[34]	IgM	$Man_3GlcNAc$	Manα-1,3-Manβ1,4-GlcNAc

Table 1. Binding specificity of monoclonal antibodies against *S. mansoni*. Comparison between reported data and results obtained employing our synthetic N-glycan array.

In conclusion, we have evaluated the utility of our synthetic N-glycan array for the characterization of mAbs generated in serum of *S. mansoni* infected mice. Ten of the seventeen analyzed mAbs found to recognize specific elements presented on the glycan array, including Lewis X, LDN, LDNF and paucimannose N-glycan structures. The remaining seven mAbs did not bind to any structure on the arrays, suggesting a different specificity of these. The reactivity of monoclonal antibodies towards ligands presented on the glycan array directly confirmed that these carbohydrate elements are present in *S. mansoni* glycome and that during the course of infection, they are presented to the immune system leading to the generation of anti-carbohydrate antibodies. It is important to highlighted that the use of glycan arrays with large numbers of glycans showing very subtle structural differences can contribute to the definition of a precise sugar recognition epitope and help to understand the impact of the structural context in which the carbohydrate epitope is presented.

More importantly, having in hands structures found not only in helminths but also being shared with food allergens (cross-reactive antigens), we could potentially apply our glycan array as a tool for example in allergy diagnosis. Recently our glycan arrays attracted attention of the Yazdanbakhsh group at Leiden Medical Center and could be possible used for a project discussing link between helminths infection and protection against peanut allergies in the Ghanese population. It has been reported[35,36], that in Europe and North America the incidence of peanut allergy has increased recently, while the prevalence is much lower in many developing countries. This phenomenon was linked to the “hygiene hypothesis” and to the much higher prevalence of parasitic infections in developing countries.[27] There are several proposed mechanisms by which helminths are believed to protect against allergies, including the activation of regulatory mechanisms, the modulation of innate immune cells as well as cross-reactivity between parasite antigens and allergens and their role in IgE sensitization. The array screening of possible cross-reactive antigens will potentially help in characterization and identification of glycan structures being the main players in the helminth infections and allergies.

3.3. Identification of anti-carbohydrate antibodies in sera of *S. mansoni* infected patients

In this section, we will describe the screening of sera from *S. mansoni* infected individuals living in the endemic area of Piida fishing community at Butiaba, along Lake Albert in Uganda employing our glycan arrays. Pooled serum samples from children (age 5 to 11) and adults (age 26 to 46)[37] were analyzed. We measured total IgG and IgM polyclonal antibody responses and we have compared the results between both age groups. In addition, sera from three healthy European adults were analyzed and used as negative controls.

3.3.1 Natural antibodies of healthy individuals

We began the serum antibody analysis with the sera from three healthy individuals employed as negative controls in our schistosomiasis study. Total IgG and IgM glycan binding profiles were recorded and compared for each donor.

The majority of antibodies in human sera are the result of natural immunizations.[38] However, during the life-time healthy individuals acquire a variety of so-called natural antibodies (nAbs) many of which are directed towards carbohydrate antigens present in normal, healthy tissues (xeno-autoantigens).[39] Natural antibodies are present in the sera of all individuals in the absence of an antigenic stimulus and they form part of the innate immune response. The best known and studied nAbs are against blood group antigens, group A (GalNAcα-1,3(Fucα-1,2)Gal) and group B (Galα-1,3(Fucα-1,2)Gal), α-Gal-epitope (Galα-1,3-Gal) first identified by Galili *et. al* and Hanganutziu-Deicher (H-D) antigen (Neu5Gcα2-3Galβ1-4Glc)(Figure 7).[38–41] The natural blood-group specific xenoactive antibodies are responsible for agglutination due to mismatched blood transfusions while anti-Galα-1,3-Gal antibodies are the major obstacle associated with pig-to-human organ transplants. Unlike humans, non-primate mammals and new world monkey possess in their tissues the Galα-1,3-Gal epitope and lack natural antibodies directed against it. Thus the transplant of organs from these species to human can caused hyperacute rejection of

transplanted tissues.[41,42] On the other hand, anti-Neu5Gc antibodies can be also found in humans, however, their levels strongly depend on the individuals. Neu5Gc can be metabolically incorporated into tissues from dietary sources, like red meat by certain commensal bacteria and may lead to generation of antibodies which are cross-reactive against Neu5Gc-containing glycans in healthy human tissue.[43] Experiments performed on the human-like Neu5Gc-deficient *Cmah*$^{-/-}$ mouse model have demonstrated that the inflammation caused by the antigen-antibody interactions can promote tumor progression, suggesting a link between red meat consumption and carcinoma risk.

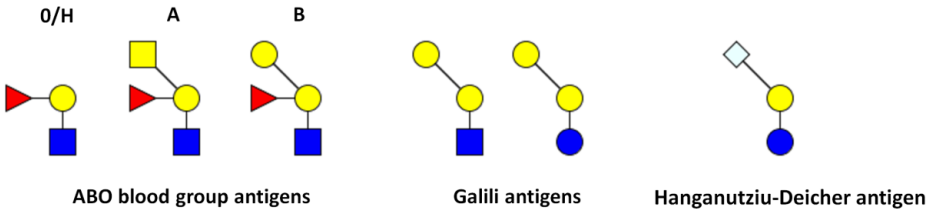


Figure 7. Blood type and Galili/isoglobo/xeno antigens of healthy human sera.

In our experiment, sera from healthy individuals were diluted in PBS (containing 1% BSA and 0.01% Tween-20) and incubated on glycan array at room temperature for 1h. Slides were washed and subsequently incubated with anti-human IgG(Fc specific)-Cy3 and anti-human IgM(μ specific)-Alexa fluor 647 simultaniously for 30 minutes in the dark. Arrays were washed from unbound secondary antibodies, dried and fluorescence emission was measured in a microarray scanner. The anti-carbohydrate IgG responses were showing generally higher fluorescence intensities (Figure 8, A) and were much more specific compared to IgM patterns (Figure 8, B). Overall IgG response recorded for each donor created a unique anti-carbohydrate "finger print" for a single, analyzed person.

The general binding tendencies were in a good agreement with the literature[39,40], the profile of human anti-glycan antibodies detects α-Gal-related structures, blood group antigens ABH (variously represented in different individuals, mainly IgM response), Lewis

and related structures (IgM response) and smaller O-type glycans. The anti-carbohydrate IgG responses of healthy donor 1 and 2 were directed towards Galili-type of structure[38] **G112** and/or **G113** with donor 2 exhibited additional responses towards Galα-1,2Gal disaccharide **G107**, blood group antigen H6 **G116** and human milk 3'sialyllactose **G120** (Neu5Acα2-3Galβ1-4Glc). The natural antibodies towards structure **G107** as well as H-type blood antigens **G116** were also observed by Bovin and coworkers, who employed glycan array to profile blood of healthy individuals and reported the presence of these common antibodies in analyzed sera.[39]

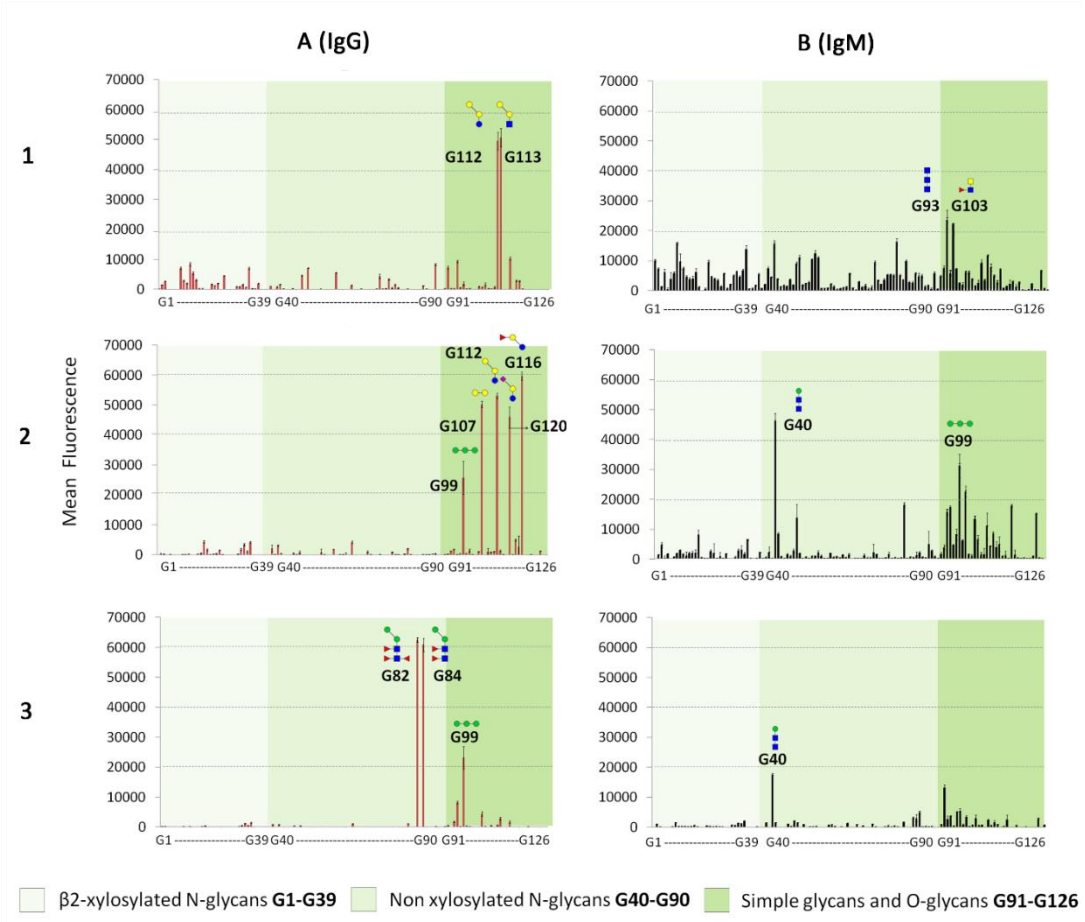


Figure 8. Anti-carbohydrates IgG (A) and IgM (B) responses in three (1-3) healthy individuals.

Additionally, individuals 2 and 3 possess IgG antibody responses towards trimannosyl structure **G107** (Manα-1,2Manα-1,2Man) and **G82** and **G84** N-glycans (Figure 8, 1-2). The Manα-1,2Man linkages are present in the branches of the pathogenic fungi[44] and in the arm 3 of Man9 high mannose N-glycans while core α-1,3 bisfucosylation is a characteristic feature of certain nematodes glycans[45]. Antibodies directed towards these sugars might suggest possible recent infection-related disease of these healthy donors.

As mentioned before, the anti-carbohydrate IgM responses were of much lower intensities and are less specific. As most of the natural antibodies directed against blood group antigens are of the IgM class it could suggest that some of them could bind to these structures on the array. Additionally, smaller O-glycans, including chitobiose-type structure **G40** and **G93** were recognized by anti-glycan IgM antibodies.

Summarizing, natural antibodies against several self-glycans as well as carbohydrate antigens encounters during the human life-time are present in the sera of every healthy individual. More importantly for our study, healthy donors with no previous history of the *S. mansoni* infection, do not elicit antibodies against any of N-glycans present on the array (with exception to compounds **G82**, **G84**) and can be further used as the negative controls for the screening of the sera of *Schistosoma* infected patients. In the next study, we used the pooled sera of healthy individuals, and compared the anti-carbohydrate responses directed against N-glycans excluding glycan fragments **G91**-**G126**.

3.3.2 Glycan binding analysis of anti-carbohydrate antibodies from sera of *S. mansoni* infected children and adults

Shotgun microarrays covering structures from the entire *S. mansoni* glycome, including natural N-glycans and glycolipids harvested at different parasite life stages are key tools to match antibody responses to parasite glycan structures. Recent studies revealed a differential glycan recognition profiles characteristic for two different age group, children and adults.[13] However, due to the nature of the shotgun array, it is difficult to identify the

responsible antigen from the mixture of the printed fraction. Here, we used our synthetic glycan array as a complementary tool, to evaluate whether an age-related differentiation of sera can be defined on the basis of antibody reactivity to our array of parasitic N-glycans. Sera were diluted 1:100 with PBS and incubations were performed as previously described in section 3.3.1 (Figure 10).

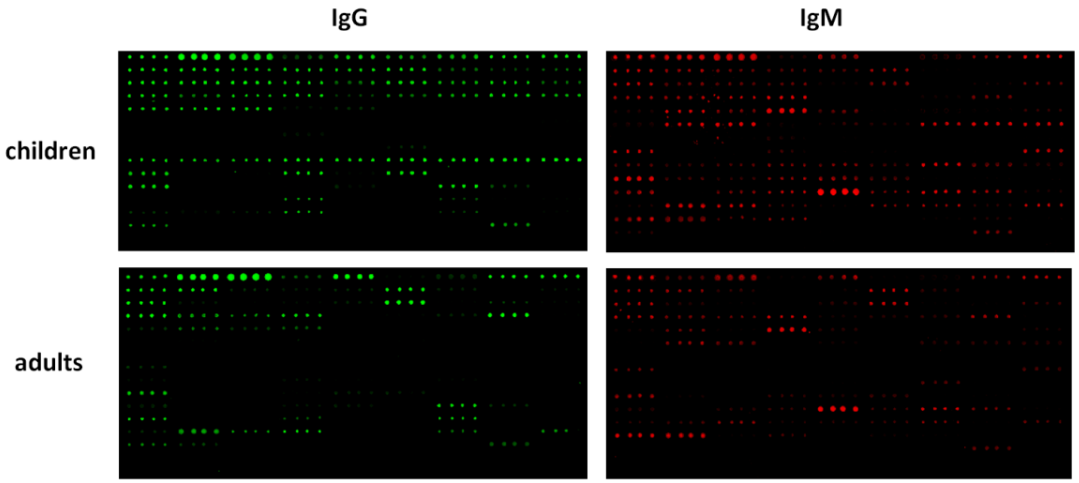


Figure 10. Fluorescence read-out of anti-carbohydrate IgG (green) and IgM (red) responses in sera of Schistosoma infected children and adults probed with anti-human IgG(Fc specific)-Cy3 and anti-human IgM(μ specific)-Alexa fluor 647 antibodies respectively.

The IgM responses to the glycans on our array (Figure 11) were of weak to moderate intensity and of low specificity in both age groups. Generally, we observed a higher anti-carbohydrate response for children than for adults but no significant differences in bound structures could be detected between the age-groups.

Figure 11. Glycan array screening of schistosomiasis infected sera followed by incubation with anti human-IgM-647; **A.** IgM from pooled sera of children (green); **B.** adults (yellow); **C.**healthy individuals (red). Error bars show the standard deviation of mean fluorescence for 4 replicate spots quantified for each ligand.

For IgG binding, however, we found a more pronounced binding pattern and a strong differentiation between the two age groups based on core β-1,2 xylosylated and core α-1,3 fucosylated glycan structures (Figure 12). The major children derived anti-carbohydrate IgG antibody response was found to be directed towards both, core β-1,2-xylosylated and core α-1,3-fucosylated structures of high mannose, hybrid and complex type N-glycans. Interestingly, the pooled adult IgG antibodies only bound to certain core β-1,2-xylosylated compounds, but not to core α-1,3-fucosylated structures lacking xylose. Importantly, the glycans recognized by either the adults or children were bound at only very low levels by serum antibodies from individuals uninfected with *S. mansoni* (Figure 11, 12 C) suggesting a specific response to *S.mansoni* infection. Structures which were truncated on the 3-arm and presenting substituents no larger than a single α-1,3 mannose residue were best bound both for children and adults. Core-xylosylated structures presenting more sterically demanding substituents on the 3-arm were still bound by the children serum pool, however, with lower intensity. In adults no antibodies against these glycans were detected.

Altogether, we were able to detect with our array an important number of anti-carbohydrate Abs in the blood of *S. mansoni* infected patients as demonstrated above, their protective potential or specific function, however, remains unclear.

Since the parasite has to developed mechanisms to evade or resist the immune response of the host in order to remain in the host for a long period of time, certain carbohydrates may down regulate immune responses in the chronically infected patients and act as a "smoke screen" subverting the immune system away from the epitopes that could provoke protective immune responses.[20] It has been demonstrated that children are generally far more susceptible to infection and re-infection than adults, even if in the area of study adults are much more exposed to the parasite. It has been shown that the acquisition of immunity against parasites is age-dependent and require maturation of the immune system by multiple infections and treatments.[20] This might suggest that certain parasitic glycans which are recognized by the children immune system during the infection are ignored by adult immune system that had been exposed to the parasite previously.

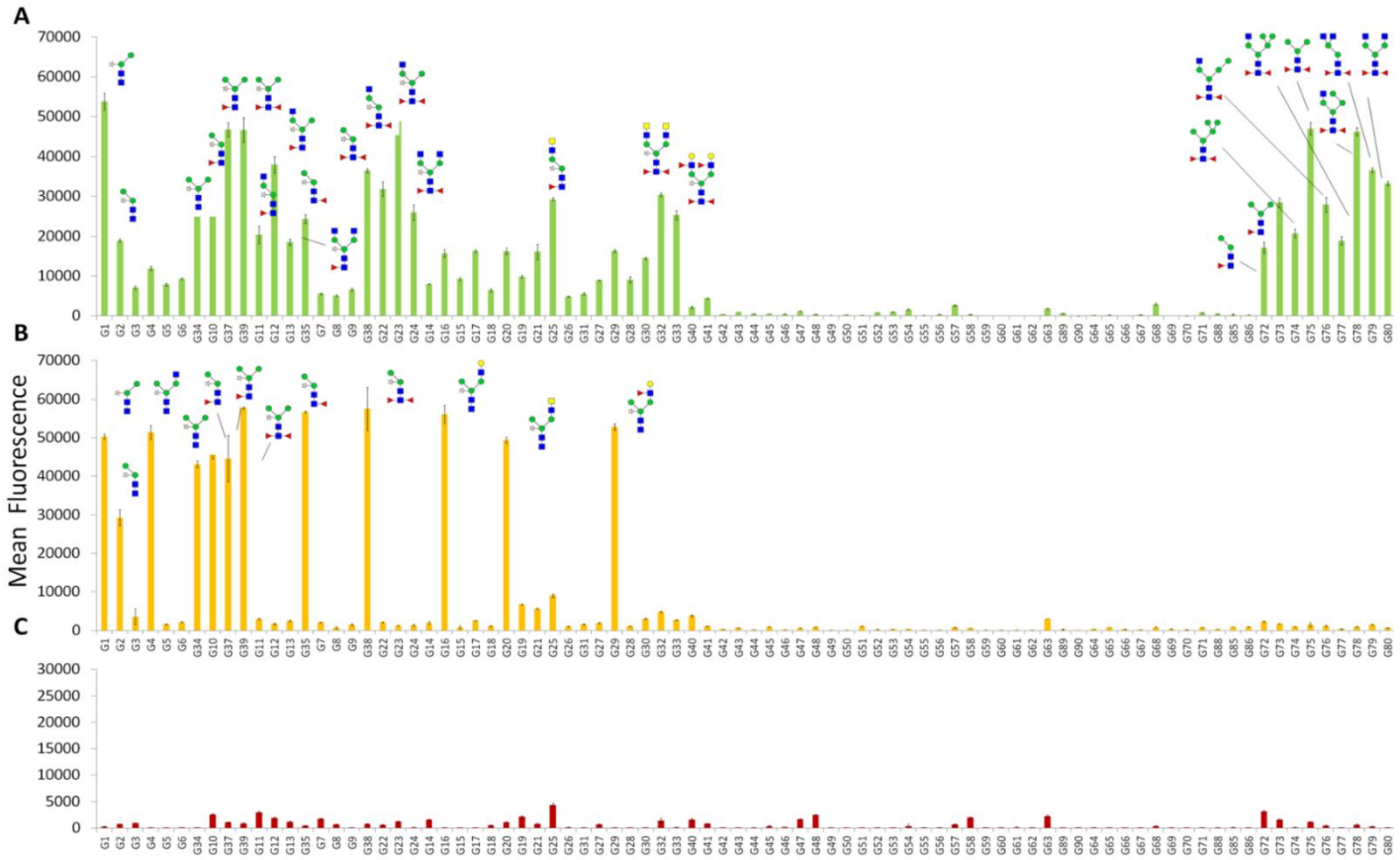


Figure 12. Glycan array screening of schistosomiasis infected sera followed by incubation with anti human-IgG--(Fc specific)-Cy3; **A.** IgG from pooled sera of children (green); **B.** adults (yellow); **C.**healthy individuals (red). Error bars show the standard deviation of mean fluorescence for 4 replicate spots quantified for each ligand.

For the selection of glycan structures as antigens of a glycoconjugate vaccine one should also consider the specificity of the encountered hits in the glycan array screen for the targeted species. A further interpretation of the serum antibody binding profiles to identify structures as potential vaccine candidates or for specific diagnostic purposes has to be made with caution, as core β-1,2-xylosylated and core α-1,3 fucosylated glycans are not unique in *S. mansoni* glycome but they are also present on several other helminths parasites, plants and insects. Nevertheless, the fact that we see strong IgG binding to core-xylosylated glycans in schistosomiasis sera and not in controls indicates that this response is indeed directed against the parasite. A protective function of the elicited antibodies can, however, not be determined by glycan array screening alone. We require additional immunological studies and an immunization trial with putative glycan conjugate vaccine candidates in a relevant animal model. Additionally, a protective immune response against *S. mansoni* may be formed by the combined action of multiple antibodies against various glycans, glycoproteins and/or glycolipids rather than against a single antigen.

3.4 Summary

This chapter described the successful application of the synthetic glycan arrays towards screening and characterization of anti-carbohydrate antibodies. We have probed a group of monoclonal antibodies induced in *S. mansoni* infected mice and evaluated their specificity towards structures present on the microarray. Subsequently, we have probed the sera of healthy Europeans and recorded their unique IgM and IgG binding profiles which were directed mostly towards blood group and possible environmental antigens. Finally, we have compared the IgG and IgM antibody responses in serum samples in *S. mansoni* infected individuals divided into two age-groups, children and adults. The overall IgM response patterns were similar for children and adults, but showed higher intensity for younger age group. For the IgG binding profile, we found a more pronounced binding pattern and an important age-related differentiation. These difference appear to be based on core β-1,2

xylosylation and core α-1,3 fucosylation as major denominators. This observation highlights the antigenic nature of both these motifs as well as the value of our synthetic core-xylosylated structures to detect specific antibodies in different serum cohorts. However, if this glycan structures might induce protective immunity during the parasitic infections cannot be still clearly evaluated. These glycan array screening of serum is ongoing and it will be focused in classifying the immunoglobulin isotype and subclass. This will help to clarify the functional role of the detected antibodies.

3.5 Experimental Part

General methods

Materials: Bovine serum albumin BSA (globulin free) was purchased from Sigma Aldrich. Anti-mouse IgG (H+L specific) Alexa fluor 647 produced in rabbit, and anti-mouse IgM-Alexa fluor 555 (μ specific) produced in goat, were purchased from Invitrogen. Secondary anti-mouse antibodies were used in 1:1000 dilution in PBS containing 1% BSA and 0.01% Tween-20. Anti-human IgG (Fc specific)-Cy3, produced in goat, was purchased from Sigma-Aldrich and Anti-human IgM-Alexa fluor 647 (μ specific) was purchased from Invitrogen, Breda, the Netherlands. Anti-human antibodies were used in 1:1000 dilution dilution in PBS containing 1% BSA and 0.01% Tween-20.
Glycan arrays (36x18) were prepared and fluorescence was analyzed as indicated in Chapter 2.

Mouse mAbs reactive to schistosome glycans were provided by Dr. C.H Hokke from the Leiden Parasite Glycobiology Group (Leiden University Medical Center, the Netherlands). They were generated from *S. mansoni* infected mice as described previously[25]. Human sera were obtained from 41 *Schistosoma mansoni* infected individuals living in the Piida community, Butiaba in Uganda where *S. mansoni* is endemic. Ethical approval for the Piida

study was obtained from the Uganda National Council for Science and Technology (UNCST). Sera were made available by Dr. Dunne, Cambridge University, United Kingdom and by N. Kabatereine, Vector Control Division, Uganda Ministry of Health, Kampala.

Incubation of *S. mansoni* specific monoclonal antibodies

Culture supernatants containing mAbs were diluted 1:50 in PBS containing 1% BSA and 0.01% Tween-20, and aliquots (180 µL per array) were used to incubate individual wells on a glycan array slide at room temperature for 1h. The slides were washed with PBST (PBS containing 0.01% Tween-20) followed by PBS. Slides were then incubated with anti-mouse IgG-647 and anti-mouse IgM-555 (1:1000) in PBS containing 1% BSA and 0.01% Tween-20 for 1h in the dark. Arrays were washed from unbound secondary antibodies with PBST, PBS and water. Microarrays were dried and scanned for the fluorescence emission in a microarray scanner.

Incubation of pooled serum samples from *S. mansoni* infected patients

Previously collected selected sera of schistosomiasis patients were pooled into two age-groups: children (21 individuals, age between 5 to 11 years) and adults (20 individuals, age between 20 to 46 years). In both groups,the ratio female/male ratio was 3:2. In addition, sera from 3 healthy Europeans was analyzed and used as negative control. Sera were diluted 1:100 with PBS containing 1% BSA and 0.01% Tween-20, and aliquots (180 µL per array) used to incubate individual wells on a glycan array slide at room temperature for 1h. The slides were washed with PBST (PBS containing 0.05% Tween-20) followed by PBS. Slides were then incubated with anti-human IgG(Fc specific)-Cy3 (Figure 12) and anti human-IgM(µ specific)-Alexa fluor 647 (Figure 11) in PBS containing 1% BSA and 0.01% Tween-20 for 30 minutes in the dark. Slides were washed subsequently with PBST, PBS and H_2O and dried in a slide spinner prior to scanning the fluorescence emission in a microarray scanner.

References

(1) Oyelaran, O.; Gildersleeve, J. C. *Expert review of vaccines* **2007**, *6*,957-969.

(2) Paulson, J. C.; Blixt, O.; Collins, B. E. *Nature chemical biology* **2006**, *2*, 238–248.

(3) Park, S.; Lee, M.-R.; Shin, I. *Chemical Communications* **2008**, *37*, 4389–4399.

(4) Liang, P.-H.; Wu, C.-Y.; Greenberg, W. A.; Wong, C.-H. *Current opinion in chemical biology* **2008**, *12*, 86–92.

(5) Wang, D.; Carroll, G. T.; Turro, N. J.; Koberstein, J. T.; Kovác, P.; Saksena, R.; Adamo, R.; Herzenberg, L. A.; Herzenberg, L. A.; Steinman, L. *Proteomics* **2007**, *7*, 180–184.

(6) Scanlan, C. N.; Pantophlet, R.; Wormald, M. R.; Saphire, E. O.; Stanfield, R.; Wilson, I. A.; Katinger, H.; Dwek, R. A.; Rudd, P. M.; Burton, D. R. *Journal of virology* **2002**, *76*, 7306–7321.

(7) Calarese, D. A.; Lee, H.-K.; Huang, C.-Y.; Best, M. D.; Astronomo, R. D.; Stanfield, R. L.; Katinger, H.; Burton, D. R.; Wong, C.-H.; Wilson, I. A. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 13372–13377.

(8) Burton, D. R.; Pyati, J.; Koduri, R.; Sharp, S. J.; Thornton, G. B.; Parren, P.; Sawyer, L.; Hendry, R. M.; Dunlop, N.; Nara, P. L.. *Science* **1994**, *266*, 1024–1027.

(9) Mascola, J. R.; Stiegler, G.; VanCott, T. C.; Katinger, H.; Carpenter, C. B.; Hanson, C. E.; Beary, H.; Hayes, D.; Frankel, S. S.; Birx, D. L.. *Nature medicine* **2000**, *6*, 207–210.

(10) Shivatare, S. S.; Chang, S.-H.; Tsai, T.-I.; Ren, C.-T.; Chuang, H.-Y.; Hsu, L.; Lin, C.-W.; Li, S.-T.; Wu, C.-Y.; Wong, C.-H. *Journal of the American Chemical Society* **2013**, *135*, 15382–15391.

(11) Tuccillo, F. M.; de Laurentiis, A.; Palmieri, C.; Fiume, G.; Bonelli, P.; Borrelli, A.; Tassone, P.; Scala, I.; Buonaguro, F. M.; Quinto, I.; others. *BioMed research international* **2014**, *2014*, 742831-742844.

(12) Huang, C.-Y.; Thayer, D. A.; Chang, A. Y.; Best, M. D.; Hoffmann, J.; Head, S.; Wong, C.-H. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 15–20.

(13) Van Diepen, A.; Smit, C. H.; van Egmond, L.; Kabatereine, N. B.; Pinot de Moira, A.; Dunne, D. W.; Hokke, C. H. *PLoS Neglected Tropical Diseases* **2012**, *6*, 1922-1931.

(14) De Boer, A. R.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Glycoconjugate journal* **2008**, *25*, 75–84.

(15) Aragon, A. D.; Imani, R. A.; Blackburn, V. R.; Cupit, P. M.; Melman, S. D.; Goronga, T.; Webb, T.; Loker, E. S.; Cunningham, C. *Molecular and biochemical parasitology* **2009**, *164*, 57–65.

(16) Faveeuw, C.; Mallevaey, T.; Paschinger, K.; Wilson, I. B. H.; Fontaine, J.; Mollicone, R.; Oriol, R.; Altmann, F.; Lerouge, P.; Capron, M.; others. *European journal of immunology* **2003**, *33*, 1271–1281.

(17) Jurberg, A. D.; Brindley, P. J. *Frontiers in genetics* **2015**, *6*, 1-156.

(18) Gryseels, B.; Polman, K.; Clerinx, J.; Kestens, L. *The Lancet* **2006**, *368*, 1106–1118.

(19) Pearce, E. J.; MacDonald, A. S.; others. *Nature Reviews Immunology* **2002**, *2*, 499–511.

(20) Van Diepen, A.; van der Velden, N. S.; Smit, C. H.; Meevissen, M. H.; Hokke, C. H. *Parasitology* **2012**, *139*, 1219–1230.

(21) Hokke, C. H.; Deelder, A. M.; Hoffmann, K. F.; Wuhrer, M. *Experimental parasitology* **2007**, *117*, 275–283.

(22) Khoo, K.-H.; Chatterjee, D.; Caulfield, J. P.; Morris, H. R.; Dell, A. *Glycobiology* **1997**, *7*, 663–677.

(23) Wuhrer, M.; Koeleman, C. A. M.; Deelder, A. M.; Hokke, C. H. *FEBS Journal* **2006**, *273*, 347–361.

(24) Kariuki, T.; Farah, I.; Wilson, R.; Coulson, P. *Parasite immunology* **2008**, *30*, 554–562.

(25) Van Remoortere, A.; Hokke, C. H.; van Dam, G. J.; van Die, I.; Deelder, A. M.; van den Eijnden, D. H. *Glycobiology* **2000**, *10*, 601–609.

(26) Alkan, S. S. *Nature Reviews Immunology* **2004**, *4*, 153–156.

(27) Scott, A. M.; Wolchok, J. D.; Old, L. J. *Nature Reviews Cancer* **2012**, *12*, 278–287.

(28) Zabriskie, J. B. *Essential clinical immunology*; Cambridge University Press:UK, 2009.

(29) Abbas, A. K.; Lichtman, A. H.; Pillai, S. *Cellular and Molecular Immunology*; Elsevier Health Sciences, 2014.

(30) Robijn, M. L.; Koeleman, C. A.; Wuhrer, M.; Royle, L.; Geyer, R.; Dwek, R. A.; Rudd, P. M.; Deelder, A. M.; Hokke, C. H. *Molecular and biochemical parasitology* **2007**, *151*, 148–161.

(31) De Boer, A. R.; Hokke, C. H.; Deelder, A. M.; Wuhrer, M. *Analytical chemistry* **2007**, *79*, 8107–8113.

(32) Van Diepen, A.; van der Plas, A.-J.; Kozak, R. P.; Royle, L.; Dunne, D. W.; Hokke, C. H. *International journal for parasitology* **2015**, *45*, 465–475.

(33) Van Die, I.; Cummings, R. D. *Glycobiology* **2010**, *20*, 2-12.

(34) Van Remoortere, A.; Bank, C. M.; Nyame, A. K.; Cummings, R. D.; Deelder, A. M.; van Die, I. *Glycobiology* **2003**, *13*, 217–225.

(35) Amoah, A. S.; Obeng, B. B.; Larbi, I. A.; Versteeg, S. A.; Aryeetey, Y.; Akkerdaas, J. H.; Zuidmeer, L.; Lidholm, J.; Fernández-Rivas, M.; Hartgers, F. C.; others. *Journal of Allergy and Clinical Immunology* **2013**, *132*, 639–647.

(36) Van der Kleij, D.; Tielens, A. G.; Yazdanbakhsh, M. *Infection and immunity* **1999**, *67*, 5946–5950.

(37) Naus, C.; Booth, M.; and Jones, F.M.; Kemijumbi, J.; Vennervald, B.J.; Kariuki, C.H.; Ouma, J.H.; Kabatereine, N.B.; Dunne, D.W. *Tropical Medicine & International Health* **2003**, *8*, 561-568.-

(38) Galili, U.; Rachmilewitz, E.; Peleg, A.; Flechner, I. *The Journal of experimental medicine* **1984**, *160*, 1519–1531.

(39) Huflejt, M. E.; Vuskovic, M.; Vasiliu, D.; Xu, H.; Obukhova, P.; Shilova, N.; Tuzikov, A.; Galanina, O.; Arun, B.; Lu, K.; Bovin, N. *Molecular Immunology* **2009**, *46*, 3037–3049.

(40) Obukhova, P.; Rieben, R.; Bovin, N. *Xenotransplantation* **2007**, *14*, 627–635.

(41) Wieslander, Jö.; Månsson, O.; Kallin, E.; Gabreilli, A.; Nowack, H.; Timpl, R. *Glycoconjugate Journal* **1990**, *7*, 85–100.

(42) Sandrin, M. S.; Vaughan, H. A.; Dabkowski, P. L.; McKenzie, I. *Proceedings of the National Academy of Sciences* **1993**, *90*, 11391–11395.

(43) Samraj, A. N.; Läubli, H.; Varki, N.; Varki, A. *Frontiers in oncology* **2014**, *4*, 1-13.

(44) Cambi, A.; Netea, M. G.; Mora-Montes, H. M.; Gow, N. A.; Hato, S. V.; Lowman, D. W.; Kullberg, B.-J.; Torensma, R.; Williams, D. L.; Figdor, C. G. *Journal of Biological Chemistry* **2008**, *283*, 20590–20599.

(45) Yan, S.; Serna, S.; Reichardt, N.-C.; Paschinger, K.; Wilson, I. B. *Journal of Biological Chemistry* **2013**, *288*, 21015–21028.

CHAPTER 4

GLYCAN-DIRECTED TARGETING OF DENDRITIC CELLS

4.1 Introduction

Targeting antigens to dendritic cell (DC) subsets is a promising strategy to enhance the efficacy of vaccines. DCs are highly specialized antigen-presenting cells (APCs) that constantly sample their environment in the search for foreign molecules.[1] Once DCs encounter an antigen, they are able to capture it, further process it and present antigen fragments to T cells (Figure 1). Since DCs stimulate T cell responses but are also capable of inducing tolerance, they are an attractive target for immunotherapy[2] and have been exploited to improve vaccine candidate efficacy.[3,4,5]

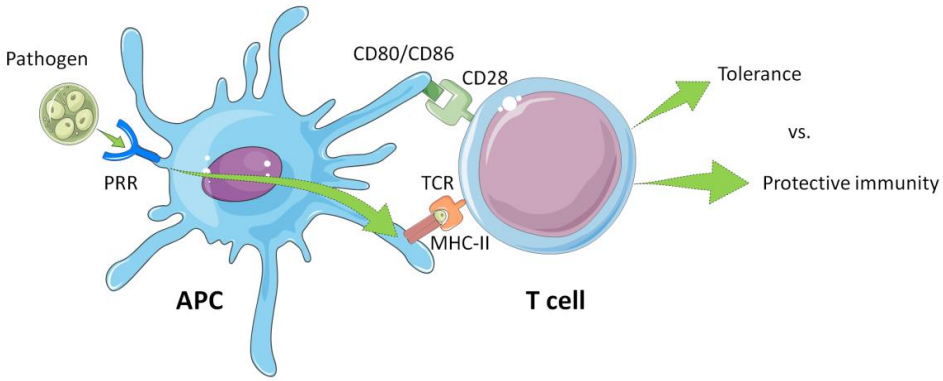


Figure 1. Dendritic cell recognizes the pathogen through the pattern recognizing receptor (PRR), uptake it, process and the fragments are presented by the major histocompatible complex (MHC-II) to the $CD4^{+}$T cells. Peptides presented by the MHC-II molecule are recognized by the T cells receptors (TCR) along with binding of the co-stimulatory molecules CD80 or CD86 by CD28.

Some DC-targeting vaccines are generated *ex vivo* using antigen-loaded monocyte-derived DCs that are further re-administered to the patient.[6,7] More direct methods involve the *in vivo* targeting of specific DC receptors, including C-type lectin receptors which were mentioned in Chapter 2. CLRs represent a large family of receptors specialized in the recognition of carbohydrates on pathogens and self-antigens through their extracellular carbohydrate recognition domain (CRD).[8,9,10] Since CLRs exhibit a differential expression on DC subsets and a number of them act as endocytic receptors, they are attractive candidates

for an efficient antigen targeting to DCs.[9,9] Still, most of the efforts towards CLR targeting utilize monoclonal antibodies[11] (for example, anti-DEC205[12,12], anti-mannose receptor mAbs) and far less examples are available for glycan-based CLR targeting.[13] However, glycan-based CLR targeting offers some advantages compared to antibody-mediated targeting, such as a low immunogenicity and the possibility to target several CLRs simultaneously. Furthermore, targeting CLRs with glycans resembles the natural CLR-ligand interactions and allows modulating binding strength by adjusting ligand densities and spatial orientation.[14] Finally, glycans show a potentially more favorable pharmacokinetic profile when compared to large protein drugs [15,15] and as described in Chapter 1, they can be prepared relatively fast and with good yields applying chemoenzymatic synthesis.[16,17] The inherently weak interactions between single carbohydrates and proteins can be substantially increased by a multivalent glycan presentation. To this end, nanoparticles, liposomes, and dendrimers have been used as different display systems for multivalent CLR targeting.[8]

The Chapter 2 described how the presence of core β-1,2-xylose disrupted the interaction of several N-glycans with mannose binding plant lectins such as *Galanthus nivalis* agglutinin (GNA) and *Concanavalin A* (ConA). More importantly, the presence of core xylose on the N-glycan **G6** completely abolished binding to the human Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin (DC-SIGN)[16], a CLR involved in the recognition of many pathogens such as HIV, *Schistosoma mansoni* or Ebola virus.[18,19,20] In this Chapter, based on the glycan array screening of DC-SIGN, we will describe the preparation of two glycoconjugates (G0-OVA and XG0-OVA) which were synthesize by conjugating the biantennary N-glycan (**G50**=**G0**) and its O-2 core xylosylated analogue N-glycan (**G6**=**XG0**) to the model protein antigen ovalbumin (OVA). We will analyze their binding to a set of murine CLR-Fc fusion proteins (including homologues of human DC-SIGN) using lectin microarray and further evaluate whether the differential binding of **G0** and **XG0** to CLRs will have an impact on dendritic cell targeting. We will perform uptake studies using Alexa Fluor® 647 labeled glycoconjugates and murine dendritic cells and finally, we will measure the effects of

ovalbumin glycoconjugates on T cell activation in a dendritic cell/T cell co-culture assay (Figure 2).

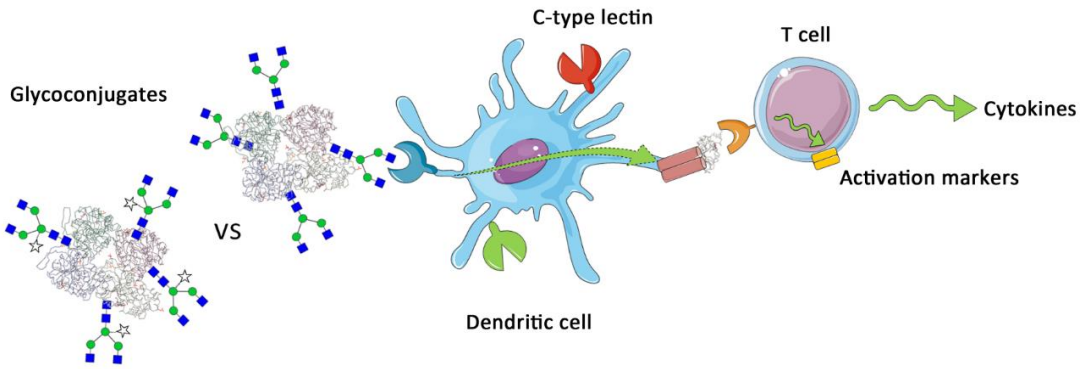


Figure 2. Schematic representation of the DCs/T cells co-culture assay applying two, chosen glycoconjugates, XG0-OVA and G0-OVA as a targeting molecule of DCs CLRs.

In the second part of the Chapter 4, inspired by the results obtained from the DCs and Alexa Fluor ® 647 labeled glycoconjugates uptake study, we will describe the preparation of an alternative to the organic dye labeled glycoproteins system, glycoprotein coated gold nanoclusters (AuNCs). AuNCs exhibit a size-dependent fluorescence and are considered to be a novel ultra-small, bio-labels.[21] Based on the progress in the macromolecule directed synthesis of clusters, we will use glycoconjugates for the generation of glycan-OVA-AuNCs. So created system will combine DCs targeting properties of glycan with antigenic properties of the protein and the imaging capability of fluorescent gold nanoclusters. An introduction describing gold nanomaterials in general illustrated with examples of synthetic strategies previously found in literature will be given. Finally, in the last part of the Chapter we will show a first example for the use of glycoprotein-based gold nanoclusters in DCs targeting.

4.2 Generation and characterization of ovalbumin glycoconjugates, XG0-OVA and G0-OVA

N-glycan structures (**G0** and **XG0**) equipped with a C5 amino linker for further conjugation were chemically synthesized as described in Chapter 1.[16,17] The corresponding OVA conjugates were generated by first activating glycan-amino linker with disuccinimidyl suberate ester (DSS) (Figure 3) and then crosslinking accessible OVA lysine residues in a second step.[22] The glycoconjugates G0-OVA and XG0-OVA were characterized by gel electrophoresis (Figure 3, B) and by MALDI-TOF mass spectrometry (Figure 3C).

The mass shift upon conjugation of OVA with activated glycans G0-DSS and XG0-DSS (Figure 3C) indicated an average conjugation of 3-4 glycans per OVA molecule. To evaluate the influence of core xylose on the recognition by CLRs and DC targeting, a similar glycan valence was required for both, XG0-OVA and G0-OVA. By carefully adjusting the glycan/OVA ratio and the protein concentration during the conjugation reaction, we were able to produce both OVA glycoconjugates displaying equal numbers of glycan moieties.

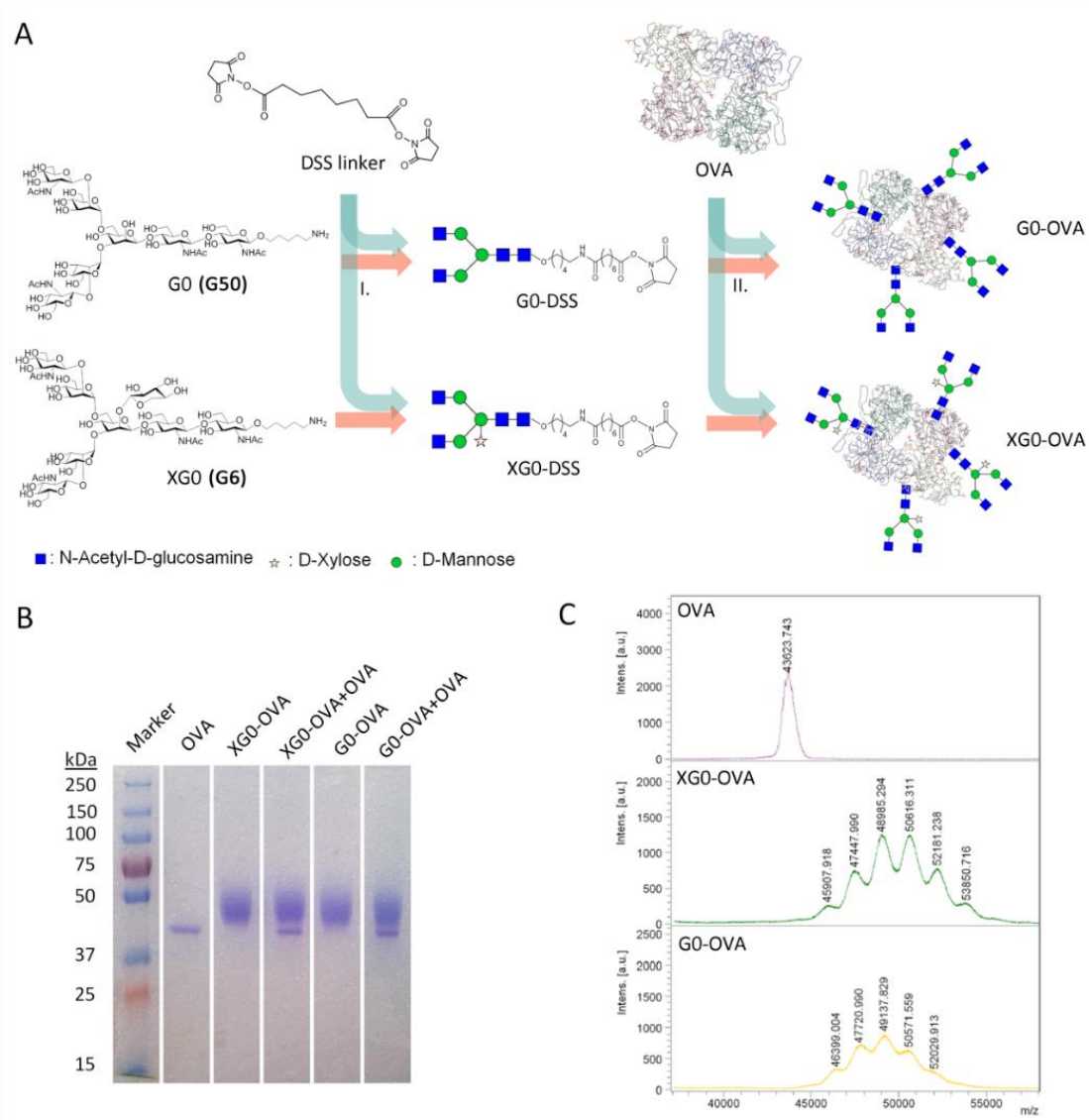


Figure 3. A. Synthesis of OVA glycoconjugates employing disuccinimidyl suberate ester (DSS) as a linker; **I.** mono-substituted NHS-activated suberate ester; **II.** Addition of activated N-glycan to OVA solution. Evaluation of the glycoconjugate synthesis; **B.** SDS-PAGE (10% acrylamide) stained with Coomassie Blue G-250; **C.** MALDI-TOF spectra of OVA (violet), XG0-OVA (green) and G0-OVA (yellow).

4.3 Binding of OVA neoglycoconjugates to CLRs printed on the array

To identify CLRs involved in the recognition of the glycoconjugates G0-OVA and XG0-OVA by murine DCs, we performed binding studies using lectin microarray technology. To this end, a library of CLR-Fc fusion proteins consisting of the extracellular part of the respective CLR and the Fc part of human IgG1 molecules was used.[23] We analyzed binding of G0-OVA and XG0-OVA towards eight recombinant murine CLRs: DC immunoactivating receptor (DCAR), DC immunoreceptor (DCIR), Dectin-2, Macrophage inducible C-type lectin (Mincle), Macrophage-restricted C-type lectin (MCL), Clec12b and the murine DC-SIGN homologues SignR1 and SignR3.

To achieve a more uniformly oriented presentation of the CLR extracellular CRDs on the microarray surface, NHS-slides were first coated with protein G before the CLR-Fc fusion proteins were subsequently spotted on top (Figure 4).

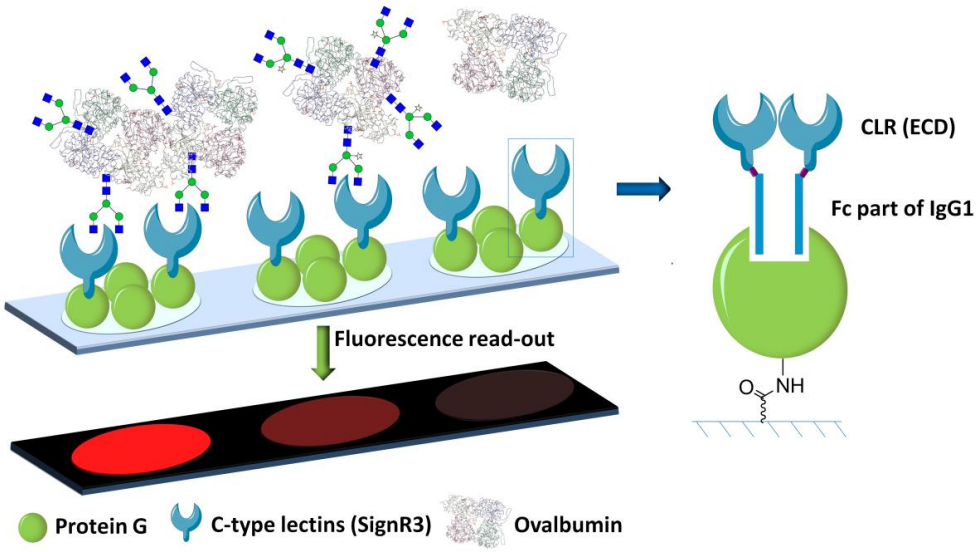


Figure 4. Schematic representation of the protein G coated lectin array-assisted binding studies with glycoconjugates and unconjugated OVA. In the zoomed view illustration of the protein G recognizing Fc part of the IgG1 antibody fused with extracellular domain (ECD) of C-type lectin receptors (CLR).

Protein G binds to the Fc region of the CLR-Fc fusion proteins thus improving the orientation of the CLR on the microarray surface.[24,24] Human Fc was included as negative control. Glycoconjugates and OVA were fluorescently labelled in primary amino groups from lysines and N-terminal employing commercially available Alexa Fluor® 647 NHS ester. Solutions of XG0-OVA, G0-OVA and OVA in sodium carbonate buffer were incubated in the dark for 1 hour with the fluorescent reagent. Excess of dye was later removed by dialysis and ultracentrifugation. The reaction conditions were adjusted to obtain a comparable degree of substitution (DOS) of dye for each glycoprotein and unmodified OVA, which was further estimated based on absorbance of dye and protein. An average of 11 dye molecule was assigned per each construct (a detailed calculation is given in the Experimental section). Subsequently glycoconjugates and OVA dilutions in PBS were incubated in the dark on the CLR microarrays at 4°C overnight. The lectin microarray was washed and the fluorescence was measured in a microarray scanner (Figure 5).

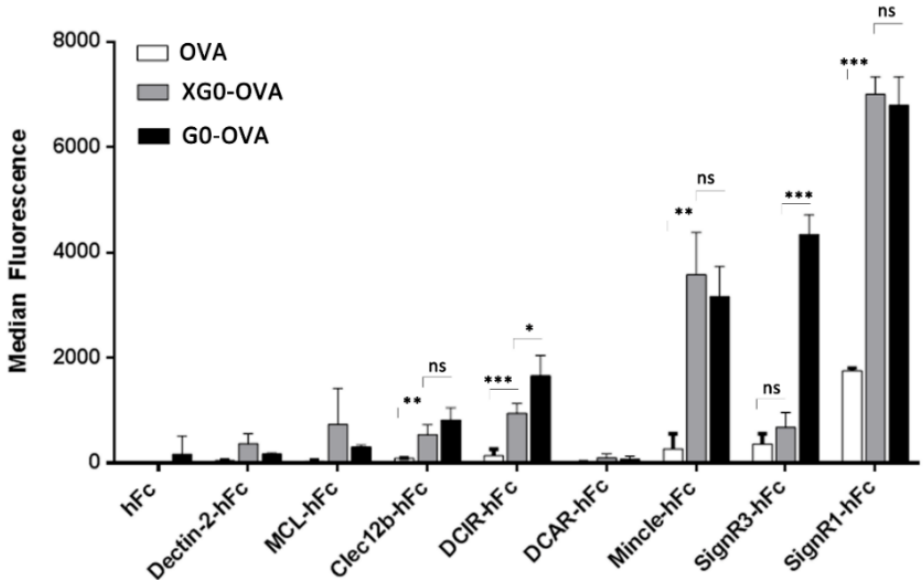


Figure 5. C-type lectin microarray-based analysis of binding of CLR-hFc fusion proteins to glycoconjugates XG0-OVA (grey) and G0-OVA (black) compared to unmodified OVA (white). Statistical analysis was performed with an unpaired student's t test, ***p<0.0001,**p<0.01, *p<0.05, ns-not significant.

