



UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI CHIMICA

CORSO DI DOTTORATO IN CHIMICA

XXIX CICLO

**Glycomimetic Antagonists of C-Type Lectins:
Improving Ligand Potency and Multivalent Presentation**

Settore Scientifico Disciplinare CHIM/06

Tutor: Prof. Anna BERNARDI

Dottorando: Giulio GOTI

Coordinatore: Prof. Emanuela LICANDRO

R10403

A.A. 2015/2016

INTRODUCTION	1
0.1 C-TYPE LECTINS: SOME GENERALITIES OF A COMPLEX WORLD	1
0.2 C-TYPE LECTINS IN IMMUNITY	3
0.3 TARGETING C-TYPE LECTINS	6
0.3.1 Monovalent ligands	6
0.3.2 Multivalent ligands.....	9
0.3.2.1 Multivalency: a theoretical point of view	9
0.3.2.2 Multivalent glycoconjugates.....	13
0.4 AIM OF THE WORK.....	20
0.5 REFERENCES.....	21
CHAPTER 1 - MBL LIGANDS AS NEUROPROTECTIVE AGENTS IN STROKE	25
1.1 MANNOSE BINDING LECTIN AND ITS ROLE IN THE COMPLEMENT SYSTEM	25
1.2 MBL IN HEALTH AND DISEASE.....	29
1.3 MBL IN ISCHEMIC INJURY - BACKGROUND.....	29
1.4 MANNO-GLYCODENDRIMERS AS MBL ANTAGONISTS IN REPERFUSION INJURY	31
1.5 CONCLUSIONS	52
1.6 EXPERIMENTAL	53
1.6.1 General methods and procedures.....	53
1.6.2 Synthesis of MBL multivalent antagonists.....	54
1.6.2.1 Products numbering for spectral assignment	54
1.6.2.2 General procedure for the CuAAC reaction.....	55
1.6.2.3 Preparation and characterization of hexavalent Polyman42 51	55
1.6.2.4 Preparation and characterization of nonavalent Polyman20 52	60
1.6.2.5 Synthesis of stabilized tetravalent glycodendrons 76a and 77a	62
1.6.2.6 Synthesis of 16-valent glycodendrimer 81	77
1.6.2.7 Isotopic distribution of dendrons	86
1.6.2.8 SPR inhibition assay experiments.....	95
1.7 REFERENCES.....	98
CHAPTER 2 - RATIONAL DESIGN OF DC-SIGN MULTIVALENT ANTAGONISTS	101
2.1 DC-SIGN IN HIV INFECTION: A PARADIGM IN ESCAPING IMMUNITY.....	101
2.2 STRUCTURE-BASED SYNTHESIS OF DC-SIGN LIGANDS.....	103
2.3 SYNTHESIS OF TETRACOORDINATING LIGANDS: GRABBING DC-SIGN.....	106
2.4 CONCLUSIONS	113
2.5 EXPERIMENTAL	114
2.5.1 General methods and procedures.....	114
2.5.2 Synthesis of a “cross-shaped” DC-SIGN antagonist	115
2.6 REFERENCES.....	126
CHAPTER 3 - GLYCOMIMETIC LEAD STRUCTURES IDENTIFICATION: EXPLOITING MALDI-TOF DETECTABLE GLYCAN MICROARRAYS.....	128
3.1 GLYCAN MICROARRAYS: GENERAL FEATURES AND APPLICATIONS.....	128
3.2 IMMOBILIZATION CHEMISTRY	129
3.2.1 Covalent immobilization	129
3.2.2 Non-covalent immobilization.....	130
3.3 DETECTION TECHNIQUES AND SURFACE CHARACTERIZATION	131
3.4 ON-CHIP GLYCOMIMETIC LIBRARY EXPANSION	134
3.4.1 On-array chemoenzymatic reactions: optimization studies	135
3.4.1.1 Glycan printing	136
3.4.1.2 Enzyme activity assay: clickable NDP sugars as glycoside donors	139

Table of contents

3.4.1.3	On-chip glycosylation.....	141
3.4.1.4	Library diversification: CuAAC of arrayed glycans	146
3.4.1.5	Array stability: fluorescent analysis.....	151
3.4.2	Synthesis of Lewis ^X mimics array: looking for lead structures	152
3.4.2.1	Synthesis of tagged GDP-fucose glycoside donors.....	153
3.5	CONCLUSIONS	161
3.6	EXPERIMENTAL	162
3.6.1	General methods and procedures.....	162
3.6.2	Array fabrication and characterization	163
3.6.2.1	Synthesis of self-assembled hydrophobic monolayer on ITO slides.....	163
3.6.2.2	Reactive double lipidic layer formation.....	163
3.6.2.3	General procedure for glycan printing and detection	163
3.6.2.4	GalT-1 and GalT-1-DM expression and purification.....	164
3.6.2.5	General procedure for GalT-1 and GalT-1-DM enzyme activity assay	165
3.6.2.6	Spectroscopic data of 6-azido-UDP-D-galactose 114	166
3.6.2.7	Enzymatic glycosylation with UDP-GalNAz. Synthesis of azido tagged 123	166
3.6.2.8	General procedure for on-chip glycosylation.....	166
3.6.2.9	General procedure for on-chip CuAAC	166
3.6.2.10	Synthesis of methyl hexynoate 104c	167
3.6.2.11	Fluorescein printing and fluorescence detection	167
3.6.2.12	General procedure for enzymatic fucosylation of biantennary 152	167
3.6.3	Synthesis of tagged GDP-L-fucose glycoside donors.....	167
3.6.3.1	Preparation of 6-azido-GDP- β -L-fucose 127	169
3.6.3.2	Preparation of 5-ethynyl-GDP-L-fucose 128	175
3.7	REFERENCES.....	180

INTRODUCTION

0.1 C-Type lectins: some generalities of a complex world

0.1 Glycans represent an essential component of living organisms, where they are present as polysaccharide chains or as conjugates with proteins and lipids.^{1,2} Notably, all cells express glycans on their surface forming a nanolayer called glycocalix,^{3,4} and the great majority of proteins are subjected to glycosylation, which is the most common protein post-translational modification. These evidences and the ubiquitous diffusion of oligosaccharides highlight their importance in cellular processes. However the relevance of carbohydrates has been neglected for a long time and is during the past fifty years that glycans started to attract attention.⁵ This underestimation of carbohydrate functions in the context of cellular biology may reside in the difficulties encountered during their investigation. Particularly, glycan and glycoconjugate biosynthesis and its regulation are far more complicated to elucidate compared to polynucleic acids or proteins, where a clear correlation between a gene and the corresponding polypeptide sequence can be identified. Moreover, a great effort from an analytical point of view is required to determine the structural diversity of glycans, which can be constituted just by one monosaccharide unit up to a polysaccharide sequence of different monomers, connected in a linear or branched array and potentially leading to different regio- and stereoisomers.^{6,7}

Nowadays it is well established that carbohydrates are extensively engaged by cells in different biological processes, being able to control embryogenesis and organism development,⁸ determine cellular adhesion and migration,⁹ promote tumour progression and modulate inflammation and immune response.¹⁰ Many of these events are mediated by carbohydrate-protein interactions, where glycans are recognized and selectively bound by proteins known as lectins. The great majority of lectins can be classified into three different superfamilies of proteins depending on the structural domain composition and the nature of the minimal sacchharide epitope recognized: galectins mainly recognize the disaccharide lactose or *N*-acetyllactosamine, while Siglecs bind preferentially to sialic acid residues and C-type lectins have certain selectivity for mannose and galactose related structures.¹¹ Among these, proteins belonging to the C-type lectin superfamily are the most abundant and are broadly expressed by animals, viruses and parasites. Apart from some exceptions, C-type lectins are either extracellular or transmembrane proteins, which can be classified into seventeen different groups based on the overall domain composition (Fig. 1).¹²

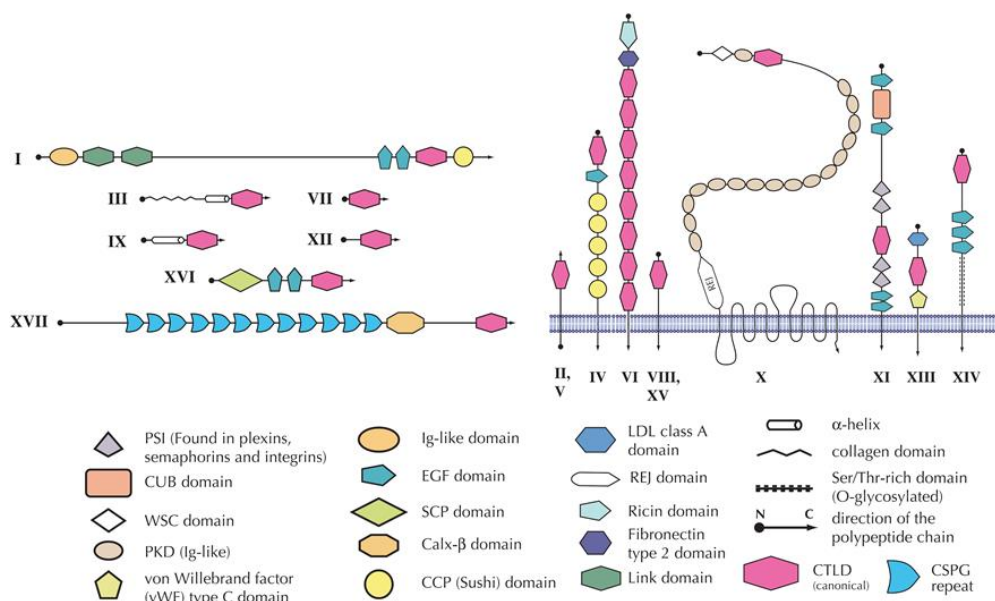


Figure 1 - Domain architecture of vertebrate C-type lectins. Group numbers **I-XVII** are indicated. Soluble C-type lectins are shown on the left, while transmembrane C-type lectins on the right. The legend for schematic representation

of structural domains is reported. **I** lecticans, **II** the ASGR group, **III** collectins, **IV** selectins, **V** NK receptors, **VI** the macrophage mannose receptor group, **VII** REG proteins, **VIII** the chondrolectin group, **IX** the tetranectin group, **X** polycystin 1, **XI** attractin, **XII** EMBP, **XIII** DGCR2, **XIV** the thrombomodulin group, **XV** Bimlec, **XVI** SEEC, **XVII** CBCP.¹²

Despite this great variety of structures, all C-type lectins contain at least one C-type lectin-like domain (CTLD).^{12,13} This globular domain is characterised by an overall disordered loop structure where few conserved secondary structures can be identified. Both the N and C terminus lie below the domain where the two β strands $\beta 1$ and $\beta 5$ forms an antiparallel β -sheet. A second β -sheet is formed by the β strands $\beta 2$, $\beta 3$, $\beta 4$, while two α helices, $\alpha 1$ and $\alpha 2$, flank the domain. Four highly conserved cysteine residues, C1-4, play an important structural role forming two disulphide bonds in close proximity of the two loop regions: C1 and C4 connect $\beta 5$ with $\alpha 1$ while C2 and C3 join $\beta 3$ with $\beta 5$. Above the whole domain is a long loop region with a variable primary structure where the carbohydrate binding site is located, as revealed by the presence of a calcium ion (Fig. 2a).

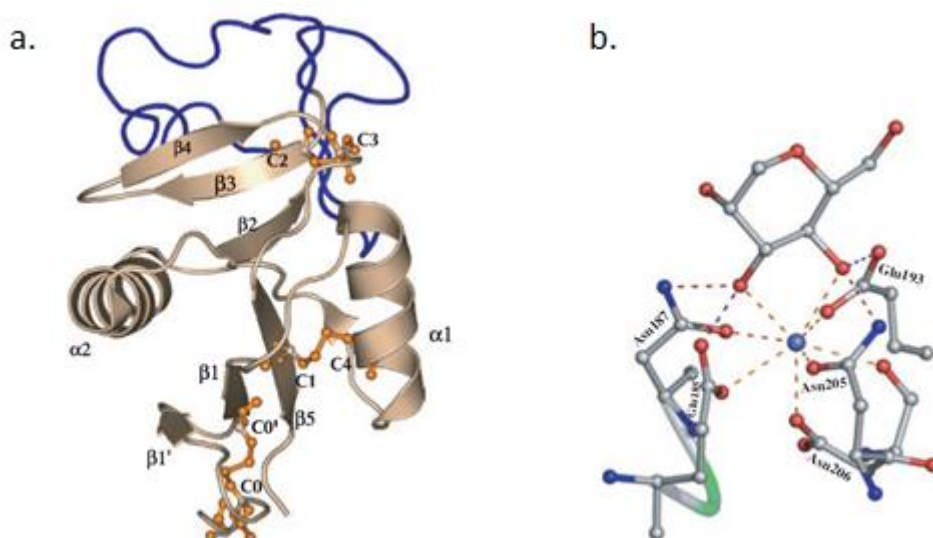


Figure 2 - C-type lectin-like domain (CTLD) structure.¹² **a)** The globular structure of CTLD is shown. The disordered loop structure where the carbohydrate binding site is located is highlighted in blue. The conserved secondary structures (α helices and β strands) and the cysteine residues (C) are indicated. **b)** A mannose residue binding to the Ca^{2+} ion at the carbohydrate binding site is shown. The sugar moiety coordinates with the metal centre (blue sphere) through equatorial vicinal hydroxyls at 3 and 4 positions. Further stabilizing interactions between the protein and both the Ca^{2+} ion and the carbohydrate moiety are shown: five ligands lie on a plane, while three others lie in a trigonal arrangement on an orthogonal plane resulting in an eightfold coordination of the Ca^{2+} ion.

The Ca^{2+} ion, from which C-type lectins take their name, is essential for the binding ability of the proteins for carbohydrates, which directly interact binding to the metal centre through two vicinal hydroxyl groups. In particular, C-type lectins are able to bind sugar moieties exposing either equatorial vicinal hydroxyls at 3 and 4 position, thus showing selectivity for mannose, glucose and *N*-acetylglucosamine (but also for fucose, binding with hydroxyls 2 and 3), or residues with a combination of equatorial/axial vicinal hydroxyls in 3 and 4 position, therefore being selective for galactose and fucose. The Ca^{2+} ion is stabilised at the binding site through interactions that it makes with carboxylic and amide groups either contained in the protein backbone or in the side chains. In particular, interactions are established between the metal and Asn, Asp, Glu and Gln residues belonging to a conserved WND sequence and either an EPN or QPD motif. The aminoacids involved in Ca^{2+} coordination also stabilize through hydrogen bonds the carbohydrate moiety bound to the metal. This creates a complex network of interactions that is always observed for CTLDs and results in a eightfold coordination to the Ca^{2+} , with five ligands lying on a plane and with the remaining three, two of which coming

from the sugar moiety, lying on an orthogonal plane in a trigonal arrangement (Fig.2b). As mentioned before, the highly conserved WND and EPN or QPD triplets have the major role of coordinating to calcium. The proline residue in EPN and QPD is in a *cis* conformation, thus allowing for optimal orientation of the flanking residues. Side chains of the triads also make contacts with bound carbohydrates, greatly influencing lectin selectivity. More in detail, EPN sequence has shown to give selectivity towards mannose, while QPD is usually associated with galactose binding selectivity. This correlation was first demonstrated by Drickamer and co-workers by an elegant mutagenesis experiment in which the replacement of the EPN triplet with the QPD in the murine MBL-A determines a switch in preferential binding from sugar with a mannose like configuration to a galactose like.¹⁴ Crystallographic analysis of the mutated MBL-A, revealed that the EPN replacement with the QPD sequence did not alter the geometry of calcium ion coordination bonds. This suggests that the switch in specificity may derive from a different network of hydrogen bonds between the sugar moiety and the aminoacidic residues coordinated to the calcium ion, thus highlighting their importance in protein specificity and in controlling the orientation and conformation of the ligand at the binding site. However, CTLDs with preferential binding for mannose related derivatives have often proved to accept also galactose related structures and vice versa and furthermore different CTLDs with the same carbohydrate selectivity have often shown different binding modes. Given all these evidences, it seems reasonable to believe that more complex and still not well defined mechanisms allowing lectins to specifically recognize their glycan target must be operative.

0.2 C-Type lectins in immunity

In animals, C-type lectins are mainly expressed at the surface of myeloid cells and engaged by the immune system, thus being involved in infectious diseases, autoimmunity and allergies. C-type lectin receptors in fact belong to the class of the so called pattern recognition receptors (PRRs), which are greatly expressed on the surface of dendritic cells (DCs) where they act as important modulators of the immune system. DCs differentiate from bone marrow stem cells, then after migration to the bloodstream they populate all tissues where they exert the role of sentinels able to recognize, process and present antigens on major histocompatibility complex (MHC) class I and II molecules to T cells, promoting their differentiation.¹⁵ Immature DCs lie in a steady-state where they recognize and scavenge pathogens and innocent self occurring antigens, deriving from necrotic and apoptotic cells. Upon pathogen infection or massive cell death, DCs are activated and after maturation they migrate into secondary lymphoid organs where they can present processed antigens as MHC class I and II to naïve CD8⁺ and CD4⁺ T cells respectively.

Depending on the infecting pathogen, DCs shape the immune response inducing alternative subsets of helper T cells (T_H). Therefore, T_H1 cells are able to activate macrophages against microbial infections producing interferon- γ , while T_H2 cells can secrete interleukin-4 (IL-4), IL-5 and IL-13 to recruit white cells against helminth and T_H17 cells express IL-17 to mobilize phagocytes and resist bacteria and fungi infections.¹⁶

DCs antigen recognition is possible through the engagement of different PRRs that can bind either non-self ligands, pathogen-associated molecular patterns (PAMPs), and self-damaged structures, damage-associated molecular patterns (DAMPs), as well as self-altered ones, tumour-associated molecular patterns (TAMPs). Among PRRs expressed by DCs there are Toll-like receptors (TLRs), nucleotide-binding oligomerization domain-like receptors (NLRs), retinoic acid-inducible gene I (RIG-I)-like receptors and C-type lectin receptors.¹⁷ C-type lectins receptors expressed by DCs recognize and bind to characteristic moieties of glycoconjugates exposed on pathogen membranes and allowing specific targeting.¹⁸

Generally speaking, C-type lectin receptors like DC-specific ICAM3-grabbing non-integrin (DC-SIGN), L-SIGN, mannose receptor, macrophage galactose-specific lectin (MGL) and langerin in Langerhans cells are highly selective for high-mannose or fucose-containing glycans (Lewis^{a,b,x,y}), GalNAc or GlcNAc. A different selectivity is shown by DC-associated C-type lectin 1 (dectin-1), a Ca²⁺-independent C-type lectin, that is

Introduction

usually engaged in fungi infection being able to bind β -glucan containing glycans.¹⁹ After recognition and binding, C-type lectin receptors trigger internalization allowing pathogen degradation and antigen processing and presentation as MHC. Furthermore, C-type lectin receptors are able to fine tune immune response through signalling pathways leading to activation and modulation of gene expression. Many details concerning C-type lectin induced signal transduction still have to be clarified. What we know, is that upon pathogen recognition, C-type lectin receptors can either directly induce gene expression, this is the case of dectin-1,²⁰ dectin-2²¹ and macrophage-inducible C-type lectin (mincle),²² while others such as DC-SIGN,²³ blood DC antigen 2 protein (BDCA2),²⁴ DC immunoreceptor (DCIR)²⁵ and myeloid C-type lectin-like receptor (MICAL) can synergistically tailor immune response by modulating gene expression signalling triggered by TLRs. Depending on the pathogenesis of the infection, different pathways can be initiated. Nevertheless, all C-type lectin receptors initiate their signalling in two alternative ways. Lectin such as mincle, dectin-2, BDCA2 and C-type lectin domain family 5, member A (CLEC5A) initiate signalling through immunoreceptor tyrosine-based activation motif (ITAM)-containing adaptor molecules.^{21,24,26,27} On the contrary, dectin-1, DC-SIGN, DCIR, MICAL induce signalling by interactions between their cytosolic domain and either kinase or phosphatase proteins.^{20,23,28,29} Generally, C-type lectin receptors induced signalling pathways produces as its final event the activation and modulation of the transcriptional factor nuclear factor- κ B (NF- κ B), the main mediator in immune system of inducible gene expression.

An explicative case of how the PRRs all together orchestrate their functions leading to a highly specific immune response is given by the DC-SIGN crosstalk with Toll-like receptors and their signalling pathway modulation (Fig. 3). More in detail, during infections of *Mycobacterium tuberculosis*, *Mycobacterium leprae*, HIV-1 and *Candida albicans*, DC-SIGN recognition of these pathogens shape TLR4 mediated immune responses.^{23,30} After infection, TLR4 binds lipopolysaccharides (LPS), initiating a signalling pathway that finally activate NF- κ B for gene expression. This response is further tuned by DC-SIGN. In fact, DC-SIGN binding to polymannosylated surfaces exposed by the infecting pathogens triggers a cascade eventually activating RAF1 protein. RAF1 is a serine/threonine kinase that phosphorylates NF- κ B p65 subunit at Ser276, thus allowing recruitment of histone acetyl-transferases CREB-binding protein (CBP) and p300 that results in acetylation of various p65 lysine residues.^{23,31} Acetylation of NF- κ B p65 subunit prolongs its binding at the promoter and increases transcriptional rate of genes encoding cytokines such as IL-8 and IL-10.^{23,32}

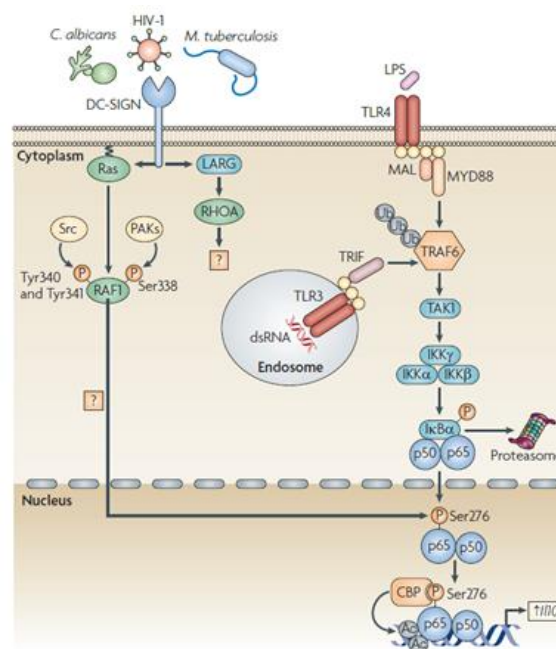


Figure 3 – PRRs crosstalk: DC-SIGN modulation of Toll-like receptors signalling.³³ DC-SIGN is able to modulate the response triggered by Toll-like receptors TLR3 and TLR4 during infections of *Mycobacterium tuberculosis*,

Mycobacterium leprae, HIV-1 and *Candida albicans*. After infection, TLR3 is activated by binding double-stranded RNA (dsRNA), while TLR4 binds lipopolysaccharides (LPS). These recognition initiate a signalling pathway involving the recruitment of the ubiquitin ligase TNF receptor-associated factor 6 (TRAF6) followed by self-induced polyubiquitylation (Ub) which triggers the recruitment of TGF β -activating kinase 1 (TAK1). TAK1 activates the I κ B kinase (IKK) complex, which is then able to phosphorylate inhibitor of NF- κ B α (I κ B α), promoting its proteasomal degradation and the release of NF- κ B. Finally, NF- κ B translocates into the nucleus, where it binds to promoter sequences triggering gene expression. This cascade of events is modulated by DC-SIGN. Indeed, DC-SIGN recognition of polyglycosylated surfaces exposed by pathogens triggers a cascade eventually activating the serine/threonine kinase RAF1 protein. RAF1 is then able to phosphorylate NF- κ B p65 subunit at Ser276, thus allowing recruitment of histone acetyl-transferases CREB-binding protein (CBP) and p300 which determines the acetylation of various lysine residues of NF- κ B p65 subunit. Acetylation prolongs NF- κ B binding at the promoter, thus increasing transcriptional rate of genes encoding cytokines such as IL-8 and IL-10.

Up to now, the concept that C-type lectins play an essential role in the first-line host defence has been stressed. However, some viruses and bacteria have evolved mechanisms to escape from immune activation, exploiting C-type lectin receptors for their entry and to modulate PRRs signalling, thus promoting survival and infection. A very well studied example of this mechanism is DC-SIGN mediated infection, since DC-SIGN ability to recognize many dangerous viruses such as HIV,^{34,35,36} Ebola virus,^{37,38} Cytomegalovirus,³⁹ Hepatitis C virus,^{40,41,42} Dengue virus^{43,44} but also bacteria like *Helicobacter pylori*, *Mycobacterium tuberculosis* as well as parasites like *Leishmania amastigotes*, *Schistosoma mansoni* and *Candida albicans*.^{45,46,47,48} In HIV infection, after binding of DC-SIGN and internalization, virus particles are able to circumvent lysosomal degradation through still unknown mechanism. Moreover, the infected DC-SIGN expressing dendritic cell is further exploited by HIV virus to interact with T cells, enabling dissemination and spreading of the infection. *Mycobacterium tuberculosis* and *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) also target DC-SIGN through the membrane mannose containing glycolipid lipoarabinomannan (Man-LAM). It was demonstrated that this interaction leads to the expression of the anti-inflammatory IL-10, while in *M. bovis* BCG it results in an inhibition of the TLR4 mediated DC differentiation signal therefore showing how these bacteria succeed in downregulating immune response.³⁰

Being able to trigger and modulate the immunological system, C-type lectins can be responsible for the pathogenesis of inflammatory diseases and dysregulation of their activity may lead to autoimmune responses, as well as allergic disorders.¹⁰ Notably, mannose binding lectin (MBL) was shown to sense damaged cells after ischemic stroke and brain injury, playing a role in the establishment of an inflammatory environment that further exacerbates the lesion.^{49,50} Blocking MBL with antibodies has proved to provide a protective effect in ischemia mouse model.⁵¹ With the same approach, a mannosylated proteolipid peptide was used to inhibit experimental autoimmune encephalomyelitis in mice, probably antagonizing mannose receptors expressed on DCs.^{52,53} C-type lectins can also effect allergic inflammation, as demonstrated for dectin 2, which recognizes house dust mite stimulating mobilization of innate immune cells.⁵⁴

Quite recently, it has been demonstrated that C-type lectins are also able to interact with characteristic glycans of malignant carcinogenic cells. Indeed, modifications of membrane polysaccharides and glycoconjugates enable tumour progression.⁵⁵ Several glycoforms of carcinoembryonic antigen and mucin 1 (MUC-1) have been found to target DC-SIGN and MGL expressed by DCs.^{56,57} In particular, the interaction of MUC-1 with MGL allows cancer cells to induce a T_H2-mediated response, promoting their survival.

For their numerous biological implications, C-type lectins clearly emerge as promising targets for development of new drug therapies. Indeed, C-type lectins targeting has been revealed as a potent approach for inhibition of first stage pathogen adhesion, treatment of inflammatory diseases and presentation of vaccine vehicle, finally stimulating tumour suppression.^{58,59}

0.3 Targeting C-type lectins

The development of carbohydrate based C-type lectin targeting drugs is a hard challenge. The polar nature of carbohydrates accounts for their poor pharmacokinetic properties. Their highly polarity impairs oral availability, since they are not passively absorbed in the small intestine. Additionally, oligosaccharides present a short half-life in plasma, being subjected to the hydrolytic activity of glycosidases and rapidly excreted by kidneys. Moreover, they are usually produced through multi-step synthetic processes with low atom economy. The rational design of C-type lectin ligands is even more demanding if we consider that most of the carbohydrate binding sites are located at superficial and accessible regions of the proteins, which also leads to low affinity interactions. Few small molecule ligands have been developed so far and this may be also due to the discouraging low success of classical drug discovery campaigns in the quest for new hit structures. However, recently Rademacher and coworkers⁶⁰ have shown that *in silico* low scores for lectin druggability prediction does not correlate with experimental ¹⁹F NMR based fragment screening against DC-SIGN, disclosing the limitations of the software analysis developed so far, which are not parametrized for carbohydrate and Ca²⁺ binding sites.⁶¹ The ideal C-type lectin antagonist is characterized by a proper pharmacokinetic and a straightforward synthesis, it has to show increased affinity than the natural competing ligand and it has to be selective. Promising results for such a goal have been achieved through rational design of monovalent glycomimetics and the synthesis of multivalent glycoconjugates.

0.3.1 Monovalent ligands

Several lectins have been successfully targeted by ligands, most of them with a carbohydrate-based structure. Starting from natural ligands, glycomimetics are carefully designed in order to obtain drug-like entities. This implies a simplification of the structure, a reduced polarity and increased stability for better bioavailability, the introduction of functional groups giving further interactions with the protein, therefore enhancing both affinity and selectivity. Remarkably, compounds with great avidity have been achieved for Siglecs and galectins, as well as molecules directed against bacterial lectins, which may function as toxins or as adhesion proteins enabling infection.⁶² Because of its engagement in first recognition of several pathogens, many of the developed ligands for C-type lectins have been addressed to DC-SIGN. Herein, an overview of DC-SIGN monovalent ligands is presented along with their corresponding IC₅₀ values. It should be noted that IC₅₀ values allow to quantify ligand potency, being however affected by the nature and the setting of the experimental technique used for their determination. For this reason, a comparison with a reference compound should always be assessed. Unfortunately, the lack of a standard reference in the literature devoted to DC-SIGN antagonists makes their comparison not easy.

Noncarbohydrate antagonists **1**⁶³ and **2**⁶⁴ with IC₅₀ values ranging from micromolar to submicromolar were identified with high-throughput screening and particularly compound **2** was found to block DC-SIGN mediated internalization of fluorescent-labelled mannosylated BSA in DC-SIGN expressing Raji cells (Fig. 4).

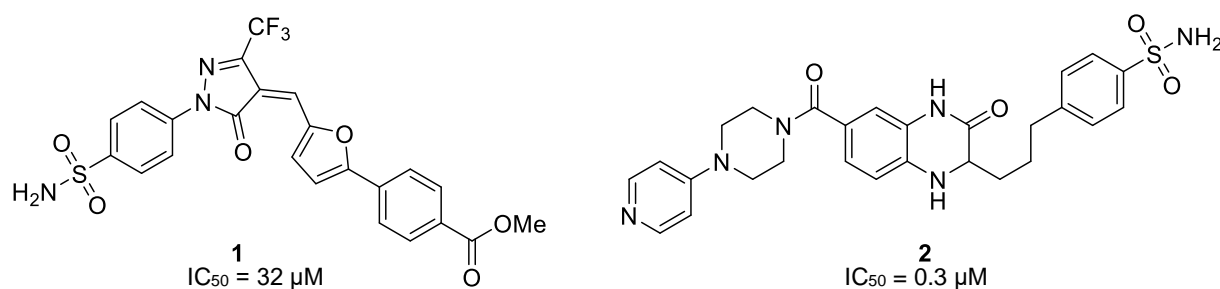


Figure 4 – Non-carbohydrate based DC-SIGN antagonists. Compound **1** showed IC₅₀ = 32 μM in inhibition assay using Man-BSA as competitor.⁶³ Compound **2** displayed IC₅₀ = 0.3 μM in a competition experiment with fluorescent mannosylated BSA.⁶⁴

DC-SIGN specifically recognizes L-fucose and D-mannose residues, therefore glycomimetic antagonists of DC-SIGN are mainly based either on the Lewis^x trisaccharide Gal β (1,4)[Fuc α (1,3)]GlcNAc or on the highly mannosylated structure of the oligosaccharide (Man)₉(GlcNAc)₂ (abbreviated as Man₉), which is often exposed in multiple presentation by several pathogen proteins. Glycomimetic **3** represents a simplified version of Lewis^x ligand, where a cyclohexane ring mimics the original GlcNAc residue.⁶⁵ This strategy is further stressed in the synthesis of **4**. NMR studies of this compound revealed that binding to DC-SIGN is ensured mainly by the fucose, and that the phenol moiety establishes further contacts with the protein (Fig. 5).⁶⁶

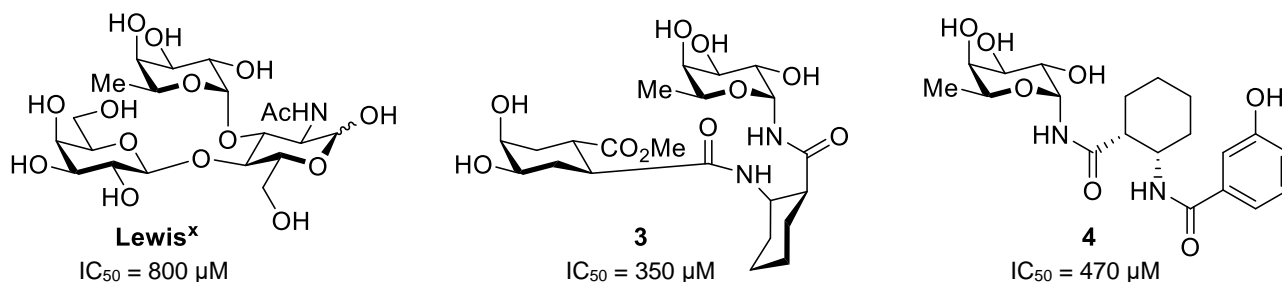


Figure 5 – L-fucose based DC-SIGN ligands. The fucose containing **3** and **4** are glycomimetic structures of Lewis^x trisaccharides. All IC₅₀ values were obtained by SPR inhibition assay in which ligands had to compete with BSA-mannotriose for DC-SIGN binding.^{65,66}

In the past few years, our group has reported the synthesis of the pseudo-dimannosides **5a,b**⁶⁷ and the pseudo-trimannosides **6a,b**,⁶⁸ whose structures are based on terminal epitopes of the natural ligand Man₉ (Fig. 6). All these ligands are characterized by the presence of a conformationally locked cyclohexane diol ring that mimics a mannose residue while conferring enzymatic stability. It is remarkable that trimannobioside **6a** is able to successfully inhibit DC-SIGN mediated HIV transfection of B-THP-1 cells,⁶⁸ while mannobioside **5b** proved to be three times more active than the corresponding α (1,2)-dimannoside in blocking virus binding for Ebola virus invasion model in Jurkat cells expressing DC-SIGN.⁶⁷

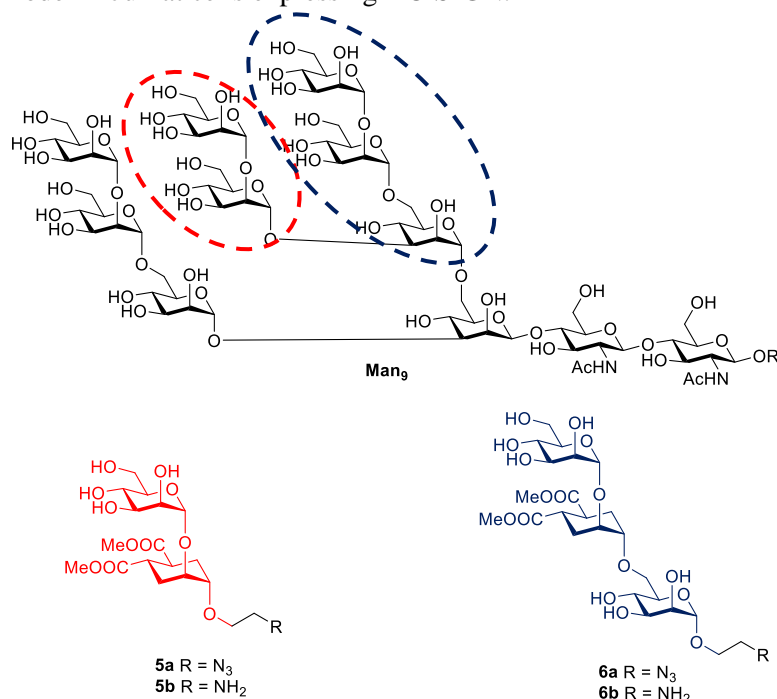


Figure 6 – Mannose based DC-SIGN glycomimetic ligands. The pseudo-dimannosides **5a,b** and pseudo trimannosides **6a,b** are glycomimetic structures of the highly mannosylated natural ligand Man₉. IC₅₀ values measured by SPR inhibition assay with BSA-mannotriose as competitor showed **5a**⁶⁷ IC₅₀ = 1.0 mM and **6a**^{69,70} IC₅₀ = 125 μM.

The less synthetically demanding **5a,b** were further optimised. A small library of bis-amido derivatives able to establish additional interactions was synthesised, with the identification of a benzyl substituted derivative as the most promising candidate.⁷¹ The library was therefore expanded with compounds belonging to the benzyl series whose increased affinity for DC-SIGN was confirmed by surface plasmon resonance (SPR) competition assay against immobilized polymannosylated-bovine serum albumin (BSA).⁷² Since Langerin is directly involved in protective mechanism against virus dissemination, selectivity is an essential requisite in the development of anti HIV adhesion molecules. Notably, almost all the library members showed low binding affinity towards Langerin, proving that achieving high affinity together with selectivity is not an impossible goal. Based on a combination of high affinity and selectivity, water solubility, affordable synthetic route, *p*-hydroxybenzylamide derivative **7** was selected as the best lead structure and the presence of a terminal azide function was exploited for the generation of multivalent glycodendrimers (Fig. 7).^{73,74}

A similar approach was adopted by Anderluh and co-workers⁷⁵ who functionalised a mannose residue with aryl substituents in order to target the Phe313 side chain, located in proximity to the DC-SIGN carbohydrate binding site. Bis-arylamides and *O*-aryl ethers were obtained tethering D-mannose either with a 1,3-diaminopropan-2-ol or with glycerol and tested *in vitro* for binding to DC-SIGN extracellular domain in competition with the Man₉ displaying HIV gp120 protein. The most active compounds 1,3-diarylglycerol **8a** and **8b** are shown in Figure 7.

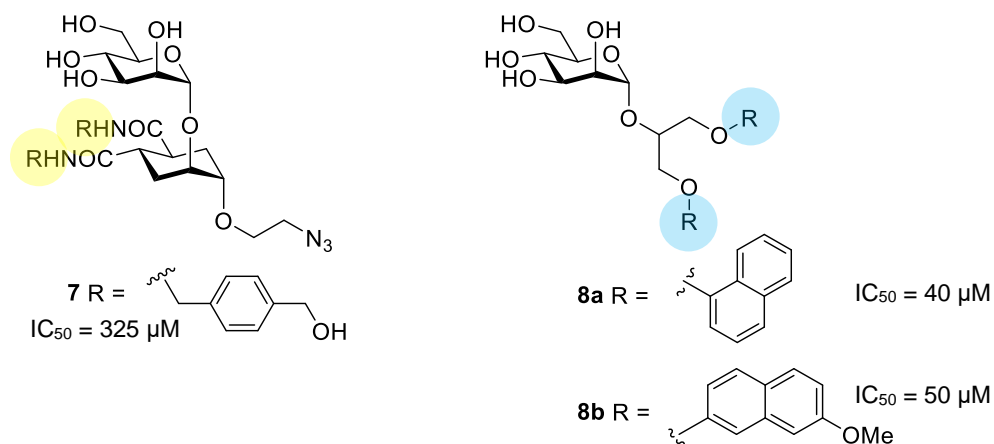
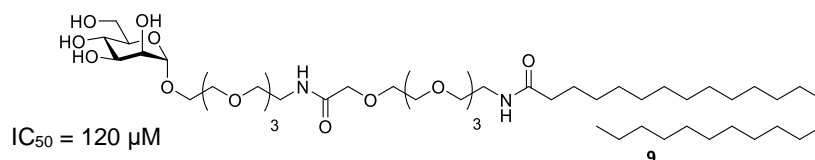


Figure 7 - Mannose based DC-SIGN glycomimetic ligands. The *p*-hydroxybenzylamide derivative **7** is a selective and potent antagonist of DC-SIGN, $IC_{50} = 325 \mu\text{M}$ measured in SPR inhibition assay versus polymannosylated immobilized-BSA.⁷² Targeting of Phe313 side chain with *O*-aryl ethers of mannosylated glycerol led to the potent DC-SIGN ligands **8a** and **8b**, IC_{50} 40 μM and 50 μM respectively. IC_{50} values were measured *in vitro* for binding to DC-SIGN extracellular domain in competition with the Man₉ displaying HIV gp120 protein.⁷⁵

DC-SIGN monovalent ligands comprise also the mannosylated glycolipid **9**, which demonstrated to be active in inhibiting HIV invasion,⁷⁶ and the carbohydrate inspired compound **10** based on the shikimic acid scaffold that matches the configuration of hydroxyl groups in 2-, 3- and 4- position of mannose (Fig. 8).⁷⁷



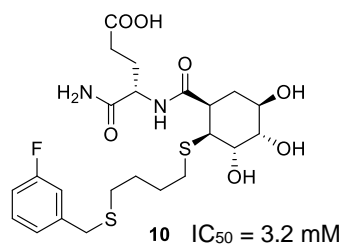


Figure 8 - Mannose based DC-SIGN glycomimetic ligands. Mannosylated glycolipid **9** performs as DC-SIGN ligand and proved to inhibit DCs mediated infection of T-Cells, $IC_{50} = 120\mu\text{M}$.⁷⁶ The shikimic acid derivative **10** is a mimic structure of mannose which showed $IC_{50} = 3.2 \text{ mM}$ in inhibition assay using Man-BSA as competitor.⁷⁷

0.3.2 Multivalent ligands

Protein-carbohydrate interactions are relatively weak, with dissociation constants often lying in the low millimolar range. This problem has been smartly overcome by Nature exploiting multivalency. Indeed, high affinity can be achieved through multiple interactions between lectin receptors and target glycans. As a consequence, lectin receptors have either multiple carbohydrate recognition domains or an oligomeric structure that allow multiple binding towards carbohydrate ligands, usually exposed in several copies forming polyglycosylated surfaces. Multivalent interactions affect lectin affinity as well as selectivity, moreover the density of multipresented glycans, their topology and orientation have emerged as important parameters.^{78,79} As Nature suggests, multivalent glycoconjugate synthesis represents an excellent strategy for targeting C-type lectins (Fig. 9), hence, for a rational design of polyvalent lectin ligands, the knowledge of the theoretical principles at the bases of multivalency is essential.

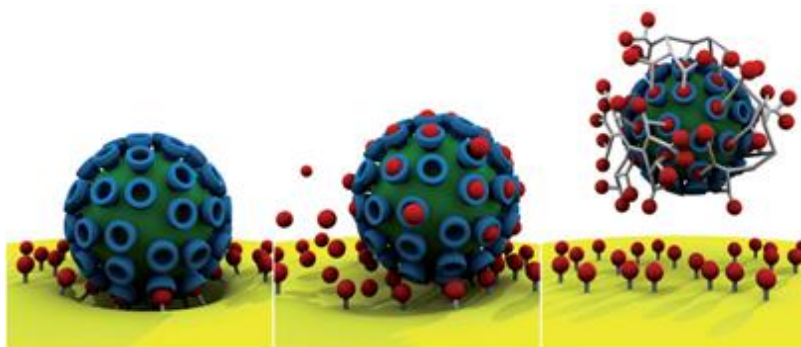


Figure 9 - Targeting protein-carbohydrate interactions. Protein-carbohydrate interactions usually take place between a multivalent receptor and a polyglycosylated target. To inhibit such interactions monovalent or multivalent ligands can be exploited. The picture depicts the comparison between the monovalent and multivalent ligand efficacy. While monovalent ligand affinity may be not strong enough for them to compete for binding with native ligands, multivalent compounds can really help in overcoming this problem.⁸⁰

0.3.2.1 Multivalency: a theoretical point of view

It is correct to speak about multivalency when a system composed by $n > 1$ entities shows a different behaviour compared to a single isolated component. In particular, a multivalent interaction requires an m -valent receptor interacting with an n -valent ligand (with m equal or not to n and both $m, n > 1$). Thus, an m -valent receptor interacting simultaneously with n isolated ligands is not a multivalent interaction.⁸⁰ To quantify the affinity enhancement in multivalent interactions, the enhancement factor β was introduced, expressed as the ratio between the binding constant of a multivalent ligand (K_{multi}) for its receptor and the binding constant of the corresponding monovalent ligand (K_{mono}).⁸¹

$$\beta = \frac{K_{multi}}{K_{mono}}$$

The enhancement of multivalent ligand affinity for a multivalent receptor can be explained by three mechanisms that can be operative simultaneously or not: chelation, statistical rebinding and protein aggregation (see below Fig. 10).⁸²

Chelation is the ability of a multivalent ligand to bind at least two binding sites of a multivalent receptor (Fig. 10a). Taking as an example a divalent receptor, the equilibrium constant for the binding of a chelating divalent ligand is far greater than the product of the two equilibrium constants for the binding of two monovalent ligands. This greater affinity is mainly due to entropic effects. Binding of a ligand to a receptor is associated with reduced mobility and loss of degrees of freedom, since several translational and rotational states are no longer accessible. The drop of entropy ($\Delta S < 0$) reduces the variation of free energy associated with the binding process ($\Delta G = \Delta H - T\Delta S$), but monovalent and chelating ligands experience it in different extents. In the case of two monovalent ligands, the free energy penalty that has to be paid is the same for the two binding events. In the case of a chelating divalent ligand, the second binding is associated with a significantly reduced drop of entropy (thus less negative ΔG), since most of the entropy penalty has already been paid during the first binding event. Chelation probably provides the major contribution to affinity enhancement that is observed in multivalent interactions. Nevertheless, an efficient chelation can only occur in the case of a multivalent antagonist with an optimized scaffold that guarantees a proper distribution of the ligands in order to bridge two different binding sites of the receptor.^{74,83,84}

Statistical rebinding can also contribute to the affinity enhancement of a multivalent ligand (Fig. 10b). This effect takes place between two or more ligands tethered to the same multivalent construct and one binding site of a (multivalent) receptor. In particular, after a first binding event, additional ligands may be located in close proximity of the binding site. In case of dissociation, the highly effective local concentration of ligands will ensure a rapid rebinding of the binding site, thus leading to a high affinity interaction.

A similar mechanism, that has been observed in multivalent interactions is described by the bind and slide model.⁸⁵ Such a behaviour was originally demonstrated for DNA-protein interactions and in the past few years it has been extended to lectin-polyglycosylated surface interactions. As the name of the model suggests, lectin receptors are depicted as entities able to bind to ligands in a dynamic way and to move along the polyglycosylated surface. This mechanism is the consequence of the detachment from a ligand and the subsequent binding to another in close proximity, which determines the sliding motion, a reduced off rate, thus an increased affinity. It has been proposed that the bind and slide phenomenon may facilitate receptor migration and their aggregation (Fig. 10c),^{86,87} therefore regulating the triggering and transduction of cell signalling.⁸⁸

Concerning affinity enhancement, the last two mechanisms described above generally account for minor contributions compared to chelation, but they gain more and more importance as the number of interactions increases.

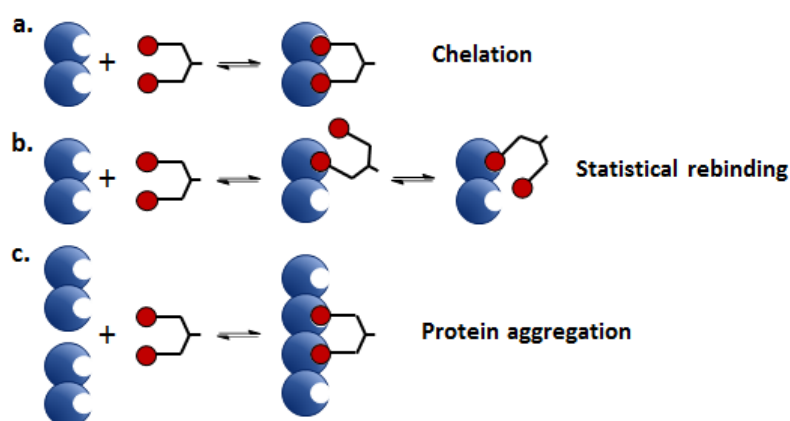


Figure 10 – Affinity enhancement mechanisms. The mechanisms governing multivalency are depicted for a divalent receptor (shown in blue) interacting with a divalent ligand (shown in red). **a)** Chelation is achieved through concomitant binding of (at least) two binding sites. A spacer of proper length with a good balance between rigidity and flexibility is required. **b)** Statistical rebinding is given by the increased local concentration of ligand, which enables high rebinding rate of one binding site upon dissociation. **c)** Protein aggregation may occur when a ligand simultaneously binds to binding sites belonging to distinct receptors.

Accurate theoretical treatments of multivalent systems from a thermodynamic point of view are extremely challenging. Interactions in a multivalent system have different properties if compared to the corresponding monovalent one and this behaviour is explained with the concept of cooperativity. Many groups have tried to identify and define the different types of cooperativity, which however remains a subject of great debate.^{80,81,89,90}

The difficulty in analysing such interactions is exemplified by the study of the simplest multivalent interaction, a divalent ligand and a divalent receptor, that nevertheless turns to be extremely complex. All the possible equilibria of the system are depicted in Figure 11.

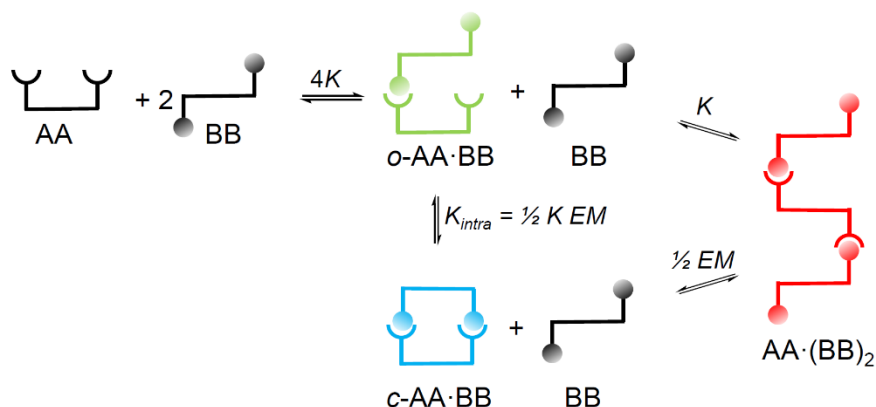


Figure 11 - Schematic representation of all the possible equilibria for interaction between a divalent receptor with a divalent ligand. Assumption of high concentration of the ligand enables to neglect states involving more than one receptor. K is the microscopic intermolecular association constant, while K_{intra} is the microscopic intermolecular association constant and EM is defined as the effective molarity.⁹⁰

For sake of simplicity, the ligand is considered in great excess relative to the receptor, which allows to neglect species presenting more than one receptor, and the only cooperativity effect taken in consideration is chelate cooperativity.⁹⁰ The accessible species for the receptor are four: the not bounded AA, the partially bounded open complex *o*-AA·BB, the chelated *c*-AA·BB and the totally bounded AA·(BB)₂. The strength of the binding interaction is defined by K , the microscopic intermolecular association constant, while the strength

of the intramolecular binding is expressed by the association constant $K_{\text{intra}} = \frac{1}{2} K \text{EM}$, where $\frac{1}{2}$ is the statistical factor associated with the cyclization and the constant EM, the microscopic effective molarity, takes into account the tendency to chelation. The relative population of the partially and fully bound species as a function of free ligand BB concentration is reported in Figure 12. In particular, Figure 12a shows the population distribution for a system with $K_{\text{intra}} = 0$, thus without chelate interaction. At this borderline situation, of course $c\text{-AA}\cdot\text{BB}$ is completely non populated, while both the population of $o\text{-AA}\cdot\text{BB}$ and $\text{AA}\cdot(\text{BB})_2$ gradually increase along with ligand concentration until $\text{AA}\cdot(\text{BB})_2$ start to be more favoured and become the only accessible. On the other hand Figure 12b present the distribution of populations in a system with high intramolecular association constant, $K_{\text{intra}} = 25$. In this case, the chelated $c\text{-AA}\cdot\text{BB}$ is highly populated, showing a bell-shaped distribution. This state is far more stable than the partially bound $o\text{-AA}\cdot\text{BB}$ therefore, at low concentration of ligand BB, is the only accessible state and its population increases along with ligand concentration until almost all the binding sites are occupied. At this point, as the ligand concentration further increases, the fully bound $\text{AA}\cdot(\text{BB})_2$ gets more and more stabilised until it becomes the only populated state.