Binding studies revealed that glycoconjugates bound strongly to three of the eight analyzed CLR-Fc fusion proteins: Mincle, and both DC-SIGN homologues, SignR1 and SignR3. Lower binding but still significant binding was detected for Clec12b and DCIR. While the presence of a core xylose in XG0-OVA showed no significant effect on glycoconjugate binding towards Mincle and SignR1, it strongly reduced the interaction with SignR3 compared to the non-xylosylated homolog G0-OVA. Unconjugated OVA was used as a control in all experiments to evaluate the impact of background glycosylation and potential protein-protein interactions on binding. With the exception of SignR1, with a known specificity towards high-mannose structures[25], we did not observe any significant binding of the control protein. Although native chicken OVA contains a single N-linked glycosylation site (Asn-292) which is predominantly substituted with high mannose ($Man_5GlcNAc_2$/$Man_6GlcNAc_2$) or less abundant hybrid ($Man_5GlcNAc_4$/$Man_5GlcNAc_5$) glycans,[26] this sugar modification seems to play a minor role in the recognition by CLRs, probably due to the monovalent presentation. Thus, the **G0/XG0** N-glycans seem to be exclusively responsible for recognition by the CLR-Fc fusion proteins tested on our array.

A significantly reduced binding of XG0-OVA compared to G0-OVA was observed for SignR3. We have shown previously that the xylose modification completely abolishes binding of **XG0** towards DC-SIGN[16]. The murine DC-SIGN family consists of several homologous type II transmembrane proteins of which each displays a single carboxyl terminal CRD.[27] Previously, it had been shown that the dominant specificities found for the DC-SIGN homologues SignR1 and SignR3 preferentially bind to mannose- and fucose motifs found in on high mannose type N-glycans and in Lewis A/B and Lewis X/Y moieties.[28] The differential carbohydrate specificities of both DC-SIGN homologues have been previously exploited for cell targeting with glycoliposomes.[29] SignR3 has been reported to bind to bacterial pathogens such as *Mycobacterium tuberculosis*, but also glycans of commensal microbiota.[30,31] Since SignR3 is considered to be the biochemically closest murine homolog of the human DC-SIGN receptor,[32,33] it might exhibit a similar glycan binding pattern as DC-SIGN thus explaining the differential binding of **G0** and **XG0** by SignR3.

4.4 DCs binding and uptake assays of glycoconjugates

4.4.1 Purification of DCs and T-cells by Magnetic Assisted Cell Sorting (MACS)

$CD11c^+$ splenic DCs as well as OT-II T cells used in the section 4.4.2, 4.4.3 and 4.10 were purified by magnetic-activated cell sorting (MACS) as shown on Figure 6. To isolate splenocytes, spleens were flushed with cell culture medium and the cell suspension was filtered through a cell strainer to remove cell aggregates. After centrifugation cell pellets were re-suspended in erythrocyte lysis buffer, washed and re-suspension in MACS buffer (PBS, 0.5% BSA, 2mM EDTA). Dendritic cells ($CD11c^+$ cells) were isolated from suspension of C57BL/6 murine spleen cells by positive sorting applying microbeads containing antibody directed against $CD11c^+$ surface marker of DCs (Figure 6, B).

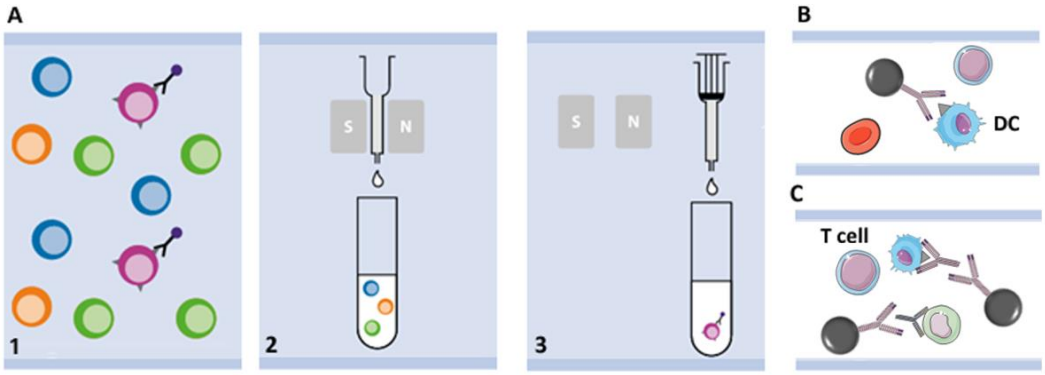


Figure 6. A. Magnetic Assisted Cell Sorting-(MACS) principles; **1.** Cells in a single-cell suspension are magnetically labeled with MACS microbeads; **2.**The sample is applied to a MACS column placed in a magnet and the unlabeled cells pass through the column; **3.** The column is removed from the separator, and the magnetically labeled cells are eluted from the column; **B.** Positive selection: target cells type (DCs) are magnetically labeled and are retained within the column; **C.** Untouched isolation: target cells type (OT-II T cells) are unlabeled and collected in the flow-through fraction. Figure A was taken from: http://www.miltenyibiotec.com.

Cells incubated with magnetic microbeads were loaded onto a MACS column placed in a magnetic field. Unbound cells passed through the column, while remaining $CD11c^+$ cells were washed with MACS buffer and eluted from the column by removing the magnetic field. To increase DC purity, the $CD11c^+$ cell purification was repeated. The cell suspension was centrifuged, resuspended in IMDM complete medium and counted. Subsequently, T cells were isolated from spleen of OT-II transgenic mice using the Pan T Cell Isolation Kit II (Figure 6, C). Non-target cells were magnetically labeled with a cocktail of biotin-conjugated antibodies followed by anti-biotin monoclonal antibodies conjugated to MicroBeads. Non-targeted cells were retained on the MACS column, while the untouched T cells passed through the column in the first fraction. As in a case of DCs, to increase T cell purity, the T cell purification was repeated. The purities of MACS-isolated DCs and T cells were analyzed by flow cytometry (Figure 7).

Flow cytometry (FC) is a technology that allows for the simultaneous measurements of multiple physical properties of particles (here cells) which flow in a fluid stream through a beam of light.[34] The characteristic cells features which can be assigned from a single FC measurement include their relative size, granularity or internal complexity and relative fluorescence intensity. These characteristics are determined using an optical-to-electronic coupling system that records how the cell scatters incident laser light and emits fluorescence. Forward-scattered light (FSC) is proportional to the cell-surface area or size while the side-scattered light (SSC) indicates the cell granularity or internal complexity. Correlated measurements of FSC and SSC can be used to differentiate cell types in a heterogeneous cell population.[34] On the other hand, fluorescently labeled monoclonal antibodies directed towards certain cell surface markers are employed to identify a particular cell type in a mixed population of cells, as well as to distinguish isolated subpopulations of cells. FSC and SSC data combined with the immunofluorescence staining pattern can be used for identification of cells present in a sample and to count their relative percentages.

Figure 7 shows scatter plots of MACS-purified DCs and T cells with correlated SSC and FSC data and fluorescence emission of mAbs-labeled surface markers. $CD11c^+$ DCs were stained with anti-CD11c APC-labeled antibody while the OT-II T splenic cells with APC-eFluor780-labeled antibody, that targets a T cell receptor associated molecule CD3e (Figure 7A-B, second and third column). The FACS analysis of MACS-purified DCs and T cells demonstrated a purity of 85.1% and 90.3% respectively. The control measurement of flow through collected during $CD11c^+$ DCs purification was also recorded (Figure 7A, first column).

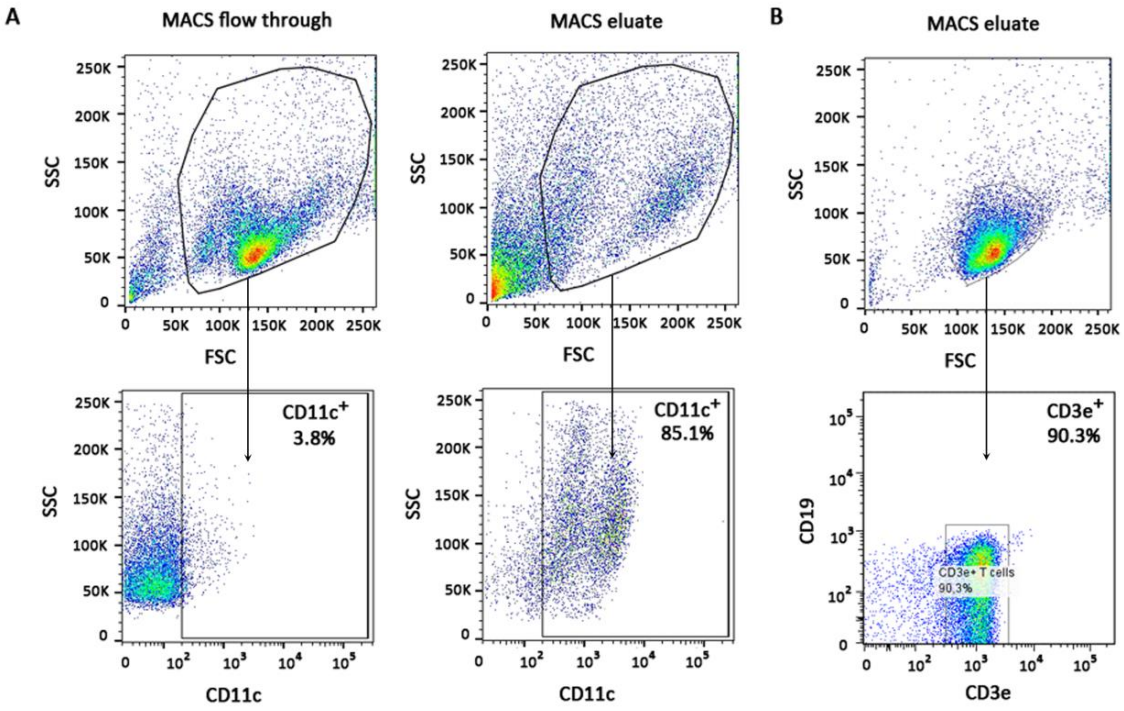


Figure 7. Representative dot plots of MACS-purified DCs and T cells **A.** splenic CD11c+ cells from wild-type C57BL/6 mice and **B.** splenic T cells from transgenic OT-II mice.

4.4.2 DCs binding and uptake assay followed by confocal microscopy and FACS

We then studied the uptake of fluorescently labeled G0-OVA, XG0-OVA and unconjugated OVA by murine DCs isolated from C57BL/6 mice by confocal microscopy (Figure 8A). Purified DCs were seeded on poly-D-lysine-coated glass cover-slips overnight, incubated with either fluorescently labeled OVA, G0-OVA or XG0-OVA, washed to remove unbound proteins and fixed.

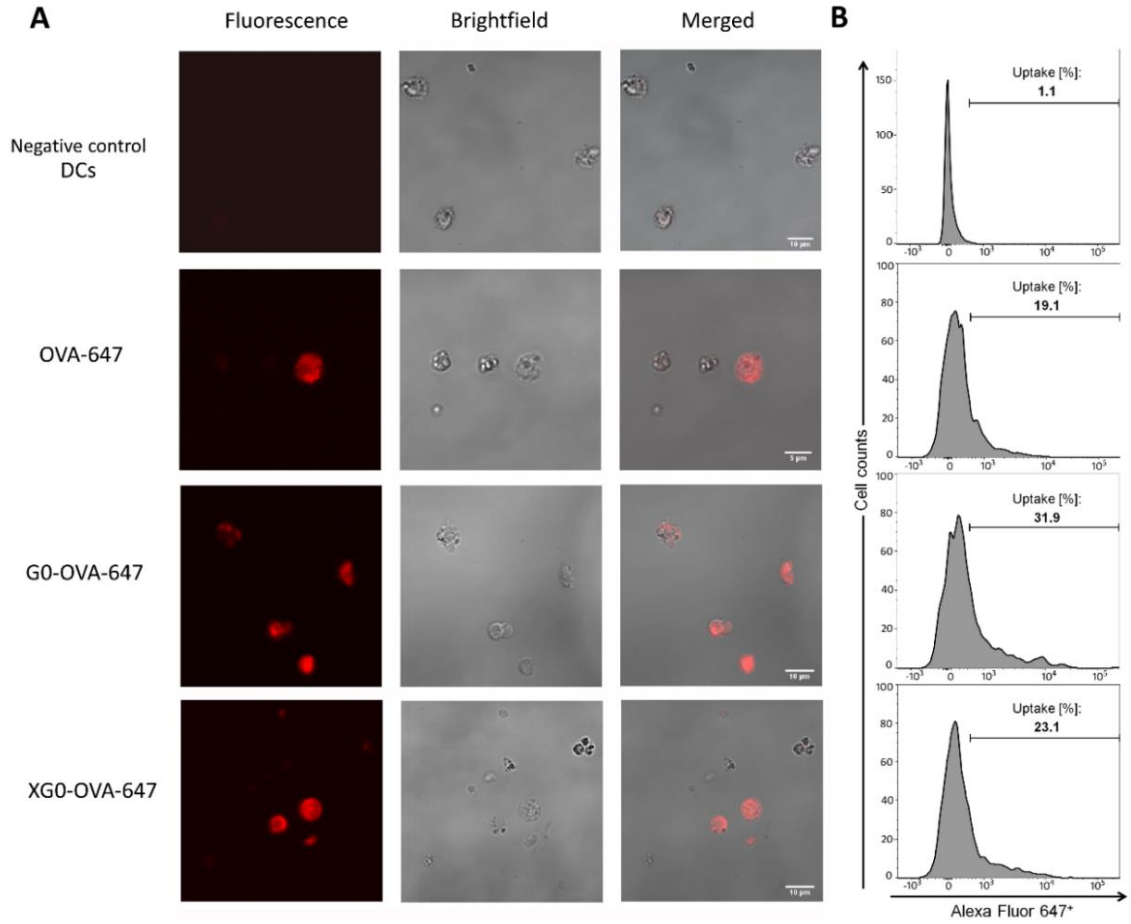


Figure 8. Uptake of the Alexa Fluor®-647 labeled glycoconjugates XG0-OVA and G0-OVA and OVA by murine DCs; **A.** Internalization of compounds measured by confocal fluorescence microscopy; **B.** Uptake of glycoconjugates and OVA analyzed by flow cytometry (data are representative for three independent experiments).

In all cases, we could observe a rapid internalization within 10 minutes (Figure 8A). To quantify potential differences in glycoconjugate uptake, which could indicate a possible influence of the core xylose modification on internalization, we analyzed the uptake by flow cytometry (Figure 8B). Purified DCs were pulsed with fluorescently labeled OVA and conjugates G0-OVA and XG0-OVA, incubated and subsequently washed to remove non-internalized fluorescent conjugates.

We observed that both sugar modifications increased OVA uptake by DCs. After 10 minutes of incubation, OVA was detected in about 17% of the murine DCs, while a slight increase was observed upon incubation with conjugate XG0-OVA (22.6%). However, incubation with the non-xylosylated glycoconjugate G0-OVA led to a marked increase in glycoconjugate internalization (in 37% of the murine DCs) indicating that the respective glycan modification indeed markedly impacted DC targeting efficacy.

4.4.3 *In vitro* stimulation and DCs/T cells co-culture

Upon internalization into DCs, antigens are processed and peptide fragments can be presented on major histocompatibility complex (MHC-II) molecules to T cells. However, besides activating T cells, DCs may also shape T cell responses, for instance by inducing $CD4^+$ T cell differentiation into Th1, Th2, or Th17 cells.[35] To analyze whether the differential DC targeting capability of the OVA glycoconjugates also impacted subsequent T cell activation, a DC/T cell co-cultivation assay was performed. To this end, $CD11c^+$ splenic DCs were pulsed with the OVA neoglycoconjugates or OVA alone and then incubated with purified $CD4^+$ T cells from OT-II transgenic mice. This transgenic system allows to evaluate antigen presentation and subsequent T cell activation since OT-II $CD4^+$ T cells have a T cell receptor specific for the $OVA_{323-339}$ peptide presented by the MHC-II molecule I-A.[36]

As a read-out, the expression of the early activation marker CD69[3738] was measured by flow cytometry. Indeed, we found a higher frequency of OT-II $CD4^+$ T cells expressing CD69 on their surface upon stimulation with the OVA glycoconjugates compared to the unconjugated

OVA (Figure 9A, B). Importantly, a significantly higher frequency of CD4$^+$ T cells expressed CD69 when DCs had been pulsed with G0-OVA compared to the xylosylated analogue XG0-OVA (Figure 8A, B). This finding indicates that the differential DC targeting efficiency of the OVA glycoconjugates observed in the DC uptake assay also correlates with subsequent T cell activation.

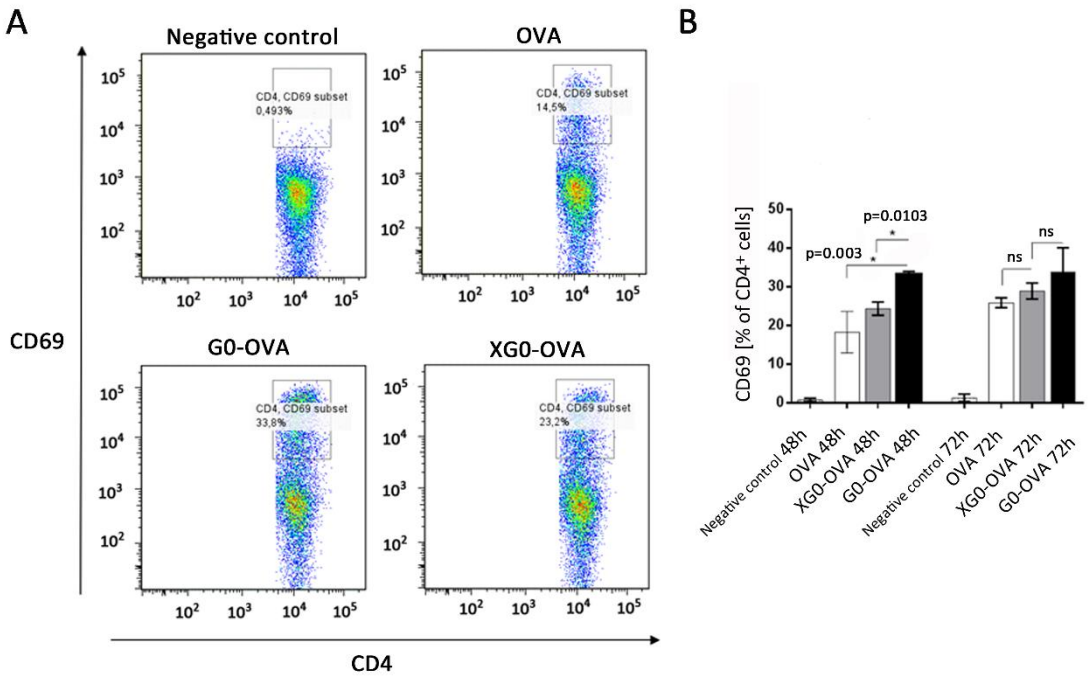


Figure 9. Frequencies of OT-II CD4$^+$ T cells expressing the early activation marker CD69 upon stimulation with glycoconjugates (G0-OVA, XG0-OVA) and OVA; **A.** Representative dot plots after 48 h; **B.** Statistical analysis for the incubation over 48 h and 72 h (data are representative for three independent experiments). Statistical analysis was performed with an unpaired student's t test, *p<0.05, ns-not significant.

Besides the expression of activation markers, activated T cells produce cytokines with different effector functions. Thus, as an additional functional read-out, we measured the concentrations of interleukin 2[39] (IL-2) and interferon gamma[40] (IFN-γ) produced by the OT-II CD4$^+$ T cells. While IL-2 is produced by activated T cells and further promotes T cell

proliferation, IFN-γ is an effector cytokine mainly produced by Th1 cells. Indeed, significantly higher levels of IL-2 were produced by T cells that had been stimulated with the OVA glycoconjugates compared to the unconjugated OVA (Figure 10A). While both glycoconjugates induced a higher IL-2 production than unconjugated OVA, the presence of the xylose modification in XG0-OVA resulted in a lower IL-2 level than for G0-OVA. A similar trend was observed for IFN-γ production (Figure 10B). Thus, the results obtained for glycoconjugate-dependent cytokine production are consistent with the CD69 expression data (Figure 9) and indicate that glycan-based DC targeting may be used to elicit subsequent T cell responses.

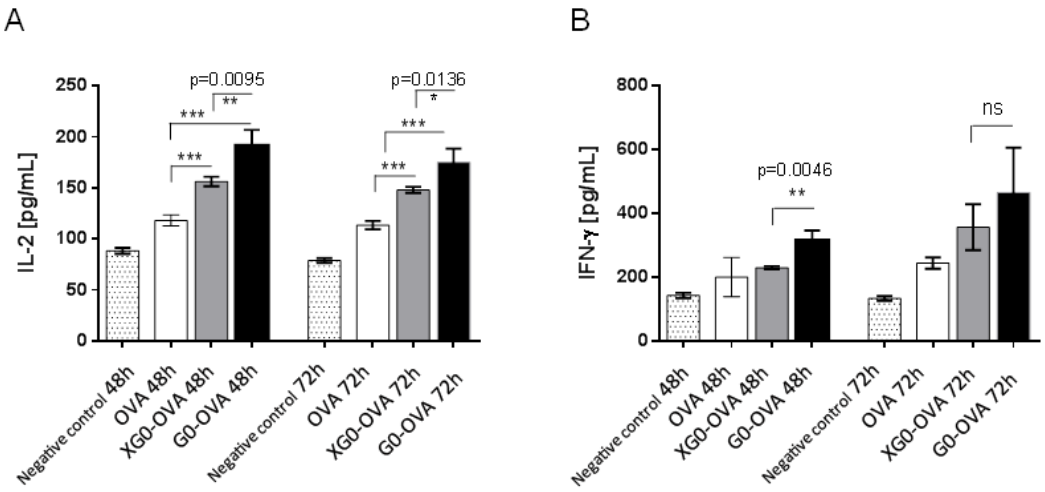


Figure 10. Cytokine production by OT-II $CD4^+$ T cells with $CD11c^+$ DCs pulsed with OVA, XG0-OVA and G0-OVA; **A.** Interleukin-2 (IL-2) production after 48 and 72 h of incubation; **B.** Interferon-γ (IFN-γ) production after 48 and 72 h (data are representative for three independent experiments). Statistical analysis was performed with an unpaired student's t test, *p<0.05, **p<0.01, ***p<0.0001, ns- not significant.

4.5 Summary

The use of protein or peptide antigens alone may not always induce an effective internalization by DCs and subsequent T cell activation. In this context, the use of glycan-modified antigens and their targeting via CLRs is a promising strategy to enhance the uptake of the construct. Recent studies indicate that different carbohydrates may even initiate differential signalling processes by engagement of the same CLR as has been shown for fucose-based DC-SIGN targeting.[41,42,43] However, targeting CLRs with glycans is still hampered by incomplete knowledge of CLR binding specificities.[44] In this Chapter, we demonstrate the utility of CLR microarrays to identify natural glycan ligands. N-glycans conjugated to the model antigen OVA were targeted to murine DCs *in vitro*. We showed that xylosylated and non-xylosylated biantennary N-glycans enhanced the uptake and presentation of OVA and, more importantly, we could demonstrate that even small structural glycan modifications can have a marked impact on uptake by DCs and subsequent T cell activation. In line with our glycan array based binding study, we found that the non-xylosylated G0-OVA glycoconjugate bound with higher affinity to SignR3 than XG0-OVA and unconjugated OVA. Consequently, G0-OVA was preferably internalized by murine DCs and led to enhanced T cell activation as demonstrated by CD69 staining. Finally, the production of IL-2 and IFN-γ measured after 48h was elevated further underlining an increase of T cell activation as well as their differentiation towards a Th1 phenotype. Ultimately, our study highlights the utility of glycan-based CLR targeting and may open up new possibilities for the specific delivery of vaccine antigens into DCs.

4.6. Glycoconjugate OVA-Gold Nanoclusters

In the second part of this Chapter, we will use G0-OVA glycoconjugate, which exhibited the highest DCs targeting, for the preparation of protein-derived gold nanoclusters, G0-OVA-AuNCs. This novel system combines the targeting properties of N-glycan with potential antigenic properties of the protein and bioimaging capabilities of fluorescent AuNCs (Figure 11).

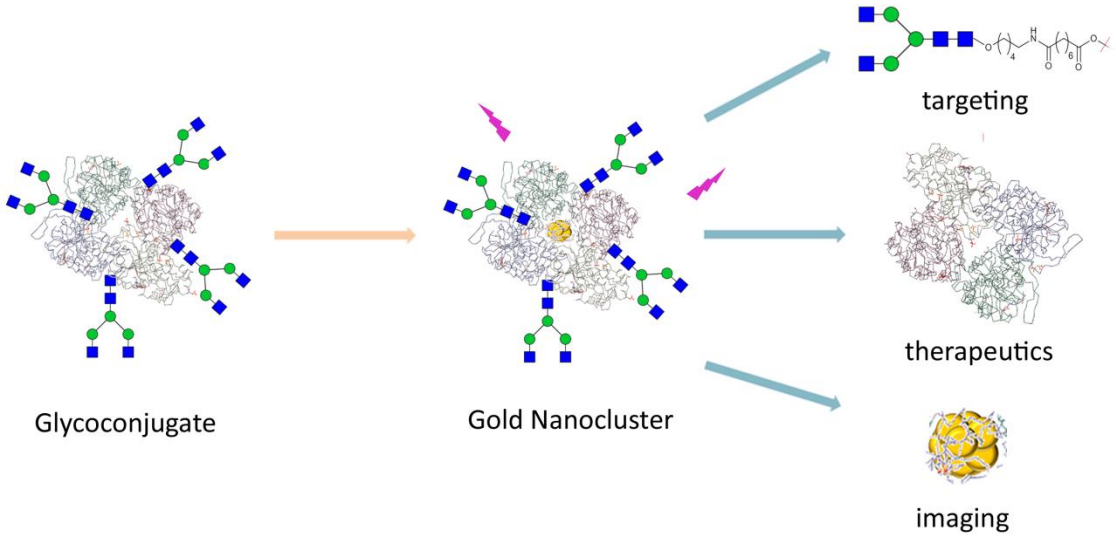


Figure 11. Schematic representation of the G0 glycan-derived OVA-AuNCs preparation with the properties assigned to each component of the system.

OVA-G0-AuNCs can be considered as an alternative system to fluorescently labeled OVA-G0. AuNCs are characterized by their biocompatible properties, like water solubility or low toxicity in a broad range of concentrations.[45] The size-tunable fluorescence of AuNCs falls into the near-IR window which is particularly well suited for *in vitro* and *in vivo* imaging as with this emission wavelength the minimal biological auto-fluorescence is induced. In addition, the presence of a gold core opens the possibility of *in vivo* X-ray computed

tomography (CT) imaging and radiotherapy application, as certain AuNCs have been found to act as radiosensitizers.[46]

4.6.1. Introduction to Gold Nanoclusters

The word *nano is* derived from Greek *nanos*, meaning dwarf and is used to describe materials or properties which occur on a nanometer scale (10^{-9} m).[47] The science dealing with nanomaterials, called nanoscience, combines multiple fields of research, including chemistry, physics, biology and medicine. One of the exciting branches of this multidisciplinary field is the generation of new (bio)materials and investigation their interaction with biological systems. The knowledge obtained can be further applied for therapy, drug delivery, imaging, biosensing and cell labeling.[21]

Among different metals, gold is one of the most extensively studied material in nanoscience. This nobel metal in its bulky state is highly unreactive and mostly associated with the jewelry or electronic industry[48] whereas in the nano scale form it has found entry into a variety of biomedical applications.[47] Gold nanomaterials can be categorized according to their dimensions, into gold nanoparticles (NPs) and gold nanoclusters (NCs).[21,49] Gold nanoparticles with sizes close to the wavelength of light (3-100 nm) exhibit distinctive properties which differs from others NPs, such as surface plasmon resonance (SPR) and their photothermal properties. Gold nanoclusters (AuNCs) on the other hand, are smaller than 3 nm in diameter which is consistent with ten to hundred gold atoms. Their ultra-small dimensions approach to the Fermi wavelength of electrons (0.7 nm), which places them between metal atoms and larger nanoparticles.[47] AuNCs exhibit molecular-like behavior and optical, electrical and chemical properties distinctive from the GNPs. Due to strong quantum confinement effects, AuNCs are too small to support surface plasmon resonance but can emit size-dependent fluorescence. Nowadays, the use of labeling agents, especially for the tagging of biomolecules has been demonstrated to be indispensable in many biological studies.[50] In this Thesis we have discussed how the fluorescence can be used to follow complex biological

processes, like carbohydrate-protein interactions, where the fluorescently-tagged lectin is required to evaluate its binding towards ligands printed on the microarray (see Chapter 2). In the first part of this Chapter, we used commercialized fluorescently-labeled antibodies to differentiate between the cell subclasses by FACS analysis and finally followed the uptake of Alexa Fluor®-647 glycoconjugates by DCs.

The fluorescent labeling commonly uses conventional organic dyes, like fluorescein and rhodamine, which are considered to be the earlier classes of fluorescent labeles.[50] They are easy to use because of the well-established standard protocols for conjugation and their availability. In spite of the popularity of organic dyes, some of them can have certain limitations, like poor photochemical stability and a low photobleaching threshold which decreases their sensitivity and limits the trafficking lifetime.[51] Stokes shifts of organic dyes, measured as a difference between the maximum of excitation and emission peaks, normally range between <50 nm.[52] The poor separation of the absorption and emission spectra together with their broad profiles may favour the cross-talk between different dye molecules. The small size of organic dyes minimizes possible steric hindrance which could interfere with the function of a labeled biomolecule. However, in the biological environment, some of even small organic dyes cause aggregation due to their high hydrophobicity. The latter issue of rhodamine or cyanine dyes was overcome via their sulfonation and generation of more hydrophilic, relatively new fluorescent molecules called Alexa dyes.[53,54]

Together with the development of the material science, new labeling agents have evolved and started gaining the popularity due to their better optical properties compered to organic dyes.[51] For example, fluorophore-tagged latex/silica nanobeads have become a good alternative to commonly used fluorophores.[55] By encapsulation of thousands of the organic dye molecules in a polymer shell, the labeled agents are less affected by environment and more stable to photobleeching. Still, several limitations can be associated with such a system, like leaking of the fluorophores through surface defects or particle agglomeration.[56] Nowadays, gold nanoclusters are considered to be another, novel ultra-small fluorophores with high biocompatible properties. They are much smaller than biological molecules in

which are entrapped and thus may have just a little effect on their functions and on their interaction with targeted biomolecules. Moreover, while entrapped by the biomolecule, gold clusters are isolated from the external environment and thus might have just a limited interplay with other molecules. Another advantage of AuNCs is their water solubility and simple preparation which follows green synthetic strategy, avoiding harmful chemicals or heavy metals. AuNCs exhibit good photostability, strong photoluminescence, large Stokes shift and size-dependent tuneable fluorescence.[21] Moreover, the long emission wavelength of AuNCs from the red visible region to the near-infrared (IR) is highly favorable in bioimaging because it overlaps the tissue transparency window (650-950 nm).

Template-based synthesis of gold clusters had been introduced in early 2000 which described the use of several biological macromolecules including dendrimers, polymers, proteins or even DNA for clusters formation. The first protein-directed synthesis of gold nanoclusters was reported by Xie and co-workers in 2009.[57] In this report, commercial bovine serum albumin (BSA) was used as a template and led to the development of a one-pot synthesis for BSA-AuNCs generation. As a result, red emitting, small clusters consistent of 25 gold atoms were obtained with a photoemission peak at around 640 nm. Based on this strategy several other groups prepared BSA-AuNCs and demonstrated examples of their possible applications.[21] Dong's group for example, showed the utility of as synthesized BSA-AuNCs as a fluorescent probe for the detection of glucose[58], while Huang and coworkers further functionalized BSA-AuNCs with folic acid for *in vitro* and *in vivo* tumor-targeted imaging.[59]
The ease of the synthesis and the attractive, potential bio applications of BSA-derived gold nanoclusters promoted the investigation of alternative proteins for the protein-template assisted synthesis of AuNCs. Geckeler[60] and one year later Yu-Chie Chen[45] prepared gold nanoclusters (AuNCs@ew) with inexpensive chicken egg white proteins (CEW) directly separated from egg and the latter have successfully applied them for the sensing of polyphosphates and Concananalin A (ConA). At this point it is known, that the preparation of nanoclusters is mainly dependent on the amino acid composition of the protein which is used

as a template. It acts as both, reducing and stabilizing agent but also restricts the size of the gold nanoclusters by preventing their further growth and expansion into larger nanoparticles.[61] Template protein which can be successfully applied for the generation of AuNCs with red photoluminescence have to carry sufficient numbers of certain amino acids, such as cysteine and tyrosine residues which are able to reduce Au (III) into Au (I) and Au (0) and thus promote cluster formation. [59,45]

In the first part of this section we will discuss the synthesis of OVA-derived AuNCs and their detailed characterization, including UV-Vis and fluorescence spectroscopy, dynamic light scattering (DLS) or transmission electron microscopy (TEM). Finally, based on the knowledge acquired during the OVA-AuNCs synthesis we will present the synthesis of first glycoprotein gold nanoclusters (OVA-G0-AuNCs) and discuss possible biological applications of glycan derived clusters exemplified by the DCs uptake study analyzed by confocal microscopy.

4.7. Synthesis of OVA-AuNCs

Here the preparation of G0-OVA-AuNCs based on prior experiments employing native OVA are described.

In a typical experiment, protein-derived gold nanoclusters are generated by the addition of tetrachloroauric (III) acid ($HAuCl_4 \cdot 3H_2O$) to a highly concentrated aqueous solution of protein under vigorous stirring. After several minutes, the pH of reaction is adjusted to 10 or higher with sodium hydroxide (NaOH) and the mixture is incubated in the dark at 37 °C (Figure 12). Generated AuNCs are subsequently dialyzed over water to remove excess of reagents.[57 62]

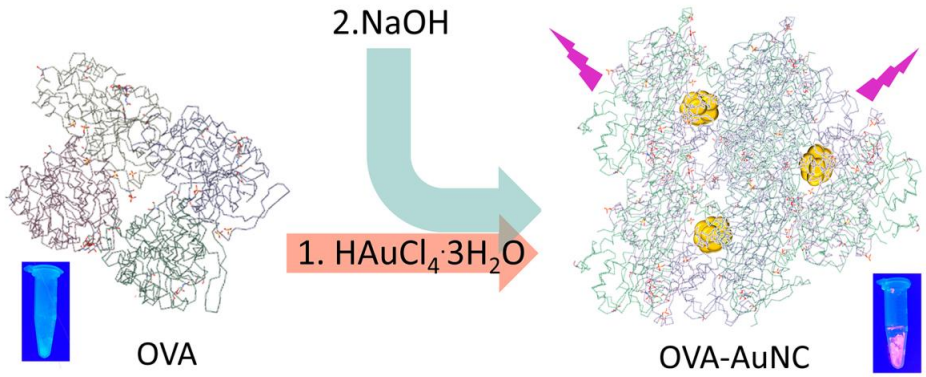


Figure 12. Schematic representation of OVA gold nanoclusters (OVA-AuNCs) synthesis starting with OVA protein and $HAuCl_4 \cdot 3H_2O$ followed by addition of sodium hydroxide; the bottom left picture correspond to OVA protein and bottom right picture to OVA gold nanoclusters, both under UV lamp irradiation (365 nm).

After revision of the literature we studied several synthetic strategies for the generation of OVA-AuNCs.[57,59,60,63] A major aim in the optimization of an experimental protocol was to reduce amount of protein or more importantly precious OVA glycoconjugates required for the efficient formation of fluorescent clusters.

To this end we investigated the effect of different OVA concentrations (15, 10, 5, 2.5, 1.25 and finally 0.65 mg mL^{-1}) on protein gold cluster formation. The final concentrations of $HAuCl_4 \cdot 3H_2O$ (4.2 mM) and NaOH (150 mM) as well as the reaction temperature (37 °C) were kept unchanged. After overnight incubation at 37 °C in the dark, the fluorescence of the reaction mixtures was measured as an indicator of clusters formation (Figure 13). The highest fluorescence intensity was recorded at 15 $mg \cdot mL^{-1}$ of OVA concertation, followed by 10 $mg \cdot mL^{-1}$, which was further considered to be the minimal amount of protein required to generate fluorescent clusters. Unlike samples 1-2 (Figure 13B), OVA-AuNCs synthesized at 5 $mg \cdot mL^{-1}$ OVA concentration exhibited no visible or very low fluorescence emission (Figure 13C, pink line).

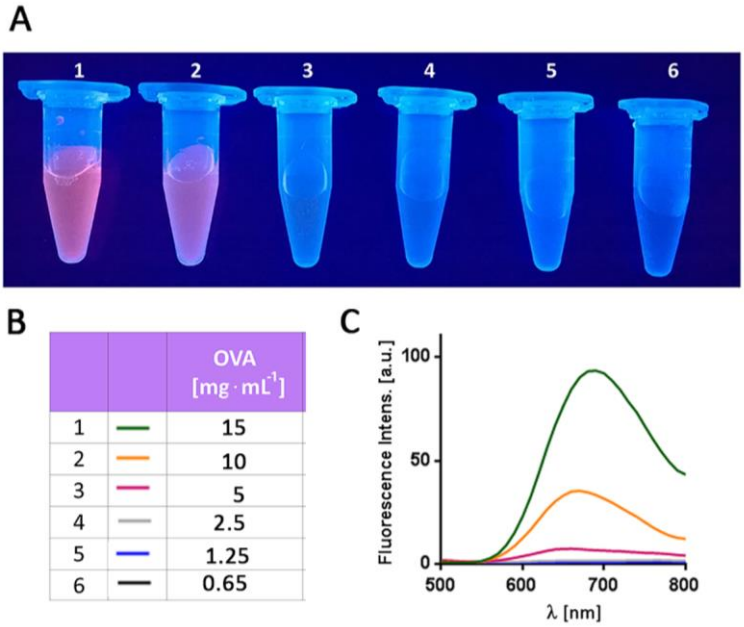


Figure 13. Optimization of the OVA-AuNCs synthesis; **A.** Picture of the OVA-AuNC solutions under UV light illumination (365nm); **B.** Amounts of OVA used during the synthesis; **C.** Fluorescence emission spectra of OVA-AuNCs at six different concentrations of OVA used.

Knowing that gold clusters exhibit size-tunable fluorescence which disappears with increased size of the gold core we have recorded the UV-vis spectra of three reactions (Figure 14).

Unlike AuNCs, the larger gold nanoparticles interact with incident electromagnetic fields resulting in the excitation of localized surface plasmon resonance (LSPR) that prompts light scattering and absorption at ~520 nm.

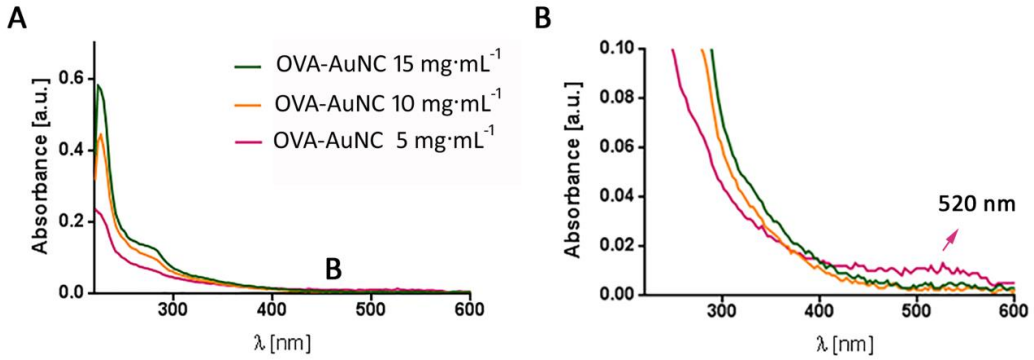


Figure 14. A. UV-visible spectra of OVA clusters generated with different amounts of protein; **B.** Zoomed view in the region of 400-600 nm shows the plasmon band at 520 nm for molecules generated with 5 $mg{\cdot}mL^{-1}$ of OVA (pink line).

The presence of LSPR observed only in the sample at 5 $mg{\cdot}mL^{-1}$ (Figure 14B) can explain the absence of fluorescence due to the increased gold core size from cluster (up to 3 nm) into nanoparticle range.

Finally, analyzed samples formed with less amount of OVA (2.5-0.65 $mg{\cdot}mL^{-1}$) lack the fluorescence suggesting no cluster formation. Moreover, after freeze-drying and re-suspension in water an important precipitate was observed, indicating lack of stability of created systems.

In order to accelerate the synthesis of OVA-AuNCs, its preparation has been alternatively investigated applying the microwave irradiation.[45,63] Based on the previous experiment (Figure 13) we have investigated the influence of the temperature, reaction time and the amount of gold (III) chloride trihydrate added to the 15 $mg{\cdot}mL^{-1}$ solution of OVA (Figure 15A, entries 1-5). Solution of protein in water was placed in a microwave vial followed by addition of 0.1M $HAuCl_4 \cdot 3H_2O$ (different concentrations, see Figure 15A). The reaction mixture was

allowed to incubate under vigorous magnetic stirring at 37 °C for 5-10 minutes. After this time 1M NaOH (150 mM) was added and the vial was placed in the microwave for 6 minutes at either 100 °C or 70 °C.[59] After heating, the fluorescence emission of different reaction mixtures was measured. As a result, the strongest fluorescence intensity was recorded in the sample 1 (Figure 15, B-C) containing 4.2 mM $HAuCl_4 \cdot 3H_2O$ and heated to 100 °C. Subsequently, in order to establish the optimal time of the microwave-assisted reaction, we have employed heating cycles of 3, 6, 9 and 18 minutes (Figure 15, D).

A

		OVA [mg·mL^{-1}]	Temp. [°C]	$HAuCl_4$ [mM]
1	—	15	100	4.2
2	—	15	70	4.2
3	—	15	100	8.4
4	—	15	70	8.4
5	—	15	100	12.6
6	—	10	100	4.2

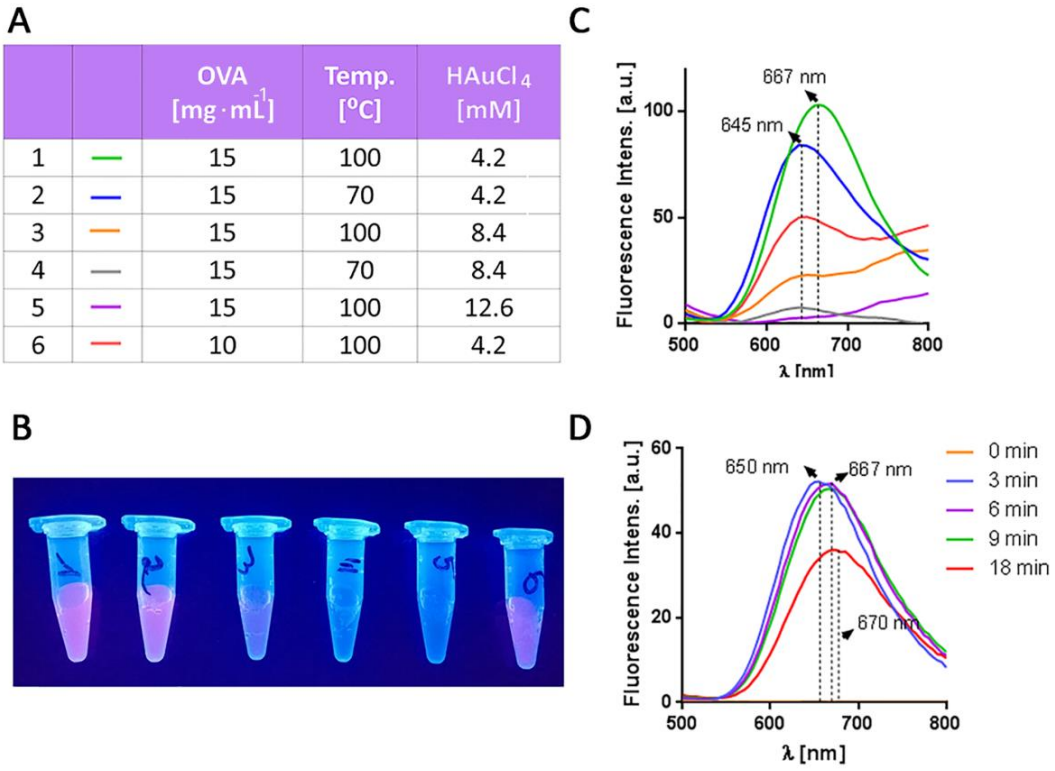


Figure 15. Microwave assisted synthesis of OVA-AuNCs, **A.** Table with conditions used during the study; **B.** Picture of the OVA-AuNCs solutions under UV light illumination (365nm); **C.** Fluorescence emission spectra recorded for six reactions with different conditions applied; **D.** Fluorescence emission spectra recorded at different reaction times.

Already after 3 minutes the fluorescence intensity reaches maximum and after another 3 minutes it shifts from 650 nm to longer wavelengths (670 nm) while the intensity remains unchanged. It is reported in the literature by Xie and coworkers, that a shift of the fluorescence emission maximum into shorter wavelengths correspond with formation of larger size AuNCs.[64] When longer time of heating was applied (Figure 15D, 18 minutes) partial decrease of the fluorescence was observed which could be associated with overheating and decomposition of part of clusters or to the partial generation of larger gold nanoparticles. As a result, in order to obtain the maximum of fluorescence and the smallest clusters possible we have chosen heating cycle of 6 minutes for the next experiments.

A mechanism of AuNCs formation has been proposed by Xie and coworkers[57] who postulated that upon addition of Au (III) ions to the aqueous protein solution, the protein molecules sequesters Au ions and entraps them. The reduction ability of protein is activated by the addition of NaOH, and the entrapped Au ions undergo progressive reduction to form AuNCs in situ. The amino acid involved in the gold clusters formation have been suggested to be cysteine, threonine and tryptophan.[60] It has been reported that tyrosine reduces Au(III) to Au(I) by their phenol groups and their reduction capability can be strongly improved by adjusting of reaction pH above the pKa of tyrosine (~10).[65] Tryptophan moieties can also reduce metal ions in alkaline medium, after deprotonation of its indole group and they have been found to be the strongest reducing amino acid for gold in protein AuNCs formation.[66] On the other hand, cysteine residues have been reported to facilitate the formation of Au(I) thiolate polymers based on their reduction capability of thiol groups.[64] But additionally thiol groups of cysteine residues have been suggested to play important role as stabilizer for the nanocluster due to the known interaction between gold and sulfur atoms.

4.8 Characterization of OVA-AuNCs

In this section the detailed characterization of AuNCs will be described. The optical properties of AuNCs were characterized by UV-vis absorption and fluorescence emission measurements. Additionally, the size of AuNCs was determined using transmission light microscopy (TEM), dynamic light scattering (DLS) and MALDI-TOF spectrometry.

In the daily light OVA-AuNCs exhibit pale brown color and under UV light illumination (365 nm) emit an intense fluorescence (670 nm). The spectral characteristics of OVA-AuNCs are given in the Figure 16. UV-vis spectrum of OVA-AuNC lacks LSPR at 520 nm suggesting the absence of gold nanoparticles larger than 3 nm. The solution of clusters exhibits an absorbance shoulder at 278 nm, which correspond to the absorption from the aromatic residues (tryptophan and tyrosine) and disulfide bond of native protein (Figure 16, A).[67,58] In comparison to OVA the absorbance band of OVA-AuNCs slightly differs by more monotonous increase towards shorter wavelengths, which can be explained by the conformational changes of OVA during OVA cluster generation.[68]

The emission spectrum of OVA-AuNCs (Figure 16, B) displays a maximum around 670 nm approaching the near-IR region (range between 700 nm to 1.1 µm) upon excitation at 350 nm. It is worth mentioning that the excitation peak of AuNCs is quite broad and the Stoke shift large (above 200 nm). These spectral properties allow for the simultaneous excitation of AuNCs with other fluorescent probes at single excitation wavelength and their further discrimination based on the different fluorescent emission spectra.[69] This spectral multicolor detection is also known as multiplexing.[52] With a violet laser diode for example, blue-fluorescent dye Alexa Fluor ® 405 and red-fluorescent OVA-AuNCs could be both excited at ~400 nm and would emit at 421 and 670 nm respectively with no spectral overlap.

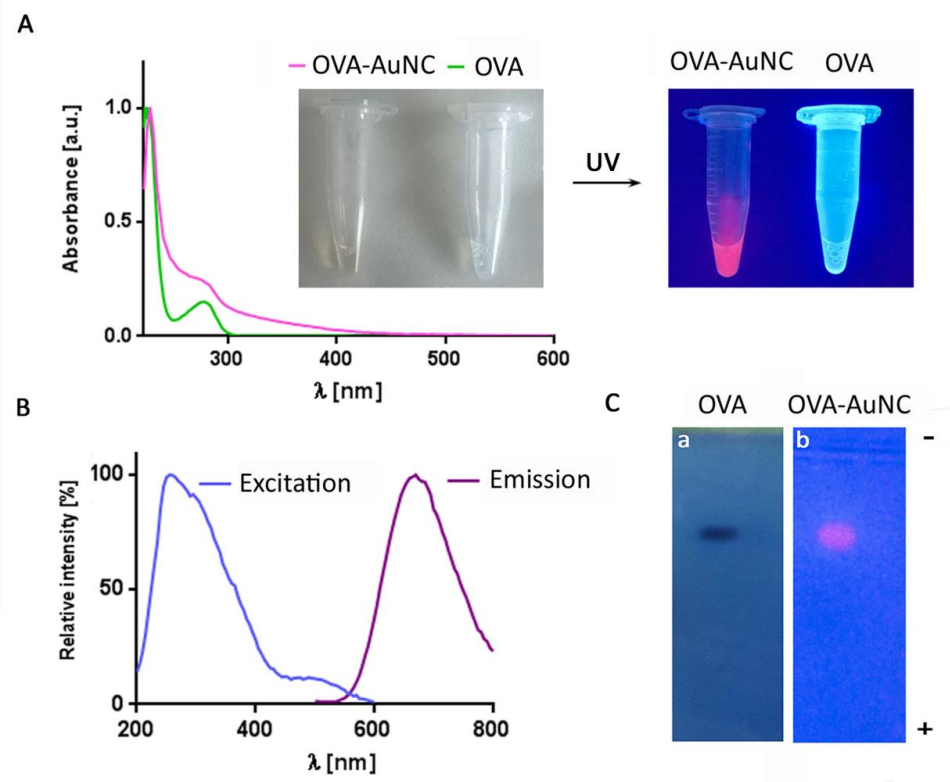


Figure 16. A. UV-visible spectra of OVA (green line) and OVA-AuNCs (violet line) and image of OVA and OVA-GNCs solutions under visible illumination (left) and UV light illumination (365 nm) (right); **B.** Fluorescence excitation (blue line) and emission (violet line) spectra of OVA-AuNCs; **C.** Agarose gel (0.75%) of **a)** OVA protein stained with Coomassie Brilliant Blue G-250; **b)** OVA-AuNCs under the UV light illumination.

OVA-AuNCs were further analyzed by agarose gel electrophoresis using as standard a solution of OVA protein. The band corresponding to OVA was stained by Coomassie Brilliant Blue G-250 and exhibited very similar electrophoretic mobility that OVA-AuNCs band visualized under UV lamp illumination (Figure 16C). This feature indicates that both OVA and OVA-AuNCs have similar size and as they migrate to the positive pole, both samples are negative in charge.

A size estimation of synthesized OVA-AuNCs was achieved by different techniques, as summarized in Figure 17. Dynamic light scattering (DLS) has been used to measure the hydrodynamic radius of OVA-AuNCs, which includes the inorganic gold core, protein coating and the solvent layer (Figure 17, A).

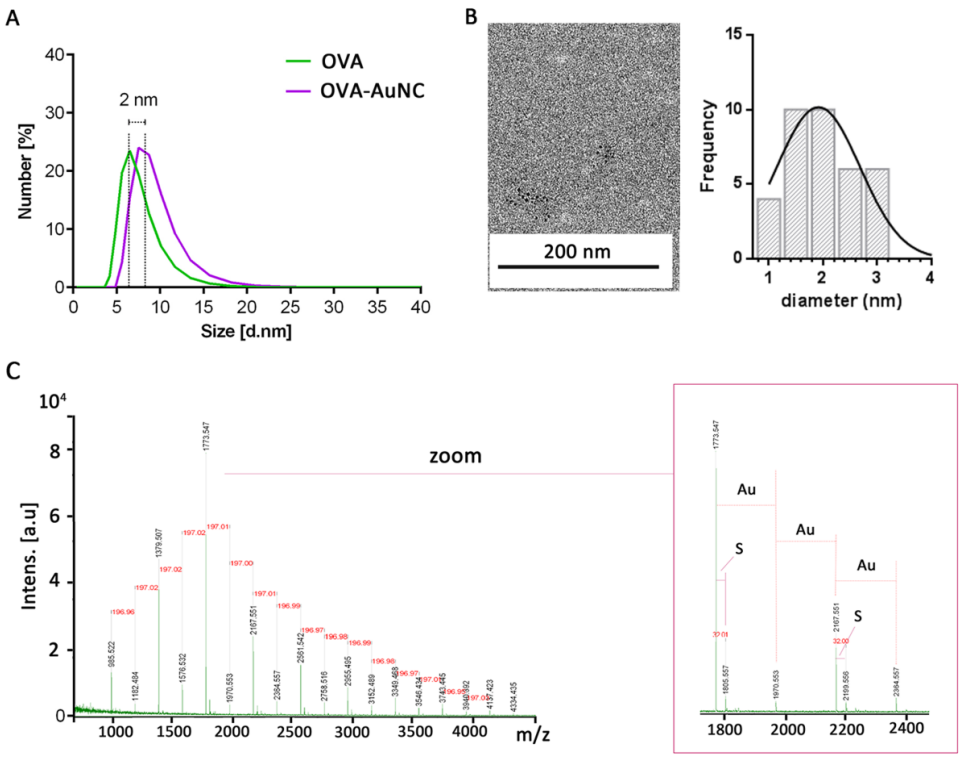


Figure 17. Characterization of the OVA-AuNCs; **A.** DLS of OVA protein (green line) and OVA-AuNCs (violet line); **B.** TEM image of OVA-AuNCs; **C.** MALDI-TOF spectrum of OVA-AuNCs with the zoom view in the region 1800-2400 m/z.