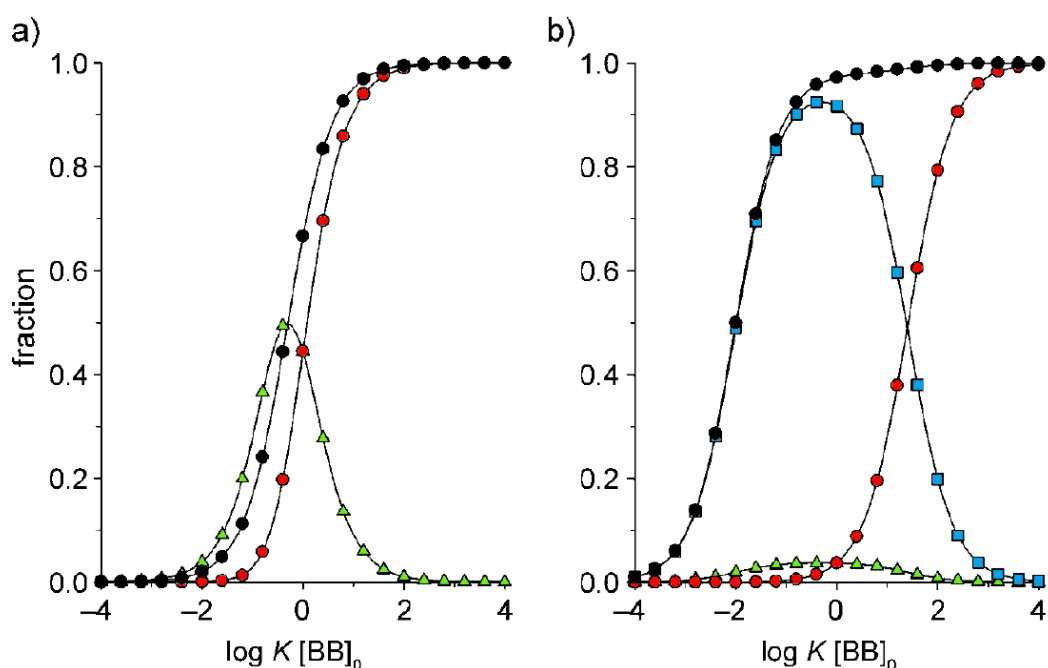


Figure 12 - Speciation profiles. a) Speciation profile of the equilibria for a system not affected by chelate interaction, $K_{\text{intra}} = 0$; $\text{EM} = 0$. b) Speciation profile of the equilibria for a system showing strong chelate interaction, $K_{\text{intra}} = 25$; $\text{EM} = 50$. The concentration scale on the abscissa is normalized by the equilibrium constant K . Fractions of the species are reported as follows: (●) = total fraction of receptor AA species with occupied binding sites; (●) = fraction of $\text{AA}\cdot(\text{BB})_2$; (▲) = fraction of $o\text{-AA}\cdot\text{BB}$; (■) = fraction of $c\text{-AA}\cdot\text{BB}$.⁹⁰

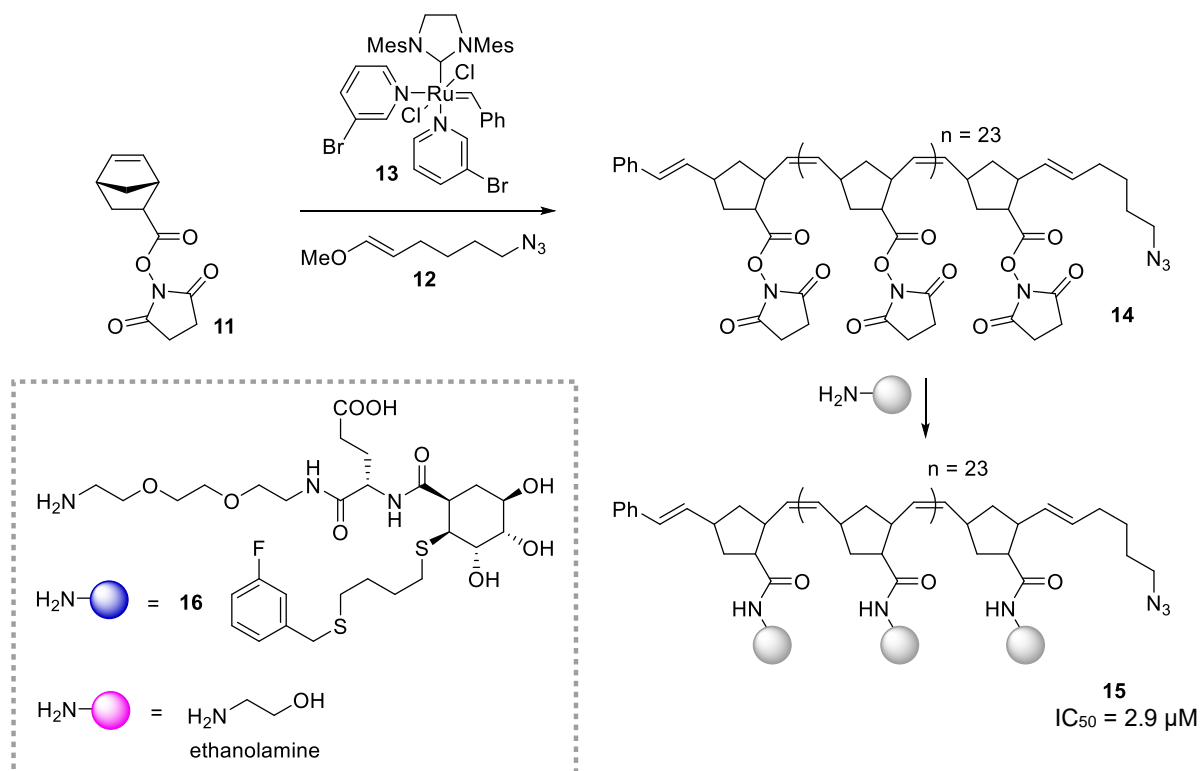
The great variability of population distribution that can be observed in a simple multivalent system gives an idea of how complex is the study and the analysis of multivalent interactions. We have just seen how ligand concentration and effective molarity (which is a constant property of the system) greatly influence the species composition of the system. Moreover, the initial assumption that ligand BB is in great excess relative to the receptor should not be forgotten. A change in ligand to receptor ratio may result in further accessible states envisaging more than one receptor in the binding, therefore increasing the complexity of the system and enabling clustering and aggregation processes.

0.3.2.2 Multivalent glycoconjugates

In the attempt of interfering with lectin mediated biological processes, many potent multivalent antagonists have been successfully synthesised. Notably, the effectiveness of a glycoconjugate as an antagonist is dictated not only by the nature of the ligands engaged, but it also depends on the architecture of the polyvalent scaffold, the valency, the ligand density, the kind of linker engaged and the flexibility of the construct. For this reason, a huge variety of glycoconjugate structures comprising glyconanoparticles, glyconanodiamonds, glycopolymers, fullerenes, self-assembled Janus glycodendrimers,⁹¹ liposomes, micelles and nanowires,⁹² cyclodextrins, calixarenes, glycodendrons and glycodendrimers have been investigated and successfully employed as antagonists for different kind of lectins such as galectins, Siglecs, bacterial lectins, bacterial toxins and C-type lectins. (Multivalent glycoconjugates have been extensively reviewed).^{59,62c,93,94,95,96,97,98}

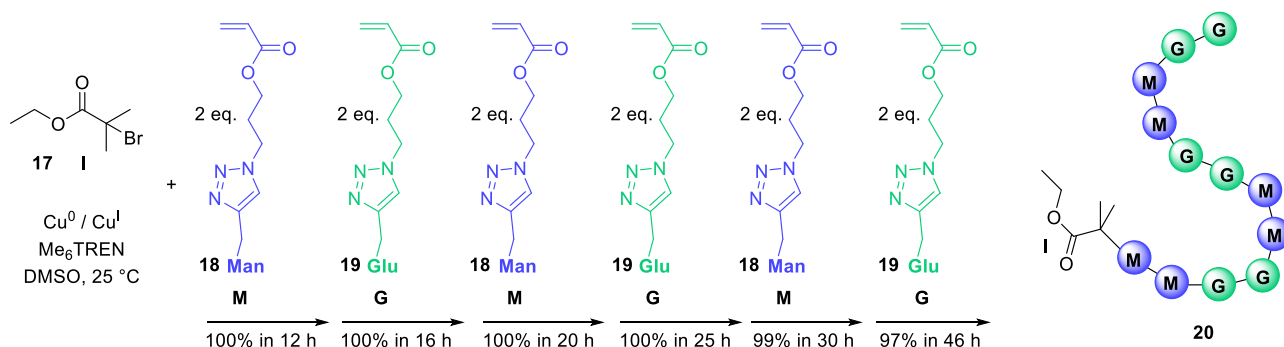
For the sake of clarity, we will focus our attention on the applications of multivalent glycoconjugates as C-type lectin antagonists. The great majority of synthetic C-type lectin multivalent ligands reported so far have been developed for DC-SIGN targeting and an overview on the great diversity of such antagonists is herein presented. On the other hand, few multivalent glycoconjugates have been developed to target MBL. This is probably mostly due to the difficulty in studying this protein, which is expressed as different isoforms in different organisms and present as scarcely characterized aggregates. Some of the few multivalent MBL antagonists so far reported will be also described.

Glycopolymers have been reported as convenient scaffolds for glycan multipresentation to lectins and new generation polymerization strategies have enabled a certain control in terms of polymer length and ramification. Kiessling and co-workers applied ring opening metathesis polymerization (ROMP) for the synthesis of a polymer loaded with a mannose mimic derived from shikimic acid.⁷⁷ ROMP of *N*-hydroxysuccinimide activated norbornene carboxylic acid monomers **11** allowed for subsequent functionalization with the amino tethered glycomimetic **16** (Scheme 1). Performing the reaction in a mixture with ethanolamine furnishes a polymer with well defined density of glycomimetics **15**: the multivalent polymers were 25 monomeric units long displaying on average seven glycomimetics per polymer chain. The binding ability of the polymer for DC-SIGN was established by a competitive assay against a mannosylated (20-25 mannose copies displayed) fluorescein-labelled BSA. The IC₅₀ value for the multivalent antagonist binding to the immobilised DC-SIGN extracellular domain is 2.9 μM, thus gaining three orders of magnitude compared to the monovalent ligand, IC₅₀ = 3.2 mM.



Scheme 1 – Glycopolymers as DC-SIGN multivalent ligands. ROMPs have been exploited to prepare activated polymers which were functionalized with multiple copies of monovalent DC-SIGN ligand **16**. The glycopolymer **15** showed an increased affinity of 3 orders of magnitude $IC_{50} = 2.9 \mu\text{M}$ compared to the monovalent ligand **16**, $IC_{50} = 3.2 \text{ mM}$.⁷⁷

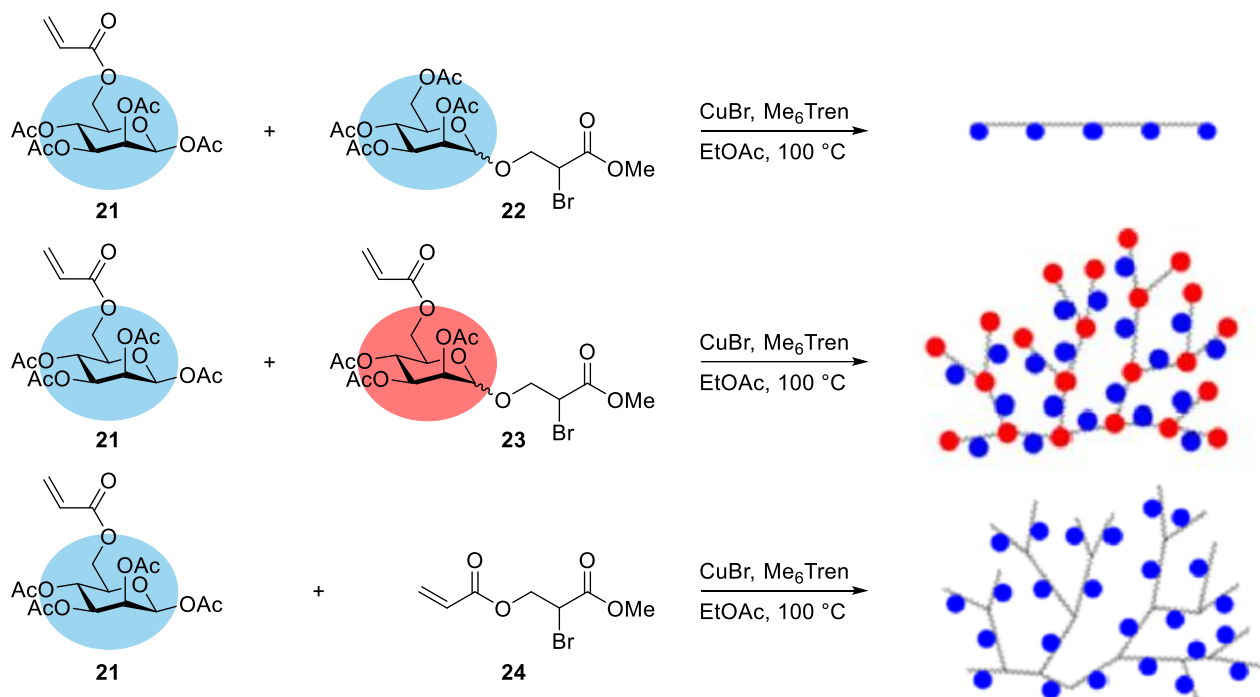
Following a similar post-polymerization modification approach, the Haddleton group reported the synthesis of even more potent mannosylated polymers as DC-SIGN antagonists.⁹⁹ The synthesis of multi-block glycopolymers has also been accomplished through single-electron transfer living radical polymerization (SET-LRP) (Scheme 2). Mannose and glucose derived glycomonomers **18** and **19** were successfully polymerized obtaining polymers with defined length and composition. In particular, the influence of the density and the distance between mannose units for DC-SIGN binding was evaluated by SPR competition assay flowing the polymers with soluble DC-SIGN over immobilised gp120 protein. The most active polymer, $IC_{50} = 153 \text{ nM}$, showed only 10 fold weaker activity compared to soluble gp120, $IC_{50} = 11 \text{ nM}$.¹⁰⁰



Scheme 2 – Multi-block glycopolymers as DC-SIGN ligands. Single-electron transfer living radical polymerization (SET-LRP) of mannosylated and glucosylated monomers allowed the synthesis of DC-SIGN antagonists with control over ligand density.¹⁰⁰

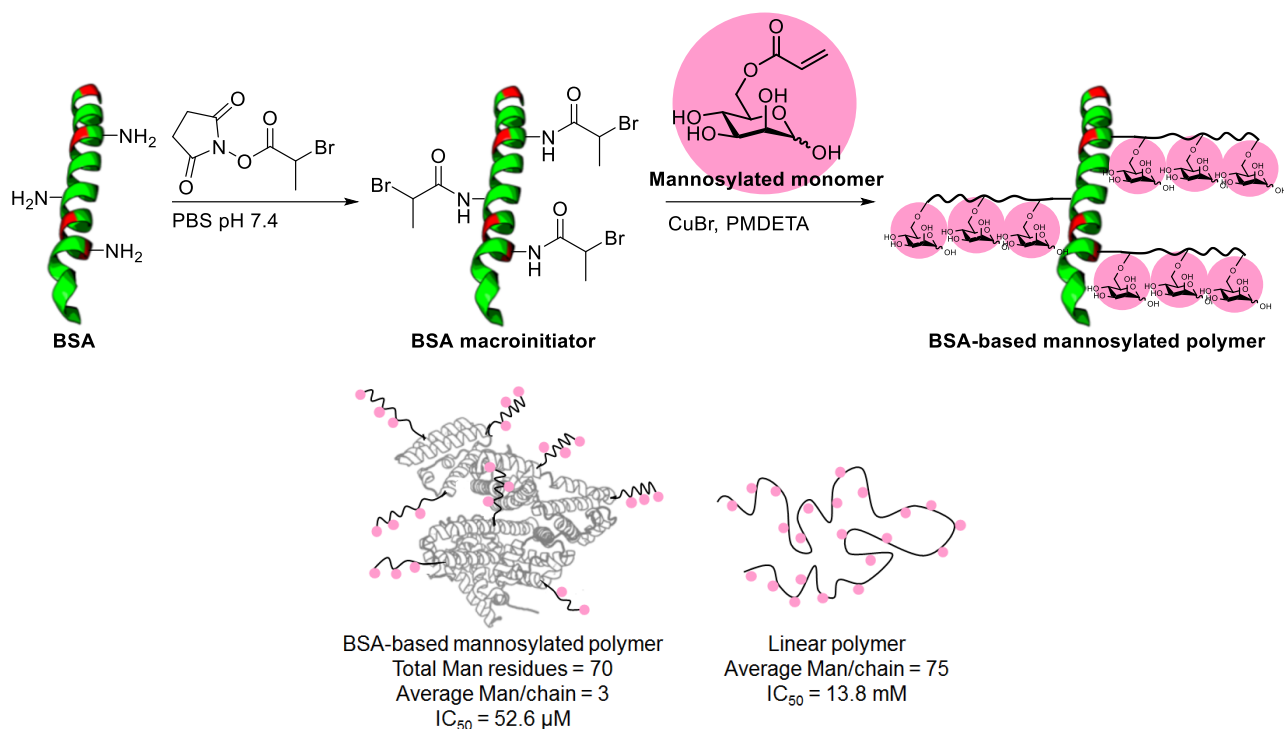
Mannosylated polymers have also shown to act as potent MBL antagonists. Exploiting atom transfer radical polymerization (ATRP), mannose monomers **21**, **22** and eventually inimer **24** and mannosylated inimer **23**

(inimers act both as monomers and as initiator units) have been polymerized to obtain linear and branched chains characterized by different lengths, mannose units and density, branch density, with or without mannose at the branching sites (Scheme 3).¹⁰¹ The ability in binding hMBL (human MBL) contained in human serum was evaluated in a competition enzyme-linked lectin assay (ELLA) with the immobilized natural polysaccharide mannan. As expected, the number of mannose residues exposed by the polymer greatly affects the binding activity. Interestingly, branched polymers were far more active than linear ones and the presence of a mannose residue at the branching site further increased the affinity.



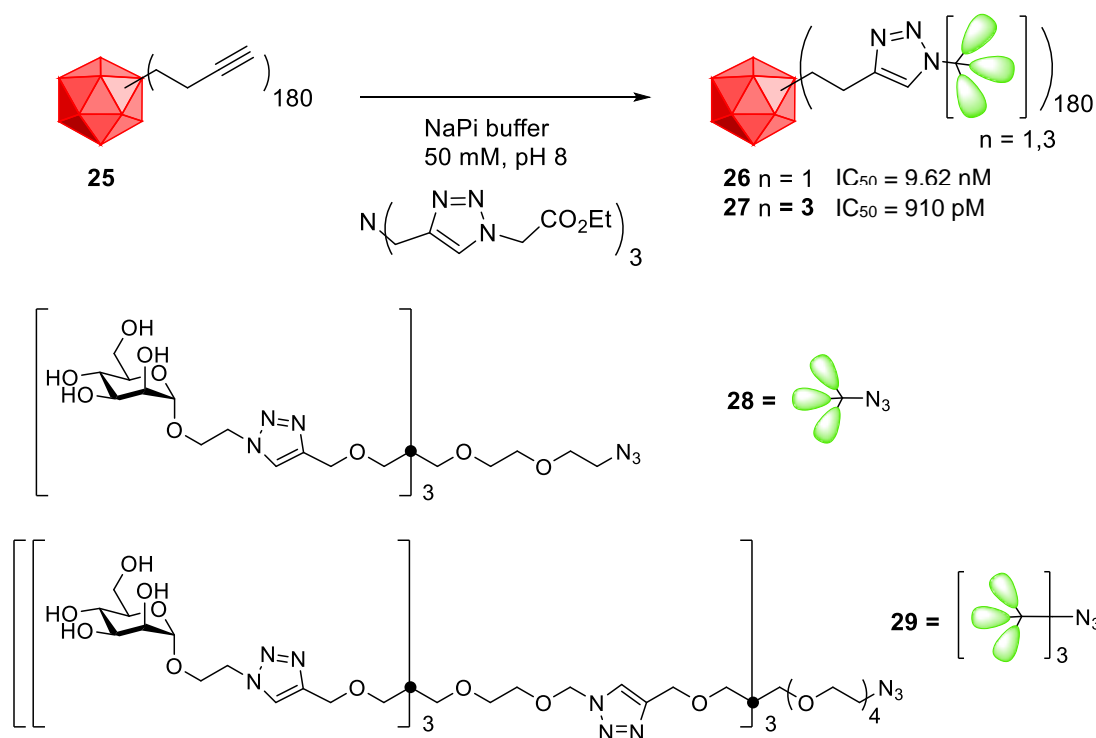
Scheme 3 - Mannosylated polymers as potent MBL antagonists. Atom transfer radical polymerization (ATRP) was applied to the synthesis of mannosylated polymers. Mannose monomers **21**, **22** and eventually inimer **24** and mannosylated inimer **23** were used. The use of inimer **24** allowed the introduction of ramifications, while inimer **23** generated branched chains with mannose residues at the ramification sites.¹⁰¹

In a recent work, the same authors further applied ATRP for the synthesis of MBL antagonists in a grafting-from approach. Direct polymerization of glycomonomers was accomplished using BSA as macroinitiator.¹⁰² The single conjugation of a glycopolymer to BSA for the development of MBL multivalent ligands had already been described,¹⁰³ but now the modification of exposed lysine residues into initiator unit allows the polymerization to start from multiple sites of the protein (Scheme 4). As a consequence, glycopolymers-BSA conjugates displaying from 70 up to 360 mannose residues per protein are formed with an overall spherical distribution of the polymeric chains. As confirmed by ELLA competition assay, this tridimensional multiple presentation of ligands confers greater affinity for rhMBL (recombinant human MBL) compared to simple polymeric structures. Indeed, the BSA conjugate exposing 70 mannose residues as an average of 3 residues per polymeric chain showed to antagonize MBL binding to immobilized mannan 260 times stronger than a corresponding glycopolymer displaying 75 mannose moieties.



Scheme 4 – Mannosylated polymer-BSA conjugates. The synthesis of BSA-based mannose polymer conjugates is shown. A representative portion of BSA structure is depicted in green. Amino residues of Lys side chains are functionalized to give a BSA macroinitiator. Atom transfer radical polymerization (ATRP) is then applied to polymerize mannose monomers giving the corresponding glycoconjugate with polymeric chains grafting out the BSA scaffold. ELLA competition assay showed that the three dimensional arrangement of the ligands of BSA-glycopolymer conjugate allowed to gain two orders of magnitude, $IC_{50} = 52.6 \mu\text{M}$, when compared to linear mannose polymer with very similar content of mannose residues, $IC_{50} = 13.8 \text{ mM}$.¹⁰²

As we have just seen, proteins are a powerful scaffold for multiple presentation. Recently, the oligomeric Ebola virus capsid was chosen as a platform for the preparation of highly multivalent DC-SIGN antagonists.¹⁰⁴ Ebola virus envelope is constituted by 180 copies of the monomeric Q β that assemble together forming an icosahedral structure of $\sim 28 \text{ nm}$ diameter. The over-expression of recombinant Q β in *E. coli* and self-assembly allowed the formation of capsid-like scaffold **25** with site specific introduction of the unnatural amino acid L-homopropargylglycine. This tag enabled the direct functionalisation through CuAAC with trivalent and nonavalent mannose glycodendrons **28** and **29**, resulting in the 540 fold Q β -(Man₃)₁₈₀ **26** and the 1620 fold Q β -(Man₉)₁₈₀ **27** mannose antagonists (Scheme 5). Q β -(Man₉)₁₈₀ **28** displays 1620 residues of mannose and is therefore the highest glycosylated structure with homogeneous presentation. The spherical symmetry of the molecule along with its dimension allowed impressive mimicking of the highly glycosylated Ebola virus envelope as it is demonstrated by DC-SIGN mediated Ebola virus infection model, where infection is inhibited with $IC_{50} = 9.62 \text{ nM}$ for Q β -(Man₃)₁₈₀ **26** and 910 pM for Q β -(Man₉)₁₈₀ **27**.



Scheme 5 – Ebola capsid is a valuable scaffold for highly valent DC-SIGN ligands. Ebola virus capsid **25** generated by self-assembly of 180 copies of mutated monomeric Q β was prepared. Trivalent and nonavalent mannosylated glycodendrons **28** and **29** were introduced by CuAAC affording the 540 fold Q β -(Man₃)₁₈₀ **26** and the 1620 fold (Q β -(Man₉)₁₈₀) **27** that strongly inhibited DC-SIGN mediated Ebola virus infection, $IC_{50} = 9.62 \text{ nM}$ and 910 pM respectively.¹⁰⁴

Self-assembly strategy was also applied by the Wagner group, showing that the inhibitory potency of mannoside glycolipids for DC-SIGN mediated HIV infection can be drastically enhanced upon formation of micelles. Particularly the best inhibition was achieved with trimannosylated glycolipid **30** that showed two inhibition rates in dose-response inhibition of HIV-1 *trans*-infection model, (Fig. 13) with the faster slope *b* reached upon formation of dynamic micelles. Moreover the authors showed that dynamic micelles of mannoside glycolipids are more effective in binding DC-SIGN compared to corresponding rigid polymeric structures.¹⁰⁵

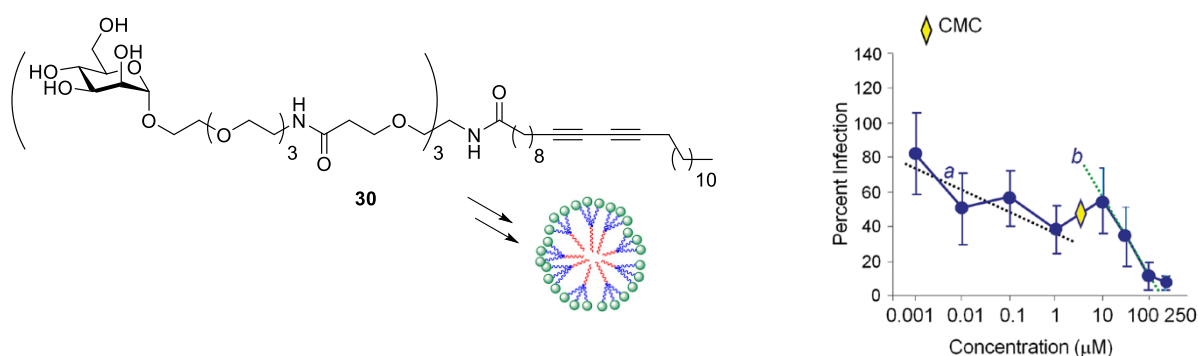


Figure 13 – DC-SIGN multivalent antagonists by self-assembly. Trimannosylated glycolipid **30** displayed two inhibition rates in dose-response inhibition of HIV-1 *trans*-infection model. The dynamic formation of highly valent micelles determines the faster slope *b*. The CMC (critic micellar concentration) is shown.¹⁰⁵

Gold nanoparticles (GNPs) have demonstrated to be a valuable scaffold for multipresentation of oligomannosides.¹⁰⁶ Different degrees of functionalisation can be achieved, with 50 % of density being

sufficient to reach great inhibition in DC-SIGN triggered HIV *trans*-infection studies. Gold nanoparticles bearing dimannoside up to heptamannoside moieties all showed inhibition ability in the low nanomolar and subnanomolar range. Notably, the best inhibition was achieved with GNP **31** displaying tetramannoside moieties, but high potency was also reached by nanoparticles tethered with the simpler dimannoside and trimannoside structures (Fig. 14).¹⁰⁷ Gold nanoparticles carrying fucose based ligands were also addressed in DC-SIGN binding (Fig. 14). The *N*- α -fucosyl- β -alanyl- β -alanine derivative **32** showed to be potent DC-SIGN ligand, able to trigger internalization of the receptor in dendritic cells (DCs) without the induction of DC maturation, thus emerging as promising DC-SIGN targeting tools for selective delivery of antigens and dyes.¹⁰⁸

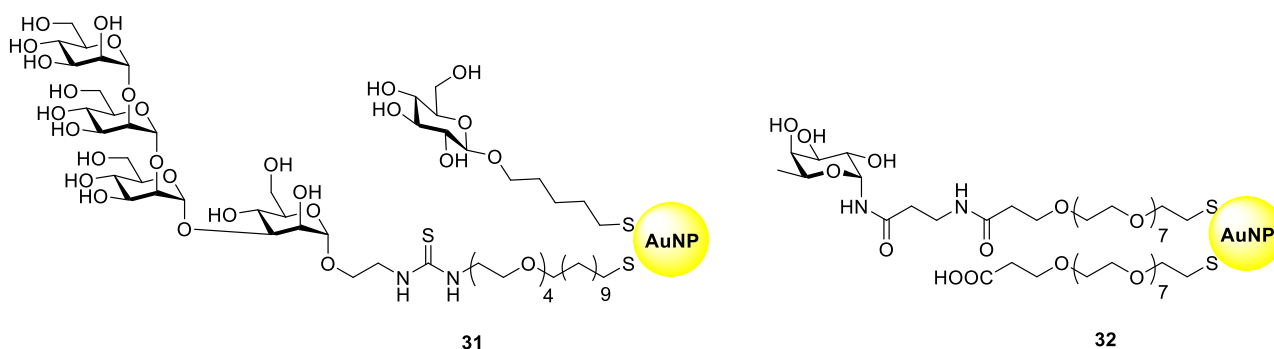
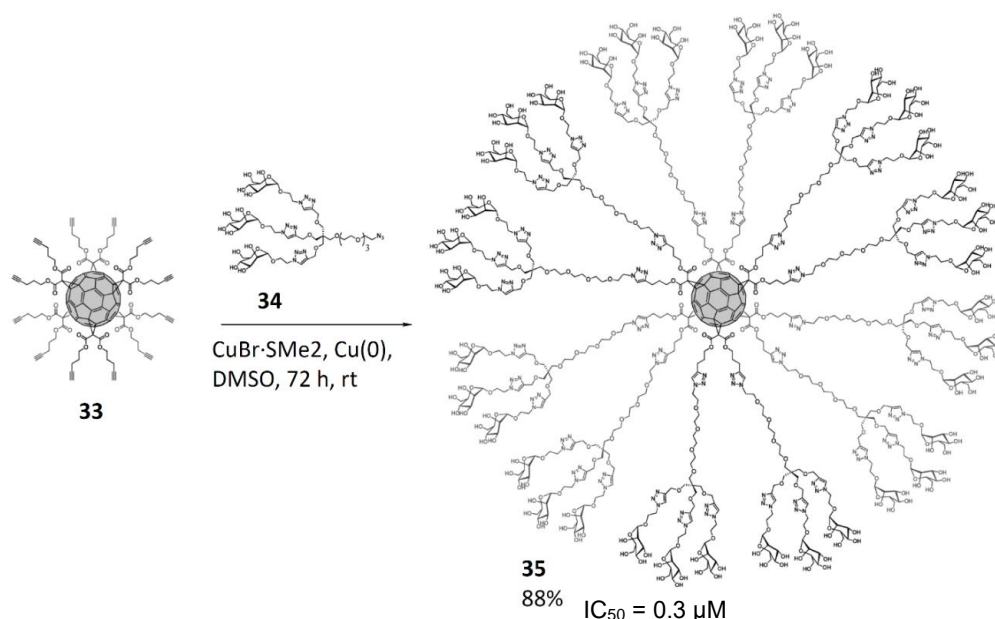


Figure 14 - Nanoparticles as platform for multipresentation of glycans. Gold nanoparticles have been described for multipresentation of DC-SIGN ligands. Gold nanoparticles **31** and **32** bearing either a tetramannoside or a fucose based ligand are shown.^{107,108}

Highly homogeneous constructs with spherical distribution of glycans can be obtained employing the rigid structure of fullerenes.¹⁰⁹ So called fullerene “sugar balls” displaying up to 36 mannose residues have been obtained through CuAAC between hexakis-adduct **33** and mannosylated glycodendrons. The resulting conjugates were tested as DC-SIGN ligands in Ebola virus infection model. Results identified glycofullerene **35** as the most active, $IC_{50} = 0.3 \mu M$, and highlighting the number of glycans exposed and especially linker length as important parameters for the control of glycan density and accessibility for DC-SIGN receptor (Scheme 6).¹¹⁰



Scheme 6 – Fullerene “sugar balls”. CuAAC have been exploited to connect mannosylated dendrons to conveniently functionalized fullerenes. This approach led to construct **35** which showed an $IC_{50} = 0.3 \mu M$ for DC-SIGN mediated Ebola virus infection.¹¹⁰

Glycodendrimers and glycodendrons are promising platforms for lectin antagonists with potential clinical applications. The synthesis of these constructs has been accomplished either through divergent or convergent routes leading to products with various scaffold structure and properties. The generation of mannosylated Boltorn-type dendrons **36** and **37** have been reported previously by our group (Fig. 15).^{111,68} These two dendrons are characterised by a common tetravalent polyester scaffold bearing four copies of either pseudobiomannoside **5b** or pseudotrimannoside **6b** ligands, properly designed to mimic the natural mannose glycan Man₉ often exposed in multiple presentation on several pathogen glycoproteins. Both dendrons act as potent inhibitors of DC-SIGN mediated *cis*- and *trans*-infection in Jurkat cell Ebola model,¹¹¹ while strong inhibition of DC-SIGN mediated *trans*-infection was further confirmed for dendron **36** by in vitro HIV infection assay.⁶⁸ Moreover, **37** proved to guarantee a protective effect in MBL triggered brain ischemic injury in acute stroke in mice.⁵¹

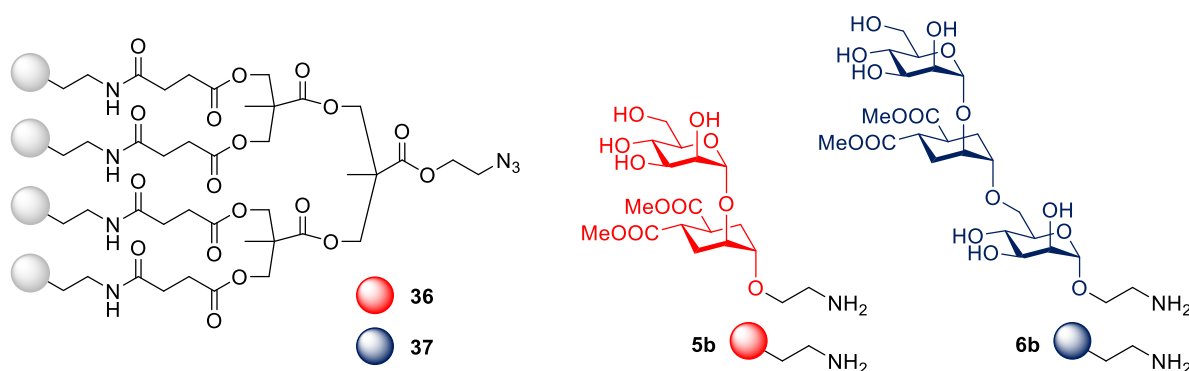


Figure 15 - Tetramannosylated dendrons 36 and 37. The polyester scaffold is either functionalized with pseudobiomannoside **5b** or pseudotrimannoside **6b** ligands.

Scaffolds based on small molecules such as aromatic compounds¹¹² and sugars¹¹³ have been largely employed. Exploiting Cu catalysed azide alkyne cycloaddition (CuAAC), pentaerythritol based glycodendrons **38** and **39** were prepared, allowing the synthesis of higher valency compounds up to the 18-valent glycodendrimers **40** and **41** (Fig. 16). Constructs bearing multiple copies of the bis-amido glycomimetic **7** revealed as the most active DC-SIGN antagonists in HIV and Dengue virus infection model.⁷³ Despite the high affinity of these compounds for DC-SIGN, their structure is too short to bridge at least two carbohydrate binding sites. The introduction of a rigid spacer, derived from PEGylated phenyl acetylene moieties, enabled chelation resulting in the extremely potent DC-SIGN antagonists **42**, **43**, with IC₅₀ values in the low nanomolar range in HIV infection studies.⁷⁴

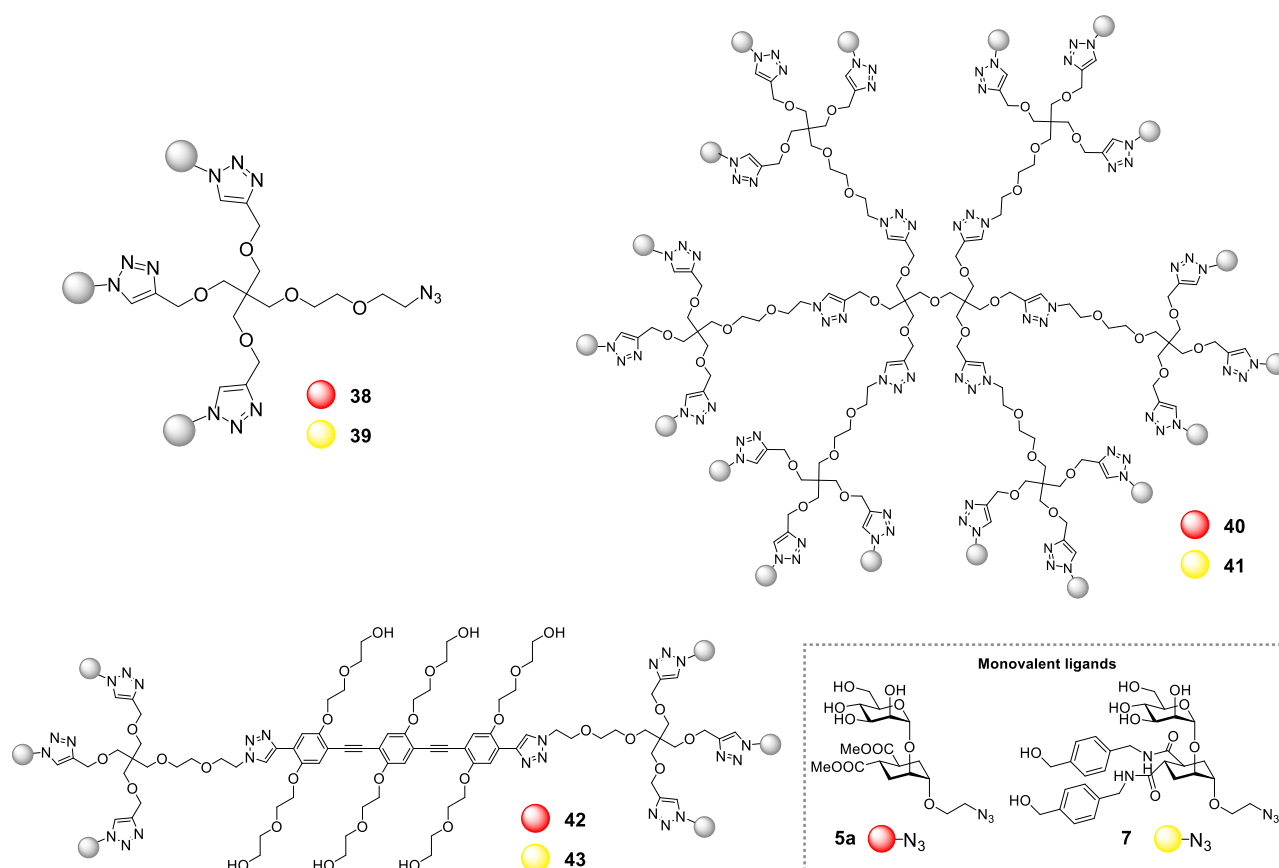


Figure 16 - Multipresentation of glycomimetics. Up to 18-valent glycodendrimers **40** and **41** have been prepared by CuAAC of azido tagged trivalent glycodendrons **38** and **39** with a dipentaerythritol based scaffold.⁷³ The same strategy was applied to functionalize a rigid spacer, thus generating chelating antagonists **42** and **43**.⁷⁴

0.4 Aim of the work

So far we have appreciated the myriad of biological processes, both physiological and pathological, that are regulated and mediated by lectins. The generation of molecular tools able to interact strongly with these receptors is a challenging goal. Remarkable progresses have been possible with the synthesis of new generation of glycomimetics and glycoconjugates along with more insights in the principles that govern multivalency. Yet, the optimization of pharmacokinetic properties, the improvement of ligand affinity and selectivity remain an issue for clinical applications.

All these problems were addressed in this PhD thesis work, whose final goal is the identification and the synthesis of new C-type lectins antagonists. Two main strategies were pursued:

- *Scaffold optimization.* New multivalent constructs have been designed for MBL and DC-SIGN binding through structure optimization of previously developed antagonists. In particular, stabilised versions of the already reported hydrolytically sensitive glycodendrons **36** and **37** (Fig. 15) were successfully prepared employing a tetramide scaffold that showed great stability and versatility for the generation of compounds with increased valency. Moreover, guided by the high affinity for DC-SIGN achieved with chelating ligands **42**, **43**, the synthesis of a tetravalent antagonist with a rigid architecture allowing simultaneous binding of the four C-type lectin like domains of DC-SIGN tetramer was undertaken.
- *Ligand optimization.* The development of new potent and selective C-type lectin antagonists is highly desirable and glycan microarrays have proved to be a useful tool for the identification of promising ligand candidates. In this context, preliminary studies for the chemoenzymatic synthesis of MALDI-

TOF detectable microarrays of monovalent glycomimetic structures were conducted. More in detail, the synthesis of clickable NDP-sugar nucleotides was attempted to enable on-chip glycosylation and subsequent library diversity expansion through CuAAC.

0.5 References

- ¹ R. Apweiler, H. Hermjakob, N. Sharon *Biochim. Biophys. Acta* **1999**, *1473*, 4-8.
- ² T. Wennekes, R. J. B. H. N. van den Berg, R. G. Boot, G. A. van der Marel, H. S. Overkleeft, J. M. F. G. Aerts *Angew. Chem. Int. Ed.* **2009**, *48*, 8848-8869.
- ³ J. W. VanTeeffelen, J. Brands, E. S. Stroes, H. Vink *Trends Cardiovasc. Med.* **2007**, *17*, 101-105.
- ⁴ M. Nieuwdorp, M. C. Meuwese, H. L. Mooij, C. Ince, L. N. Broekhuizen, J. J. P. Kastelein, E. S. G. Stroes, H. Vink *J. Appl. Physiol.* **2008**, *104*, 845-852.
- ⁵ C. R. Bertozzi, L. L. Kiessling *Science* **2011**, *291*, 2357-2364.
- ⁶ R. D. Cummings *Mol. BioSyst.* **2009**, *5*, 1087-1104.
- ⁷ H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero, H. Rüdiger *ChemBioChem* **2004**, *5*, 740-764.
- ⁸ R. S. Haltiwanger, J. B. Lowe *Annu. Rev. Biochem.* **2004**, *73*, 491-537.
- ⁹ K. Ohtsubo, J. D. Marth *Cell* **2006**, *126*, 855-867.
- ¹⁰ Y. van Kooyk, G. A. Rabinovich *Nat. Immunol.* **2008**, *9*, 593-601.
- ¹¹ H. Lis, N. Sharon *Chem. Rev.* **1998**, *98*, 637-674.
- ¹² A. N. Zelensky, J. E. Gready *FEBS J.* **2005**, *272*, 6179-6217.
- ¹³ K. Drickamer, M. E. Taylor *Curr. Opin. Struct. Biol.* **2015**, *34*, 26-34.
- ¹⁴ K. Drickamer *Nature* **1992**, *360*, 183-186.
- ¹⁵ J. Banchereau, R. M. Steinman *Nature* **1998**, *392*, 245-252.
- ¹⁶ R. M. Steinman, J. Banchereau *Nature* **2007**, *449*, 419-426.
- ¹⁷ S. Akira, S. Uematsu, O. Takeuchi *Cell* **2006**, *124*, 783-801.
- ¹⁸ Selected reviews for the role of C-type lectin in immunity: a) T. B. H. Geijtenbeek, S. J. van Vliet, A. Engering, B. A. Hart, Y. van Kooyk *Annu. Rev. Immunol.* **2004**, *22*, 33-54; b) T. B. H. Geijtenbeek, S. I. Gringhuis *Nature Reviews* **2009**, *9*, 465-479; c) I. M. Dambuza, G. D. Brown *Curr. Opin. Immunol.* **2015**, *32*, 21-27.
- ¹⁹ S. E. Hardison, G. D. Brown *Nat. Immunol.* **2012**, *13*, 817-822.
- ²⁰ N. C. Rogers, E. C. Slack, A. D. Edwards, M. A. Nolte, O. Schulz, E. Schweighoffer, D. L. Williams, S. Gordon, V. L. Tybulewicz, G. D. Brown, C. Reis e Sousa *Immunity* **2005**, *22*, 507-517.
- ²¹ K. Sato, X. Yang, T. Yudate, J.-S. Chung, J. Wu, K. Luby-Phelps, R. P. Kimberly, D. Underhill, P. D. Cruz, Jr., K. Ariizumi *J. Biol. Chem.* **2006**, *281*, 38854-38866.
- ²² S. Yamasaki, M. Matsumoto, O. Takeuchi, T. Matsuzawa, E. Ishikawa, M. Sakuma, H. Tateno, J. Uno, J. Hirabayashi, Y. Mikami, K. Takeda, S. Akira, T. Saito *Proc. Nat. Acad. Sci. U.S.A.* **2009**, *106*, 1897-1902.
- ²³ S. I. Gringhuis, J. den Dunnen, M. Litjens, B. van het Hof, Y. van Kooyk, T. B. H. Geijtenbeek *Immunity* **2007**, *26*, 605-616.
- ²⁴ W. Cao, L. Zhang, D. B. Rosen, L. Bover, G. Watanabe, M. Bao, L. L. Lanier, Y.-J. Liu *PLoS Biol* **2007**, *5*, e248.
- ²⁵ F. Meyer-Wentrup, A. Cambi, B. Joosten, M. W. Looman, I. J. M. de Vries, C. G. Figdor, G. J. Adema *J. Leukocyte Biol.* **2009**, *85*, 518-525.
- ²⁶ S. Yamasaki, E. Ishikawa, M. Sakuma, H. Hara, K. Ogata, T. Saito *Nat. Immunol.* **2008**, *9*, 1179-1188.
- ²⁷ A. B. H. Bakker, E. Baker, G. R. Sutherland, J. H. Phillips, L. L. Lanier *Proc. Nat. Acad. Sci. U.S.A.* **1999**, *96*, 9792-9796.
- ²⁸ M. Richard, N. Thibault, P. Veilleux, G. Gareau-Page, A. D. Beaulieu *Mol. Immunol.* **2006**, *43*, 1716-1721.
- ²⁹ A. S. J. Marshall, J. A. Willment, H.-H. Lin, D. L. Williams, S. Gordon, G. D. Brown *J. Biol. Chem.* **2004**, *279*, 14792-14802.
- ³⁰ T. B. H. Geijtenbeek, S. J. van Vliet, E. A. Koppel, M. Sanchez-Hernandez, C. M. J. E. Vandenbroucke-Grauls, B. Appelmelk, Y. van Kooyk *J. Exp. Med.* **2003**, *197*, 7-17.
- ³¹ L.-F. Chen, S. A. Williams, Y. Mu, H. Nakano, J. M. Duerr, L. Buckbinder, W. C. Greene *Mol. Cell. Biol.* **2005**, *25*, 7966-7975.
- ³² L.-F. Chen, Y. Mu, W. C. Greene *EMBO J.* **2002**, *21*, 6539-6548.
- ³³ T. B. H. Geijtenbeek, S. I. Gringhuis *Nat. Rev. Immunol.* **2009**, *9*, 465-479.
- ³⁴ T. B. H. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. F. van Duijnhoven, J. Middel, I. L. M. H. A. Cornelissen, H. S. L. M. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor, Y. van Kooyk *Cell* **2000**, *100*, 587-597.
- ³⁵ S. Pohlmann, F. Baribaud, R. W. Doms *Trends Immunol.* **2001**, *22*, 643-646.
- ³⁶ N. Sol-Foulon, A. Moris, C. Nobile, C. Boccaccio, A. Engering, J.-P. Abastado, J.-M. Heard, Y. Van Kooyk, O. Schwartz *Immunity* **2002**, *16*, 145-155.

- ³⁷ G. Simmons, J. D. Reeves, C. C. Grogan, L. H. Vandenberghe, F. Baribaud, J. C. Whitbeck, E. Burke, M. J. Buchmeier, E. J. Soilleux, J. L. Riley, R. W. Doms, P. Bates, S. Pöhlmann *Virology* **2003**, *305*, 115-123.
- ³⁸ C. P. Alvarez, F. Lasala, J. Carrillo, O. Muñiz, A. L. Corbí, R. Delgado *J. Virol.* **2002**, *76*, 6841-6844.
- ³⁹ F. Halary, A. Amara, H. Lortat-Jacob, M. Messerle, T. Delaunay, C. Houlès, F. Fieschi, F. Arenzana-Seisdedos, J.-F. Moreau, J. Déchanet-Merville *Immunity* **2002**, *17*, 653-664.
- ⁴⁰ J. P. Gardner, R. J. Durso, R. R. Arrigale, G. P. Donovan, P. J. Maddon, T. Dragic, W. C. Olson *Proc. Nat. Acad. Sci. U.S.A.* **2003**, *100*, 4498-4503.
- ⁴¹ S. Pöhlmann, J. Zhang, F. Baribaud, Z. Chen, G. J. Leslie, G. Lin, A. Granelli-Piperno, R. W. Doms, C. M. Rice, J. A. McKeating *J. Virol.* **2003**, *77*, 4070-4080.
- ⁴² P.-Y. Lozach, H. Lortat-Jacob, A. de Lacroix de Lavalette, I. Staropoli, S. Foug, A. Amara, C. Houlès, F. Fieschi, O. Schwartz, J.-L. Virelizier, F. Arenzana-Seisdedos, R. Altmeyer *J. Biol. Chem.* **2003**, *278*, 20358-20366.
- ⁴³ E. Navarro-Sanchez, R. Altmeyer, A. Amara, O. Schwartz, F. Fieschi, J.-L. Virelizier, F. Arenzana-Seisdedos, P. Despres *EMBO Rep.* **2003**, *4*, 723-728.
- ⁴⁴ B. Tassaneeritthep, T. H. Burgess, A. Granelli-Piperno, C. Trumpfheller, J. Finke, W. Sun, M. A. Eller, K. Pattanapanyasat, S. Sarasombath, D. L. Birx, R. M. Steinman, S. Schlesinger, M. A. Marovich *J. Exp. Med.* **2003**, *197*, 823-829.
- ⁴⁵ B. J. Appelmek, I. van Die, S. J. van Vliet, C. M. J. E. Vandembroucke-Grauls, T. B. H. Geijtenbeek, Y. van Kooyk *J. Immunol.* **2003**, *170*, 1635-1639.
- ⁴⁶ N. Maeda, J. Nigou, J.-L. Herrmann, M. Jackson, A. Amara, P. H. Lagrange, G. Puzo, B. Gicquel, O. Neyrolles *J. Biol. Chem.* **2003**, *278*, 5513-5516.
- ⁴⁷ M. Colmenares, A. Puig-Kroger, O. M. Pello, A. L. Corbi, L. Rivas *J. Biol. Chem.* **2002**, *277*, 36766-36769.
- ⁴⁸ A. Cambi, K. Gijzen, I. J. M. de Vries, R. Torensma, B. Joosten, G. J. Adema, M. G. Netea, B.-J. Kullberg, L. Romani, C. G. Figdor *Eur. J. Immunol.* **2003**, *33*, 532-538.
- ⁴⁹ R. Gesuete, C. Storini, A. Fantin, M. Stravalaci, E. R. Zanier, F. Orsini, H. Vietsch, M. L. M. Manesse, B. Ziere, M. Gobbi, M.-G. De Simoni *Ann. Neurol.* **2009**, *66*, 332-342.
- ⁵⁰ A. Cervera, A. M. Planas, C. Justicia, X. Urrea, J. C. Jensenius, F. Torres, F. Lozano, A. Chamorro *PLoS ONE* **2010**, *2*, e8433.
- ⁵¹ F. Orsini, P. Villa, S. Parrella, R. Zangari, E. R. Zanier, R. Gesuete, M. Stravalaci, S. Fumagalli, R. Ottria, J. J. Reina, A. Paladini, E. Micotti, R. Ribeiro-Viana, J. Rojo, V. I. Pavlov, G. L. Stahl, A. Bernardi, M. Gobbi, M.-G. De Simoni *Circulation* **2012**, *126*, 1484-1494.
- ⁵² J. M. Kel, J. Oldenampsen, M. E. Luca, J. W. Drijfhout, F. Koning, L. Nagelkerken *Am. J. Pathol.* **2007**, *170*, 272-280.
- ⁵³ M. E. Luca, J. M. Kel, Junda M.; W. van Rijs, J. Wouter Drijfhout, F. Koning, L. Nagelkerken *J. Neuroimmunol.* **2005**, *160*, 178-187.
- ⁵⁴ N. A. Barrett, A. Maekawa, O. M. Rahman, K. F. Austen, Y. Kanaoka *J. Immunol.* **2009**, *182*, 1119-1128.
- ⁵⁵ S. Hakomori *Proc. Nat. Acad. Sci. U.S.A.* **2002**, *99*, 10231-10233.
- ⁵⁶ C. A. Aarnoudse, J. J. G. Vallejo, E. Saeland, Y. van Kooyk *Curr. Opin. Immunol.* **2006**, *18*, 105-111.
- ⁵⁷ E. Saeland, S. J. van Vliet, M. Baeckstroem, V. C. M. van den Berg, T. B. H. Geijtenbeek, G. A. Meijer, Y. van Kooyk *Cancer Immunol. Immunother.* **2007**, *56*, 1225-1236.
- ⁵⁸ J. Idoyaga, C. Cheong, K. Suda, N. Suda, J. Y. Kim, H. Lee, C. G. Park, R. M. Steinman *J. Immunol.* **2008**, *180*, 3647-3650.
- ⁵⁹ S. Sattin, A. Bernardi *Trends Biotechnol.* **2016**, *34*, 483-495.
- ⁶⁰ J. Aretz, E.-C. Wamhoff, J. Hanske, D. Heymann, C. Rademacher *Front. Immunol.* **2014**, *5*, 1-12.
- ⁶¹ A. Volkamer, A. Griewel, T. Grombacher, M. Rarey *J. Chem. Inf. Model.* **2010**, *50*, 2041-2052.
- ⁶² Selected reviews: a) A. Bernardi, P. Cheshev *Chem. Eur. J.* **2008**, *14*, 7434-7441; b) B. Ernst, J. L. Magnani *Nat. Rev. Drug Discov.* **2009**, *8*, 661-677; c) S. Cecioni, A. Imberty, S. Vidal *Chem. Rev.* **2015**, *115*, 525-561.
- ⁶³ M. J. Borrok, L. L. Kiessling *J. Am. Chem. Soc.* **2007**, *129*, 12780-12785.
- ⁶⁴ S. L. Mangold, L. R. Prost, L. L. Kiessling *Chem. Sci.* **2012**, *3*, 772-777.
- ⁶⁵ G. Timpano, G. Tabarani, M. Anderlueh, D. Invernizzi, F. Vasile, D. Potenza, P. M. Nieto, J. Rojo, F. Fieschi, A. Bernardi *ChemBioChem* **2008**, *9*, 1921-1930.
- ⁶⁶ M. Andreini, D. Doknic, I. Sutkeviciute, J. J. Reina, J. Duan, E. Chabrol, M. Thepaut, E. Moroni, F. Doro, L. Belvisi, J. Weiser, J. Rojo, F. Fieschi, A. Bernardi *Org. Biomol. Chem.* **2011**, *9*, 5778-5786.
- ⁶⁷ J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martínez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo, A. Bernardi *ChemMedChem* **2007**, *2*, 1030-1036.
- ⁶⁸ S. Sattin, A. Daghetti, M. Thépaut, A. Berzi, M. Sánchez-Navarro, G. Tabarani, J. Rojo, F. Fieschi, M. Clerici, A. Bernardi *ACS Chem. Biol.* **2010**, *5*, 301-312.
- ⁶⁹ S. Mari, I. Sanchez-Medina, P. Mereghetti, L. Belvisi, J. Jimenez-Barbero, A. Bernardi *Carbohydr. Res.* **2007**, *342*, 1859-1868.
- ⁷⁰ I. Sutkeviciute, M. Thépaut, S. Sattin, A. Berzi, J. McGeagh, S. Grudinin, J. Weiser, A. Le Roy, J. J. Reina, J. Rojo, M. Clerici, A. Bernardi, C. Ebel, F. Fieschi *ACS Chem. Biol.* **2014**, *9*, 1377-1385.