Remarkably, the size difference between OVA (6.5 nm) and OVA-AuNCs (8.7 nm) of about 2 nm corresponds well with the expected size of the gold core. This also suggests that the ratio of gold core to OVA is around 1:1. The samples were measured at high salt concentrations (0.1 M NaCl) and adjusted protein concentration (5 $mg \cdot mL^{-1}$) in order to avoid OVA aggregation.

The size of the inorganic core was further confirmed by transmission electron microscopic (TEM) image, giving an average size of gold clusters of 1.91 ± 0.76 nm (Figure 17, B). Additionally, OVA-AuNCs were studied by MALDI-TOF spectrometry (Figure 17, C) using sinapic acid as matrix and the spectra were acquired in reflection mode. MALDI-TOF spectrum showed signals with m/z values between 985.5 to 4334.4 corresponding up to 22 gold atoms per gold cluster, stabilized by cysteine thiolate groups of the protein (Figure 17, C). The gold core size corresponds to the described "magic number" sizes for AuNCs, with 2, 3, 5, 8, 10, 15, 18, 22, 25, 29, 33, *etc.* gold atoms/cluster depending on the synthetic strategy applied.[67] Intact OVA-AuNCs could not been detected by MALDI-TOF, suggesting that the protein may be detached from the core upon ion formation. This behavior of the protein-derived AuNCs had been reported on the example of clusters generated with molecules containing low numbers of cysteines like RNAse (8 Cys).[67] In contrary, AuNCs stabilized with BSA carrying 35 cystein residues have been previously detected by MALDI-TOF.[59] OVA contains four free cysteine residues and one disulfide bond, perhaps an insufficient number to stabilize the OVA-AuNCs during ion formation.

To clarify the binding between OVA protein and gold cluster as well as to establish the oxidation state of the inorganic gold core of the clusters, X-ray photoelectron spectroscopy (XPS) was performed (Figure 18). This technique allows the measurement of binding energies for all elements present in the cluster and the comparison with binding energies of reagents used during the reaction, such as $HAuCl_4 \cdot 3H_2O$ and native OVA. Figure 18 shows XPS spectra of $HAuCl_4 \cdot 3H_2O$ (A) and OVA-GNC (B). As expected, $HAuCl_4 \cdot 3H_2O$ shows the typical binding energy for gold 4f Au(III) at 85.5 eV (Figure 18, A). Once OVA-stabilized AuNCs are formed, the binding energy of gold 4f Au(III) shifts to 84.3 eV indicating reduced state of gold in OVA-AuNCs (Figure 18, B). According to the literature this value can be assigned to a mixture of Au (I) and Au (0) species.[57,70]

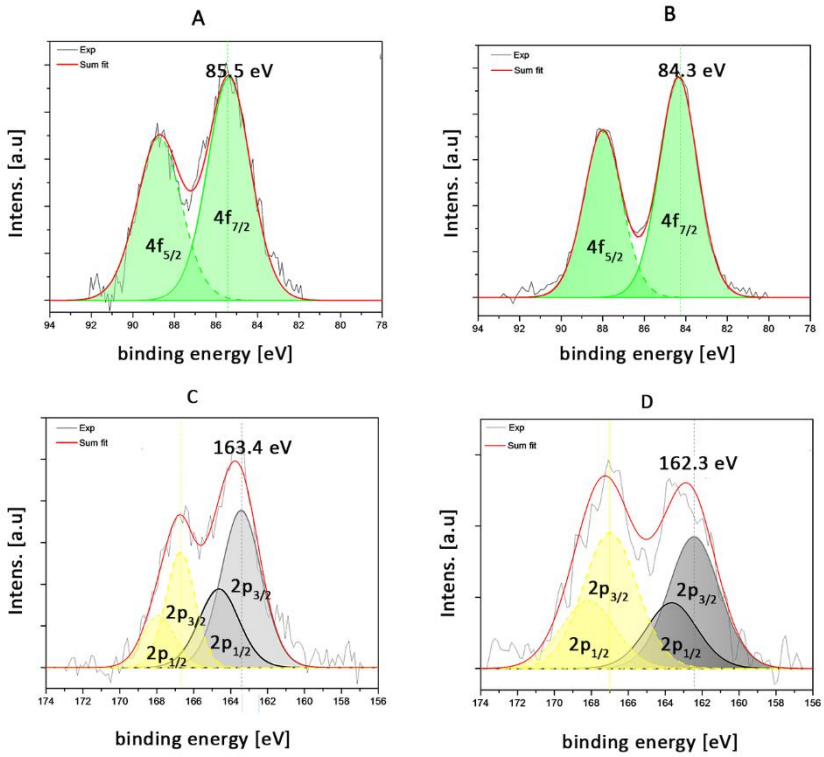


Figure 18. X-ray photoelectron spectra (XPS) of $HAuCl_4 \cdot 3H_2O$, OVA solution and OVA-AuNCs **A.** Au 4f region for solid $HAuCl_4 \cdot 3H_2O$ **B.** Au 4f region of OVA-AuNCs solution **C.** S 2p region of OVA solution **D.** OVA-AuNCs solutions.

Regarding the protein alone, sulfur S 2p region in native OVA (Figure 18, C) is resolved into two species: one with binding energies at 163.4 eV assigned to S-H (S-C) groups of cysteines and other at 166.7 eV, which corresponds to the oxidized sulfur species, like disulfide bound (S-S).[71] For AuNCs (Figure 18, D) there is a visible shift observed in the binding energy assigned to the first species (S-H) into the lower values of 162.3 eV which can indicate Au-S-C-S bound formation.[67,72,73] The maximum corresponding to the oxidized sulfur species stays unchanged. To summarize, the XPS analysis revealed the presence of Au(I) and Au(0) atoms and an involvement of the thiol groups of cysteine moieties in reducing Au during formation of AuNCs.

Knowing that OVA protein is attached to the gold core of clusters via cysteine resides as suggested by MALDI-TOF and XPS analysis, we have investigated a potential influence of the cluster generation on the secondary structure of OVA protein by circular dichroism (CD, Figure 19). Each regular secondary structure found in the protein has its own characteristic CD spectrum.[74] The broad negative band recorded for native OVA contains two peaks at around 208 and 222 nm characteristic for the alpha-helical structure of protein. Generation of OVA-AuNCs (Figure 19, violet line) causes irreversible changes in the structure of OVA demonstrated by a blue shift in the CD spectrum, a very low ellipticity above 210 nm and negative band around 195 nm.[75] Upon gold reduction and cluster formation OVA adopts a more irregular and disordered structure.

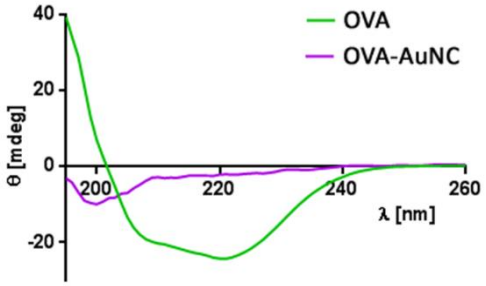


Figure 19. Circular dichroism (CD) spectra of OVA solution (green line) and OVA-AuNCs (violet line) indicating conformational changes in the native protein compared to the OVA bound and stabilizing gold clusters.

Previous studies on gold nanoparticle have also reported a partial loss of the albumin secondary structure upon absorption on the surface of nanoparticle.[68,67] Thus the observed changes in the CD of OVA-AuNCs might indicate an attachment of protein to the gold surface but also could be caused by the harsh reaction conditions involving 1M NaOH and heating for clusters formation. Recently, a preparation of protein stabilized gold nanoclusters under mild reaction conditions that maintains the secondary structure and functionality of proteins attached to gold core was reported, which opens new possibilities regarding synthesis of OVA stabilized AuNCs.[76]

In this section, we have optimized the synthesis and characterized OVA-AuNCs using different techniques. Spectroscopic characterization of OVA-AuNCs was performed by UV-vis and fluorescence measurements. The size of OVA-AuNCs was estimated by agarose gel and DLS.

The size of gold core was assigned by TEM and MALDI-TOF. Oxidation state of gold core was assigned to a mixture of Au(I) and Au(0) species based on XPS measurements. Protein secondary structure and protein oxidation state was studied by CD and by XPS, respectively.

4.9 Preparation of glycoconjugate OVA-AuNCs

Having in hands fully characterized OVA-AuNCs and having established a synthetic strategy towards their generation, we investigated the preparation of AuNCs with OVA glycoconjugates. Based on the DCs cell uptake studies described in the first part of this Chapter, we have chosen G0-OVA conjugates as first candidate towards glycoprotein-based AuNCs generation (Figure 20). Applying microwave-assisted synthesis optimized in section 4.7, we efficiently synthesized sufficient amount of G0-OVA-AuNCs on a mg scale which allowed further characterization and a first biological assay.

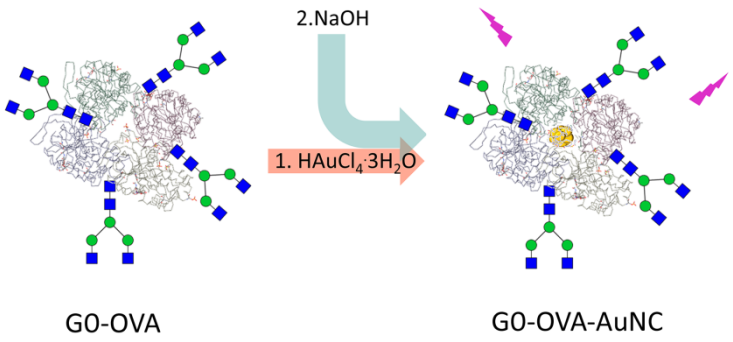


Figure 20. Schematic representation of the G0 glycan-derived OVA-AuNCs generation.

To investigate a possible influence of the sugar on the physical properties of G0-OVA-AuNC, we have measured its UV-visible absorption spectrum, the fluorescence emission spectrum and determined the gold core size by TEM. The results are summarized in Figure 21.

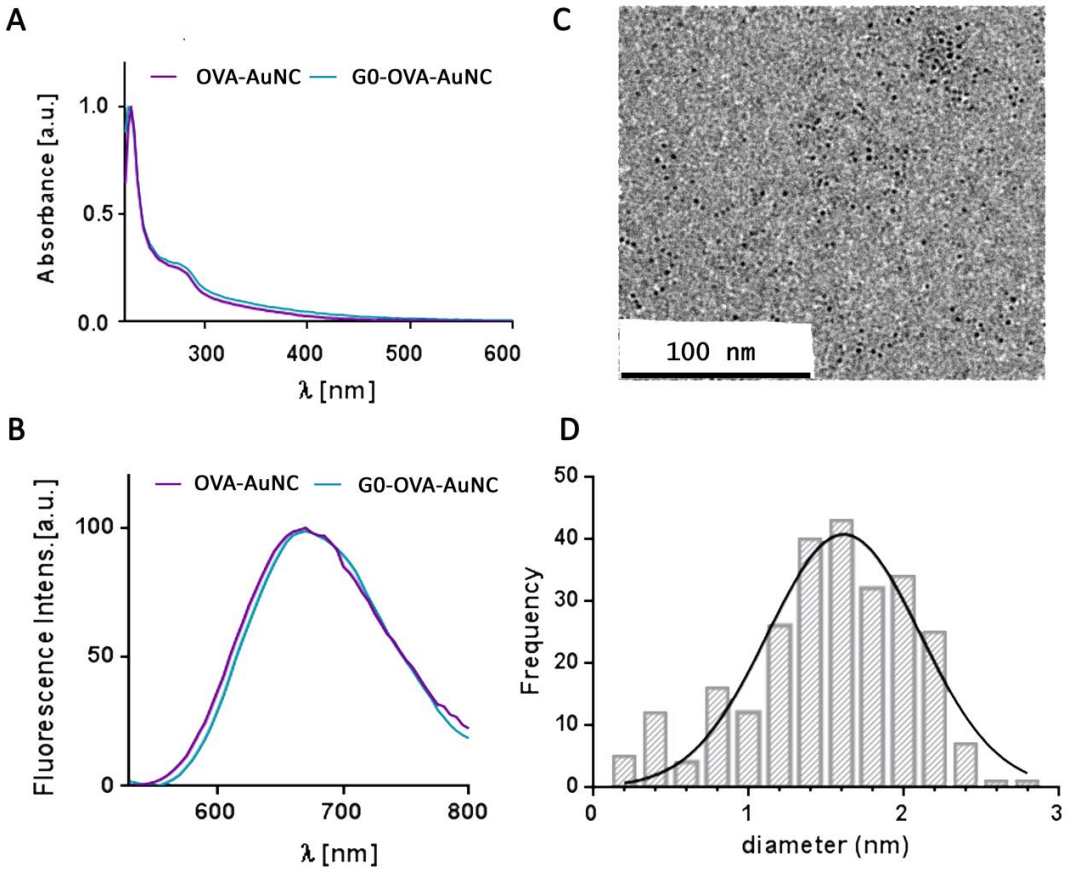


Figure 21. Characterization of G0-OVA-AuNCs; **A.** UV-visible spectra of OVA-AuNC (violet line) and G0-OVA-AuNC (blue line); **B.** Fluorescence emission spectra of OVA-AuNCs (violet line) and G0-OVA-AuNCs (blue line); **C.** TEM image of G0-OVA-AuNCs **D.** Size distribution of gold core diameter of G0-OVA-AuNCs.

The absorption UV-vis and fluorescence emission of G0-OVA-AuNCs were identical to OVA-AuNCs with characteristic absorbance at 278 nm and red fluorescence emission with maximum around 670 nm (Figure 21 A-B). The size of G0-OVA-AuNC gold core was established based on TEM image showed an average diameter of 1.61 ± 0.49 nm (Figure 21 C-D) comparable to the size of OVA-AuNCs (1.91 ± 0.76 nm). The spectral characteristics and size of glycan derived AuNCs demonstrated that the glycan conjugation to OVA did not influence gold nanocluster formation.

To further investigate possible impact of the sugar structure and/or the number of glycan attached to OVA on the generation and optical properties of protein stabilized gold nanoclusters, we performed the synthesis with different numbers of **G0** glycan conjugated to OVA protein (2-3 or 5-6 copies) and applying XG0-OVA glycoprotein, with the same ligand valency as starting G0-OVA (3-4) but with additional 2-O-xylose core modification on N-glycan (Figure 22).

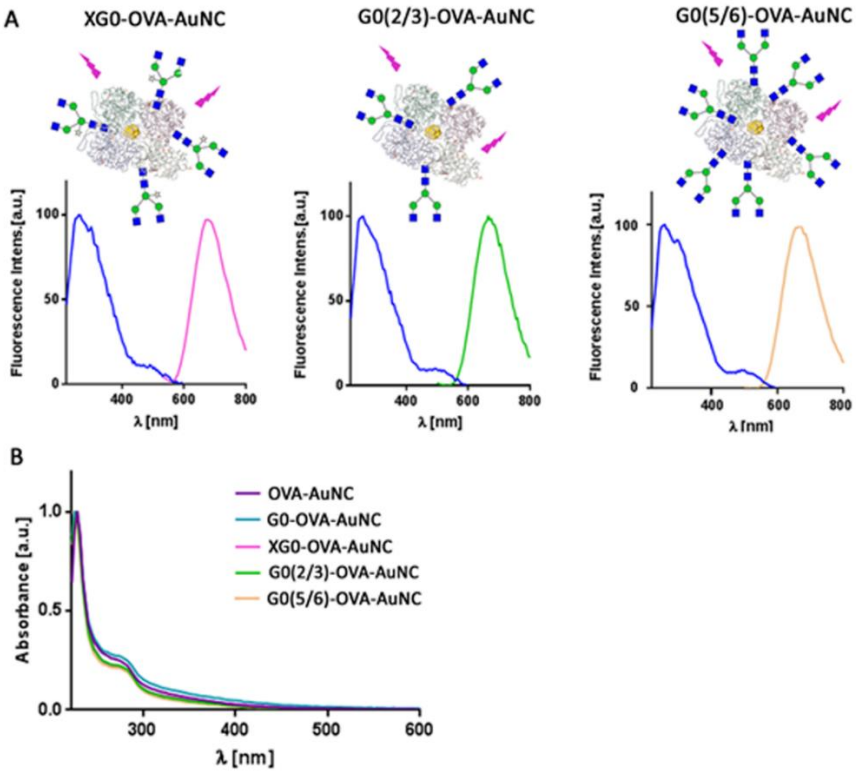


Figure 22. Spectral characterization of glycan modified OVA-AuNCs; **A.** Excitation (blue lines) and emission spectra of XG0-OVA-AuNCs (pink), G0-OVA-AuNCs with 3-4 sugars (green) and G0-OVA-AuNCs with 5-6 sugars (orange); **B.** Comparison of UV-visible spectra of glycan modified OVA-AuNCs.

As shown on Figure 22 we did not observe any measurable effect of the sugar structure or its valency (measured up to 6 ligands) on the formation and optical properties of the gold nanoclusters. It can be speculated that the glycan moieties which are conjugated to OVA by

DSS linker through the free amine groups of either lysine moieties or OVA N-terminus do not affect the reduction properties of glycoproteins which, as mentioned before, are based on tryptophan, tyrosine and cysteine residues. Additionally, sugars seem not to hamper formation of S-Au bond even while presented in higher number (5-6 copies).
It is also important to remember that beside N-glycans occupied lysine moieties, other important functional groups, like hydroxyl, carboxylic and remaining amino groups are still available for site-specific labeling. It opens whole new possibilities of conjugations other components to the system and its further evolution and improvement.

Similar to protein-derived AuNCs, glycan-OVA-AuNCs generated during the study are highly fluorescent and biocompatible due to their size, water solubility and emission in the range of near-IR. Interestingly, unlike the purely protein-stabilized clusters, glycoprotein-AuNCs can potentially benefit from the presence of sugar moieties by exhibition of intracellular targeting properties mediated through the C-type lectin receptor endocytosis. The ability to track the biomolecule inside the cell is an important property for any fluorescent dye. Additionally, use of glycan containing AuNCs can find a direct application as a sensing probes for glycan binding proteins, like plant lectins. A first successful example of use AuNCs as a sensing probe for plant lectins was given by Chen and co-workers who described the quantitative analysis of ConA applying OVA-AuNCs by measuring the loss of fluorescence upon creation of the OVA-AuNCs-ConA complexes.[45] ConA is specific towards high mannose structures which predominantly substitute the Asn-292 glycosylation site of OVA protein.[26]

4.10 Biological application of glycoconjugate OVA-AuNC. Targeting of DCs

Prior to biological assays the stability and solubility of OVA-AuNCs were evaluated in a solution of fetal bovine serum (FBS), the most common used serum-supplement for *in vitro* cell culture. Freeze-dried OVA-AuNCs were easily re-suspended in 10% solution of FBS in PBS buffer and left overnight at 37 °C. After this time, fluorescence emission was recorded (Figure 23). The spectrum of OVA-AuNCs did not show any significant change in the emission, neither in the fluorescence peak maximum nor in the fluorescent intensity when compared to starting OVA-AuNCs, suggesting that cluster are stable over the studied time in the presence of FBS.

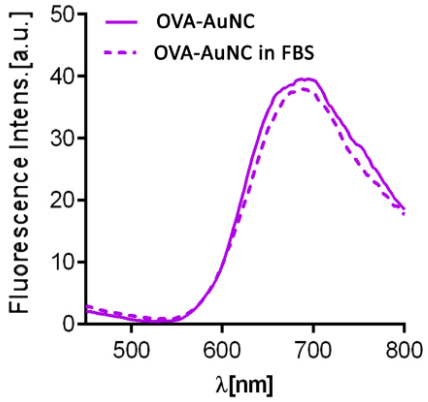


Figure 23. Fluorescence emission spectra of OVA-AuNC (violet continuous line) and OVA-AuNCs after incubation with 10% FBS (violet dashed line).

Taking advantage of the good bio-stability of OVA-AuNC in FBS solution, their imaging properties and previously demonstrated DCs targeting capability of G0 N-glycan, we were encouraged to study a possible *in vitro* biological application of G0-OVA-AuNCs by their uptake by murine DCs.

We have isolated murine $CD11c^+$ DCs from C57BL/6J mice following procedure described in section 4.4.1, seeded them on poly-D-lysine-coated glass cover-slips and pulsed with G0-OVA-AuNCs. The glass slide was washed to remove unbound clusters, DCs were fixed and the confocal measurements were performed. Figure 24 shows fluorescence images taken after 40 minutes of incubation. A negative control representing $CD11c^+$ DCs was also included.

Strong fluorescence could be clearly observed in the sample incubated with G0-OVA-AuNCs, suggesting their uptake by murine DCs. Similar results had been obtained with analogue, Alexa Fluor ® 647 labeled glycoconjugate. In this study, it was demonstrated that the developed system is suitable for *in vitro* uptake assays and that the glycoprotein gold nanoclusters are an interesting alternative to the use of organic dyes in cell assays.

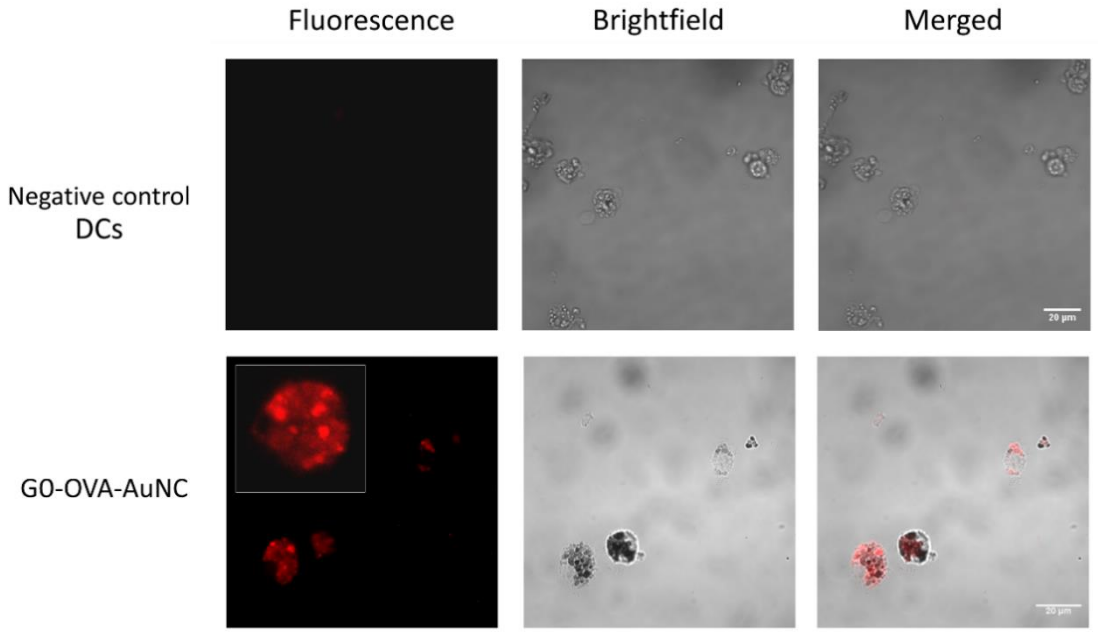


Figure 24. Uptake of fluorescent G0-OVA-AuNCs by murine DCs measured by confocal fluorescence microscopy.

Changes in the secondary structure of OVA protein during the AuNCs formation (CD) did not influence the DCs targeting capability of G0-OVA-AuNCs. Native OVA can be internalized by the DCs through receptor-mediated endocytosis, including mannose receptors recognizing sugars on OVA surface and through scavenger receptors which recognize negatively charged proteins.[77] Moreover, sugar modified clusters G0-OVA-AuNCs are targeted to the murine DCs through C-type lectins receptors described in the section 4.3 of this Chapter and thus conformational changes of protein should not play a role in the cluster internalization.

4.11 Summary

In this Chapter, we have shown the utility of glycoconjugates decorated with sugars selected on the basis of the microarray study described in the Chapter 2 (section 2.3.3, screening of DC-SIGN) for DCs targeting. A biantennary N-glycan **G0** and its analogous O-2 core xylosylated N-glycan **XG0** were synthesized, covalently conjugated to the model antigen ovalbumin (OVA) and analyzed for binding to a set of murine CLR-Fc fusion proteins using lectin microarray technology. We further evaluated if the differential binding of **G0** and **XG0** to murine SignR3 receptor has an impact on dendritic cell targeting. First, the uptake of fluorescently labeled glycoconjugates by DCs was followed by confocal microscopy and second, effects of the ovalbumin glycoconjugates on T cell activation were measured in a DC/T cell co-cultivation assay. Our results showed that both xylosylated and non-xylosylated biantennary N-glycans enhanced the uptake and presentation of OVA and, more importantly, G0-OVA conjugates were internalized better than the other two. This demonstrates that even small structural glycan modifications can have a marked impact on uptake by DCs and subsequent T cell activation. Knowing that the G0-OVA glycoconjugates enhanced the uptake by DCs, in the second part of this Chapter, we described the preparation and characterization of OVA-G0 AuNCs that combine glycan mediated targeting, the immunogenic properties of OVA protein and imaging properties of fluorescent gold clusters in a single entity. The preparation and characterization of AuNCs was optimized using unconjugated OVA protein as a reducing and stabilizing agent. Later, the optimized conditions were applied to G0-OVA-AuNCs generation. We have subsequently synthesized three other OVA-glycan-derived AuNCs, applying XG0-OVA and G0-OVA with different sugar valency. We did not observe an influence of the glycans on the measured optical properties of the clusters. Finally, we had successfully applied G0-OVA-AuNCs as a fluorescent probe in an *in vitro* uptake study utilizing murine DCs.

Based on the C-type lectin receptors specifically recognizing certain glycan ligands, glycan modified AuNCs have a big potential as an efficient and specific targeting fluorescent probe

for *in vitro* and *in vivo* imaging. Moreover, they can find an application as a sensing probes for glycan binding proteins, including plant lectins but also other biologically relevant molecules.

4.12 Experimental Part

General Methods

Materials: LPS-free ovalbumin was purchased from Hyglos (Bernried, Germany). Gold (III) chloride trihydrate and sodium hydroxide were purchased from Sigma Aldrich. Glycoconjugates were labeled with Alexa Fluor® 647 NHS ester from Thermo Fisher Scientific. Dialysis of glycoconjugates was performed on Slide-Z-lyser ™ dialysis cassettes (10K) from Thermo Fisher Scientific. All aqueous solutions were prepared with nanopure water produced won a Diamond UV water purification system (Branstead International, IA, Madrid, Spain). Murine-hFc C-type lectins were construct and expressed in the laboratory of Prof. Dr. Bernd Lepenies as previously described. [23]

MALDI-TOF mass spectrometry was controlled by FlexControl 3.3 software (Bruker Daltonics). Lectin microarrays were prepared using NHS activated glass slides, Nexterion® H (Schott AG, Mainz, Germany). Lectin microarrays were printed employing a robotic non contact piezoelectric SciFLEXARRAYER spotter S11 (Scienion, Berlin, Germany). Lectin microarrays were compartmentalized with 16 Well ProPlate Module 7 x 16 mm Well from Electron Microscopy Sciences. Fluorescence was measured using an Agilent G265BA microarray scanner system (Agilent Technologies) and quantification was performed with ProScanArray Express software (Perkin Elmer) with an adaptive circle quantification method (spot diameter: 50-300 µm). Histograms represent median RFU values with local background subtraction.

IMDM medium was supplemented with 2 mM L-glutamine, 100 U/mL penicillin, 100-µg/mL streptomycin and 10% fetal calf serum (FCS) purchased from PAN Biotech. Dendritic cells

(CD11c+ cells) were isolated from a suspension of C57BL/6 murine spleen cells by magnetic-activated cell sorting using CD11c+ MicroBeads (Miltenyi, Bergisch-Gladbach, Germany) and OT-II T cells using the Pan T Cell Isolation Kit II (Miltenyi Bergisch-Gladbach, Germany). Monoclonal antibodies against CD11c (APC), CD3e (APC), CD19 (PE), CD4 (FITC), CD8 (APC-H7) were purchased from eBioscience. Fluorochrome-conjugated monoclonal antibodies against CD69 (PE), CD4 (Alexa Fluor 700) were purchased from eBioscience and BD Pharmingen (Cowley, UK), respectively. The flow cytometric analysis was performed on a FACS Canto II flow cytometer (BD Biosciences, Heidelberg, Germany). Data were analyzed using FlowJo analysis software (Tree Star, Ashland, OR). Fluorescent images were taken using a Zeiss LSM 510 laser scanning confocal microscope (Carl Zeiss, Oberkochen, Germany) equipped with a red helium-neon (HeNe633) laser and UV laser (364 nm) and a 63×oil immersion objective. The measurement of cytokines IL-2 and IFN-γ in the supernatants after 48 and 72h was performed according to the manufacturer's instructions (PeproTech, Hamburg, Germany). Plates were developed with a solution of tetramethyl benzidine (Ultra TMB-Elisa Substrate Solution, Thermo Scientific). After a few minutes, 2M sulfuric acid solution was added to stop the reaction. Absorbance was measured at 450 nm using an ELISA reader (Tecan).

Microwave-assisted synthesis of gold nanoclusters was carried out in a Biotage Initiator microwave reactor with microwave reaction vials of a total volume 0.2-0.5 mL. Fluorescence spectra of OVA-AuNCs and OVA-glycan-AuNCs were measured on a Varioskan Flash (Thermo Scientific) operating with a ScanIt Software, using NUNC MaxiSorp Polystyrene Black 96 or 384 Well plates. UV-visible spectra were obtained on a NanoDrop ND-1000 spectrophotometer. Circular Dichroism was recorded using Jasco J815 spectrometer.

XPS Surface analysis by XPS was performed in a SPECS SAGE HR 100 system spectrometer in an ultrahigh vacuum (UHV) chamber. The X-ray sources employed for this analysis were an non monochromatic Mg Kα (1253.6 eV) and 250 W or Al Kα operated at 1.25 kV and 300W,

calibrated using the $3d_{5/2}$ line of Ag with a full width at half maximum (FWHM) of 1.1 eV. The take- off angle was fixed at 90º and the analysis was conducted at a pressure of ~10^{-6} Pa. Surfaces were brought into the XPS chamber within 5 min after cleaning/preparation. The selected resolution for the spectra was 30 eV of Pass Energy and 0.5 eV/step for the general survey spectra and 15 eV of Pass Energy and 0.15 eV/step for the detailed spectra of the different elements. Spectra were analyzed with the CasaXPS 2.3.15dev87 software. The analysis consisted of satellite removal, Shirley background subtraction, calibration of the binding energies related to the C 1s C-C peak at 285 eV, and asymmetric peak fitting with Gaussian-Lorentizan line shapes where the FWHM of all the peaks were constrained while the peak positions and areas were set free. Samples are prepared by dehydration on clean titania substrates where titania is selected as a substrate to avoid any possible overlap of gold signal with signal of silica coming from wafer or glass. $HAuCl_4$ salt is measured as received in custom made powder holder.

TEM imaging was conducted on a JOEL JEM-2100F field emission TEM with an accelerating voltage of 200 kV. AuNCs solutions were drop-cast onto copper TEM grids coated with ultrathin carbon support (Ted Pella, Redding, USA). The average gold nanoclusters diameter was quantify using ImageJ (Java). **Dynamic light scattering** (DLS) was performed at fixed scattering angle of 90º on a Malvern Zetasizer Nano ZS applying 0.1M NaCl and solutions were sonicated before measurement.

Mice

Dendritic cells isolated from C57BL/6 mice and T cells isolated from OT-II transgenic mice (own breeding at the Federal Institute of Risk Assessment (BfR), Berlin, Germany) were used in all experiments except for the confocal microscopy studies where dendritic cells isolated from C57BL/6J mice (CIC biomaGUNE, San Sebastian, Spain) were used.

Synthesis of ovalbumin-glycoconjugates

G0-OVA: Disuccinimidyl suberate (DSS) linker (6.8 mg, 18.46 µmol, 13 eq.) was dissolved in 1 mL of DMSO (18.4 mM) and activated with 15 µL of triethylamine. The biantennary N-glycan (2 mg, 1.42 µmol) was dissolved in 0.3 mL of DMSO and added dropwise to the linker solution. The reaction was stirred at RT for 90 min. The product was extracted with 800 µL PBS (pH 7.4), washed 3 times with chloroform and centrifuged at 3000g to facilitate phase separation. Attachment of DSS linker to N-glycan was confirmed by MALDI-TOF mass spectrometry. G0-DSS (40 eq.) was added to 1.25 mg of OVA and the reaction was incubated overnight at RT. Sample was concentrated in a 10K Amicon filter device and dissolved in sterile PBS.

XG0-OVA: DSS linker (7.4 mg, 20 µmol, 17 eq) was dissolved in 1.1 mL of DMSO (22 mM) and activated with 15 µL of triethylamine. The xylosylated biantennary N-glycan (2.15 mg, 1.2 µmol) was dissolved in 0.3 mL of DMSO and added dropwise to solution of linker. The reaction was stirred at RT for 90 min. The product was extracted with 800 µL PBS (pH 7.4), washed three times with chloroform and centrifuged at 3000g. The attachment of DSS linker to N-glycan XG0 was confirmed by MALDI-TOF mass spectrometry. XG0-DSS (40 eq.) was subsequently added to 1.25 mg of OVA and reaction was left stirring overnight at RT. The sample was concentrated in a 10 kDa Amicon filter device and dissolved in sterile PBS.

Alexa Fluor®-647 labeling of ovalbumin and ovalbumin glycoconjugates

400 µL of 10 [mg/mL] solutions of OVA, 70 µL of 10 [mg/mL] XG0-OVA and 86 µL of 10 [mg/mL] G0-OVA in sodium bicarbonate buffer (100 mM, pH=9) were prepared. Subsequently 50, 12 and 15 µL of 10 [mg/mL] Alexa Fluor®-647 NHS ester were added and reaction mixtures were gently shaken at RT for a 1h (ovalbumin) or 2h (OVA-glycoconjugates). Excess dye was removed by dialysis (10k) over water (12h) and glycoconjugates were concentrated by ultracentrifugation with 10kDa Amicon devices. The

ratio of Alexa Fluor®-647 to protein (degree of substitution, DOS) was estimated based on the absorbance of dye (ε_{dye}, 250 000 $cm^{-1}M^{-1}$) and OVA (ε_{OVA}, 35 900 $cm^{-1}M^{-1}$) and calculated based on the equation:

$DOS = \frac{[Dye]}{[OVA]}$, where

$[Dye] = \frac{(A_{Max} * dilution\ factor)}{\varepsilon_{dye}}$ (molar concentration of Alexa Fluor®-647)

$[OVA] = \frac{(A_{280} - 0.05 * A_{Max}) dilution\ factor}{\varepsilon_{OVA}}$ (molar concentration of OVA or glycoproteins)

Calculation of the DOS of OVA-647, G0-OVA-647 and XG0-OVA-647:

	A_{max}	A_{280}	*[Dye]*	*[OVA]*	*DOS*
OVA	7.98	0.48	6.38e-5	5.28e-6	**12.1**
XG0-OVA	8.52	0.52	6.82e-5	6.12e-6	**11.1**
G0-OVA	8.88	0.54	7.10e-5	6.25e-6	**11.3**

Preparation of CLR microarray

NHS-activated glass slides were incubated overnight with protein G solution (20 µg/mL) in PBS at 4°C. Each slide was washed three times with PBST (PBS containing 0.05% Tween-20), followed by PBS and water. Slides were dried in a slide spinner. CLR-Fc solutions were prepared by dilution with printing buffer (PBS, 0.01% Tween-20 and 0.1% glycerol) to a final concentration of 0.2 mg/mL and placed into a 384 well source plate (Scienion, Berlin, Germany). 0.6 nL of CLR-Fc solutions were spatially arrayed onto protein G functionalized glass slides in five replicate spots per C-type lectin. After printing, the slides were placed in a 75 % humidity chamber (saturated NaCl solution) at 4°C for 2 hours. To reduce unspecific

binding, the slides were incubated with bovine serum albumin solution (3% BSA in PBS) for 1h followed by the standard washing procedures with PBST, PBS and water.

Incubation of microarrays with OVA and OVA glycoconjugates

CLR microarrays were incubated with solutions of Alexa Fluor®-647 labeled OVA and glycoconjugates (XG0-OVA and G0-OVA) at a final concentration of 50 µg/mL in binding buffer (PBS, 0.01% Tween-20, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 1% BSA). After overnight incubation at 4˚C, arrays were washed with PBS, water and dried in a slide spinner. Fluorescence was measured in a microarray scanner.

Isolation of spleen cells subset

To isolate splenocytes, spleens were flushed with IMDM medium supplemented with 2 mM L-glutamine, 100 U/mL penicillin, 100-µg/mL streptomycin and 10% fetal calf serum (FCS). The cell suspension was kept cold and filtered through a 40 µM cell strainer to remove cell aggregates. After centrifugation (300 g, 5 min, 4°C), cell pellets were resuspended in 5 mL of freshly prepared erythrocyte lysis buffer (10% 100 mM Tris pH 7.5 + 90% 160 mM NH_4Cl), mixed gently and incubated at RT for 2 min. Cells were washed twice in complete IMDM medium and centrifuged before re-suspension in MACS buffer (PBS, 0.5% BSA, 2mM EDTA). Dendritic cells (CD11c+ cells) were isolated from suspension of C57BL/6 murine spleen cells by magnetic-activated cell sorting using CD11c+ functionalized MicroBeads. Cells incubated with magnetic microbeads were loaded on a MACS column placed in a magnetic field. Unbound cells passed through the column, while remaining CD11c+ cells were washed with MACS buffer and eluted from the column. To increase DC purity, the CD11c+ cell purification was repeated. The cell suspension was centrifuged, resuspended in IMDM complete medium and counted. T cells were isolated from spleen of OT-II transgenic mice using the Pan T Cell Isolation Kit II. Non-target cells were magnetically labeled with a cocktail of biotin-conjugated

antibodies against CD11b, CD11c, CD19, CD45R (B220), CD49b (DX5), CD105, Anti-MHC-class II, and Ter-119 markers followed by anti-biotin monoclonal antibodies conjugated to MicroBeads. Non-targeted cells were retained on the MACS column, while the unlabeled T cells passed through the column in the first fraction. To increase T cell purity, the purification procedure was repeated. Isolated T cells were subsequently centrifuged, resuspended in IMDM complete medium and counted. The purity of DCs and T cells isolated by MACS was analyzed by flow cytometry. Cells were pre-incubated with CD16/32 (Fc-Block, dilution 1:100) blocking antibody in 100 µL FACS buffer (PBS, 0.5% BSA, 2mM EDTA) at 4°C for 10 min. Subsequently, purified DCs and flow through were incubated with CD11c (APC) monoclonal antibody (1:100 dilution) for 30 min at 4°C in the dark. Purified T cells were incubated with anti-CD3e (APC) and anti-CD19 (PE) antibodies (1:100 dilution) for 30 min at 4°C in the dark. As a control, unstained samples were also prepared. Cells were washed with 1 mL of FACS buffer re-suspend in 100 µL FACS buffer and vortex before each measurement.

Uptake of labeled glycoconjugates by DCs followed by confocal fluorescence microscopy

DCs from C57BL/6J mice ($2x10^6$ cells) were seeded on poly-D-lysine-coated glass cover-slips and incubated at 5% CO_2 at 37ºC overnight. Alexa Fluor®-647 labeled OVA, XG0-OVA and G0-OVA were added to cells (10 µg/mL) and incubated for 10 min at 37°C. After incubation, cells were carefully washed with cooled PBS and fixed with 3% paraformaldehyde at RT for 20 min. After washing with PBS and water, the cover-slips were mounted on slides using Vectashield® mounting medium (Vector Labs, Burlingame, USA) and the fluorescent images were taken.

Uptake of labeled glycoconjugates by DCs followed by flow cytometry

Purified DCs were seeded in 96-well round bottom plates ($2x10^4$ cells/well) in complete IMDM medium. After 30 min (37°C, 5% CO_2), cells were pulsed with Alexa Fluor®-647 labeled

OVA, XG0-OVA and G0-OVA (10 µg/mL). After 10 min of incubation (37°C, 5% CO_2), cells were carefully washed with cold PBS (2x) and resuspended in 100 µL FACS buffer containing CD16/32 (Fc-Block, dilution 1:100) blocking antibody at 4°C for 10 min. Subsequently, a control sample was incubated with CD11c (APC) monoclonal antibody (1:100 dilution) for 30 min at 4°C in the dark. As a control, unstained sample was also prepared. Cells were washed with 1 mL of FACS buffer, resuspended in 100 µL FACS buffer and vortexed before each measurement.

Co-cultivation of OT-II T cells with OVA-glycoconjugate pulsed DCs

Purified DCs were seeded in 96-well round bottom plates ($2x10^4$ cells/well) in complete IMDM medium. After 30 min (37°C, 5% CO_2), cells were pulsed with OVA, XG0-OVA and G0-OVA (30 µg/mL). After 1h of incubation, 60 µL of purified T cells ($7x10^4$ cells/well) were added and incubated at 37°C, 5% CO_2 for 48h. The expression of the early T cell activation marker CD69 and cytokine levels (IL-2, IFN-γ) in the supernatant were determined after 48 and 72 hours of stimulation.

Detection of CD69 by flow cytometry

Expression of the T cell activation marker CD69 was analyzed by flow cytometry 48 and 72 hours after stimulation. Fc receptors were blocked using anti-CD16/32 (dilution 1:100) in 100 µL FACS buffer at 4°C for 10 min. Subsequently, cells were incubated with anti-CD4-AF700 and anti-CD69-PE (1:100 dilution) for 30 min at 4°C in the dark. Cells were washed with 1 mL FACS buffer and analyzed in 100 µL FACS buffer. Data was analyzed using FlowJo analysis software.

Optimized synthesis of OVA-AuNCs and glycan-OVA-AuNCs

General procedure for OVA-AuNCs synthesis:

OVA-AuNCs: To a solution of OVA protein (15 mg/mL, 124 µL) in a microwave vial (0.1-0.5 mL) 0.1 M $HAuCl_4 \cdot 3H_2O$ (5.4 µL, 4.2 mM) was added under vigorous stirring. After 5-10 minutes, 1M NaOH was added (24 µL, 150mM) and the reaction mixture was heated under microwave irradiation for one cycle of 6 minutes at 100ºC. After the microwave heating reaction mixture turns pale brown. The solution of OVA-AuNCs was subsequently dialysed against water and freeze dried until use.

G0-OVA-AuNCs were generated following the general procedure using a solution of G0-OVA glycoprotein (13.5 mg/mL, 78 µL), 0.1M $HAuCl_4 \cdot 3H_2O$ (2.6 µL, 4.2 mM) and 1M NaOH (15 µL, 150mM).

XG0-OVA-AuNCs were generated following the general procedure using a solution of XG0-OVA glycoprotein (13.9 mg/mL, 98 µL), 0.1M $HAuCl_4 \cdot 3H_2O$ (4 µL, 4.2 mM) and 1M NaOH (19 µL, 150mM).

G0-OVA-AuNCs (2-3) were generated following the general procedure using a solution of G0-OVA (2-3) glycoprotein (15 mg/mL, 25 µL), 0.1M $HAuCl_4 \cdot 3H_2O$ (1 µL, 4.2 mM) and 1M NaOH (4.8 µL, 150mM).

G0-OVA-AuNCs (5-6) were generated following the general procedure using a solution of G0-OVA (5-6) glycoprotein (15 mg/mL, 37 µL), 0.1M $HAuCl_4 \cdot 3H_2O$ (1.5 µL, 4.2 mM) and 1M NaOH (7.1 µL, 150mM).

Uptake of OVA-AuNCs and glycan-OVA-AuNCs by DCs followed by confocal fluorescence microscopy

DCs from C57BL/6J mice ($2x10^6$ cells) were seeded on poly-D-lysine-coated glass cover-slips overnight. OVA-AuNCs, XG0-OVA-AuNCs and G0-OVA-AuNCs were added to cells (150 µg/mL) and incubated for 40 min at 37°C. After incubation, cells were carefully washed with cooled PBS and fixed with 3% paraformaldehyde at RT for 20 min. After washing with PBS and water, the cover-slips were mounted on slides using Vectashield® mounting medium. Fluorescent images were taken using confocal microscope equipped with a UV laser (365 nm) and an 63× oil immersion objective.

References

(1) Banchereau, J.; Palucka, A. K. *Nature Reviews Immunology* **2005**, *5*, 296–306.

(2) Figdor, C. G.; de Vries, I. J. M.; Lesterhuis, W. J.; Melief, C. J. *Nature medicine* **2004**, *10*, 475–480.

(3) Shortman, K.; Lahoud, M. H.; Caminschi, I. *Experimental & molecular medicine* **2009**, *41*, 61–66.

(4) Kreutz, M.; Tacken, P. J.; Figdor, C. G. *Blood* **2013**, *121*, 2836–2844.

(5) Caminschi, I.; Lahoud, M. H.; Shortman, K. *European journal of immunology* **2009**, *39*, 931–938.

(6) Fong, L.; Brockstedt, D.; Benike, C.; Wu, L.; Engleman, E. G. *The Journal of Immunology* **2001**, *166*, 4254–4259.

(7) Lu, W.; Arraes, L. C.; Ferreira, W. T.; Andrieu, J.-M. *Nature medicine* **2004**, *10*, 1359–1365.

(8) Lepenies, B.; Lee, J.; Sonkaria, S. *Advanced drug delivery reviews* **2013**, *65*, 1271–1281.

(9) Figdor, C. G.; van Kooyk, Y.; Adema, G. J. *Nature Reviews Immunology* **2002**, *2*, 77–84.

(10) Drickamer, K.; Taylor, M. E. *Current opinion in structural biology* **2015**, *34*, 26–34.

(11) Pereira, C. F.; Torensma, R.; Hebeda, K.; Kretz-Rommel, A.; Faas, S. J.; Figdor, C. G.; Adema, G. J. *Journal of Immunotherapy* **2007**, *30*, 705–714.

(12) Kreutz, M.; Giquel, B.; Hu, Q.; Abuknesha, R.; Uematsu, S.; Akira, S.; Nestle, F. O.; Diebold, S. S. *PloS one* **2012**, *7*, e40208.

(13) Singh, S. K.; Stephani, J.; Schaefer, M.; Kalay, H.; Garcia-Vallejo, J. J.; den Haan, J.; Saeland, E.; Sparwasser, T.; van Kooyk, Y. *Molecular immunology* **2009**, *47*, 164–174.

(14) Van Kooyk, Y.; Unger, W. W.; Fehres, C. M.; Kalay, H.; Garcia-Vallejo, J. J. *Molecular immunology* **2013**, *55*, 143–145.

(15) Lin, J. H. *Current drug metabolism* **2009**, *10*, 661–691.

(16) Brzezicka, K.; Echeverria, B.; Serna, S.; van Diepen, A.; Hokke, C. H.; Reichardt, N.-C. *ACS Chemical Biology* **2015**, *10*, 1290–1302.

(17) Serna, S.; Kardak, B.; Reichardt, N. C.; Martin-Lomas, M. *Tetrahedron: Asymmetry* **2009**, *20*, 851–856.

(18) Garcia-Vallejo, J. J.; van Kooyk, Y. *Trends in immunology* **2013**, *34*, 482–486.

(19) Zhang, F.; Ren, S.; Zuo, Y. *International reviews of immunology* **2014**, *33*, 54–66.

(20) Cambi, A.; Koopman, M.; Figdor, C. G. *Cellular microbiology* **2005**, *7*, 481–488.

(21) Xiaochao, Q.; Yichen, L.; Lei, L.; Yanran, W.; Jingning, L.; Jimin, L. *Journal of Nanomaterials* **2015**, *2015*, 1–24.

(22) Eriksson, M.; Serna, S.; Maglinao, M.; Schlegel, M. K.; Seeberger, P. H.; Reichardt, N.-C.; Lepenies, B. *Chembiochem* **2014**, *15*, 844–851.

(23) Maglinao, M.; Eriksson, M.; Schlegel, M. K.; Zimmermann, S.; Johannssen, T.; Götze, S.; Seeberger, P. H.; Lepenies, B. *Journal of Controlled Release* **2014**, *175*, 36–42.

(24) Guo, S.-L.; Chen, P.-C.; Chen, M.-S.; Cheng, Y.-C.; Lin, J.-M.; Lee, H.-C.; Chen, C.-S. *PloS one* **2012**, *7*, e51370.

(25) Takahara, K.; Arita, T.; Tokieda, S.; Shibata, N.; Okawa, Y.; Tateno, H.; Hirabayashi, J.; Inaba, K. *Infection and immunity* **2012**, *80*, 1699–1706.

(26) Harvey, D.; Wing, D.; Küster, B.; Wilson, I. *Journal of the American Society for Mass Spectrometry* **2000**, *11*, 564–571.

(27) Park, C. G.; Takahara, K.; Umemoto, E.; Yashima, Y.; Matsubara, K.; Matsuda, Y.; Clausen, B. E.; Inaba, K.; Steinman, R. M. *International immunology* **2001**, *13*, 1283–1290.

(28) Galustian, C.; Park, C. G.; Chai, W.; Kiso, M.; Bruening, S. A.; Kang, Y.-S.; Steinman, R. M.; Feizi, T. *International immunology* **2004**, *16*, 853–866.

(29) Kawauchi, Y.; Kuroda, Y.; Kojima, N. *Glycoconjugate journal* **2012**, *29*, 481–490.

(30) Lightfoot, Y. L.; Selle, K.; Yang, T.; Goh, Y. J.; Sahay, B.; Zadeh, M.; Owen, J. L.; Colliou, N.; Li, E.; Johannssen, T.; others. *The EMBO journal* **2015**, *34*, 881-895.

(31) Tanne, A.; Ma, B.; Boudou, F.; Tailleux, L.; Botella, H.; Badell, E.; Levillain, F.; Taylor, M. E.; Drickamer, K.; Nigou, J.; others. *The Journal of experimental medicine* **2009**, *206*, 2205–2220.

(32) Takahara, K.; Yashima, Y.; Omatsu, Y.; Yoshida, H.; Kimura, Y.; Kang, Y.-S.; Steinman, R. M.; Park, C. G.; Inaba, K. *International immunology* **2004**, *16*, 819–829.

(33) Eriksson, M.; Johannssen, T.; von Smolinski, D.; Gruber, A. D.; Seeberger, P. H.; Lepenies, B. *Frontiers in immunology* **2013**, *4*, 196.

(34) Biosciences, B. *Introduction to flow cytometry: A learning guide*; BD Biosciences: San Jose, 2000; 1–54.

(35) Unger, W. W.; van Kooyk, Y. *Current opinion in immunology* **2011**, *23*, 131–137.

(36) Barnden, M. J.; Allison, J.; Heath, W. R.; Carbone, F. R. *Immunology and cell biology* **1998**, *76*, 34–40.

(37) Testi, R.; D'Ambrosio, D.; De Maria, R.; Santoni, A. *Immunology Today* **1994**, *15*, 479–483.

(38) Ziegler, S. F.; Ramsdell, F.; Alderson, M. R. *Stem cells* **1994**, *12*, 456–465.

(39) Gaffen, S. L.; Liu, K. D. *Cytokine* **2004**, *28*, 109–123.

(40) Schroder, K.; Hertzog, P. J.; Ravasi, T.; Hume, D. A. *Journal of leukocyte biology* **2004**, *75*, 163–189.

(41) Geijtenbeek, T. B.; Gringhuis, S. I. *Nature Reviews Immunology* **2009**, *9*, 465–479.

(42) Gringhuis, S. I.; den Dunnen, J.; Litjens, M.; van der Vlist, M.; Geijtenbeek, T. B. *Nature immunology* **2009**, *10*, 1081–1088.

(43) Gringhuis, S. I.; Kaptein, T. M.; Wevers, B. A.; Mesman, A. W.; Geijtenbeek, T. B. *Nature communications* **2014**, *5*: 3898.

(44) Dambuza, I. M.; Brown, G. D. *Current opinion in immunology* **2015**, *32*, 21–27.

(45) Selvaprakash, K.; Chen, Y.-C. *Biosensors and Bioelectronics* **2014**, *61*, 88–94.

(46) Zhang, X.-D.; Luo, Z.; Chen, J.; Song, S.; Yuan, X.; Shen, X.; Wang, H.; Sun, Y.; Gao, K.; Zhang, L.; others. *Scientific reports* **2015**, *5*, 8669.

(47) Dreaden, E. C.; Alkilany, A. M.; Huang, X.; Murphy, C. J.; El-Sayed, M. A. *Chemical Society Reviews* **2012**, *41*, 2740–2779.

(48) Goodman, P. *Gold Bulletin* **2002**, *35*, 21–26.

(49) Zheng, J.; Nicovich, P. R.; Dickson, R. M. *Annual review of physical chemistry* **2007**, *58*, 409.

(50) Wang, F.; Tan, W. B.; Zhang, Y.; Fan, X.; Wang, M. *Nanotechnology* **2005**, *17*, 1-13.

(51) Demchenko, A. P. *Luminescent Metal Nanoclusters*. In *Advanced fluorescence reporters in chemistry and biology II: Fundamentals and molecular design*; Springer Science & Business Media: New York, 2010.

(52) Resch-Genger, U.; Grabolle, M.; Cavaliere-Jaricot, S.; Nitschke, R.; Nann, T. *Nature methods* **2008**, *5*, 763–775.

(53) Panchuk-Voloshina, N.; Haugland, R. P.; Bishop-Stewart, J.; Bhalgat, M. K.; Millard, P. J.; Mao, F.; Leung, W.-Y.; Haugland, R. P. *Journal of Histochemistry & Cytochemistry* **1999**, *47*, 1179–1188.

(54) Berlier, J. E.; Rothe, A.; Buller, G.; Bradford, J.; Gray, D. R.; Filanoski, B. J.; Telford, W. G.; Yue, S.; Liu, J.; Cheung, C.-Y.; others. *Journal of Histochemistry & Cytochemistry* **2003**, *51*, 1699–1712.

(55) Santra, S.; Wang, K.; Tapec, R.; Tan, W. *Journal of biomedical optics* **2001**, *6*, 160–166.

(56) Santra, S.; Zhang, P.; Wang, K.; Tapec, R.; Tan, W. *Analytical chemistry* **2001**, *73*, 4988–4993.

(57) Xie, J.; Zheng, Y.; Ying, J. Y. *Journal of the American Chemical Society* **2009**, *131*, 888–889.

(58) Jin, L.; Shang, L.; Guo, S.; Fang, Y.; Wen, D.; Wang, L.; Yin, J.; Dong, S. *Biosensors and Bioelectronics* **2011**, *26*, 1965–1969.

(59) Zhang, P.; Yang, X. X.; Wang, Y.; Zhao, N. W.; Huang, C. Z.; others. *Nanoscale* **2014**, *6*, 2261–2269.

(60) Joseph, D.; Geckeler, K. E. *Colloids and Surfaces B: Biointerfaces* **2014**, *115*, 46–50.

(61) Chevrier, D. M.; Chatt, A.; Zhang, P. *Journal of Nanophotonics* **2012**, *6*, 0645041-06450416.

(62) Le Guével, X.; Hötzer, B.; Jung, G.; Schneider, M. *Journal of Materials Chemistry* **2011**, *21*, 2974–2981.

(63) Yoshimoto, J.; Tanaka, N.; Inada, M.; Arakawa, R.; Kawasaki, H. *Chemistry Letters* **2014**, *43*, 793–795.

(64) Luo, Z.; Yuan, X.; Yu, Y.; Zhang, Q.; Leong, D. T.; Lee, J. Y.; Xie, J. *Journal of the American Chemical Society* **2012**, *134*, 16662–16670.

(65) Xie, J.; Lee, J. Y.; Wang, D. I.; Ting, Y. P. *ACS Nano* **2007**, *1*, 429–439.

(66) Si, S.; Mandal, T. K. *Chemistry-A European Journal* **2007**, *13*, 3160–3168.

(67) Kong, Y.; Chen, J.; Gao, F.; Brydson, R.; Johnson, B.; Heath, G.; Zhang, Y.; Wu, L.; Zhou, D. *Nanoscale* **2013**, *5*, 1009–1017.

(68) Shang, L.; Wang, Y.; Jiang, J.; Dong, S. *Langmuir* **2007**, *23*, 2714–2721.

(69) Wang, L.-L.; Qiao, J.; Liu, H.-H.; Hao, J.; Qi, L.; Zhou, X.-P.; Li, D.; Nie, Z.-X.; Mao, L.-Q. *Analytical chemistry* **2014**, *86*, 9758–9764.

(70) Kitagawa, H.; Kojima, N.; Nakajima, T. *J. Chem. Soc., Dalton Trans.* **1991**, *11*, 3121–3125.

(71) Techane, S.; Baer, D. R.; Castner, D. G. ACS *Analytical Chemistry* **2011**, *83*, 6704–6712.

(72) Gobbo, P.; Biesinger, M. C.; Workentin, M. S. *Chemical Communications* **2013**, *49*, 2831–2833.

(73) Gobbo, P.; Mossman, Z.; Nazemi, A.; Niaux, A.; Biesinger, M. C.; Gillies, E. R.; Workentin, M. S. *Journal of Materials Chemistry B* **2014**, *2*, 1764–1769.

(74) Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2005**, *1751*, 119–139.

(75) Greenfield, N. J. *Nature protocols* **2006**, *1*, 2876–2890.

(76) Couleaud, P.; Adan-Bermudez, S.; Aires, A.; Mejias, S. H.; Sot, B.; Somoza, A.; Cortajarena, A. L. *Biomacromolecules* **2015**, *16*, 3836–3844.

(77) Shakushiro, K.; Yamasaki, Y.; Nishikawa, M.; Takakura, Y. *Immunology* **2004**, *112*, 211–218.

FINAL CONCLUSIONS

Final Conclusions

Glycans together with nucleic acids, proteins and lipids are one of the four fundamental macromolecular components of all cells. The tremendous complexity of glycans reflects their multiple and diverse biological functions as well as evolutionary fight between the host and the invading pathogen. From all possible glycan functions, their role in host-pathogen interactions, by involvement in the host receptors binding and the modulation of the host immune responses were studied in this Thesis. We have focused our research on the development and use of glycan microarrays for studying the interaction of synthetic N-glycans predominately found in parasites and plants with plant lectins, C-type lectins receptors and antibodies from *S. mansoni* infected mice and individuals living in schistosomal endemic areas, respectively. Based on the results obtained in the microarray screening, we have also explored the DC targeting potential of selected N-glycans conjugated to the model protein ovalbumin and studied the influence of xylose modification on the uptake and further activation of T cells. We have finally prepared first glycan modified gold nanoclusters as an attractive novel system for imaging of glycoconjugates.

In the first part of this Thesis, we have demonstrated the chemoenzymatic synthesis of N-glycans and showed the efficiency of this strategy towards the generation of complex and well defined ligands. The use of recombinant glycosyltransferases and glycosidases limited the numbers of required synthetic scaffolds to six structures (**G1**-**G6**) and allowed for the preparation of 39 core β-1,2-xylosylated N-glycans decorated with α-1,3 and/or α-1,6 core fucose and terminal structural elements, including LN, LDN and Lewis X motifes.

In the second Chapter, we describe the construction of a focused, synthetic glycan microarray, containing 126 diverse glycans including core xylosylated structures

synthesized during this Thesis. This high throughput screening tool has allowed us to study carbohydrate-protein interactions using minimal amount of both, sugars and proteins in a relatively short time.

The first microarray-assisted screening was performed using plant and fungi lectins of known general specificity but this screening already revealed the great potential of our glycan arrays. We could observe the influence of core xylose on the interaction of N-glycans with mannose binding lectins such as, C*oncanavalin A* and *Galanthus nivalis*. The presence of core xylose reduced the interactions with ConA and completely abolished binding towards GNA. As this β-1,2-xylose is a common structure found on plant N-glycans, observed lack of binding could be rationalized by a tolerance mechanism avoiding interactions between plant lectins and self-glycans.

Additionally, we have analyzed the carbohydrate specificity of murine and human C-type lectins receptors using glycan microarrays. Here, we were mainly interested in host immune receptors which are involved in the interactions with glycoconjugates of the helminth parasite *S. mansoni*. We could clearly observe the binding of human MGL-1 towards LDN and LDNF glycan elements and DC-SIGN to set of Lewis X terminated glycan structures. Our results were with a good agreement to the previous reports describing MGL-1 and DC-SIGN as receptors able to bind *S. mansoni* soluble egg antigens (SEAs) expressing LDN, LDNF and Lewis X type of glycans. As a major result of the binding studies, however, we noted the difference in the binding of DC-SIGN receptor towards xylosylated and non-xylosylated biantennary N-glycans. Similar to the behavior of GNA lectin, the presence of core xylose completely abolished binding towards DC-SIGN on the glycan microarrays. Inspired by this finding, we have further constructed glycoconjugates bearing either xylosylated or non-xylosylated N-glycans and analyzed their DCs targeting properties *in vitro*. The study was explored in the Chapter 4.

It is important to mention, that besides MGL-1 and DC-SIGN, other analyzed C-type lectin receptors did not show any binding towards ligands present on the microarray. Especially Dectin-2 which is believed to play a role in the host-*S. mansoni* interactions and was found to recognize components of SEAs. This behavior might be explained either by the absence of a relevant ligand on the array or by insufficient binding strength perhaps due to non optimal presentation of glycans and protein analyzed.

In Chapter 3, we described the application of our glycan microarrays for the screening and characterization of anti-carbohydrate antibodies, as a complementary tool to shotgun arrays derived from natural glycans of different development stages of *S. mansoni*. The majority of antibodies generated during schistosomiasis are directed against antigenic parasite glycans. Therefore, their identification and characterization might be essential for the development of future carbohydrate-based vaccines or to understand the role of certain sugars in the course of parasitic infection. The protective immune responses may require action of multiple antibodies against wide range of glycan elements not tested so far. However, it is important to remember that high antibody titers against glycan elements do not certainly correlate with protection and some anti-carbohydrate Abs might block the protective functions of other antibodies or are simply part of the natural antibody repertoire of healthy individuals. Thus the comparison of the profiles of Ab responses coming from sera of infected, healthy and resistant individuals using the glycan arrays is believed to be a promising strategy towards the identification of important glycan antigens.
First, we have evaluated the specificity of group of monoclonal antibodies induced in *S. mansoni* infected mice and confirmed the functionality of synthetic glycan arrays towards screening of antibodies. We have observed the known specificities of certain MAbs towards Lewis X, LDN or LDNF glycan motifs highlighting their antigenic character during the *S. mansoni* infection. Moreover, taking advantage of the nature of the synthetic glycan array we could additionally draw conclusions about the importance of the presentation of these

elements required for antibody recognition. Certain Lewis X or LDN glycan elements could be recognized as a simple tri- or disaccharide structures directly coupled to the array surface while other in the context of larger N-glycan structures. Additionally, we established the minimal epitope for100-4G11 IgM MAb binding and identified new good ligand structures.

Next, we have recorded the anti-carbohydrate IgG and IgM profiles of healthy individuals with no previous history of *S. mansoni* infection. We observed the unique anti-carbohydrate fingerprints for each individual reflecting the natural antibodies and antibodies raised against recent infections circulating in the blood stream of each donor. More importantly for our study, healthy donors did not elicit antibodies against any of relevant N-glycans present on the array and could be further used as the negative controls for the screening of the sera of *S. mansoni* infected patients.

We have compared the anti-carbohydrate responses elicited from pooled sera of two age-groups (children and adults) coming from the *S. mansoni* endemic area. It has been demonstrated in the literature that children are far more susceptible for the infection and re-infection compared to adults, in who the acquisition of the protective immune responses towards parasite occurs however it takes many years of exposure, multiple infections and treatments to develop.

Applying the glycan microarrays, we have observed interesting differences in the profiles of anti-carbohydrate IgG antibody responses whereas the IgM did not significantly differ depending on the age-group. We found a more pronounced specific IgG binding pattern and important differentiation between children and adults based on core β-1,2-xylosylated and core α-1,3-fucosylated N-glycan structures, with the adult pool IgG antibodies bounding only to certain core β-1,2-xylosylated compounds and none of the core α-1,3-fucosylated structures that lack xylose. This observation highlighted the antigenic nature of both motifs as well as the value of our synthetic arrays to detect specific antibodies in different serum cohorts.

Discovered differences in the IgG profiles of children and adults may raise a question about the role of certain glycans in the modulation of immune responses and acting as a smoke screen during the parasitic infections. While some of the glycans are ignored by the mature immune system of adults, they can still induce the anti-carbohydrate responses in the younger age group. However, based only on the performed microarray-assisted screening it is not possible to take clear conclusions about the exact glycan structures involved in the modulation or the induction of protective immunity during parasitic infections. However, our findings identified important antigenic structures and more importantly have raised the interest in continuing with glycan array screening of serum this time with a focus on immunoglobulin isotype and subclass that can help to clarify the functional role of the detected antibodies.

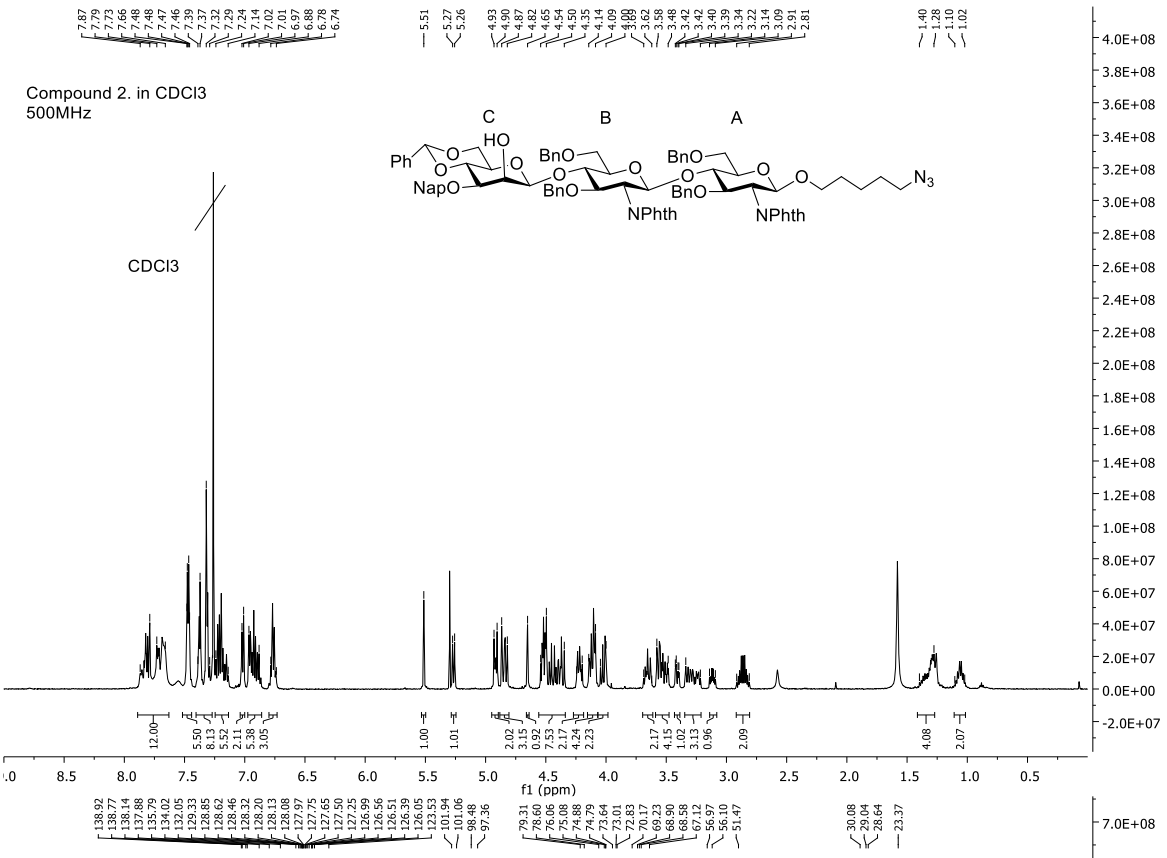
In Chapter 4, we explored the use of glycans as dendritic cells antigen targeting molecules utilizing the endocytic C-type lectins receptors expressed on the cells surface. Targeting antigens to dendritic cell subsets is a promising strategy which may enhance the efficacy of vaccines and further stimulate antigen presentation and subsequent T cell activation. Based on the difference in the recognition of xylosylated versus non-xylosylated biantennary compounds **G0** and **XG0** by the DC-SIGN assigned by the microarray-assisted study mentioned before, we have covalently conjugated both N-glycans to the model antigen ovalbumin (G0-OVA, XG0-OVA) and studied their targeting properties to murine DCs *in vitro*. An initial binding screening was performed using C-type lectin microarrays. Subsequently, we studied the binding of glycoconjugates constructed with fluorescently labeled OVA bearing average of 3-4 copies of non-xylosylated and xylosylated biantennary glycan G0. As result, we recorded interactions with several C-type lectins, including homologues of DC-SIGN, SignR1 and SignR3 previously not detected on glycan microarrays. These results highlighted the importance of the multivalent presentation of glycans and proteins in the strengthening of the relatively weak interactions between single molecules

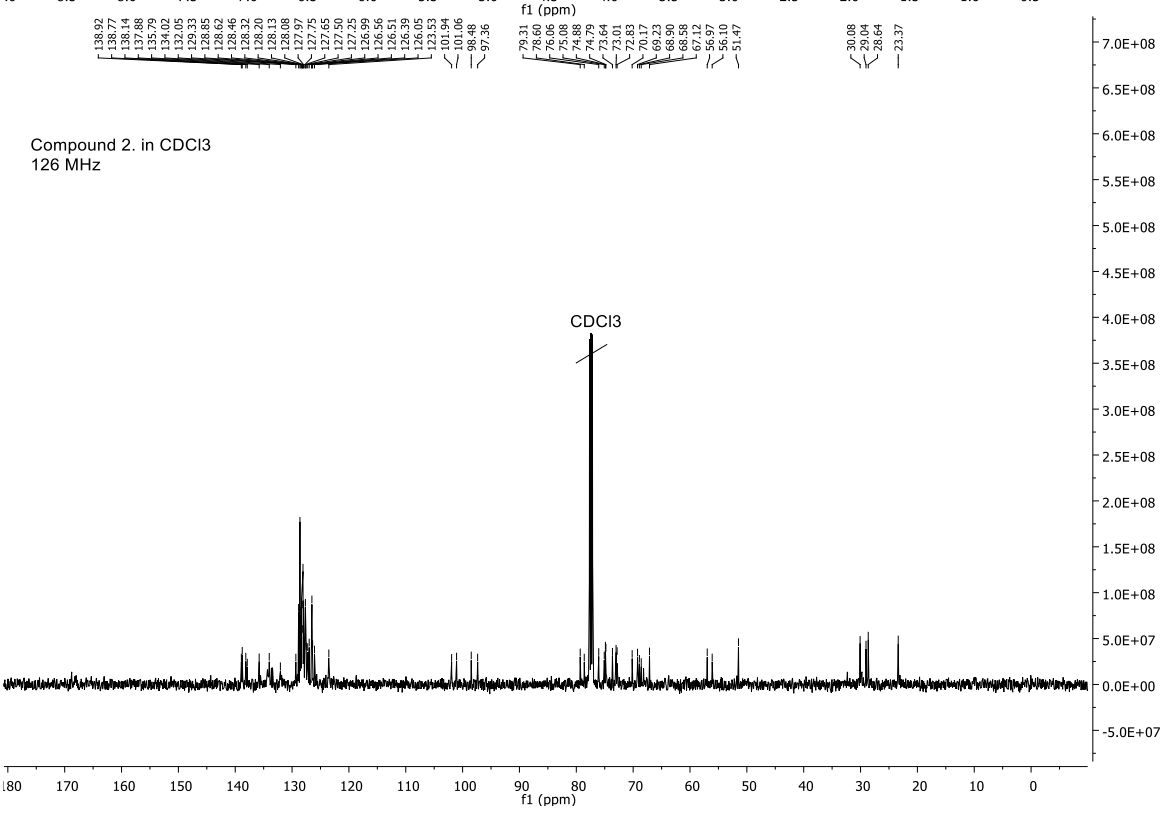
as well as importance of their special orientation and presentation. We also demonstrate the utility of CLR microarrays to identify natural glycan ligands.

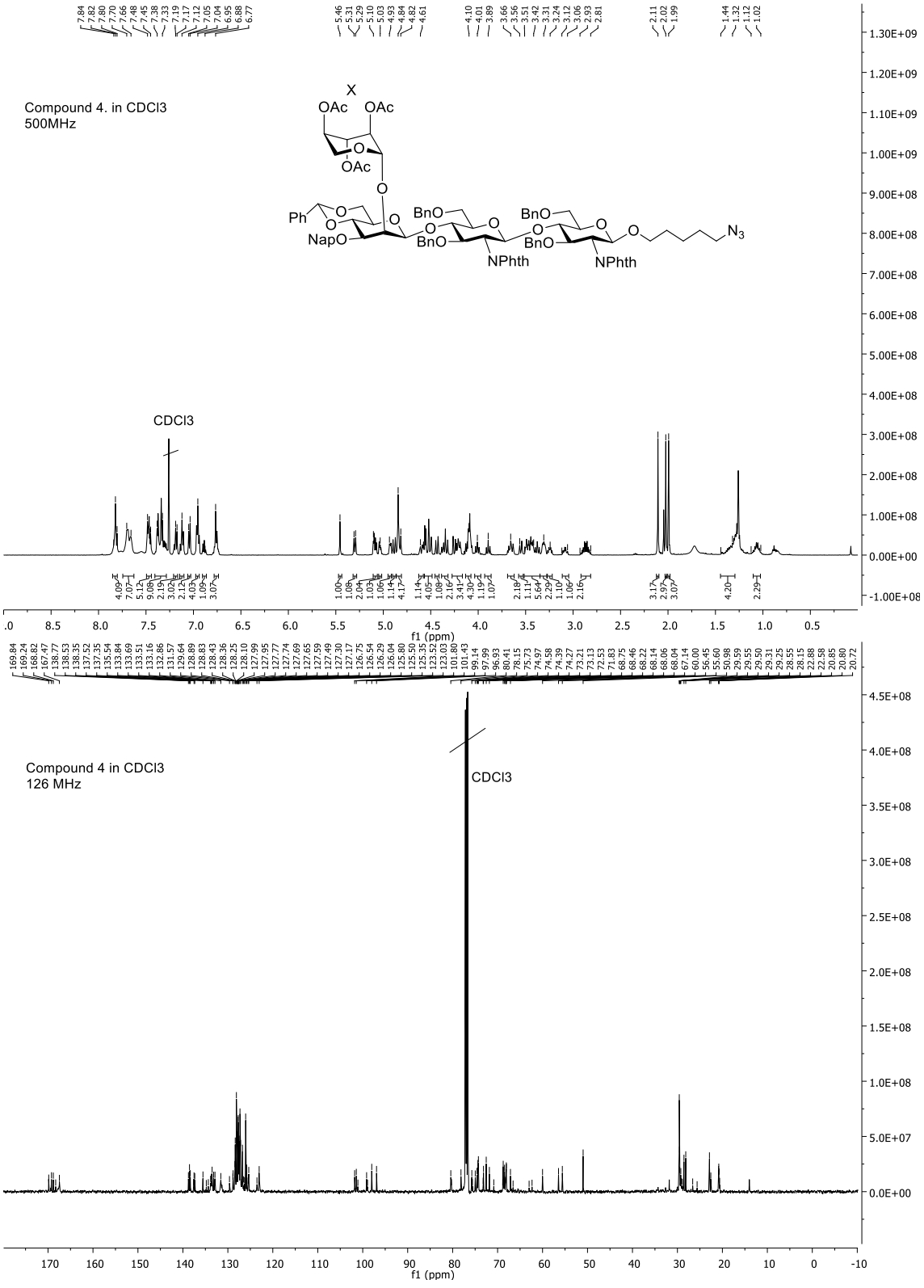
We demonstrated that both N-glycans, **G0** and **XG0** enhanced the uptake and presentation of OVA compared to unconjugated protein opening new possibilities for the specific delivery of vaccine antigens into DCs. Moreover, presence of the xylose modification had a marked impact on the uptake by DCs and subsequent T cell activation as demonstrated by CD69 expression and production of IL-2 and IFN-γ measured after 48h. In line with our glycan and lectin microarray binding studies, we found that non-xylosylated G0-OVA was preferably internalized by murine DCs and led to enhanced T cell activation. In the future, the use of this glycan could be explored as targeting molecule which in conjugation to model antigen structures could enhance the uptake of all construct by the internalization through the interaction of CLRs.

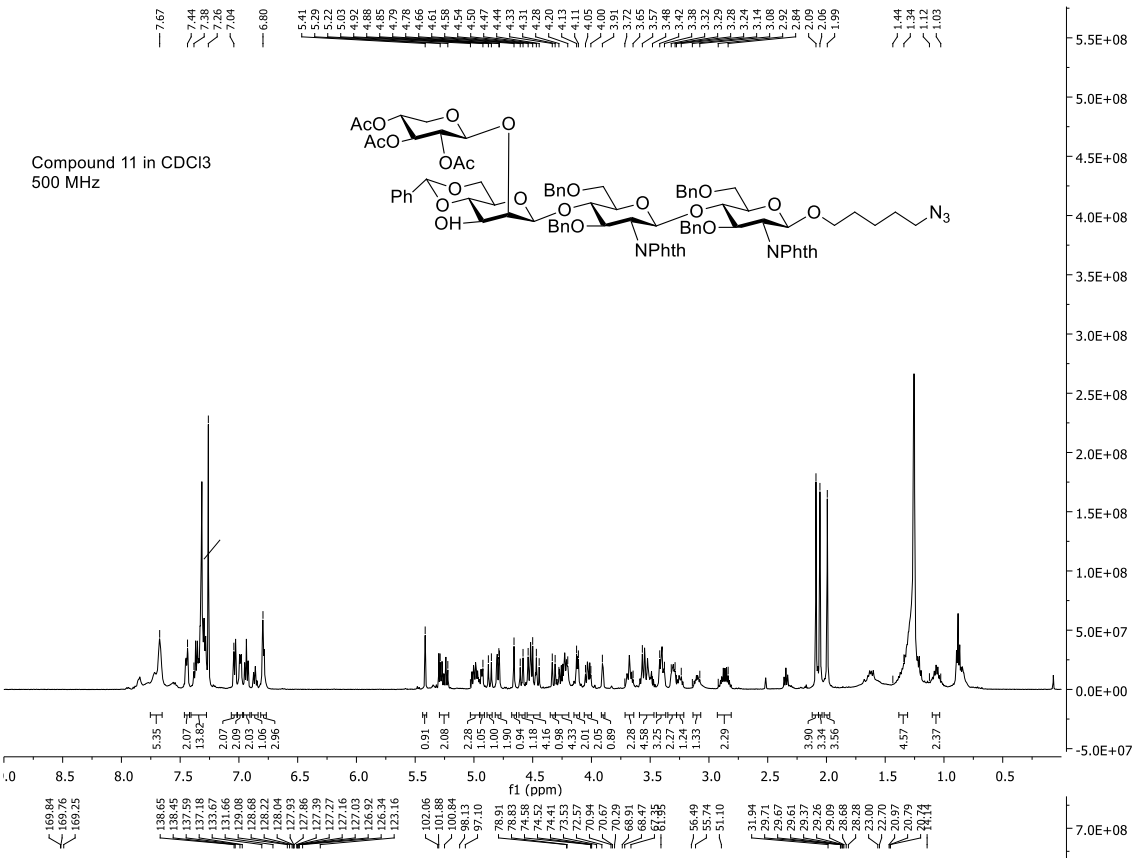
Knowing that **G0** N-glycan enhanced the uptake of glycoconjugate by DCs compared to the unconjugated protein or the XG0-OVA, in the last part of this Thesis we prepared gold nanoclusters utilizing G0-OVA as a stabilizing and reducing agent. Compared to the glycoconjugate alone, G0-OVA-AuNC combines targeting properties of glycan with the antigenic properties of OVA and the imaging properties of fluorescent gold clusters in a single construct. These glycoprotein gold nanoclusters are highly biocompatible and emit size dependent fluorescence in the red spectra. Additionally, the presence of the gold core opens the possibility of *in vivo* X-ray computed tomography (CT) imaging and radiotherapy. These nanostructures could be used as a promising candidates for dendritic cell targeting or as a specific fluorescent probes.

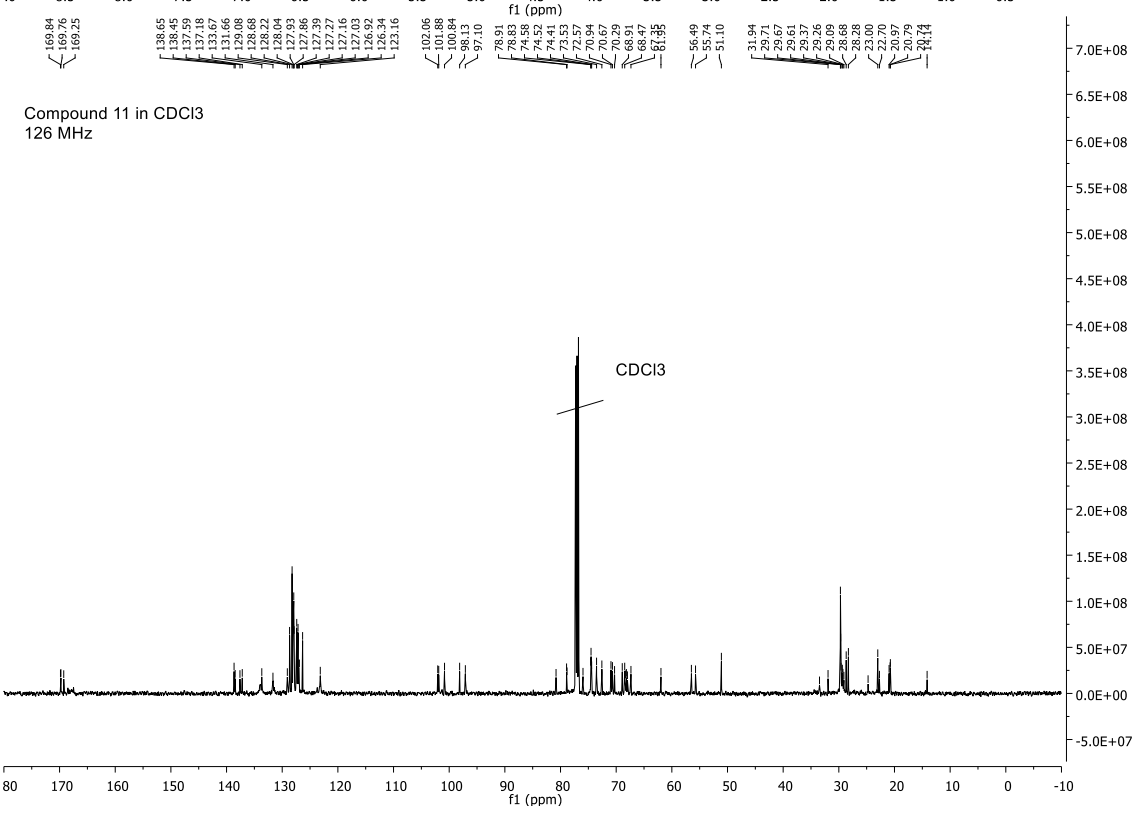
APPENDIX

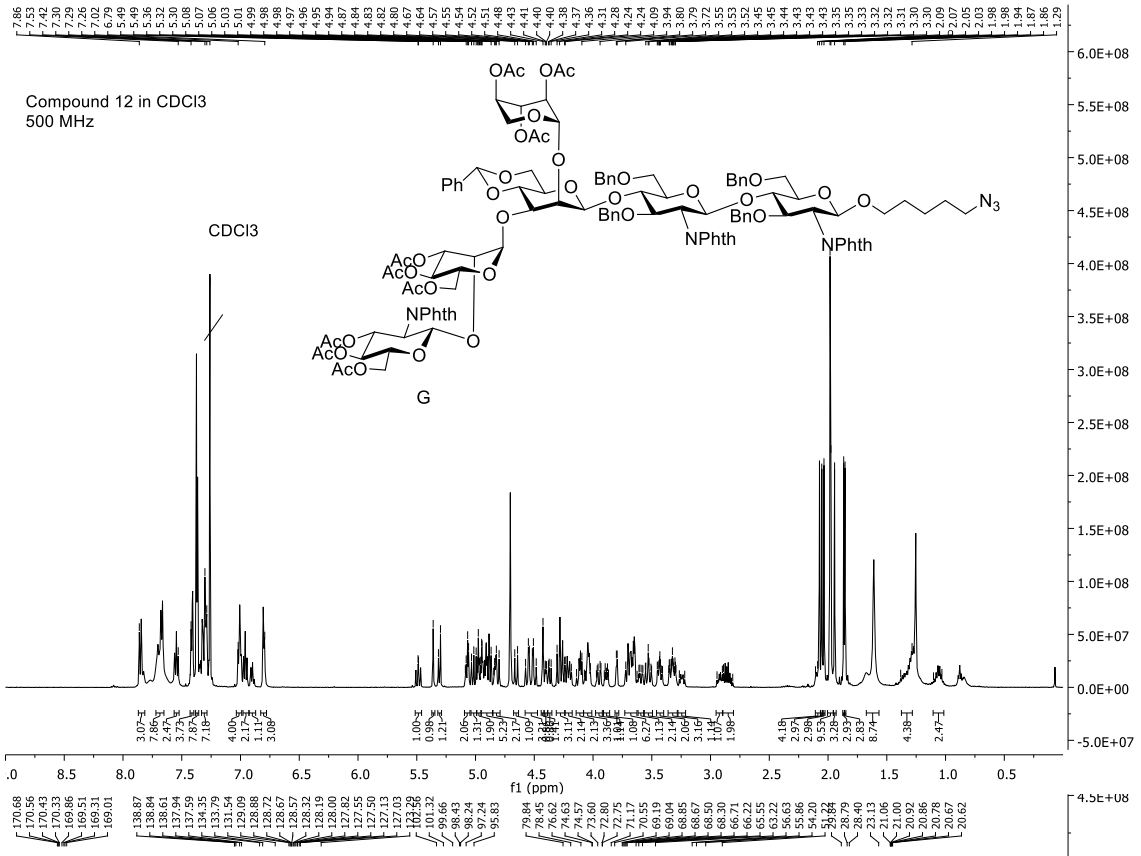

Compound 2. in CDCl3
500MHz
C
B
A
HO
Ph
O
NapO
BnO
BnO
BnO
BnO
NPhth
NPhth
N3
CDCl3
f1 (ppm)

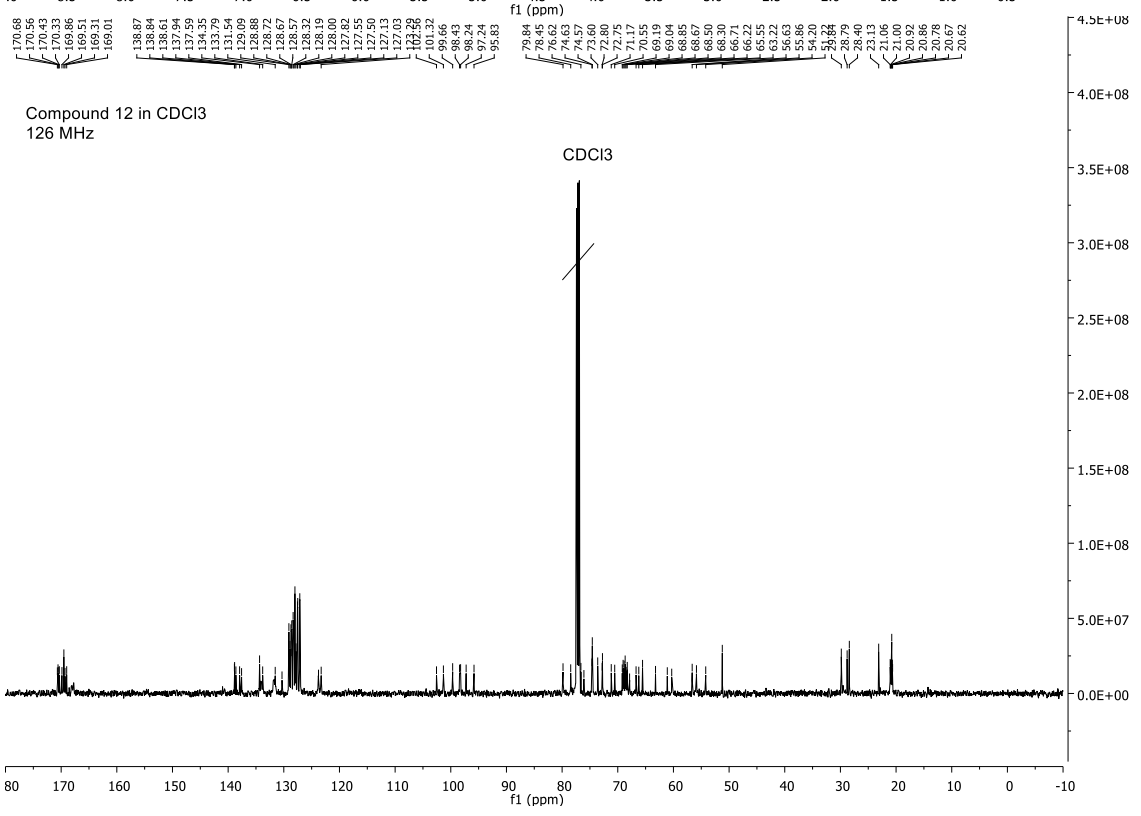

Compound 2. in CDCl3
126 MHz
CDCl3
f1 (ppm)

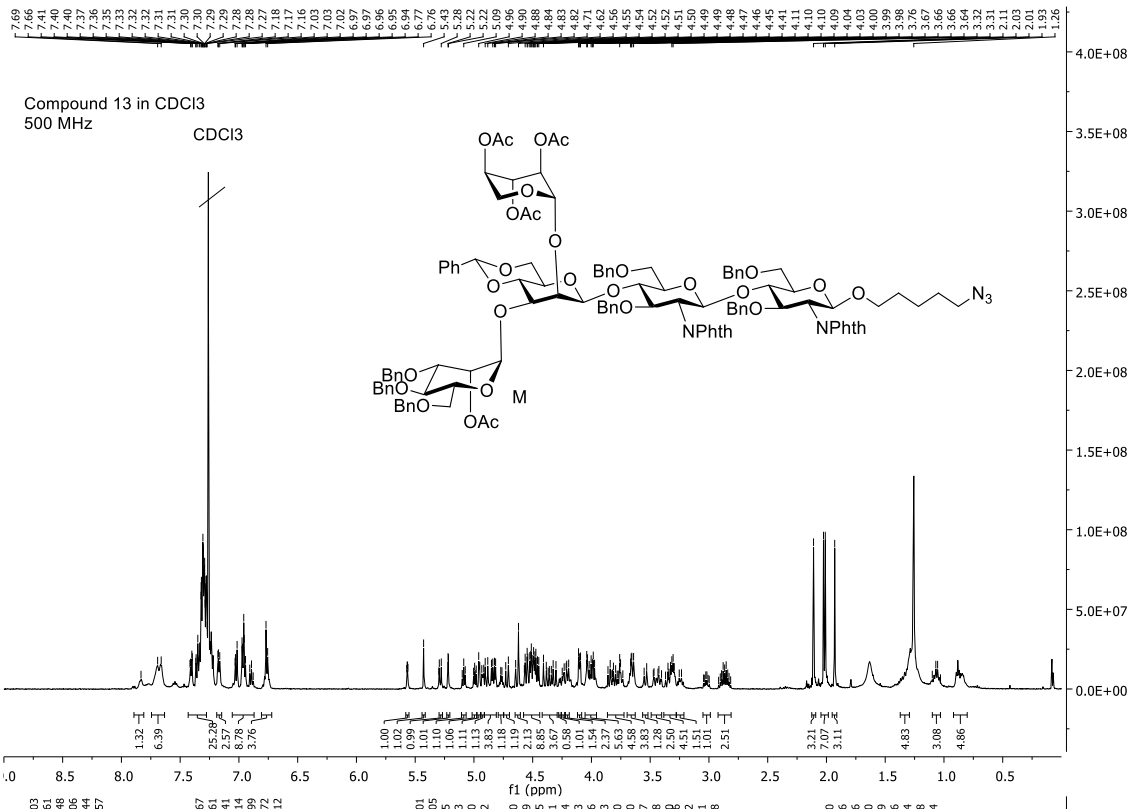

Compound 4. in CDCl3
500MHz
CDCl3
f1 (ppm)
Compound 4 in CDCl3
126 MHz
CDCl3
f1 (ppm)

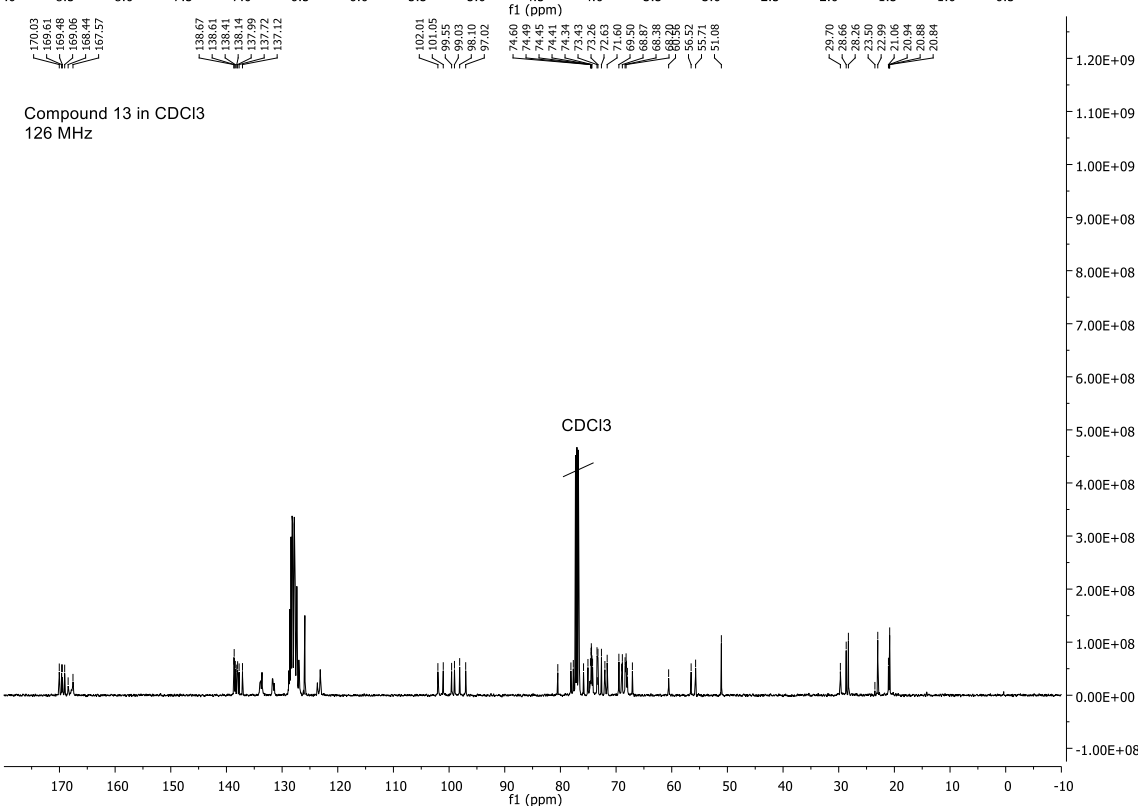

Compound 11 in CDCl3
500 MHz
AcO
AcO
OAc
Ph
OH
BnO
BnO
NPhth
BnO
BnO
NPhth
N3
f1 (ppm)

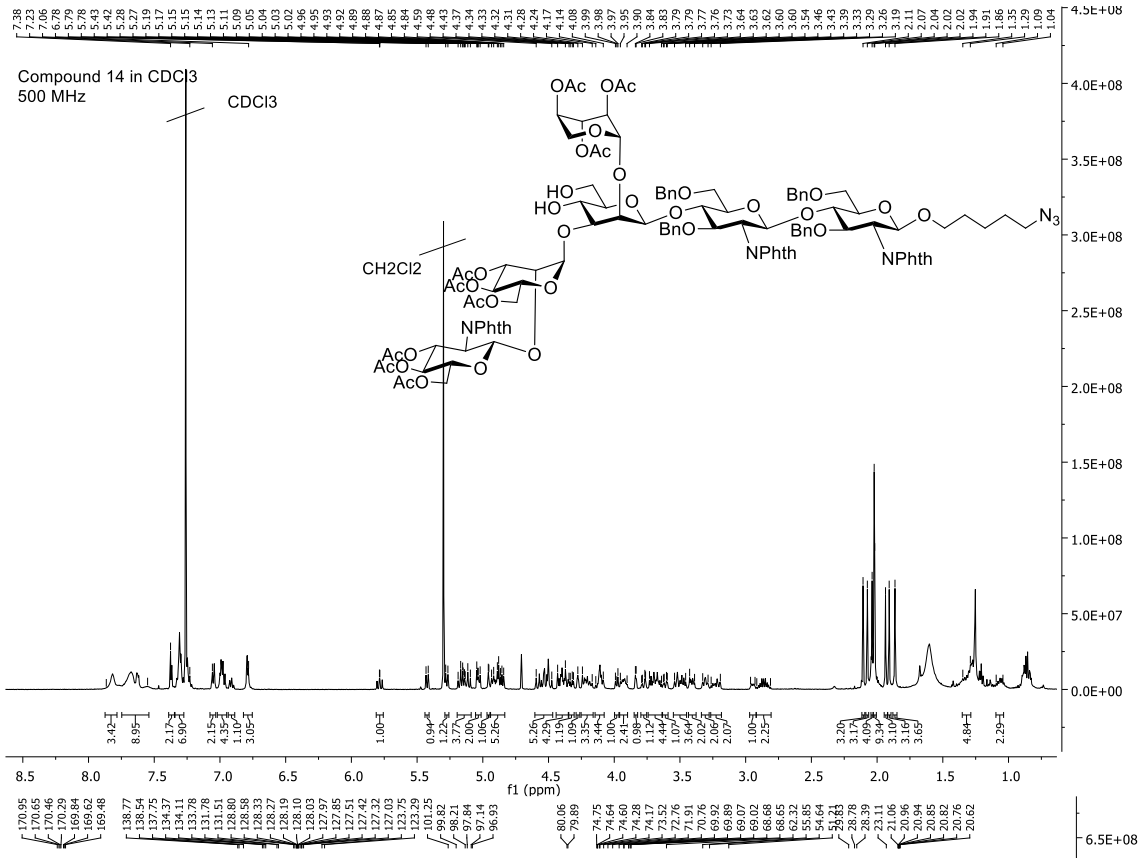

Compound 11 in CDCl3
126 MHz
CDCl3
f1 (ppm)

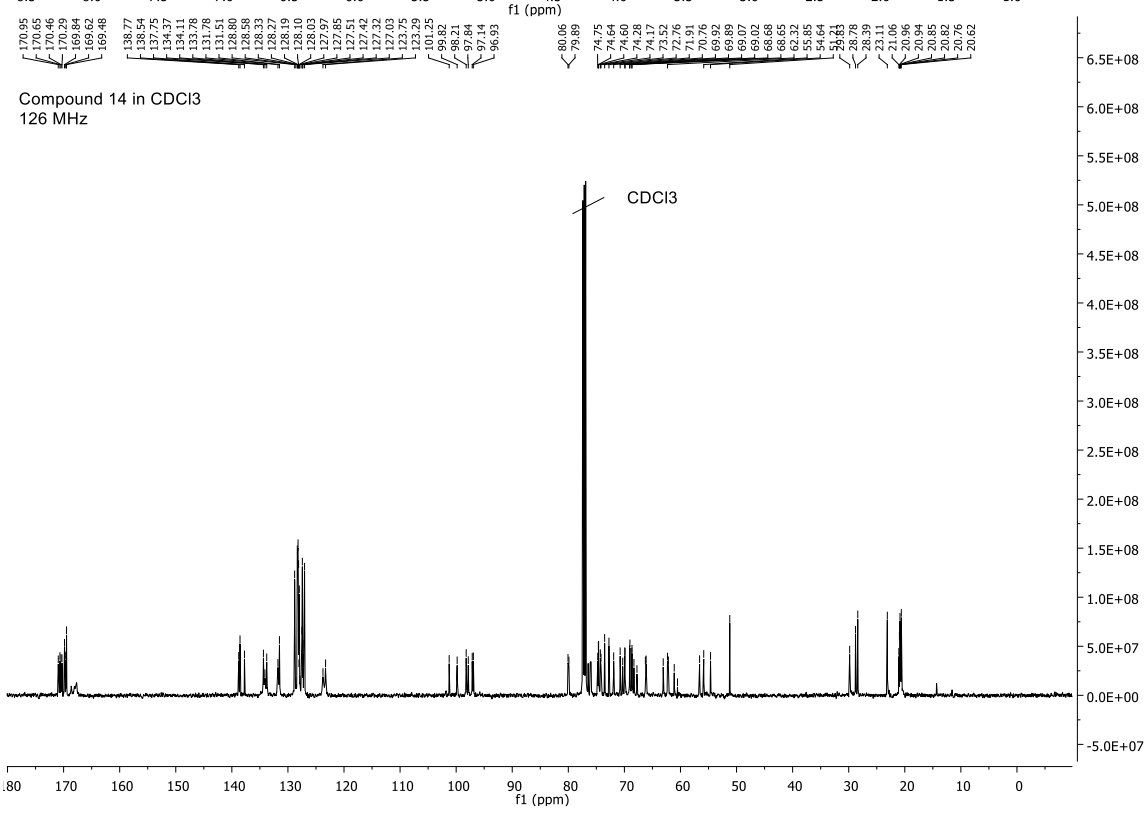

Compound 12 in CDCl3
500 MHz
CDCl3
G
f1 (ppm)

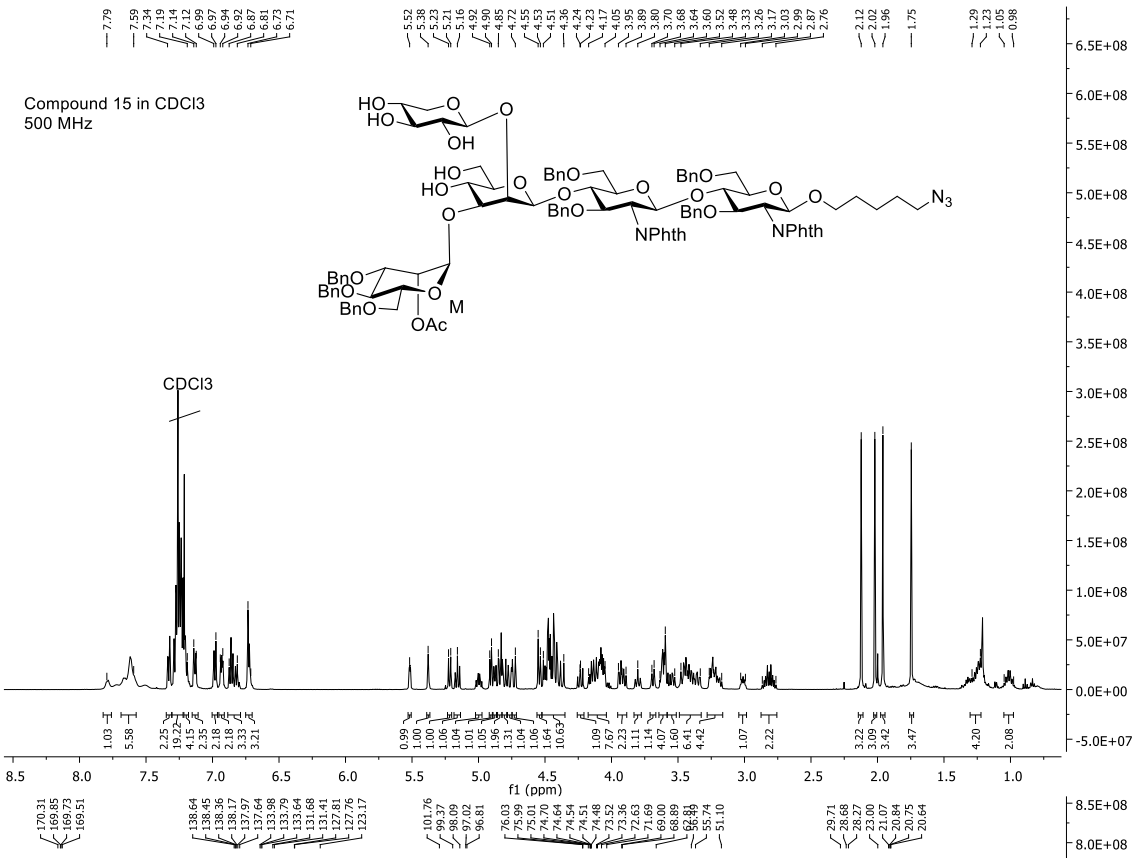

Compound 12 in CDCl3
126 MHz
CDCl3
f1 (ppm)