- ⁷¹ N. Obermajer, S. Sattin, C. Colombo, M. Bruno, U. Svajger, M. Anderluh, A. Bernardi *Mol. Diversity* **2011**, *15*, 347-360.
- ⁷² N. Varga, I. Sutkeviciute, C. Guzzi, J. McGeagh, I. Petit-Haertlein, S. Gugliotta, J. Weiser, J. Angulo, F. Fieschi, A. Bernardi *Chem. Eur. J.* **2013**, *19*, 4786-4797.
- ⁷³ N. Varga, I. Sutkeviciute, R. Ribeiro-Viana, A. Berzi, R. Ramdasi, A. Daggetti, G. Vettoretti, A. Amara, M. Clerici, J. Rojo, F. Fieschi, A. Bernardi *Biomaterials* **2014**, *35*, 4175-4184.
- ⁷⁴ S. Ordanini, N. Varga, V. Porkolab, M. Thépaut, L. Belvisi, A. Bertaglia, A. Palmioli, A. Berzi, D. Trabattoni, M. Clerici, F. Fieschi, A. Bernardi *Chem. Commun.* **2015**, *51*, 3816-3819.
- ⁷⁵ T. Tomasic, D. Hajsek, U. Svajger, J. Luzar, N. Obermajer, I. Petit-Haertlein, F. Fieschi, M. Anderluh *Eur. J. Med. Chem.* **2014**, *75*, 308-326.
- ⁷⁶ L. Dehuyser, E. Schaeffer, O. Chaloin, C. G. Mueller, R. Baati, A. Wagner *Bioconjugate Chem.* **2012**, *23*, 1731-1739.
- ⁷⁷ K. C. A. Garber, K. Wangkanont, E. E. Carlson, L. L. Kiessling *Chem. Commun.* **2010**, *46*, 6747-6749
- ⁷⁸ T. Weber, V. Chandrasekaran, I. Stamer, M. B. Thygesen, A. Terfort, T. K. Lindhorst *Angew. Chem. Int. Ed.* **2014**, *53*, 14583-14586.
- ⁷⁹ Y. Zhang, Q. Li, L. G. Rodriguez, J. C. Gildersleeve *J. Am. Chem. Soc.* **2010**, *132*, 9653-9662.
- ⁸⁰ C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Kokschi, J. Dervede, C. Graf, E.-W. Knapp, R. Haag *Angew. Chem. Int. Ed.* **2012**, *51*, 10472-10498.
- ⁸¹ M. Mammen, S.-K. Choi, G. M. Whitesides *Angew. Chem. Int. Ed.* **1998**, *37*, 2754-2794.
- ⁸² R. J. Pieters *Org. Biomol. Chem.* **2009**, *7*, 2013-2025.
- ⁸³ V. Wittmann, R. J. Pieters *Chem. Soc. Rev.* **2013**, *42*, 4492-4503.
- ⁸⁴ T. R. Branson, T. E. McAllister, J. Garcia-Hartjes, M. A. Fascione, J. F. Ross, S. L. Warriner, T. Wennekes, H. Zuilhof, W. B. Turnbull *Angew. Chem. Int. Ed.* **2014**, *53*, 8323-8327.
- ⁸⁵ T. K. Dam, C. F. Brewer *Biochemistry* **2008**, *47*, 8470-8476.
- ⁸⁶ C. W. Cairo, J. E. Gestwicki, M. Kanai, L. L. Kiessling *J. Am. Chem. Soc.* **2002**, *124*, 1615-1619.
- ⁸⁷ C. F. Brewer *Trends Glycosci. Glycotechnol.* **1997**, *9*, 155-178.
- ⁸⁸ B. N. Stillman, D. K. Hsu, M. Pang, C. F. Brewer, P. Johnson, F.-T. Liu, L. G. Baum *J. Immunol.* **2006**, *176*, 778-789.
- ⁸⁹ C. A. Hunter, H. L. Anderson *Angew. Chem. Int. Ed.* **2009**, *48*, 7488-7499.
- ⁹⁰ G. Ercolani, L. Schiaffino *Angew. Chem. Int. Ed.* **2011**, *50*, 1762-1768.
- ⁹¹ S. Zhang, Q. Xiao, S. E. Sherman, A. Muncan, A. D. M. Ramos Vicente, Z. Wang, D. A. Hammer, D. Williams, Y. Chen, D. J. Pochan, S. Vértesy, S. André, M. L. Klein, H.-J. Gabius, V. Percec *J. Am. Chem. Soc.* **2015**, *137*, 13334-13344.
- ⁹² Y. Liu, Y. Zhang, Z. Wang, J. Wang, K. Wei, G. Chen, M. Jiang *J. Am. Chem. Soc.* **2016**, *138*, 12387-12394.
- ⁹³ P. Bojarová, V. Křen *Biomater. Sci.* **2016**, *4*, 1142-1160.
- ⁹⁴ I. Nierengarten, J.-F. Nierengarten *Chem. Asian J.* **2014**, *9*, 1436-1444.
- ⁹⁵ J. L. Jiménez Blanco, C. Ortiz Mellet, J. M. García Fernández *Chem. Soc. Rev.* **2013**, *42*, 4518-4531.
- ⁹⁶ J. E. Hudak, C. R. Bertozzi, *Chem. Biol.* **2014**, *21*, 16-37.
- ⁹⁷ A. Bernardi, J. Jiménez-Barbero, A. Casnati, C. De Castro, T. Darbre, F. Fieschi, J. Finne, H. Funken, K.-E. Jaeger, M. Lahmann, T. K. Lindhorst, M. Marradi, P. Messner, A. Molinaro, P. V. Murphy, C. Nativi, S. Oscarson, S. Penadés, F. Peri, R. J. Pieters, O. Renaudet, J.-L. Reymond, B. Richichi, J. Rojo, F. Sansone, C. Schäffer, W. B. Turnbull, T. Velasco-Torrijos, S. Vidal, S. Vincent, T. Wennekes, H. Zuilhofxy, A. Imberty *Chem. Soc. Rev.* **2013**, *42*, 4709-4727.
- ⁹⁸ R. Roy *Trends Glycosci. Glycotechnol.* **2003**, *15*, 291-310.
- ⁹⁹ C. R. Becer, M. I. Gibson, J. Geng, R. Ilyas, R. Wallis, D. A. Mitchell, D. M. Haddleton *J. Am. Chem. Soc.* **2010**, *132*, 15130-15132.
- ¹⁰⁰ Q. Zhang, J. Collins, A. Anastasaki, R. Wallis, D. A. Mitchell, C. R. Becer, D. M. Haddleton *Angew. Chem. Int. Ed.* **2013**, *52*, 4435-4439.
- ¹⁰¹ K. Lin, A. M. Kasko *Biomacromolecules* **2013**, *14*, 350-357.
- ¹⁰² K. Lin, A. M. Kasko *Bioconjugate Chem.* **2015**, *26*, 1504-1512.
- ¹⁰³ J. Geng, G. Mantovani, L. Tao, J. Nicolas, G. Chen, R. Wallis, D. A. Mitchell, B. R. G. Johnson, S. D. Evans, D. M. Haddleton *J. Am. Chem. Soc.* **2007**, *129*, 15156-15163.
- ¹⁰⁴ R. Ribeiro-Viana, M. Sánchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo, B. G. Davis *Nat. Commun.* **2012**, *2302/1-2302/8*.
- ¹⁰⁵ E. Schaeffer, L. Dehuyser, D. Sigwalt, V. Flacher, S. Bernacchi, O. Chaloin, J.-S. Remy, C.G. Mueller, R. Baati, A. Wagner *Bioconjugate Chem.* **2013**, *24*, 1813-1823.
- ¹⁰⁶ O. Martínez-Ávila, K. Hijazi, M. Marradi, C. Clavel, C. Champion, C. Kelly, S. Penadés *Chem. Eur. J.* **2009**, *15*, 9874-9888.
- ¹⁰⁷ O. Martínez-Ávila, L. M. Bedoya, M. Marradi, C. Clavel, J. Alcamí, S. Penadés *ChemBioChem* **2009**, *10*, 1806-1809.
- ¹⁰⁸ D. Arosio, F. Chiodo, J. J. Reina, M. Marelli, S. Penadés, Y. van Kooyk, J. J. Garcia-Vallejo, A. Bernardi *Bioconjugate Chem.* **2014**, *25*, 2244-2251.
- ¹⁰⁹ J.-F. Nierengarten, J. Iehl, V. Oerthel, M. Holler, B. M. Illescas, A. Muñoz, N. Martín, J. Rojo, M. Sánchez-Navarro, S. Cecioni, S. Vidal, K. Buffet, M. Durkae, S. P. Vincent *Chem. Commun.* **2010**, *46*, 3860-3862.

¹¹⁰ J. Luczkowiak, A. Muñoz, M. Sánchez-Navarro, R. Ribeiro-Viana, A. Ginieis, B. M. Illescas, N. Martín, R. Delgado, J. Rojo *Biomacromolecules* **2013**, *14*, 431-437.

¹¹¹ J. Luczkowiak, S. Sattin, I. Sutkevičiūtė, J. J. Reina, M. Sánchez-Navarro, M. Thépaut, L. Martínez-Prats, A. Daggetti, F. Fieschi, R. Delgado, A. Bernardi, J. Rojo *Bioconjugate Chem.* **2011**, *22*, 1354-1365.

¹¹² Y. M. Chabre, R. Roy *Chem. Soc. Rev.* **2013**, *42*, 4657-4708.

¹¹³ M. Dubber, O. Sperling, T. K. Lindhorst *Org. Biomol. Chem.* **2006**, *4*, 3901-3912.

CHAPTER 1

MBL Ligands As Neuroprotective Agents In Stroke

1.1 Mannose Binding Lectin and its role in the complement system

The Mannose Binding Lectin (MBL) is a circulating soluble C-type lectin that plays a crucial role in innate immunity. Indeed, MBL is fundamental for initiation of the complement system, which guarantees a front line protection against infecting pathogens and self-altered structures.^{1,2}

The complement system machinery is the evolution result of an ancestral defence system. It comprises more than 30 plasma and membrane proteins and allows for protection in the lag period that is required for the adaptive immune system to develop a specific immunological response. The complement system fulfils its functions through three main pathways, the classical pathway, the lectin pathway and the alternative pathway that are interconnected by a complex signalling network.¹ For the sake of simplicity, herein we will focus just on the lectin pathway and particularly on its initiation mediated by MBL (Fig. 1.1).

Beyond the classical pathway, the existence of the lectin pathway was demonstrated only after the discovery of MBL, which was identified as the solely activating pattern recognition receptor. Nowadays, we know that the situation is far more complex and at least six different PRRs were recognized as initiators of the lectin pathway: MBL, collectin-10, collectin-11, ficolin-1, ficolin-2 and ficolin-3. The PRRs, upon initial selective recognition of PAMPs and DAMPs motives, trigger lectin pathway through autoactivation of associated proteases, called MBL-associated serine proteases (MASPs). To date, three MASPs have been identified, MASP-1, MASP-2 and MASP-3 that can be found associated in initiation complexes to two non-catalytic proteins, MAp19 and MAp44. The regulatory mechanisms governing the formation of the initiation complex and the roles and mode of actions of recruited proteins are still unclear and object of investigation.³ Nevertheless, it is proven that activation of MASPs finally triggers a proteolytic cascade, in which factors C4 and C2 are initially cleaved with the formation of convertase C4b2a that activates C3 into the C5 convertase C4b2a3b, that cleaves C5 into the anaphylatoxins C5a and C5b eventually leading to the formation of the membrane attack complex (MAC).⁴

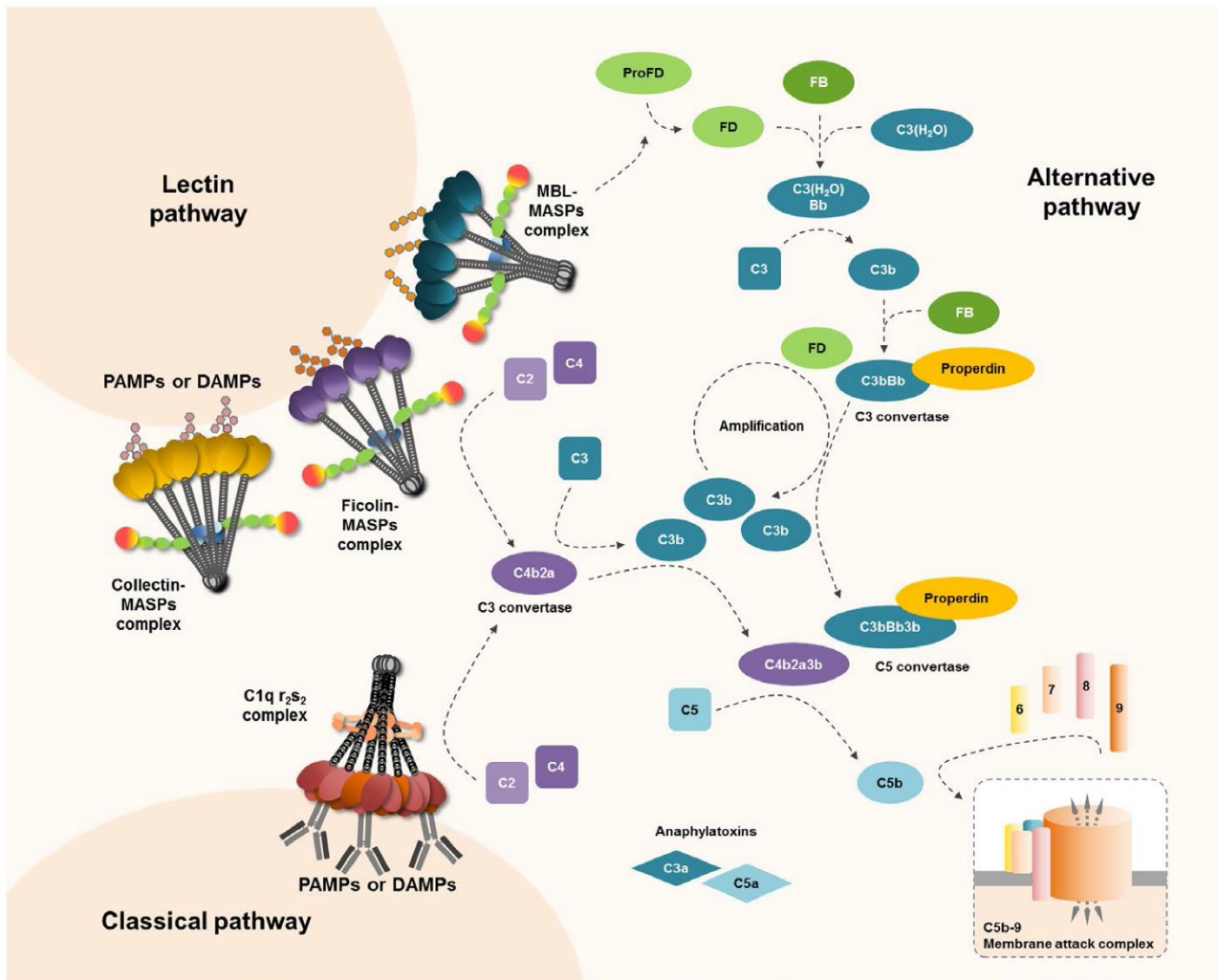


Figure 1.1 - The complement system pathway. So far, three pathways have been described that accounts for complement system functions. Beside the classical pathway, engaging adaptive immunity to form complex C1qr₂s₂, the lectin pathway has been identified MBL, ficolin-1, ficolin-2, ficolin-3, collectin-10, collectin-11 are all possible initiators of the signalling, which is triggered upon binding to PAMPs or DAMPs exposed by cells. Activation of MASP results in the cleavage of both C2 and C4 factors that combine together to form the C3 convertase C4b2a. C3 conversion finally leads to conversion of C5 into anaphylatoxins C5a and C5b promoting the formation of the MAC.⁴

Recent insights established MAC formation to take place with the assembly of C5b protein with C6, C7, and C8 defining a structure that is elongated with 18 copies of C9, therefore resulting in an asymmetric barrel shape complex (Fig. 1.2).^{5,6,7} For long time, the membrane attack complex was supposed to act as a simple pore able to insert into the phospholipidic membrane of target cells and to cause lysis due to uncontrolled osmotic gradient. However, recent studies suggests that precise and fine regulations of the process must be operative, thus further highlighting the complexity of the complement system defence response.⁴

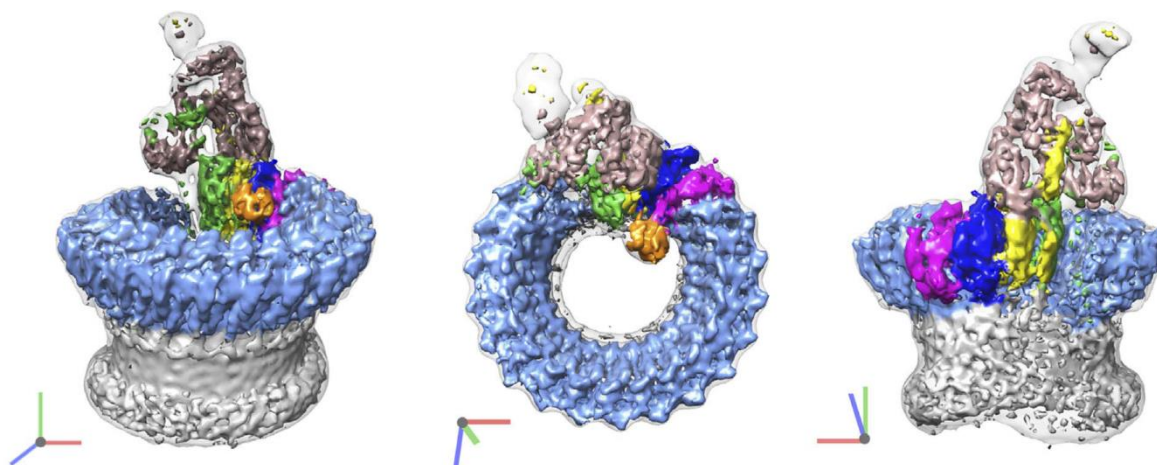


Figure 1.2 – The MAC structure. Sharpened map coloured according to complement proteins: C5b (tan), C6 (green), C7 (yellow), C8a (magenta), C8b (dark blue), C8g (orange) and C9 (light blue). Apart from C5b, all proteins of the MAC complex possess the cholesterol-dependent cytolysin/MAC/perforin-like (CDC/MACPF) domain, which associate into a β barrel determining pore formation (shown in grey). Axes are shown.⁶

MBL is a homotrimeric, soluble C-type lectin belonging to the collectin family that is mainly expressed in hepatocytes and secreted as a soluble protein in the serum, where is normally found in 1 $\mu\text{g}/\text{mL}$ concentration in humans.⁸ In animals, two gene sequences encoding for MBL can be identified on the same chromosome in close proximity, suggesting that the two isoforms arise from gene duplication of an ancestral gene.⁹ In rodents two MBL isoforms originate from the expression of *Mbla* and *Mblc*. In primates the corresponding ortholog gene are present, *MBL1P1* and *MBL2*, but in higher primates only the latter is functional, while *MBL1P1* is found as a pseudogene.¹⁰

The functional structure of the protein is a homotrimer, formed by three subunits in which four different domains can be recognized: an N-terminal cysteine rich domain, with three cysteine residues responsible for intra and inter-chain disulphide bonds formation; the characteristic collagen like domain; a hydrophobic neck domain; a C-terminal C-type lectin-like domain. Subunits association process may be guided by interactions between the collagen regions, however experiments with truncated recombinant MBL showed that the presence of the neck region alone linked to the CTLD is sufficient to give, in solution and in crystals, a homo trimeric complex.^{11,12} The assembly of three subunits is strongly favoured and stabilised by interactions between the hydrophobic neck regions leading to formation of a α -helical coiled-coil motif. The resulting structure is depicted in Figure 1.3, where the three CTLDs clearly show a flat and trigonal arrangement, with calcium binding sites Ca2 (Ca2 is the calcium binding site involved in sugar recognition and binding) averagely spaced by 45 Å.

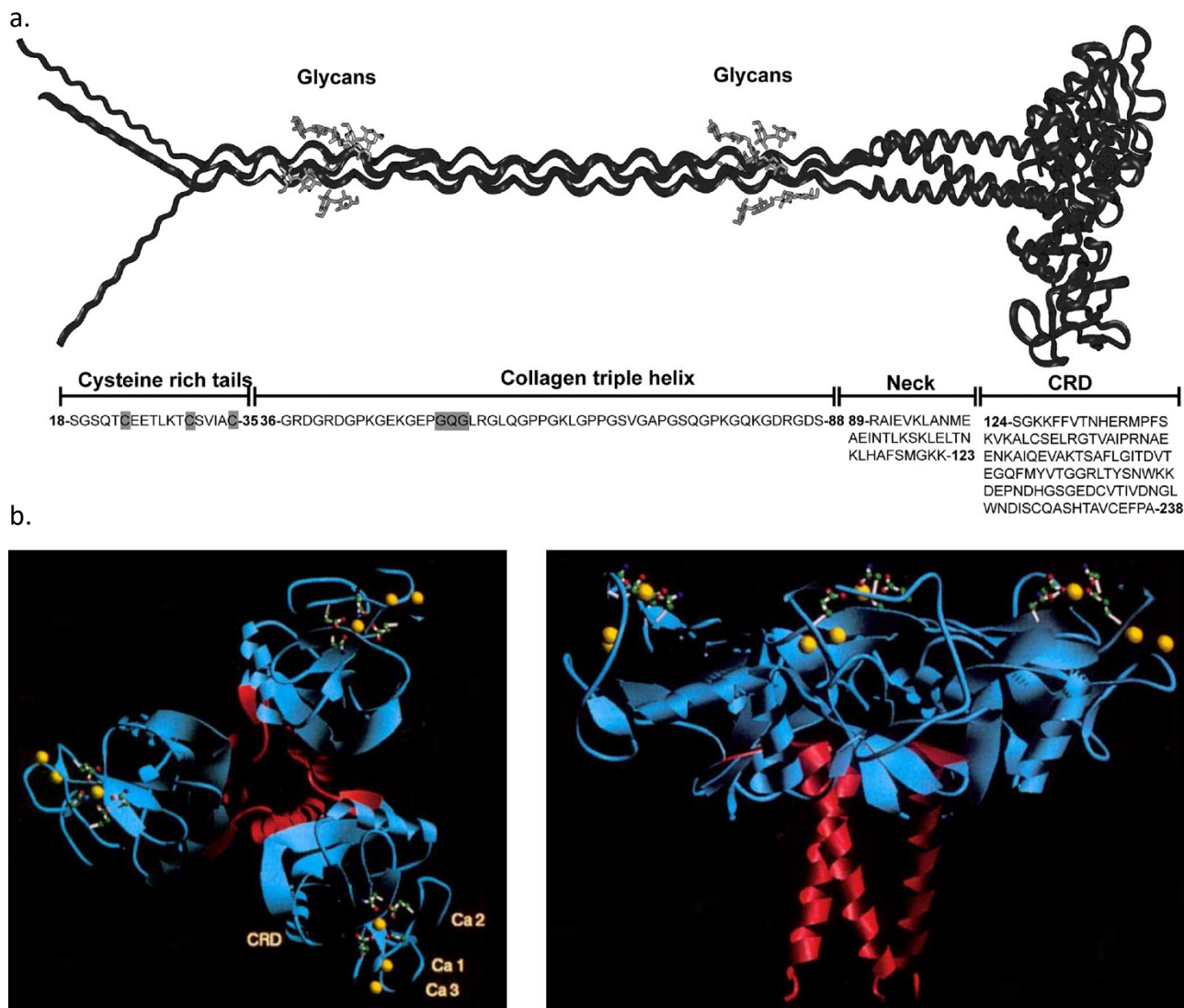


Figure 1.3 – MBL structure. a) Model structure of rat MBL homotrimer. The model was created by connection of a collagen region to crystal structures of the CRDs and the neck regions. The model was completed by addition of three *N*-terminal cysteine-rich peptides at the *N*-terminus of collagen regions and by the addition of four glycosylgalactosyl disaccharides to residues Lys-44, Lys-47, Lys-79, and Lys-82 in each of the three collagen chains. At the bottom, the rat MBL sequence is aligned with the four regions of MBL.¹³ b) Crystal structure of truncated human MBL. The structure of MBL homotrimer lacking the collagen region is shown. The three CRDs displayed in a trigonal arrangement are shown in blue. Ca²⁺ ions are shown as yellow spheres, while the α -helices of the three neck domains are shown in red.¹²

The trivalent nature of MBL allows for multivalent protein-carbohydrate interactions, thus enhancing binding affinity. Notably, MBL homotrimers (thus displaying 3 CRDs) in solution assemble together forming higher oligomers, ranging from dimers up to octamers (6-24 CRDs; 3 CRDs per MBL homotrimer). In particular in mice MBL is present in serum ranging from monomers up to tetramers, whereas in humans, the main fraction is constituted by trimers and tetramers, with lower and higher forms occurring, from dimers up to hexamers.^{14,15} Some evidences suggests that higher MBL oligomers may increase complement activation,¹⁶ hence revealing oligomerization process as a plausible regulation level in lectin pathway signalling. However, the study of oligomerization mechanisms and oligomer structures characterization are frustrated by the multidomain structure of the protein, which hampers X-ray crystallography determinations. Recently, a combination of X-ray scattering analysis and analytical ultracentrifugation data allowed for the determination of murine MBL dimers, trimers and tetramers, which showed a common near-planar V-shaped structure, with CTLDs separated on average by 20 nm (within the oligomers) (Fig. 1.4).¹³

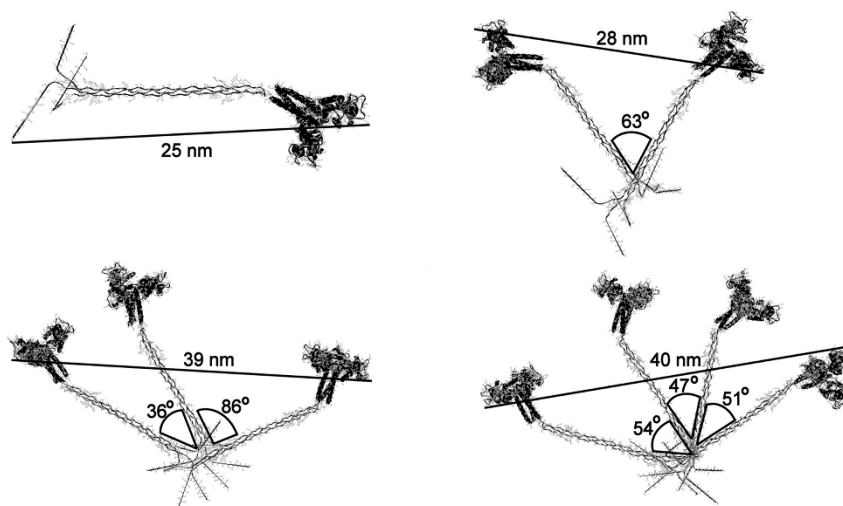


Figure 1.4 – Structures of murine MBL oligomers. MBL homotrimers are known to associate together in solution forming aggregates. The structures for the dimer, trimer and tetramer of trimeric MBL have been calculated from X-ray scattering analysis and analytical ultracentrifugation data. The structure of MBL homotrimer is depicted along with the best-fit structures for the corresponding dimer, trimer and tetramer. The approximate angle between the two collagen helices is also shown.¹³

1.2 MBL in health and disease

As a C-type lectin, MBL selectively binds to D-mannose, *N*-acetyl-D-glucosamine, D-glucose and L-fucose containing polysaccharides that are exposed either by invading pathogens, pathogen associated molecular patterns (PAMPs), and by self-altered cells, damage associated molecular patterns (DAMPs). As a consequence of binding, different responses can be triggered. Activation of complement is the most investigated pathway, but MBL is also able to promote apoptosis, to modulate inflammation and to induce opsonophagocytosis, even if it is not clear whether MBL directly acts as an opsonin or if phagocytosis is stimulated through initiation of other pathways.

Recognition of pathogens and activation of defence mechanisms is a complex event that is orchestrated in a cooperative fashion by several PRRs. Among these, MBL is able to recognize and target bacteria, viruses and fungi. Both the surface lipoarabinomannan glycan of *Mycobacterium avium*¹⁷ and peptidoglycan and lipoteichoic acid from *Staphylococcus aureus*^{18,19} are natural MBL ligands. A study revealed that MBL-null mice inoculated with *S. aureus* are far more susceptible to infections compared to wild type mice. Notably, an inversion of the phenotype was obtained treating MBL-null mice with recombinant human MBL (rhMBL) prior to inoculation, thus showing the important role of MBL in contrasting the infection.²⁰

MBL easily detects high-mannose glycans exposed on viral proteins. Studies have shown that MBL binds to influenza A virus inhibiting hemagglutination.²¹ Moreover, a higher susceptibility to herpes simplex virus-2 infection was revealed in MBL-null mice.²²

Mannan oligosaccharide found as major component of fungal cell wall is another target for MBL. Indeed MBL is known to bind *Saccharomyces cerevisiae*, *Aspergillus fumigatus* and *Candida albicans*. Immune response against *C. albicans* infections are modulated by MBL, which is able to induce phagocytosis and respiratory burst and to stimulate monocytes to synthesise TNF- α .^{23,24}

1.3 MBL in ischemic injury - Background

Being part of the innate immune system, MBL plays an important role in the detection of self-altered cells. Nauta and coworkers reported that MBL recognition of apoptotic and necrotic cells induces activation of macrophages and phagocytosis allowing non-inflammatory clearance of dying cells.²⁵ Altered motives may

also be displayed by self damaged cells. MBL recognition of such DAMPs have proved to prompt complement activation, which aggravates the pathogenesis of ischemia and reperfusion injury. Indeed, MBL activity was found to exacerbate lesions in mice models of ischemia and reperfusion injury in heart,^{26,27,28,29} skeletal muscle,³⁰ gut^{31,32} and kidney.³³

In vitro experiments in mice models also revealed the deleterious effects of MBL in brain ischemic injury. Ischemic stroke occurs when the blood supply is severely reduced or completely blocked. The restoration of blood circulation is followed by the establishment of an inflammatory environment that worsen the lesion, which is referred to as reperfusion damage. In this context, it has been show that MBL senses damaged cells activating the complement cascade and contributing to inflammatory responses leading to recruitment of inflammatory cells, synthesis of inflammatory molecules, activation of phagocytosis, induction of cell death and endothelial damage. Particularly, MBL binding to endothelial cells generates cytotoxic damage and increased MBL deposition in endothelium has been observed after ischemic stroke. Hence, MBL targeting emerges as a promising therapeutic strategy to limit reperfusion injuries after stroke. Recent studies have demonstrated that recombinant human C1 inhibitor (rhC1-INH), an oligomannosylated 478 aminoacids protein, colocalizes with deposited MBL, suggesting a pivotal role in silencing complement activation, thus reducing inflammation responses in ischemic brain mice models.³⁴ Moreover, Lozano, Chamorro et al. observed reduced complement deposition and neutrophils recruitment in ischemic brain tissues of MBL-null mice, along with smaller infarction volume and better prognosis compared to wild type mice. Remarkably, the same study was also addressed to patients with acute stroke presenting single nucleotide polymorphism in the promoter and encoding region of *MBL2* gene. This genotype is characterised by a low level of serum circulating MBL and was revealed to be associated with 11 times more probability of positive outcome after stroke.³⁵

The disclosure of MBL as a potential therapeutic target in acute stroke is of great interest. Indeed, in 2013 stroke represented the fifth cause of death in the United States, with about 795 000 people per year experiencing a new or recurrent stroke (ischemic or haemorrhagic) and estimated to be responsible for \$33 billion of direct and indirect annual costs.³⁶ Ischemic stroke is caused by thrombotic or embolic occlusions and represents on his own the 87% cases of strokes. The only therapies available, including thrombolysis with tissue plasminogen activator, are limited in application due to the narrow therapeutic window (from 4.5 to 6 hours form stroke event) dictated by safety concerns. For this reason, the development of alternative treatment agents is really desirable.

Furthermore, the therapeutic relevance of MBL as promising drug target was corroborated by *in vivo* experiments. Blocking the most abundant mouse isoform, MBL-A, with anti-MBL-A monoclonal antibodies in a mouse model of brain ischemic injury mice models resulted in protecting effects with a very wide therapeutic window: mice treated with mAb up to 18 hours after stroke showed 58% reduction of overall ischemic volume by MRI analysis.³⁷

Following a similar approach, aiming to block MBL with synthetic antagonists, multivalent constructs have been investigated. Promising results were obtained for Polyman2 **37**, a glycodendron based on a tetravalent polyester scaffold functionalised with pseudo-trisaccharide moieties **6**. Indeed, *in vivo* experiments with transient middle cerebral artery occlusion (tMCAo) mice models, revealed remarkable protecting effect of **37**, that was able to reduce ischemic functional deficits and ischemic volume when given up to 24 hours after stroke (Fig. 1.5a). Maximum efficacy was achieved after 6 hours from ischemic event, with a 68% reduction of ischemic volume. Moreover, Polyman2 **37** treated mice resulted in MBL deposition abrogation in the endothelium (Fig. 1.5b), suggesting a plausible correlation between MBL deposition and tissue damage.³⁷

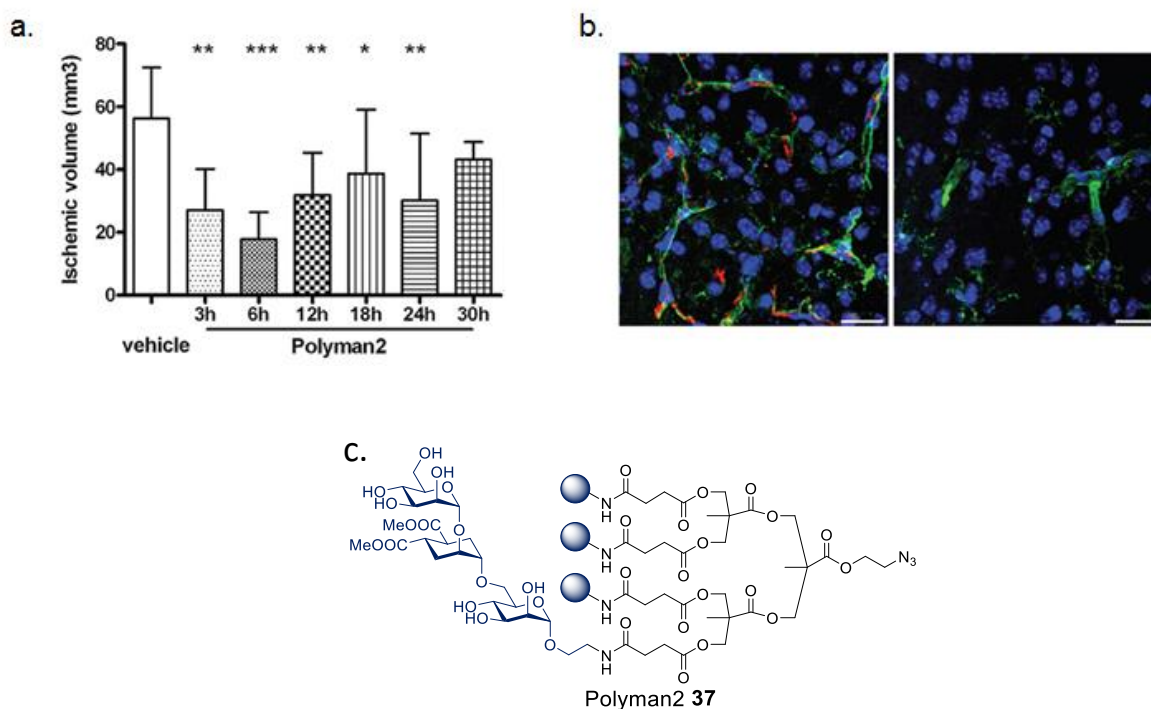


Figure 1.5 - *In vivo* studies of Polyman2 37 dendron protecting effects. **a)** Polyman2 37 was given 3, 6, 12, 18, 24, or 30 hours from transient middle cerebral artery occlusion (tMCAo) in mice. Infarct volume was assessed 48 hours from injury. Maximum reduction of ischemic volume was obtained when Polyman2 37 was given after 6 hours from injury. **b)** Representative images of MBL-A staining in the ipsilateral cortex 24 h after tMCAo in a vehicle treated mouse (left) and in a Polyman2 37 treated mouse (right). Polyman2 37 given 6 h after injury). Nuclei are stained in blue, blood vessels in green and MBL in red. **c)** Structure of tetraivalent Polyman2 37 dendron.³⁷

1.4 Manno-glycodendrimers as MBL antagonists in reperfusion injury

The binding ability of Polyman2 37 and its extremely wide therapeutic window make it a promising candidate for stroke treatment. In fact, Polyman2 37 possesses good solubility in water and minor cytotoxicity. However, 37 suffers from chemical instability ascribed to the unhindered succinyl ester bonds in the dendron scaffold (Fig. 1.5). The presence of these moieties makes the constructs really sensitive to nucleophiles, both under basic and mild acidic catalysis, impeding purification by common chromatographic techniques (both direct and reverse phase), preventing the scale up of the synthesis and hampering pharmacokinetic studies. Moreover, Polyman2 37 hydrolysis was observed even in water solution at physiological pH (7.4, PBS buffer), and quantified as 30% after 6 h by NMR analysis.³⁸ MS analysis of the hydrolysed product confirmed that the cleavage takes place at the succinyl ester bond, while the remaining hindered α,α -disubstituted ester bonds contained in the molecule remained untouched. For these reasons, the design and preparation of more stable and easily accessible glycodendrons able to perform as MBL ligands were sought after.

The nature of MBL targets in both physiological and pathological processes are still not clear, but X-ray structures of MBL complexes with mono and oligosaccharides have been determined.^{39,40,41} A rational design of synthetic multivalent ligands based on the structural data has never been attempted. However, by testing previously developed DC-SIGN pseudo-mannosylated ligands using Surface Plasmon Resonance (SPR),^{38,42,43} we demonstrated that they also perform as MBL antagonists.

SPR (Surface Plasmon Resonance) is a physical phenomenon that is nowadays commonly applied as a biophysical method for measuring kinetic aspects of interactions between macromolecules and their ligands. From a theoretical point of view, the SPR is the collective excitation of electrons of a metal surface that is observed upon irradiation and absorption of a polarized light. The surface plasmon resonance is maximum

when the angle of incidence of the radiation, Θ , reaches the value of the resonance angle, Θ_{res} . Resonance angle, Θ_{res} , is dependent from the refracting indexes of the materials crossed by the incident light, thus the polarizing glass prism and the metal surface (glass/metal). When $\Theta = \Theta_{\text{res}}$ the radiation absorption of electrons is maximum and, as a consequence, the intensity of the reflected light reaches a minimum.

SPR was revealed as a valuable tool for the determination of dissociation constants for small molecule-protein interactions. Devices based on this technique consists of a chip coated with a metal surface, usually gold, on which the protein of interest (or either the ligand) is immobilized. A solution of the ligand (or either the protein) is then flowed through channels over the gold surface and the extent of the ligand-protein interaction is measured irradiating with a polarized light. The refracting index of the glass prism is a constant, while the refracting index of the gold layer is sensitive to changes at its surface. Therefore, association and dissociation processes that takes place over the chip will result in a modulation of SPR signal, which can be detected as a variation of Θ_{res} measuring the intensity of reflected light (Fig. 1.6).

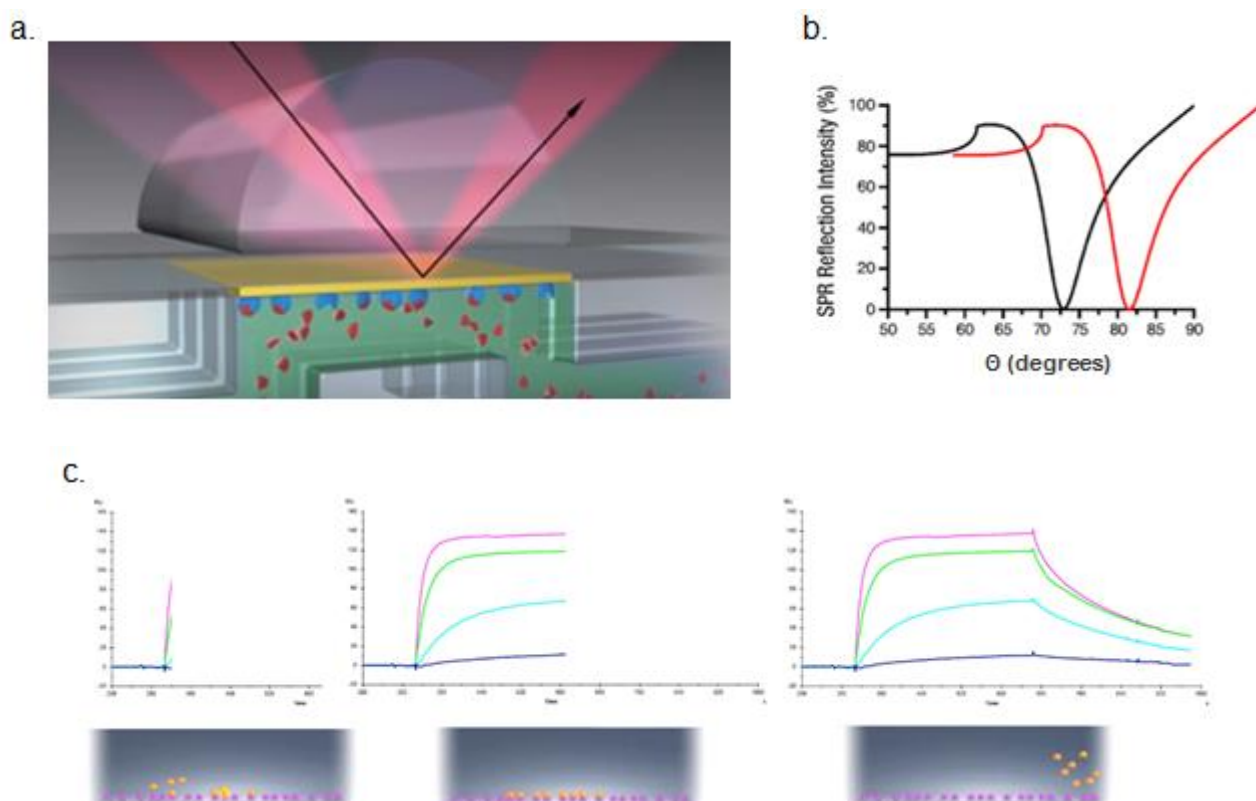
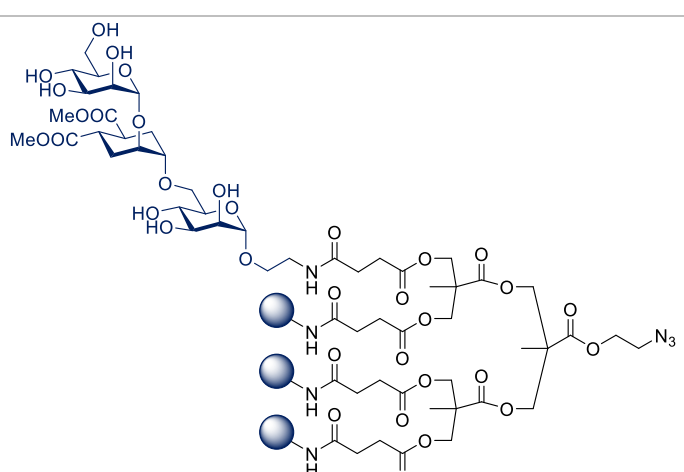
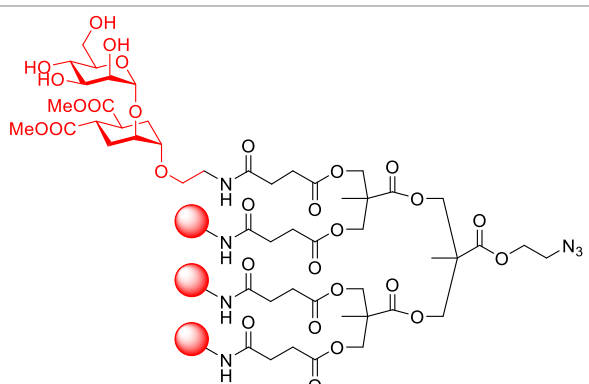
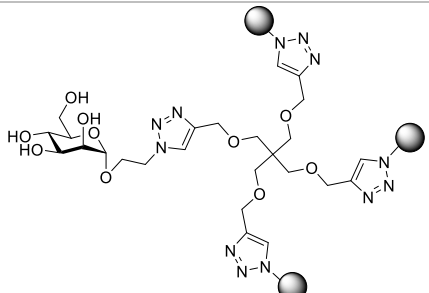


Figure 1.6 - SPR based analysis of carbohydrate-protein interactions. **a)** Schematic representation of an SPR based device. In the picture, the extent of the binding of ligands (in red) towards proteins (in blue), immobilized on a gold surface, is revealed irradiating with light polarized by a glass prism. **b)** The typical SPR signal is shown. The intensity of reflected radiation is monitored varying the incident angle Θ . When $\Theta = \Theta_{\text{res}}$ the plasmon resonance is maximum and the reflected radiation intensity reaches a minimum. Θ_{res} is characteristic for each surface and sensitive to modifications that are detected as a shift in Θ_{res} value. **c)** Association and dissociation at the gold surface of SPR devices are translated into a sensogram. Curves corresponding to solutions of different concentration of ligands are reported. As the ligand is binding, the signal rapidly increases until it reaches a maximum, forming a plateau. Washing out the ligand, the signal rapidly decreases. From the sensograms, the dissociation k_{off} and association k_{on} rate constants can be extrapolated and the corresponding dissociation constant K_D can be calculated.

For our screening of putative MBL antagonists, multivalent constructs with different geometries and valencies, bearing either simple mannose residues or the glycomimetic dimannosides **5** and trimannosides **6** were tested (Table 1.1), while the already investigated Polyman2 **37** was screened as reference compound. All

SPR analysis were performed by the Gobbi group at “Mario Negri - Istituto di Ricerche Farmacologiche” in Milan.

The Polyman2 **37** analogue, Polyman1 **36** was also included in the screening.⁴⁴ In fact, although this dendron presents the same drawbacks of Polyman2 **37**, it allows to evaluate the impact on MBL binding when pseudo trisaccharide **6** units are switched with the simplified structure of pseudo-disaccharide **5**. The more robust structures of Polyman5 **44**,⁴⁵ Polyman8 **45**,⁴² Polyman17 **46**,⁴² still belonging to the tetravalent series, were tested. These glycodendrons display respectively four copies of D-mannose and of glycomimetic ligands **5** and **6**. The relationship between the amount of displayed ligands and potency enhancement was evaluated screening the related hexavalent Polyman9 **47**⁴² and Polyman21 **48**⁴². Moreover, hexavalent compounds Polyman30 **49**,⁴³ Polyman32 **50**,⁴³ Polyman31 **42**⁴³ and Polyman42 **51**⁴⁶ with a rigid rod-like structure of different length were tested to study chelation of MBL binding sites. Finally, the nonavalent presentation of glycomimetic **5** was evaluated with Polyman 20 **52**.⁴²

Compound	Structure	Ligand/ Valency
Polyman2 ^a 37		Pseudo-trimannoside 6 / 4
Polyman1 ^b 36		Pseudo-dimannoside 5 / 4
Polyman5 ^c 44		D-Mannose/ 4

Polyman8 ^d 45		Pseudo-dimannoside 5 / 4
Polyman17 ^d 46		Pseudo-trimannoside 6 / 4
Polyman9 ^d 47		Pseudo-dimannoside 5 / 6
Polyman21 ^d 48		Pseudo-trimannoside 6 / 6
Polyman30 ^e 49		Pseudo-dimannoside 5 / 6
Polyman32 ^e 50		Pseudo-dimannoside 5 / 6

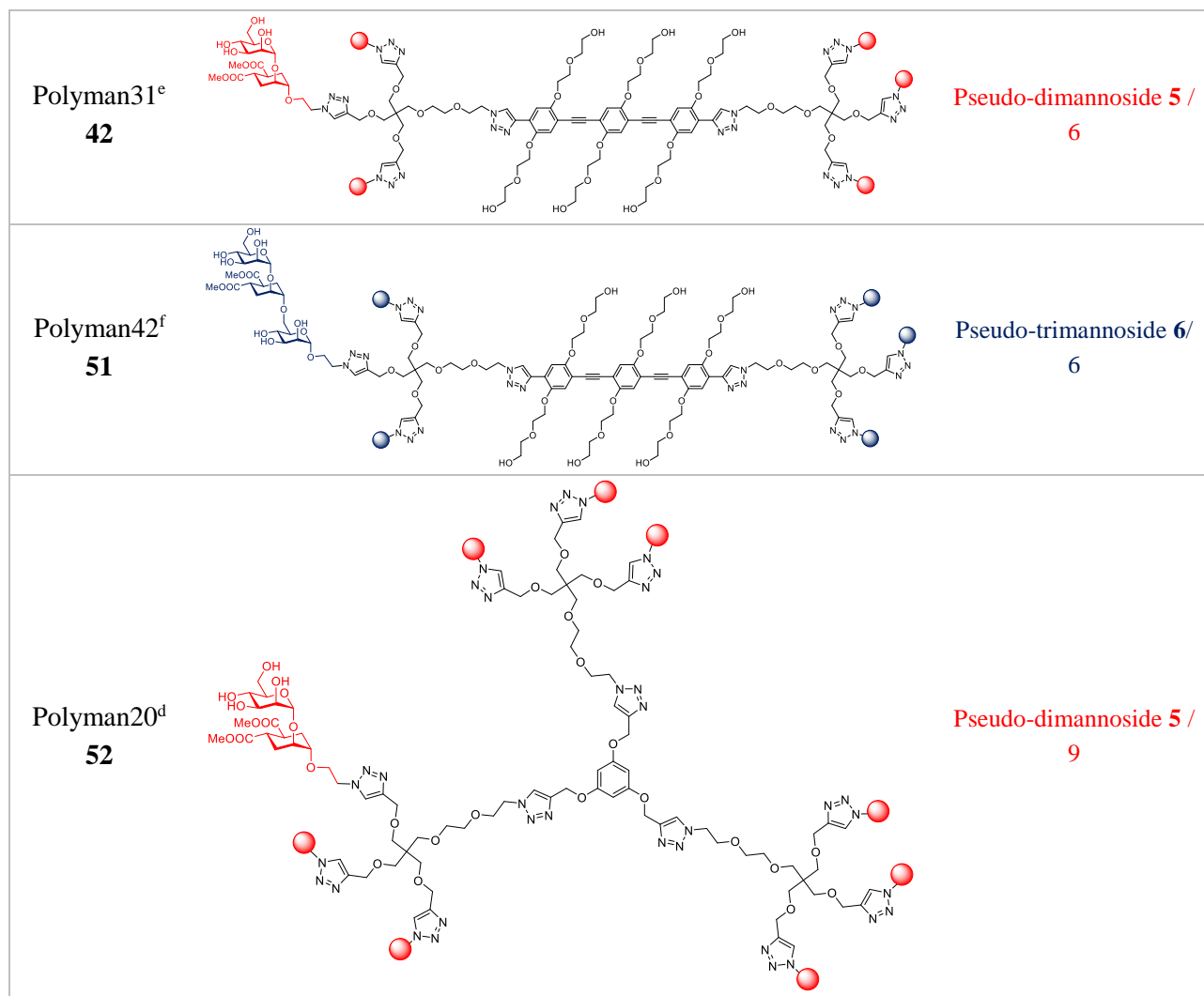


Table 1.1 - Glycodendrons and glycodendrimers as potential MBL antagonists. The multivalent constructs tested as MBL ligands are reported. Ligands deriving from D-mannose are depicted in black, from pseudo-dimannoside **5** in red and from pseudo-trimannoside **6** in blue. The corresponding valency is also shown. Names of the compounds are reported with a superscript specifying the literature reference for synthesis and characterisation: a) S. Sattin et al. *ACS Chem. Biol.* **2010**, 5, 301-312; b) J. Luczkowiak et al. *Bioconjugate Chem.* **2011**, 22, 1354-1365; c) M. Touaibia et al. *ChemMedChem* **2007**, 2, 1190-1201; d) N. Varga et al. *Biomaterials* **2014**, 35, 4175-4184. Preparation and characterisation of Polyman42 **51** is also described in the Experimental part of Chapter1; e) S. Ordanini et al. *Chem. Commun.* **2015**, 51, 3816-3819; f) Preparation and characterisation of Polyman42 **51** is described in the Experimental part of Chapter1.