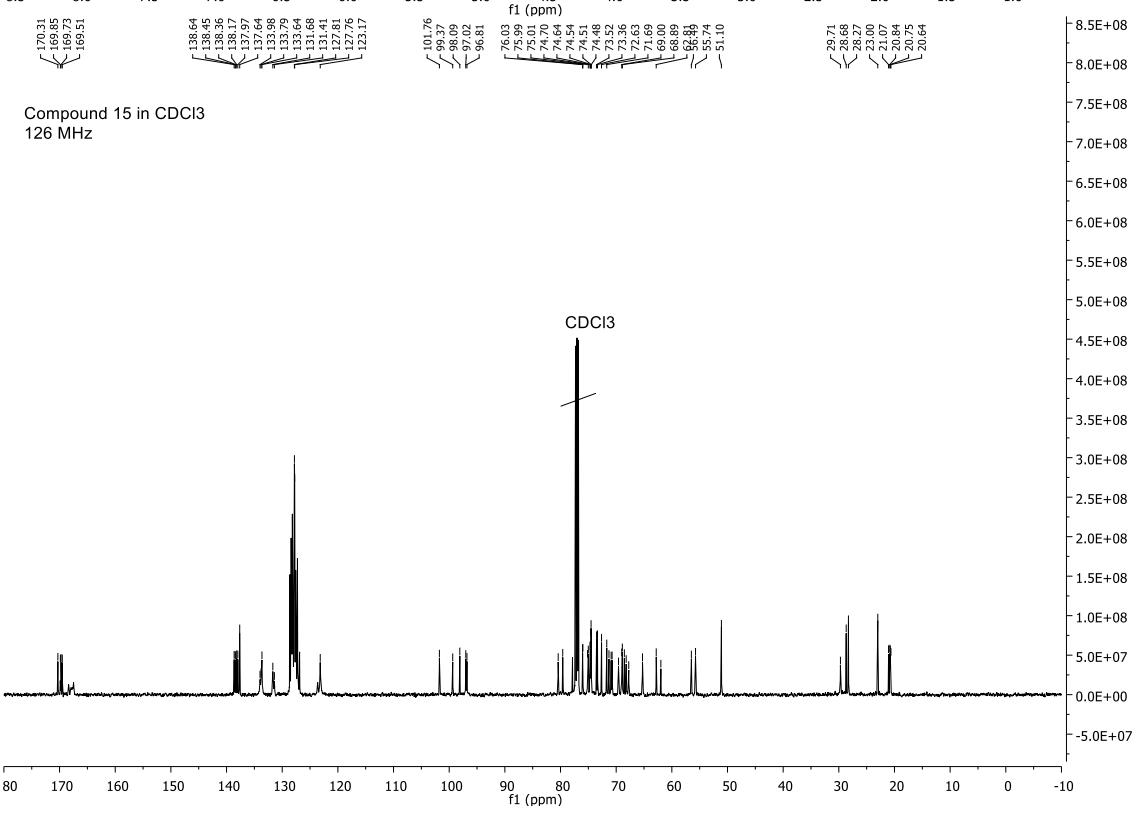

Compound 13 in CDCl3
500 MHz
CDCl3
OAc
OAc
OAc
O
Ph
O
O
O
O
BnO
BnO
O
O
BnO
BnO
O
O
N3
NPhth
NPhth
BnO
BnO
BnO
O
M
OAc
f1 (ppm)

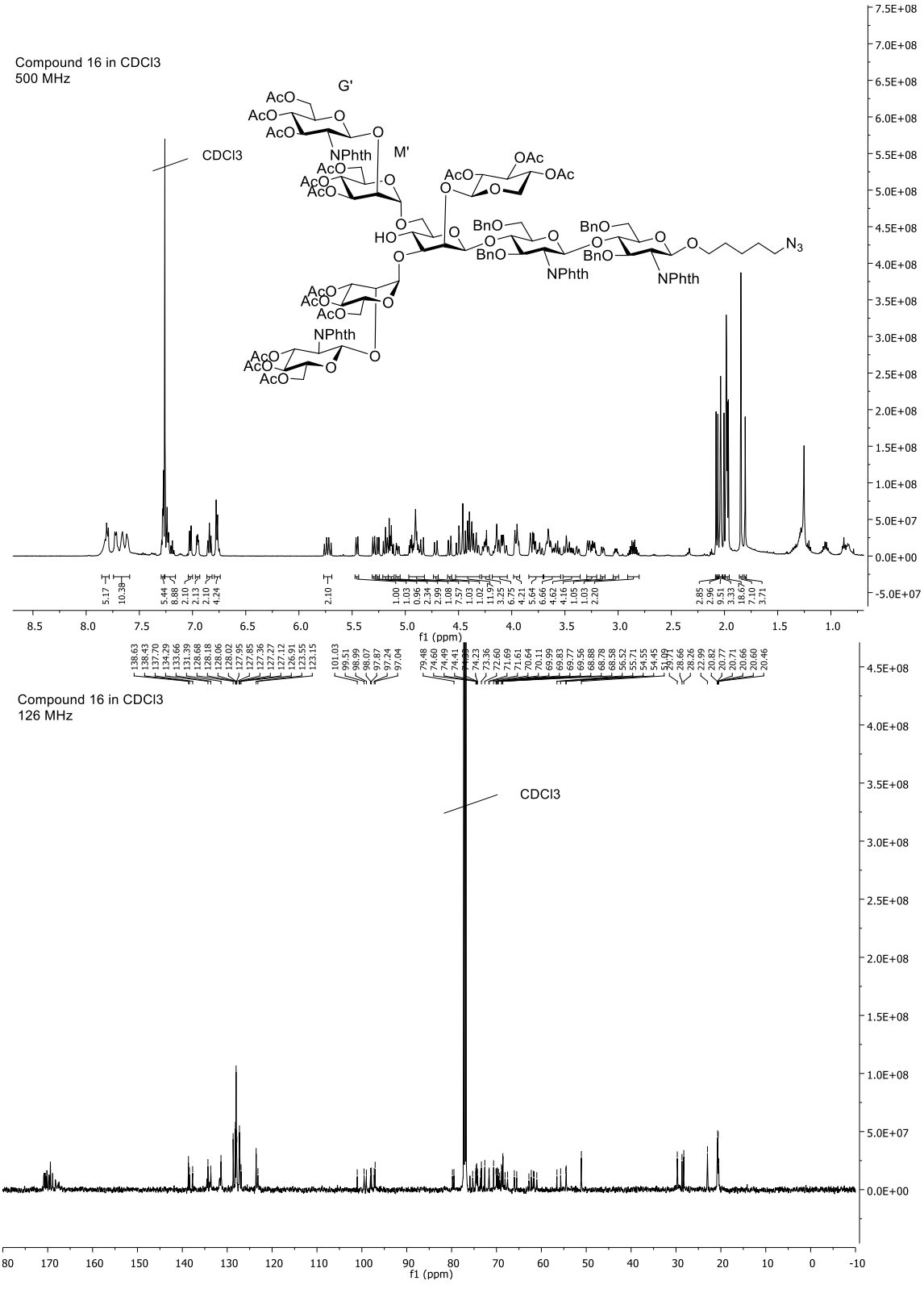

Compound 13 in CDCl3
126 MHz
CDCl3
f1 (ppm)

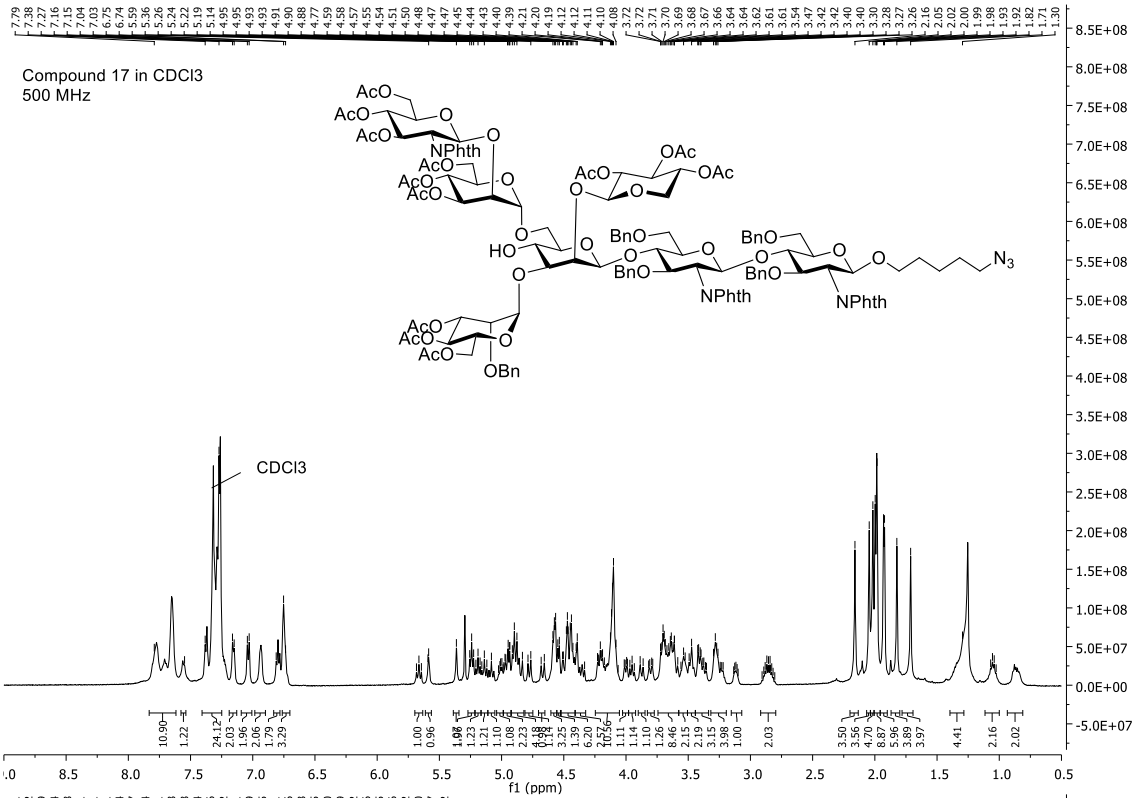

Compound 14 in CDCl3
500 MHz
CDCl3
CH2Cl2
f1 (ppm)

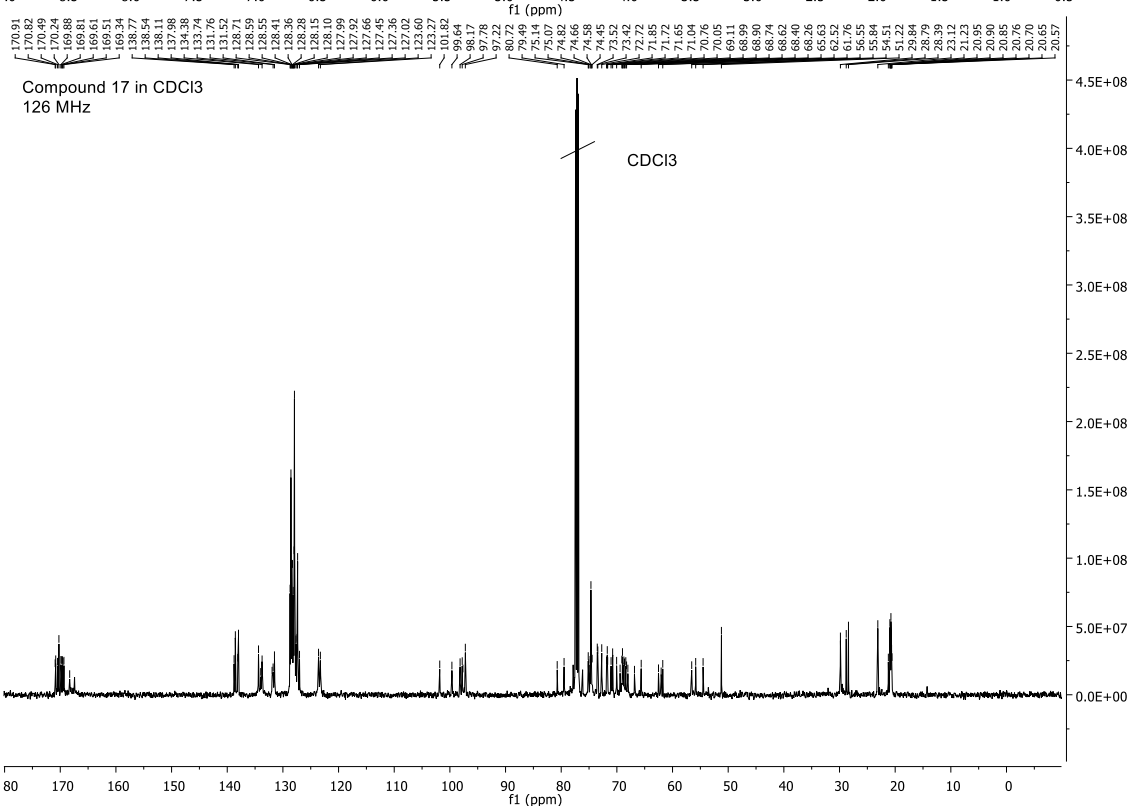

Compound 14 in CDCl3
126 MHz
CDCl3
f1 (ppm)

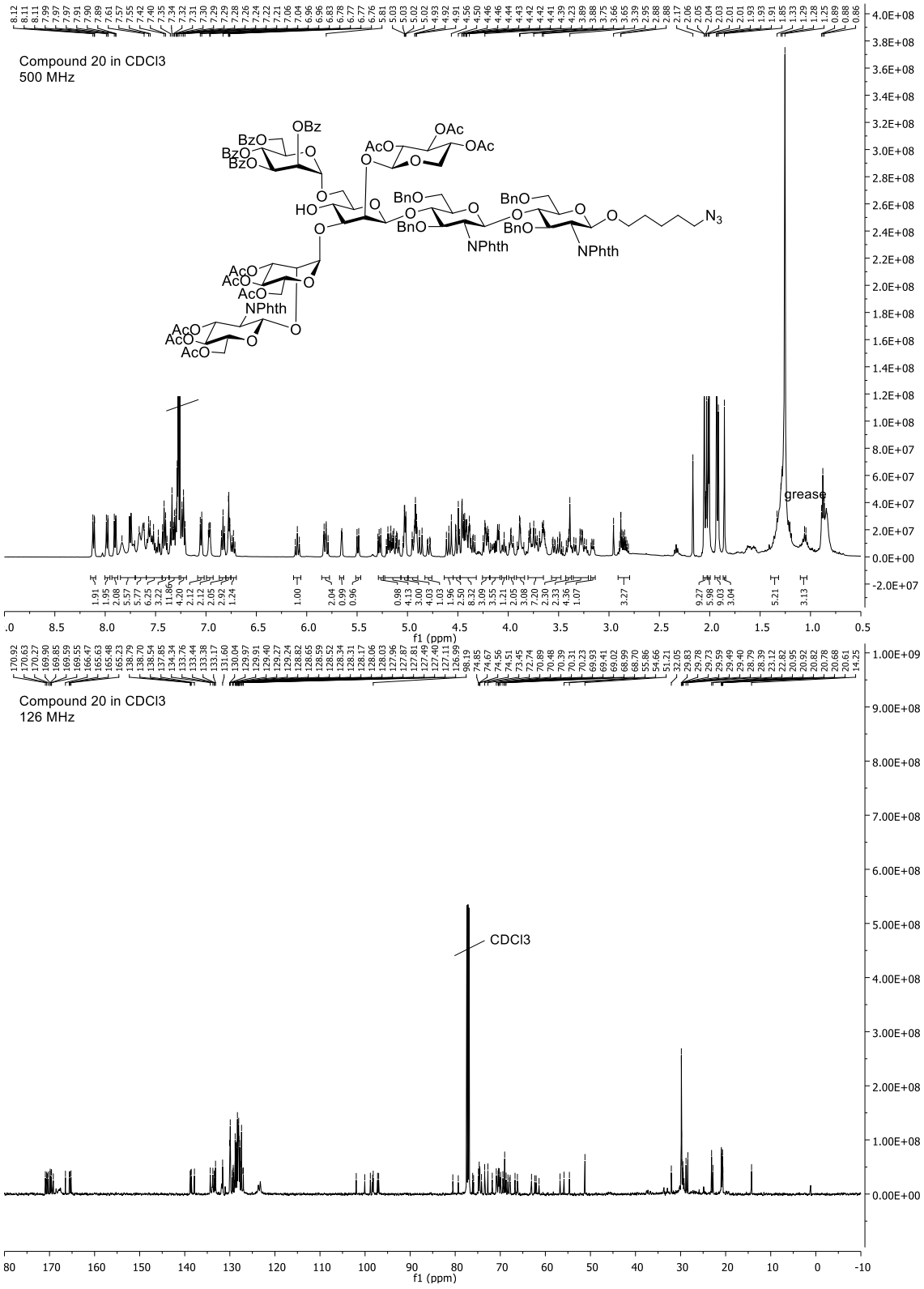

Compound 15 in CDCl3
500 MHz
CDCl3
f1 (ppm)

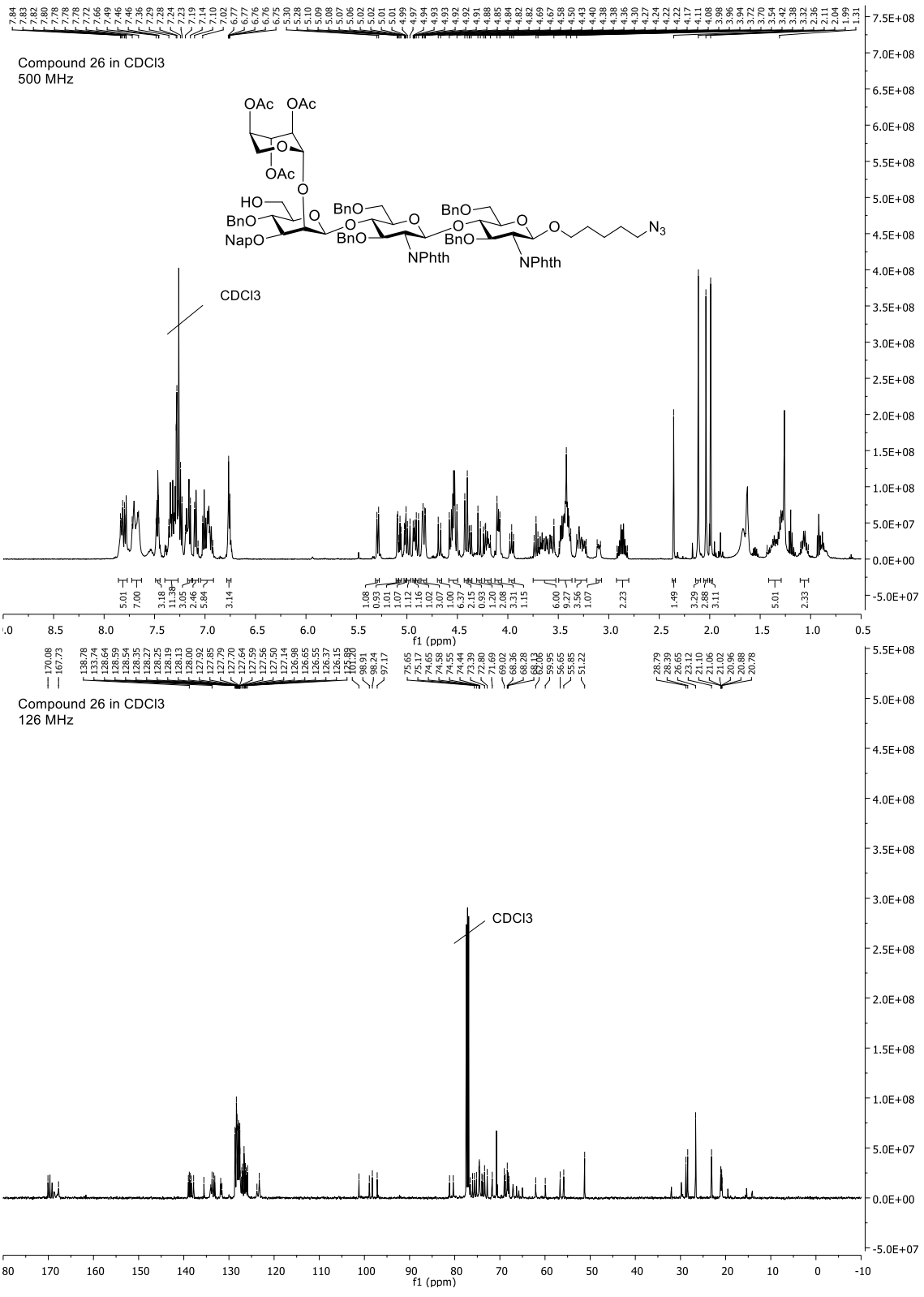

Compound 15 in CDCl3
126 MHz
CDCl3
f1 (ppm)

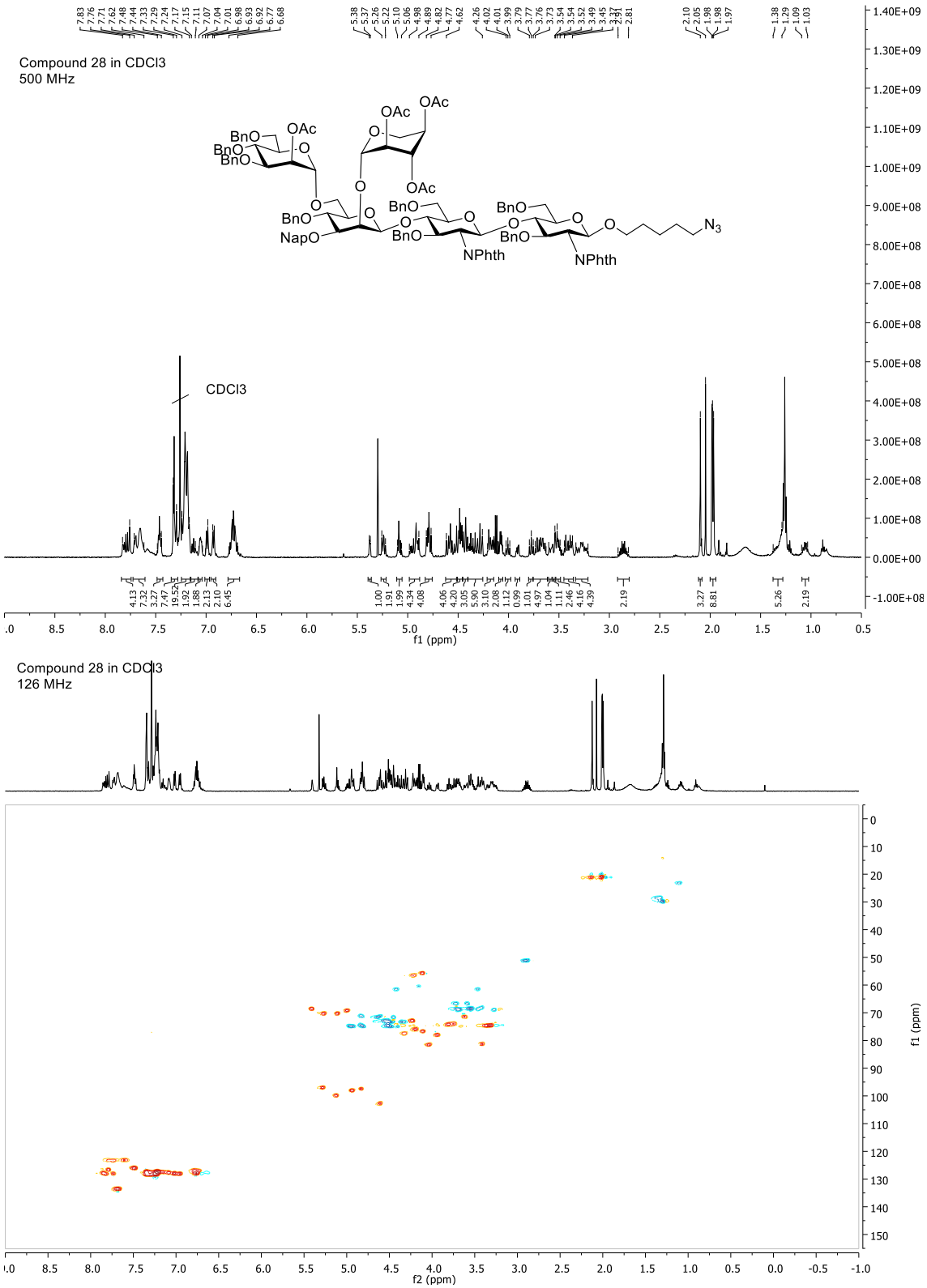

Compound 16 in CDCl3
500 MHz
G'
M'
CDCl3
NPhth
BnO
OAc
AcO
HO
N3
f1 (ppm)
Compound 16 in CDCl3
126 MHz
CDCl3
f1 (ppm)

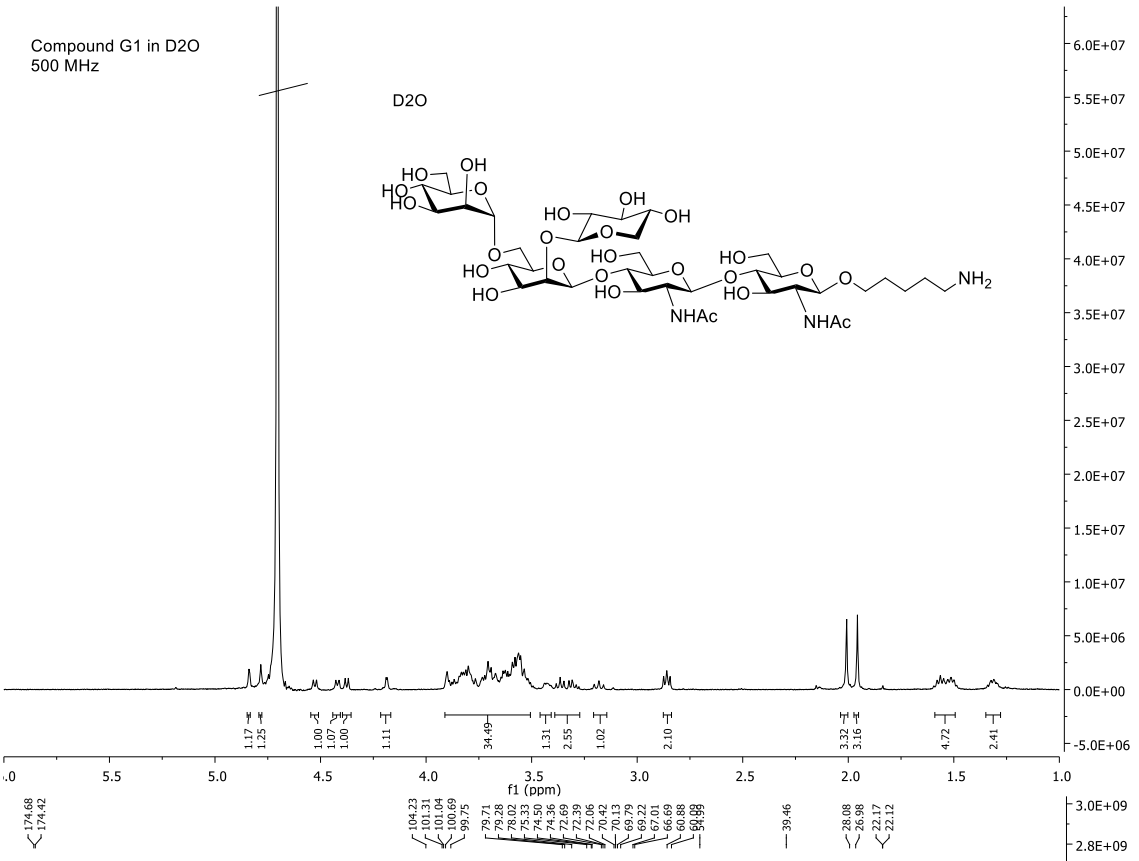

Compound 17 in CDCl3
500 MHz
CDCl3
f1 (ppm)

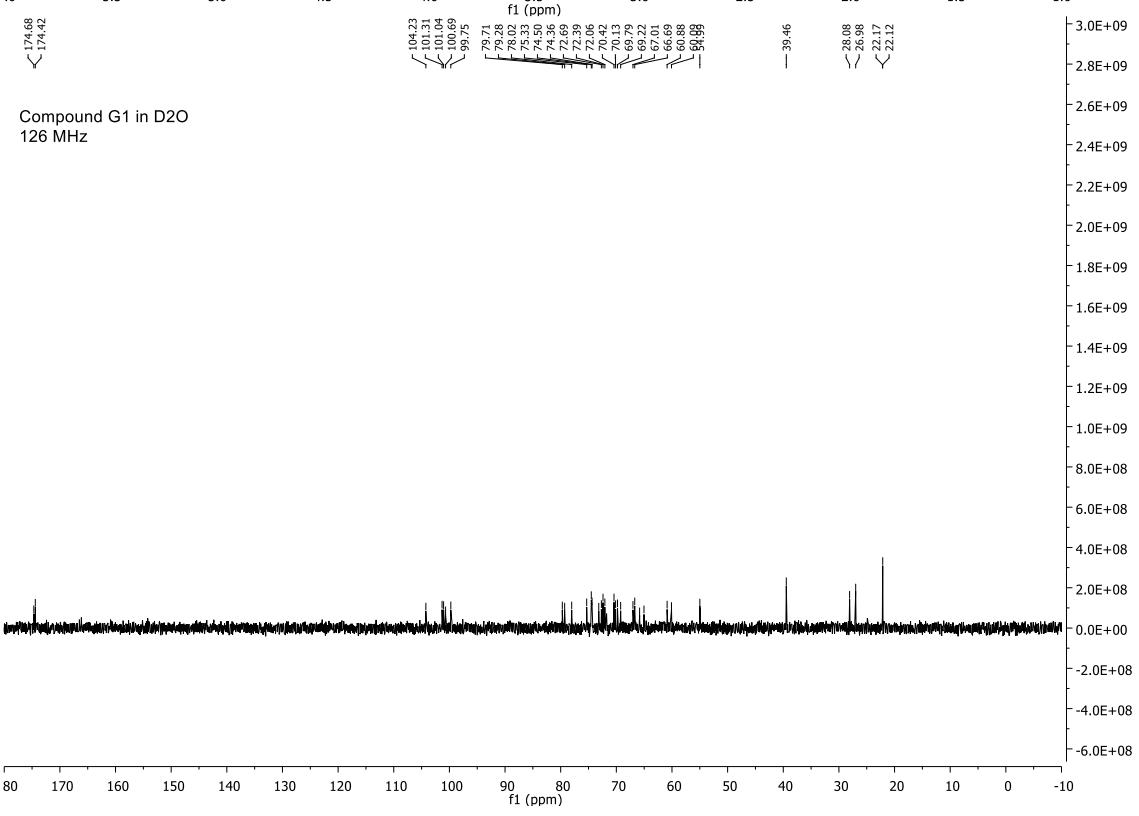

Compound 17 in CDCl3
126 MHz
CDCl3
f1 (ppm)

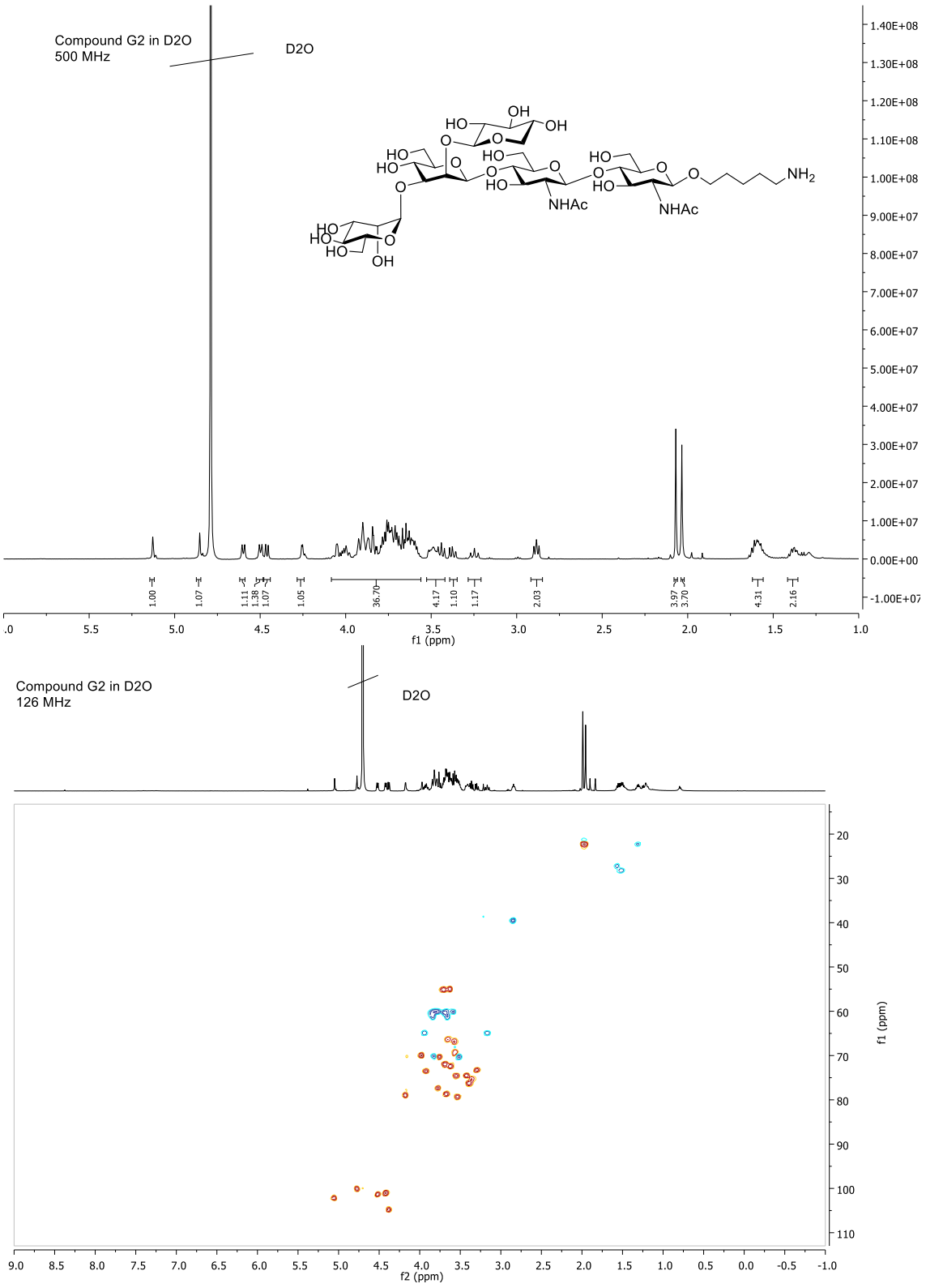

Compound 20 in CDCl3
500 MHz
grease
f1 (ppm)
Compound 20 in CDCl3
126 MHz
CDCl3
f1 (ppm)

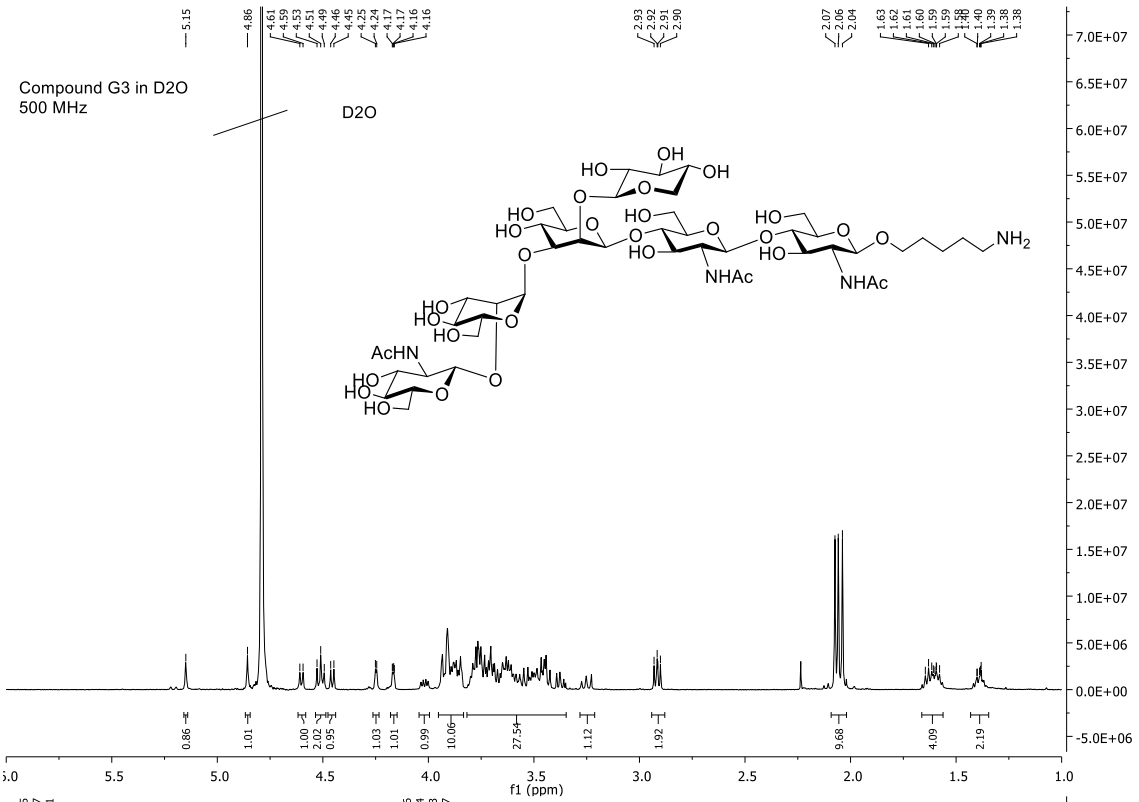

Compound 26 in CDCl3
500 MHz
OAc
OAc
OAc
O
HO
BnO
NapO
BnO
BnO
BnO
BnO
BnO
NPhth
NPhth
N3
CDCl3
f1 (ppm)
Compound 26 in CDCl3
126 MHz
CDCl3
f1 (ppm)

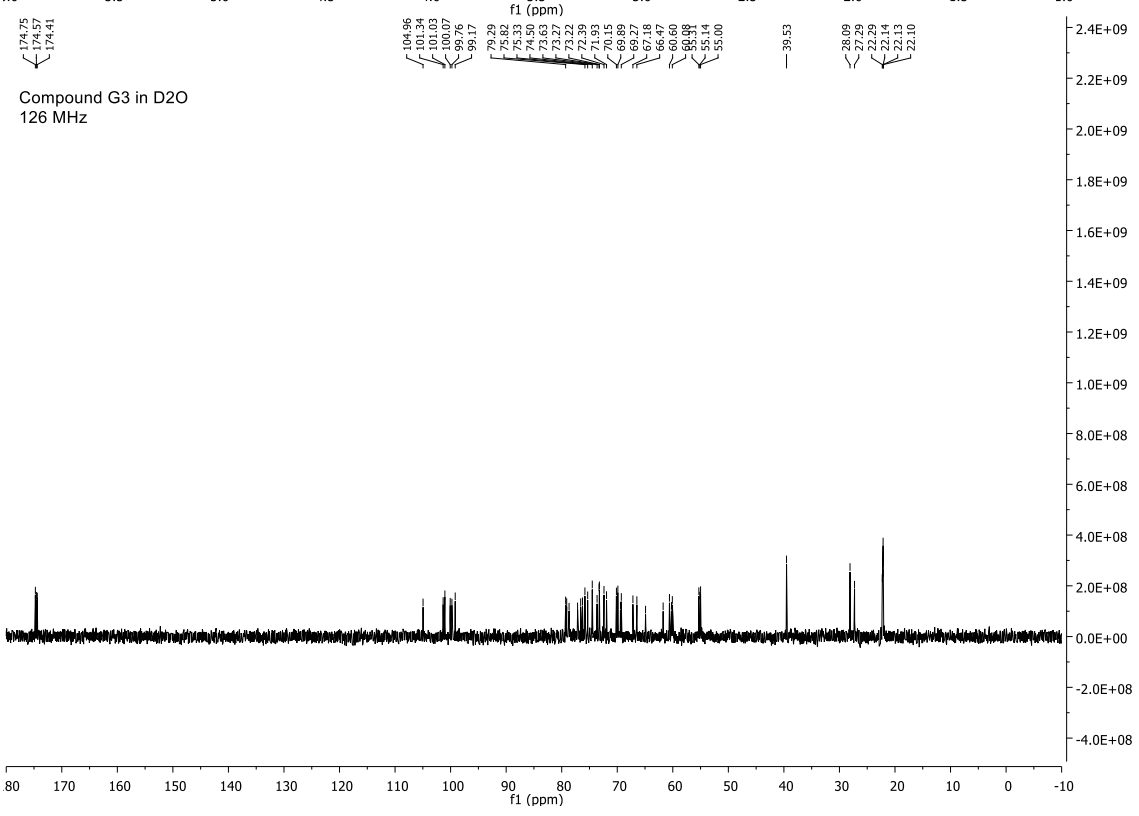

Compound 28 in CDCl3
500 MHz
CDCl3
f1 (ppm)
Compound 28 in CDCl3
126 MHz
f2 (ppm)
f1 (ppm)

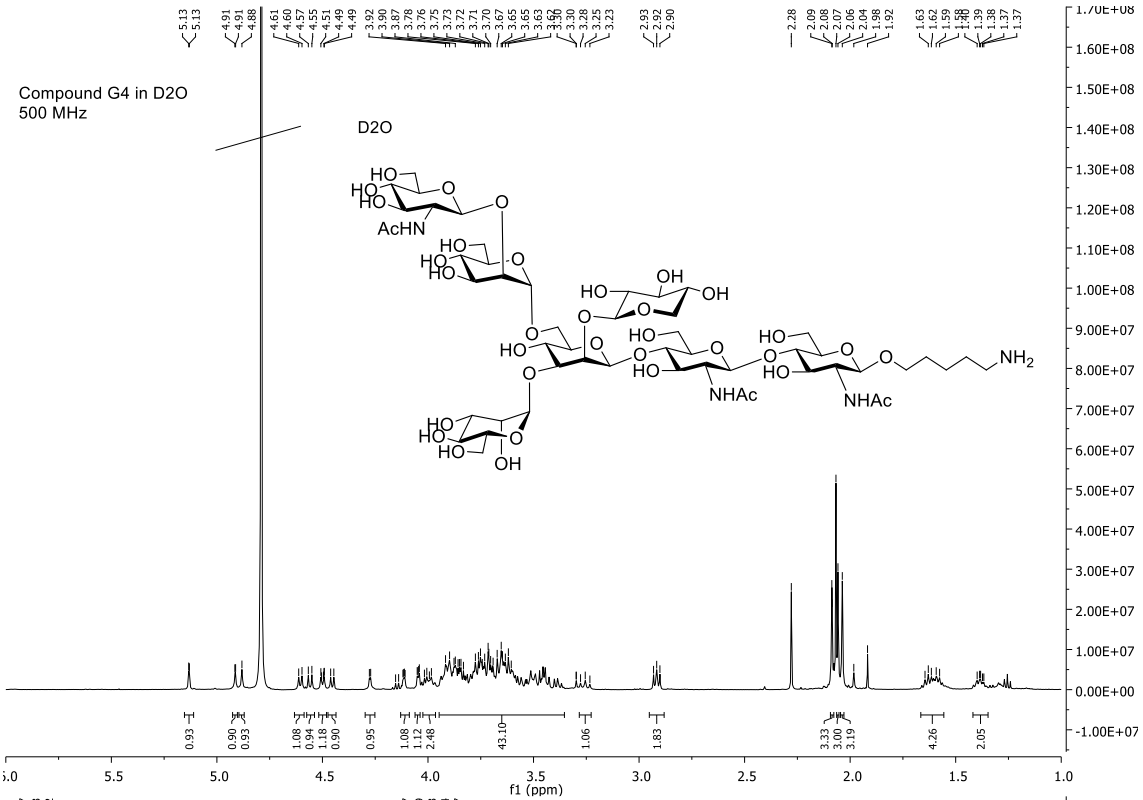

Compound G1 in D2O
500 MHz
D2O
NHAc
NHAc
NH2
f1 (ppm)

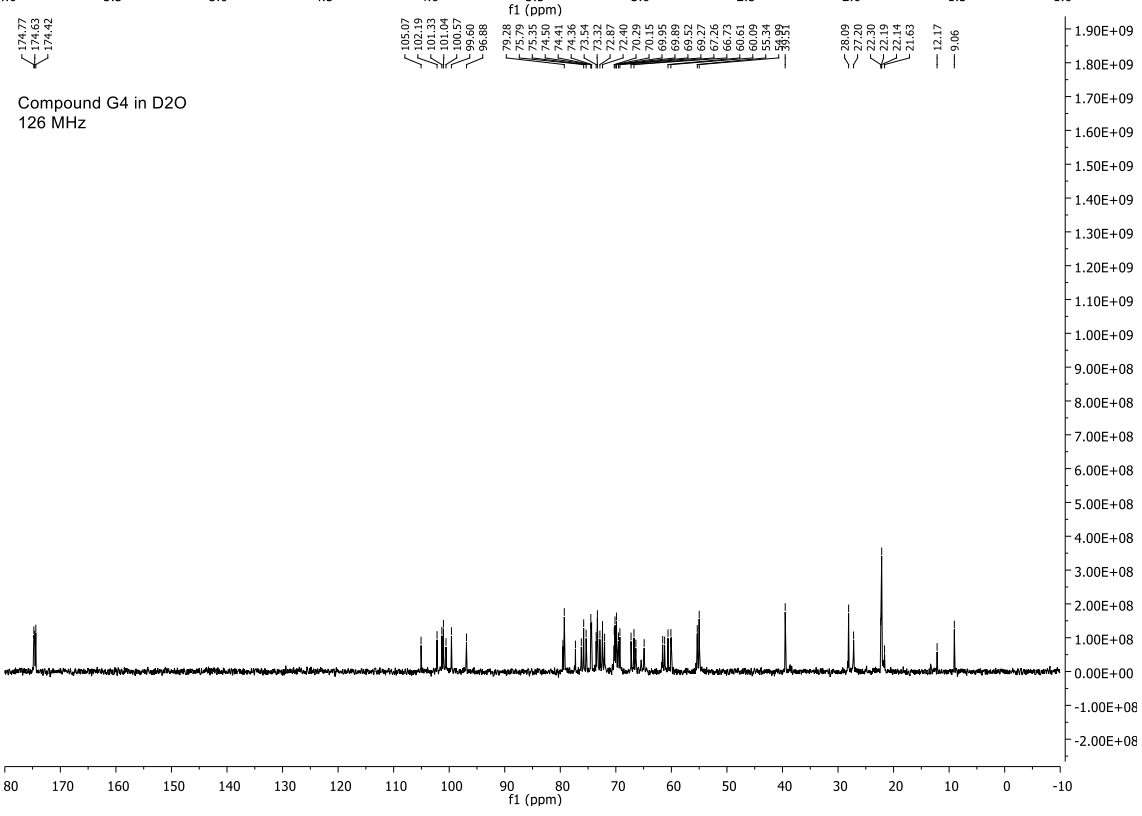

Compound G1 in D2O
126 MHz
f1 (ppm)

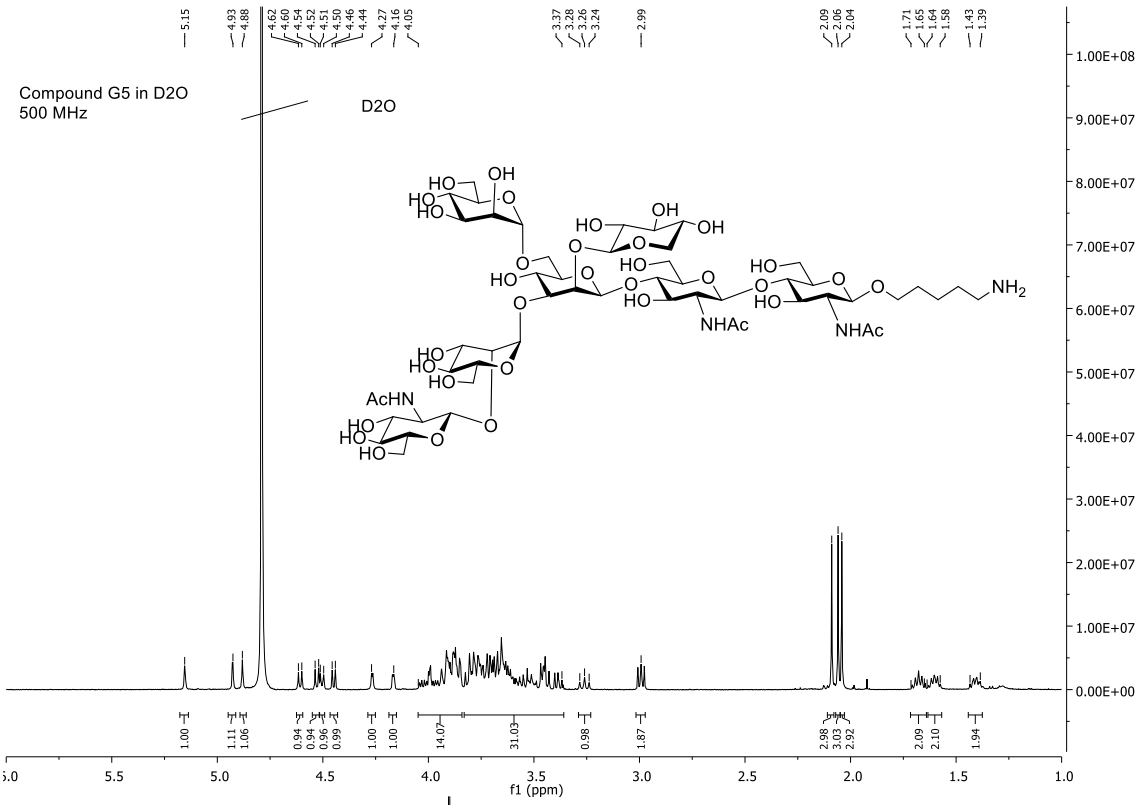

Compound G2 in D2O
500 MHz
D2O
NHAc
NHAc
NH2
1.00
1.07
1.11
1.38
1.07
1.05
36.70
4.17
1.10
1.17
2.03
3.97
3.70
4.31
2.16
f1 (ppm)
Compound G2 in D2O
126 MHz
D2O
f1 (ppm)
f2 (ppm)

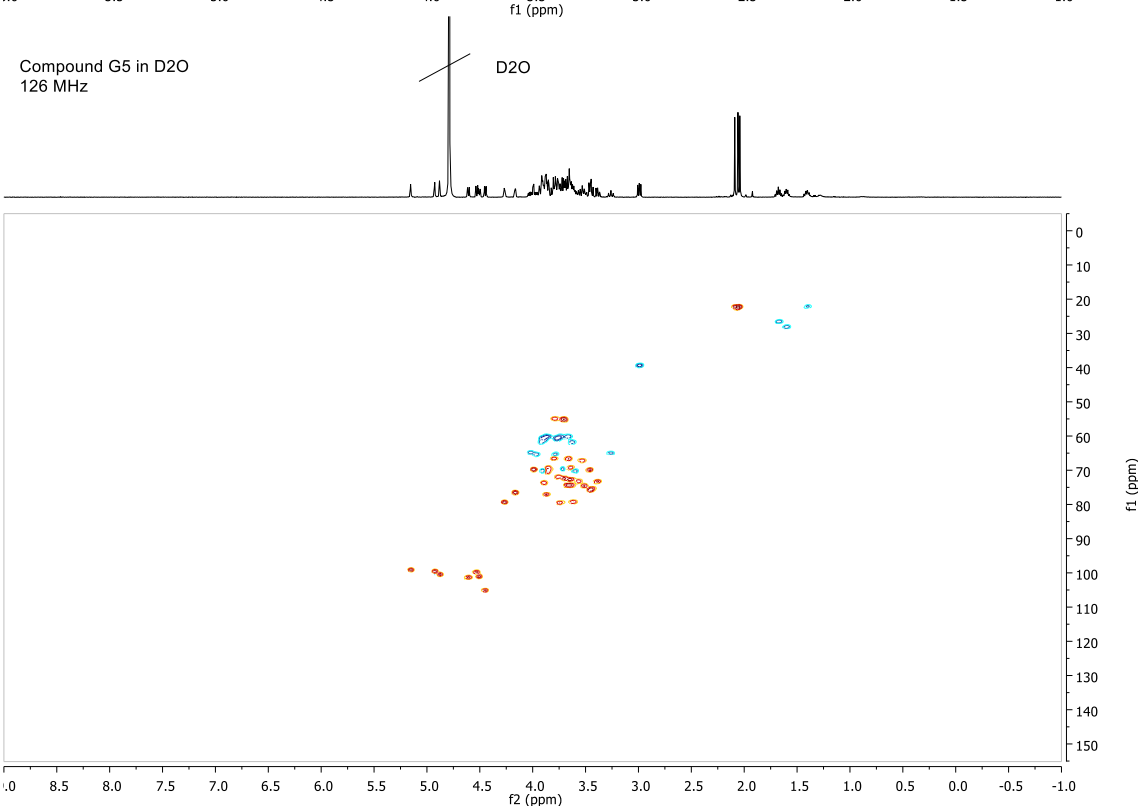

Compound G3 in D2O
500 MHz
D2O
NHAc
NHAc
AcHN
NH2
f1 (ppm)