Two preliminary screening campaigns to identify multivalent ligands for MBL were first conducted by non-competitive SPR assay. MBL was immobilized on the sensor chip, then solutions of putative multivalent ligands were flowed on the gold surface to measure the entity of the binding. As a first analysis, only single point measurements were registered.

Compounds belonging to the tetravalent series were tested in binding assay with the murine isoform MBL-A (Fig. 1.7a). The results suggests that binding ability of this multivalent constructs is enhanced when copies of both glycomimetics **5** and **6** are displayed instead of D-mannose residues, with **6** performing as the best. Moreover, the higher avidity of **36** and **37**, when compared to constructs displaying the corresponding monovalent ligands **45** and **46**, identifies Boltorn glycodendron scaffold as a platform with better ligand presentation for the arrangement of MBL-A binding sites. Polyman8 **45** and Polyman2 **37** were further investigated along with hexavalent and a nonavalent compounds, all displaying dimannoside monovalent

glycomimetic **5**, in interactions with either recombinant human MBL (RHmbL) and the two murine isoforms MBL-A, MBL-C (Fig. 1.7b). In some cases a certain degree of selectivity for the different protein forms was observed. For example, while no selectivity was shown by the hexavalent Polyman9 **47**, Polyman2 **37**, which bears four copies of the more potent trimannoside **6**, achieved binding for rhMBL but presented significantly decreased affinity for the murine isoforms. The highest binding affinity was exhibited by rigid rod containing compounds **50** and **42**, suggesting that chelation mechanisms may be operative with these extended scaffolds.

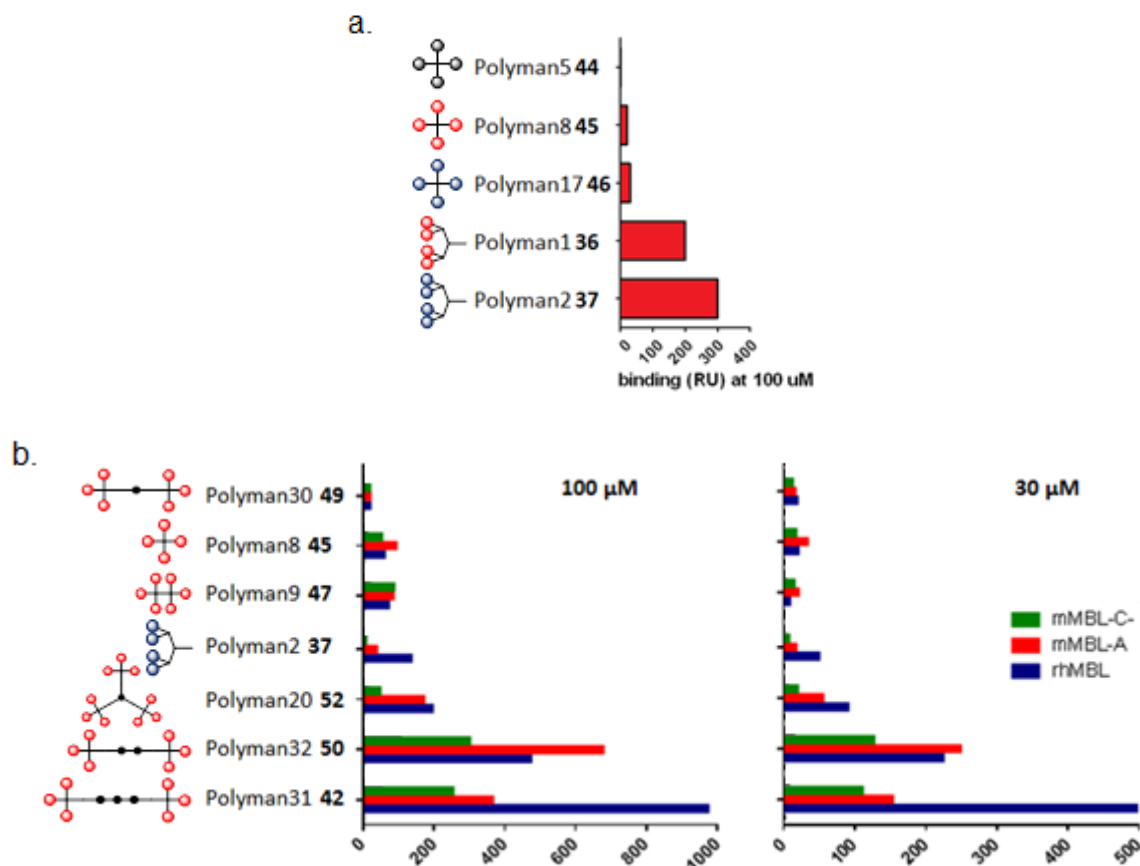


Figure 1.7 - SPR direct interaction experiments with MBL-A. **a)** Compounds in the tetravalent series were tested for direct interaction with the murine isoform MBL-A. MBL was immobilized on the surface of sensor chip, whereas the dendrimers were injected for 3 minutes at a flow rate of 100 μ L/min. The maximal binding for each dendrimer is expressed in resonance units, RU. Single point measurements were recorded. **b)** The graphics show results for the second screening campaign. Selected constructs were tested in direct interaction with both murine MBL isoforms MBL-A and MBL-C and with recombinant human MBL, rhMBL. Single point measurements were recorded immobilizing MBL isoforms over the sensor chip and flowing the multivalent ligands at two different concentrations (100 and 30 μ M).

These initial studies allowed the selection of compounds with the most promising geometries, therefore the affinities of Polyman20 **52**, Polyman31 **42**, Polyman9 **47** and Polyman2 **37** for rhMBL were further determined by a more accurate SPR inhibition assay. In this experiment, IC_{50} values of multivalent ligands can be calculated flowing the compounds at different concentrations together with rhMBL over the sensor chip, where they have to compete for binding against an immobilized polymannosylated BSA (Fig. 1.8). Additionally, the two constructs Polyman42 **51** and Polyman21 **48**, both functionalised with six copies of the trimannoside **6**, were also tested.

To contain the loss of materials and the costs associated with this kind of analysis, monovalent ligands were not screened and experiments were restricted to multivalent compounds. Thus, a rigorous calculation of the relative inhibition potency (β) was not possible but compounds affinity was evaluated as relative potency γ , taking the already investigated Polyman2 **37** as a reference. Relative potency weighted for the valency γ_w is

also reported (Table 1.2). Compared with the affinity of Polyman2 **37** a 3.5 fold higher potency (2.3 corrected per ligand) was achieved with the higher functionalized Polyman21 **48**, while the extended structure of Polyman42 **51** possibly allows chelation, resulting in a compound with 8.5 increased relative inhibition potency corrected per number of monovalent ligands. In order to achieve comparable inhibition potency with more accessible compounds, constructs bearing multiple copies of the synthetically less demanding ligand **5** were screened. Polyman9 **47** provided comparable affinity for rhMBL as Polyman2 **37** and approaching half the inhibition potency of the corresponding hexavalent Polyman21 **48**. An IC_{50} value of 60 μM was achieved by the chelating compound Polyman31 **42**, which is four time less potent than the dendron of analogous structure Polyman42 **51**, but reaches the potency showed by the non chelating Polyman21 **48**. Finally, further increased potency was achieved with the higher extent of functionalization of Polyman20 **52**, that showed an IC_{50} value of 21 μM corresponding to a γ_w value 4 times higher than both Polyman9 **47** and Polyman2 **37**. From this screening, Polyman31 **42** was chosen as the best compromise between synthetic accessibility, loading of monovalent ligand and potency and was therefore selected for further *in vitro* and *in vivo* investigations.

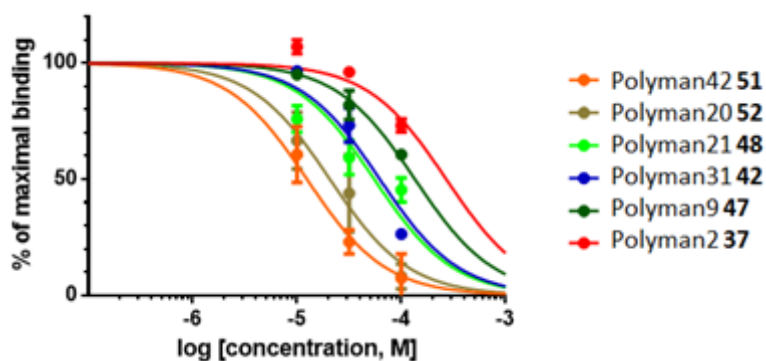


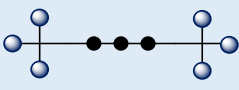



Figure 1.8 - SPR dose-response inhibition curves for multivalent compounds in interaction with rhMBL. The curves show the inhibition profiles for the multivalent compounds screened in the interaction with rhMBL. Multivalent ligands were preincubated with rhMBL (8 nM) at different concentrations (10, 30, 100 μM) and the solutions were flowed over an SPR sensor chip coated with previously immobilized mannosylated-BSA. The concentration dependence of SPR signal allows fitting of model curve to the experimental data for each multivalent ligand and determination of the corresponding IC_{50} value (data showed in Table 1.2).

Compound	IC_{50} (μM)	Relative potency γ	γ_w
Polyman2 37 	191	1	1
Polyman21 48 	55	3.5	2.3
Polyman42 51 	15	12.7	8.5
Polyman9 47 	126	1.5	1.0

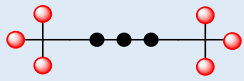
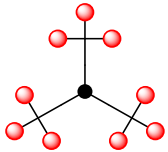
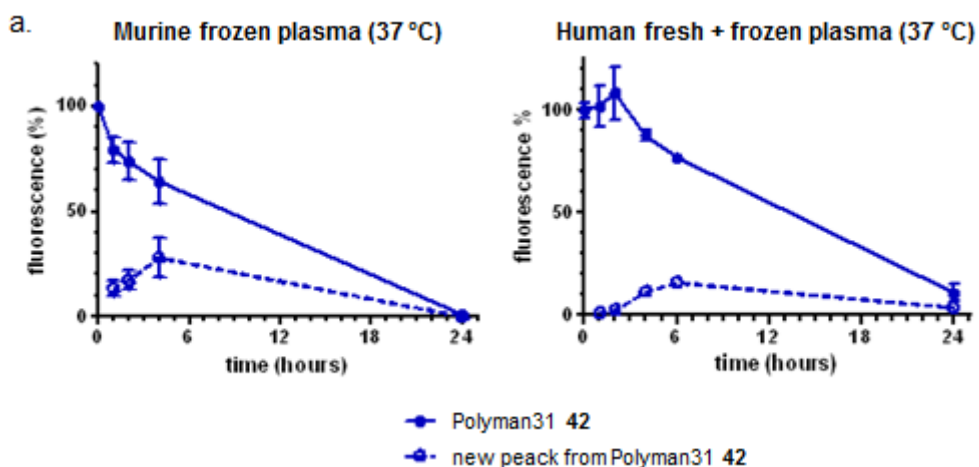
Polyman31 42 	60	3.2	2.1
Polyman20 52 	21	9	4.0

Table 1.2 - Compounds affinities for rhMBL. From inhibition curves, IC_{50} values of multivalent ligands for rhMBL were calculated. The relative potency γ for each constructs was calculated taking Polyman2 **37** as a reference ($\gamma=1$). Taking into consideration the number of monovalent ligands displayed by each rhMBL antagonist the relative potency weighted for the valency γ_w was defined.

The intrinsic fluorescence of Polyman31 **42** allowed to assess its stability in plasma by simply measuring intensity variations of its emission. Tests with both murine and human plasma at 37 °C revealed that the molecule is rather unstable showing a reduction in fluorescence intensity around 10-30% after 4 hours and complete disappearance of the signal after 24 hours (Fig. 1.9a). However, an additional fluorescence signal appeared (Fig. 1.9a) and increased inversely to the fluorescence of **42**. This new signal can likely be ascribed to a degradation product, possibly arising from hydrolysis of methyl ester groups of the dendron glycomimetic moieties. *In vivo* stability was also assessed in mice, monitoring both by fluorescence and MS spectrometry. The susceptibility of the dendron to hydrolysis was confirmed by analysis of the mass spectra (Fig. 1.9b). Moreover, the disappearance of signals of both Polyman31 **42** and hydrolysed metabolites was evaluated by fluorescence and mass analysis and was established to occur completely in just 1 hour in mice. Combining the *in vivo* data with the analysis in plasma, we conclude that the decrease of fluorescence observed *in vitro* (stability tests in plasma) is probably due to hydrolytic lability of the molecule, while *in vivo* clearance of the molecule is likely due to a fast excretion rate. (Fig. 1.9b) Indeed, fluorescence measurements of the mice urine suggests that most of **42** is probably expelled with first renal pass in mice (5 minutes after intravenous administration), which may also justify the inefficacy of Polyman31 **42** in providing beneficial effects in *in vivo* experiments with transient middle cerebral artery occlusion (tMCAo) mice models (Fig. 1.9c).



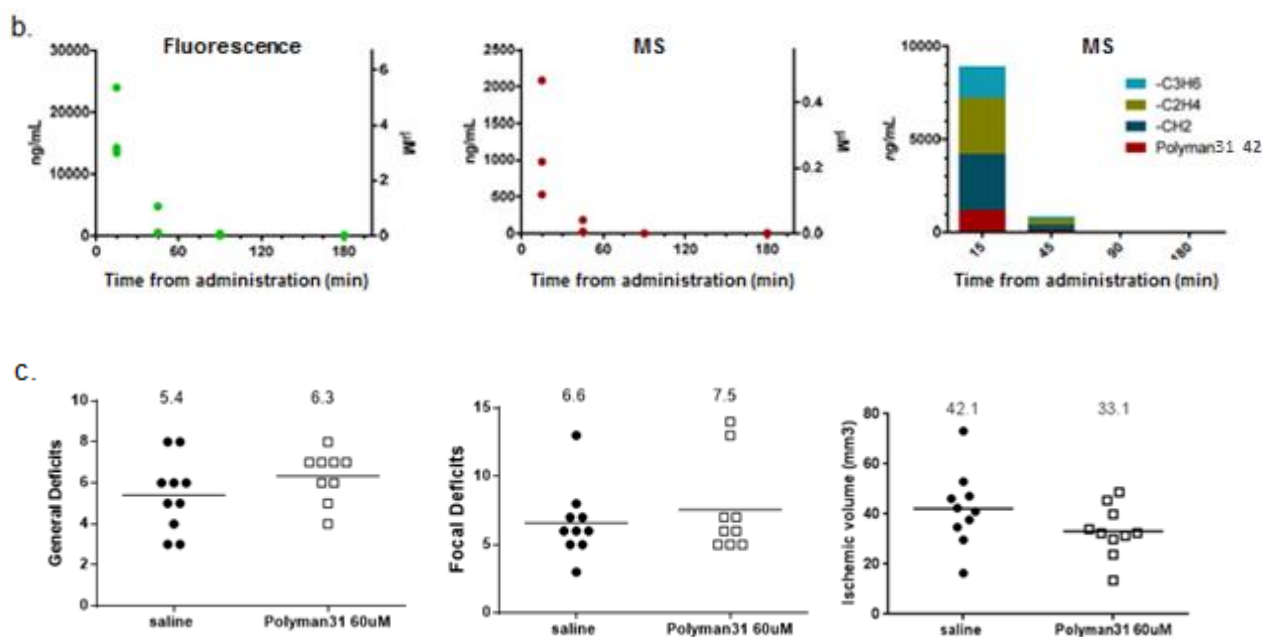


Figure 1.9 - *In vitro* and *In vivo* tests of Polyman31 42. **a)** *In vitro* stability test using both murine and human plasma at 37 °C. Fluorescence signal of Polyman31 42 and formation of a second peak ascribed to partial hydrolysis of 42 are monitored over time up to 24 hours. **b)** Concentration of Polyman31 42 in mice at different time from injection. The presence of the dendrimer was evaluated monitoring both with fluorescence and mass spectroscopy. The presence of hydrolysed metabolites of Polyman31 42 was also monitored by mass spectroscopy. **c)** Polyman31 42 protecting effects were evaluated in transient middle cerebral artery occlusion (tMCAo) mice models. No significant improvement of general and focal deficits, neither reduction of ischemic volume were observed upon administration comparing with a control experiment in which a saline buffer solution was administered.

The drawbacks of Polyman31 42 brought us to reconsider previously screened compounds with alternative architecture. We stated that Polyman2 37 combines good affinity for MBL ($K_D = 6 \mu\text{M}$) and good activity *in vivo* with relatively low loading of monovalent sugar ligand, while Polyman20 52 presents high avidity *in vitro* displaying multiple copies of the synthetically less demanding dimannoside 5. Thus, additional characterization of Polyman20 52 in *ex vivo* and *in vivo* studies will be pursued at the Mario Negri Institute. For our part, we focused on the possibility of optimizing the structure of 37 with two main goals: 1) chemical stabilization of the dendron moiety; 2) replacement of the ps-trisaccharide ligand 6 with the more accessible ps-di counterpart 5.

The design of a stabilised version 53 of previously developed dendrons 36 and 37 was based on minor changes of the previously developed scaffold, switching the hydrolytically unstable unhindered ester bonds into more robust amide functions (Fig. 1.10). To avoid solubility issues and for synthetic reasons explained below, we sought to replace the succinyl moiety of 37 with a slightly longer and more hydrophilic linker (blue box in Fig. 1.10). Additionally, a proper functionalization of the dendron focal point was adopted to allow the preparation of higher valency compounds by multipresentation.

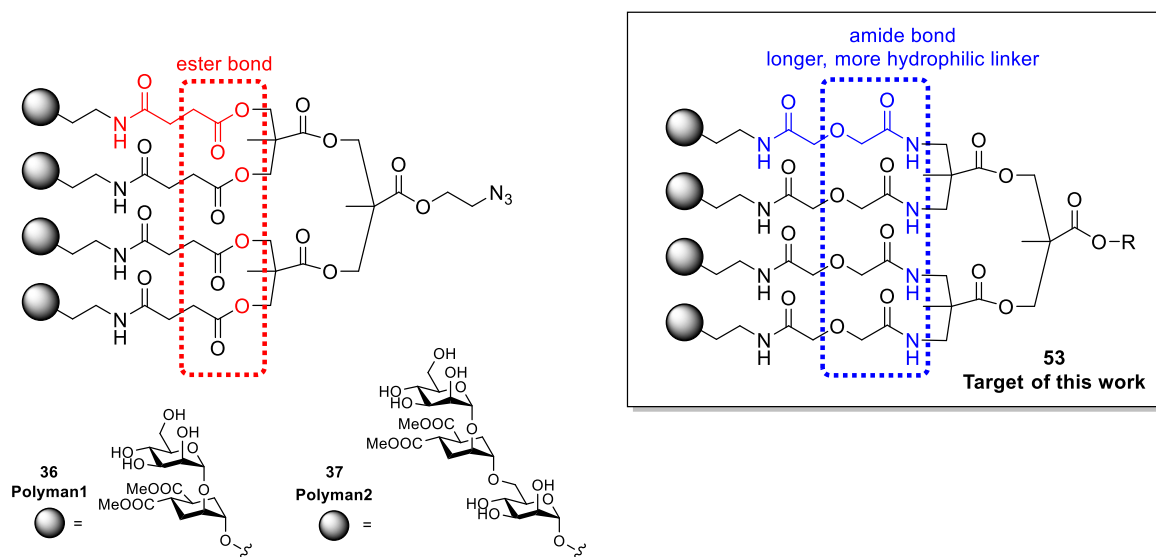
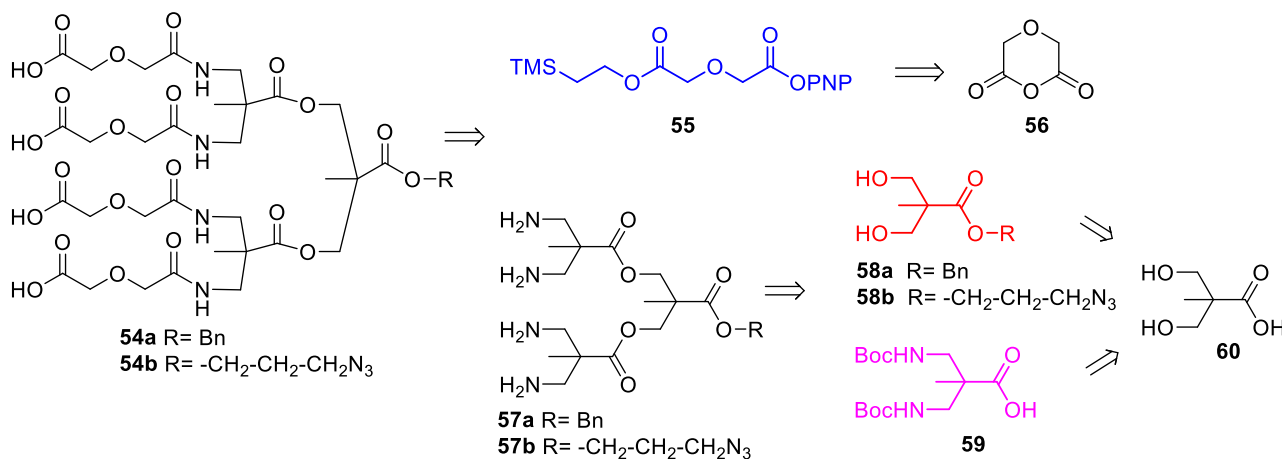


Figure 1.10 - Stabilised tetraamide MBL antagonists. Our aim is to attempt to stabilize previously developed dendrons **36** and **37** generating analogues with general structure of **53**. The feeble unhindered ester bonds of succinyl linker (red box) will be replaced with amido functions using a longer and more hydrophilic linker (blue box).

The retrosynthetic analysis that we adopted for the preparation of **54a** is depicted in Scheme 1.1 and three building blocks can be identified. The dendron structure **54a** is formed by reaction of tetraamine **57a** with the capping unit **55**. Building block **55** can be derived from ring opening reaction of diglycolic anhydride **56**. It was selected to replace the shorter succinate linker in order to avoid the possible intramolecular cyclization of amide groups over the activated carboxylic acid during the final functionalization of **54a**. In fact, an analogous reactivity has been described for aspartic acid derivatives, where 5-member ring imides can form upon activation.⁴⁷ Moreover, a linker containing an ether oxygen atom in the chain was chosen to favour water solubility of the final constructs.

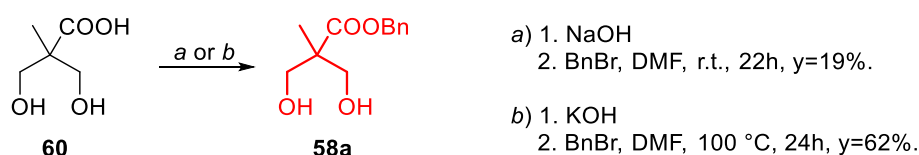
On the other hand, the tetraamine **57a** can be prepared by connecting two units of 2,2-bis(aminomethyl)propionic acid **59** to the 2,2-bis(hydroxymethyl)propionate ester focal unit **58a**. Introduction of reactive groups at the ester focal point allows for further functionalization of the final glycodendrons, enabling access to higher valency compounds and multiple presentation of the multivalent ligand. For this reason, the azido containing glycodendron **54b** was also prepared employing building block **58b**. Finally both **58a,b** and **59** can originate in a divergent approach from the common precursor 2,2-bis(hydroxymethyl)propionic acid **60**.



Scheme 1.1 - Retrosynthetic analysis of scaffold 54. Retrosynthetic analysis allowed to identify three building blocks: **55** can be obtained from ring opening reactions of diglycolic anhydride, while the ester **58** and the Boc-protected aminoacid **59** can both originate from the common precursor 2,2-bis(hydroxymethyl)propionic acid **60**.

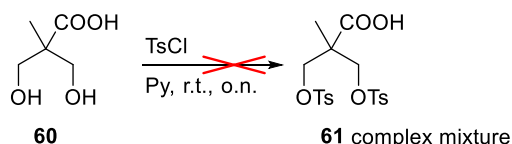
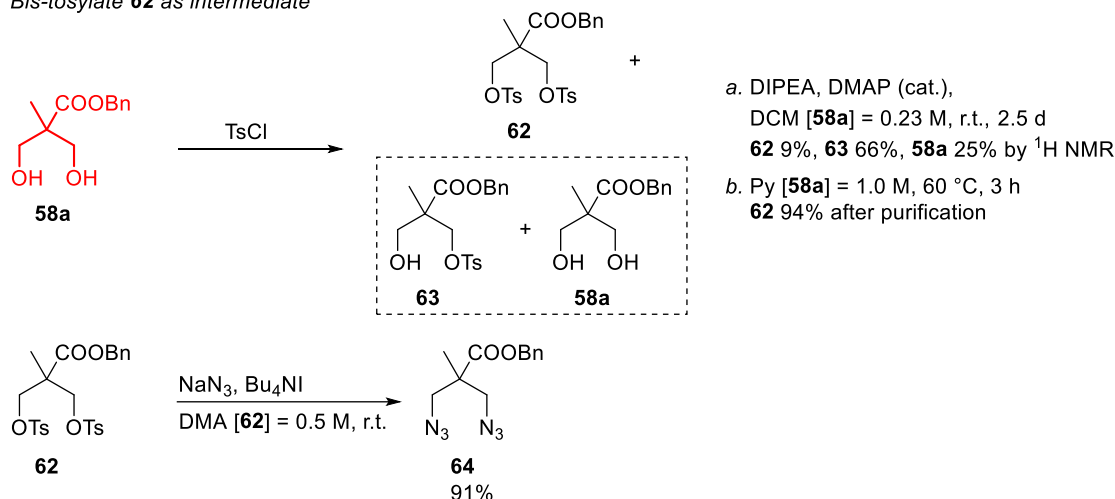
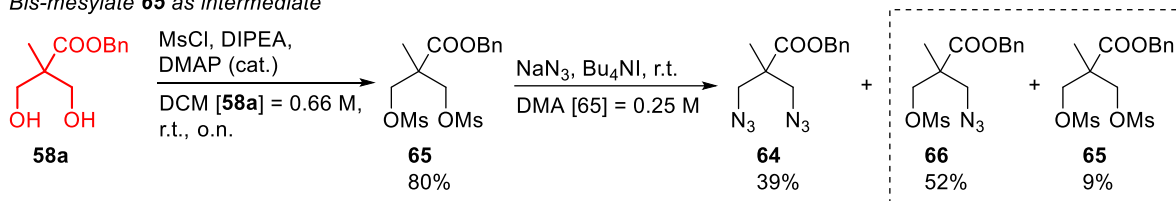
Herein we describe the synthesis of stabilised tetraamide antagonists with general structure as **53**. As a preliminary study, to explore the feasibility of the synthesis, dendrons with a benzyl ester group at the focal point **54a** were selected and prepared. The building blocks **58a**, **59** and **55**, identified by the retrosynthetic analysis, were prepared and connected to one another following a synthetic route aimed to minimizing protection/deprotection steps, which instead were required in the preparation of the polyester scaffold of the corresponding dendrons **36** and **37**.^{44,48}

The focal unit **58a** was obtained from 2,2-bis(hydroxymethyl)propionic acid **60** (Scheme 1.2). Deprotonation using NaOH followed by addition of BnBr afforded **58a** in low yield, 19%. However, changing for KOH in harsher conditions furnished **58a** in good yield (62%), which is comparable to an already reported procedure.⁴⁹



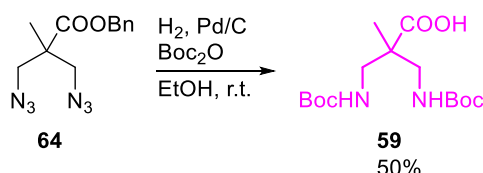
Scheme 1.2 - Synthesis of building block 58a. The two different procedures adopted for esterification of **60** to give the corresponding benzyl ester are reported.

The preparation of synthon **59** requires the transformation of the two hydroxyl groups of **60** into leaving groups by double nucleophilic substitution with NaN₃. Since direct tosylation of **60** as free acid using TsCl and pyridine at room temperature was unsuccessful (Scheme 1.3a), the ester **58a** was conveniently employed as intermediate in the synthesis of **59**. Therefore, **58a** was successfully transformed into the bis-tosylate **62**. As a first attempt, mild conditions (TsCl, DIPEA, DMAP cat., DCM, rt) were tried achieving only partial conversion (Composition of the crude revealed by ¹H NMR analysis: bis-tosylate **62** 9%, mono-tosylate **63** 66%, unreacted starting material **58a** 25%). This reaction showed a strong concentration dependence. Indeed complete conversion was obtained performing the reaction with TsCl at 60 °C with pyridine as the solvent at higher concentration, which furnished **62** as unique product in very high yield (94%). Pleasantly, the following double nucleophilic substitution step gave the desired bis-azido **63** in 91% yield. An alternative route envisaging mesylate **65** as intermediate was also explored. As opposed to tosylation, complete conversion was obtained performing mesylation in mild reaction conditions at increased concentration (MsCl, DIPEA, DMAP cat., DCM, rt). However, low conversion of **65** was achieved in the subsequent nucleophilic substitution reaction with NaN₃, which led, after 3.5 days, to a mixture of the product **64** (39%) along with mono functionalised **66** (52%) and unreacted starting material **62** (9%) (percentages calculated by ¹H NMR analysis of the crude) (Scheme 1.3b). Again, excessive dilution is likely responsible for the observed partial conversion.

Direct tosylation of **60**Bis-tosylate **62** as intermediateBis-mesylate **65** as intermediate

Scheme 1.3 - Synthesis of bis-azido intermediate 64. a) Unsuccessful direct tosylation of **60** is reported. b) Mild reaction conditions were not efficient for the synthesis of **62**, which was readily obtained under harsher and more concentrated reaction conditions. From **62**, double nucleophilic substitution afforded the desired **64**.

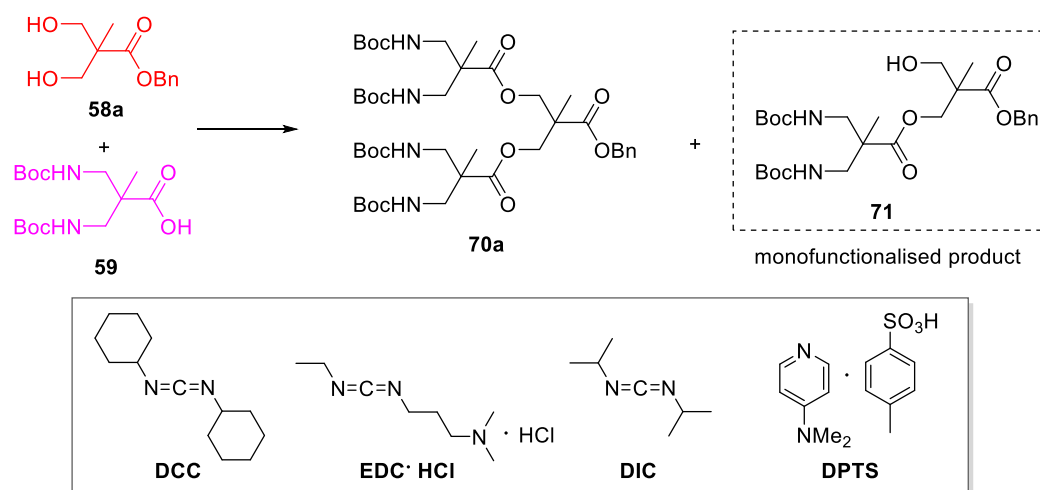
One-pot catalytic hydrogenation of **64** in the presence of di-*tert*-butyl dicarbonate finally afforded the *N*-Boc-protected bisaminoacid **59** (Scheme 1.4). This procedure enables subsequent reduction of azido groups into amines, protection of amines as Boc derivatives and removal of the benzylester, affording the carboxylic acid **59** in overall 50% yield, after flash column chromatography.



Scheme 1.4 - Synthesis of building block 59. The protected aminoacid **59** was obtained by catalytic hydrogenation of **64** in presence of di-*t*-butyl dicarbonate.

With the building blocks **58a** and **59** in hand, we started to synthesize the tetravalent scaffold. A double condensation reaction between diol **58a** and two molar equivalents of bisaminoacid **59** allowed the formation of the Boc-protected tetraamine **70a**. Initial trials using DCC (*N,N'*-dicyclohexylcarbodiimide) as condensing agent (Table 1.3, entry 1 and 3) identified DCM and room temperature as better reaction conditions compared to THF under reflux (Table 1.3, entry 2). The use of DPTS (dimethylaminopyridinium *p*-toluenesulfonate) was

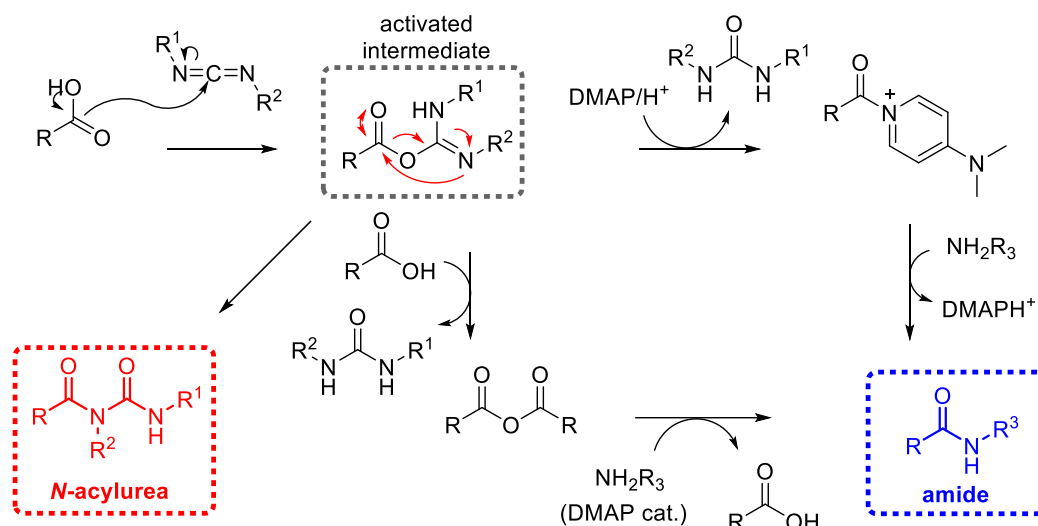
investigated, as it was reported to minimize the formation of *N*-acylurea side products, which are usually formed in this kind of reactions by rearrangement of the activated carboxylic acid intermediate (Scheme 1.7).⁵⁰ However, in our case, a slightly detrimental effect was observed using DPTS instead of DMAP (Table 1.3, entry 3). A better reaction outcome was finally obtained using a purer source of compound **59** (Table 1.3, entry 4), which allowed to obtain **70a** in 80% yield. Alternative condensing agents were also tested. EDC·HCL (*N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide) was less effective for the reaction (Table 1.3, entry 5), while the use of DIC (*N,N*'-diisopropylcarbodiimide) resulted in low yield, 37%, due to difficult isolation of the product from the corresponding *N*-acylurea side-product (Table 1.3, entry 6).



Scheme 1.6 - Synthesis of protected tetraamine 70a. Condensation of focal unit **58a** with an excess of building block **59** afforded the Boc-protected tetraamine **70a**. The possibly occurring monosubstituted side product **71** is shown along with the structures of the coupling reagents tested for the reaction.

Entry	Conditions:	Products:
1	59 (3.5 mol. eq.), DCC (2.8 mol. eq.), DMAP (1.0 mol. eq.), DCM, r.t.	70a 55% and 71 (isolated but not quantified)
2	59 (2.5 mol. eq.), DCC (2.8 mol. eq.), DMAP (1.0 mol. eq.), THF reflux	70a 16% and 71 48%
3	59 (3.0 mol. eq.), DCC (2.8 mol. eq.), DPTS (2.0 mol. eq.), DCM reflux	70a 39%
4	59 ^a (3.0 mol. eq.), DCC (3.0 mol. eq.), DMAP (1.0 mol. eq.), DCM, r.t.	70a 80%
5	59 ^a (3.0 mol. eq.), EDC (3.0 mol. eq.), DMAP (1.0 mol. eq.), NEt ₃ (3.0 mol. eq.), DCM, r.t.	70a and 71 in comparable amount (by TLC)
6	59 ^a (3.0 mol. eq.), DIC (3.0 mol. eq.), DMAP (1.0 mol. eq.), DCM, r.t.	70a 37%

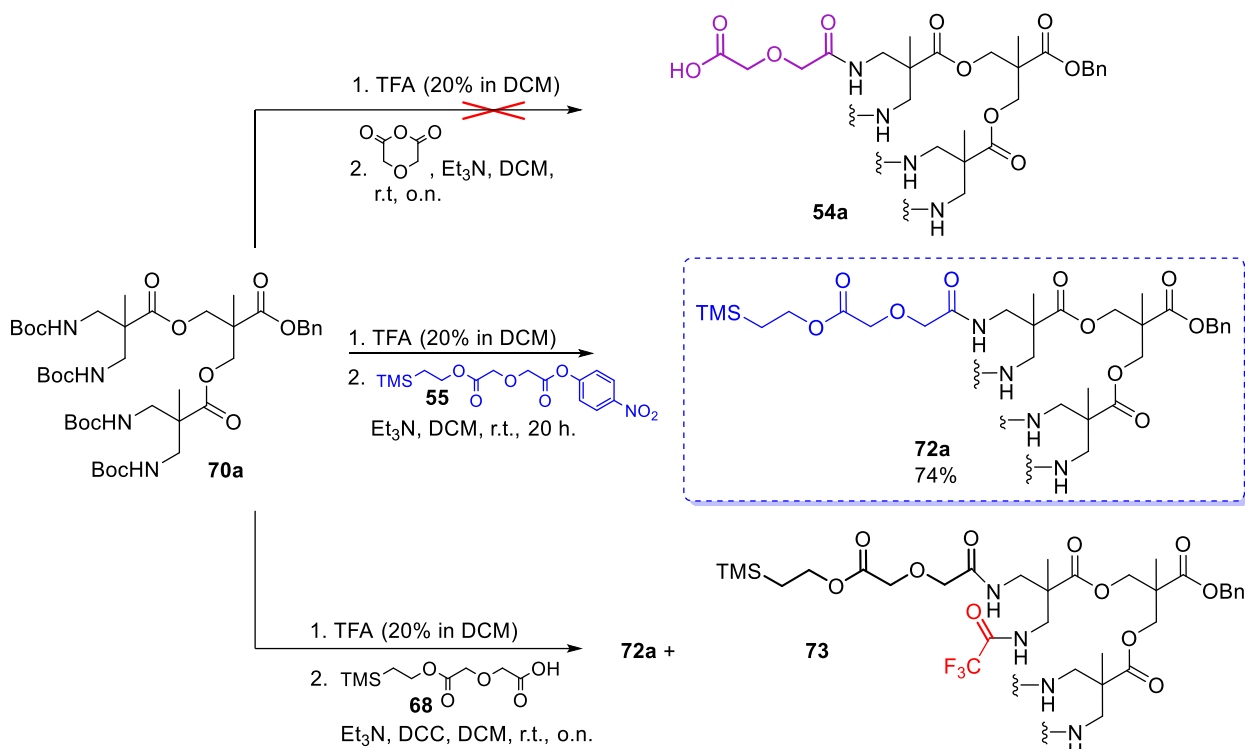
Table 1.3 - Reaction conditions attempted for the synthesis of 70a. The conditions that have been tested are reported: the amounts of **59**, the coupling agents and the solvent were varied. a) Reactions performed with a better purified source of **59**.



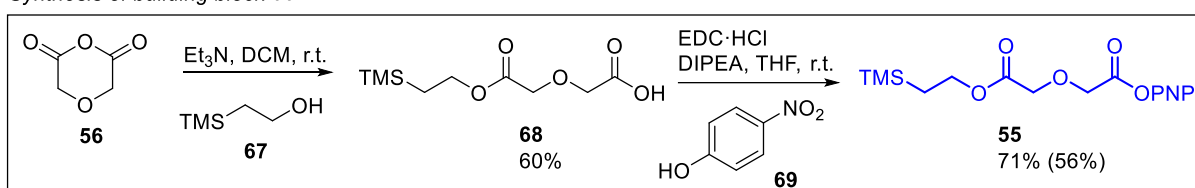
Scheme 1.7 - *N*-acylurea: a side-product in carbodiimide mediated condensation reactions. The general mechanism for amide formation envisaging activation of a carboxylic acid with a carbodiimide derivative and subsequent reaction with an amine is depicted. The attack of the acid over the carbodiimide lead to an activated intermediate that can either evolve toward the formation of the desired amide (blue box) or rearrange to give an *N*-acylurea as side-product (red box).

In order to directly obtain the tetraacid **54a** from **70a**, *N*-Boc removal in a 4:1 TFA/DCM mixture leading to tetraamine **57a** followed by ring opening reaction of diglycolic anhydride **56** was attempted. However, this procedure was abandoned due to difficulties in monitoring the formation of the tetra carboxylic acid **54a** (Scheme 1.9). Thus, we decided to prepare the related scaffold **72a** protected as trimethylsilylethyl ester. As a first attempt, *N*-Boc removal from **70a** and subsequent condensation reaction of the tetraamine with acid **68**, deriving from ring opening reaction of diglycolic anhydride with trimethylsilylethanol (60% yield after chromatographic purification), was tested. Using DCC as coupling agent, **72a** was obtained as judged by MS analysis, albeit the formation of many side-products was also observed. Indeed, residual TFA deriving from deprotection step is also activated towards nucleophilic attack, leading to undesired functionalization of the tetraamine scaffold as trifluoroacetamides. Up to doubly functionalized side-products were detected by ESI-MS analysis.

Elongation of the scaffold to give **72a** was cleanly achieved employing *p*-nitrophenyl ester **55**. This building block was synthesized by EDC·HCl mediated condensation reaction of carboxylic acid **68** and *p*-nitrophenol and was pure enough to be employed in the following step without further purification (71% yield was determined by ¹H NMR) (Scheme 1.5). However, chromatographic purification for characterisation purposes gave pure **55** as a waxy white solid in 56% yield. Finally, *N*-Boc deprotection of carbamate **70a** in a 4:1 TFA/DCM mixture and subsequent reaction of the tetraamine with an excess of the activated ester **55** afforded dendron **72a** as a crude that was successfully purified (74% yield) by automated chromatography using a high performance column (Biotage SNAP Ultra).

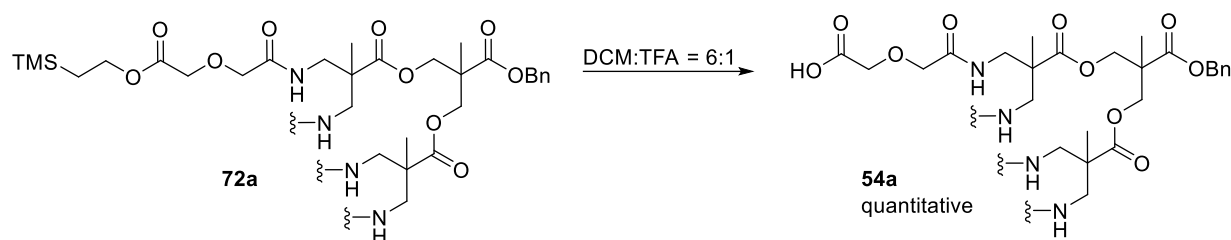


Synthesis of building block 55



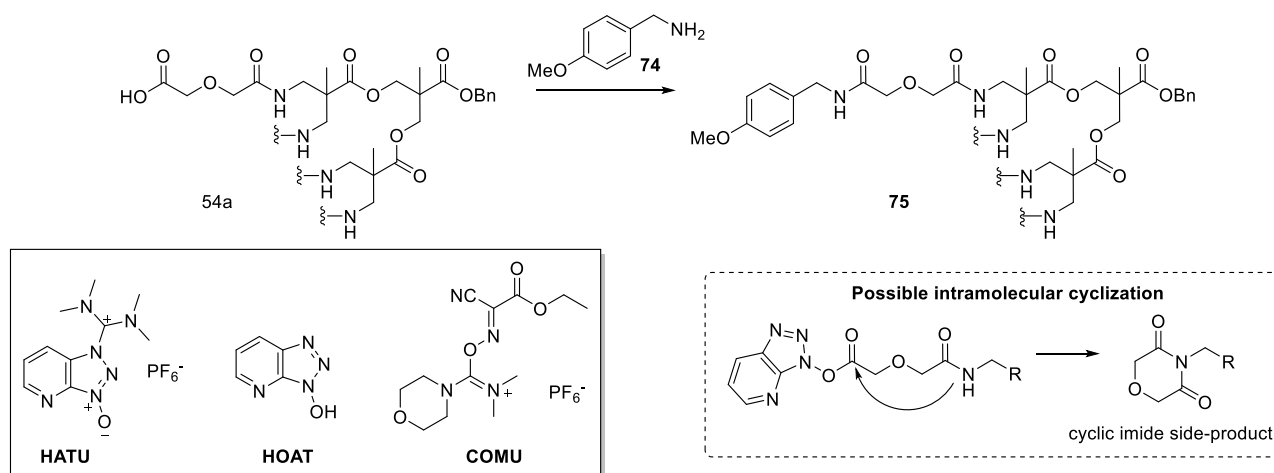
Scheme 1.8 - Synthesis of dendron 72a employing building block 55. Treatment of **70a** with TFA/DCM mixture and subsequent reaction with the activated building block **55** gave the protected **72a** dendron. Attempts aimed to directly obtain the deprotected derivative **54a** by ring opening of diglycolic anhydride failed. Deprotection of **70a** and subsequent DCC mediated condensation with intermediate **68** afforded the product **72a** in a mixture with side-products deriving from unwanted condensation with residual TFA. The synthesis of building block **55** is also shown. Starting from diglycolic anhydride **56**, a ring opening reaction with trimethylsilylethanol affords carboxylic acid **68**, which is activated as *p*-nitrophenyl ester giving the desired **55**.

The tetravalent trimethylsilylethyl ester dendron **72a** turned to be a very useful intermediate. This scaffold presented high stability, it could be stored for months without degradation and conveniently deprotected in acidic conditions only once needed. Treatment with a 6:1 DCM/TFA mixture quantitatively afforded the tetraacid **54a**. Coevaporation of TFA with toluene and diethylether is a critical operation since esterification of **54a** was observed under acidic conditions in the presence of nucleophilic solvents such as methanol. Generally the product could be employed directly as a crude in the next synthetic step, but an analytical sample was purified by automated reverse chromatography (Biotage SNAP C18, gradient elution: from 100% H₂O to 100% MeOH) for characterisation purposes, (72%).



Scheme 1.9 - Deprotection of dendron 72a. The synthesis of tetraacid **54a** was achieved treating the corresponding protected scaffold with a mixture of DCM/TFA.

The reaction conditions for the final functionalization of dendron **54a** were optimized employing the commercially available *p*-methoxybenzylamine **74** as a more accessible amine counterpart. (Scheme 1.10). Treating **54a** with an excess of the amine **74** in presence of HATU and DIPEA in DMA at 30 °C allowed the formation of the tetrasubstituted **75** in 47% yield (Table 1.4, entry 1). Since ESI-MS analysis of the crude assessed full conversion, the low yield may be due to purification issues and the small scale of the reaction with consequent loss of material during purification. Monitoring by ESI-MS analysis, no significant differences were obtained varying the HATU coupling agent with a combination of HATU and HOAT or using COMU (Table 1.4, entry 2 and 3). On the contrary, increasing the temperature is detrimental for the reaction outcome (Table 1.4, entry 4), since the amide functionality of **54a** can compete for nucleophilic attack over the activated carboxylic acid by an intramolecular process giving cyclic imide side-products (detected by ESI-MS; see Scheme 1.10). Moreover, when the amine excess was decreased to 4.8 mol.eq. a worse conversion and formation of impurities were detected (Table 1.4, entry 5).



Scheme 1.10 - Optimization study for efficient functionalization of scaffold 54a as a tetraamide. The reaction conditions for the functionalization of **54a** as glycodendrons were previously optimized using the commercially available amine **74**. The structures of the coupling reagents employed are reported in the box. The key step of the mechanism responsible for the formation of possible intramolecular cyclization side-products is also shown.

Entry	Conditions	Yield
1	74 (7.6 mol. eq.), HATU (8.0 mol. eq.), DIPEA (16.0 mol. eq.), DMA, 30 °C.	75 47%
2	74 (7.6 mol. eq.), HATU (4.4 mol. eq.), HOAT (4.4 mol. eq.), DIPEA (4.4 mol. eq.), DMA, 30 °C.	Comparable conversion to entry 1 by ESI-MS
3	74 (7.6 mol. eq.), COMU (4.8 mol. eq.), DIPEA (9.6 mol. eq.), DMA, 30 °C.	Comparable conversion to entry 1 by ESI-MS

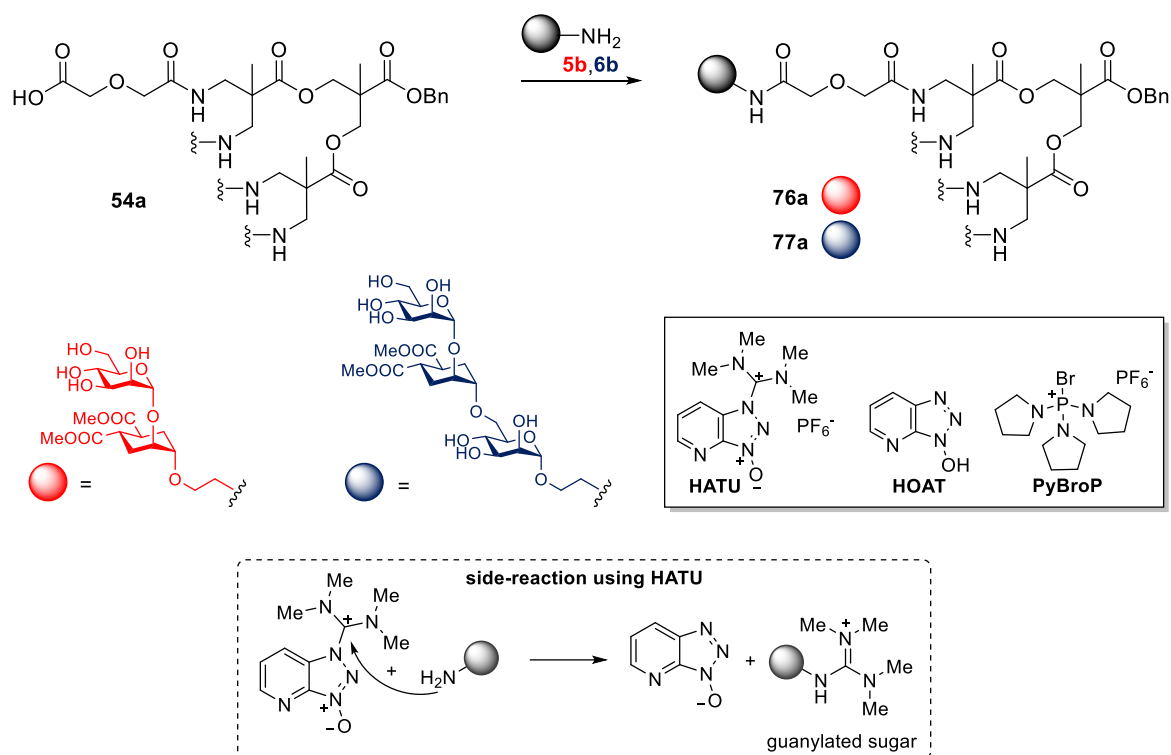
4	74 (7.6 mol. eq.), HATU (8.0 mol. eq.), DIPEA (16.0 mol. eq.), DMA, 70 °C.	Undesired cyclization
5	74 (4.8 mol. eq.), HATU (4.4 mol. eq.), DIPEA (8.8 mol. eq.), DMA, 30 °C.	Impurities

Table 1.4 - Condensation between scaffold 54a and *p*-methoxybenzylamine 74. The table summarizes the reaction conditions tested. Three different coupling reagents were screened, all giving good conversions (entry 1-3). Increasing the temperature was detrimental for the reaction (entry 4) as well as reducing the amine excess employed (entry 5).