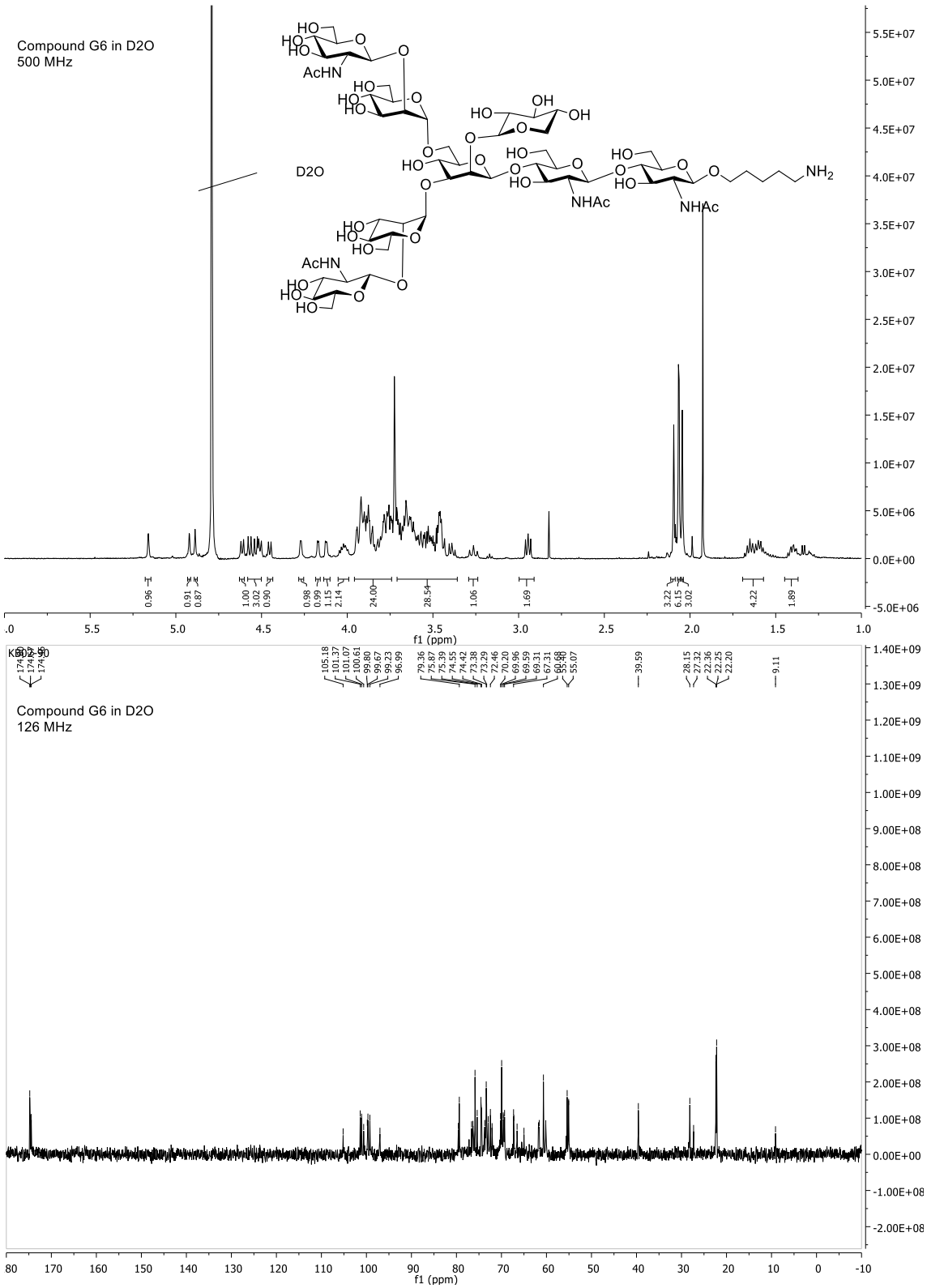

Compound G3 in D2O
126 MHz
f1 (ppm)

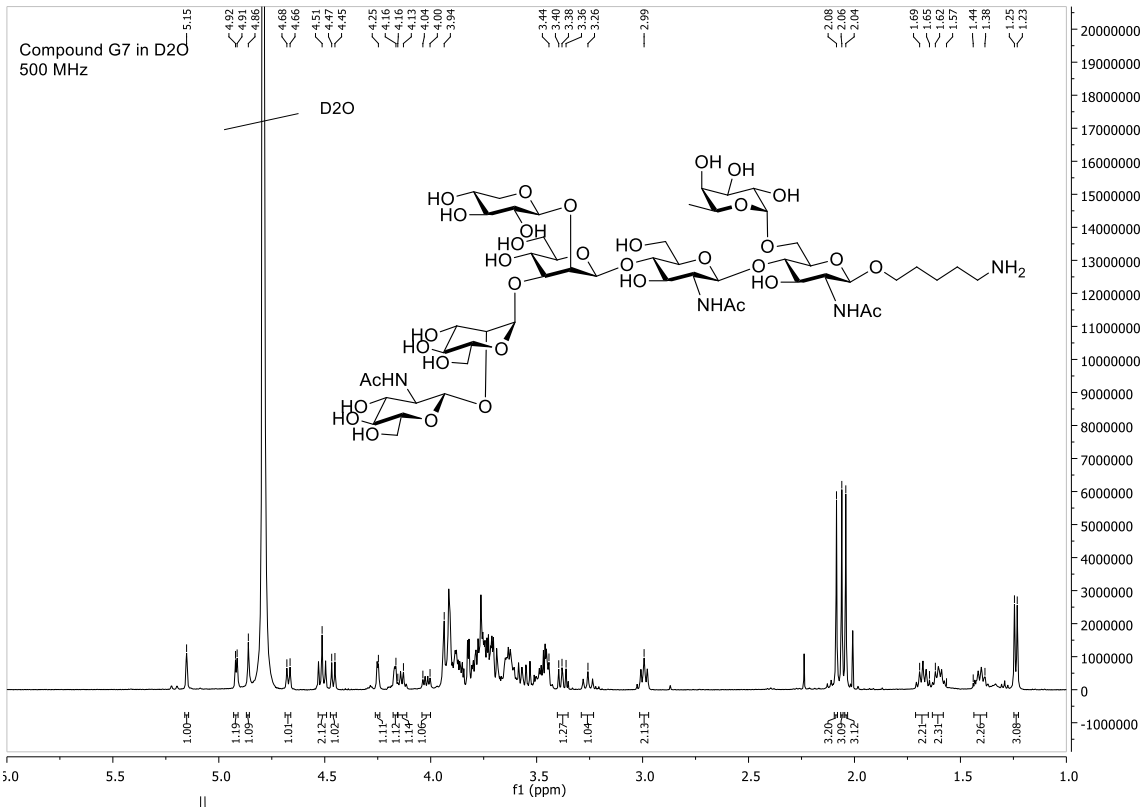

Compound G4 in D2O
500 MHz
D2O
AcHN
NHAc
NHAc
NH2
f1 (ppm)

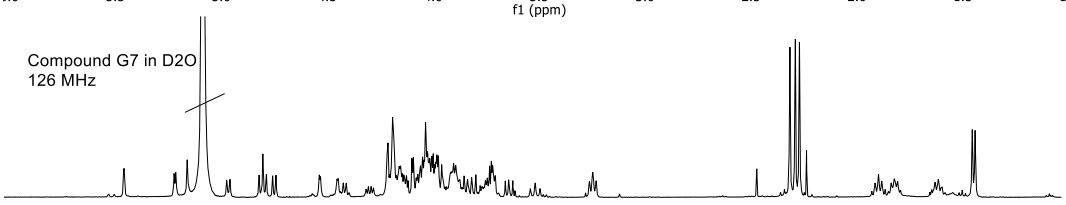

Compound G4 in D2O
126 MHz
f1 (ppm)

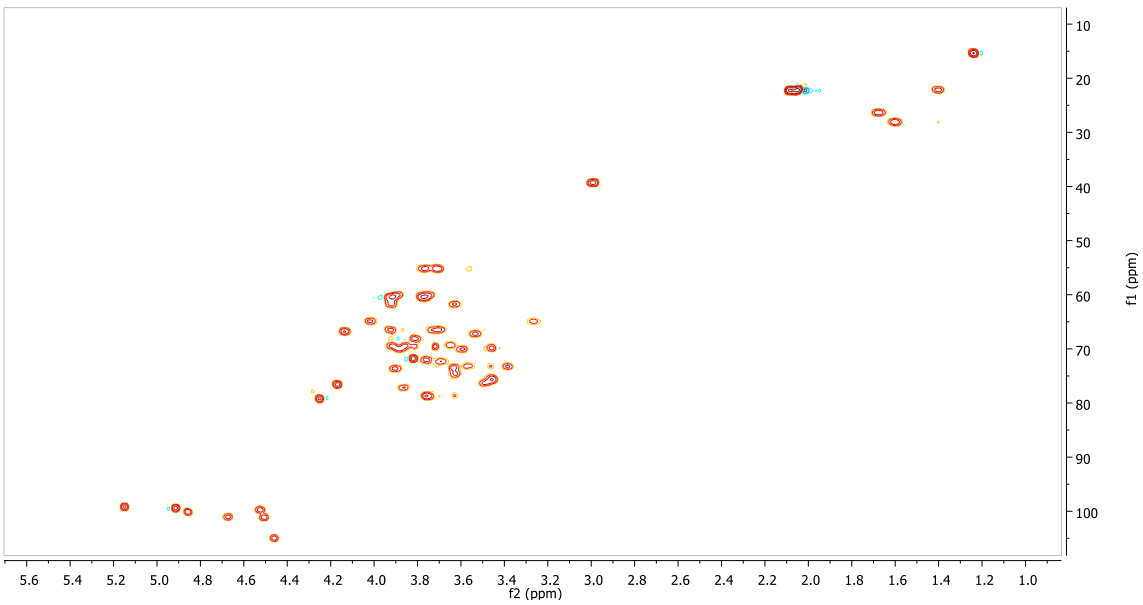

Compound G5 in D2O
500 MHz
D2O
f1 (ppm)

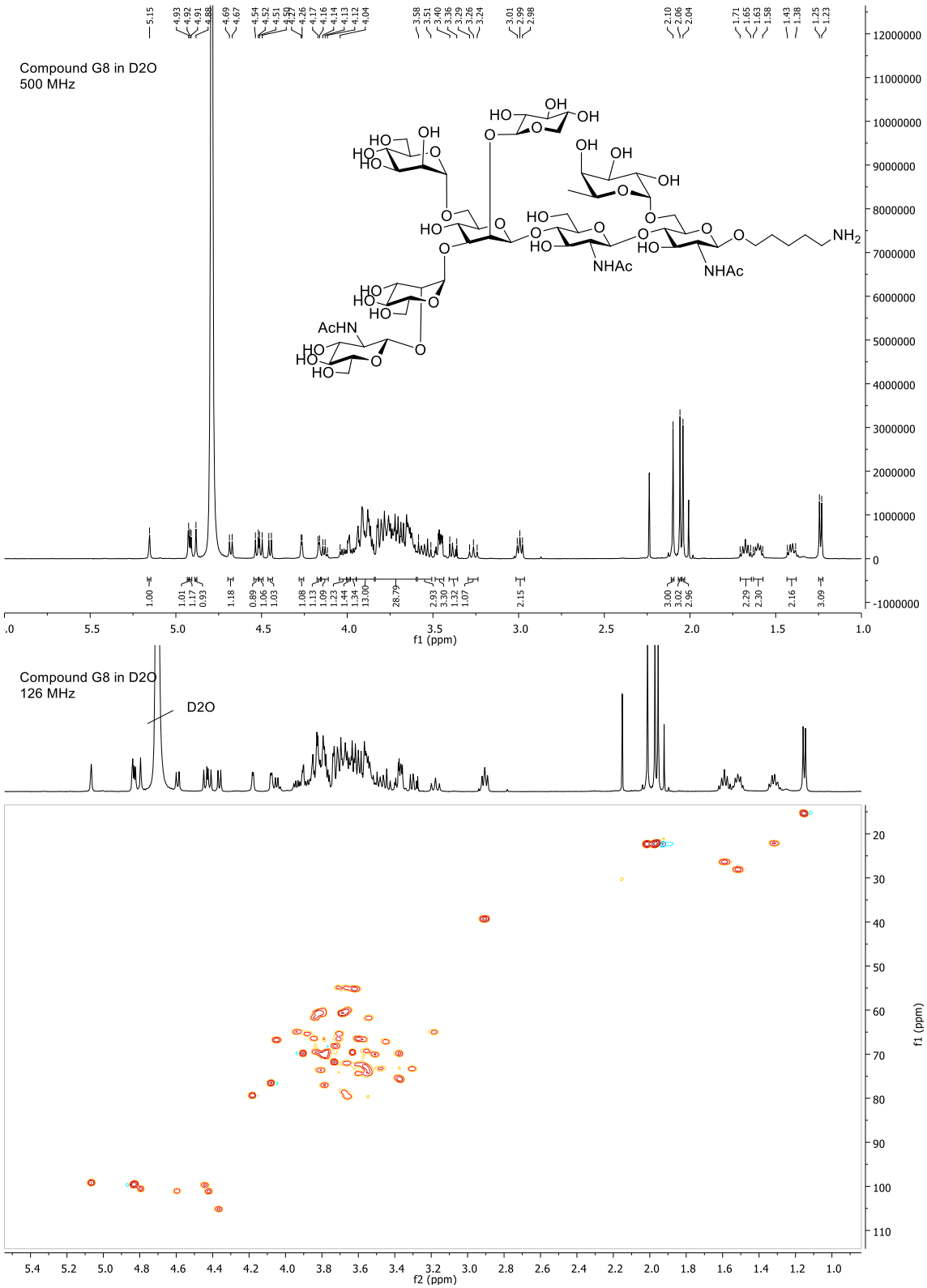

Compound G5 in D2O
126 MHz
D2O
f2 (ppm)
f1 (ppm)

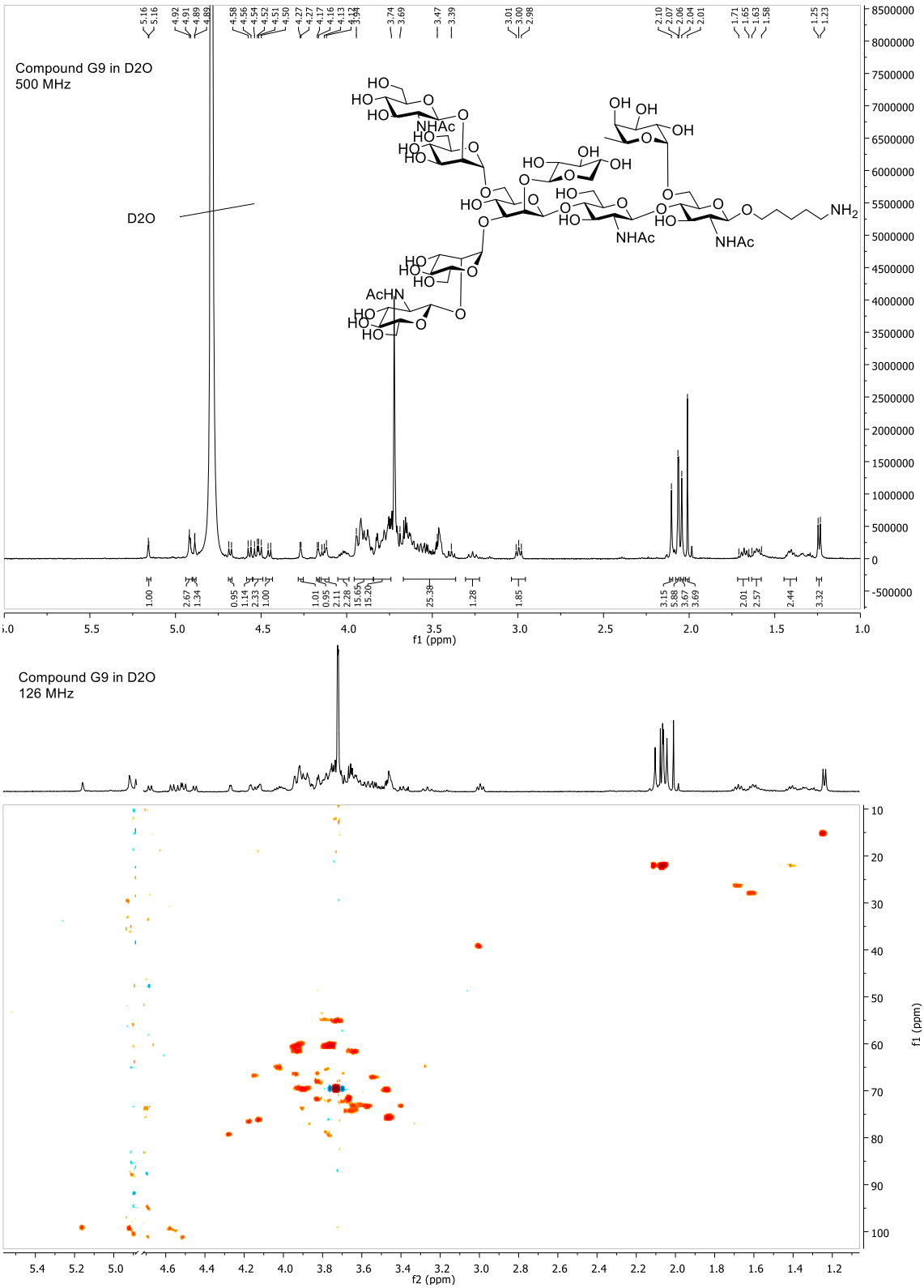

Compound G6 in D2O
500 MHz
D2O
NHAc
NHAc
NH2
f1 (ppm)
Compound G6 in D2O
126 MHz
f1 (ppm)

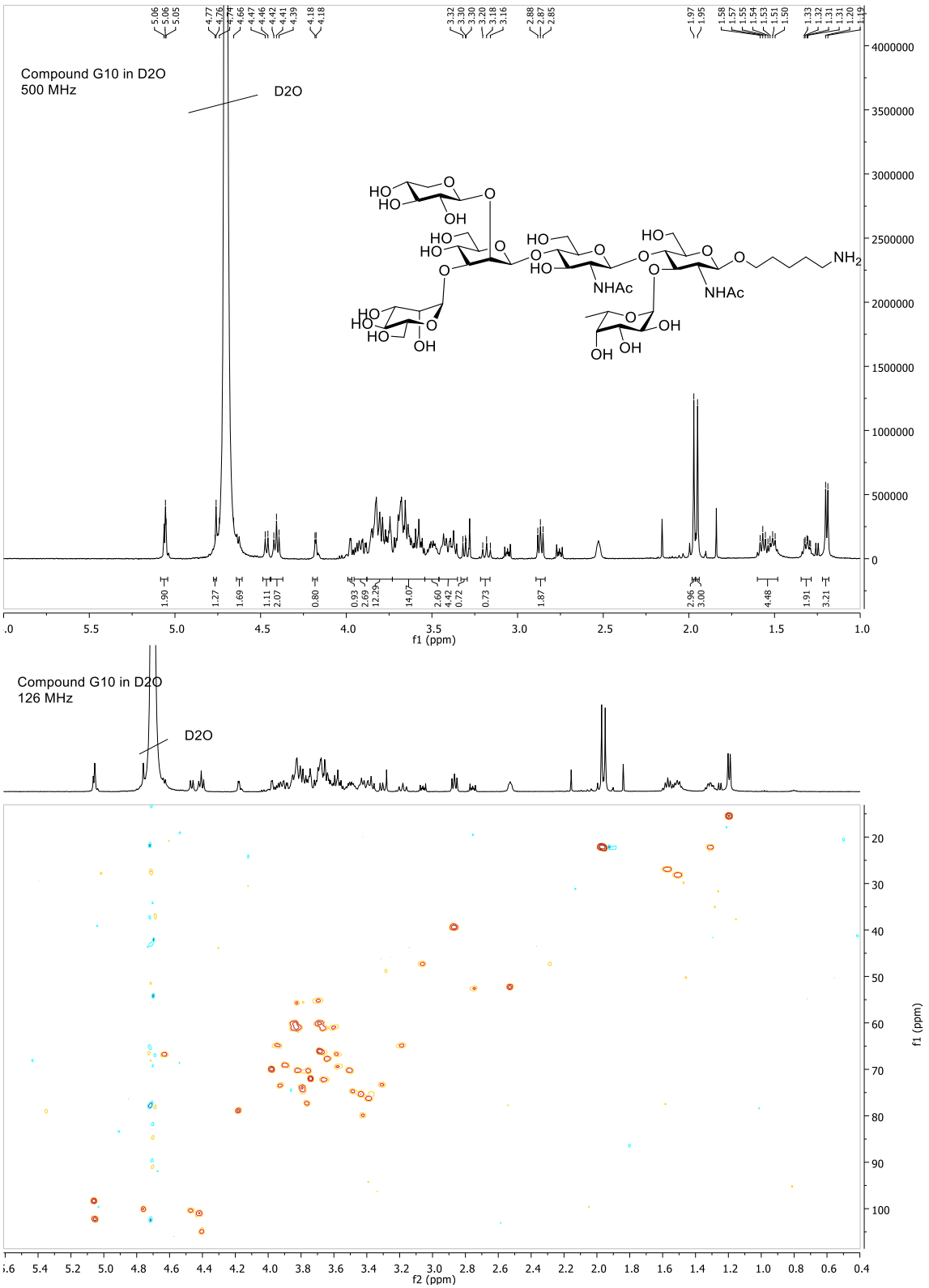

Compound G7 in D2O
500 MHz
D2O
f1 (ppm)

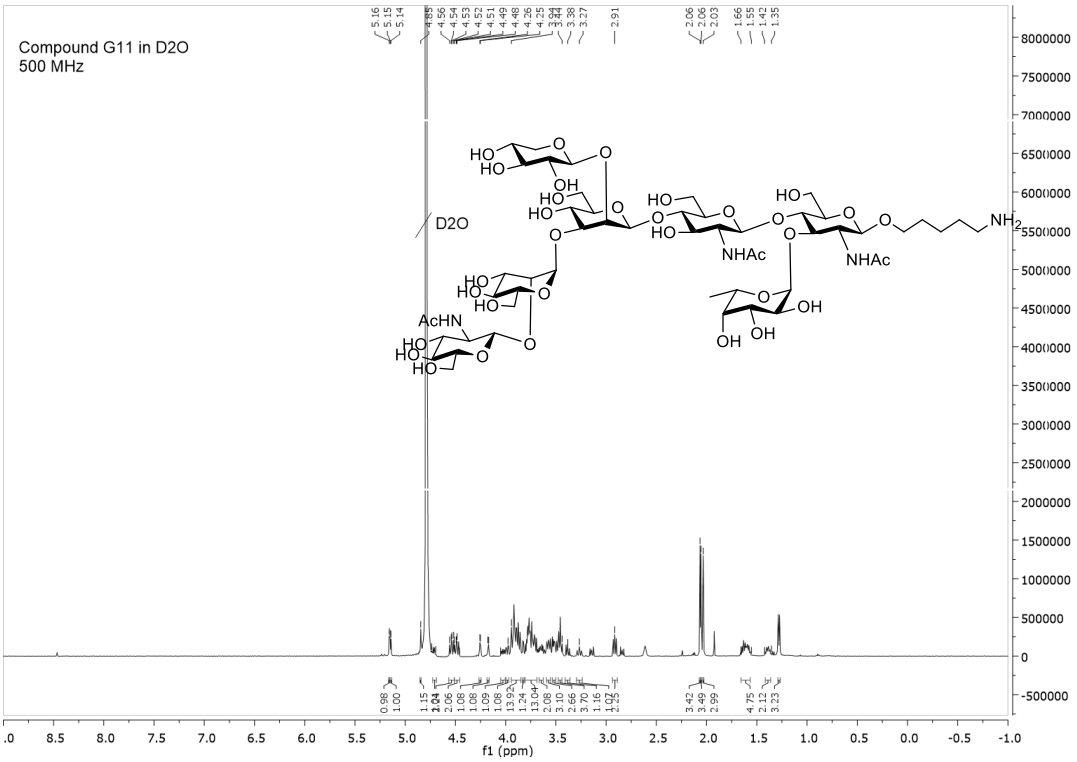

Compound G7 in D2O
126 MHz

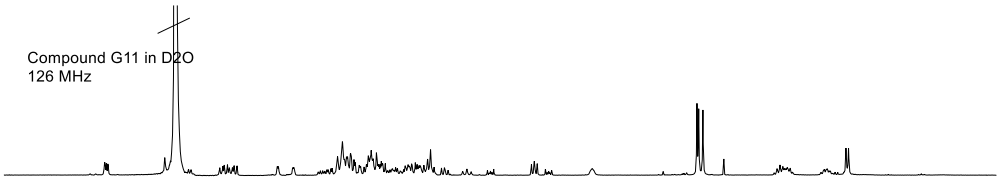

f2 (ppm)
f1 (ppm)

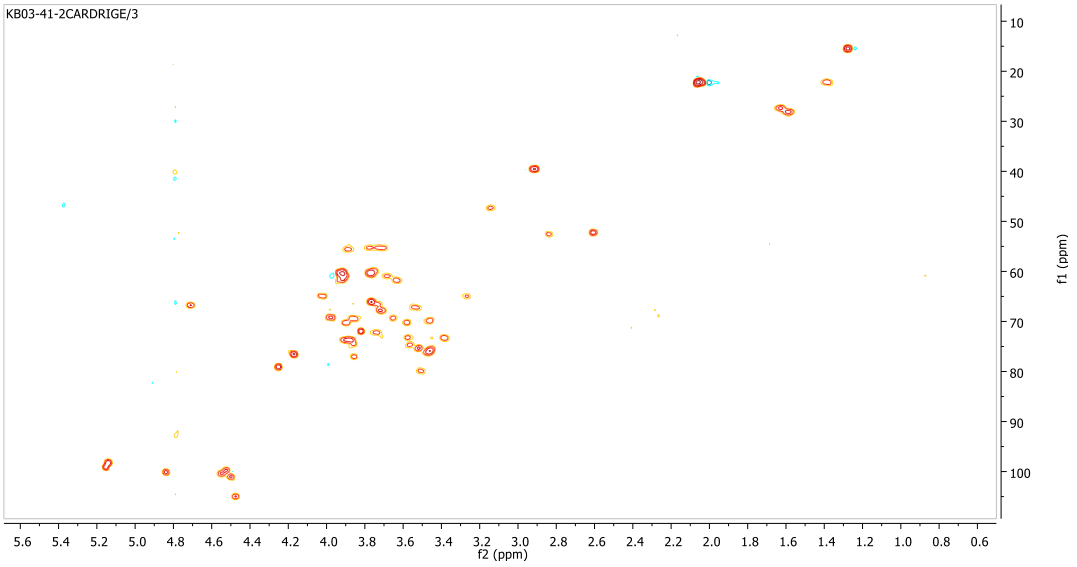

Compound G8 in D2O
500 MHz
f1 (ppm)
Compound G8 in D2O
126 MHz
D2O
f2 (ppm)
f1 (ppm)

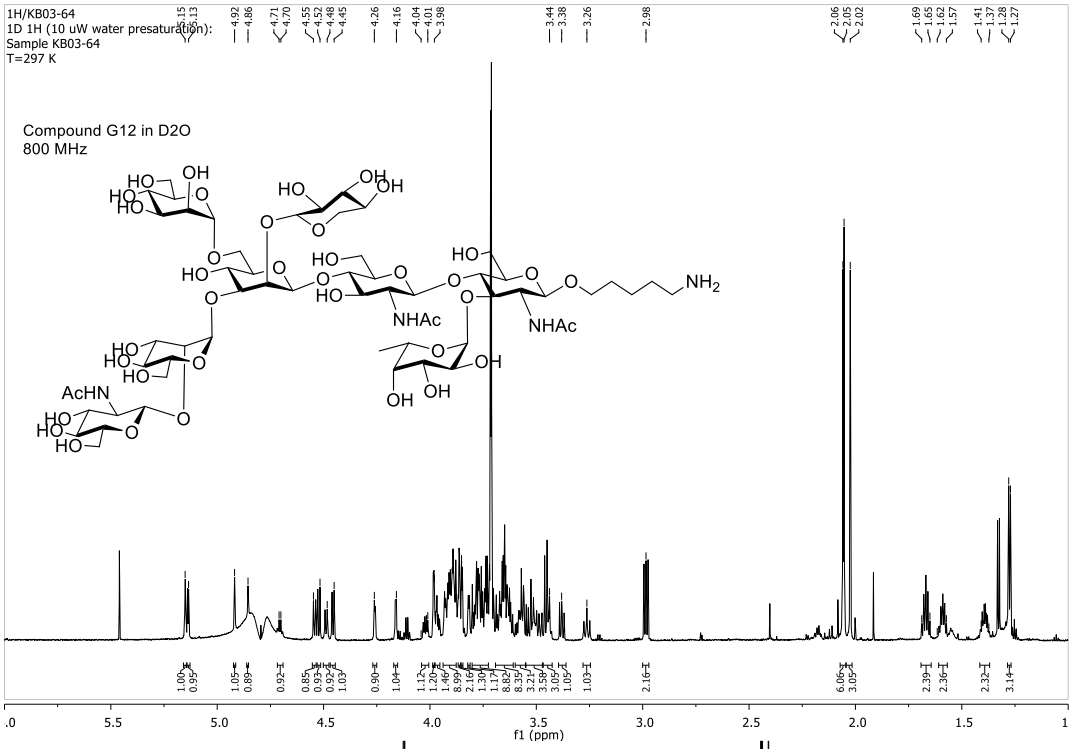

Compound G9 in D2O
500 MHz
D2O
f1 (ppm)
Compound G9 in D2O
126 MHz
f2 (ppm)
f1 (ppm)

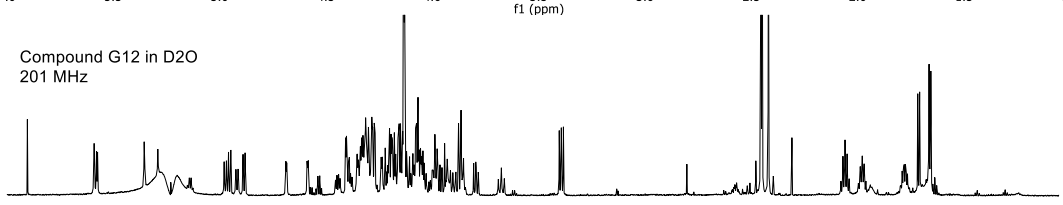

Compound G10 in D2O
500 MHz
D2O
f1 (ppm)
Compound G10 in D2O
126 MHz
D2O
f2 (ppm)
f1 (ppm)

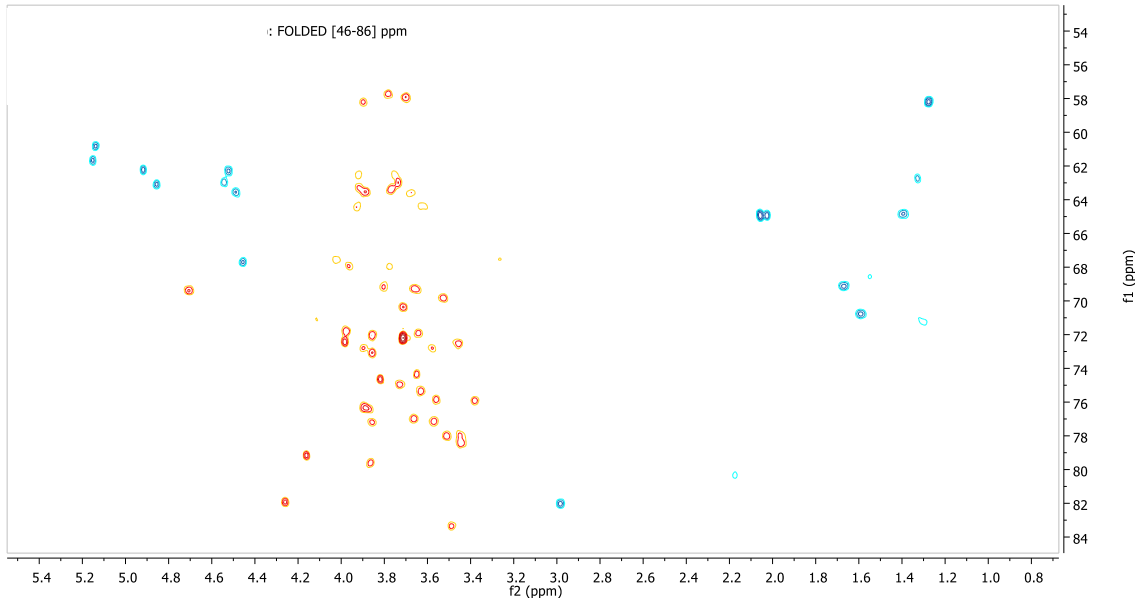

Compound G11 in D2O
500 MHz
D2O
NHAc
NHAc
AcHN
NH2
f1 (ppm)

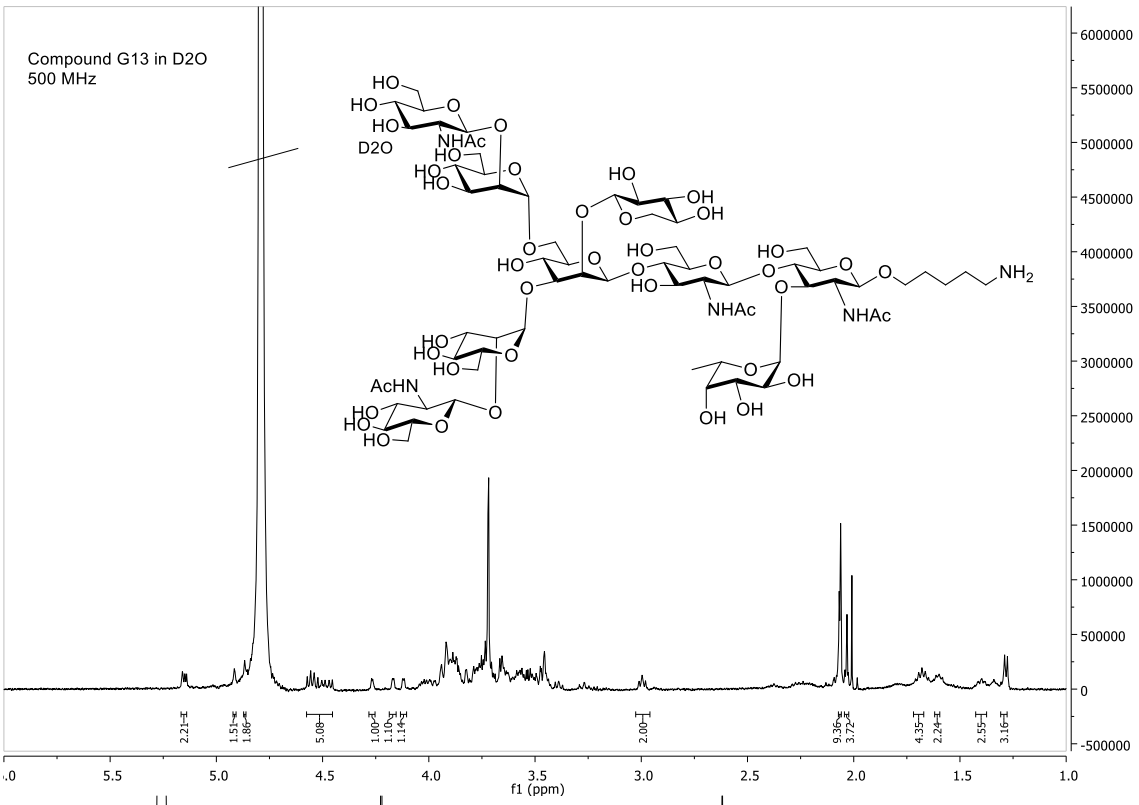

Compound G11 in D2O
126 MHz

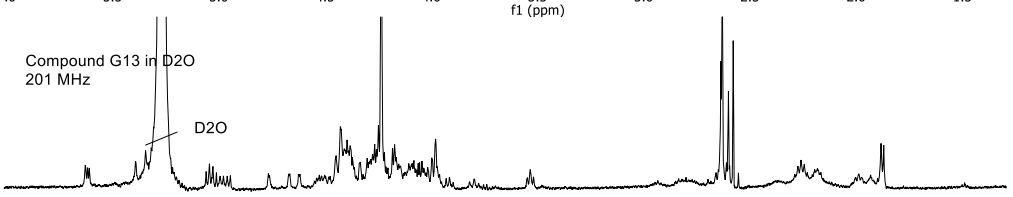

KB03-41-2CARDRIGE/3
f2 (ppm)
f1 (ppm)

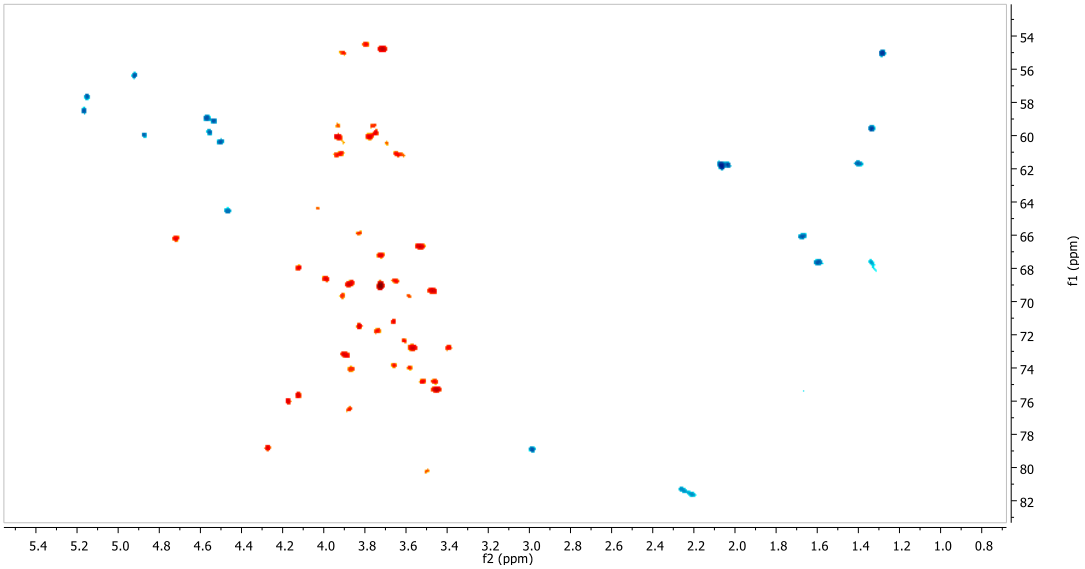

1H/KB03-64
1D 1H (10 uW water presaturation):
Sample KB03-64
T=297 K
Compound G12 in D2O
800 MHz
f1 (ppm)

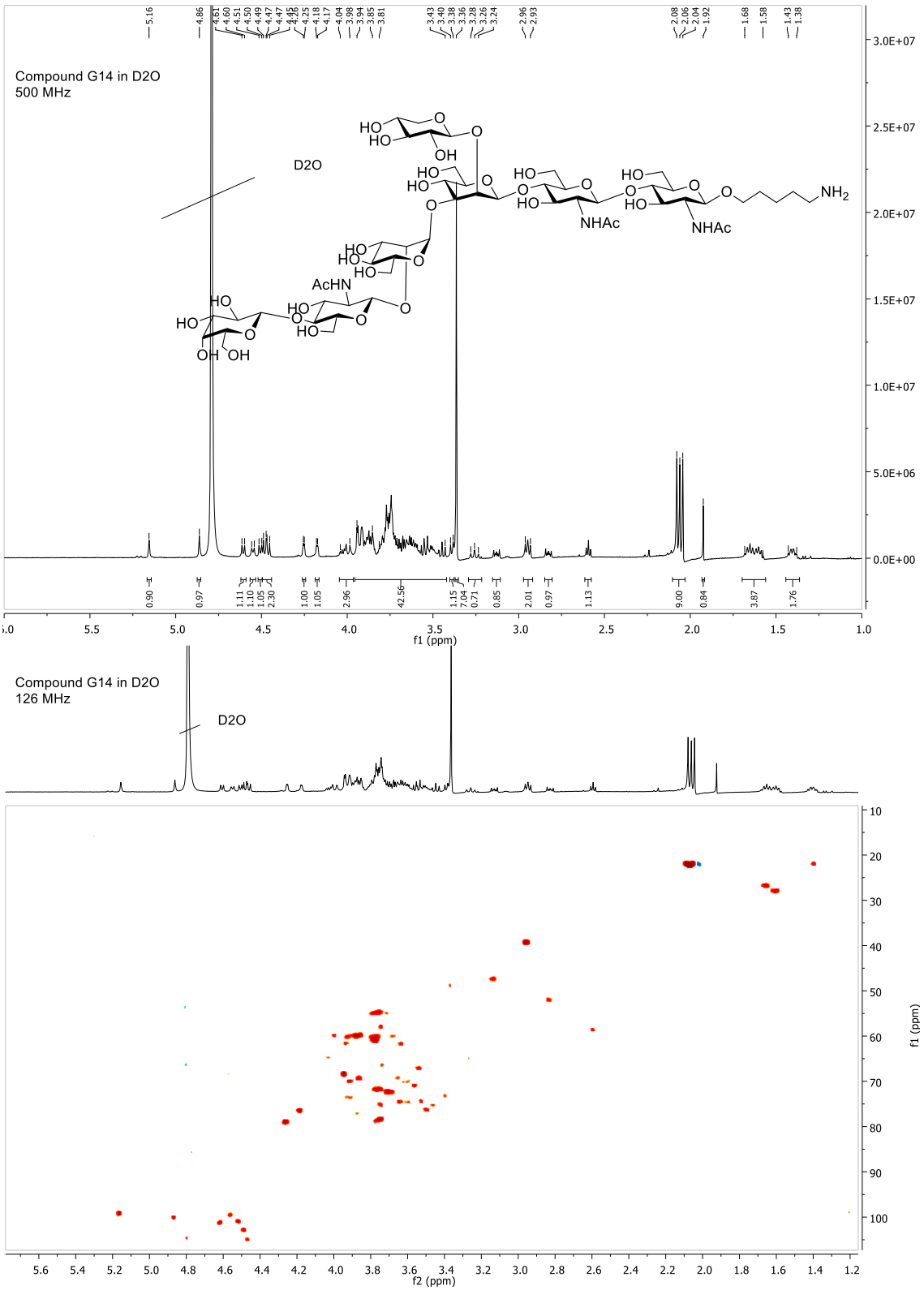

Compound G12 in D2O
201 MHz

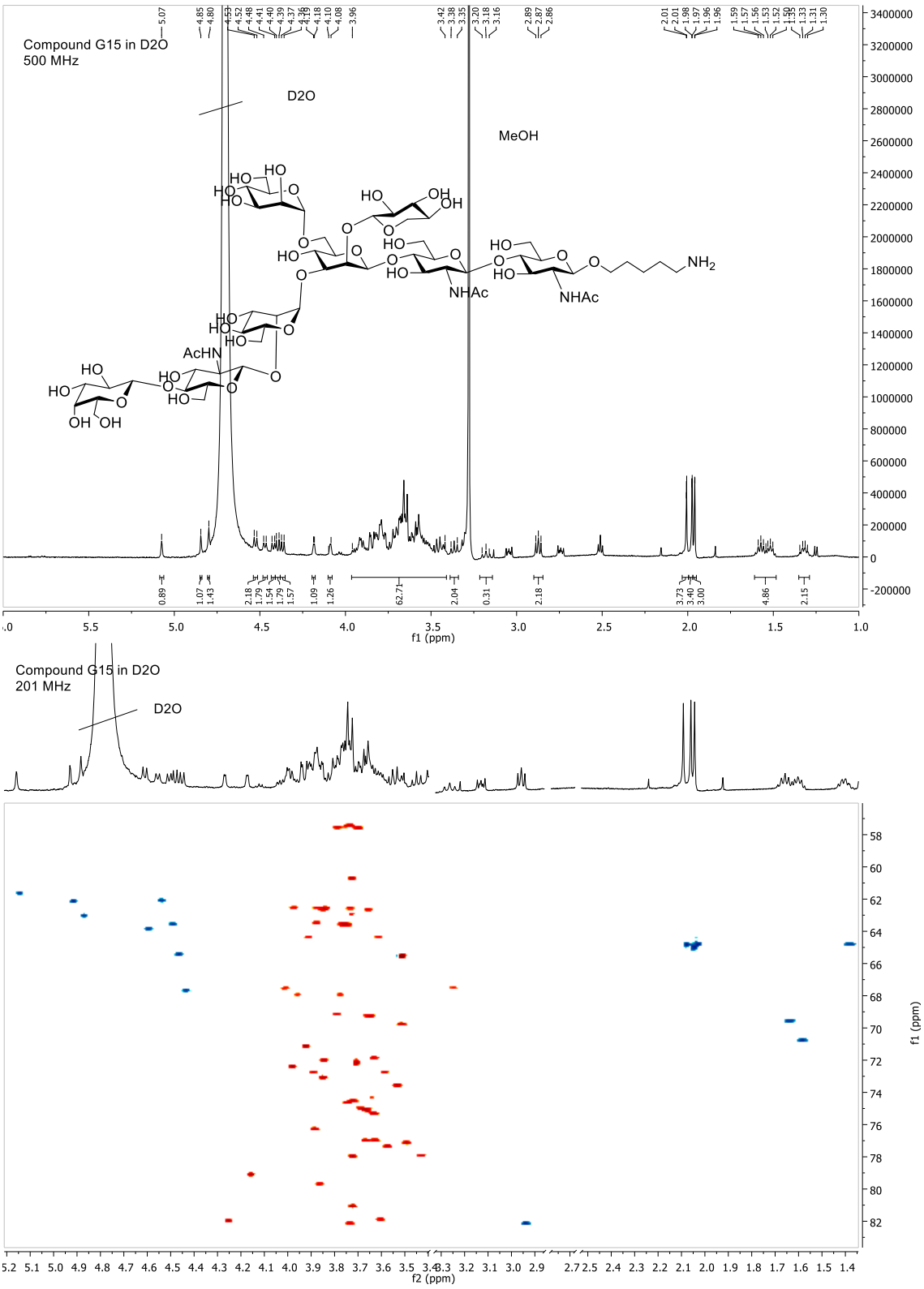

: FOLDED [46-86] ppm
f2 (ppm)
f1 (ppm)

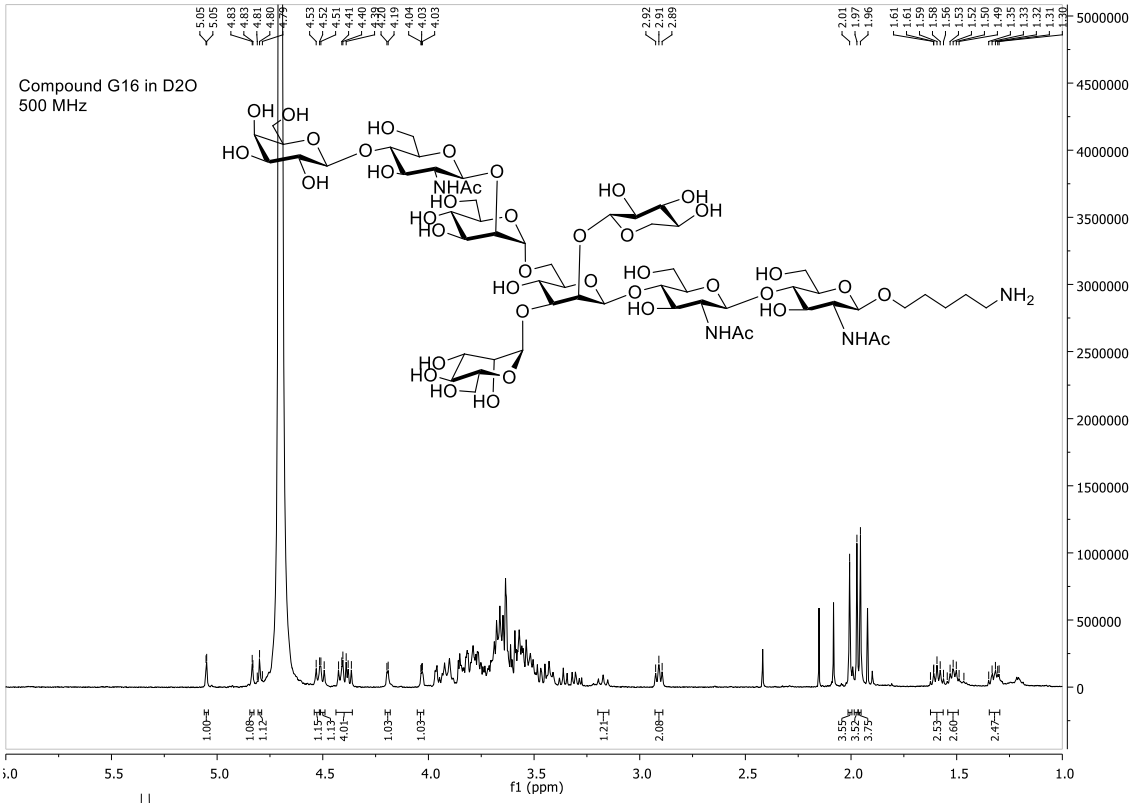

Compound G13 in D2O
500 MHz
D2O
NHAc
NHAc
NHAc
AcHN
OH
NH2
f1 (ppm)

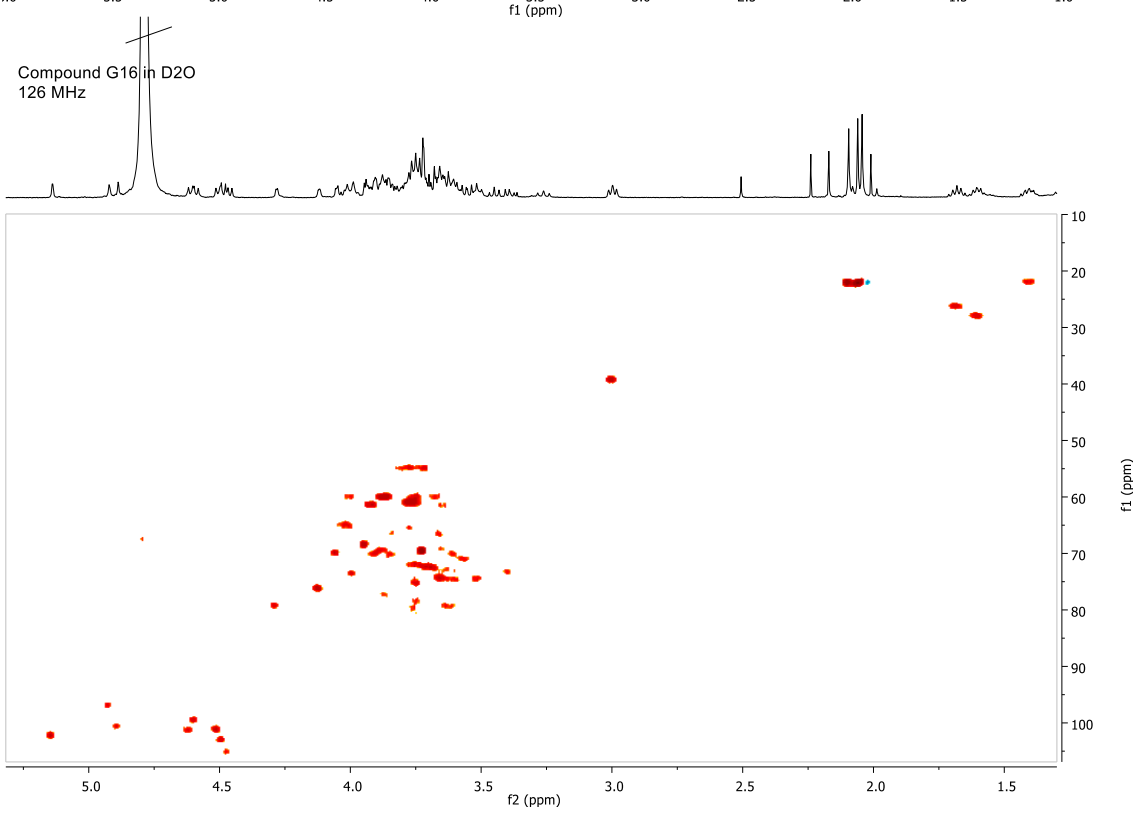

Compound G13 in D2O
201 MHz
D2O

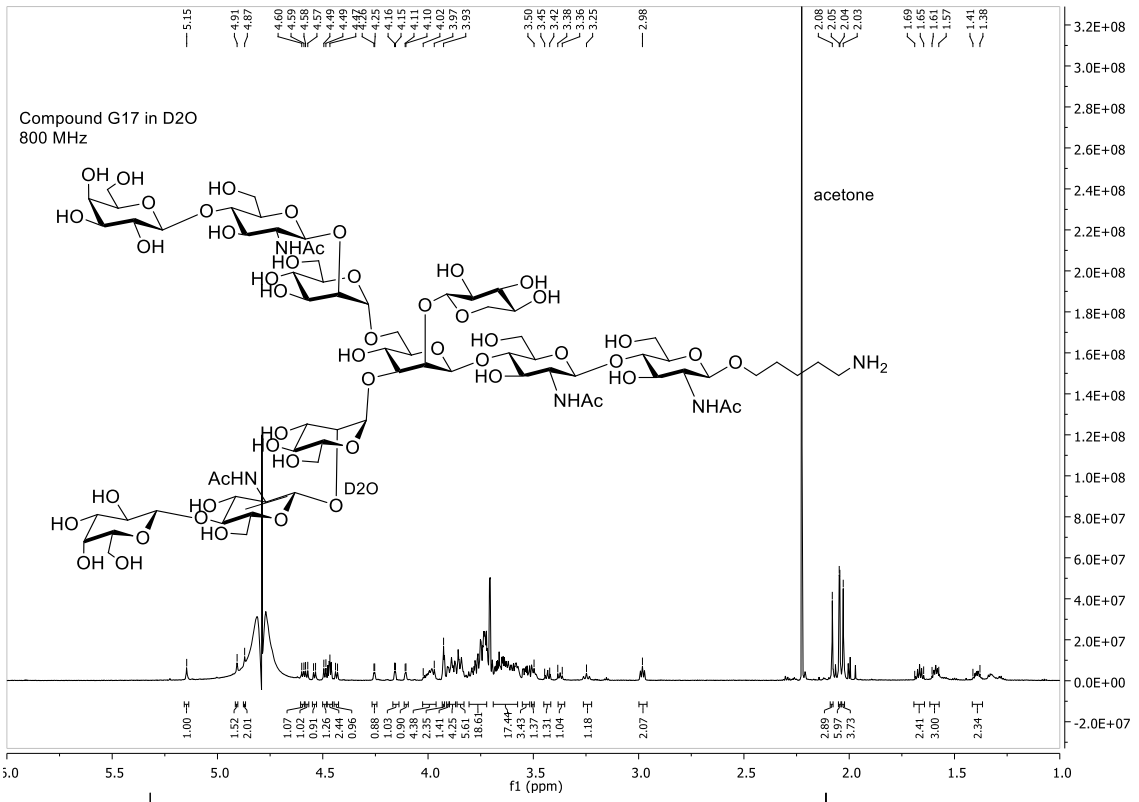

f2 (ppm)
f1 (ppm)