For the final conjugation of dendron **54a** with glycomimetics **5b**, the reaction conditions previously reported in Table 1.4, entry 1 were selected (HATU, DMA, 30 °C, 12 h) affording the tetravalent glycodendron **76a** in 82% (Table 1.5, entry 1). On the other hand, the functionalization of scaffold **54a** with the pseudo-trisaccharide **6b** required slightly harsher conditions (Table 1.5, entry 2). Both PyBroP and the combination HATU/HOAT were used as coupling agent at 37 °C, with the latter condition being the best and affording the corresponding glycodendron **77a** in 74% yield after purification. Notably, despite the increase of temperature no intramolecular cyclization was observed by MALDI analysis (HCCA, MeOH) of the reaction crude.

Both dendrons **76a** and **77a** were first purified from low molecular weight impurities by size exclusion chromatography (Sephadex LH-20) and, as opposed to **36** and **37**, could be submitted to a second direct phase flash chromatography which afforded the pure products. Moreover, **76a** proved to be stable under reverse phase HPLC purification conditions (H₂O/CH₃CN + 0.1% of formic acid). The excess of glycomimetics **5b** and **6b** could be recovered after size exclusion chromatography as guanidine derivatives, coming from side-reaction with the excess HATU (Scheme 1.11).

The stability of **77a** towards hydrolysis was corroborated by ¹H NMR: no hydrolysis was observed in water solution at physiological pH (PBS buffer, pH 7.4), monitoring the dendron for up to four days. Both glycodendrons were fully characterised by NMR, HRMS ESI and MALDI (HCCA) mass spectroscopy. Notably, the signals of both dendrons **76a** and **77a** in MALDI MS analysis are always associated with an additional signal at lower molecular weight (*m/z*: 1947.7 and 1090.1 respectively) that arises from dendron decomposition upon ionization.

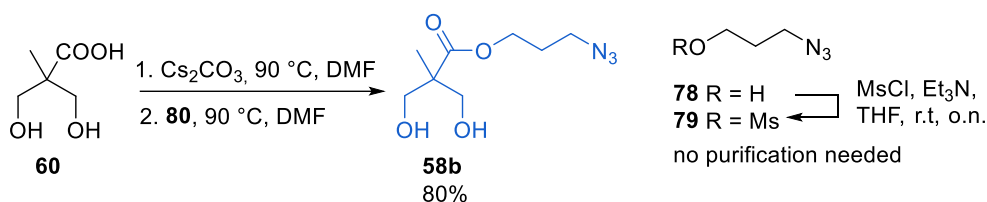


Scheme 1.11 - Synthesis of glycodendrons 76a and 77a. Scaffold **54a** was functionalised either with four copies of the amino tethered pseudo-mannoside **5b** and in analogous fashion with the pseudo-trimannoside **6b**.

Entry	Conditions	Yield
1	5b (7.6 mol. eq.), HATU (8.0 mol. eq.), DIPEA (16.0 mol. eq.), DMA, 30 °C.	76a 82%
2	6b (7.6 mol. eq.), HATU (8.0 mol. eq.), DIPEA (16.0 mol. eq.), DMA, 30 °C.	77a 15%
3	6b (7.6 mol. eq.), PyBroP (8.0 mol. eq.), DIPEA (16.0 mol. eq.), DMA, 37 °C.	77a 56%
4	6b (7.6 mol. eq.), HATU (6.0 mol. eq.), HOAT (6.0 mol. eq.), DIPEA (16.0 mol. eq.), DMA, 37 °C.	77a 74%

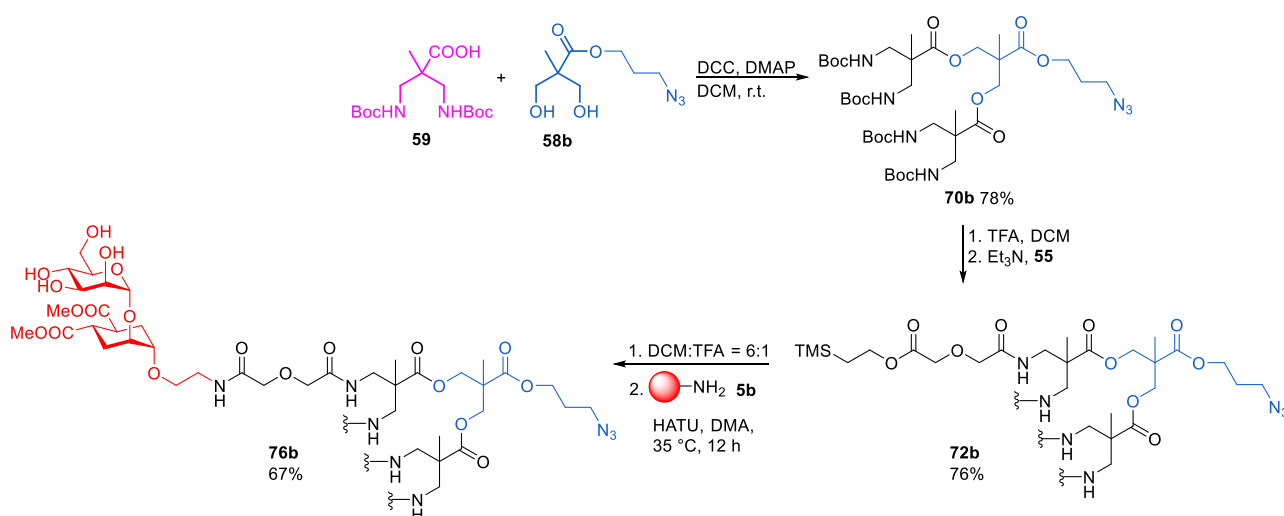
Table 1.5 - Conditions for the synthesis of glycodendrons 76a and 77a. The table reports the conditions used for condensation between the tetraacid **54a** and the linker amino group of the properly functionalised pseudo-mannoside **5b** (entry 1). Different coupling reagents were tested for the analogous conjugation of **54a** with pseudo-trimannoside **6b** (entry 2-4), revealing the combination of HATU/HOAT to perform as the best.

The increased stability of the new generation constructs brought us to consider the opportunity of orthogonally functionalize the dendron focal point allowing further diversification and the generation of higher valency compounds. For this purpose, the azido containing focal unit **58b** was prepared (80%) by esterification of **60** treating with mesylate **79** derived from 3-azidopropanol **78** (Scheme 1.12).⁵¹



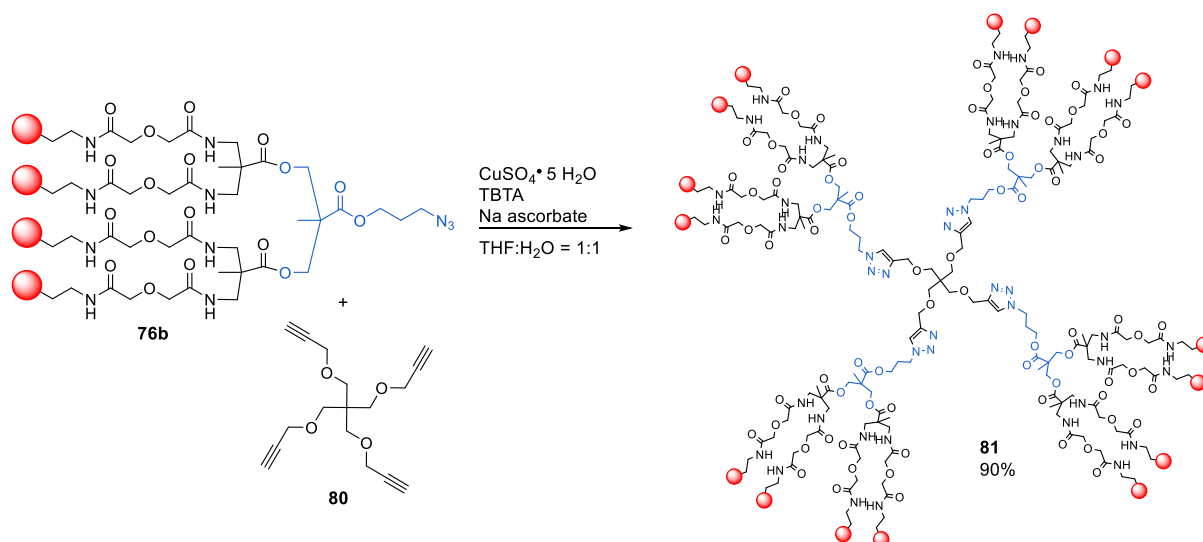
Scheme 1.12 - Synthesis of the azido containing building block 58a. Esterification of **60** with the mesylate derivative **79** obtained from 3-azidopropanol **78** afforded the azido armed **58b** that was then used as dendron focal unit.

The modularity and reproducibility of the synthesis that we have designed allowed us to follow analogous synthetic steps to combine building blocks **58b**, **59** and **55** and to readily generate the azide-armed tetraivalent dendron **72b** (Scheme 1.13). Finally, deprotection and conjugation with **5b** (HATU, DMA, 35 °C, 12 h, 67%) afforded the glycodendron **76b** that was purified with both size exclusion chromatography and direct phase flash chromatography in good yields, (67% over 2 steps).



Scheme 1.13 - Synthesis of the azido armed glycodendron 76b. Using focal unit **58b** and following the same synthetic strategy already described for the synthesis of the glycodendrons, construct **76b** was readily obtained in a reproducible way in good yields (40% over 5 steps).

The ease of preparation and increased stability of glycodendron **76b** makes it a valuable and versatile scaffold that can be used to generate compounds with higher ligand presentation. Indeed, exploiting Cu(I) catalysed azide-alkyne cycloaddition (CuAAC) reaction between **76b** and the propargyl ether derivative of pentaerythritol **80**, the hexadecavalent glycodendrimer **81** was obtained in high yield (90%) after purification with size exclusion chromatography (Sephadex LH-20) (Scheme 1.14). The purity of the dendrimer was assessed by reverse phase HPLC and the compound was fully characterised by NMR and HRMS ESI.



Scheme 1.14 - Synthesis of the 16-valent dendrimer **81.** Exploiting CuAAC reaction, the azido tethered dendron **76b** was reacted in presence of the propargyl ether derivative of pentaerythritol scaffold **80** finally affording dendrimer **81** in very good yield (90%).

The inhibition potency of the stabilised dendrons was evaluated by a recently established SPR competition assay that was developed by the group of Dr. M. Gobbi at Mario Negri Institute (Milan).⁵² In this assay dendron affinity was evaluated both for murine MBL-A and MBL-C isoforms contained in murine plasma. Therefore, the dendrons were incubated at different concentrations for 30 min at 25 °C with diluted murine plasma and then directly flowed over the sensor chip coated with immobilized mannosylated BSA (Man-BSA). The ability to compete with mannosylated BSA for the binding of MBL isoforms was compared with the previously developed Polyman2 **37**. The amount of MBL-A and MBL-C bound to the surface was detected sequentially flowing anti-mMBL-A and anti-mMBL-C antibodies, which bind to the proteins giving a clear SPR signal. The higher is the dendron concentration, the lower is the amount of MBL bound to the surface which will be recognized by the antibodies resulting in a concentration-dependent decrease of the SPR signal (Fig. 1.11).

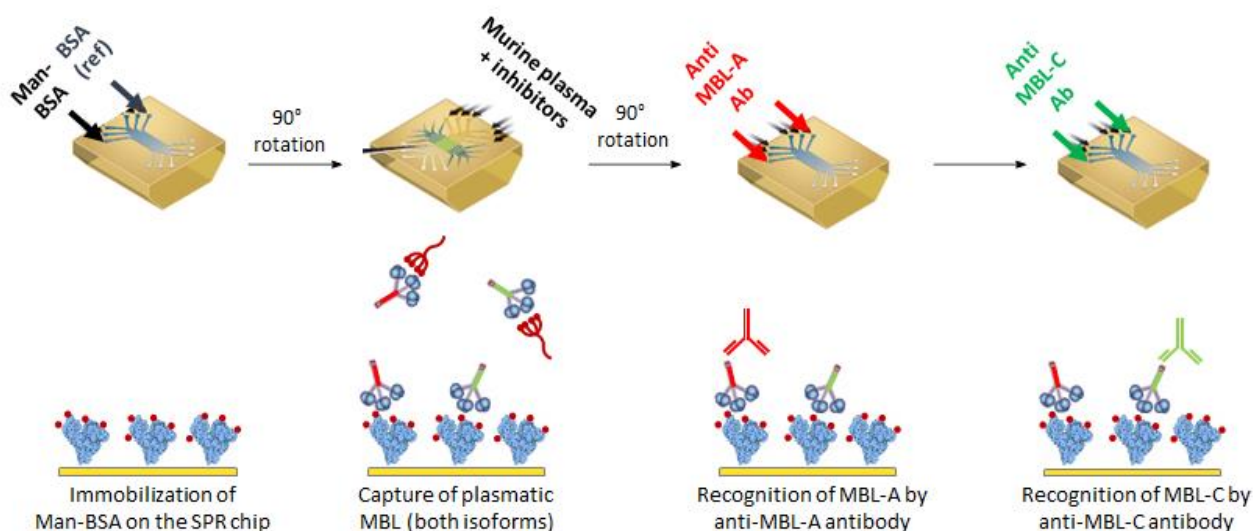


Figure 1.11 - SPR competition assay of newly synthesized dendrons for MBL-A and MBL-C isoforms in murine plasma. Dendrons inhibition potency was evaluated by a competition assay with immobilized mannosylated-BSA as competitor. BSA was also immobilized on the chip as negative control. Dendrons were incubated for 30 min at 25 °C at different concentrations with diluted murine plasma containing both MBL-A and MBL-C isoforms. The solutions were

then flowed over the sensor chip coated with immobilized Man-BSA. Anti-MBL-A and anti-MBL-C antibodies were subsequently flowed to detect MBL isoforms bound to the surface.

From the SPR experiments inhibition curves of multivalent ligands were obtained (Fig. 1.12) and the corresponding IC_{50} values for inhibition of MBL-A and MBL-C were calculated (Table 1.6).

The results obtained for inhibition of murine MBL-A binding were of difficult interpretation. Indeed, while glycodendron **76a**, bearing the pseudo-disaccharide **5**, presented comparable affinity with Polyman2 **37**, only half of the affinity was shown by Polyman50 **77a**, which is functionalised with four copies of the more potent monovalent pseudo-trisaccharide **6**. Even more unexpected behaviour was shown by the 16-valent dendrimer Polyman54 **81**, that was found to be far less potent than the others constructs ($IC_{50} \gg 200 \mu\text{M}$). Although it has been demonstrated that high ligand density can limit lectin accessibility hampering protein-carbohydrate interactions,⁵³ this seems not to be the case. Indeed, a positive correlation between valency and binding affinity was shown in the inhibition experiments with MBL-C. With this isoform comparable avidity with Polyman2 **37** was observed for Polyman49 **76a**. The Polyman50 **77a**, bearing the pseudo-trisaccharide **6**, show a two times increased avidity, while the higher ligand density of Polyman54 **81** enhanced affinity up to two orders of magnitude ($IC_{50} = 1.0 \mu\text{M}$).

The contrast between experimental data obtained for dendron binding affinity highlight the contrast of valency and avidity correlation for the inhibition of MBL-A (negative correlation) and MBL-C (positive correlation). However, it cannot be excluded that our results may be affected by the complexity of the adopted SPR assay. Control experiments incubating the dendrons in buffered solution with either pure MBL-A or MBL-C isoforms are under investigation.

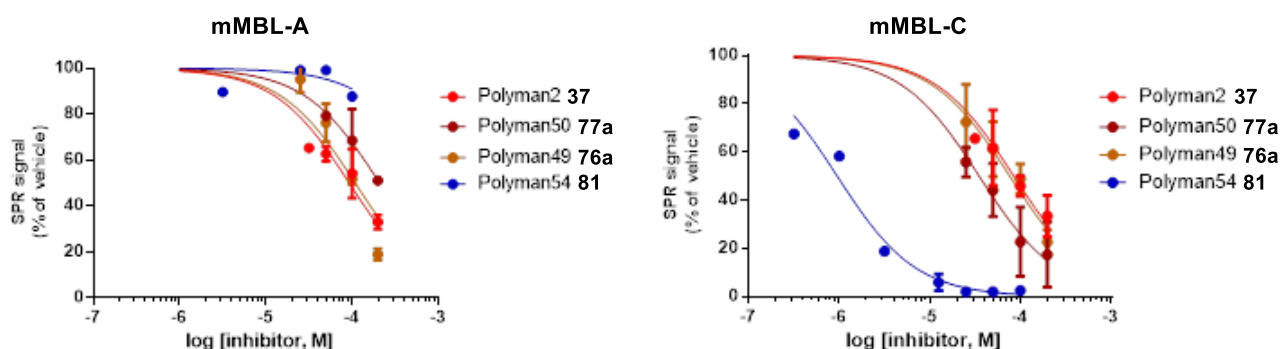


Figure 1.12 - Concentration-dependent inhibition curves for the newly synthesised dendrons 76a and 77a and the highly valent dendrimer 80. The inhibition curves were obtained by SPR inhibition assay. Multivalent ligands were previously incubated at different concentrations with murine plasma containing both MBL-A and MBL-C isoforms. The solutions were then flowed over a sensor chip where mannosylated-BSA was previously immobilised. The binding of MBL-A and MBL-C to Man-BSA was revealed subsequently flowing anti-MBL-A and anti-MBL-C antibodies on the chip which resulted in a clear concentration-dependent SPR signal. From the inhibition curves the corresponding IC_{50} values were calculated. Polyman2 **37** dendron was screened as reference compound.

Compound	MBL-A inhibition			MBL-C inhibition		
	IC_{50} (μM) (95% conf.) ^a	Relative potency γ^b	γ^c	IC_{50} (μM) (95% conf.) ^a	Relative potency γ^b	γ^c

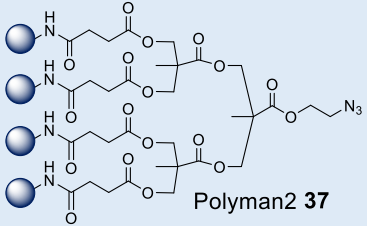
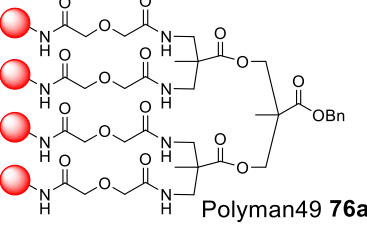
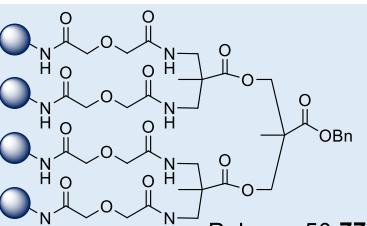
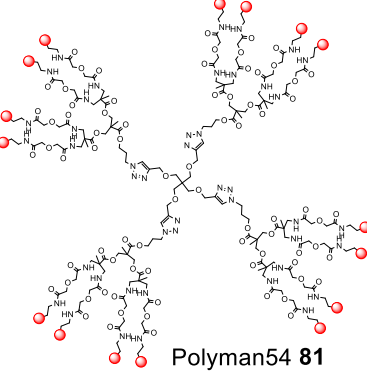
 <p>Polyman2 37</p>	93 (77-113)	1	1	84 (61-118)	1	1
 <p>Polyman49 76a</p>	109 (66-180)	0.85	0.85	77 (57-102)	1.1	1.1
 <p>Polyman50 77a</p>	209 (154-285)	0.44	0.44	35 (25-50)	2.4	2.4
 <p>Polyman54 81</p>	>> 200	-	-	1.0 (0.7-1.3)	84	21

Table 1.6 - Inhibition potencies of 76a, 77a and 80 multivalent ligands against murine MBL-A and MBL-C. The table summarizes the structures of the tested compounds and the corresponding IC_{50} values obtained in SPR competition assay for both MBL-A and MBL-C. **a)** 95% confidence intervals of IC_{50} values are reported in brackets. **b)** The relative potency γ was calculated for each compound taking IC_{50} value of Polyman2 **37** as a reference. **c)** The relative potency corrected per number of monovalent ligands γ_w was also calculated.

1.5 Conclusions

In recent years, the role of MBL in the pathogenesis of ischemic brain damage has been disclosed, identifying it as an attracting and promising pharmaceutical target. Recently we have demonstrated that the tetra-valent glycodendron Polyman2 **37** can perform as protecting agent in transient middle cerebral artery occlusion (tMCAo) mice models. Nevertheless this construct presented high sensitivity to hydrolysis, which aggravates its preparation and hampers pharmacokinetic studies.³⁷ Previously developed polymannosylated DC-SIGN antagonists have been synthesized and tested as MBL ligands showing IC_{50} values in the micromolar range. In this study we have presented *in vitro* and *in vivo* results obtained with one of the most active of these

dendrons, Polyman31 **42**. This ligand possess a hexavalent structure functionalized with pseudo-dimannoside glycomimetic moieties, which constitute a good compromise between synthetic accessibility and potency. Moreover, the extended rigid core allows this construct to perform as a chelating agent, able to simultaneously bind two carbohydrate recognition domains of MBL homotrimer. Good binding affinity towards rhMBL (60 μ M) was revealed by SPR inhibition assay. However, the poor pharmacokinetic properties of this compound brought us to select the highly active nonavalent Polyman20 **52** as an alternative candidate for *in vivo* investigations. Moreover, since Polyman2 **37** showed protecting effect in brain ischemic injury with wide therapeutic window, our aim was to stabilize the scaffold structure of this construct. Therefore dendrons **76a** and **77a**, where the labile ester bond of succinyl linkers are replaced with more robust amide bonds using longer and more hydrophilic linkers, were successfully prepared. These constructs showed remarkable increased stability, good solubility in water and proved to be able to bind MBL-C murine isoform in murine plasma in comparable way compared to the analogue **37** by SPR inhibition assay. The versatility of this newly designed scaffold and the robustness of the synthetic strategy was confirmed and exploited to prepare the azido-armed glycodendron **76b**, which gave access to the highly functionalised dendrimer **81**. The binding ability **81** for MBL-C was also assessed by SPR inhibition assay in murine plasma, where it showed an increased potency of two orders of magnitude. The efficacy of the developed dendron architecture in displaying the sugar ligands and the ease of being functionalized by CuAAC allow for a rapid access towards high valency constructs of various geometries. These features will guarantee an alternative strategy to target MBL if on going *in vivo* studies with Polyman20 **52** dendrimer will reveal it not to be effective as a protective agent in ischemic injury.⁵⁴

1.6 Experimental

1.6.1 General methods and procedures

Chemicals were purchased from commercial sources and used without further purification, unless otherwise indicated. When anhydrous conditions were required, the reactions were performed under nitrogen atmosphere. Anhydrous solvents were purchased from Sigma-Aldrich® with a content of water ≤ 0.005 %. *N,N'*-diisopropylethylamine (DIPEA), triethylamine were dried over calcium hydride, THF was dried over sodium/benzophenone and freshly distilled. Reactions were monitored by analytical thin-layer chromatography (TLC) performed on Silica Gel 60 F₂₅₄ plates (Merck) with UV detection (254 nm and 365 nm) and/or staining with ammonium molybdate acid solution, potassium permanganate alkaline solution or ninhydrin. Flash column chromatography was performed according to the method of Still and co-workers⁵⁵ using silica gel 60 (40-63 μ m) (Merck). Automated flash chromatography was performed with Biotage Isolera Prime system, Biotage SNAP KP-Sil cartridges were employed unless otherwise indicated. Size-exclusion chromatography was performed using Sephadex LH-20 from GE Healthcare Life Science. HPLC purifications were performed on Dionex Ultimate 3000 equipped with Dionex RS Variable Wavelength Detector (column: Atlantis Prep T3 OBD™ 5 μ m 19 x 100 mm; flow 15 mL/min unless stated otherwise). NMR experiments were recorded on a Bruker AVANCE-400 MHz instrument at 298 K. Chemical shifts (δ) are reported in ppm. The ¹H and ¹³C NMR resonances of compounds were assigned with the assistance of COSY and HSQC experiments. Multiplicities are assigned as s (singlet), d (doublet), t (triplet), q (quartet), quint (quintet), m (multiplet). Mass spectra were recorded on Apex II ICR FTMS (ESI ionization-HRMS), Waters Micromass Q-TOF (ESI ionization-HRMS), ThermoFischer LCQ apparatus (ESI ionization) or Bruker Daltonics Microflex LT (MALDI-TOF apparatus). Specific optical rotation values were measured using a Perkin-Elmer 241, at 589 nm in a 1 dm cell. The following abbreviations are used: DCC (*N,N'*-dicyclohexylcarbodiimide), DCM (CH₂Cl₂), DMA (*N,N'*-dimethylacetamide), DMF (*N,N'*-dimethylformamide), DIPEA (*N,N'*-diisopropylethylamine), DMAP (4-dimethylaminopyridine), EDC (*N*-(3-dimethylaminopropyl)-*N'*-

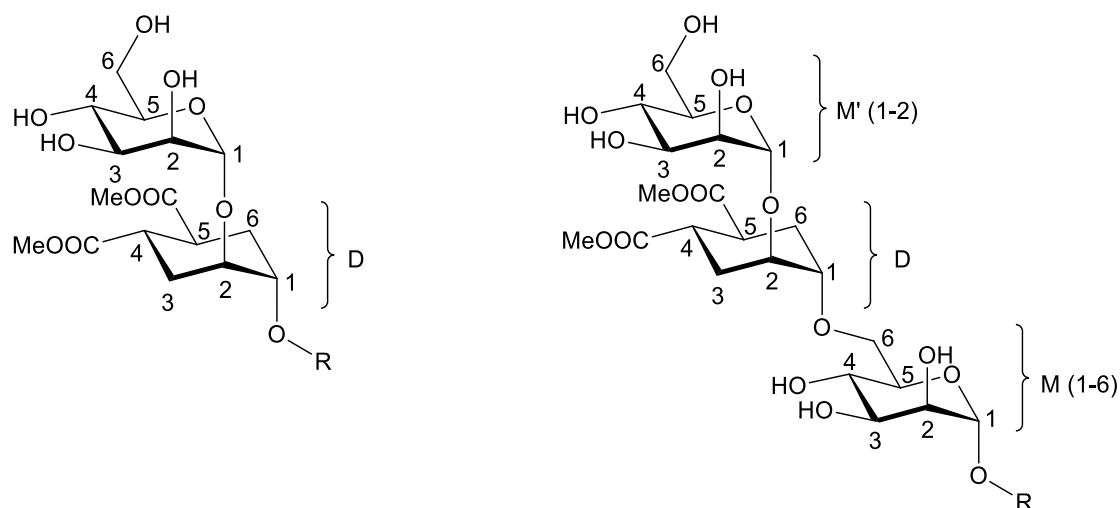
ethylcarbodiimide), HATU (1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate), HOAT (1-hydroxy-7-azabenzotriazole), TBTA (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine), TFA (trifluoroacetic acid), THF (tetrahydrofuran). Compounds **5b**,⁵⁶ **6b**,³⁸ **79**,⁵¹ **80**⁴⁵ were prepared as previously described in the literature. Compounds **56**, **60**, **67** and **69** are commercially available. Mouse plasma was obtained from eight-week-old C57Bl/6 mice (wild-type, WT, 20-25g; Harlan Laboratories) and from C57Bl/6 mice with a target mutation of both MBL-A and MBL-C genes (MBL^{-/-}, 20-25g, purchased from Jackson Laboratories-USA and colonized at the Mario Negri Institute). Blood samples were collected from the *vena cava* in 10 mM ethylenediaminetetracetic acid and 0.125% polybrene (Sigma-Aldrich). Plasma, separated by centrifugation for 15 min at 2000 g at 4 °C, was immediately stored and kept at -80 °C until use.

The IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, adheres to the principles set out in the following laws, regulations, and policies governing the care and use of laboratory animals: Italian Governing Law (D.lgs 26/2014; Authorisation n.19/2008-A issued March 6, 2008 by Ministry of Health); Mario Negri Institutional Regulations and Policies providing internal authorisation for those conducting animal experiments (Quality Management System Certificate – UNI EN ISO 9001:2008 – Reg. N° 6121); the NIH Guide for the Care and Use of Laboratory Animals (2011 edition) and EU directives and guidelines (EEC Council Directive 2010/63/UE). The Statement of Compliance (Assurance) with the Public Health Service (PHS) Policy on Human Care and Use of Laboratory Animals has been recently reviewed (9/9/2014) and will expire on September 30, 2019 (Animal Welfare Assurance #A5023-01). The results reported here were obtained within a project specifically authorized by the Italian Ministry of Health (Decree n° 161/2014B). SPR studies were carried out using a ProteOn XPR36 Protein Interaction Array apparatus (Bio-Rad Laboratories, Hercules, CA). The instrument is characterized by six parallel flow channels that can immobilize up to six ligands on the same sensor chip.⁵⁷ Mannosylated bovine serum albumin (Man-BSA, Dextra Laboratories, Reading, UK) was immobilized on one channel of the sensor chip (GLC, Bio-Rad) by amine coupling chemistry;⁵⁸ BSA (Sigma) was immobilized, as a reference, in a parallel different channel. Immobilization levels were typically 3000 and 4000 resonance units (RU, 1RU 1pg protein/mm²), respectively.

1.6.2 Synthesis of MBL multivalent antagonists

1.6.2.1 Products numbering for spectral assignment

The unusual numbering used for pseudo-disaccharide and pseudo-trisaccharide derivatives are shown below. The cyclohexanediol unit is indicated as D. Numbering of this ring does not follow IUPAC rules, but it was adopted in analogy to the numbering of mannose. In pseudo-trisaccharide derivatives, the non reducing end Man residue is indicated as M' and the reducing end one as M.

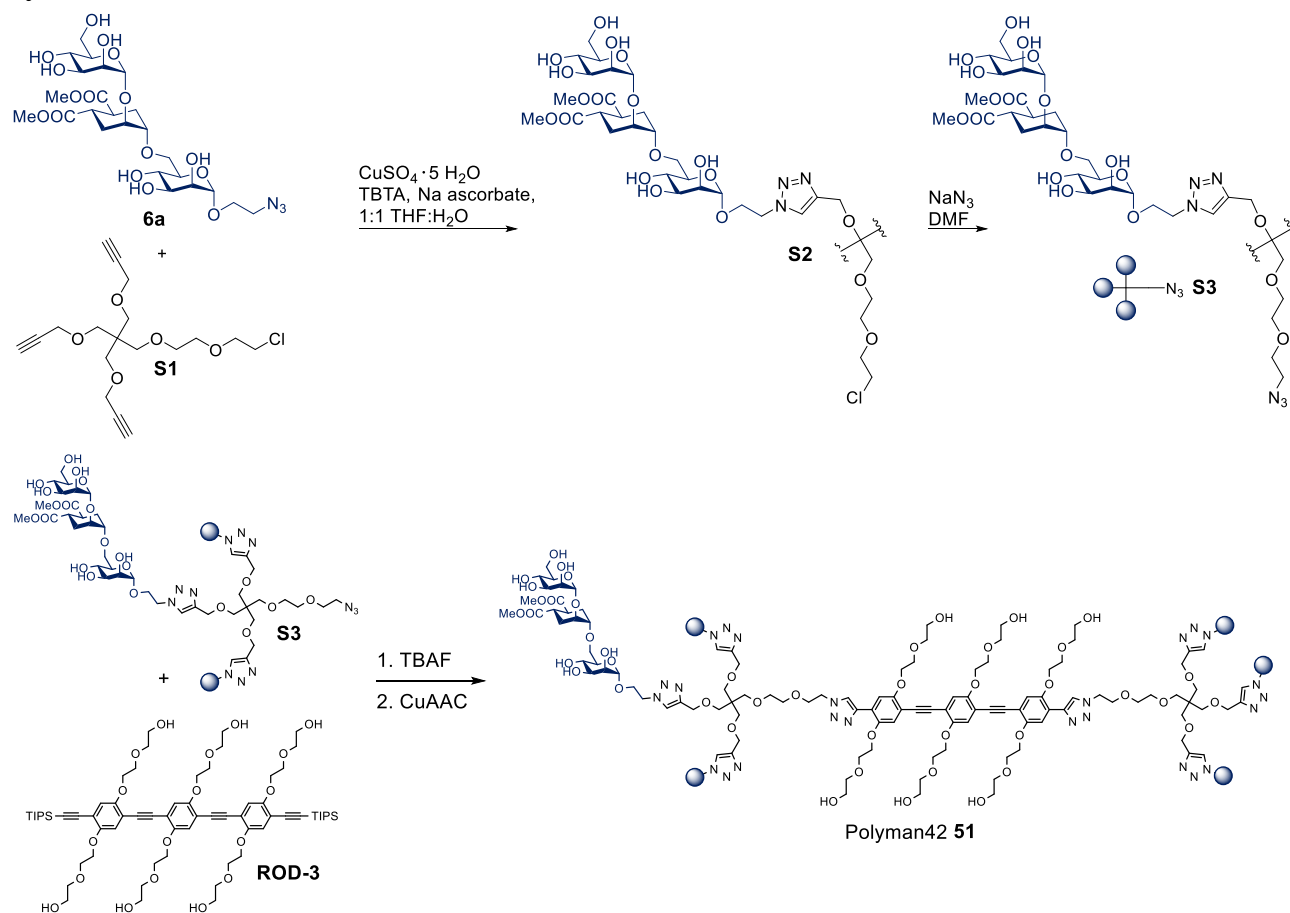


1.6.2.2 General procedure for the CuAAC reaction⁴²

In the optimized procedure for copper(I) catalyzed 1,3-dipolar cycloaddition, the starting materials and reagents were added to the reaction mixtures as solutions in water (degassed by bubbling with nitrogen) or THF (freshly distilled). Monovalent ligands **5a**, **6a** and dendrons **38**, **S3** (see below) with azide groups were added as solids. The reagents were added to the reaction in the following order: multivalent scaffold (1 eq. in THF), TBTA (1 eq. in THF), $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ (0.1 eq. in H_2O), sodium ascorbate (0.4 eq. in H_2O) and finally the azide derivative (1.1 eq. per alkyne). After the addition of all the reagents, the solvent ratio was adjusted to 1:1 by addition of THF and/or water ($c = \sim 0.03 \text{ M}$). The reactions were stirred under nitrogen atmosphere, and shielded from light. The reaction progress was followed by TLC (silica, eluent: *n*-hexane - EtOAc, 8:2 and C18, eluent: H_2O - MeOH, 1:1) or mass spectrometry (MALDI or ESI ionization). Usually, in order to achieve reaction completion, an additional 0.4 eq. of sodium ascorbate was added (2-4 h after reaction start). After reaction completion the mixtures were loaded directly on Sephadex LH-20 column ($\varnothing = 3 \text{ cm}$, height = 50 cm; eluent: MeOH) to purify the products by size exclusion chromatography. Copper residues were removed either by automated reverse phase chromatography (Biotage SNAP C18 cartridges, with gradient elution using $\text{H}_2\text{O}/\text{CH}_3\text{CN}$ or $\text{H}_2\text{O}/\text{MeOH}$ as eluents) or by dissolving the product in MeOH adding a metal scavenger (such as QuadrasilTM MP) and stirring for 5 min. The scavenger was filtered off through a cotton pad and the filtrate was concentrated to obtain the product.

1.6.2.3 Preparation and characterization of hexavalent Polyman42 51

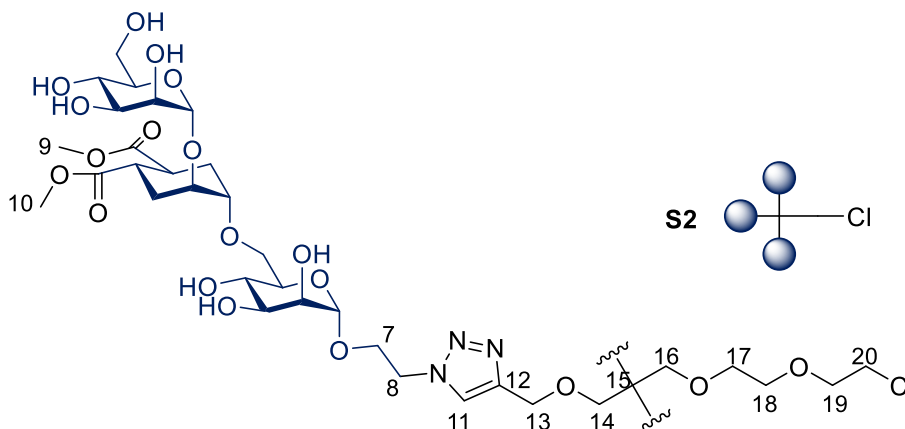
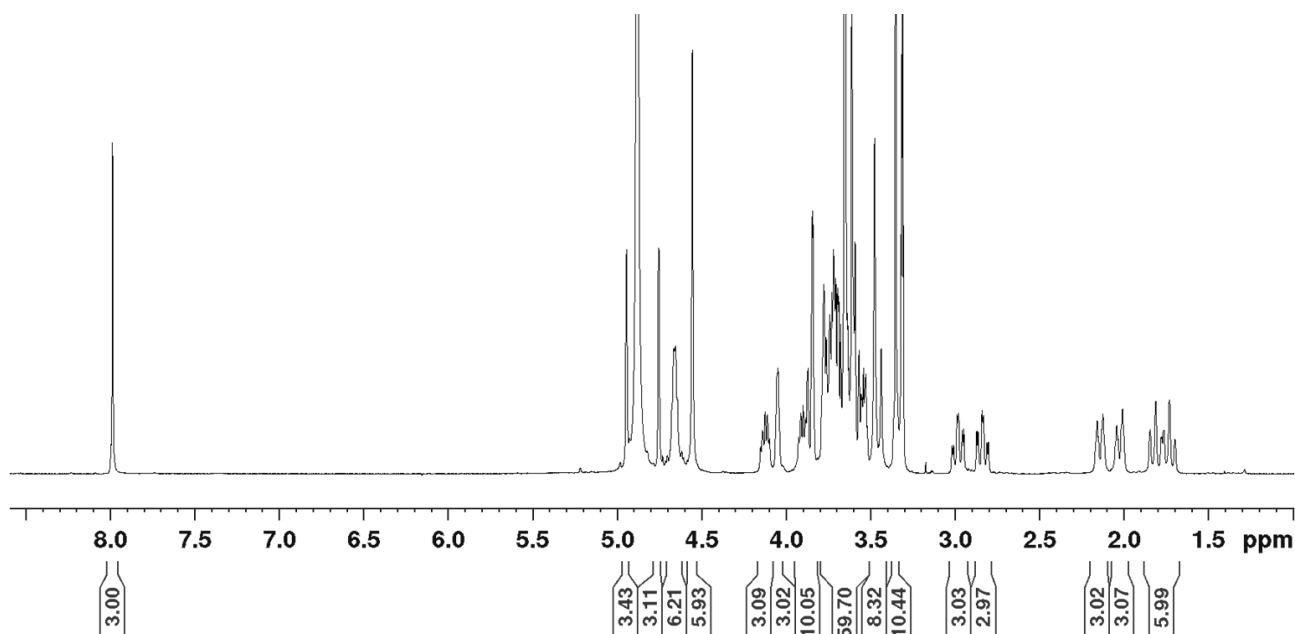
Synthetic route:

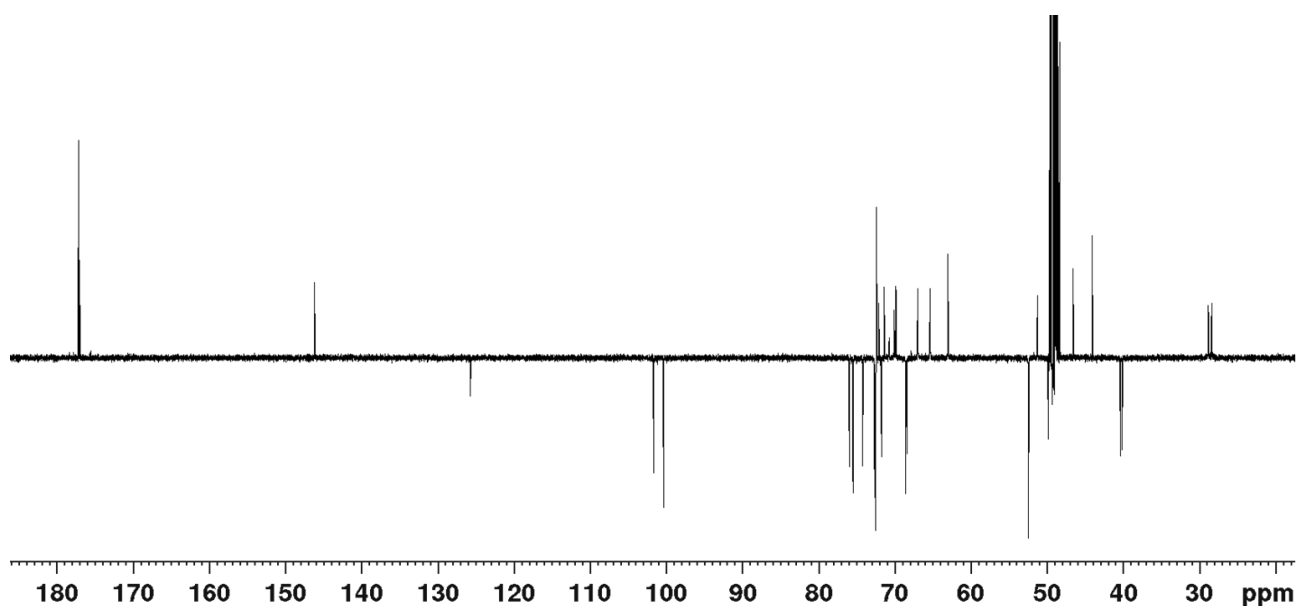


Scheme 1.15 – Synthetic strategy for Polyman42 **51**. Hexavalent glycodendrimer **51** can be obtained starting from the trivalent scaffold **S1** which is functionalized with three copies of monovalent ligand **6a** to give dendron **S2**. Substitution of the chloride atom with NaN_3 affords dendron **S3** which is connected to the linear **ROD-3** scaffold to achieve the desired Polyman42 **51**.

Synthesis of trivalent glycodendron S2

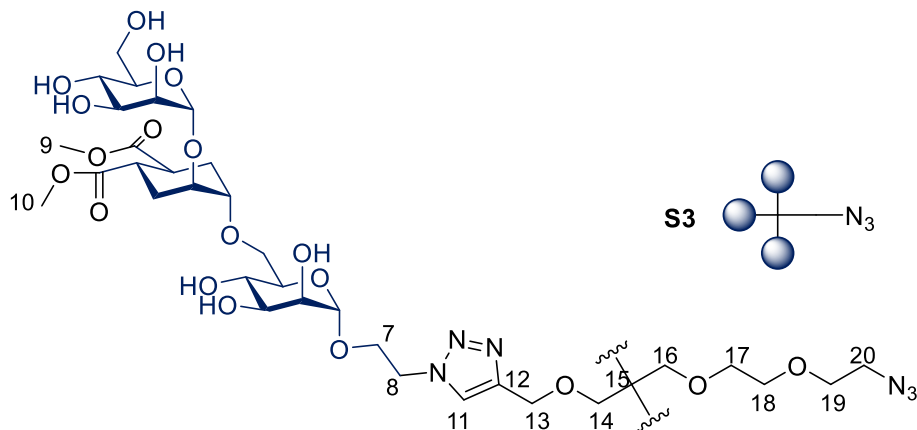
Prepared according to the general procedure for CuAAC starting from **S1**⁵⁹ and **6a**³⁸. Reaction time: 20 h. Yield: 91%. $[\alpha]_D^{18} = +45.3$ ($c = 0.93$ in MeOH). ¹H NMR (400 MHz, CD₃OD) δ (ppm): 7.99 (s, 3 H, 11-H), 4.95 (d, $J = 1.4$ Hz, 3 H, 1-H_M), 4.76 (d, $J = 1.4$ Hz, 3 H, 1-H_{M'}), 4.69-4.63 (m, 6 H, 8-H), 4.55 (s, 6 H, 13-H), 4.16-4.08 (m, 3 H, 7-H), 4.07-4.02 (m, 3 H, 2-H_D), 3.94-3.82 (m, 9 H, 7-H', 2-H_M, 6-H_{M'}), 3.79-3.67 (m, 20 H, 2-H_{M'}, 1-H_D, 6-H_{M'}, 3-H_M, 6-H_M, 6-H_{M'}, 19-H), 3.65 (s, 9 H, 9-H), 3.65 (s, 9 H, 10-H), 3.64-3.63 (m, 2 H, 20-H), 3.62-3.51 (m, 16 H, 17-H, 18-H, 4-H_M, 4-H_{M'}, 5-H_M, 5-H_{M'}), 3.47 (s, 6 H, 14-H), 3.43 (s, 2 H, 16-H), 3.37-3.34 (m, 3 H, 3-H_M), 2.98 (m, 3 H, 5-H_D), 2.89-2.78 (m, 3 H, 4-H_D), 2.19-2.09 (m, 3 H, 6-H_{Dax}), 2.07-1.97 (m, 3 H, 3-H_{Dax}), 1.87-1.77 (m, 3 H, 3-H_{Deq}), 1.77-1.67 (m, 3 H, 6-H_{Deq}). ¹³C NMR (100 MHz, CD₃OD) δ (ppm): 177.2, 177.0 (C=O), 146.2 (C12), 125.7 (C11), 101.7 (C1_{M'}), 100.4 (C1_M), 76.0 (C1_D), 75.5 (C5_{M'}), 74.2 (C3_{M'}), 72.7, 72.6, 72.5, 72.5 (C2_D, C2_M, C3_M, C5_M), 72.4 (C19), 72.1 (C17), 71.7 (C2_{M'}), 71.4 (C18), 70.7 (C16), 70.1 (C14), 69.8 (C6_M), 68.6, 68.4 (C4_M, C4_{M'}), 67.0 (C7), 65.3 (C13), 63.0 (C6_{M'}), 52.5, 52.4 (C9, C10), 51.3 (C8), 46.5 (C15), 44.1 (C20), 40.4 (C5_D), 40.1 (C4_D), 28.8 (C6_D), 28.4 (C3_D). MS (MALDI) m/z : calcd for C₉₀H₁₄₂ClN₉O₅₃ 2231.8; found 2235.1 [M+H]⁺, 2257.4 [M+Na]⁺, 2298.1 [M+Cu]⁺ (matrix DHB).

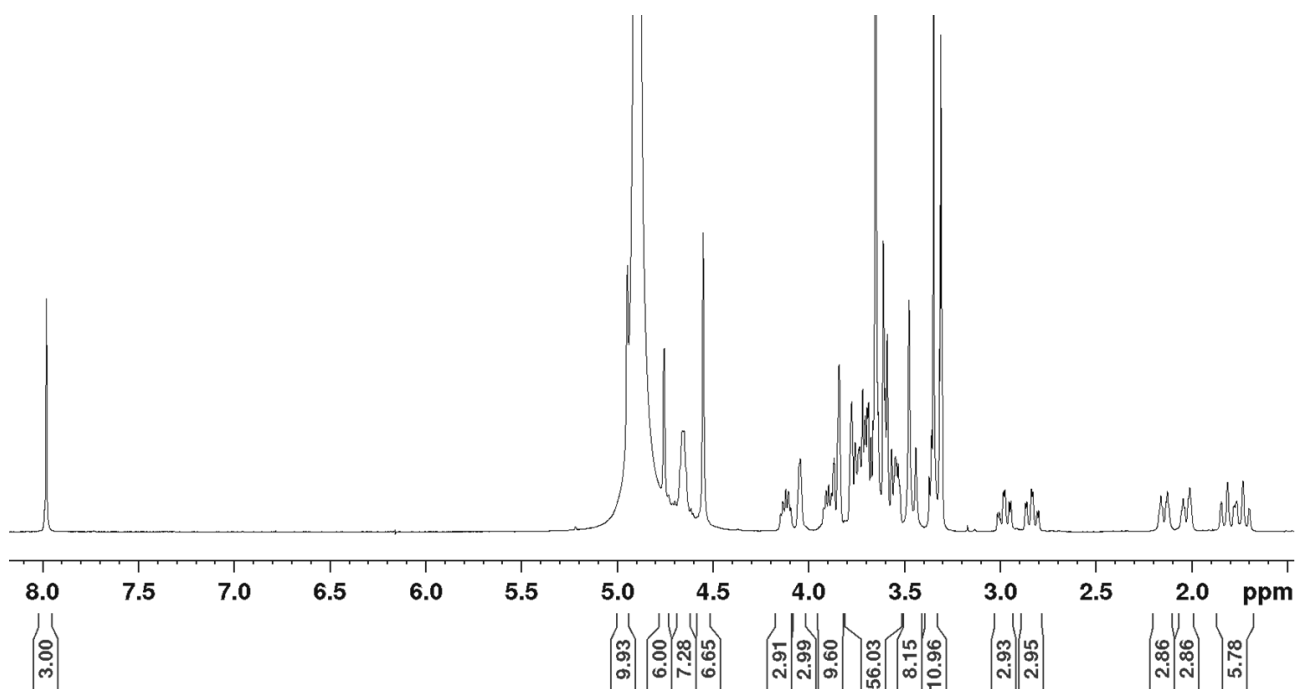
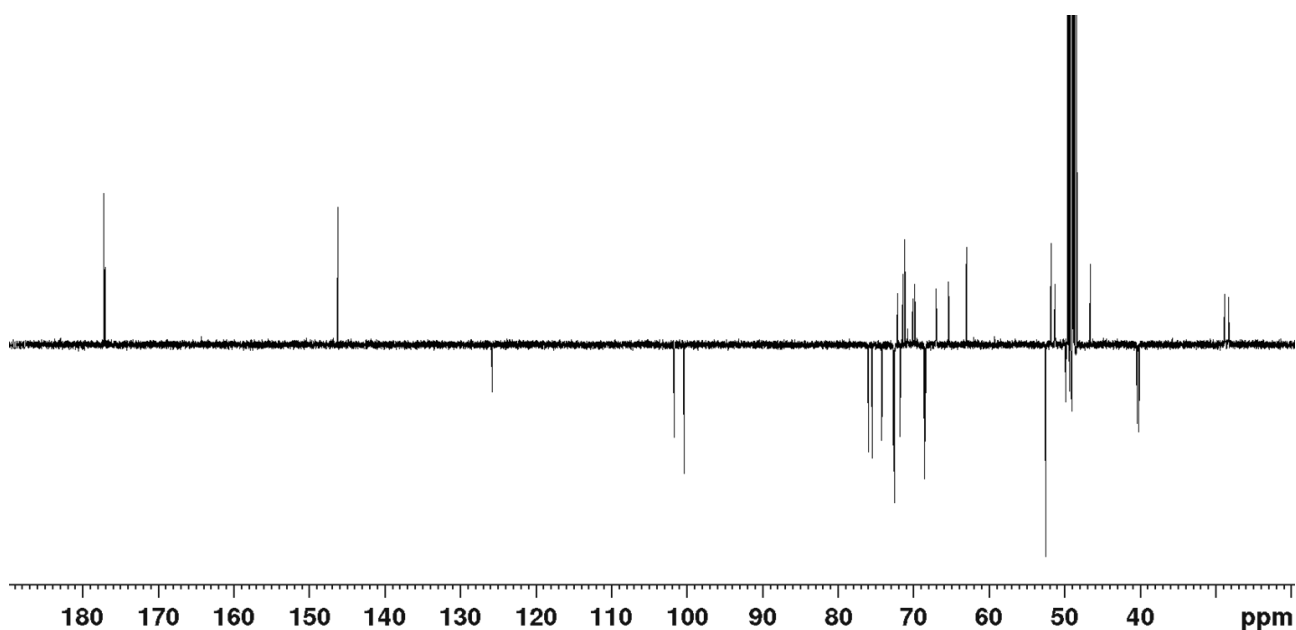
¹H NMR spectrum of S2 in CD₃OD (400 MHz)

¹³C NMR spectrum of S2 in CD₃OD (100 MHz)

Synthesis of trivalent glycodendron S3

To a solution of **S2** (75 mg, 33.56 μ mol) in dry DMF (1 mL) sodium azide (17.5 mg, 269 μ mol) was added. The reaction was stirred at 60 °C for 2.5 d monitoring by MALDI MS. The solvent was removed under reduced pressure and the resulting crude was purified by automated reverse phase chromatography (Biotage SNAP C18) to afford pure **S3** (70.8 mg, 94%). $[\alpha]_D^{18} = +45.3$ ($c = 0.93$ in MeOH). ¹H NMR (400 MHz, CD₃OD) δ (ppm): 7.98 (s, 3 H, 11-H), 4.95 (d, $J = 1.3$ Hz, 3 H, 1-H_M), 4.76 (d, $J = 1.3$ Hz, 3 H, 1-H_{M'}), 4.68-4.62 (m, 6 H, 8-H), 4.55 (s, 6 H, 13-H), 4.16-4.08 (m, 3 H, 7-H), 4.07-4.02 (m, 3 H, 2-H_D), 3.95-3.82 (m, 9 H, 7-H', 2-H_M, 6-H_M), 3.80-3.67 (m, 21 H, 2-H_{M'}, 1-H_D, 3-H_M, 3-H_{M'}, 6-H'_M, 6-H_{M'}, 6-H'_{M'}), 3.67-3.63 (m, 22 H, 9-H, 10-H, 19-H), 3.62-3.51 (m, 16 H, 4-H_M, 5-H_M, 4-H_{M'}, 5-H_{M'}, 17-H, 18-H), 3.48 (s, 6 H, 14-H), 3.44 (s, 2 H, 16-H), 3.39-3.33 (m, 2 H, 20-H), 3.04-2.92 (m, 3 H, 5-H_D), 2.90-2.76 (m, 3 H, 4-H_D), 2.21-2.10 (m, 3 H, 6-H_{Dax}), 2.07-1.98 (m, 3 H, 3-H_{Dax}), 1.87-1.77 (m, 3 H, 3-H_{Deq}), 1.77-1.66 (m, 3 H, 6-H_{Deq}). ¹³C NMR (100 MHz, CD₃OD) δ (ppm): 177.2, 177.1 (C=O), 146.2 (C12), 125.7 (C11), 101.7 (C1_M), 100.4 (C1_M), 76.0 (C1_D), 75.5 (C5_{M'}), 74.2 (C3_M), 72.7, 72.6, 72.5, 72.5 (C2_D, C2_M, C3_M, C5_M), 72.1 (C17), 71.7 (C2_{M'}), 71.4 (C18), 71.1 (C19), 70.8 (C16), 70.1 (C14), 69.8 (C6_M), 68.6, 68.4 (C4_M, C4_{M'}), 67.0 (C7), 65.3 (C13), 63.0 (C6_{M'}), 52.5, 52.5 (C9, C10), 51.8 (C20), 51.3 (C8), 46.6 (C15), 40.4 (C5_D), 40.1 (C4_D), 28.8 (C6_D), 28.4 (C3_D). MS (MALDI) m/z : calcd for C₉₀H₁₄₂N₁₂O₅₃ 2238.9; found 2264.0 [M+Na]⁺ (matrix DHB).