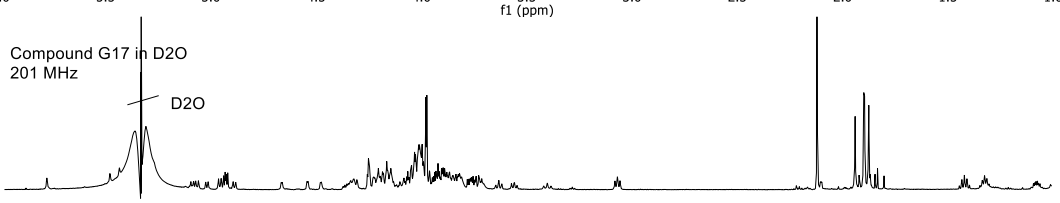

Compound G14 in D2O
500 MHz
D2O
NHAc
NHAc
AcHN
NH2
f1 (ppm)
Compound G14 in D2O
126 MHz
D2O
f1 (ppm)
f2 (ppm)

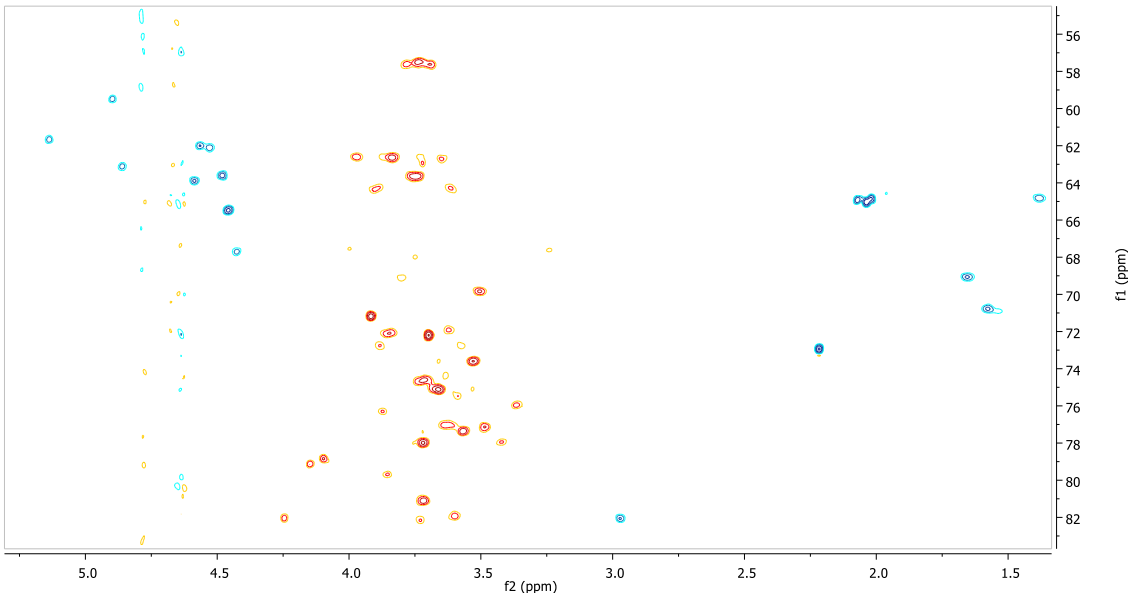

Compound G15 in D2O
500 MHz
D2O
MeOH
NHAc
NHAc
AcHN
NH2
f1 (ppm)
Compound G15 in D2O
201 MHz
D2O
f2 (ppm)
f1 (ppm)

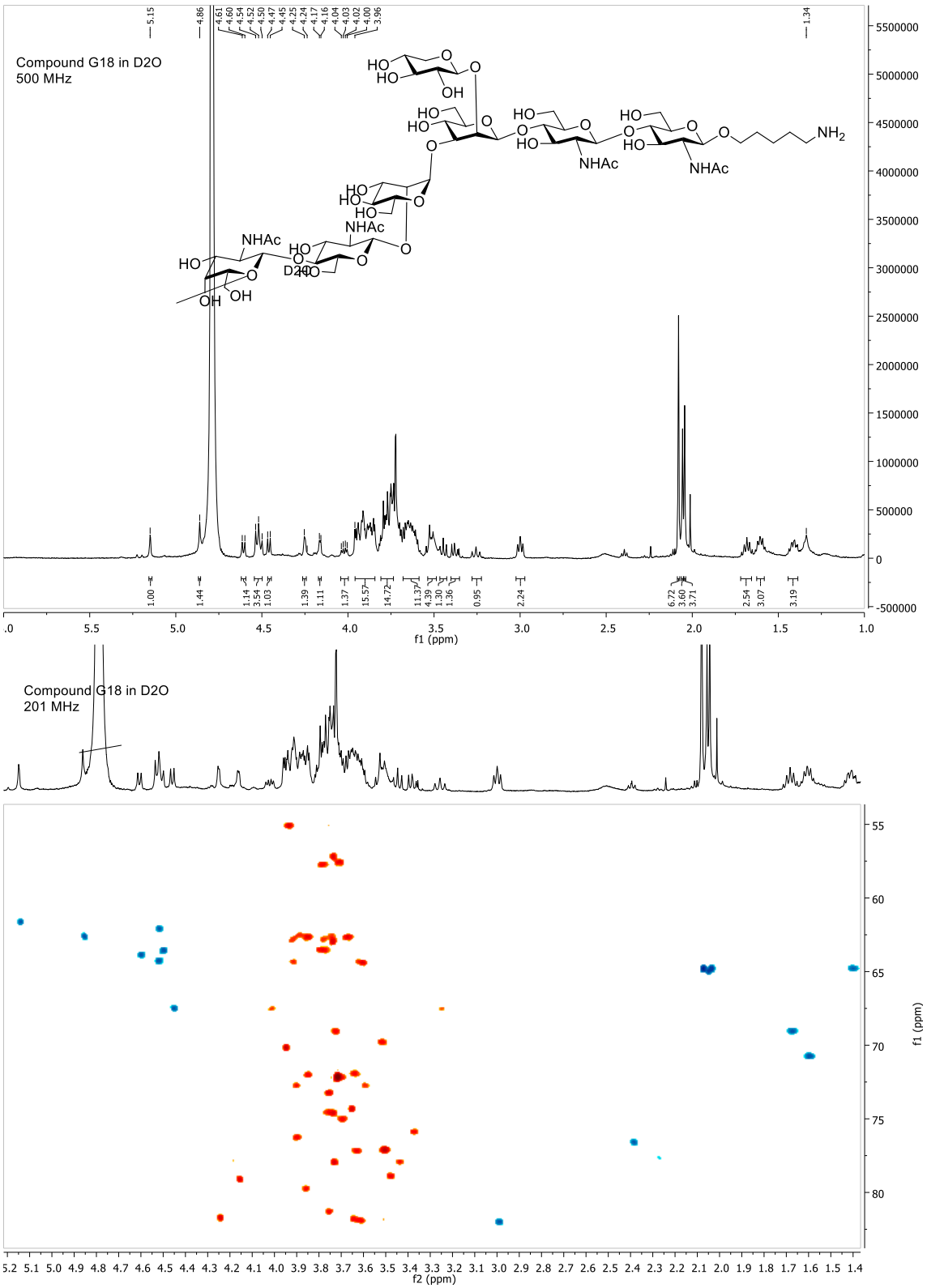

Compound G16 in D2O
500 MHz
f1 (ppm)

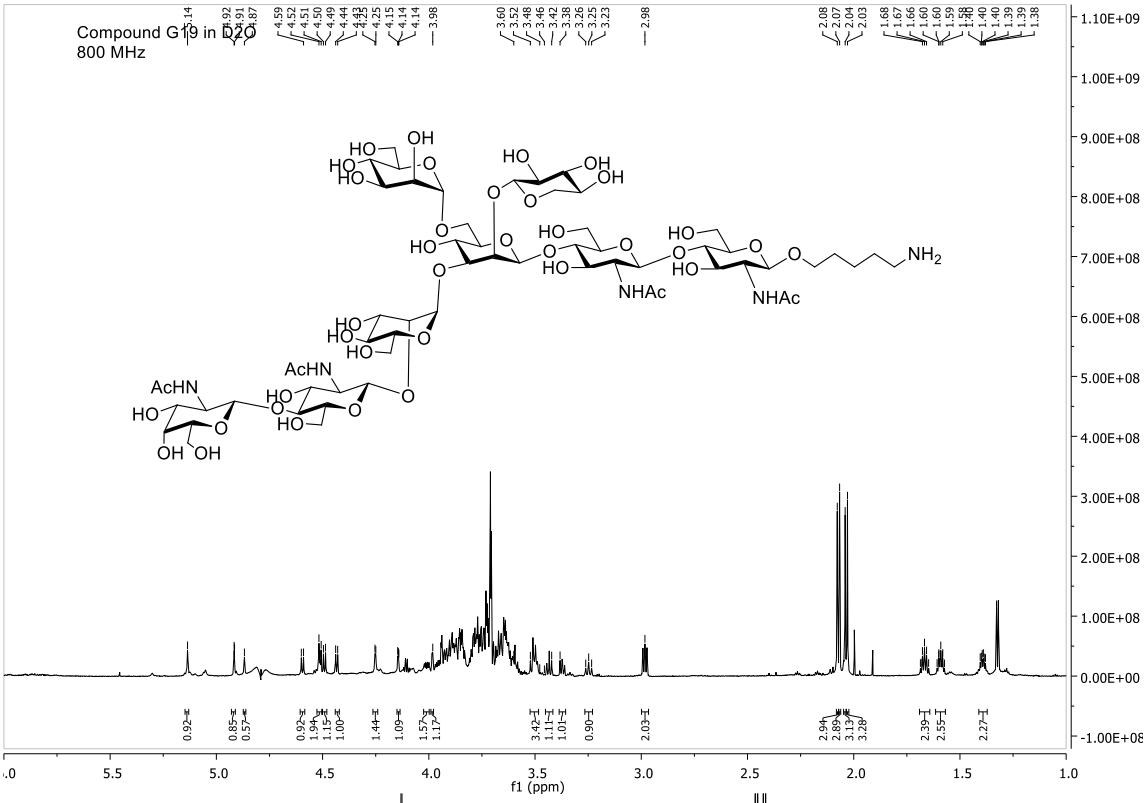

Compound G16 in D2O
126 MHz
f2 (ppm)
f1 (ppm)

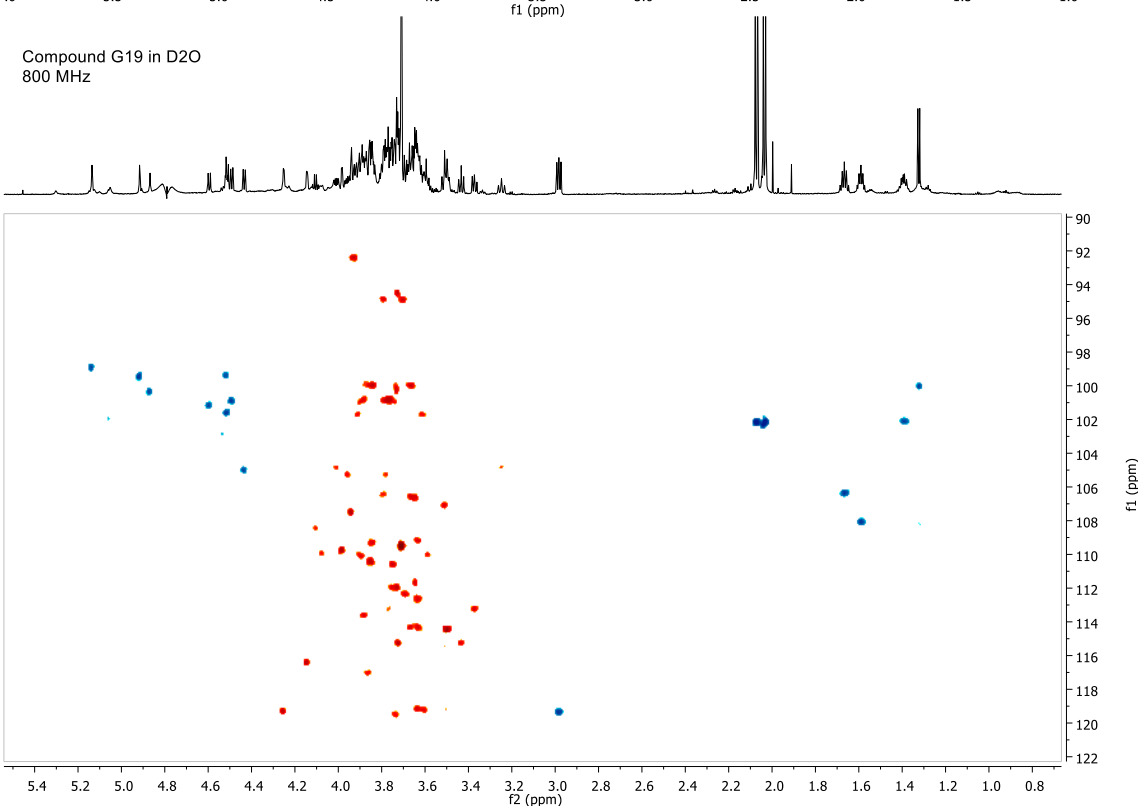

Compound G17 in D2O
800 MHz
acetone
D2O
AcHN
NHAc
NH2
f1 (ppm)

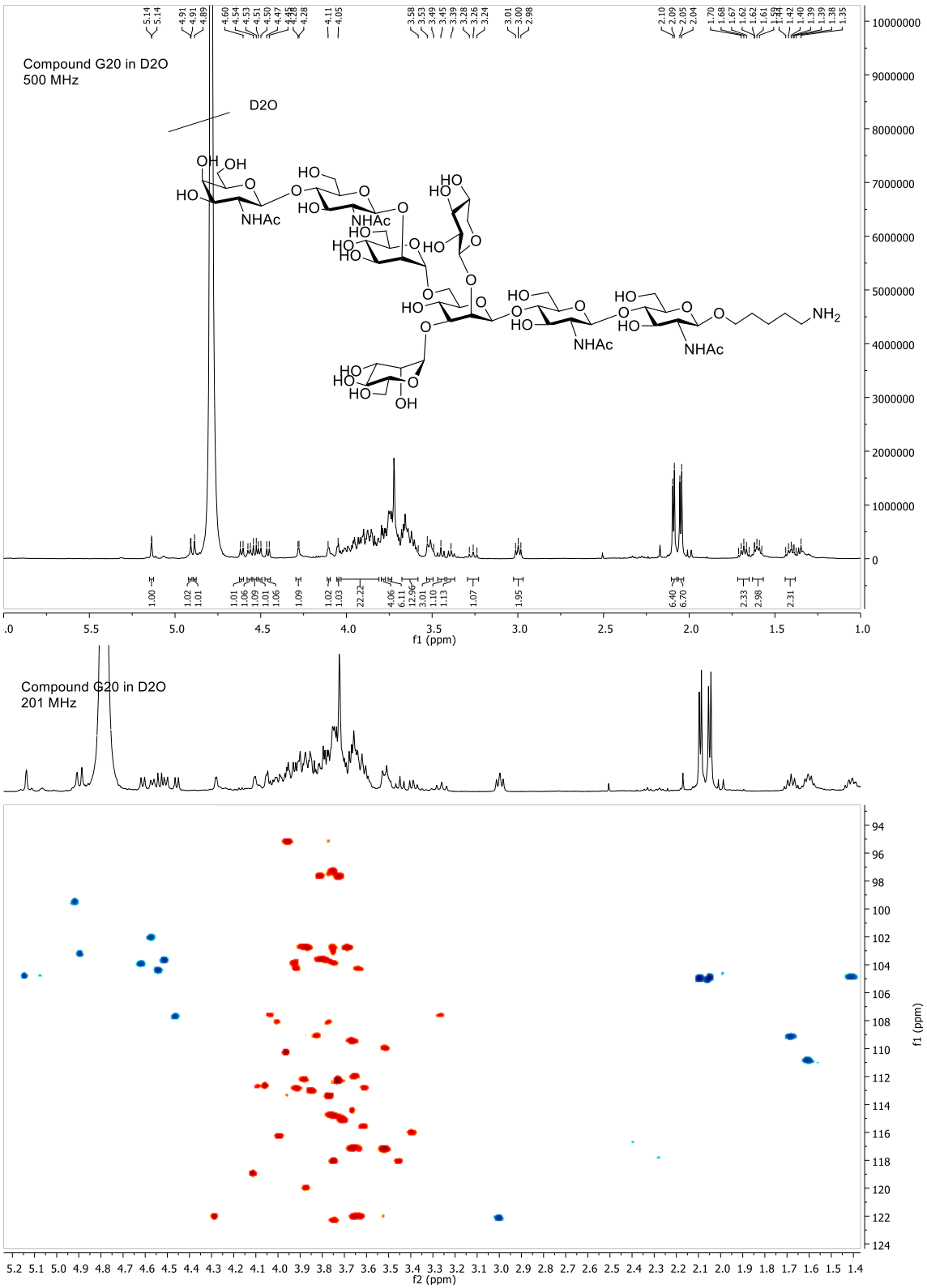

Compound G17 in D2O
201 MHz
D2O

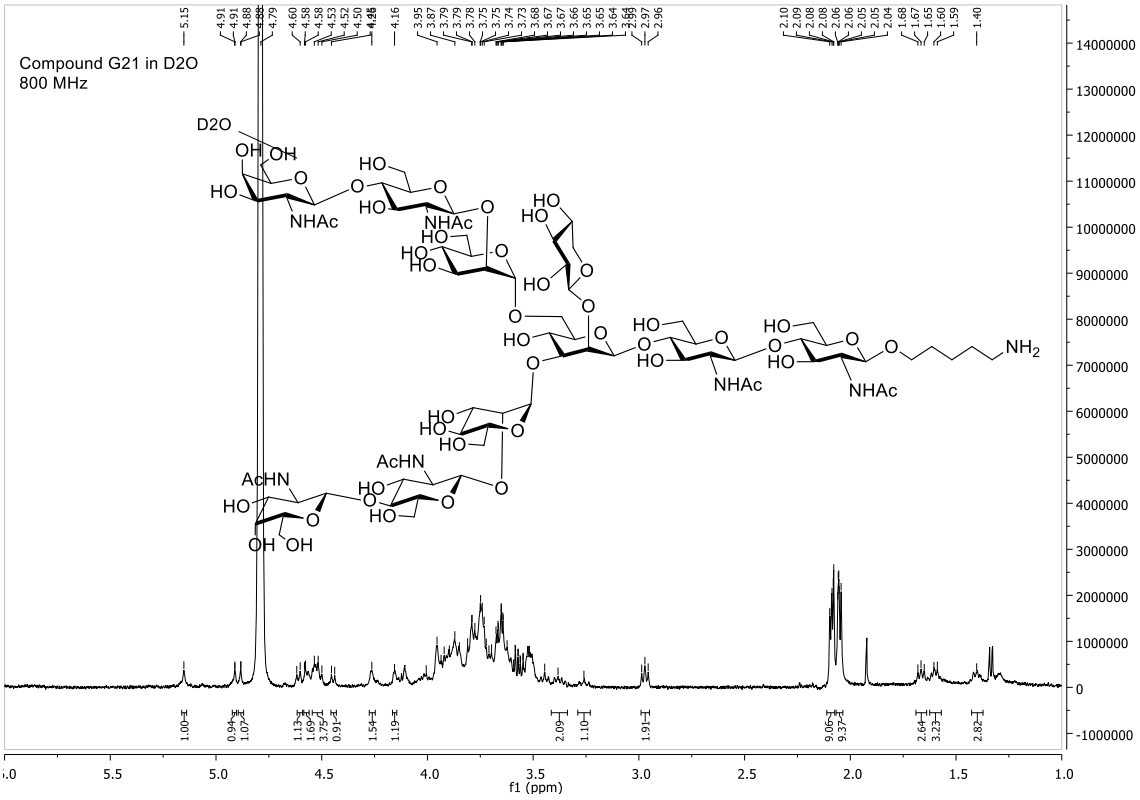

f2 (ppm)
f1 (ppm)

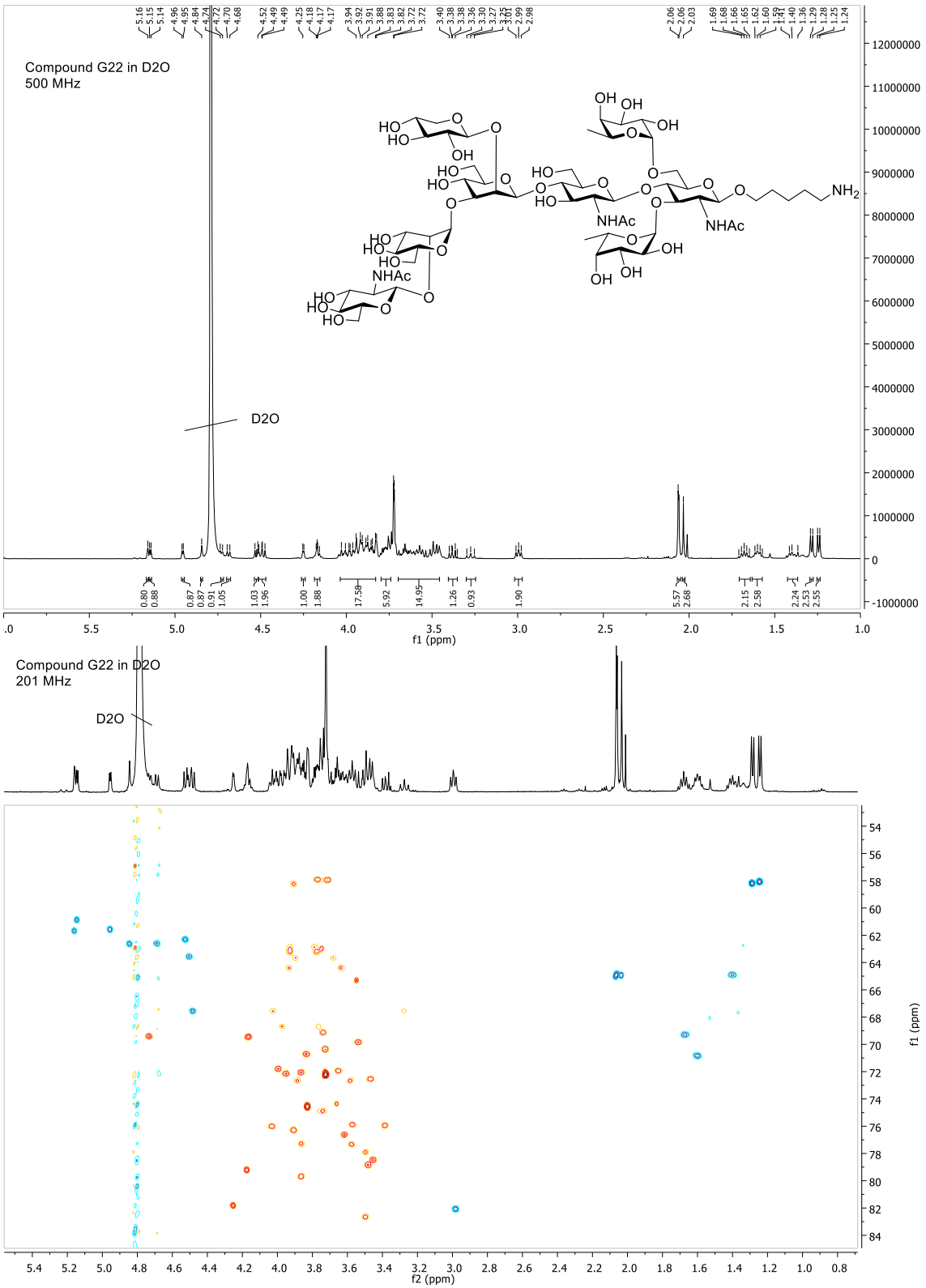

Compound G18 in D2O
500 MHz
Compound G18 in D2O
201 MHz
f1 (ppm)
f2 (ppm)

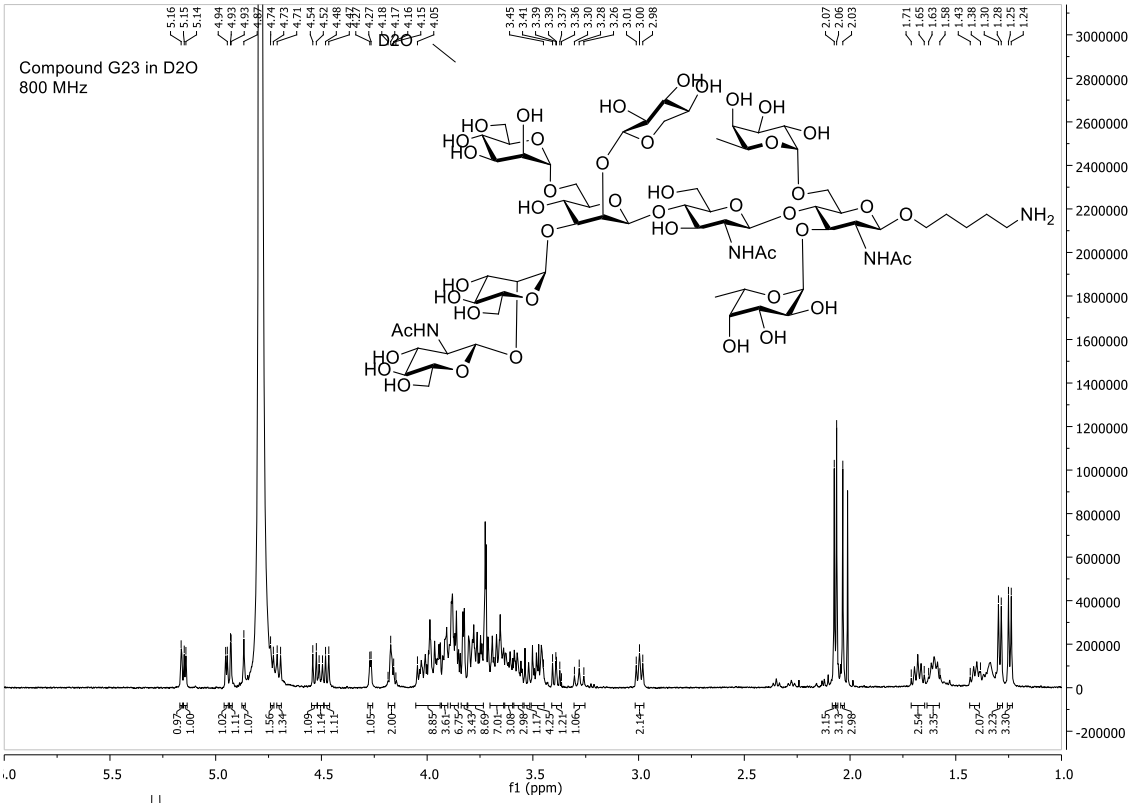

Compound G19 in D2O
800 MHz
f1 (ppm)

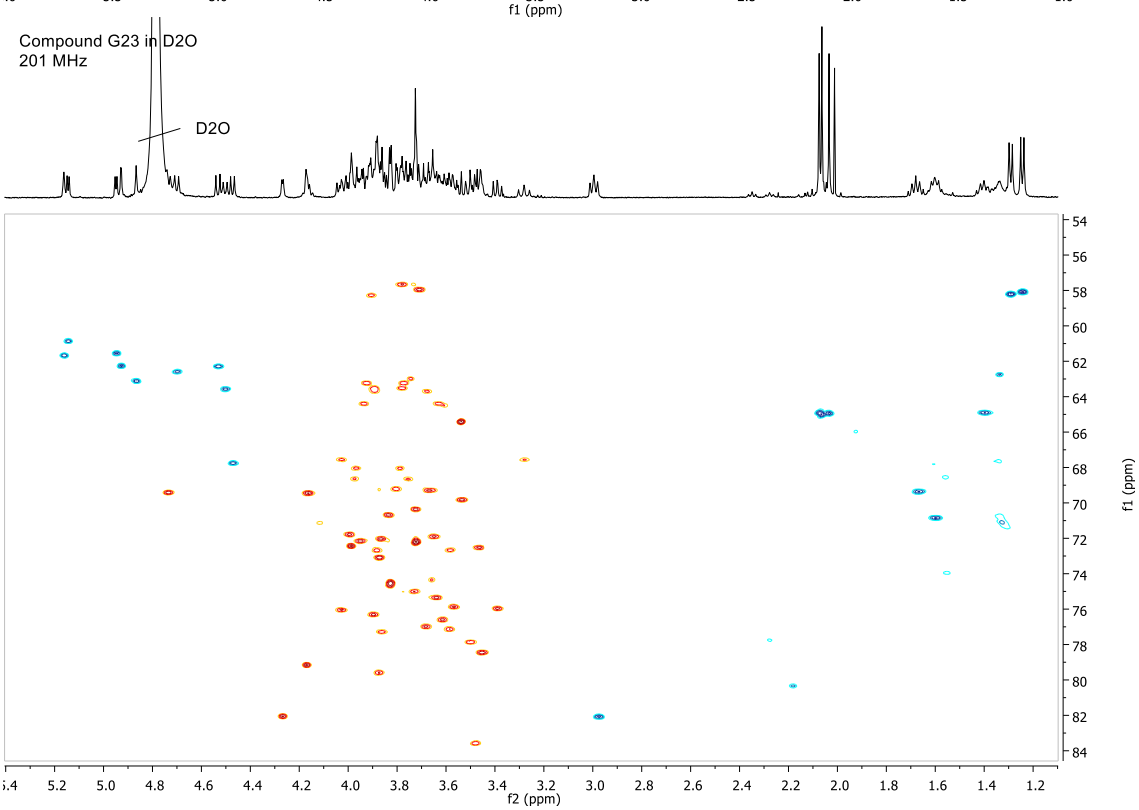

Compound G19 in D2O
800 MHz
f2 (ppm)
f1 (ppm)

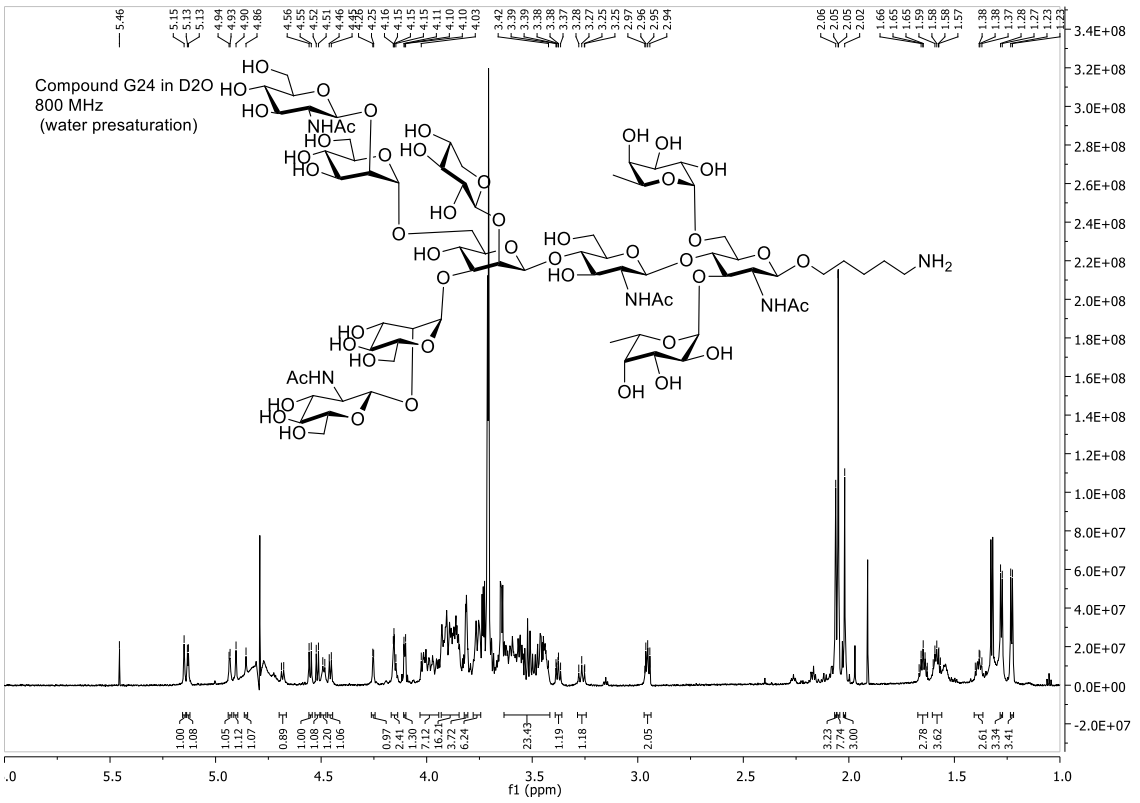

Compound G20 in D2O
500 MHz
D2O
NHAc
NHAc
NHAc
NHAc
NH2
OH
f1 (ppm)
Compound G20 in D2O
201 MHz
f1 (ppm)
f2 (ppm)

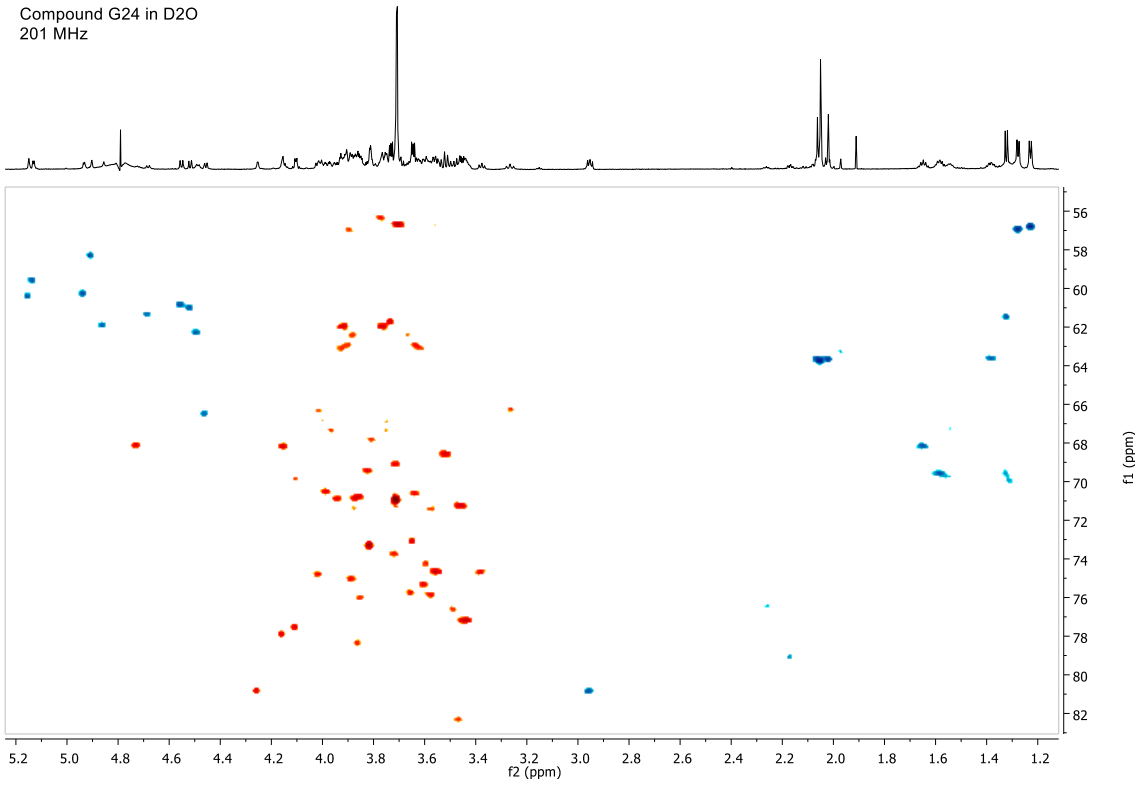

Compound G21 in D2O
800 MHz
D2O
NHAc
AcHN
NH2
f1 (ppm)

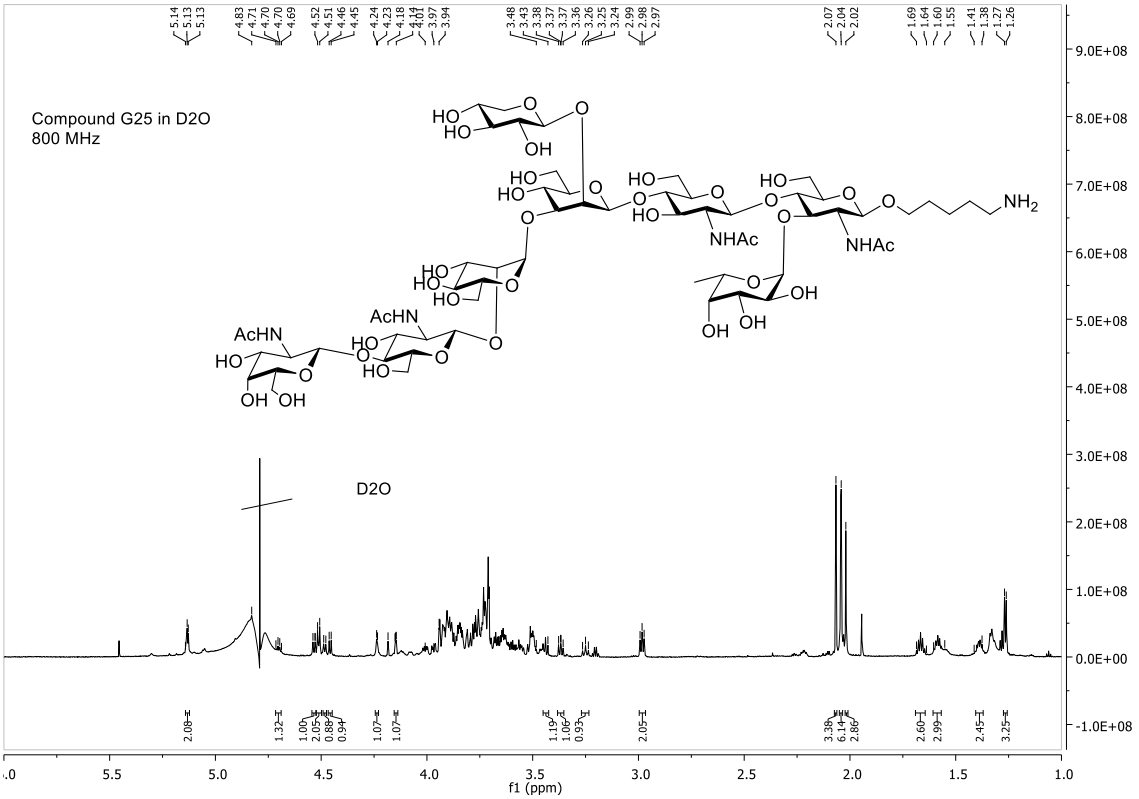

Compound G22 in D2O
500 MHz
D2O
f1 (ppm)
Compound G22 in D2O
201 MHz
D2O
f2 (ppm)
f1 (ppm)

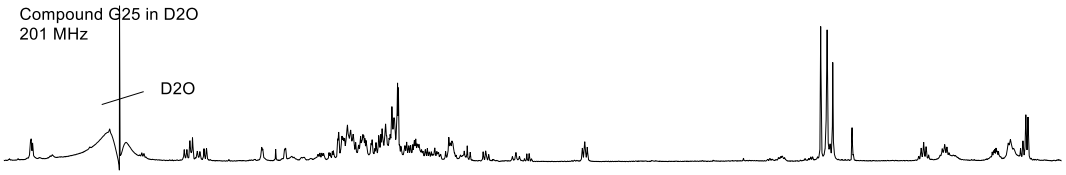

Compound G23 in D2O
800 MHz
D2O
f1 (ppm)

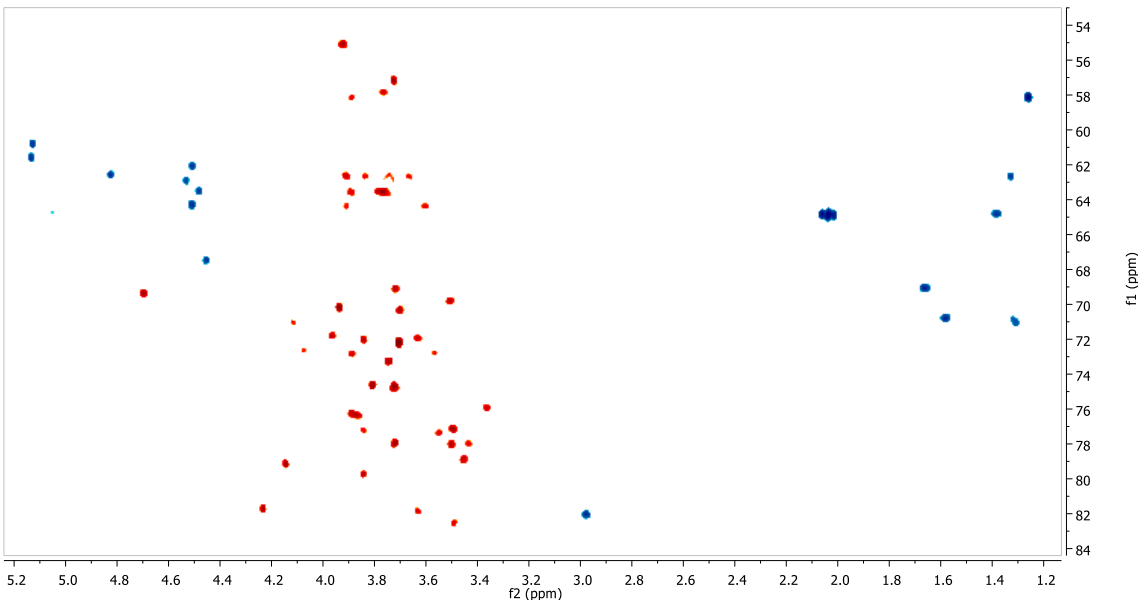

Compound G23 in D2O
201 MHz
D2O
f2 (ppm)
f1 (ppm)

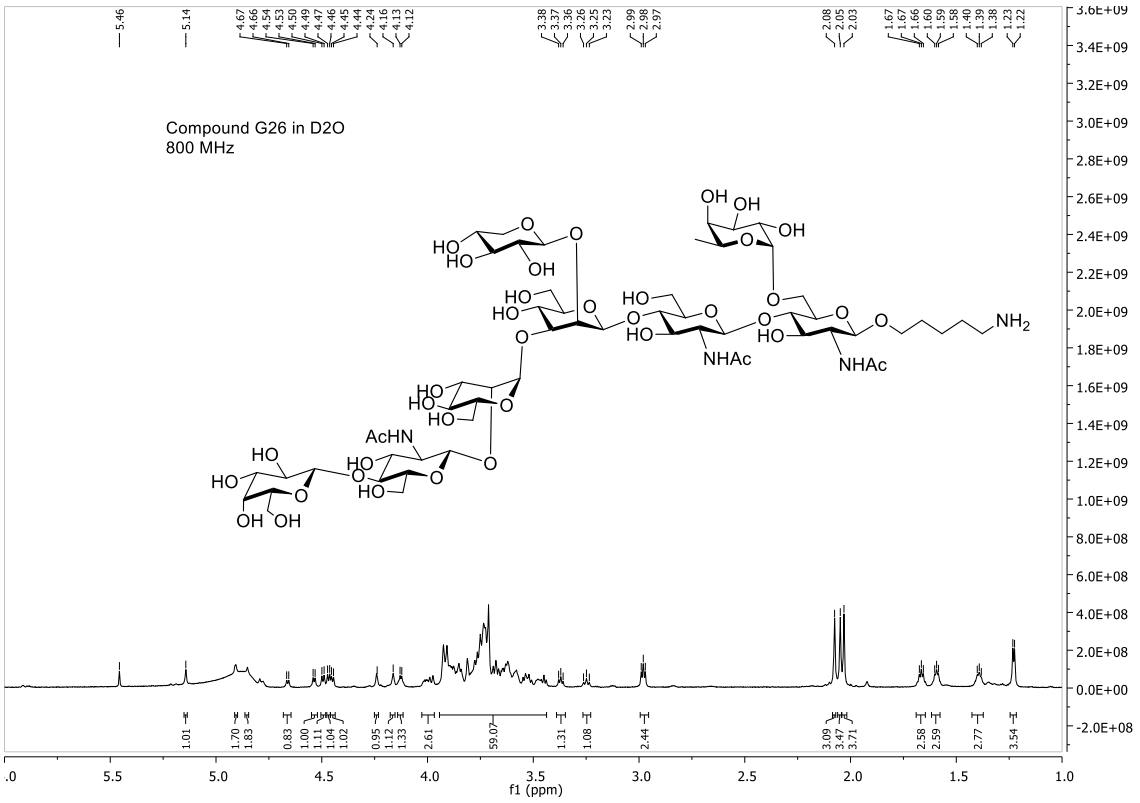

Compound G24 in D2O
800 MHz
(water presaturation)
f1 (ppm)

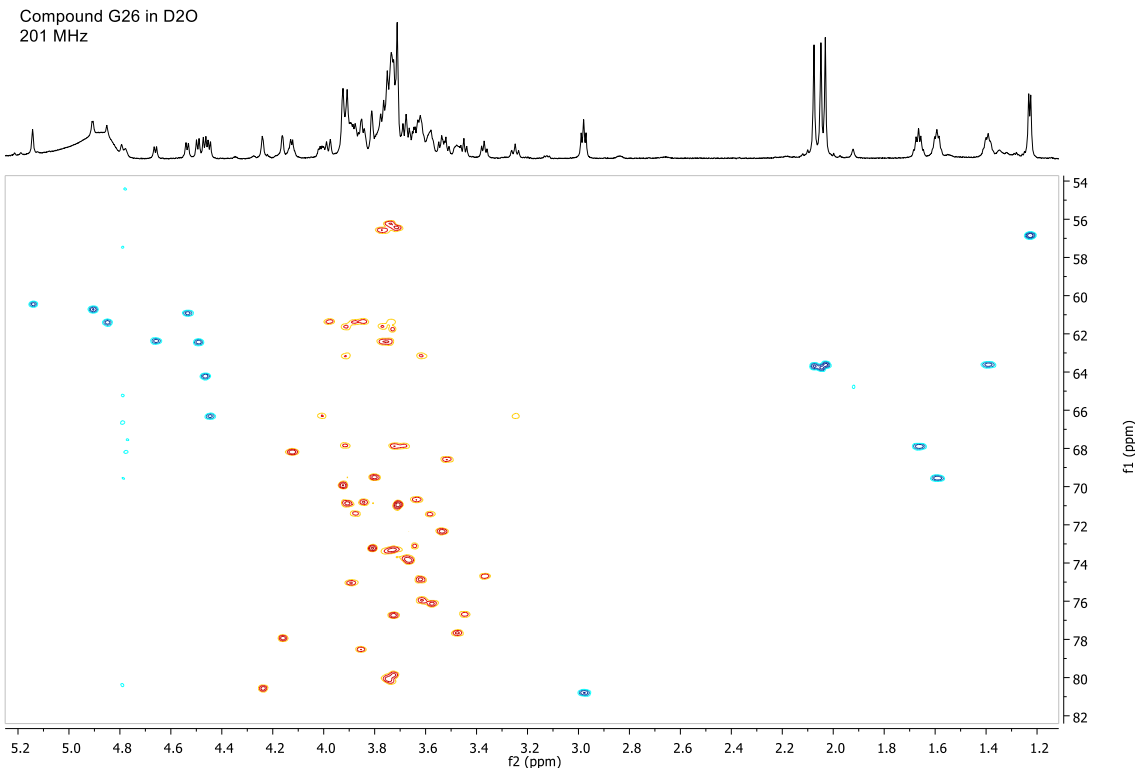

Compound G24 in D2O
201 MHz
f2 (ppm)
f1 (ppm)

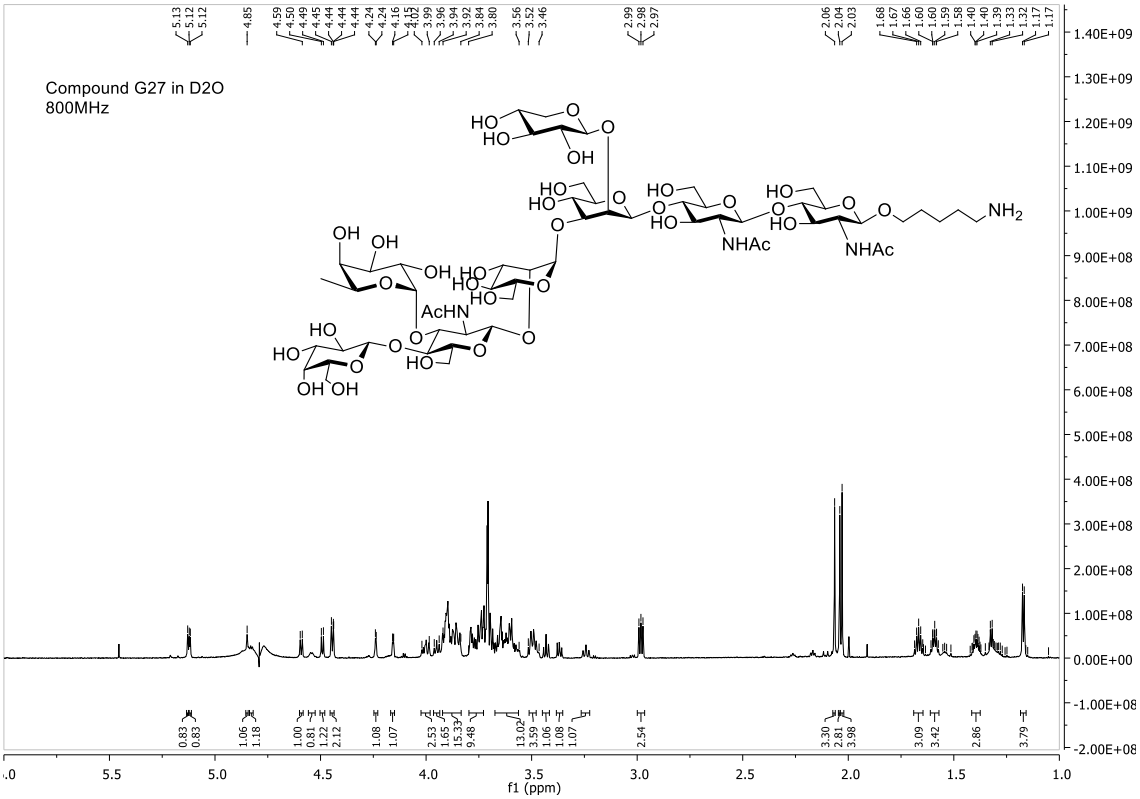

Compound G25 in D2O
800 MHz
D2O
f1 (ppm)