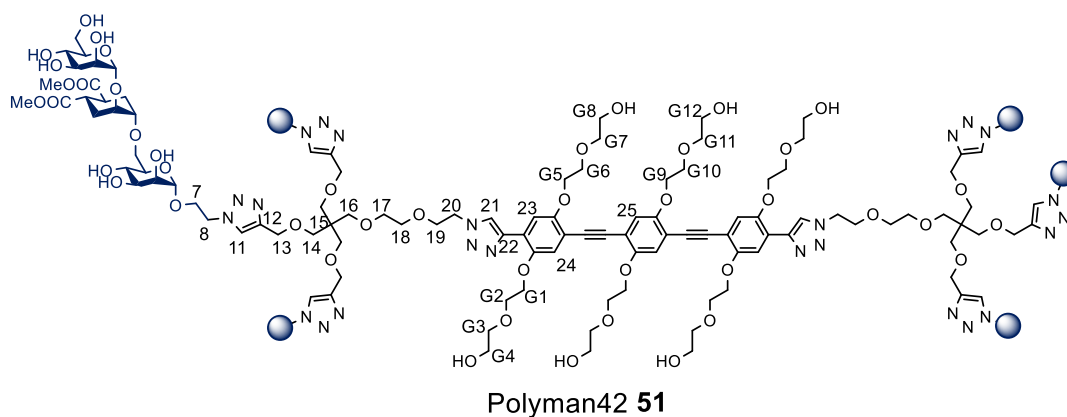


^1H NMR spectrum of S3 in CD_3OD (400 MHz) ^{13}C NMR spectrum of S3 in CD_3OD (100 MHz)

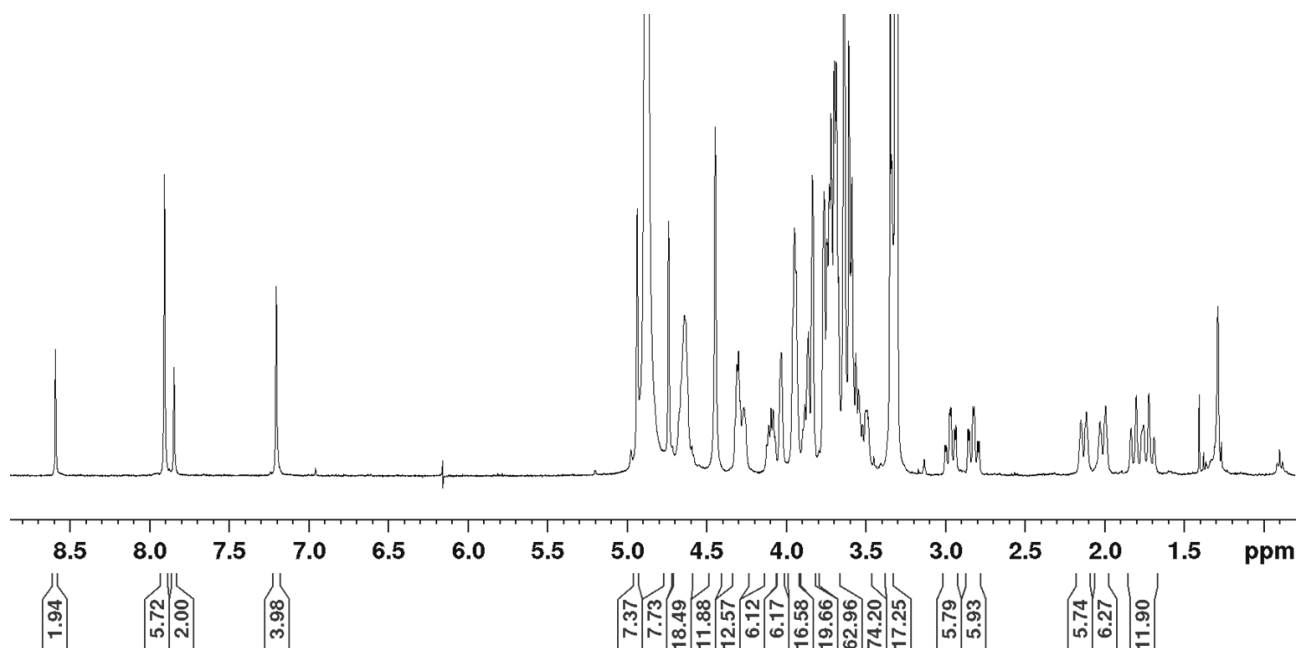
Synthesis of Polyman42 51

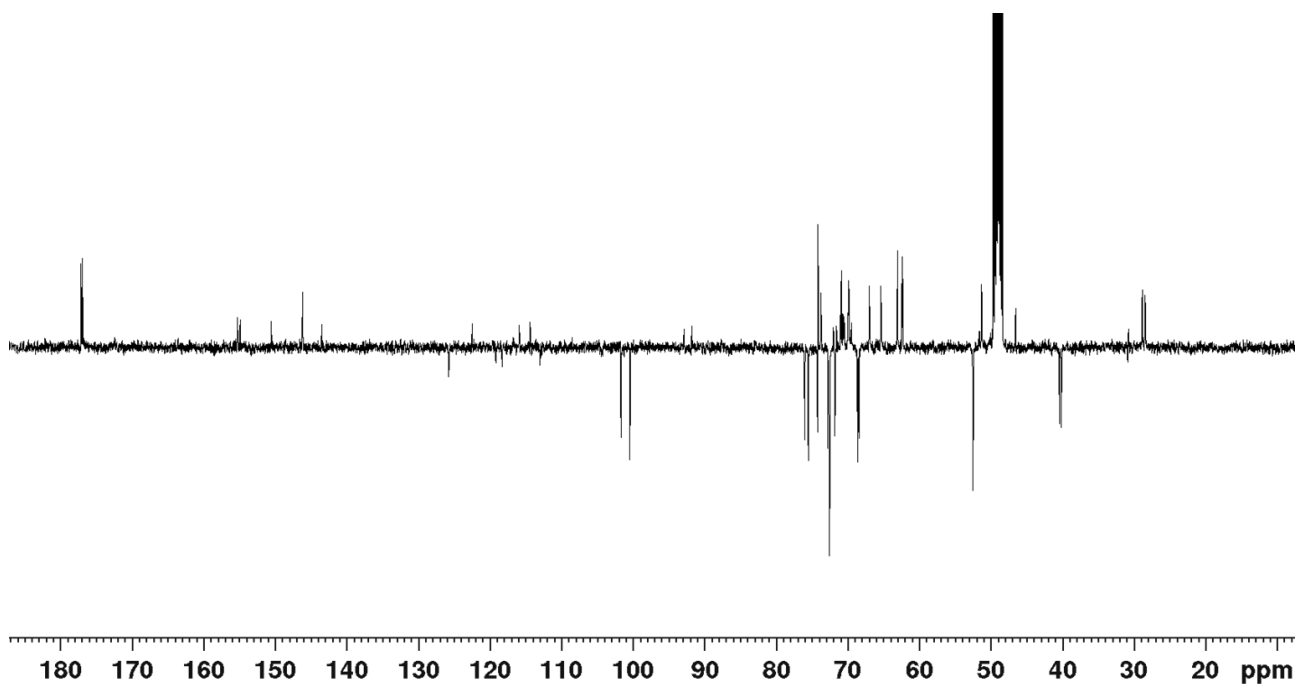
The rod-like scaffold **ROD-3⁶⁰** (10 mg, 7.91 μmol) was dissolved in freshly distilled THF (300 μL) under nitrogen atmosphere. TBAF (18 μL) was added as a 1 M solution in THF and the reaction was stirred at room temperature for 1 h. Complete deprotection was assessed by TLC analysis (eluent: DCM - MeOH, 9:1). A solution of TBTA (840 μg , 1.58 μmol) in distilled THF (84 μL) was added, followed by 29 μL of a solution of $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ (197 μg , 0.79 μmol) and 88 μL of a solution of sodium ascorbate (627 μg , 3.16 μmol) both in degassed H_2O . The mixture was stirred for 10 min and finally dendron **S3** (39 mg, 17.4 μmol) was added. The reaction was stirred at room temperature, under nitrogen atmosphere for 15 h. The complete conversion into the desired product was assessed by TLC analysis (reverse phase plates, eluent: MeOH - H_2O , 6:4) and by MALDI-TOF MS (matrix DHB). The solvent was evaporated and the crude was purified by automated

reverse phase chromatography (Biotage SNAP C18) obtaining pure Polyman42 **51** dendrimer (36.7 mg, 85%). $[\alpha]_D^{18} = +41.5$ ($c = 0.46$ in MeOH). $^1\text{H NMR}$ (400 MHz, CD_3OD) δ (ppm): 8.59 (s, 2 H, 21-H), 7.91 (s, 6 H, 11-H), 7.85 (s, 2 H, 25-H), 7.21 (s, 4 H, 23-H, 24-H), 4.94 (s, 6 H, 1-H_M), 4.74 (s, 6 H, 1-H_{M'}), 4.64 (bs, 16 H, 8-H, 20-H), 4.45 (s, 12 H, 13-H), 4.35-4.22 (m, 12 H, H_{G1}, H_{G5}, H_{G9}), 4.14-4.06 (m, 6 H, 7-H), 4.03 (s, 6 H, 2-H_D), 3.98-3.91 (m, 16 H, H_{G2}, H_{G6}, H_{G10}, 19-H), 3.90-3.81 (m, 18 H, 2-H_M, 6-H_M, 7'-H), 3.79-3.66 (m, 52 H, 2-H_{M'}, 1-H_D, 3-H_M, 6-H_{M'}, 6-H_{M'}, 6-H_{M'}, H_{G3}, H_{G4}, H_{G7}, H_{G8}, H_{G11}, H_{G12}), 3.64 (s, 36 H, 9-H, 10-H), 3.62-3.48 (m, 32 H, 4-H_M, 4-H_{M'}, 5-H_M, 5-H_{M'}, 17-H, 18-H), 3.36-3.33 (m, 10 H, 3-H_M, 14-H, 16-H), 2.97 (td, $J = 12.5$, 3.5 Hz, 6 H, 5-H_D), 2.82 (td, $J = 12.3$, 3.6 Hz, 6 H, 4-H_D), 2.20-2.08 (m, 6 H, 6-H_{Dax}), 2.07-1.96 (m, 6 H, 3-H_{Dax}), 1.86-1.64 (m, 12 H, 3-H_{Deq}, 6-H_{Deq}). $^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ (ppm): 177.22, 177.03 (C=O), 155.36, 154.99, 150.53, 146.16 (C12), 143.48 (C22), 125.73 (C11), 122.45, 121.94, 119.16 (C21), 118.28 (C23, C24), 115.92, 114.35, 113.00 (C25), 110.32, 108.53, 108.01, 101.67 (C1_M), 100.42 (C1_{M'}), 92.83, 91.76, 76.00 (C5_M), 75.50 (C5_{M'}), 74.21 (C3_M), 74.13, 73.74, 72.74, 72.59, 72.53, 72.52 (C1_D, C2_D, C2_M, C2_{M'}), 71.98 (C18), 71.77 (C3_M), 71.55 (C17), 71.05, 70.93, 70.87, 70.68, 70.49, 69.92, 69.83, 69.49, 68.59, 68.42 (C4_M, C4_{M'}), 66.97 (C7), 65.34 (C13), 63.03 (C6_M), 62.40, 62.38, 62.30, 52.51, 52.46 (C9, C10), 51.28 (C8, C20), 46.54 (C15), 40.40, 40.14 (C4_D, C5_D), 28.83, 28.43 (C3_D, C6_D). HRMS (ESI) m/z : calcd for $\text{C}_{230}\text{H}_{346}\text{N}_{24}\text{O}_{124}$ 5431.15975; found 1109.25359 $[\text{M}+5\text{Na}]^{5+}$, 1380.82980 $[\text{M}+4\text{Na}]^{4+}$. MS (MALDI) m/z : calcd for $\text{C}_{230}\text{H}_{346}\text{N}_{24}\text{O}_{124}$ 5431.2; found 5434.7 $[\text{M}+\text{H}]^+$ (matrix DHB).



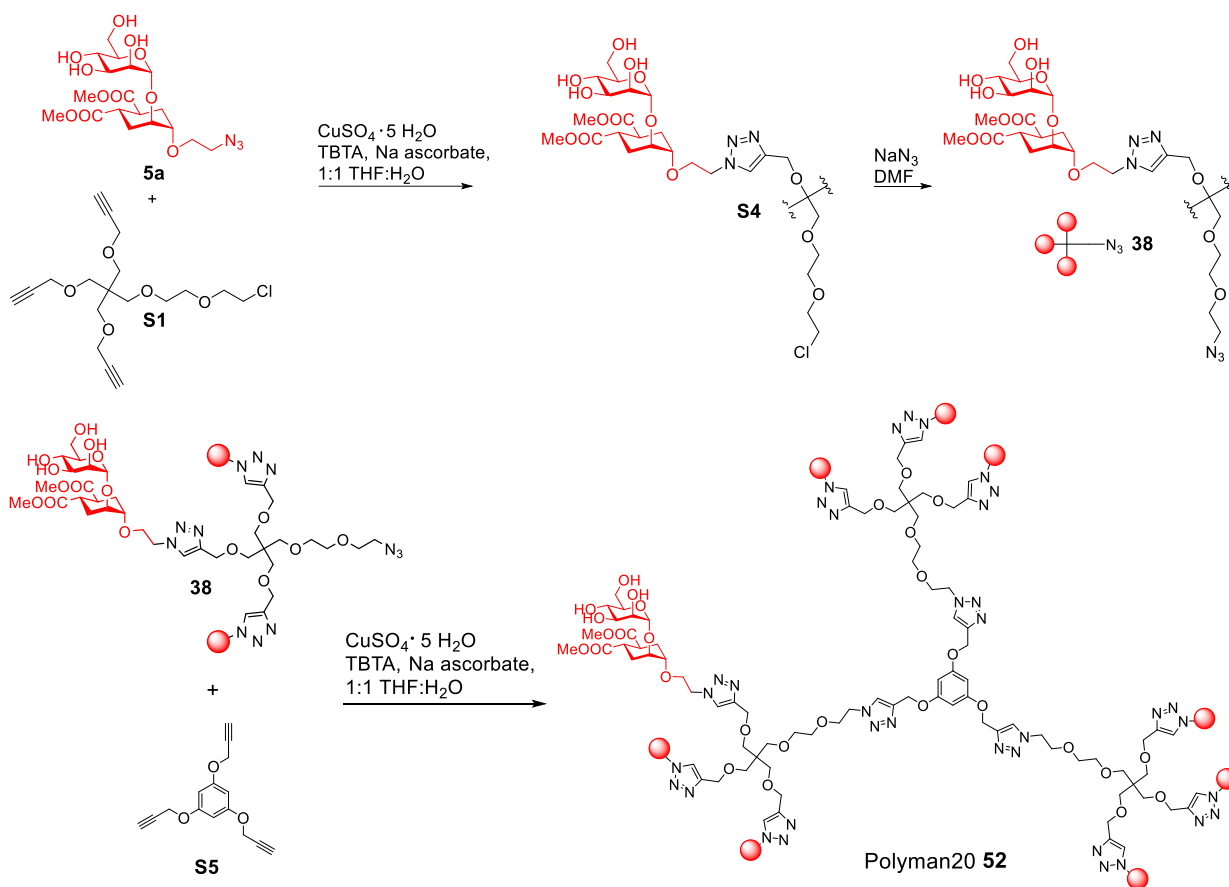
$^1\text{H NMR}$ spectrum of Polyman42 **51** in CD_3OD (400 MHz)



^{13}C NMR spectrum of Polyman42 51 in CD_3OD (100 MHz)

1.6.2.4 Preparation and characterization of nonavalent Polyman20 52

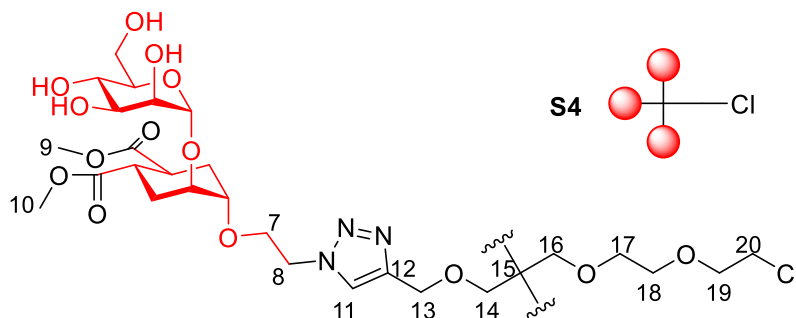
Synthetic route:



Scheme 1.15 – Synthetic strategy for Polyman20 52. Nonavalent glycodendrimer **52** can be obtained starting from the trivalent scaffold **S1** which is functionalized with three copies of monovalent ligand **5a** to give dendron **S4**. Substitution

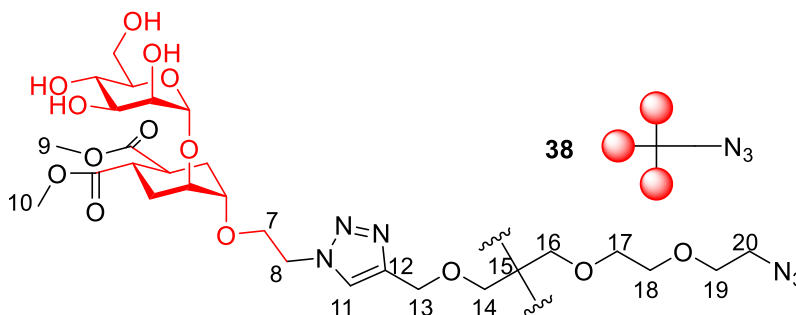
of the chloride atom with NaN_3 affords dendron **38** which is used in further CuAAC reaction with the trivalent propargyl derivative **S5** to give the desired Polyman20 **52**.

Synthesis of trivalent glycodendron **S4**⁶¹

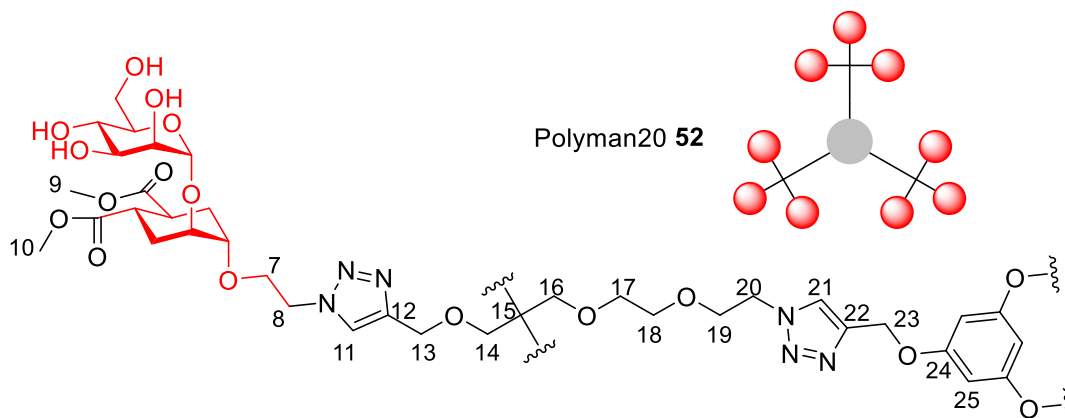


Prepared according to the general procedure for CuAAC starting from **S1**⁵⁹ and **5a**⁵⁶. Reaction time: 4 h. Yield: 88%. ¹H NMR (400 MHz, CD_3OD) δ (ppm): 7.98 (s, 3 H, 11-H), 4.90 (bs, 3 H, 1-H), 4.60 (t, $J = 5$ Hz, 6 H, 8-H), 4.55 (s, 6 H, 13-H), 4.00-3.79 (m, 15 H, 2-H, 6-H, 2- H_D , 7-H), 3.75-3.45 (m, 37 H, 6-H', 1- H_D , 3-H, 9-H, 10-H, 4-H, 5-H, 17-H, 18-H, 19-H, 20-H), 3.48 (s, 6 H, 14-H), 3.43 (s, 2 H, 16-H), 2.83-2.63 (m, 6 H, 4- H_D , 5- H_D), 2.04-1.92 (m, 6 H, 3- H_Deq , 6- H_Deq), 1.78-1.51 (m, 6 H, 3- H_Dax , 6- H_Dax). ¹³C NMR (100 MHz, CD_3OD) δ (ppm): 177.0, 176.8 (C=O), 146.4 (C12), 125.9 (C11), 100.7 (C1), 75.8 (C3), 75.8 (C5), 72.7 (C1_D), 72.7 (C19), 72.6 (C2), 72.3 (C2_D), 72.2 (C17), 71.5 (C16), 70.8 (C14), 70.2 (C18), 68.8 (C4), 68.5 (C7), 65.6 (C13), 63.2 (C6), 52.6 (C9, C10), 51.7 (C8), 46.8 (C15), 44.2 (C20), 40.4, 40.3 (C4_D, C5_D), 29.0, 28.5 (C3_D, C6_D). MS (ESI) m/z : calcd for $\text{C}_{72}\text{H}_{112}\text{ClN}_9\text{O}_{38}$ 1745.7; found 1768.7 $[\text{M}+\text{Na}]^+$.

Synthesis of trivalent glycodendron **38**⁶¹



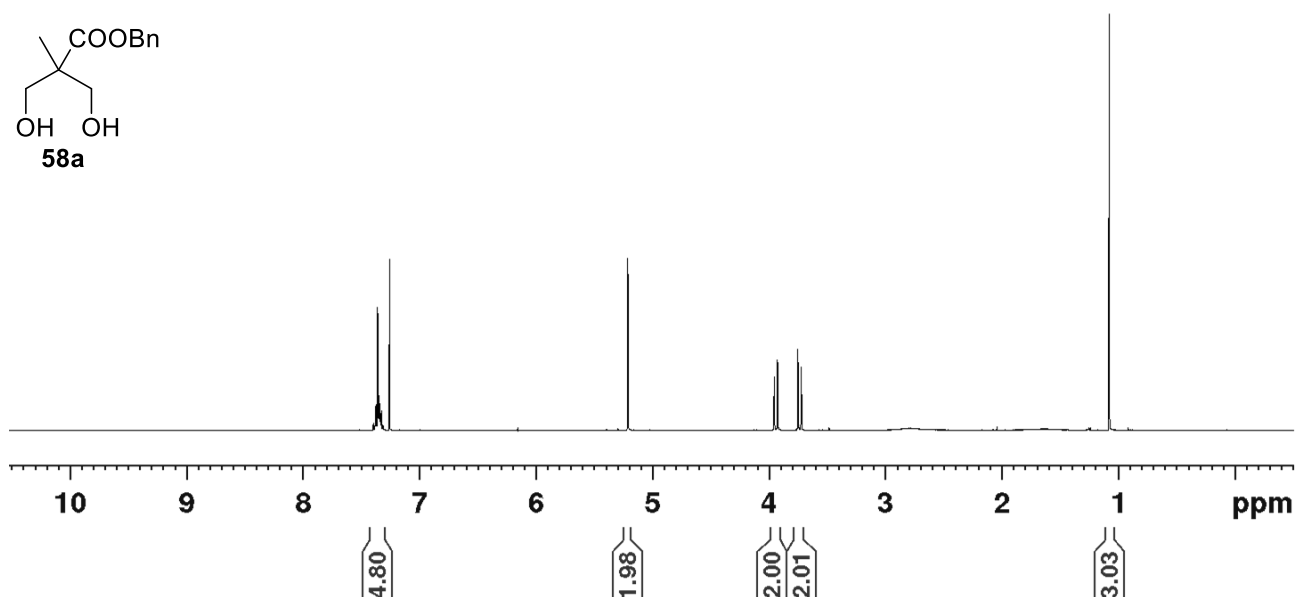
To a solution of **S4** (150 mg, 0.086 mmol) in dry DMF (1 mL) sodium azide (44.5 mg, 0.684 mmol) was added. The reaction was stirred at 60 °C for 4 d. The solvent was removed under reduced pressure and the resulting crude was purified by size exclusion chromatography using a Sephadex LH-20 column ($\varnothing = 3$ cm, height = 50 cm; eluent: MeOH) affording pure **38** (130 mg, 87%). ¹H NMR (400 MHz, CD_3OD) δ (ppm): 7.97 (s, 3 H, 11-H), 4.90 (bs, 3 H, 1-H), 4.60 (t, $J = 5$ Hz, 6 H, 8-H), 4.55 (s, 6 H, 13-H), 4.01-3.80 (m, 15 H, 2-H, 6-H, 2- H_D , 7-H), 3.72-3.49 (m, 35 H, 6-H', 1- H_D , 3-H, 9-H, 10-H, 4-H, 5-H, 17-H, 18-H, 19-H), 3.48 (s, 6 H, 14-H), 3.44 (s, 2 H, 16-H), 3.40-3.32 (m, 2 H, 20-H), 2.84-2.61 (m, 6 H, 4- H_D , 5- H_D), 2.05-1.95 (m, 6 H, 3- H_Deq , 6- H_Deq), 1.77-1.46 (m, 6 H, 3- H_Dax , 6- H_Dax). ¹³C NMR (100 MHz, CD_3OD) δ (ppm): 177.0, 176.8 (C=O), 146.4 (C12), 125.9 (C11), 100.7 (C1), 75.8 (C3), 75.7 (C5), 72.7 (C1_D), 72.5 (C2), 72.3 (C2_D), 72.3 (C17), 71.6 (C16), 71.3 (C19), 70.9 (C14), 70.2 (C18), 68.8 (C4), 68.5 (C7), 65.6 (C13), 63.2 (C6), 52.6 (C9, C10), 52.0 (C20), 51.7 (C8), 46.8 (C15), 40.4, 40.2 (C4_D, C5_D), 29.0, 28.5 (C3_D, C6_D). MS (ESI) m/z : calcd for $\text{C}_{72}\text{H}_{112}\text{N}_{12}\text{O}_{38}$ 1752.7; found 1775.5 $[\text{M}+\text{Na}]^+$.

Synthesis of Polyman20 52⁶¹

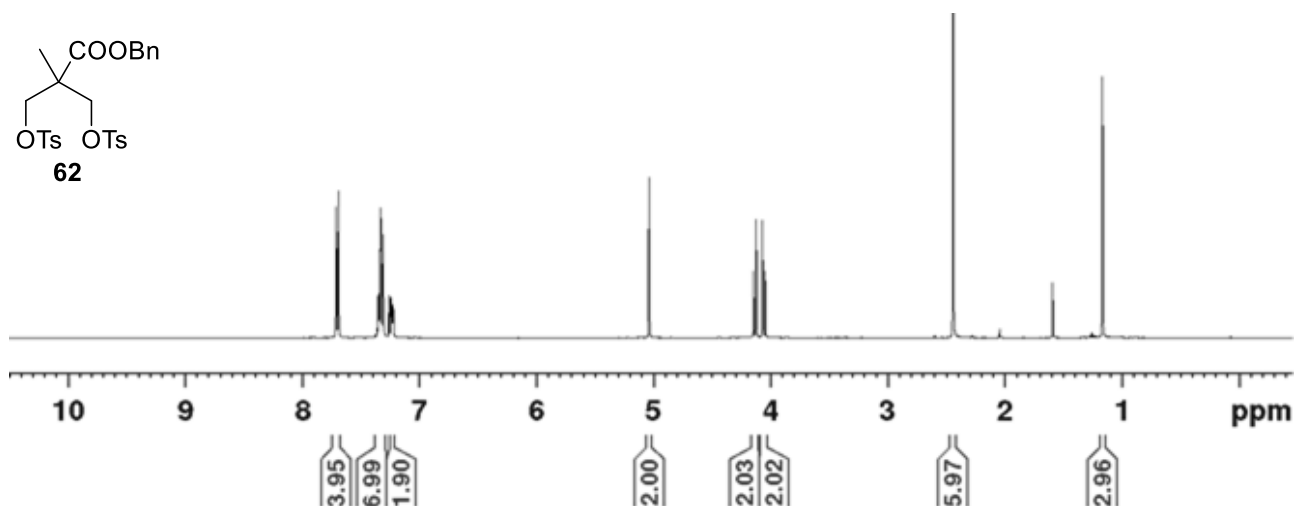
Prepared according to the general procedure for CuAAC starting from **S5**⁶² and **38**.⁴² Reaction time: 18 h. Yield: 54%. $[\alpha]_{\text{D}}^{25} = +22$ ($c = 0.22$ in MeOH). ¹H NMR (400 MHz, CD₃OD) δ (ppm): 8.10 (s, 3 H, 21-H), 7.94 (s, 9 H, 11-H), 6.33 (s, 3 H, 25-H), 5.11 (s, 6 H, 23-H), 4.90 (bs, 9 H, 1-H), 4.62-4.49 (m, 42 H, 8-H, 13-H, 20-H), 3.99-3.78 (m, 45 H, 7-H, 2-H, 6-H, 2-H_D), 3.71-3.46 (m, 117 H, 6-H', 1-H_D, 3-H, 9-H, 10-H, 4-H, 5-H, 17-H, 18-H, 19-H), 3.44 (s, 18 H, 14-H), 3.38 (s, 6 H, 16-H), 2.83-2.58 (m, 18 H, 4-H_D, 5-H_D), 2.05-1.90 (m, 18 H, 3-H_{Deq}, 6-H_{Deq}), 1.77-1.49 (m, 18 H, 3-H_{Dax}, 6-H_{Dax}). ¹³C NMR (100 MHz, CD₃OD) δ (ppm): 177.0, 176.8 (C=O), 146.3 (C12), 144.8 (C22), 126.4 (C21), 125.9 (C11), 100.7 (C1) 96.6 (C24), 75.8, 75.8 (C5, C3), 72.7 (C1_D), 72.6 (C2), 72.3 (C2_D); 72.2, 71.5, 70.8, 70.5, 70.2 (C14, C16, C17, C19, C18), 68.8 (C4), 68.5 (C7), 65.6 (C13), 63.2 (C23, C6), 52.7 (C9, C10), 51.7 (C8), 46.8 (C15), 40.4, 40.3 (C4_D, C5_D), 29.0, 28.5 (C3_D, C6_D). MS (MALDI) m/z : calcd for C₂₃₁H₃₄₈N₃₆O₁₁₇: 5500.2; found = 5502.5 [M+1]⁺, 5525.7 [M+Na]⁺ (matrix DHB).

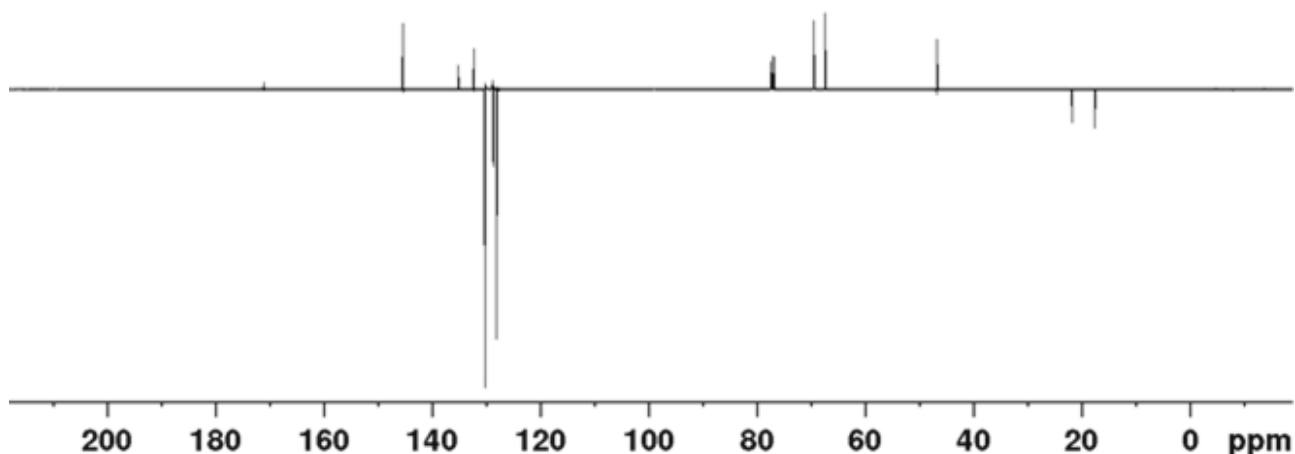
1.6.2.5 Synthesis of stabilized tetravalent glycodendrons **76a** and **77a**Synthesis of benzyl 2,2-bis(hydroxymethyl)propanoate **58a**⁴⁹

2,2-Bis(hydroxymethyl) propionic acid **60** (8.25 g, 61.5 mmol) was dissolved in dry DMF (61.5 mL) under nitrogen atmosphere. KOH (4.87 g, 73.8 mmol) was added to the solution at 0 °C and the reaction was stirred at 100 °C for 1.5 h. Benzyl bromide (8.0 mL, 67.6 mmol) was slowly added to the reaction, which was heated at 100 °C and stirred for 20 h. The formation of a white precipitate was observed. After having co-evaporated DMF with toluene, H₂O (40 mL) was added, the product was extracted with EtOAc (6x30 mL) and reunited organic phases were dried over anhydrous Na₂SO₄. After filtration and evaporation of the solvent the crude was purified by flash chromatography (eluent: *n*-hexane - EtOAc, 1:1; $R_f = 0.27$) obtaining pure **58a** as a white solid (9.94 g, 62%). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.40-7.32 (m, 5 H, *Ph*), 5.22 (s, 2 H, CH₂Ph), 3.95 (d, $J = 11.3$ Hz, 2 H, 1'-H, 2'-H), 3.72 (d, $J = 11.3$ Hz, 2 H, 1'-H', 2'-H'), 1.08 (s, 3 H, CH₃).

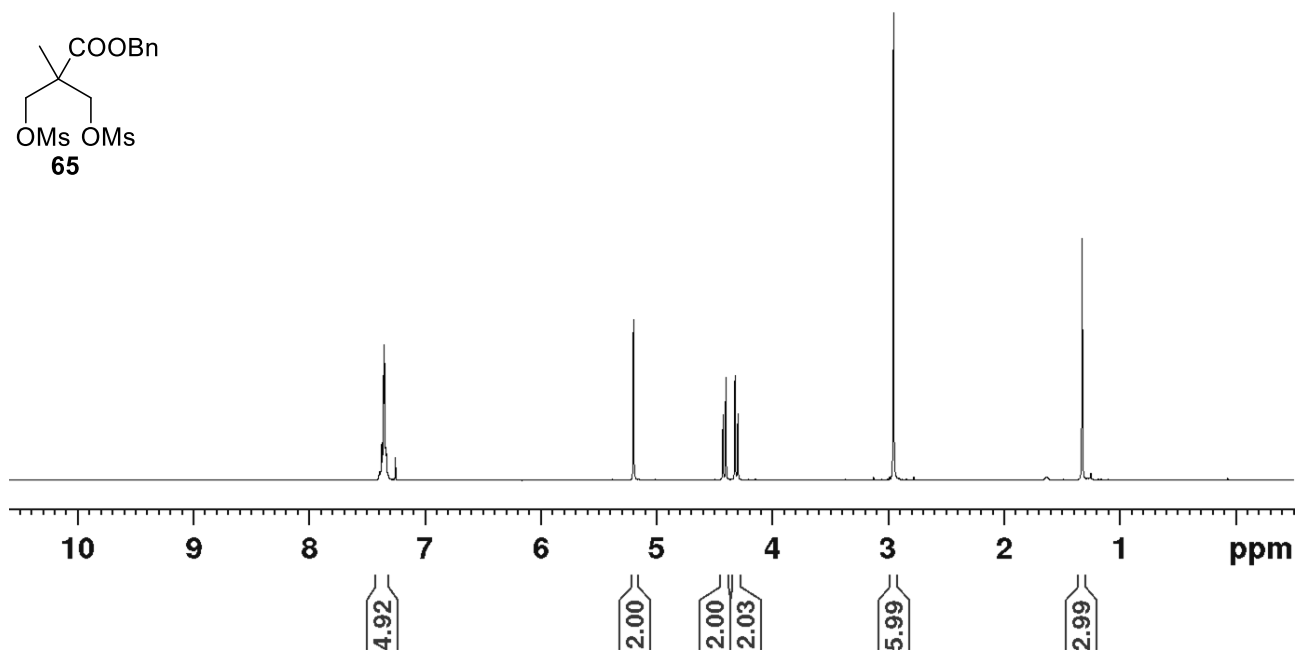
¹H NMR spectrum of **58a** in CDCl₃ (400 MHz)**Synthesis of benzyl 2,2-bis(*p*-toluensulfonylmethyl)propanoate **62****

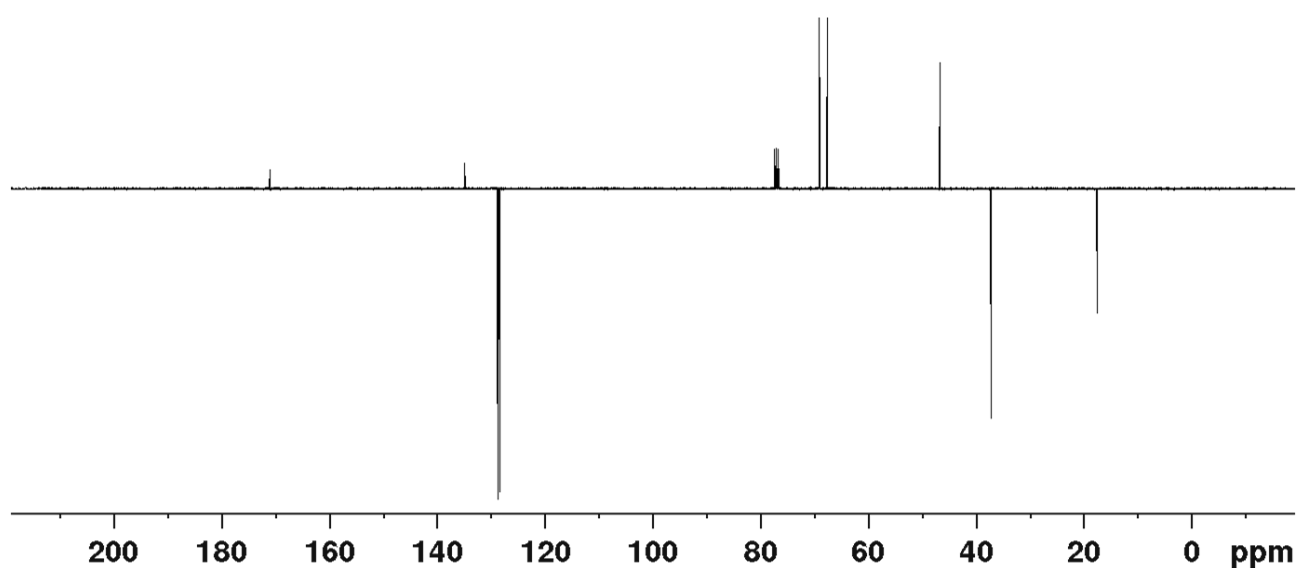
Benzyl 2,2-bis(hydroxymethyl)propanoate **58a** (5.0 g, 22.30mmol) was dissolved in dry pyridine (22 mL) under nitrogen atmosphere. Tosyl chloride (21.0 g, 110.0 mmol) was slowly added at 0 °C and the reaction was heated at 60 °C stirring for 3 h. After evaporation of pyridine, EtOAc was added (40 mL) and washed with HCl aqueous solution (2x30 mL). The aqueous phases were extracted with fresh EtOAc (2x40 mL) and finally the combined organic phases were washed with brine (40 mL) and dried over anhydrous Na₂SO₄. After filtration and evaporation of the solvent, the crude was purified with automated chromatography (gradient elution: from 90% *n*-hexane - 10% EtOAc to 40 % *n*-hexane - 60% EtOAc) obtaining pure **62** as a white solid (11.2 g, 94%); *R_f* = 0.24 (eluent: *n*-hexane - EtOAc, 4:1). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.70 (d, *J* = 8.3 Hz, 4 H, *Ar*), 7.35-7.31 (m, 7 H, *Ph* and *Ar*), 7.25-7.22 (m, 2 H, *Ph*), 5.04 (s, 2 H, CH₂Ph), 4.14 (d, *J* = 9.6 Hz, 2 H, 1'-H, 2'-H), 4.06 (d, *J* = 9.6 Hz, 2 H, 1'-H', 2'-H'), 2.44 (s, 6 H, ArCH₃), 1.17 (s, 3 H, CH₃). ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 171.0 (C=O), 145.3 (2 C, *Ar*), 135.1 (*Ph*), 132.1 (2 C, *Ar*), 130.1 (4 C, *Ar*), 128.7 (2 C, *Ph*), 128.6 (3 C, *Ph*), 127.1 (4 C, *Ar*), 69.4 (2 C, C1', C2'), 67.3 (CH₂Ph), 46.7 (C2), 21.8 (2 C, ArCH₃), 17.5 (CH₃). MS (ESI) *m/z*: calcd for C₂₆H₂₈O₈S₂ 532.1; found 555.4 [M+Na]⁺, 1086.9 [2M+Na]⁺.

¹H NMR spectrum of **62** in CDCl₃ (400 MHz)

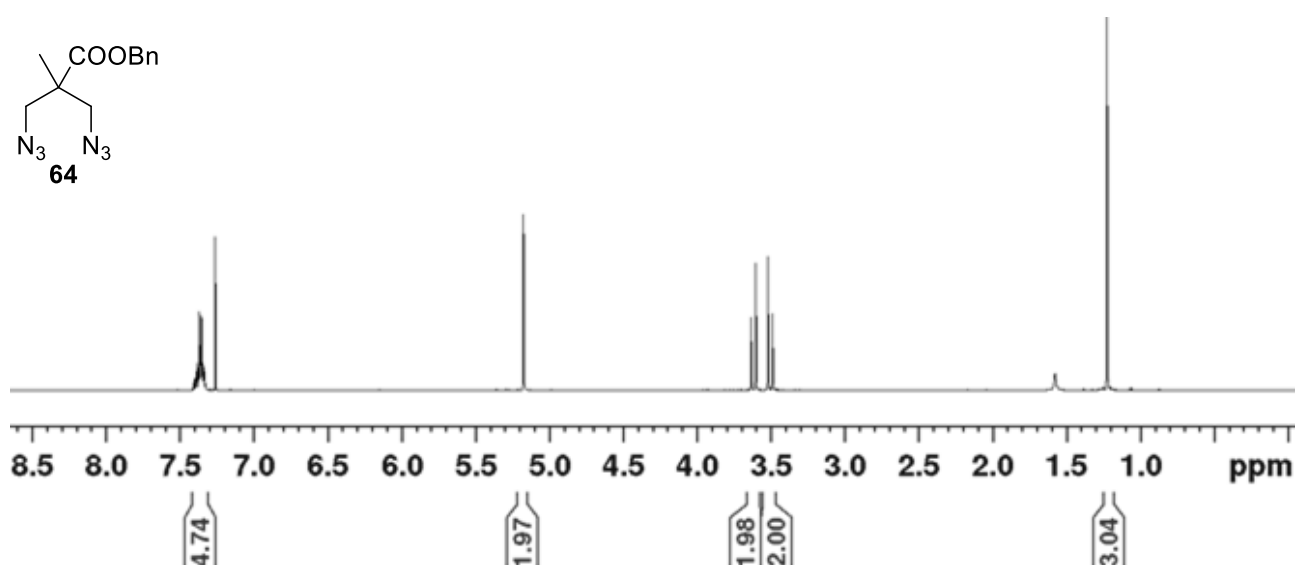
^{13}C NMR spectrum of **62** in CDCl_3 (100 MHz)Synthesis of benzyl 2,2-bis(methanesulfonylmethyl)propanoate **65**

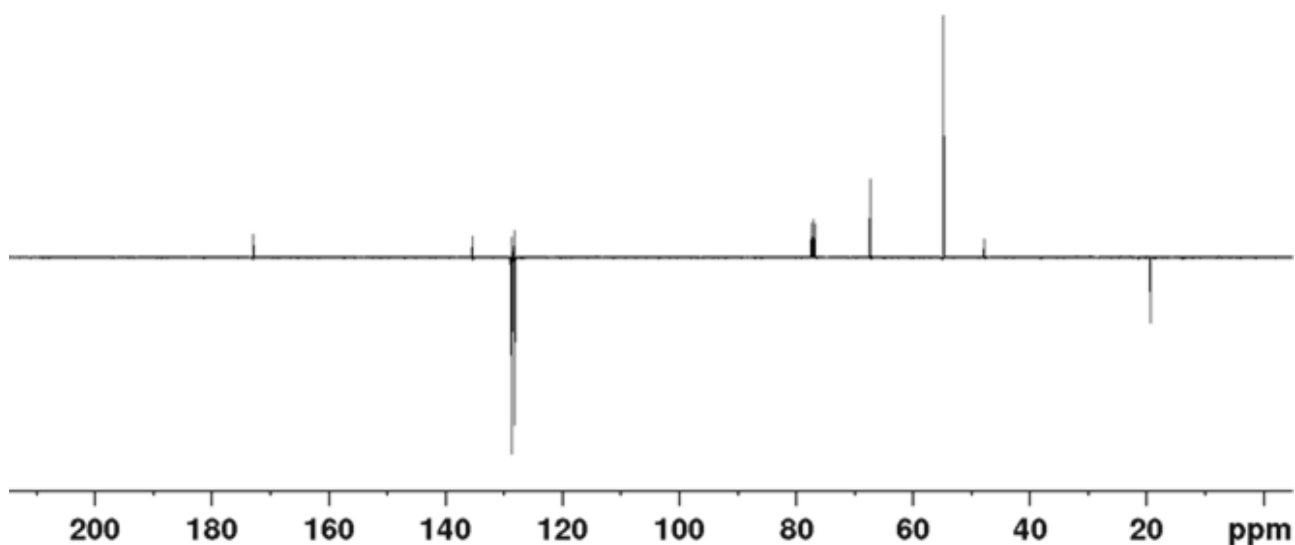
DIPEA (170 μL , 0.98 mmol) and DMAP (8 mg, 0.07 mmol) were added to a solution of **58a** (74 mg, 0.33 mmol) in dry DCM (500 μL) under stirring and nitrogen atmosphere. MsCl (75 μL , 0.98 mmol) was then added at 0 $^\circ\text{C}$. The reaction mixture was stirred at room temperature overnight. H_2O (10 mL) was added and the product was extracted in EtOAc (6x3 mL). The combined organic extracts were washed with brine (10 mL) and dried over anhydrous Na_2SO_4 . After filtration and evaporation of the solvent, the crude was purified by flash chromatography (eluent: *n*-hexane - EtOAc, 3:2; $R_f = 0.29$) and pure **65** (100 mg, 80%) was obtained as a yellow pale oil. ^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.40-7.33 (m, 5 H, *Ph*), 5.19 (s, 2 H, CH_2Ph), 4.42 (d, $J = 9.9$ Hz, 2 H, 1'-H, 2'-H), 4.31 (d, $J = 9.9$ Hz, 2 H, 1'-H', 2'-H'), 2.96 (s, 6 H, SO_3CH_3), 1.33 (s, 3 H, CH_3). ^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 171.1 ($\text{C}=\text{O}$), 134.9 (*Ph*), 128.8 (2 C, *Ph*), 128.7 (2 C, *Ph*), 128.5 (*Ph*), 69.1 (2 C, C1', C2'), 67.6 (CH_2Ph), 46.7 (C2), 37.2 (2 C, SO_3CH_3), 17.6 (CH_3). MS (ESI) m/z : calcd for $\text{C}_{14}\text{H}_{20}\text{O}_8\text{S}_2$ 380.060; found 403.018 $[\text{M}+\text{Na}]^+$.

 ^1H NMR spectrum of **65** in CDCl_3 (400 MHz)

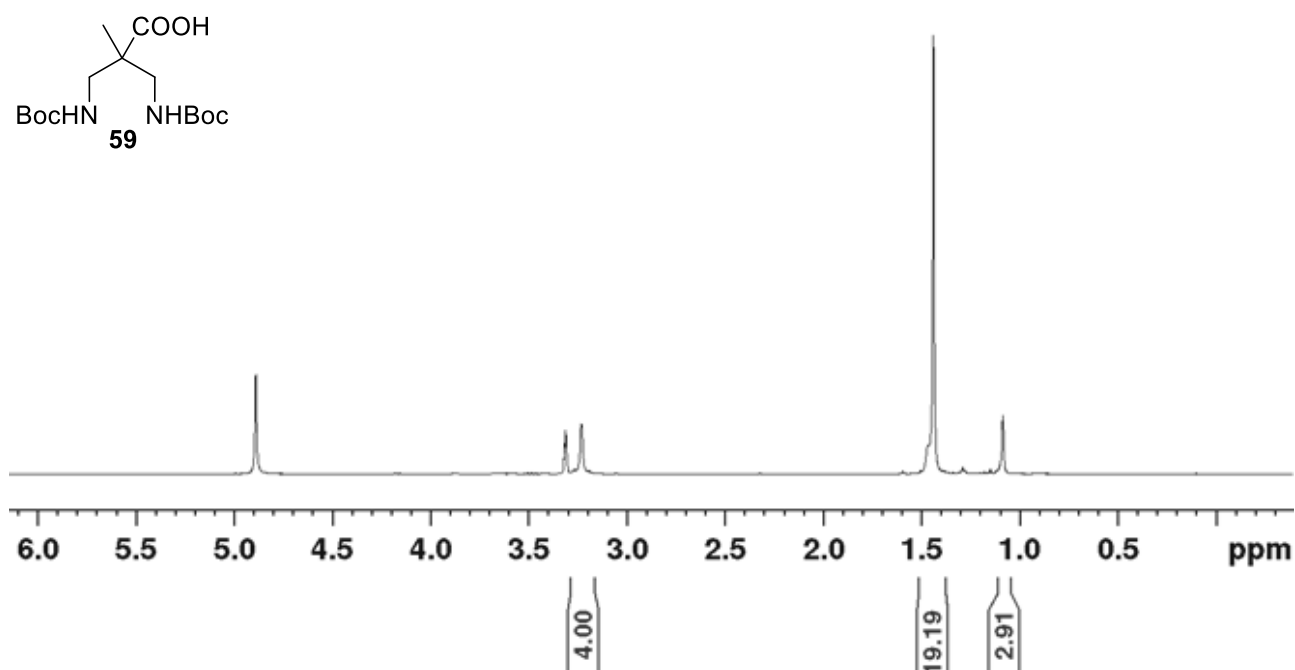
^{13}C NMR spectrum of **65** in CDCl_3 (100 MHz)Synthesis of benzyl 2,2-bis(azidomethyl)propanoate **64**

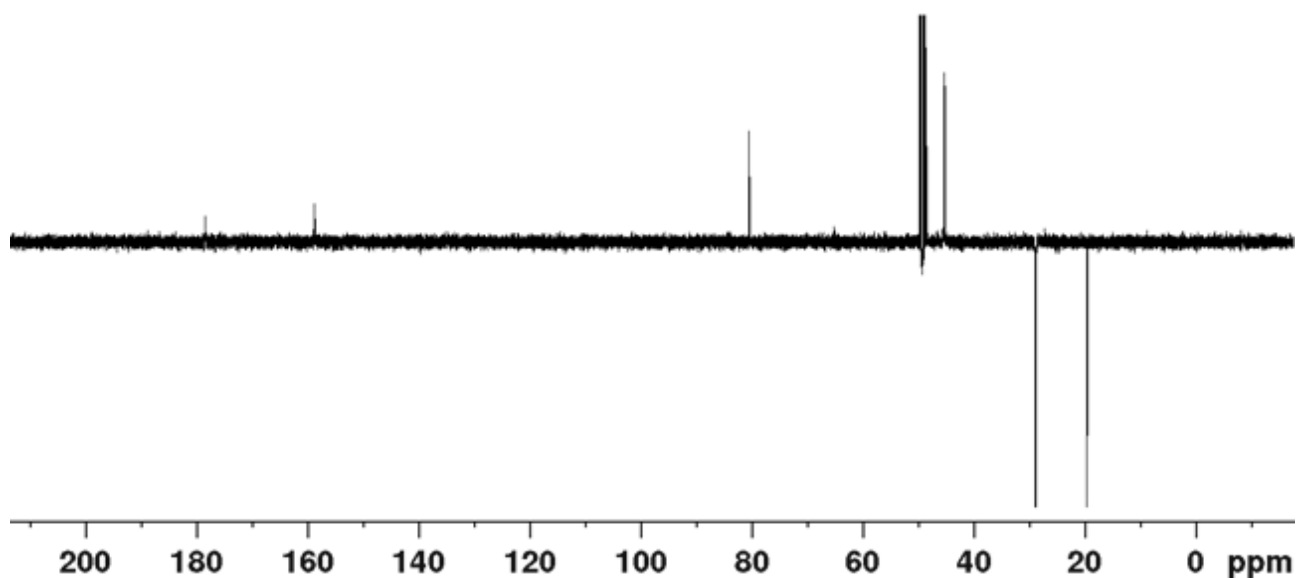
Sodium azide (9.9 g, 152 mmol) was added to a solution of benzyl 2,2-bis(*p*-toluensulfonylmethyl)propanoate **62** (10.1 g, 19.0 mmol) in dry DMA (38 mL). Bu_4NI (70.0 mg, 0.19 mmol) was added to the reaction that was stirred under nitrogen atmosphere at room temperature for 3 d until completion. DMA was evaporated, H_2O (30 mL) was added and the product was extracted with EtOAc (3x50 mL). Organic phases were dried on anhydrous Na_2SO_4 and, after filtration and evaporation of the solvent, the crude was purified with automated chromatography (gradient elution: from 100% *n*-hexane to 80% *n*-hexane - 20% EtOAc) obtaining pure **64** as a colourless liquid (4.7 g, 91%); $R_f = 0.75$ (eluent: *n*-hexane - EtOAc, 3:1). ^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.40-7.33 (m, 5 H, CH_2Ph), 5.18 (s, 2 H, CH_2Ph), 3.62 (d, $J = 12.2$ Hz, 2 H, 1'-H, 2'-H), 3.51 (d, $J = 12.2$ Hz, 2 H, 1'-H', 2'-H'), 1.22 (s, 3H, CH_3). ^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 173.1 ($\text{C}=\text{O}$), 135.4 (*Ph*), 128.8 (2 C, *Ph*), 128.6 (*Ph*), 128.3 (2 C, *Ph*), 67.3 (CH_2Ph), 54.9 (2 C, C1', C2'), 47.8 (C2), 19.4 (CH_3). MS (ESI) m/z : calcd for $\text{C}_{12}\text{H}_{14}\text{N}_6\text{O}_2$ 274.1; found 242.5 [$\text{M}-2\text{N}_2+\text{Na}$] $^+$, 297.1 [$\text{M}+\text{Na}$] $^+$.

 ^1H NMR spectrum of **64** in CDCl_3 (400 MHz)

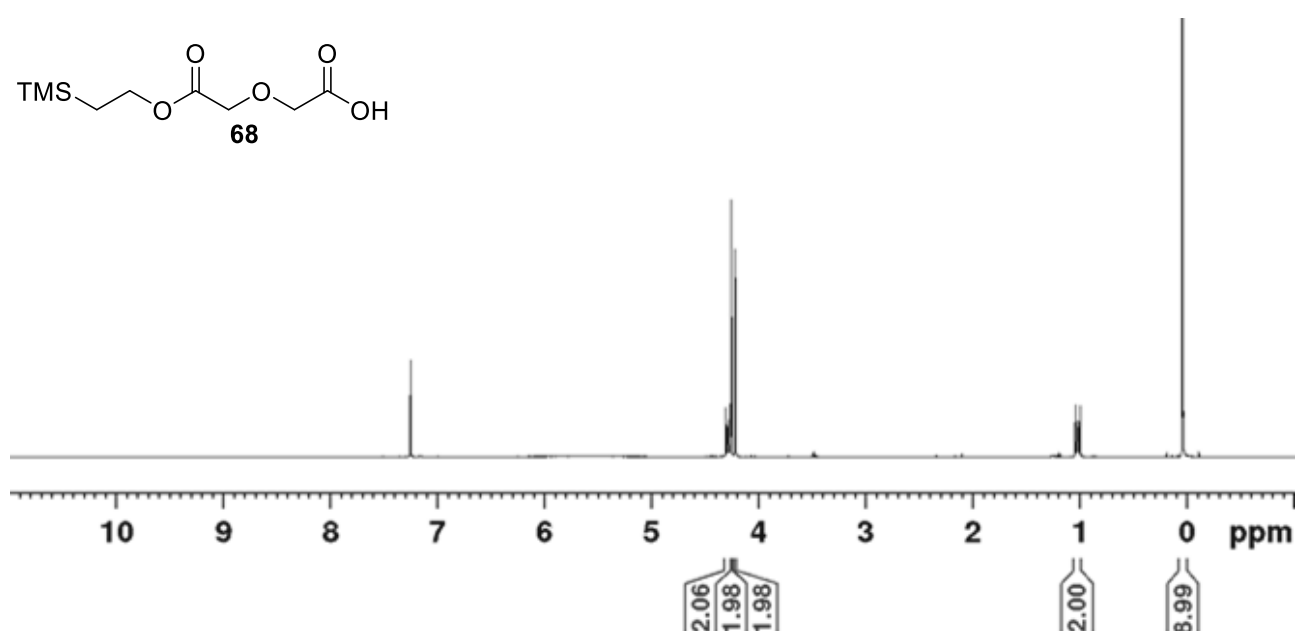
^{13}C NMR spectrum of **64** in CDCl_3 (100 MHz)Synthesis of 2,2-bis((*tert*-butoxycarbonyl)aminomethyl)propanoic acid **59**

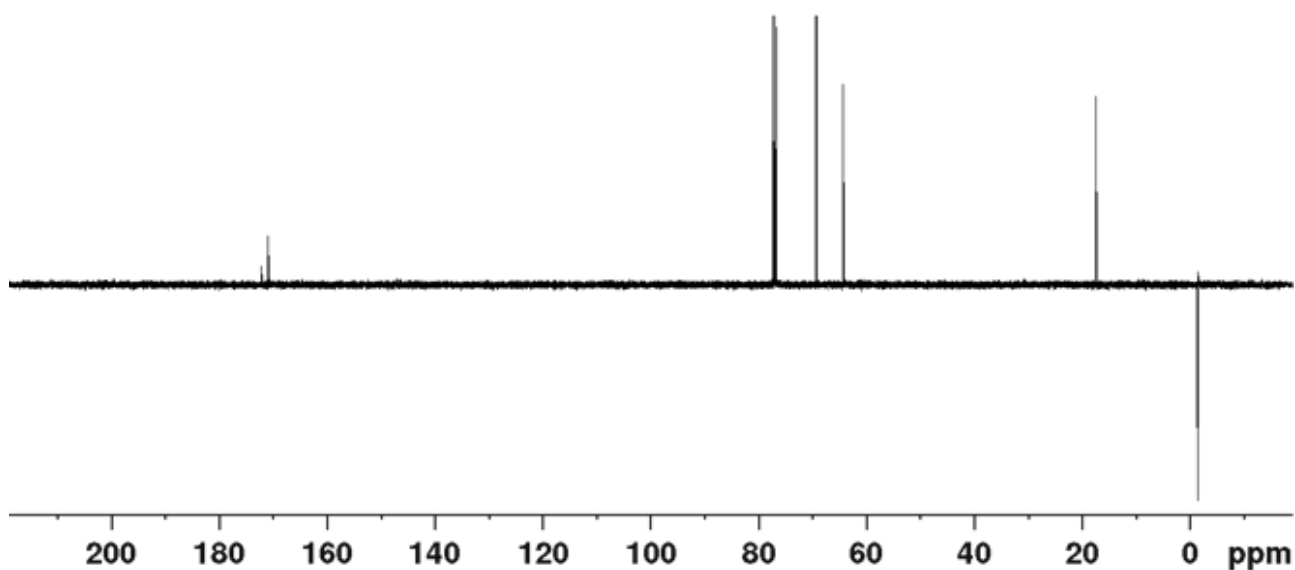
Benzyl 2,2-bis(azidomethyl)propanoate **64** (4.72 g, 17.2 mmol) was dissolved in dry MeOH (43 mL) under nitrogen atmosphere. Di-*tert*-butyl dicarbonate (9.40 g, 43.0 mmol) was added to the solution along with a catalytic amount (13% w/w) of 10% Pd/C. The reaction was stirred under hydrogen atmosphere at room temperature for 18 h until TLC (eluent: *n*-hexane - EtOAc, 3:1 + 1% AcOH) showed complete consumption of starting material. The reaction was then diluted in MeOH and filtered over a celite pad. After evaporation of the solvent the crude was purified with flash chromatography (eluent: *n*-hexane - EtOAc, 4:1 + 1% AcOH; $R_f = 0.22$) affording pure **59** as a foamy white solid (2.3 g, 50%). ^1H NMR (400 MHz, CD_3OD) δ (ppm): 3.23 (AB system, 4 H, CH_2NH), 1.44 (s, 18 H, $(\text{CH}_3)_3\text{CO}$), 1.09 (s, 3 H, CH_3). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 178.4 (C1), 158.7 (2 C, $(\text{CH}_3)_3\text{COC}=\text{O}$), 80.3 (2 C, $(\text{CH}_3)_3\text{CO}$), 49.3 (C2), 45.1 (2 C, CH_2NH), 28.7 (6 C, $(\text{CH}_3)_3\text{CO}$), 19.3 (CH_3). MS (ESI) m/z : calcd for $\text{C}_{15}\text{H}_{28}\text{N}_2\text{O}_6$ 332.195; found 355.214 $[\text{M}+\text{Na}]^+$, 687.449 $[\text{2M}+\text{Na}]^+$.

 ^1H NMR spectrum of **59** in CD_3OD (400 MHz)

¹³C NMR spectrum of **59** in CD₃OD (100 MHz)Synthesis of compound **68**

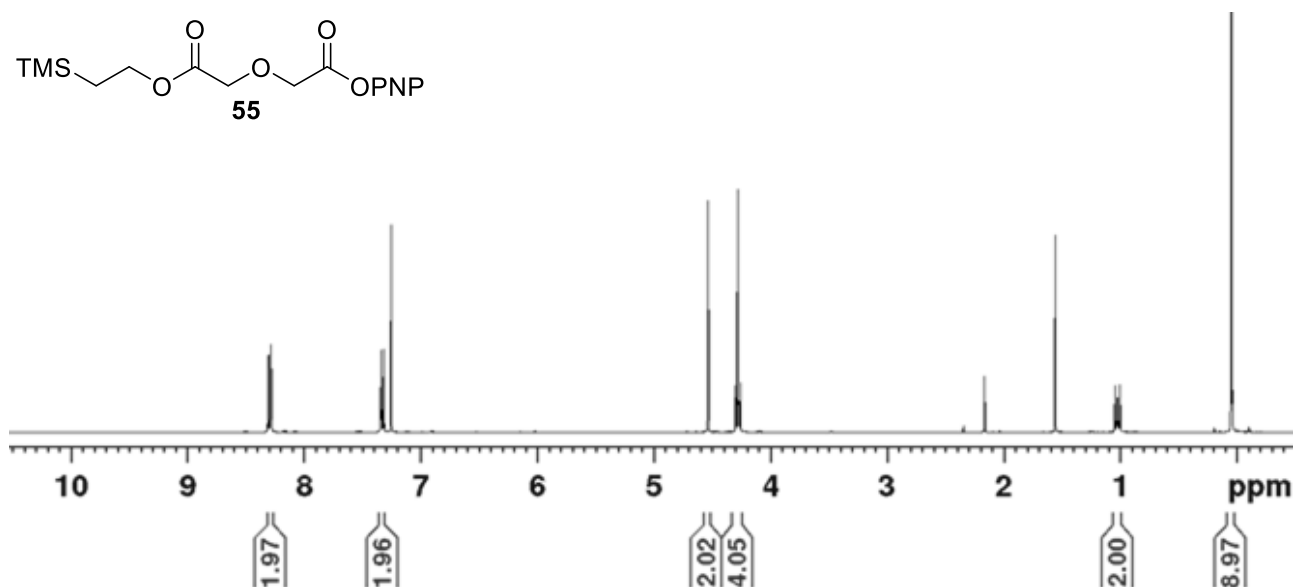
To a stirred solution of diglycolic anhydride **56** (90% purity) (505.5 mg, 4.36 mmol) in dry DCM (8.7 mL) freshly distilled Et₃N (600 μL, 4.36 mmol) and 2-(trimethylsilyl)ethanol **67** (680 μL, 4.80 mmol) were added at 0 °C under nitrogen atmosphere. The reaction was stirred for 22 h at room temperature. DCM was evaporated, an HCl aqueous solution (8 mL, 0.6 M) was added and the product was extracted with EtOAc (4x12 mL). Collected organic phases were washed with brine (12 mL) and dried over anhydrous Na₂SO₄. The crude was purified with flash chromatography (eluent: *n*-hexane - EtOAc, 2:1 + 1% AcOH; R_f = 0.24) affording a colourless oil as a pure product **68** (552.3 mg, 60%). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 4.31-4.27 (m, 2 H, CH₂CH₂TMS), 4.26 (s, 2 H, CH₂COOCH₂), 4.23 (s, 2 H, CH₂COOH), 1.05-1.01 (m, 2 H, CH₂CH₂TMS), 0.05 (s, 9 H, (CH₃)₃Si). ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 172.3 (C=O), 171.0 (C=O), 69.4 (CH₂COOCH₂), 69.3 (CH₂COOH), 64.4 (CH₂CH₂TMS), 17.5 (CH₂CH₂TMS), -1.4 (3 C, (CH₃)₃Si).

¹H NMR spectrum of **68** in CDCl₃ (400 MHz)

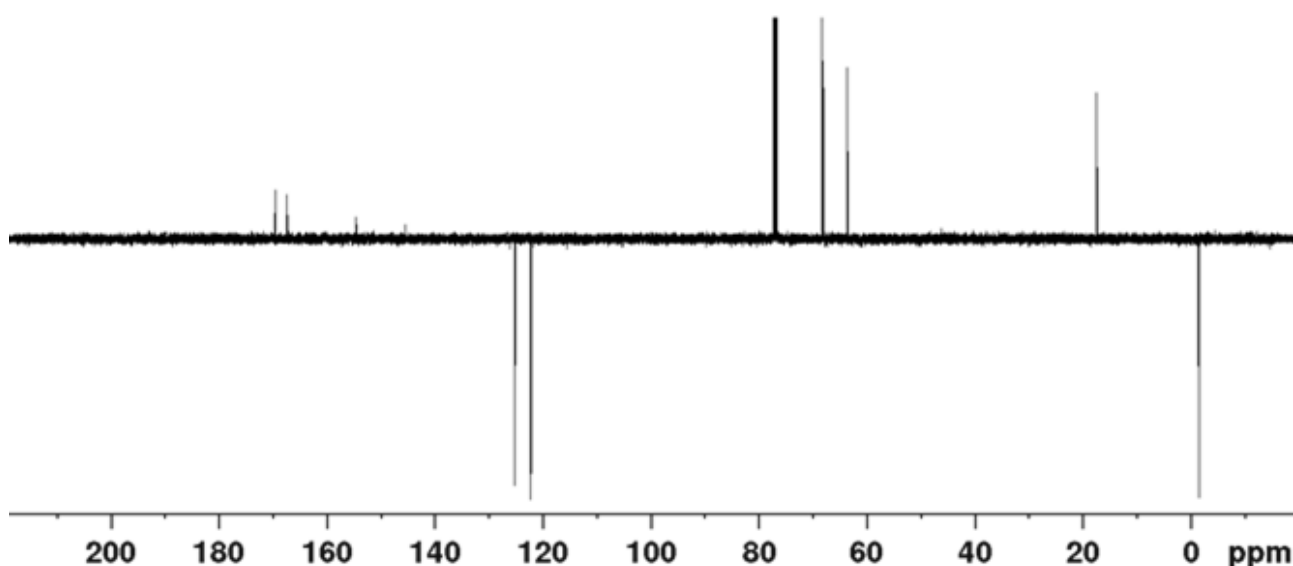
¹³C NMR spectrum of **68** in CDCl₃ (100 MHz)**Synthesis of compound 55**

The carboxylic acid **68** (225.0 mg, 0.96 mmol) was dissolved in dry THF (9.6 mL) and stirred under nitrogen atmosphere. Upon activation adding EDC hydrochloride (220.0 mg, 1.15 mmol) at 0 °C and stirring for 10 min, 4-nitrophenol **69** (150.0 mg, 1.10 mmol) and DIPEA (50 μL, 0.29 mmol) were added. The reaction was allowed to reach room temperature and stirred for 2 h until complete conversion. THF was evaporated and the crude was diluted in EtOAc (50 mL). The organic phase was washed with a 0.5 M HCl aqueous solution (2x24mL), with cold saturated NaHCO₃ aqueous solution (2x15 mL), with H₂O (2x 15 mL) and was finally dried over anhydrous Na₂SO₄. The crude consisted of the desired activated ester **55** (71% by ¹H NMR) and 4-nitrophenol **69** as the only impurity and could be employed in the next step without further purification. For characterisation purposes automated chromatography (gradient elution: from 90% *n*-hexane - 10% EtOAc to 50% *n*-hexane - 50% EtOAc) of the mixture afforded pure **55** as a white waxy solid (6.0 g, 56%); *R_f* = 0.56 (eluent: *n*-hexane - EtOAc, 2:1 + 1% AcOH). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 8.31-8.27 (m, 2 H, *Ar*), 7.36-7.32 (m, 2 H, *Ar*), 4.54 (s, 2 H, CH₂COOAr), 4.31-4.27 (m, 2 H, CH₂CH₂TMS), 4.29 (s, 2 H, CH₂COOCH₂), 1.05-1.01 (m, 2 H, CH₂CH₂TMS), 0.05 (s, 9 H, (CH₃)₃Si). ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 169.8 (C=O), 167.6 (C=O), 154.8 (*Ar*), 145.7 (*Ar*), 125.5 (2 C, *Ar*), 122.4 (2 C, *Ar*), 68.5 (CH₂COOAr), 68.2 (CH₂COOCH₂), 63.8 (CH₂CH₂TMS), 17.6 (CH₂CH₂TMS), -1.4 (3 C, (CH₃)₃Si). MS (ESI) *m/z*: calcd for C₁₅H₂₁NO₇Si 355.1; found 378.0 [M+Na]⁺, 732.7 [2M+Na]⁺.

^1H NMR spectrum of **55** in CDCl_3 (400 MHz)



^{13}C NMR spectrum of **55** in CDCl_3 (100 MHz)

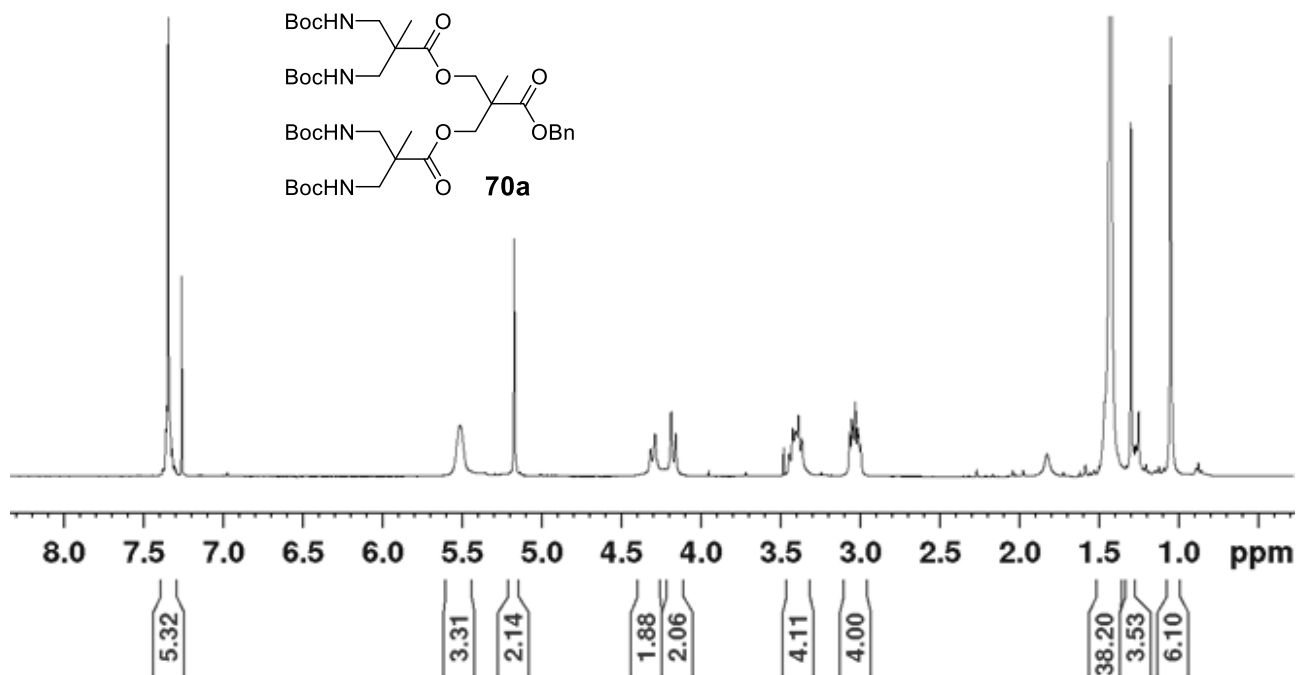


Synthesis of compound **70a**

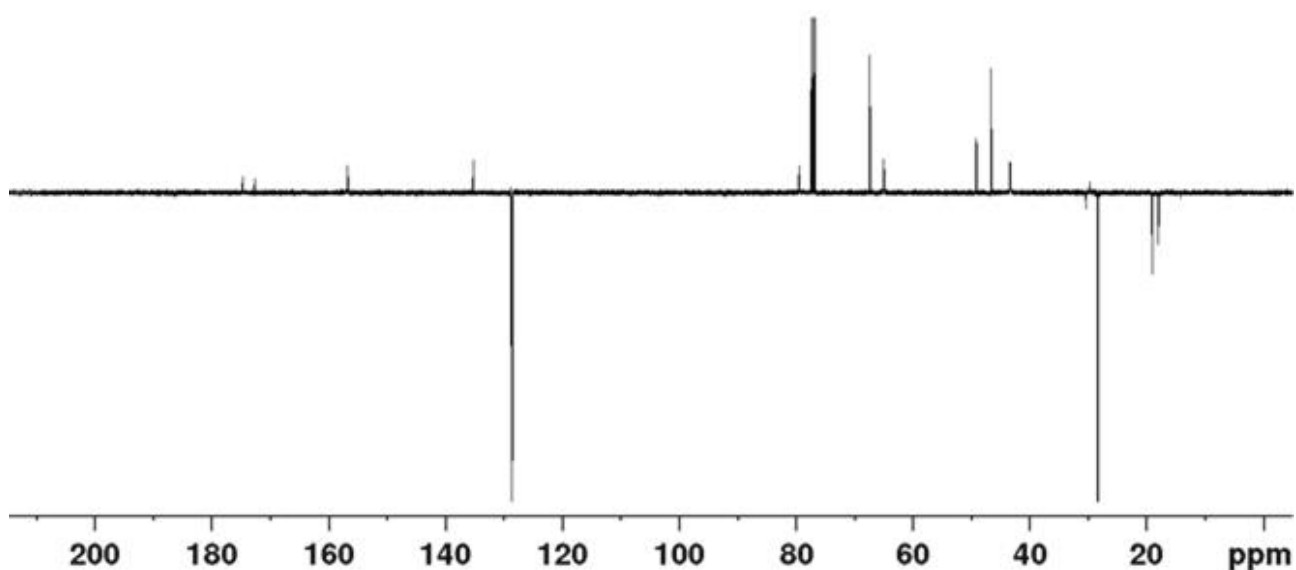
2,2-Bis(*tert*-butoxycarbonyl)aminomethyl)propanoic acid **59** (305.0 mg, 0.9 mmol) was dissolved in dry DCM (1.0 mL) and stirred under nitrogen atmosphere. Upon addition of diol **58a** (68.0 mg, 0.3 mmol) and DMAP (37.0 mg, 0.3 mmol) reagents were completely dissolved and the clear solution turned to light yellow. Finally DCC (185.0 mg, 0.9 mmol) was added at 0 °C and the formation of a white precipitate was immediately observed. The reaction was stirred overnight under nitrogen atmosphere at room temperature. The solution was diluted in DCM and filtered over a celite pad, the solvent was evaporated and the crude purified with flash chromatography (eluent: *n*-hexane - EtOAc, 4:1; R_f = 0.15). Pure product **70a** was obtained as a foamy white solid (210.2 mg, 80%). ^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.36-7.32 (m, 5 H, CH_2Ph), 5.51 (bs, 4 H, NH), 5.17 (s, 2 H, CH_2Ph), 4.30 (d, J = 11.2 Hz, 2 H, $\text{CHH}'\text{O}$), 4.17 (d, J = 11.2 Hz, 2 H, $\text{CHH}'\text{O}$), 3.44-3.36 (m, 4 H, $\text{CHH}'\text{N}$), 3.06-2.99 (m, 4 H, $\text{CHH}'\text{N}$), 1.43 (s, 36 H, $(\text{CH}_3)_3\text{CO}$), 1.30 (s, 3 H, CH_3), 1.05 (s, 6 H, CH_3). ^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 174.8, 172.7 (3 C, $\text{C}=\text{O}$), 156.8 (4 C, $(\text{CH}_3)_3\text{COC}=\text{O}$), 135.3 (*Ph*), 128.8 (*Ph*), 128.7 (2 C, *Ph*), 128.6 (2 C, *Ph*), 79.5 (4 C, $(\text{CH}_3)_3\text{CO}$), 67.4 (CH_2Ph), 65.1 (2 C, CH_2O), 49.2 (2 C,

CCH₂N), 46.8 (CCH₂O), 43.4 (4 C, CH₂N), 28.5 (12 C, (CH₃)₃CO), 19.1 (2 C, CH₃), 18.0 (CH₃). MS (ESI) m/z : calcd for C₄₂H₆₈N₄O₁₄ 852.5; found 875.6 [M+Na]⁺.

¹H NMR spectrum of **70a** in CDCl₃ (400 MHz)



¹³C NMR spectrum of **70a** in CDCl₃ (100 MHz)

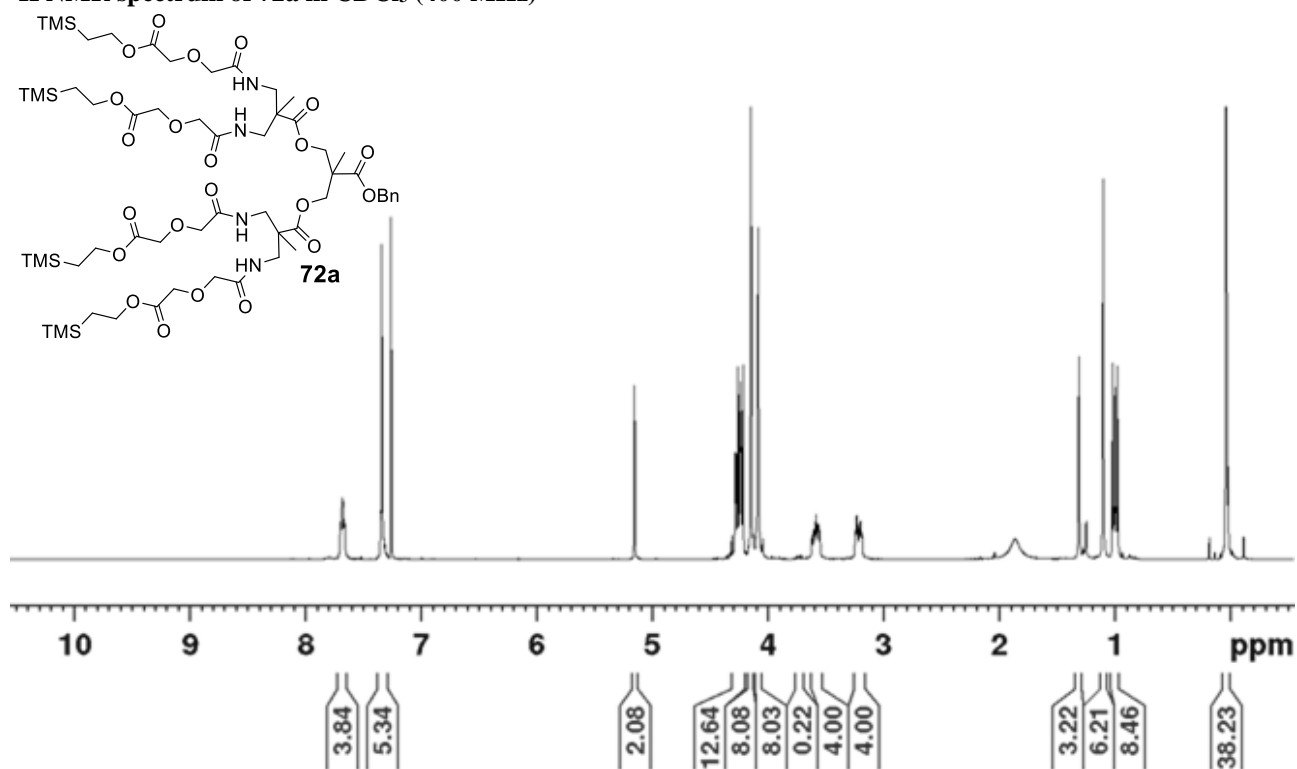


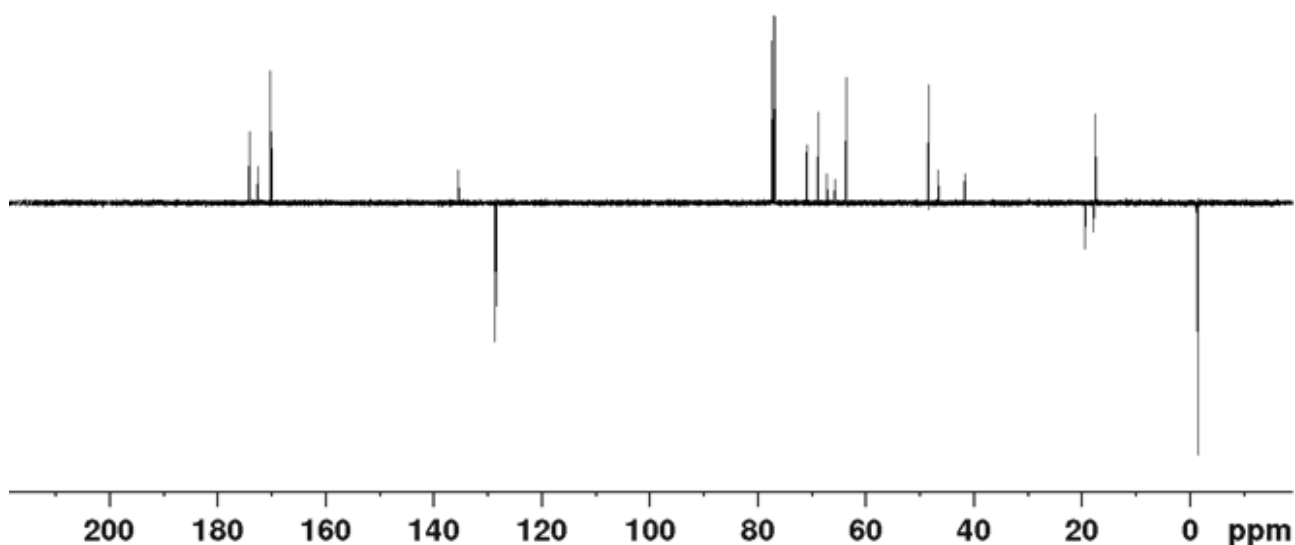
Synthesis of compound 72a

Under nitrogen atmosphere, TFA (3.1 mL, 41.0 mmol) was slowly added to a stirred solution of **70a** (350.0 mg, 0.41 mmol) in dry DCM (10.5 mL) which was reacted at room temperature for 40 min. DCM and TFA were co-evaporated with toluene, the crude was dried under high vacuum and then dissolved again in dry DCM (4.1 mL). Et₃N (2.1 mL, 12.3 mmol) was then added under nitrogen atmosphere with complete dissolution of the crude and the resulting solution was stirred for 10 min. Upon addition of the activated ester **55** (875.0 mg, 2.5 mmol) the solution turned to bright yellow and was stirred overnight at room temperature. DCM was evaporated, the crude was dissolved in EtOAc (50 mL) and washed with a saturated NaHCO₃ aqueous solution

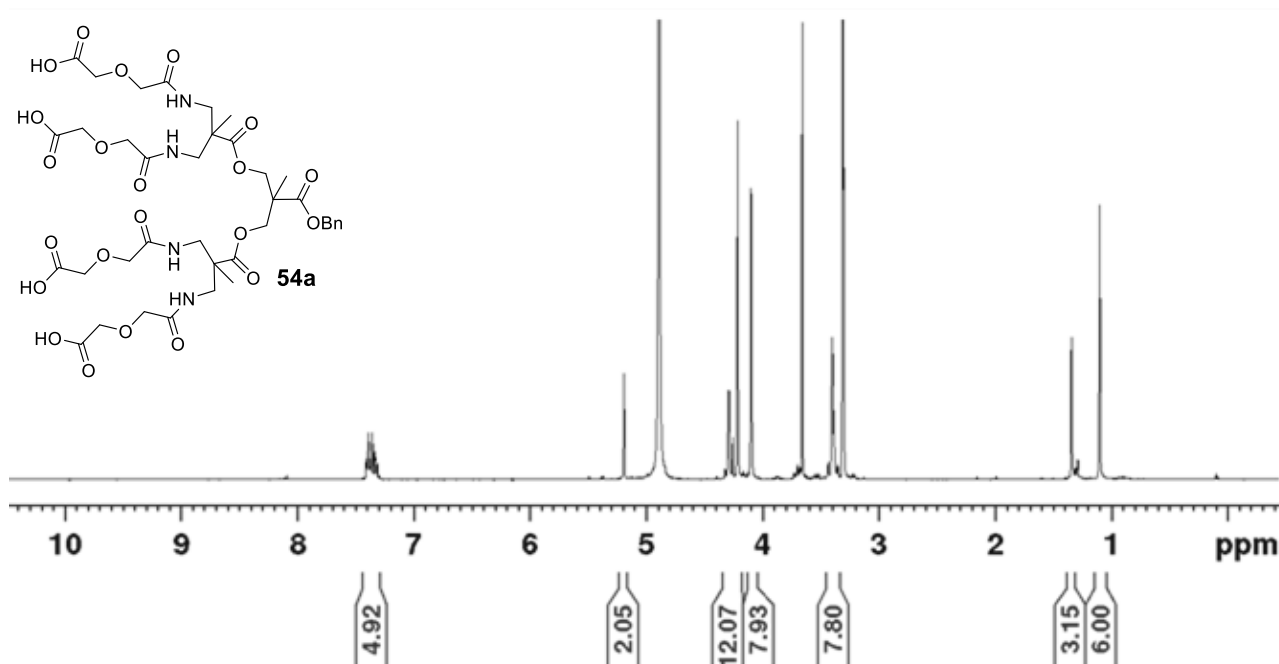
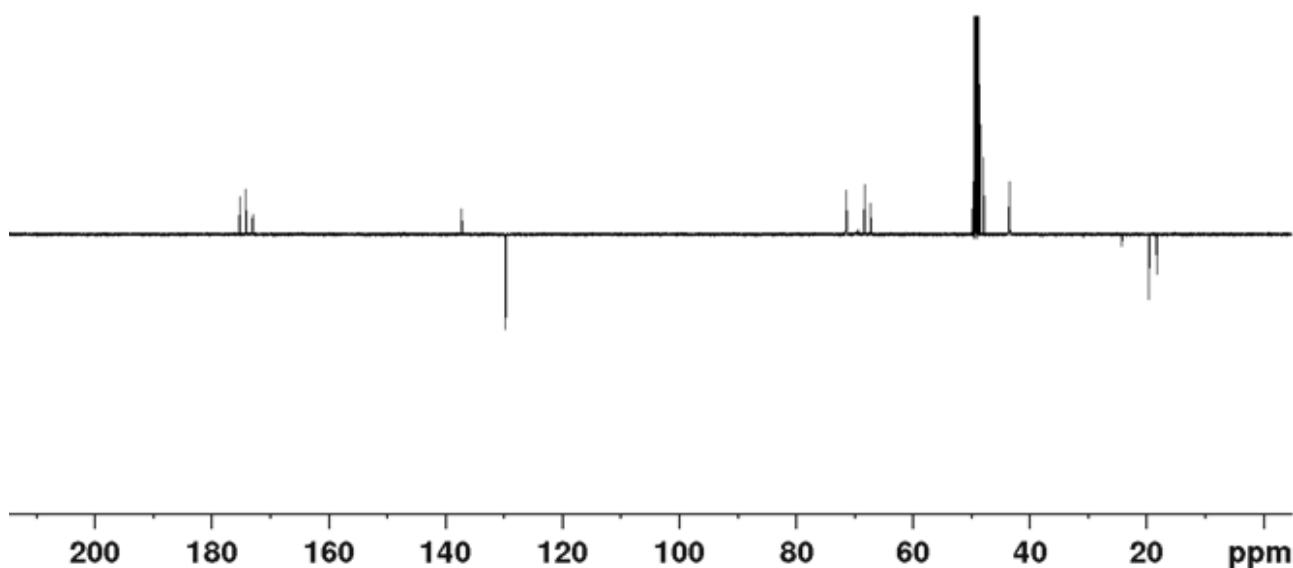
(2x15 mL) and with brine (15 mL). The organic phase was dried over anhydrous Na_2SO_4 and concentrated. The crude was purified with automated chromatography (Biotage SNAP Ultra 25g cartridge, gradient elution: from 100% *n*-hexane - EtOAc, 2:1 to 92% *n*-hexane - EtOAc, 2:1 8% MeOH) giving **72a** in high purity as a colourless oil (399.0 mg, 74%); $R_f = 0.13$ (eluent: *n*-hexane - EtOAc, 2:1 + 3% MeOH). ^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.67 (t, $J = 6.6$ Hz, 4 H, NH), 7.34-7.32 (m, 5 H, Ph), 5.15 (s, 2 H, CH_2Ph), 4.29 (d, $J = 11.2$ Hz, 2 H, $\text{CHH}'\text{O}$), 4.25 (d, $J = 11.2$ Hz, 2 H, $\text{CHH}'\text{O}$), 4.23 (m, 8 H, $\text{CH}_2\text{CH}_2\text{TMS}$), 4.14 (s, 8 H, $\text{CH}_2\text{COOCH}_2$), 4.07 (s, 8 H, CH_2CONH), 3.58 (ddd, $J = 13.6, 8.4, 4.8$ Hz, 4 H, $\text{CHH}'\text{N}$), 3.21 (ddd, $J = 14.0, 5.2, 4.4$ Hz, 4 H, $\text{CHH}'\text{N}$), 1.30 (s, 3 H, CH_3), 1.09 (s, 6 H, CH_3), 0.99 (m, 8 H, $\text{CH}_2\text{CH}_2\text{TMS}$), 0.03 (s, 36 H, $(\text{CH}_3)_3\text{Si}$). ^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 174.1, 172.6, 170.3, 170.0 (11 C, C=O), 135.5 (Ph), 128.8 (2 C, Ph), 128.7 (Ph), 128.5 (2 C, Ph), 71.0 (4 C, CH_2CONH), 68.9 (4 C, $\text{CH}_2\text{COOCH}_2$), 67.3 (CH_2Ph), 65.8 (2 C, CH_2O), 63.7 (4 C, $\text{CH}_2\text{CH}_2\text{TMS}$), 48.4 (2 C, CCH_2N), 46.7 (CCH_2O), 41.7 (4 C, CH_2N), 19.4 (2 C, CH_3), 17.9 (CH_3), 17.5 (4 C, $\text{CH}_2\text{CH}_2\text{TMS}$), -1.1 (12 C, $(\text{CH}_3)_3\text{Si}$). HRMS (ESI) m/z : calcd for $\text{C}_{58}\text{H}_{100}\text{N}_4\text{O}_{22}\text{Si}_4$ 1316.59063; found 681.28400 $[\text{M}+2\text{Na}]^{2+}$, 1339.58006 $[\text{M}+\text{Na}]^+$. MS (MALDI) m/z : calcd for $\text{C}_{58}\text{H}_{100}\text{N}_4\text{O}_{22}\text{Si}_4$ 1316.5; found 1340.2 $[\text{M}+\text{Na}]^+$ (matrix HCCA), 1340.3 $[\text{M}+\text{Na}]^+$ (matrix DHB), 1341.2 $[\text{M}+\text{Na}]^+$ (matrix SIN).

^1H NMR spectrum of **72a** in CDCl_3 (400 MHz)



¹³C NMR spectrum of **72a** in CDCl₃ (100 MHz)**Synthesis of compound 54a**

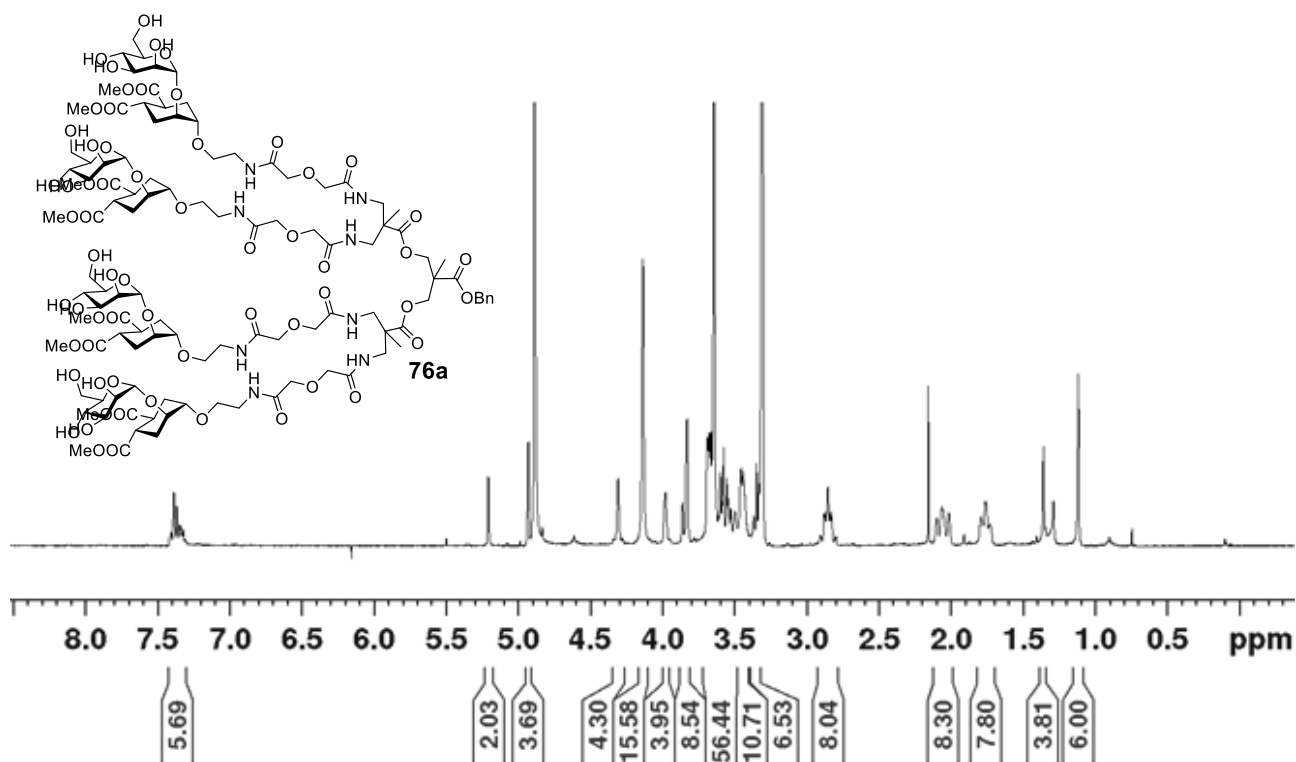
A solution of **72a** (40.1 mg, 0.03 mmol) in anhydrous DCM (1.5 mL) was stirred under nitrogen atmosphere and TFA (250 μ L, 3.26 mmol) was added dropwise. The reaction was performed stirring for 6 h at room temperature, monitoring by MS (ESI). The solvent was evaporated and TFA was co-evaporated with toluene (3x1.5 mL). The crude was washed with diethylether and **54a** was then dried under high vacuum and stored in the reaction vessel as a white foamy solid (28.0 mg, quantitative); $R_f = 0.00$ (eluent: *n*-hexane - EtOAc, 2:1 + 3% MeOH). No further purification was needed to employ the product in the next synthetic step. For characterisation purposes, an analytical amount was purified with automated reverse phase chromatography (Biotage SNAP C18 12g cartridge, gradient elution: from 100% H₂O to 100% MeOH) and was recovered as a white foamy solid (19.7 mg, 72%). However, no differences were appreciated comparing the product before and after purification by ¹H NMR analysis. ¹H NMR (400 MHz, CD₃OD) δ (ppm): 7.41-7.30 (m, 5 H, *Ph*), 5.20 (s, 2 H, CH₂Ph), 4.29 (AB system, 4 H, CH₂O), 4.22 (s, 8 H, CH₂COOH), 4.10 (s, 8 H, CH₂CONH), 3.44-3.35 (m, 8 H, CH₂N), 1.34 (s, 3 H, CH₃), 1.10 (s, 6 H, CH₃). ¹³C NMR (100 MHz, CD₃OD) δ (ppm): 175.2 (C=O), 174.2 (C=O), 172.9 (2 C, C=O), 137.2 (*Ph*), 129.7 (*Ph*), 129.6 (*Ph*), 129.5 (*Ph*), 71.3 (4 C, CH₂CONH), 69.2 (4 C, CH₂COOH), 68.2 (CH₂Ph), 67.2 (2 C, CH₂O), 49.4 (2 C, CCH₂N), 47.8 (2 C, CCH₂O), 43.4 (4 C, CH₂N), 19.5 (2 C, CH₃), 18.2 (CH₃). MS (ESI) m/z : calcd for C₃₈H₅₂N₄O₂₂ 916.307; found 917.443 [M+H]⁺, 939.369 [M+Na]⁺, 955.376 [M+K]⁺. MS (MALDI) m/z : calcd for C₃₈H₅₂N₄O₂₂ 916.3; found 917.7 [M+H]⁺, 939.8 [M+Na]⁺ (matrix HCCA).