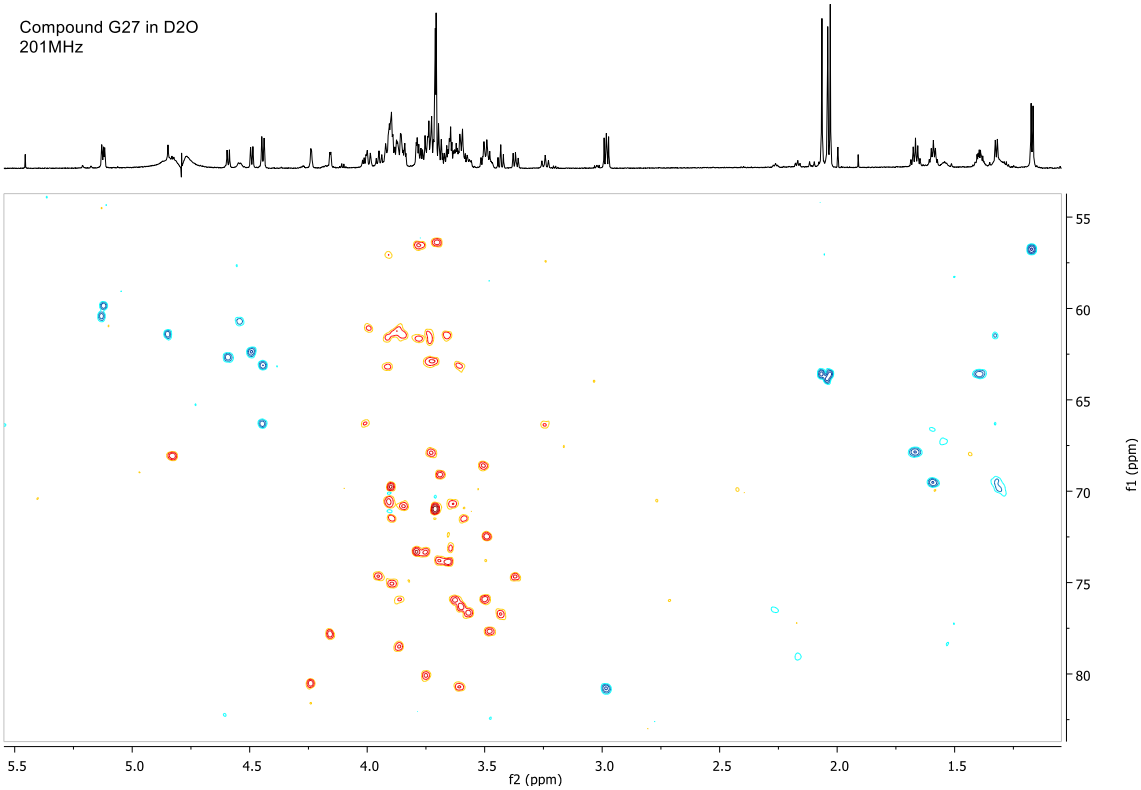

Compound G25 in D2O
201 MHz
D2O

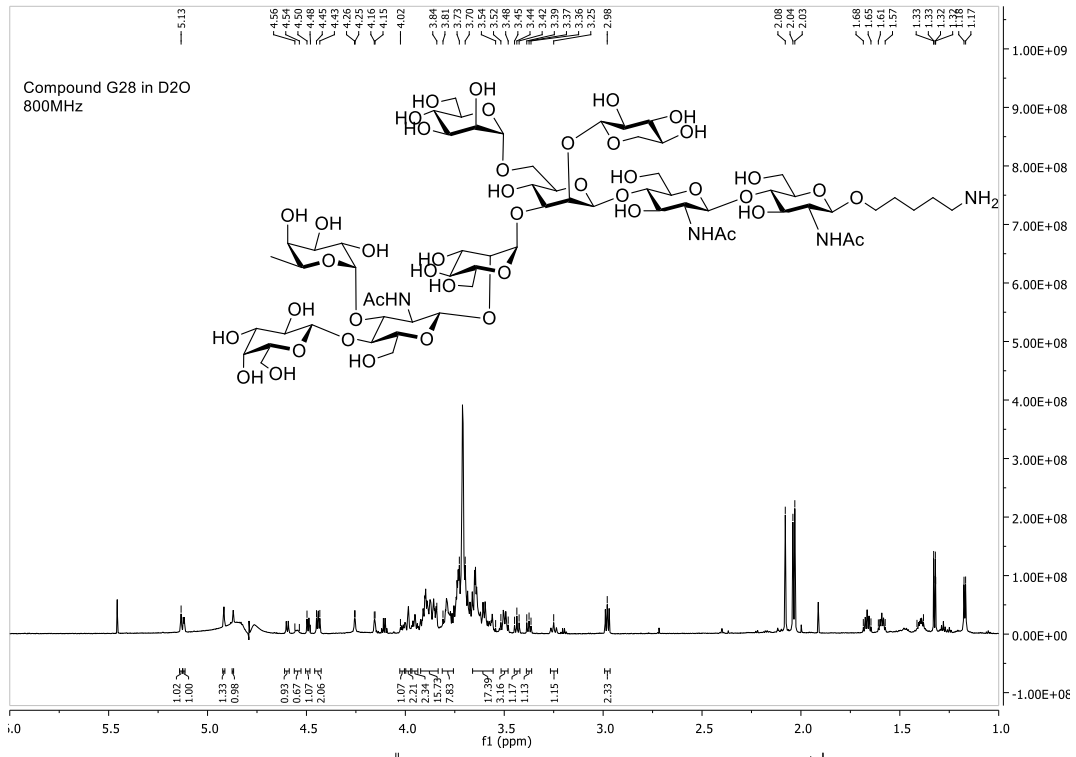

f2 (ppm)
f1 (ppm)

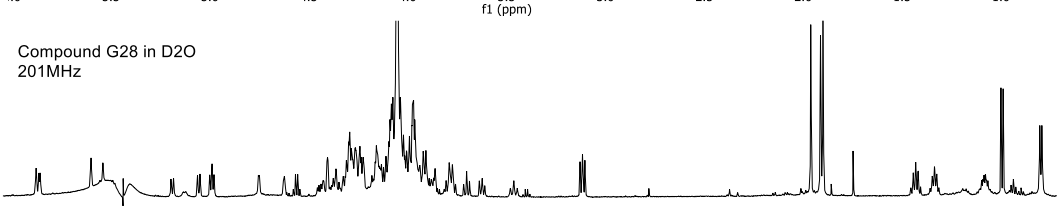

Compound G26 in D2O
800 MHz
f1 (ppm)

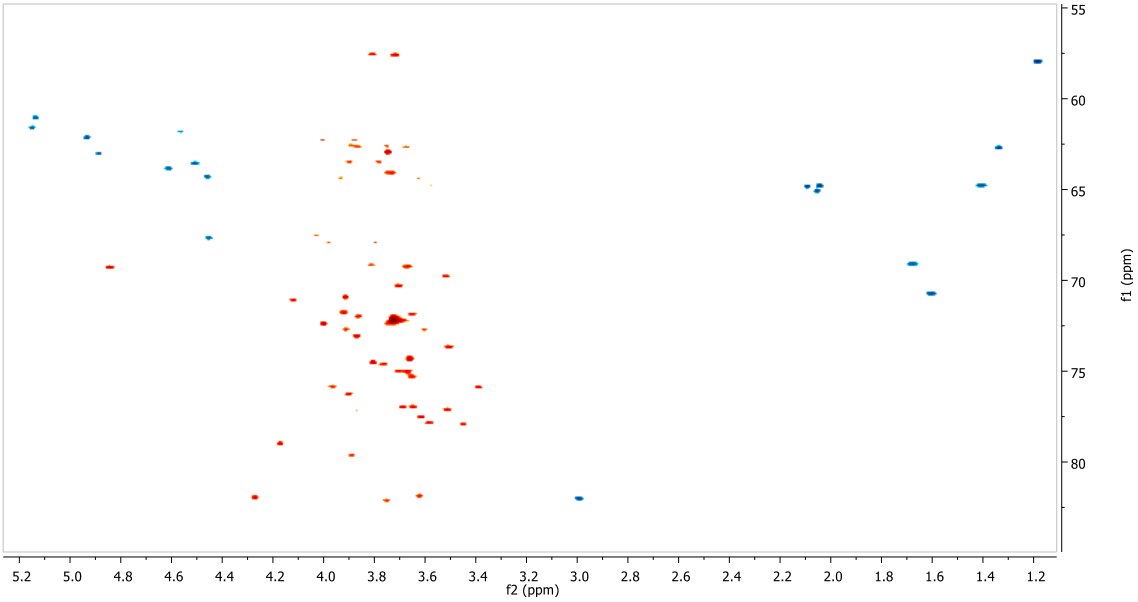

Compound G26 in D2O
201 MHz
f2 (ppm)
f1 (ppm)

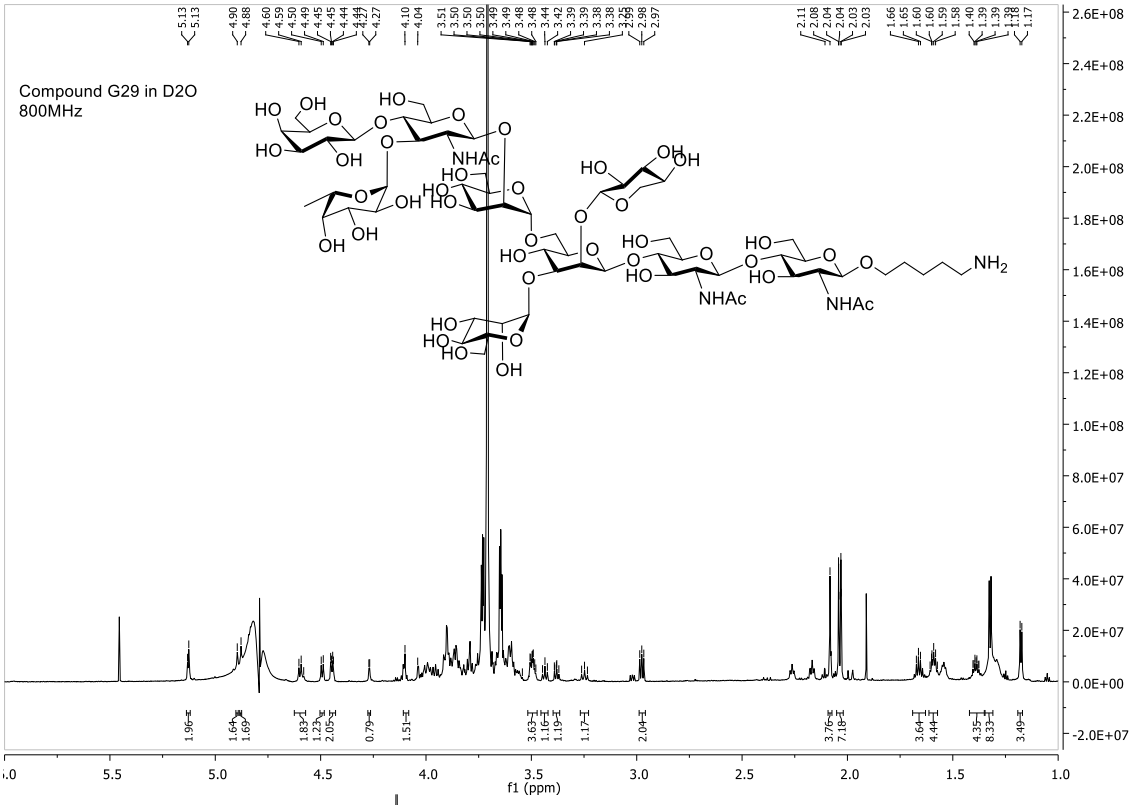

Compound G27 in D2O
800MHz
f1 (ppm)

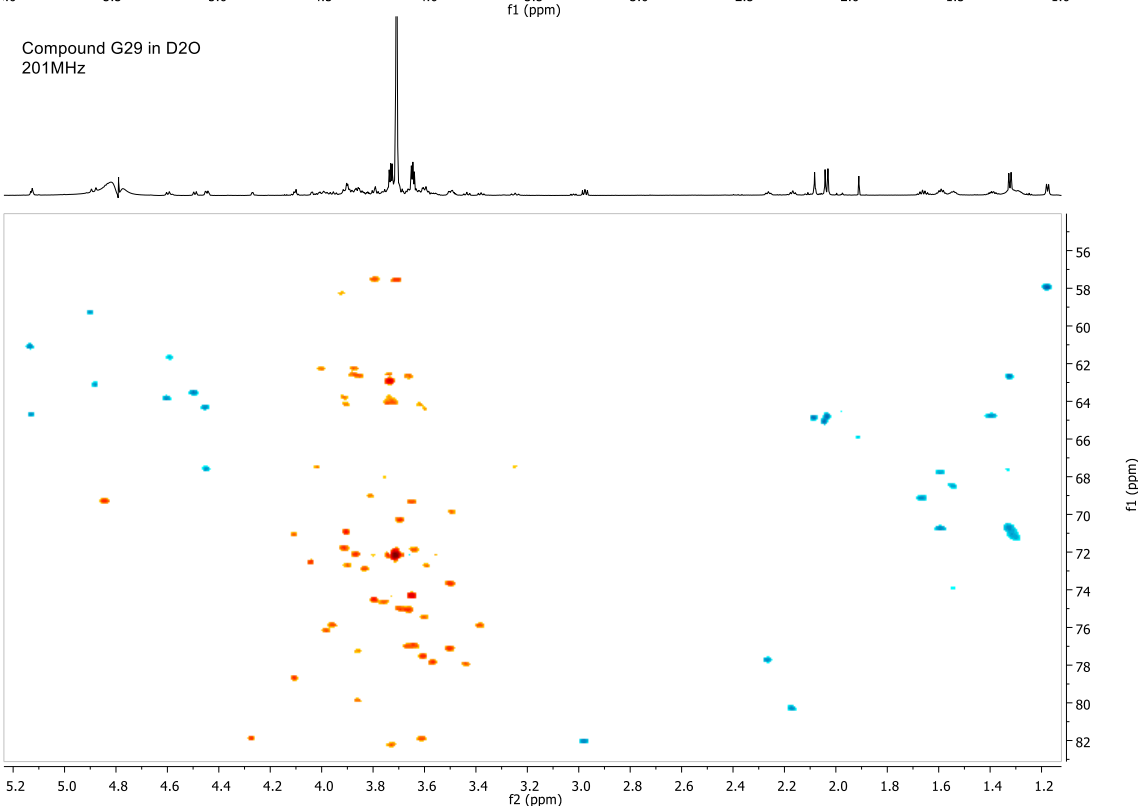

Compound G27 in D2O
201MHz
f2 (ppm)
f1 (ppm)

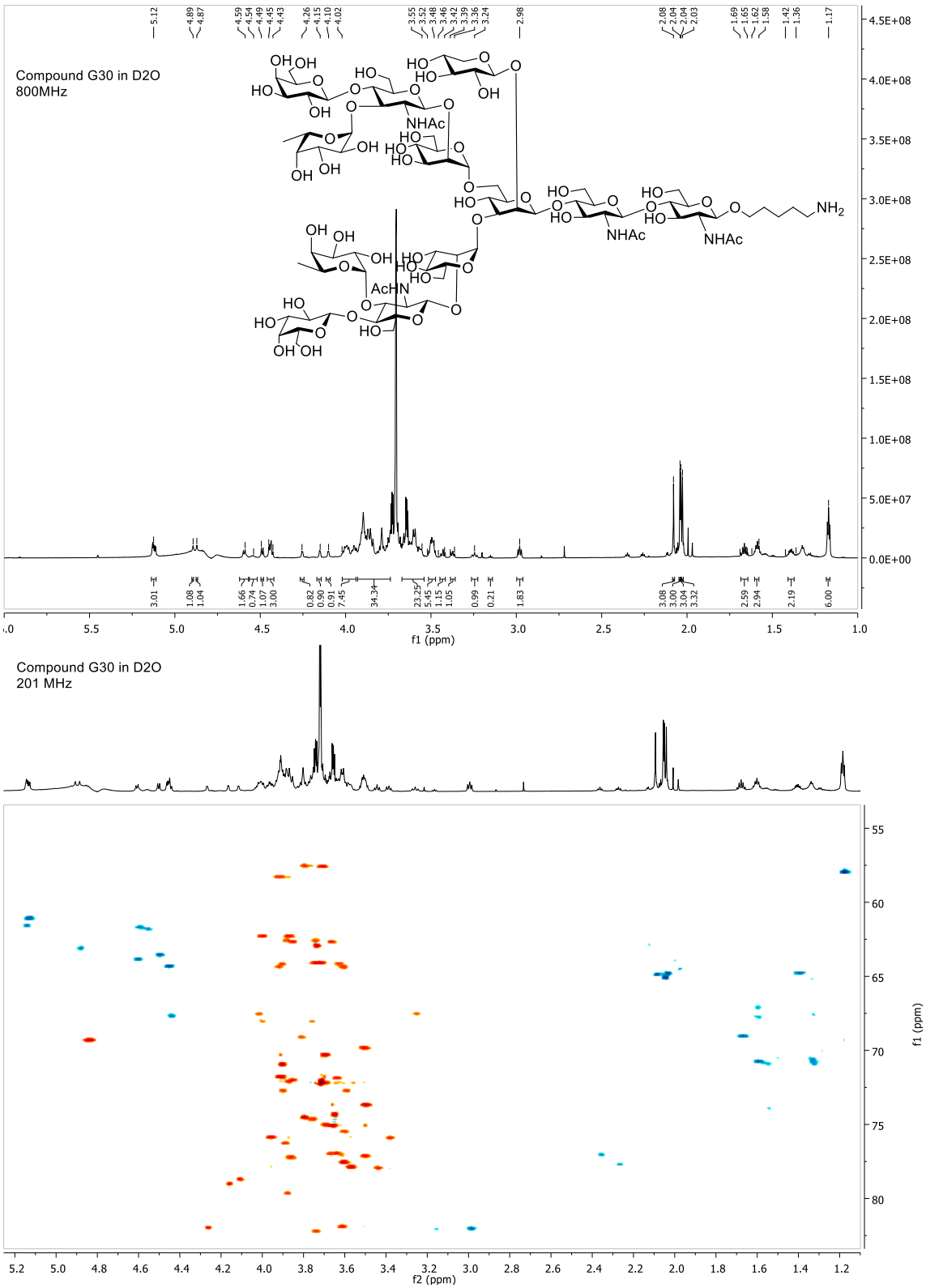

Compound G28 in D2O
800MHz
f1 (ppm)

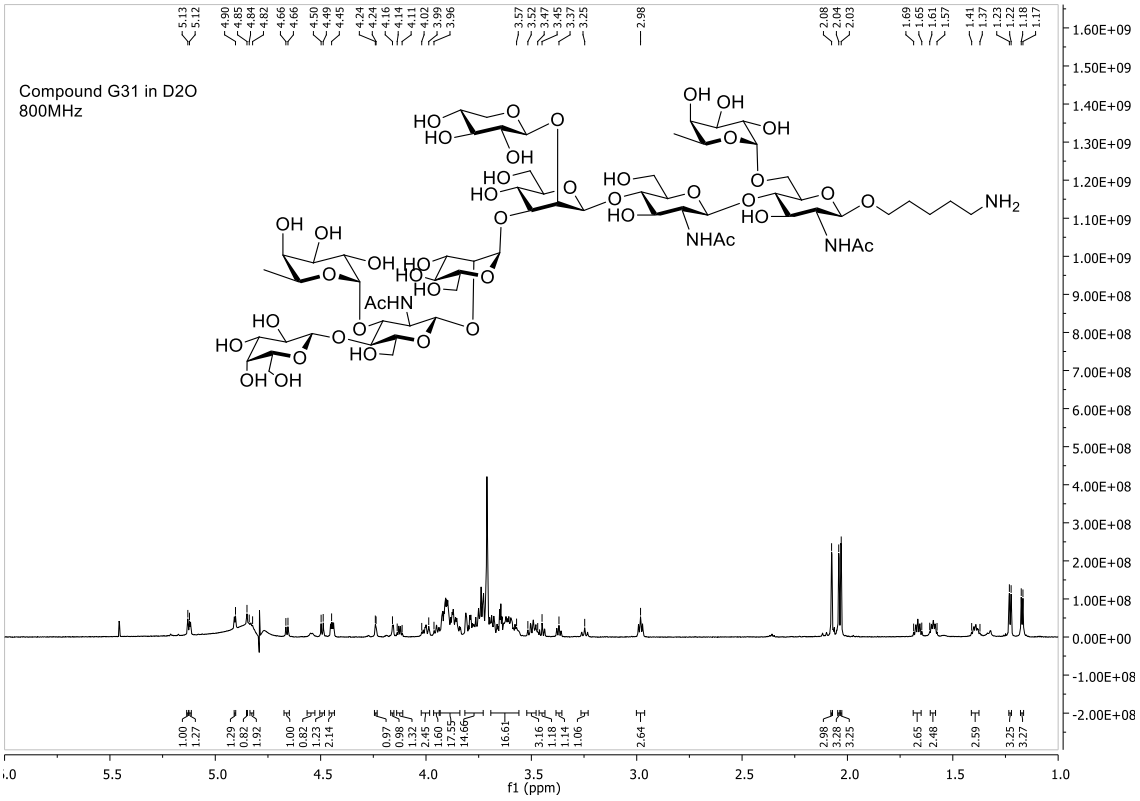

Compound G28 in D2O
201MHz

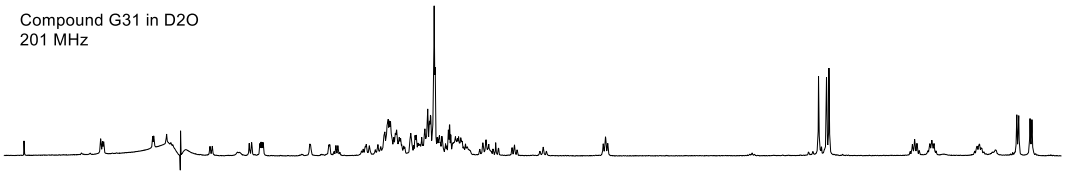

f2 (ppm)
f1 (ppm)

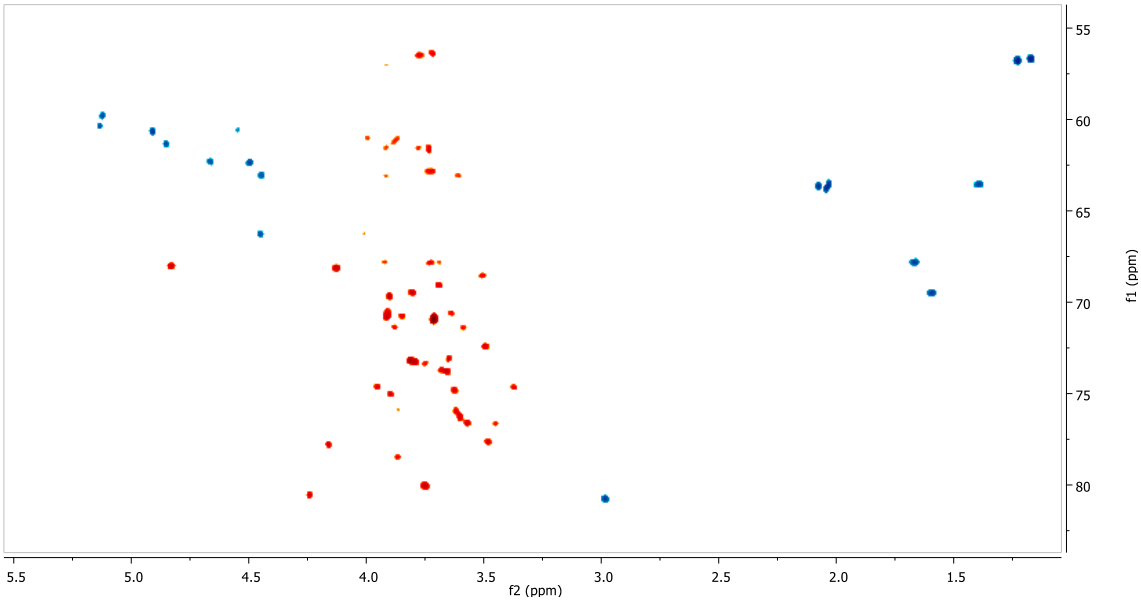

Compound G29 in D2O
800MHz
f1 (ppm)

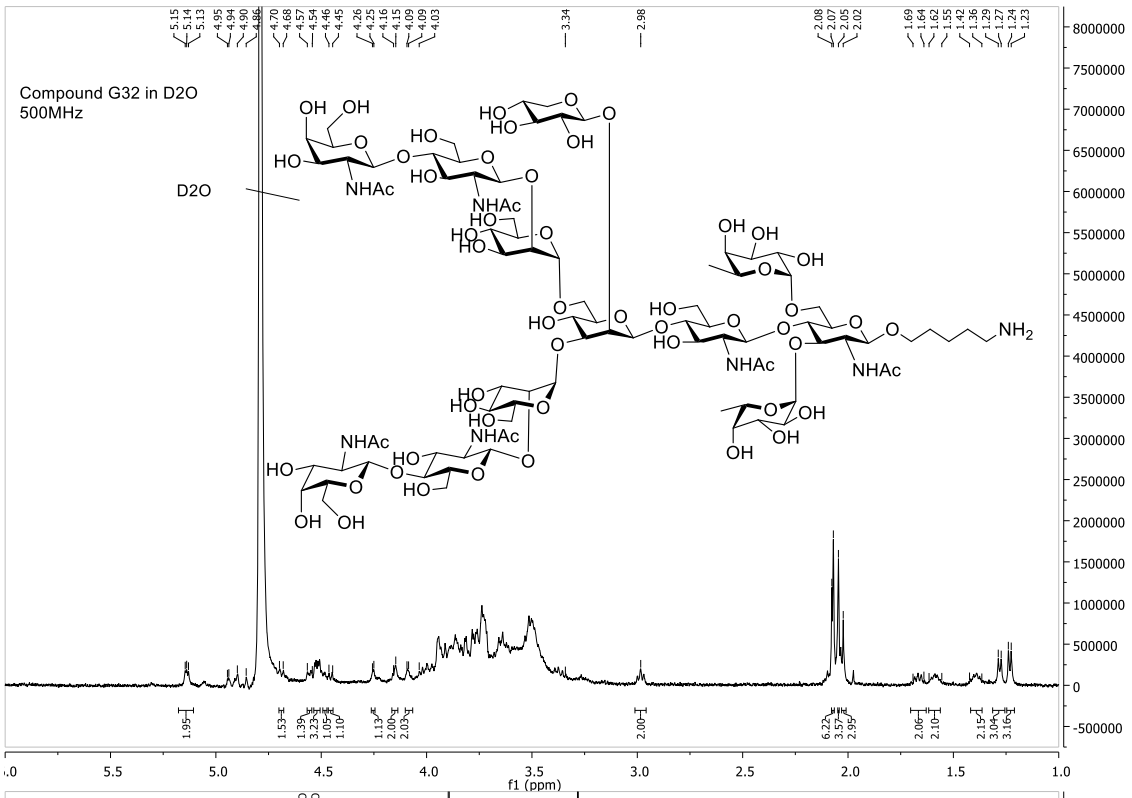

Compound G29 in D2O
201MHz
f2 (ppm)
f1 (ppm)

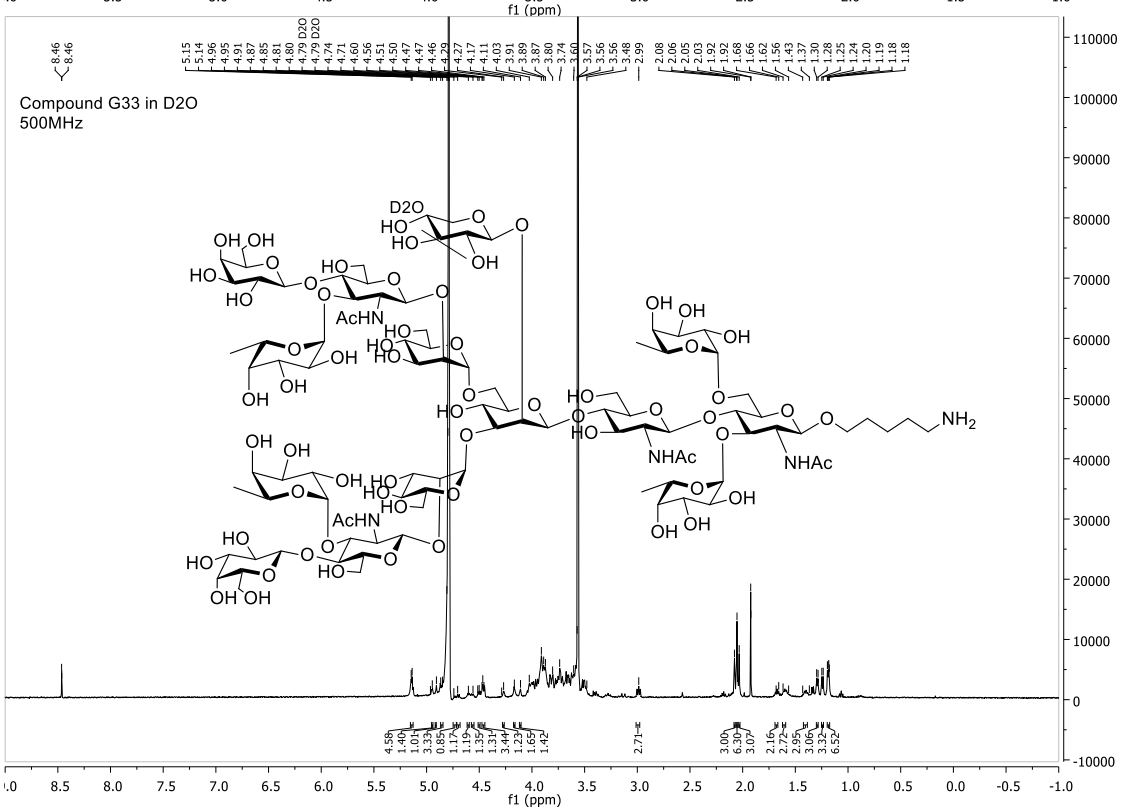

Compound G30 in D2O
800MHz
NHAc
NHAc
NHAc
AcHN
NH2
f1 (ppm)
Compound G30 in D2O
201 MHz
f1 (ppm)
f2 (ppm)

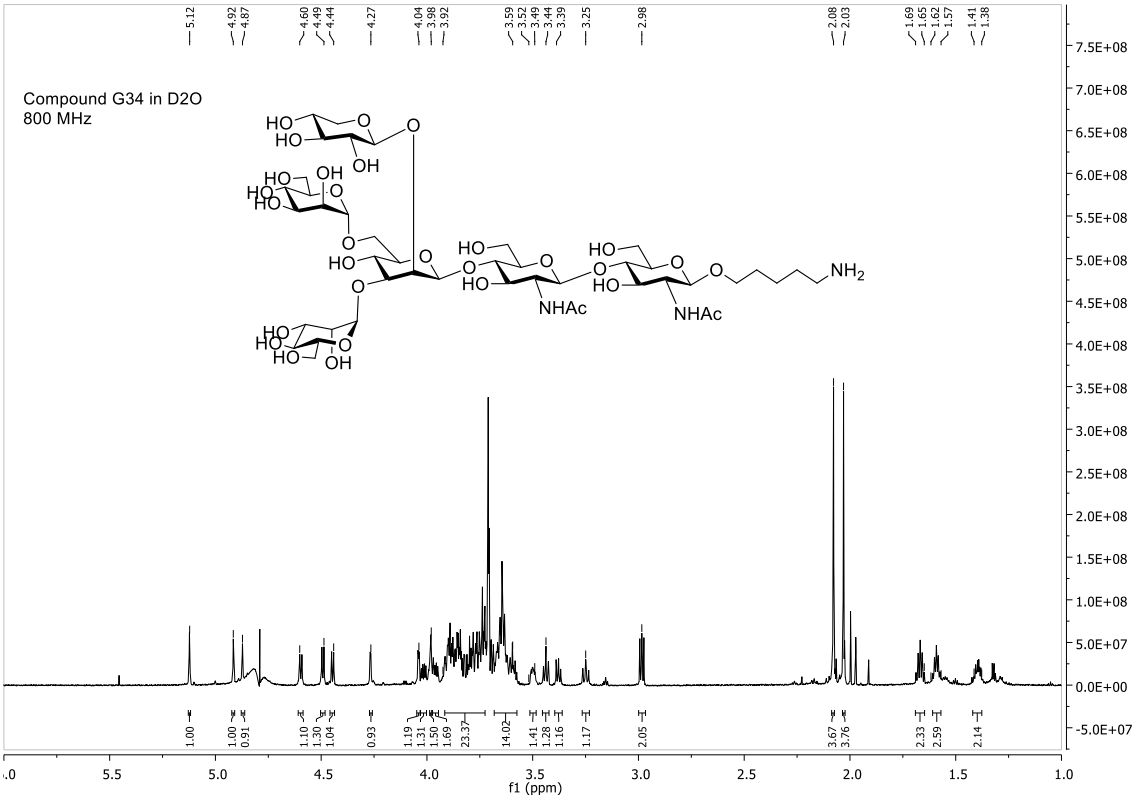

Compound G31 in D2O
800MHz
f1 (ppm)

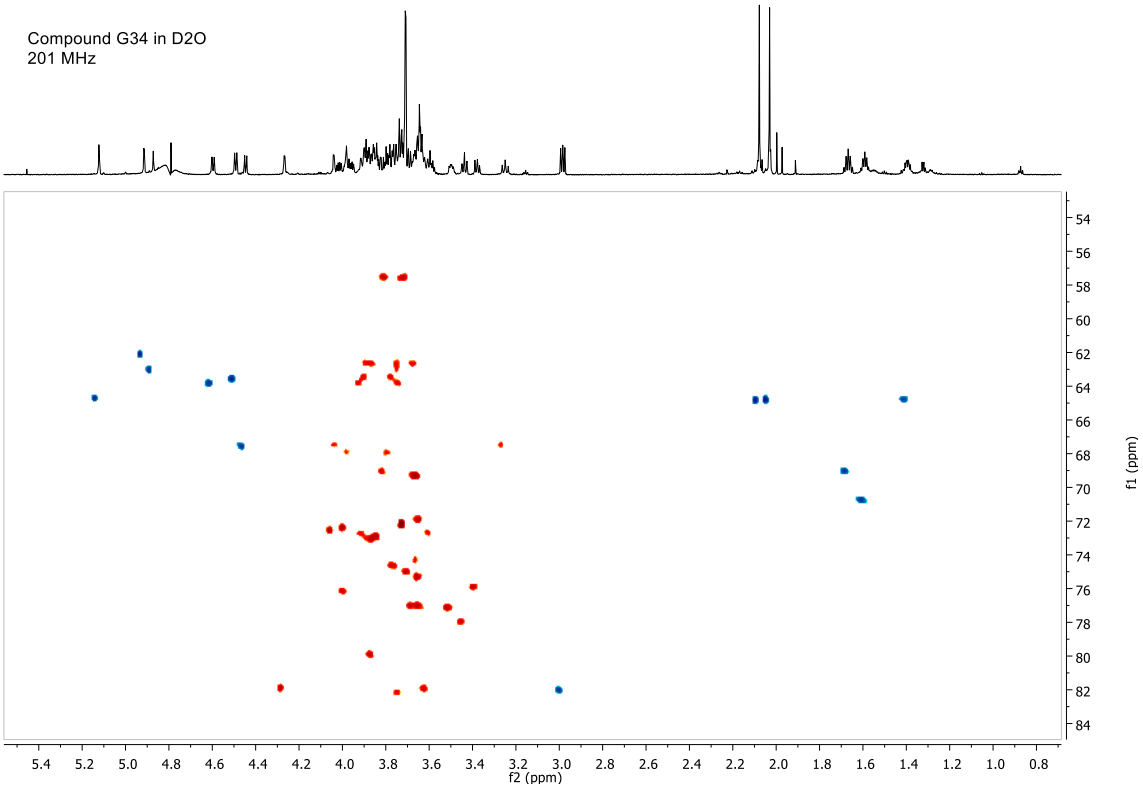

Compound G31 in D2O
201 MHz

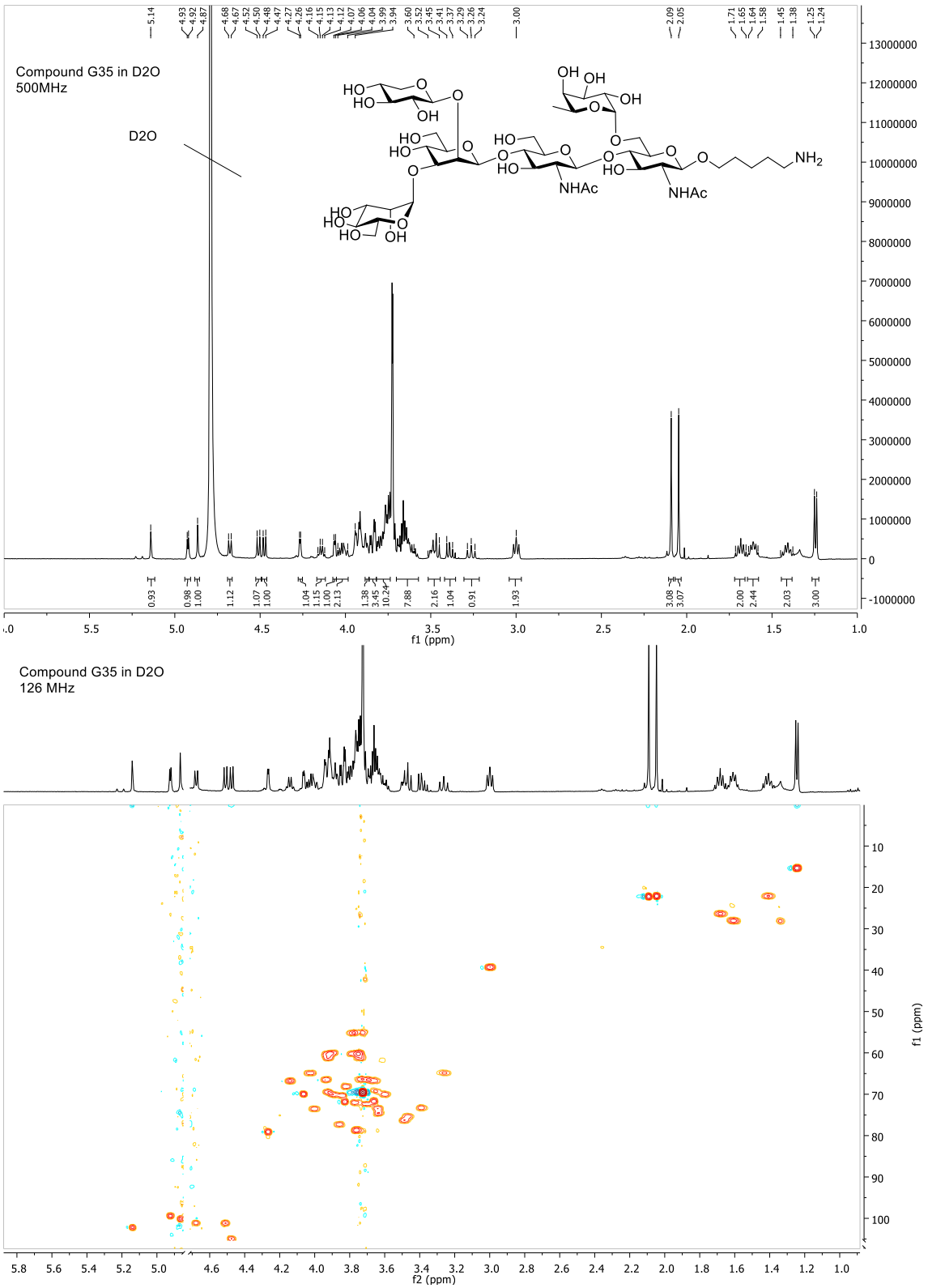

f2 (ppm)
f1 (ppm)

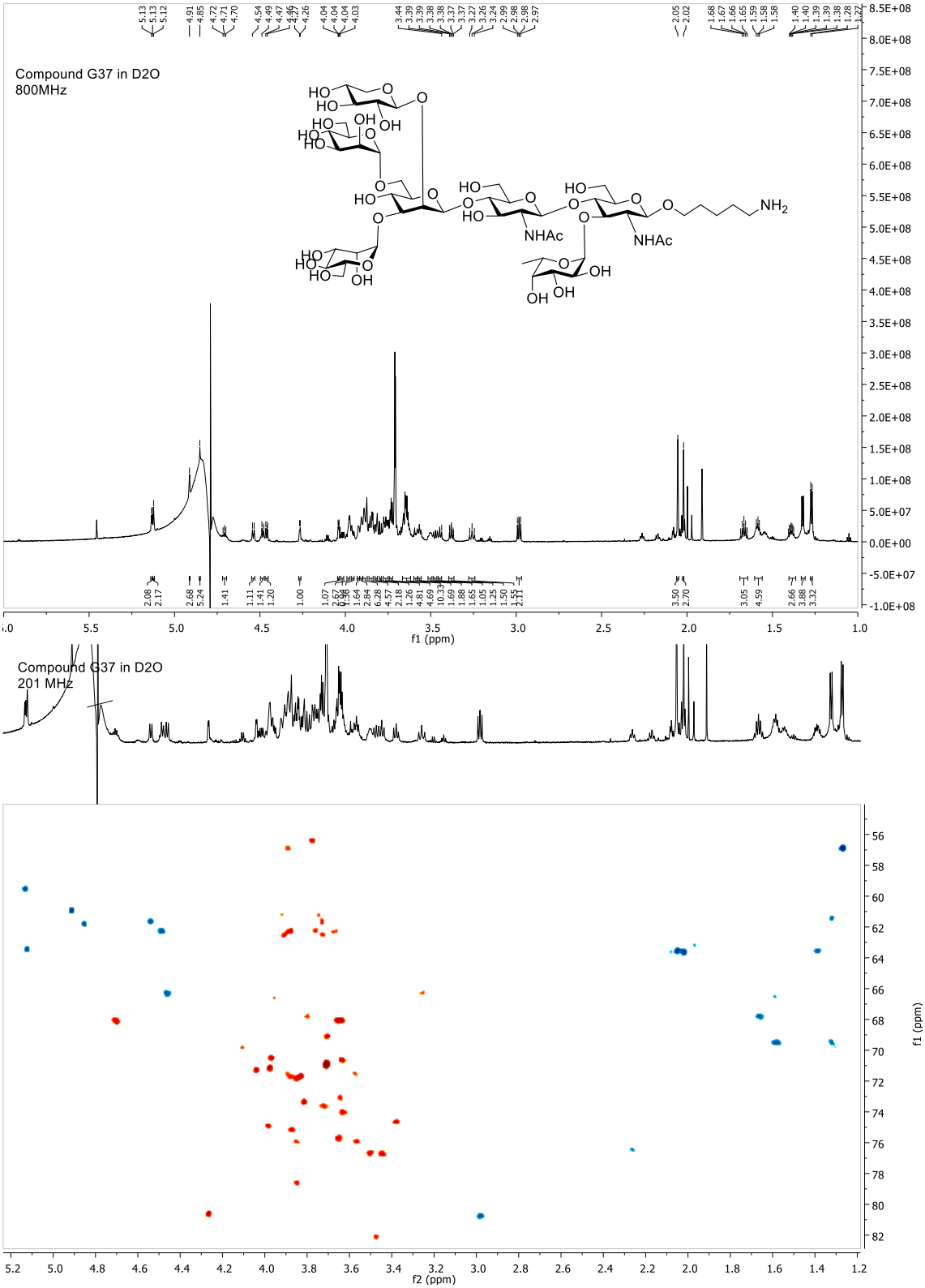

Compound G32 in D2O
500MHz
D2O
f1 (ppm)

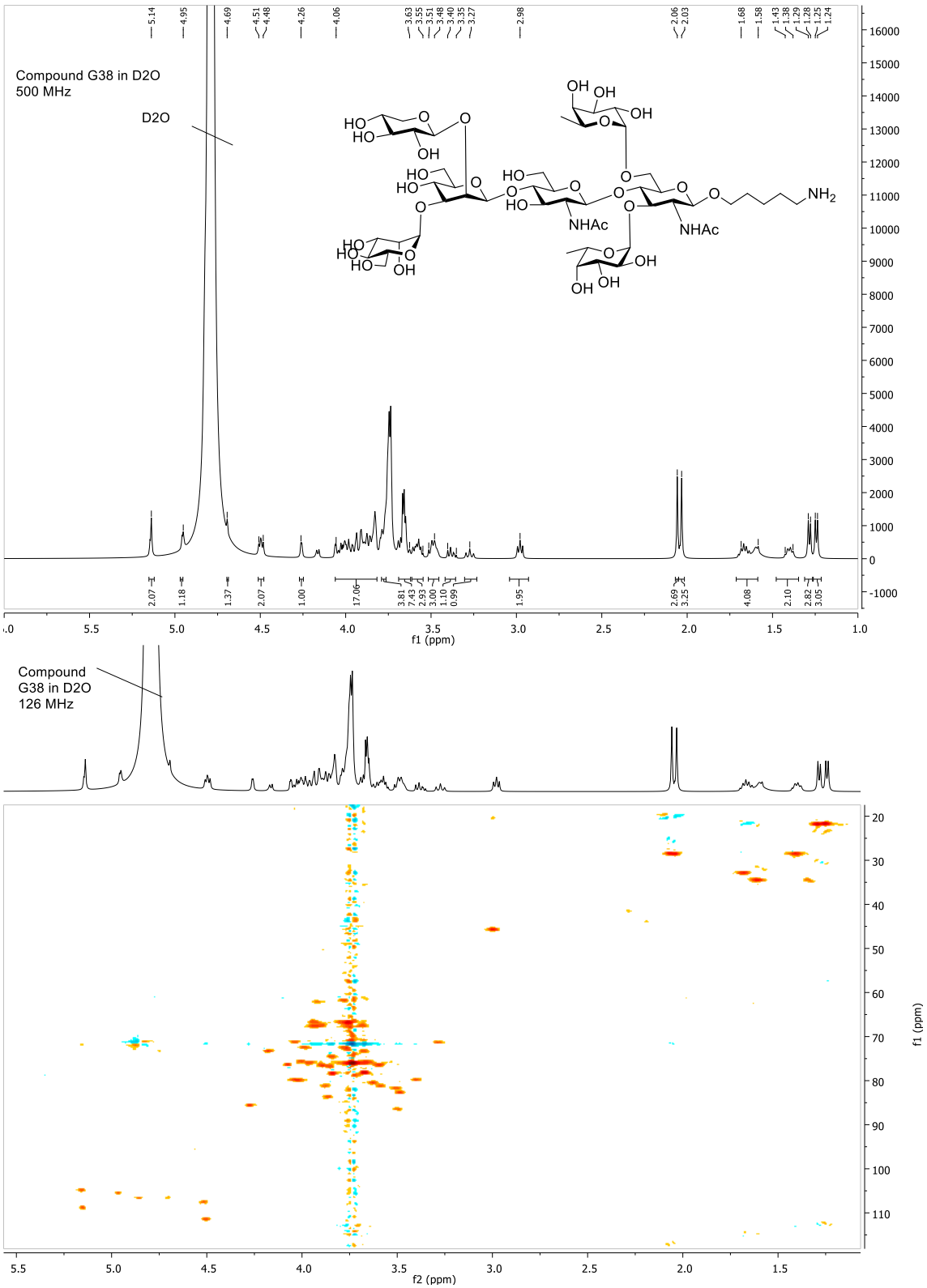

Compound G33 in D2O
500MHz
D2O
f1 (ppm)

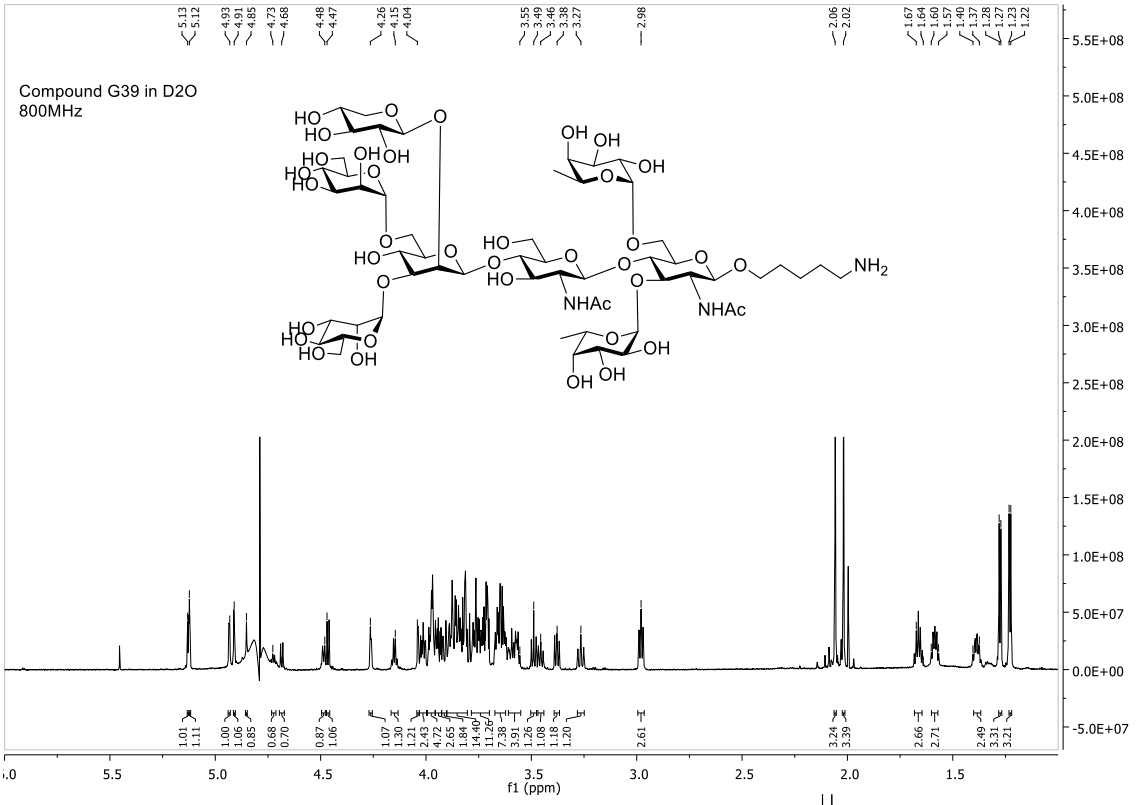

Compound G34 in D2O
800 MHz
NHAc
NHAc
NH2
f1 (ppm)

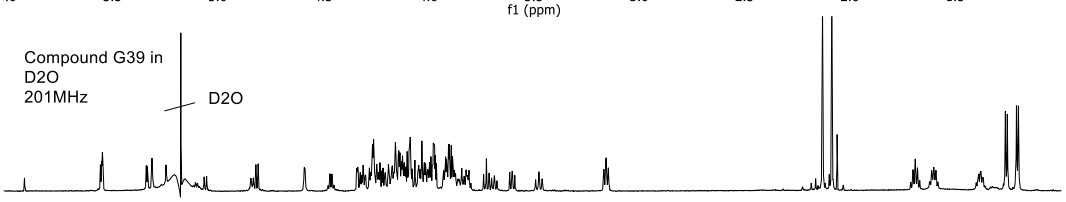

Compound G34 in D2O
201 MHz
f2 (ppm)
f1 (ppm)

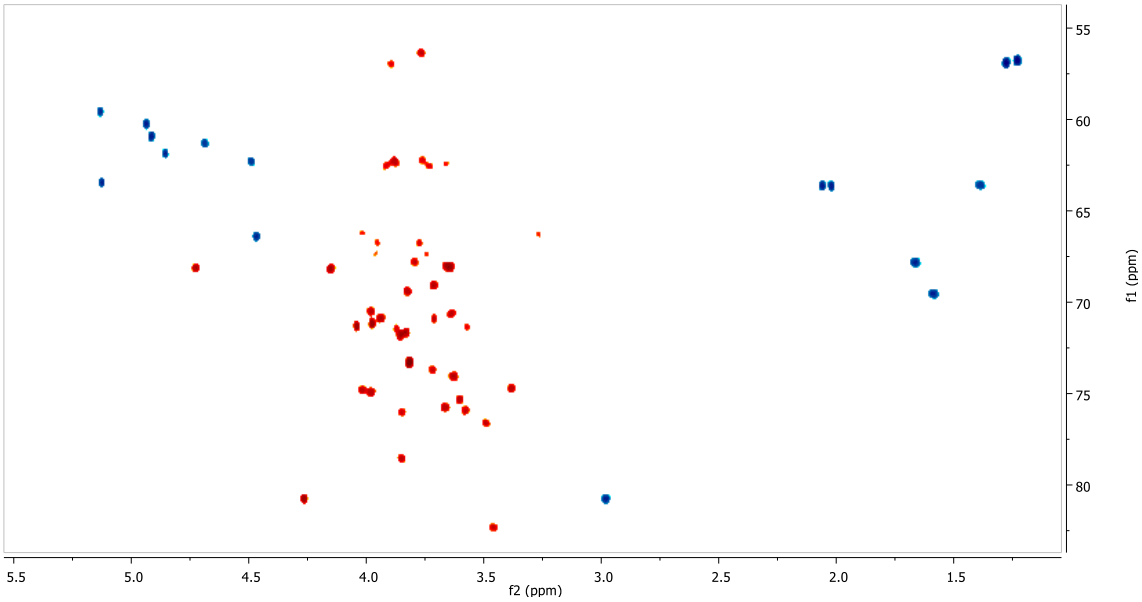

Compound G35 in D2O
500MHz
D2O
NHAc
NHAc
NH2
f1 (ppm)
Compound G35 in D2O
126 MHz
f1 (ppm)
f2 (ppm)


Compound G37 in D2O
800MHz
f1 (ppm)
Compound G37 in D2O
201 MHz
f2 (ppm)
f1 (ppm)


Compound G38 in D2O
500 MHz
D2O
Compound
G38 in D2O
126 MHz
f1 (ppm)
f2 (ppm)


Compound G39 in D2O
800MHz
f1 (ppm)


Compound G39 in D2O
201MHz
D2O


f2 (ppm)
f1 (ppm)