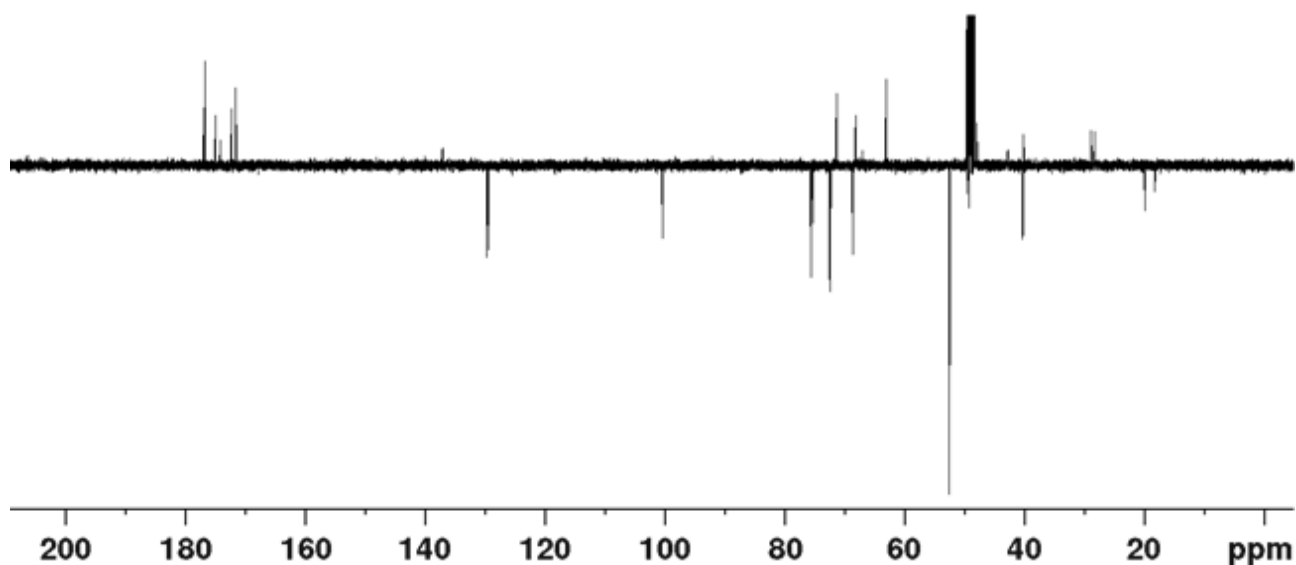
¹H NMR spectrum of 54a in CD₃OD (400 MHz)**¹³C NMR spectrum of 54a in CD₃OD (100 MHz)****Synthesis of compound 76a**

Tetraacid **54a** (4.8 mg, 5 μ mol) and the pseudo-disaccharide **5b** (18.9 mg, 39 μ mol) were dissolved together under nitrogen atmosphere in dry DMA (260 μ L). Freshly distilled DIPEA (15 μ L, 85 μ mol) was added and the reaction stirred for 15 min, then HATU (16.0 mg, 42 μ mol) was added and the solution immediately turned to bright yellow. The reaction was heated overnight at 30 $^{\circ}$ C and monitored with MS (MALDI) instrument. The solution turned to brownish-red. DMA was removed by a rotatory evaporator under high vacuum and the product was isolated by size-exclusion chromatography using a Sephadex LH-20 column (\varnothing = 3 cm, height = 50 cm; eluent: MeOH). A further purification with flash chromatography (eluent: CHCl₃ - MeOH - H₂O, 7:3:0.25; R_f = 0.31) afforded pure **76a** as a white foam (10.6 mg, 82%). An analytical sample was additionally purified by HPLC (Waters Atlantis 21 mm x 10 cm column, gradient from 100% (H₂O + 0.1% HCOOH) to 100% (CH₃CN + 0.1% HCOOH) in 12 min, t_R (product) = 6.6 min). $[\alpha]_D^{23}$ = + 30.0 (c = 0.17 in MeOH). ¹H

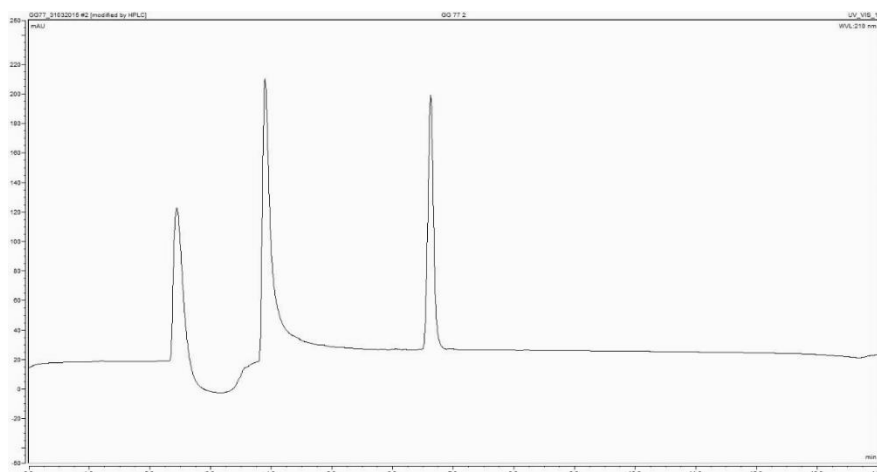
NMR (400 MHz, CD₃OD) δ (ppm): 7.39-7.33 (m, 5 H, *Ph*), 5.20 (s, 2 H, CH₂Ph), 4.93 (d, $J = 1.0$ Hz, 4 H, 1-H), 4.30 (AB system, 4 H, CH₂O), 4.13 (s, 16 H, CH₂CONH), 3.99-3.96 (m, 4 H, 2-H_D), 3.87-3.82 (m, 8 H, 6-H, 2-H), 3.70-3.52 (m, 28 H, 1-H_D, 3-H, 6-H', OCH₂CH₂N, 5-H, 4-H), 3.55 (s, 12 H, OCH₃), 3.54 (s, 12 H, OCH₃), 3.51-3.40 (m, 12 H, CHH'N, OCH₂CH₂N), 3.39-3.33 (m, 4 H, CHH'N), 2.91-2.80 (m, 8 H, 5-H_D, 4-H_D), 2.10-2.01 (m, 8 H, 6-H_{Deq}, 3-H_{Deq}), 1.79-1.73 (m, 8 H, 3-H_{Dax}, 6-H_{Dax}), 1.36 (s, 3 H, CH₃), 1.12 (s, 6 H, CH₃). ¹³C NMR (100 MHz, CD₃OD) δ (ppm): 176.9 (C=O), 176.9 (C=O), 175.1 (C=O), 174.3 (C=O), 172.4 (C=O), 171.7 (C=O), 137.2 (*Ph*), 129.8 (2 C, *Ph*), 129.6 (*Ph*), 129.5 (2 C, *Ph*), 100.4 (4 C, C1), 75.7 (4 C, C4), 75.3 (4 C, C1_D), 72.5 (4 C, C3), 72.4 (4 C, C2), 72.3 (4 C, C2_D), 71.4 (8 C, CH₂CONH), 68.6 (4 C, C5), 68.2 (CH₂Ph), 68.1 (4 C, OCH₂CH₂N), 67.1 (2 C, CH₂O), 63.1 (4 C, C6), 52.5 (8 C, OCH₃), 49.1 (2 C, CCH₂N), 47.9 (CCH₂O), 42.8 (4 C, CH₂N), 40.3 (8 C, C4_D, C5_D), 40.2 (4 C, OCH₂CH₂N), 28.9 (4 C, C6_D), 28.3 (4 C, C3_D), 19.9 (2 C, CH₃), 18.2 (CH₃). HRMS (ESI) m/z : calcd for C₁₁₀H₁₆₈N₈O₆₂ 2594.02717; found 887.66579 [M+3Na]³⁺, 1320.00119 [M+2Na]²⁺. MS (MALDI) m/z : calcd for C₁₁₀H₁₆₈N₈O₆₂ 2594.0; found 2616.6 [M+Na]⁺ (matrix HCCA).

¹H NMR spectrum of 76a in CD₃OD (400 MHz)



^{13}C NMR spectrum of 76a in CD_3OD (100 MHz)

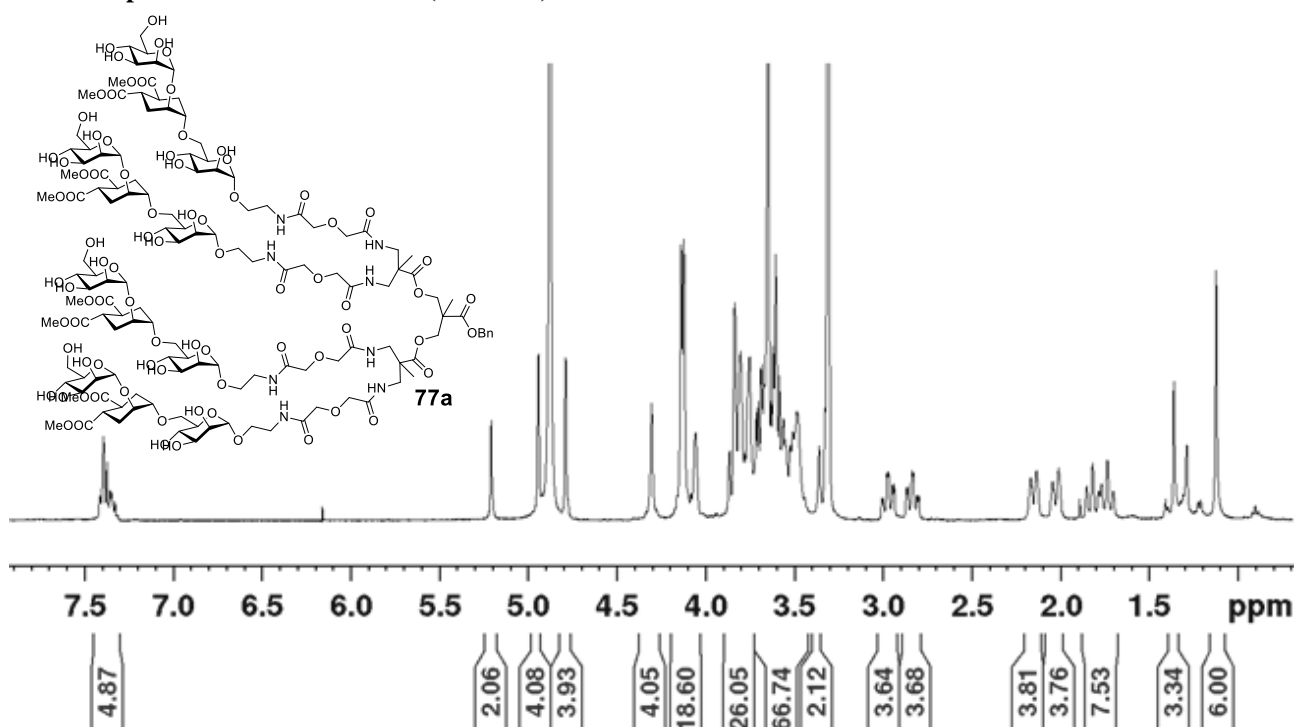
HPLC purification of analytical sample of 76a: HPLC trace. Waters Atlantis 21 mm x 10 cm column, gradient from 100% (H_2O + 0.1% HCOOH) to 100% (CH_3CN + 0.1% HCOOH) in 12 min, $t_R = 6.6$ min.

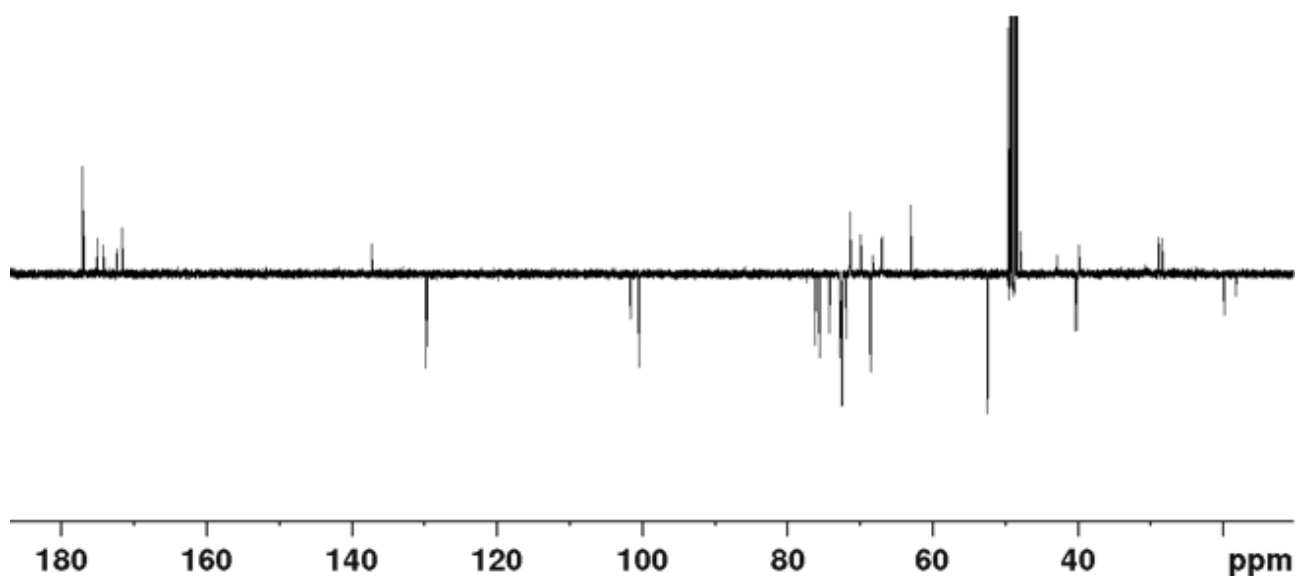
**Synthesis of compound 77a**

The tetracarboxylic acid **54a** (3.6 mg, 4 μmol) and compound **6b** (19.3 mg, 30 μmol) were dissolved together in dry DMA (200 μL) at 0 $^\circ\text{C}$. In the order HATU (9.1 mg, 24 μmol), HOAT (3.3 mg, 24 μmol) and freshly distilled DIPEA (11 μL , 63 μmol) were added at 0 $^\circ\text{C}$ to the stirred solution under nitrogen atmosphere. The solution turned to bright yellow. The reaction was heated at 37 $^\circ\text{C}$ overnight and monitored with MS (MALDI) instrument. The solution turned to brownish-red. DMA was removed by a rotatory evaporator under high vacuum and the crude was first purified with size-exclusion chromatography using a Sephadex LH-20 column ($\text{O} = 3$ cm, height = 50 cm; eluent: MeOH). A further purification with flash chromatography (eluent: CHCl_3 - MeOH - H_2O , 7:3:0.5; $R_f = 0.19$) afforded pure **77a** as a pale yellow foamy solid (9.4 mg, 74%). $[\alpha]_D^{21} = +44.9$ ($c = 0.47$ in MeOH). ^1H NMR (400 MHz, CD_3OD) δ (ppm): 7.41-7.33 (m, 5 H, *Ph*), 5.21 (s, 2 H, CH_2Ph), 4.94 (s, 4 H, 1- H_M), 4.79 (s, 4 H, 1- H_M), 4.30 (AB system, 4 H, CH_2O), 4.14 (s, 8 H, CH_2CONH), 4.12 (s, 8 H, CH_2CONH), 4.08-4.04 (m, 4 H, 2- H_D), 3.87-3.73 (m, 28 H, 6- H_M , 2- H_M , 2- H_M , 6- H_M , $\text{OCHH}'\text{CH}_2\text{N}$, 1- H_D , 6- H'_M), 3.72-3.42 (m, 44 H, 6- H'_M , 3- H_M , 3- H'_M , 5- H_M , 5- H'_M , 4- H_M , 4- H'_M , $\text{OCHH}'\text{CH}_2\text{N}$, $\text{OCHH}'\text{CH}_2\text{N}$,

^1H NMR), 3.65 (s, 24 H, OCH_3), 3.38-3.32 (m, 4 H, $\text{CH}'\text{N}$), 2.96 (td, $J = 12.4, 3.5$ Hz, 4 H, 5- H_D), 2.83 (td, $J = 12.6, 3.5$ Hz, 4 H, 4- H_D), 2.15 (dt, $J = 13.2, 2.9$ Hz, 4 H, 6- H_Deq), 2.03 (dt, $J = 13.2, 2.9$ Hz, 4 H, 3- H_Deq), 1.82 (td, $J = 13.7, 1.4$ Hz, 4 H, 3- H_Dax), 1.73 (td, $J = 13.7, 1.4$ Hz, 4 H, 6- H_Dax), 1.36 (s, 3 H, CH_3), 1.12 (s, 6 H, CH_3). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 177.2 ($\text{C}=\text{O}$), 177.0 ($\text{C}=\text{O}$), 175.1 ($\text{C}=\text{O}$), 174.3 ($\text{C}=\text{O}$), 172.4 ($\text{C}=\text{O}$), 171.7 ($\text{C}=\text{O}$), 137.2 (*Ph*), 129.8 (2 C, *Ph*), 129.6 (*Ph*), 129.5 (2 C, *Ph*), 101.6 (4 C, C1_M), 100.4 (4 C, C1_M), 76.1 (4 C, C1_D), 75.5, 74.1 (8 C, C4_M , $\text{C4}_\text{M}'$), 72.6 (4 C, C2_D), 72.5 (12 C, C3_M , $\text{C3}_\text{M}'$, $\text{C2}_\text{M}'$), 71.9 (4 C, C2_M), 71.4 (8 C, CH_2CONH), 69.9 (4 C, C6_M), 68.6 (8 C, C5_M , $\text{C5}_\text{M}'$), 68.3 (CH_2Ph), 67.0 (6 C, CH_2O , $\text{OCH}_2\text{CH}_2\text{N}$), 63.0 (4 C, $\text{C6}_\text{M}'$), 52.5 (12 C, OCH_3), 52.4 (12 C, OCH_3), 49.1 (2 C, CCH_2N), 47.9 (CCH_2O), 42.9 (4 C, CH_2N), 40.4 (4 C, C4_D), 40.2 (4 C, C5_D), 39.9 (4 C, $\text{OCH}_2\text{CH}_2\text{N}$), 28.9 (4 C, C6_D), 28.4 (4 C, C3_D), 19.9 (2 C, CH_3), 18.2 (CH_3). HRMS (ESI) m/z : calcd for $\text{C}_{134}\text{H}_{208}\text{N}_8\text{O}_{82}$ 3242.23849; found 1103.73571 $[\text{M}+3\text{Na}]^{3+}$, 1644.10830 $[\text{M}+2\text{Na}]^{2+}$. MS (MALDI) m/z : calcd for $\text{C}_{134}\text{H}_{208}\text{N}_8\text{O}_{82}$ 3242.2; found 3265.1 $[\text{M}+\text{Na}]^+$ (matrix HCCA).

^1H NMR spectrum of 77a in CD_3OD (400 MHz)

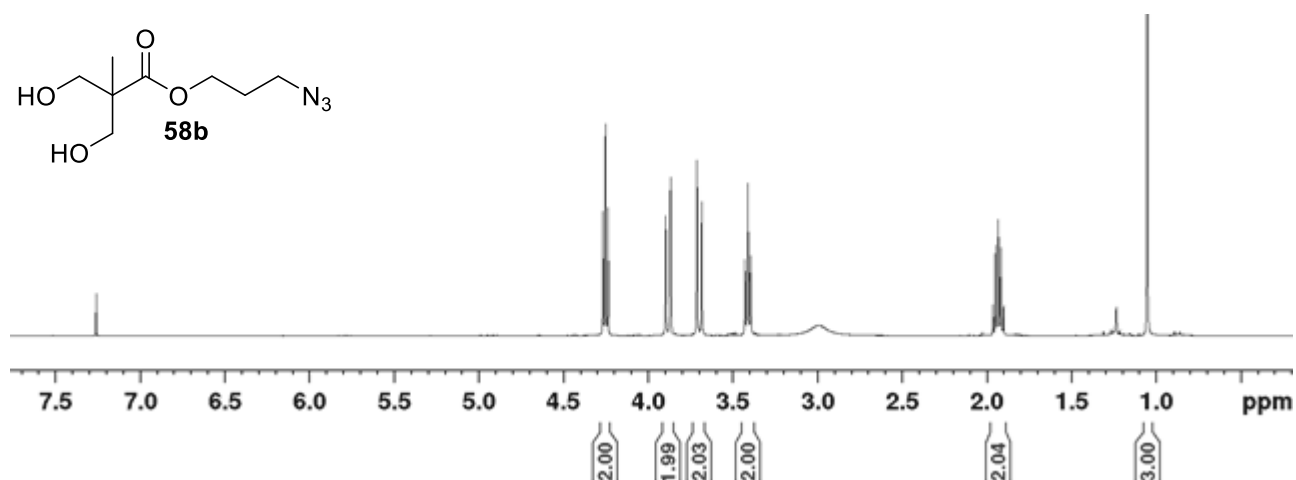


^{13}C NMR spectrum of **77a** in CD_3OD (100 MHz)

1.6.2.6 Synthesis of 16-valent glycodendrimer **81**

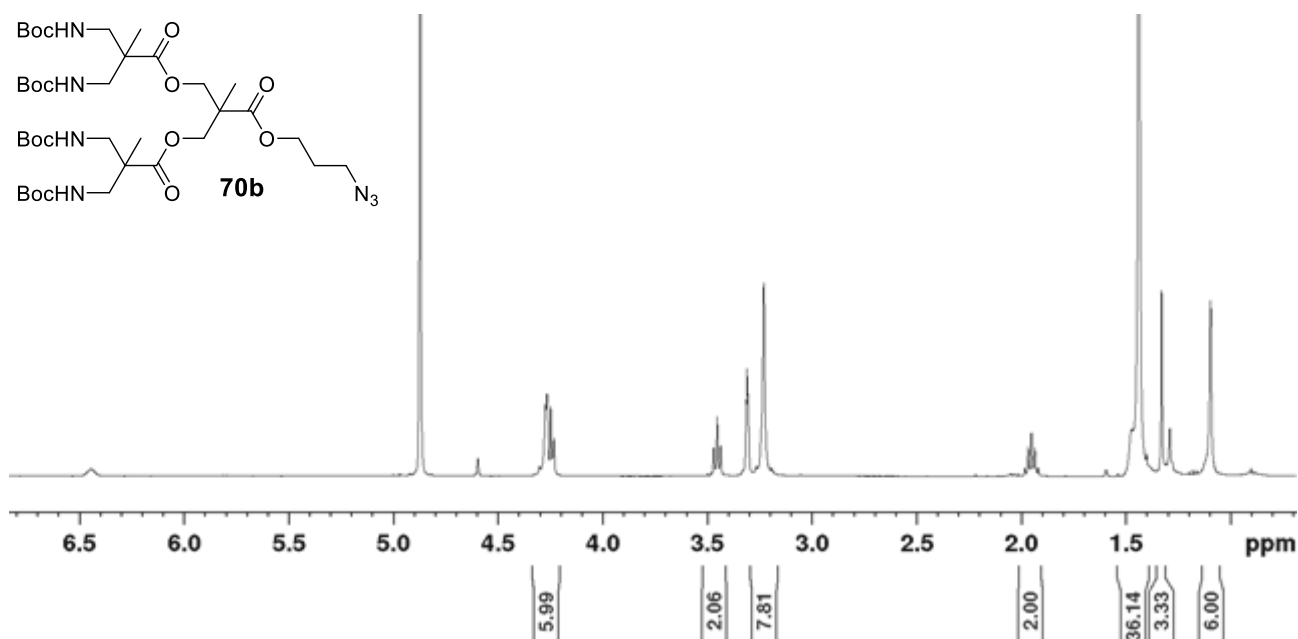
Synthesis of 3-azidopropyl 2,2-bis(hydroxymethyl)propanoate **58b**

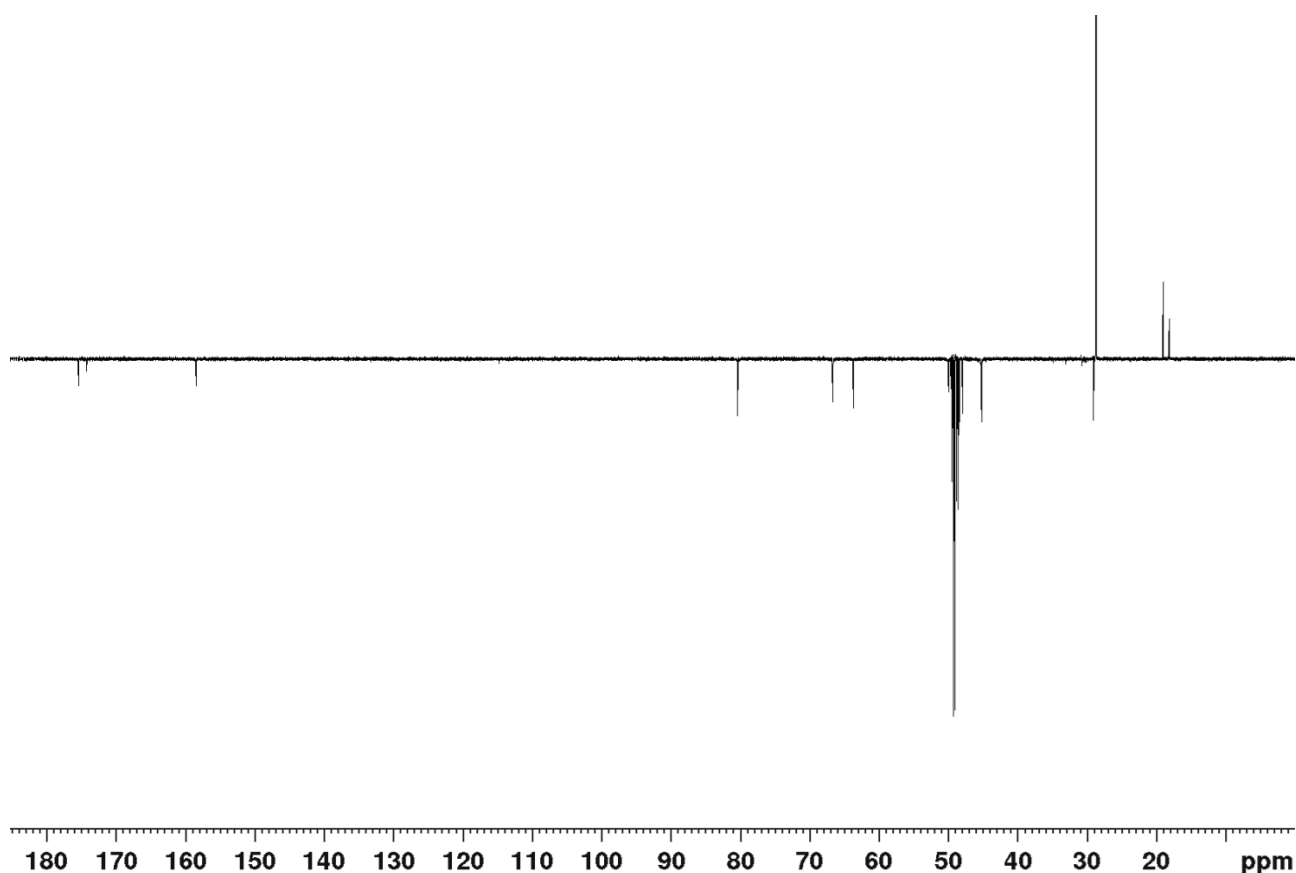
2,2-bis(hydroxymethyl)propionic acid **60** (366.0 mg, 2.73 mmol) was dissolved in dry DMF (900 μL) under nitrogen atmosphere. Cs_2CO_3 (534.0 mg, 1.64 mmol) was added and the reaction mixture stirred for 2 h at 90 $^\circ\text{C}$, with the formation of a white precipitate. The reaction was cooled down at room temperature and a solution of **79** (163.0 mg, 0.91 mmol) in dry DMF (500 μL) was added. The reaction was stirred for 3 h at 90 $^\circ\text{C}$, monitoring by TLC ($R_f = 0.22$; eluent: toluene - EtOAc, 1:1), with complete consumption of **79**. The reaction was cooled to room temperature, H_2O (15 mL) was added and the product was extracted with EtOAc (4x20 mL). Organic phases were dried over anhydrous Na_2SO_4 and, after filtration and evaporation of the solvent, the crude was purified by automated chromatography (gradient elution: from 80% *n*-hexane - 20% EtOAc to 100% EtOAc) giving **58b** as a yellow liquid (157.4 mg, 80%). The ^1H NMR spectroscopic data are in accordance with those previously reported in the literature.⁶³ ^1H NMR (400 MHz, CDCl_3) δ (ppm): 4.25 (t, $J = 6.1$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 3.88 (d, $J = 11.2$ Hz, 2 H, 1'-H, 2'-H), 3.70 (d, $J = 11.3$ Hz, 2 H, 1'-H', 2'-H'), 3.41 (t, $J = 6.6$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 1.93 (quint, $J = 6.3$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 1.05 (s, 3 H, CH_3).

 ^1H NMR spectrum of **58b** in CDCl_3 (400 MHz)

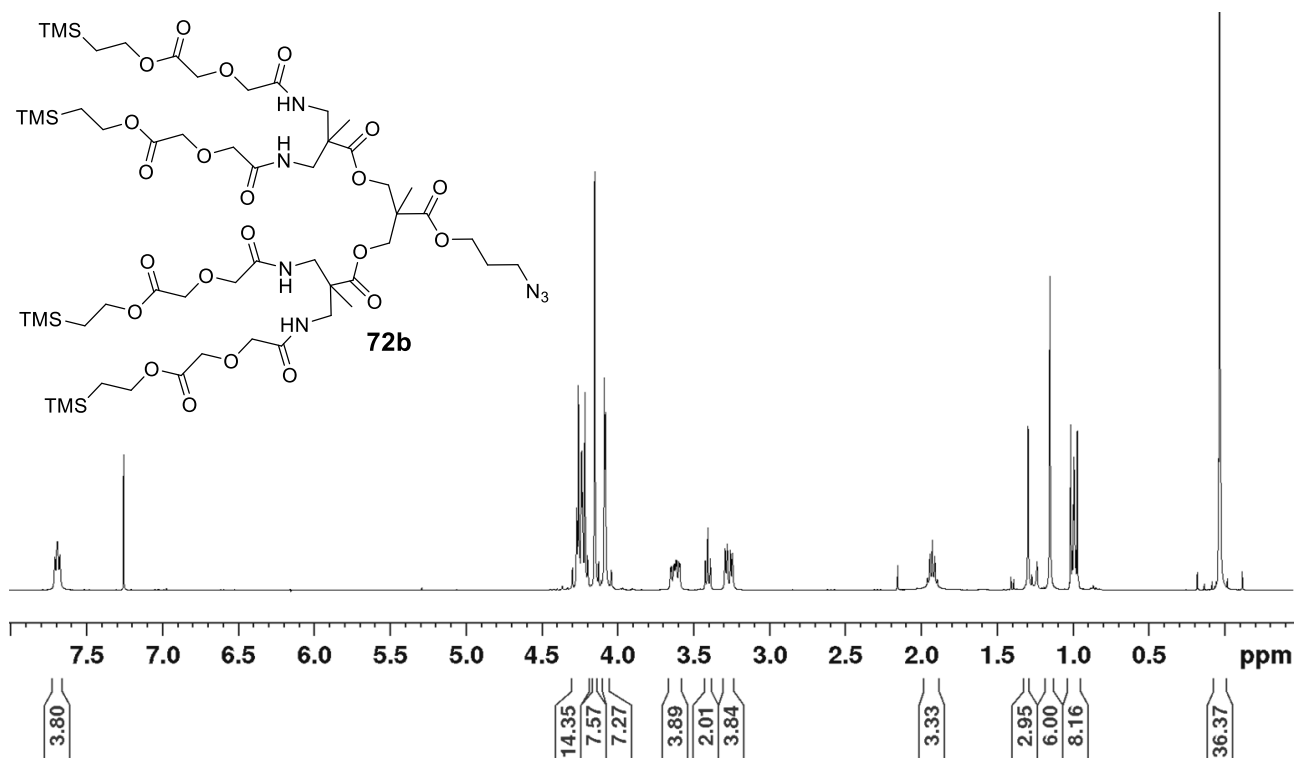
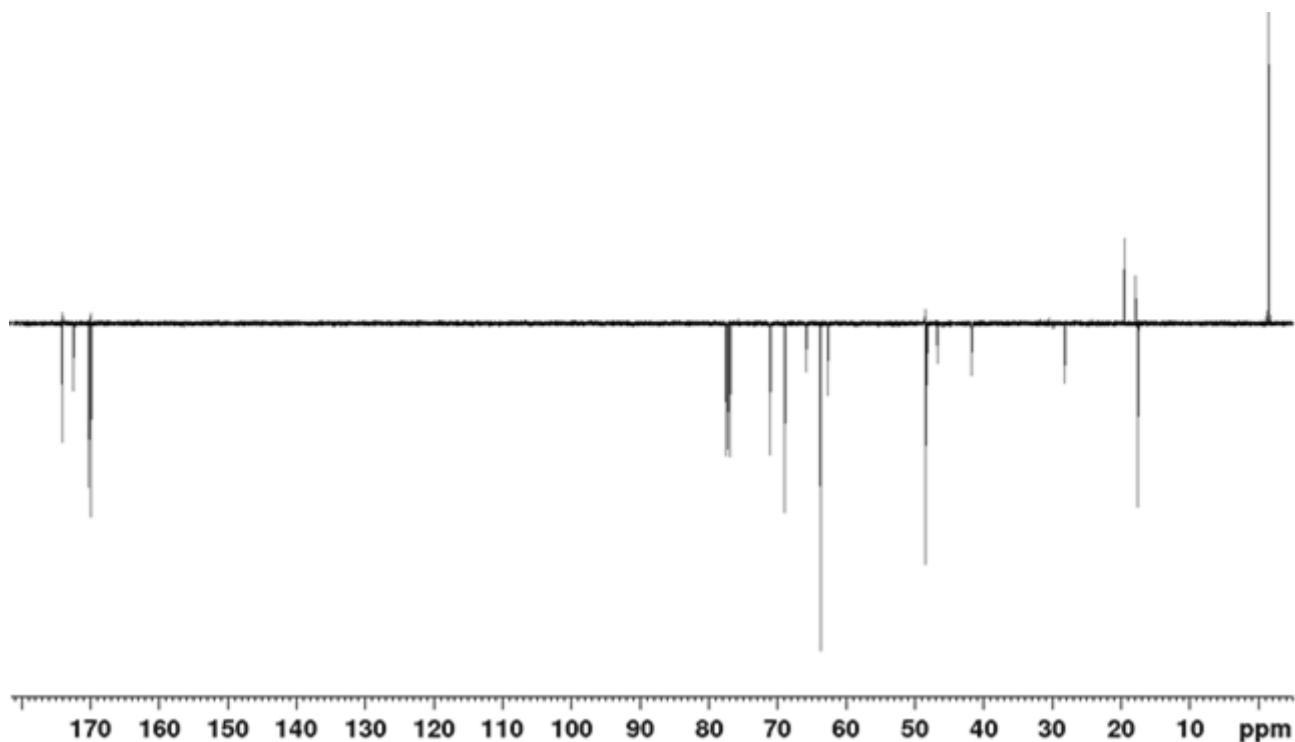
Synthesis of compound 70b

The diol **58b** (45.0 mg, 0.21 mmol) was dissolved in dry DCM (700 μ L) under nitrogen atmosphere. Compound **59** (206.0 mg, 0.62 mmol) was then added followed by DMAP (25.3 mg, 0.21 mmol). At 0 $^{\circ}$ C, DCC (128.0 mg, 0.62 mmol) was added with the formation of a white precipitate. The mixture was allowed to reach room temperature and stirred for 16 h. The mixture was then diluted with DCM and filtered over a celite pad. The solvent was evaporated and a flash chromatography (eluent: *n*-hexane - EtOAc, 4:1; R_f = 0.14) afforded pure **70b** as a white foamy solid (135.6 mg, 78%). ^1H NMR (400 MHz, CD_3OD) δ (ppm): 6.44 (bs, 4 H, NH), 4.27 (AB system, 4 H, CH_2O), 4.25 (t, J = 6.1 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 3.45 (t, J = 6.6 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 3.23 (AB system, 8 H, CH_2N), 1.95 (quint, J = 6.4 Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 1.44 (s, 36 H, $(\text{CH}_3)_3\text{CO}$), 1.33 (s, 3 H, CH_3), 1.09 (s, 6 H, CH_3). ^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 175.5, 174.3 (3 C, C=O), 158.6 (4 C, $(\text{CH}_3)_3\text{COC}=\text{O}$), 80.4 (4 C, $\text{CH}_3)_3\text{CO}$), 66.6 (2 C, CH_2O), 63.7 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 50.0 (2 C, CCH_2N), 49.3 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 47.9 (2 C, CCH_2O), 45.2 (4 C, CH_2N), 29.1 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 28.7 (12 C, $(\text{CH}_3)_3\text{CO}$), 19.0 (2 C, CH_3), 18.2 (CH_3). MS (ESI) m/z : calcd for $\text{C}_{38}\text{H}_{67}\text{N}_7\text{O}_{14}$ 845.5; found 868.5 $[\text{M}+\text{Na}]^+$.

 ^1H NMR spectrum of 70b in CD_3OD (400 MHz)

¹³C NMR spectrum of **70b** in CD₃OD (100 MHz)**Synthesis of compound 72b**

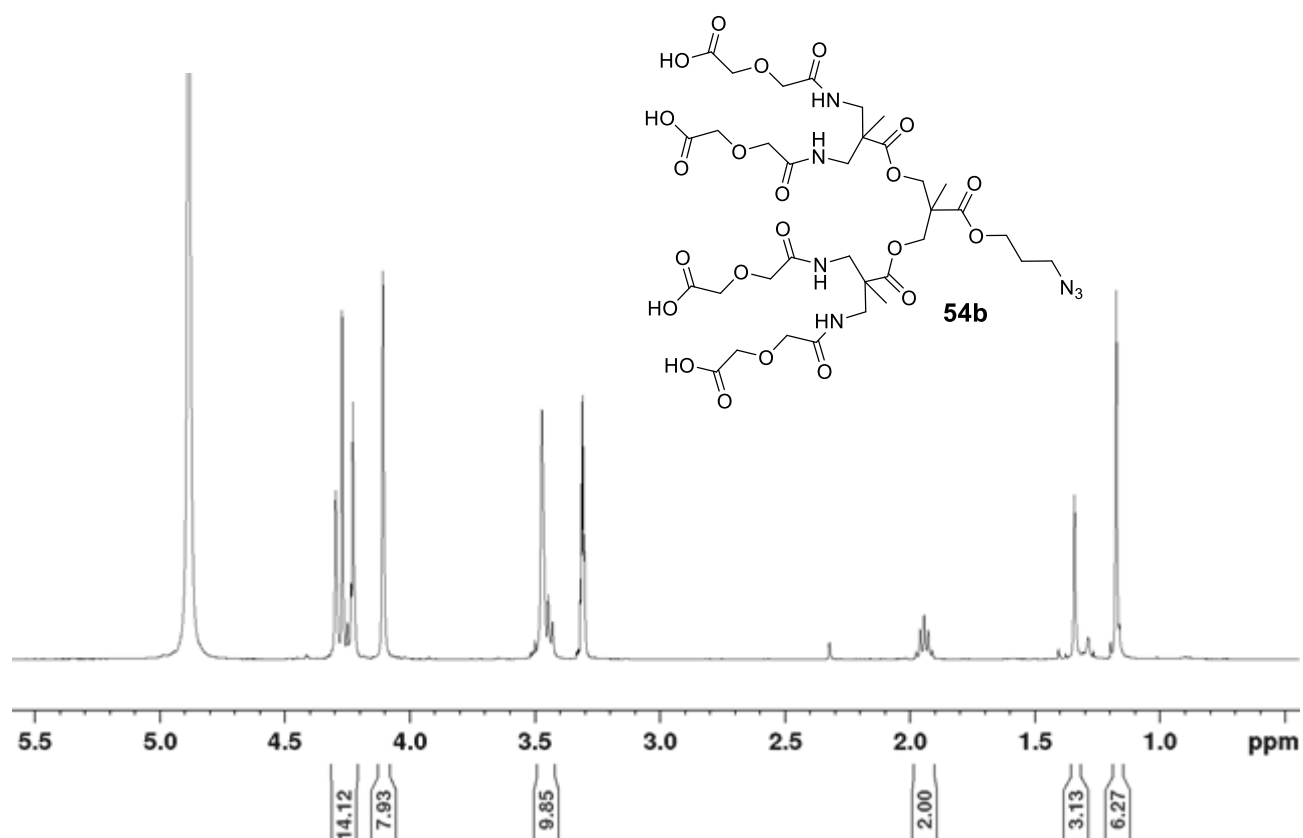
Under nitrogen atmosphere, **70b** (95.3 mg, 0.113 mmol) was dissolved in dry DCM (2.8 mL) and TFA (860 μ L, 11.3 mmol) was slowly added at 0 °C. The solution was stirred at room temperature for 40 min, then DCM and TFA were co-evaporated with toluene and diethylether. The crude was dried under high vacuum and redissolved under nitrogen atmosphere in DCM (1.1 mL). Freshly distilled triethylamine (590 μ L, 3.38 mmol) was added at 0 °C. The mixture was stirred for 10 min, then the activated ester **55** (240.0 mg, 0.68 mmol) was added. The bright yellow solution was allowed to reach room temperature and stirred overnight. DCM was evaporated from the crude, which was dissolved in EtOAc (12 mL) and washed with a saturated NaHCO₃ aqueous solution (2x4 mL) and brine (4 mL). The organic phase was dried over anhydrous Na₂SO₄, the solvent was evaporated under reduced pressure and the crude was purified with flash chromatography (eluent: CHCl₃ - acetone, 10:1; R_f = 0.18). Compound **72b** was obtained as a pure colourless oil (112.3 mg, 76%). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.70 (dd, J = 7.4, 5.8 Hz, 4 H, NH), 4.30-4.19 (m, 14 H, CH₂O, CH₂CH₂TMS, OCH₂CH₂CH₂N₃), 4.15 (s, 8 H, CH₂COOCH₂), 4.10-4.07 (m, 8 H, CH₂CONH), 3.62 (ddd, J = 14.0, 8.1, 2.7 Hz, 4 H, CHH'N), 3.41 (t, J = 6.6 Hz, 2 H, OCH₂CH₂CH₂N₃), 3.28 (dd, J = 14.0, 5.8 Hz, 4 H, CHH'N), 1.93 (quint, J = 6.4 Hz, 2 H, OCH₂CH₂CH₂N₃), 1.30 (s, 3 H, CH₃), 1.16 (s, 6 H, CH₃), 1.02-0.98 (m, 8 H, CH₂CH₂TMS), 0.03 (s, 36 H, (CH₃)₃Si). ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 174.1, 172.6, 170.3, 170.0 (11 C, C=O), 71.0 (4 C, CH₂CONH), 68.9 (4 C, CH₂COOCH₂), 65.8 (2 C, CH₂O), 63.7 (4 C, CH₂CH₂TMS), 62.6 (OCH₂CH₂CH₂N₃), 48.5 (2 C, CCH₂N), 48.2 (OCH₂CH₂CH₂N₃), 46.7 (2 C, CCH₂O), 41.7 (4 C, CH₂N), 28.1 (OCH₂CH₂CH₂N₃), 19.5 (2 C, CH₃), 17.8 (CH₃), 17.5 (4 C, CH₂CH₂TMS), -1.4 (12 C, (CH₃)₃Si). HRMS (ESI) m/z : calcd for C₅₄H₉₉N₇O₂₂Si₄ 1309.59202; found 677.78621 [M+2Na]²⁺, 1332.58236 [M+Na]⁺.

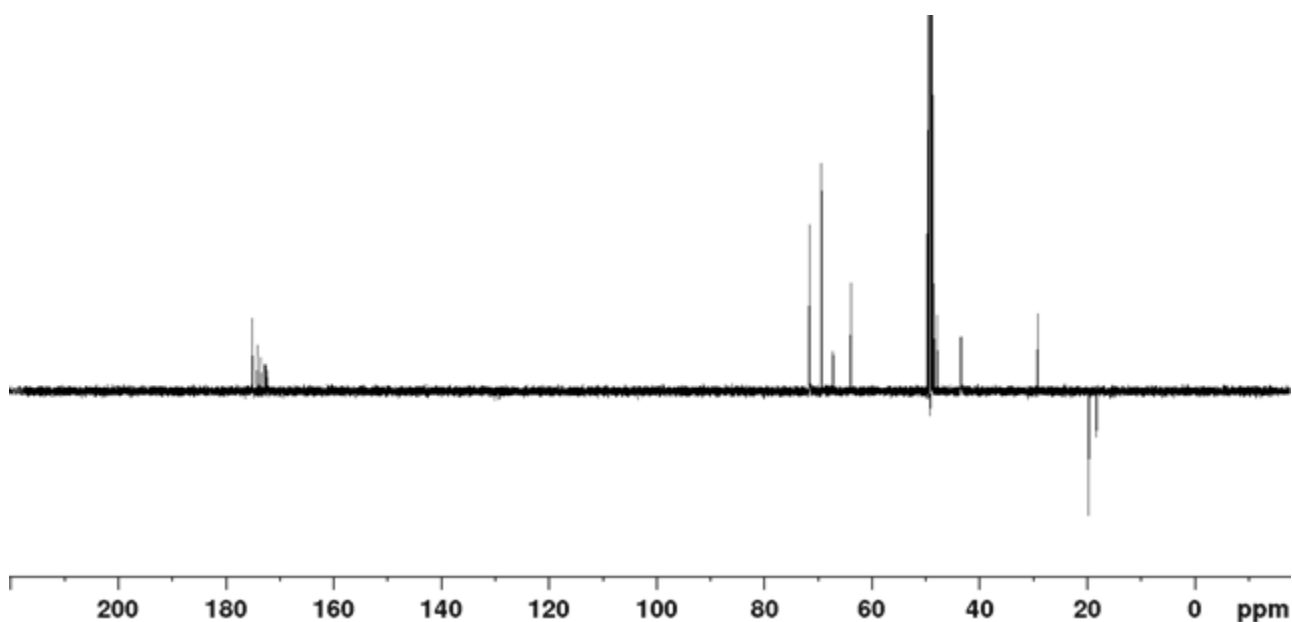
¹H NMR spectrum of 72b in CDCl₃ (400 MHz)**¹³C NMR spectrum of 72b in CDCl₃ (100 MHz)****Synthesis of compound 54b**

To a solution of **72b** (22.8, 17.4 μmol) in dry DCM (870 μL), TFA (145 μL, 1.9 mmol) was slowly added at 0 °C under nitrogen atmosphere. The reaction is stirred at room temperature for 6 h, monitoring with TLC (eluent: CHCl₃ - acetone, 10:1) until full conversion of the starting reagent. DCM was evaporated and TFA co-evaporated with toluene. The crude was then washed with diethylether and dried under high vacuum to

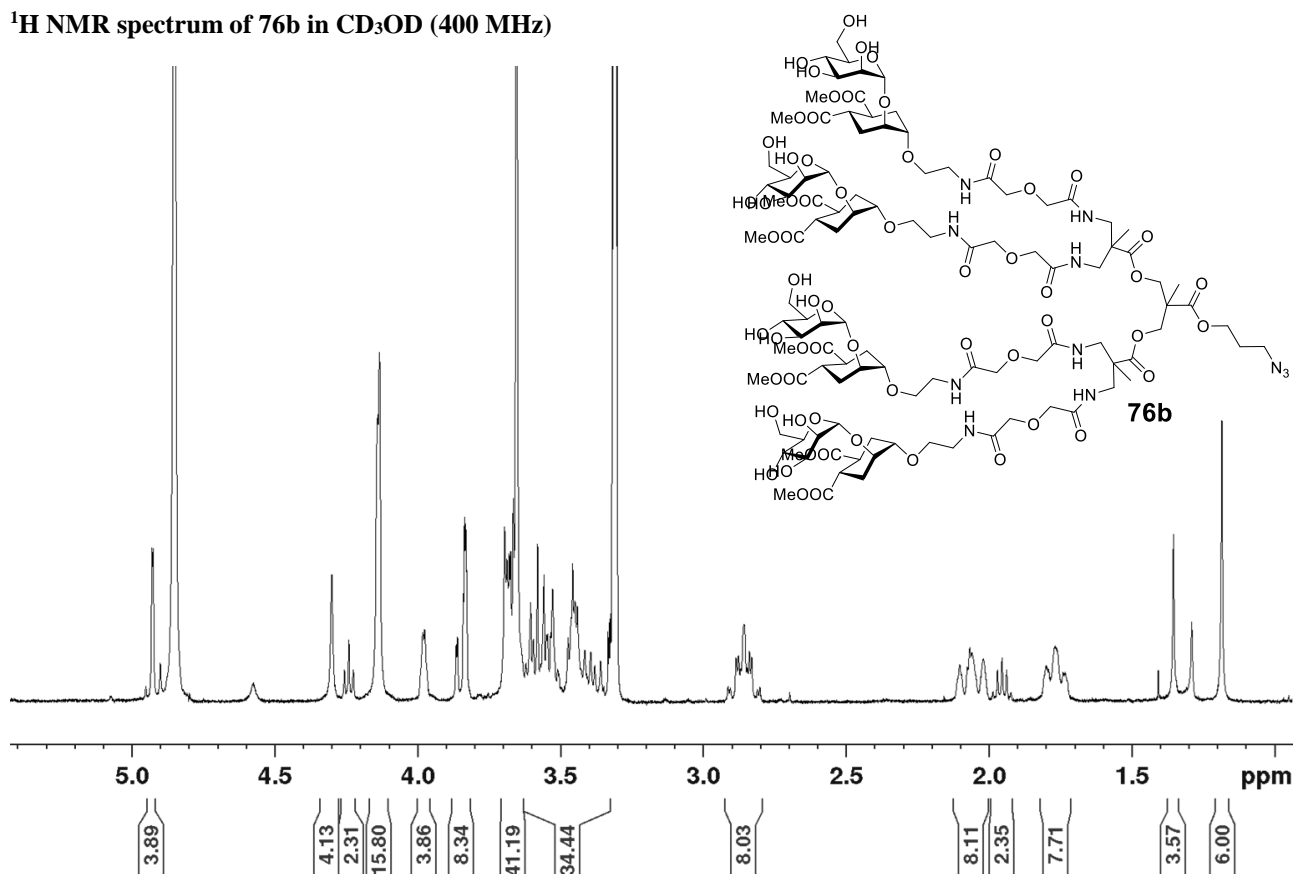
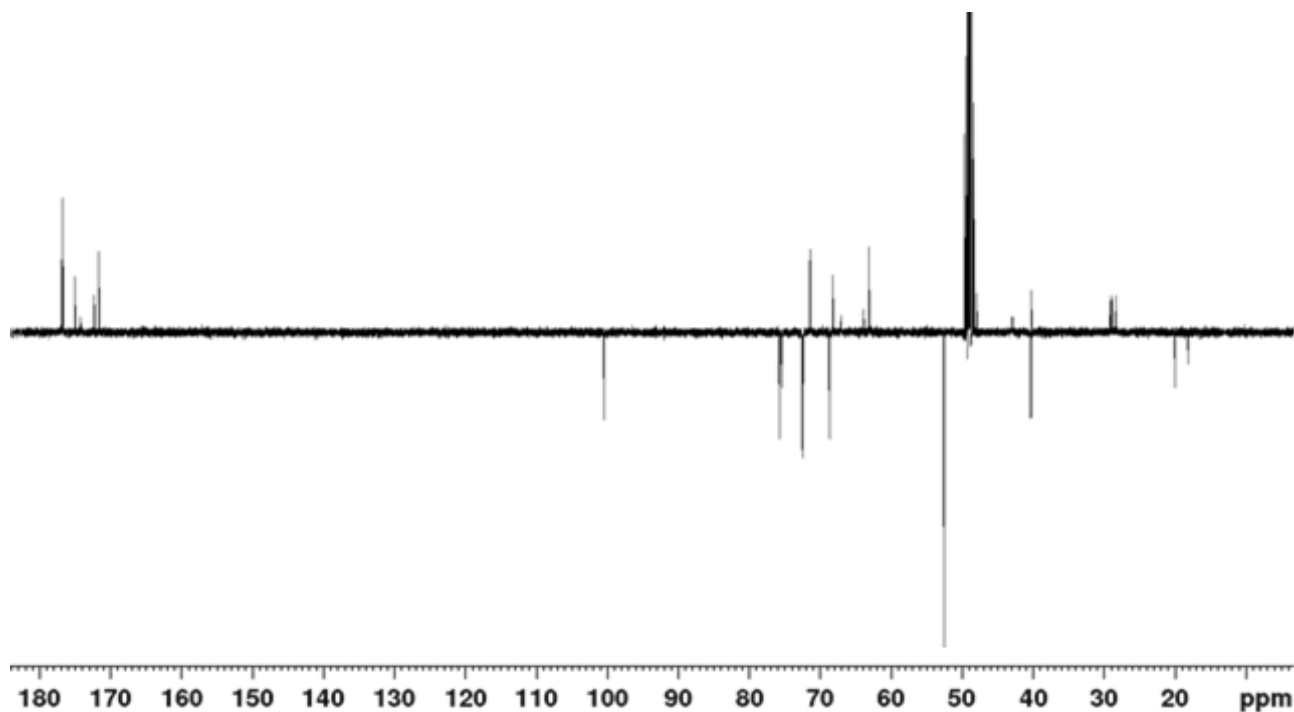
obtain a white foamy solid (quantitative). ^1H NMR (400 MHz, CD_3OD) δ (ppm): 4.30 (s, 4 H, CH_2O), 4.24 (t, $J = 6.2$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 4.23 (s, 8 H, CH_2COOH), 4.11 (s, 8 H, CH_2CONH), 3.49-3.46 (m, 8 H, CH_2N), 3.45 (t, $J = 6.5$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 1.94 (quint, $J = 6.4$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 1.34 (s, 3 H, CH_3), 1.18 (s, 6 H, CH_3). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 175.2, 174.3, 173.6 (4 C, $\text{C}=\text{O}$), 71.5 (4 C, CH_2CONH), 69.2 (4 C, CH_2COOH), 67.1 (2 C, CH_2O), 63.8 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 49.3 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 47.8 (2 C, CCH_2O), 43.3 (4 C, CH_2N), 29.1 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 19.6 (2 C, CH_3), 18.1 (CH_3). MS (ESI) m/z : calcd for $\text{C}_{34}\text{H}_{51}\text{N}_7\text{O}_{22}$ 909.309; found 477.650 $[\text{M}+2\text{Na}]^{2+}$, 932.293 $[\text{M}+\text{Na}]^+$, 948.288 $[\text{M}+\text{K}]^+$.

^1H NMR spectrum of 54b in CD_3OD (400 MHz)

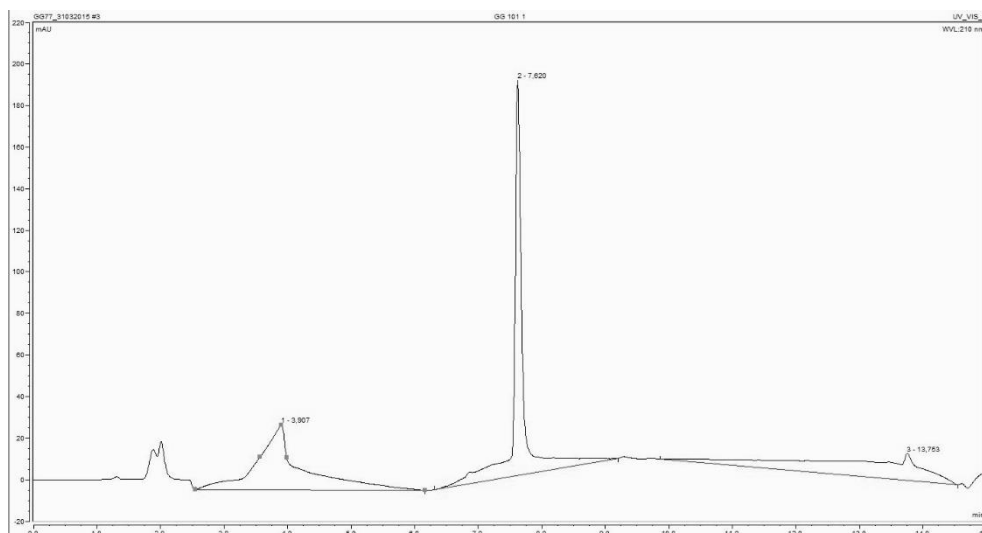


^{13}C NMR spectrum of **54b** in CD_3OD (100 MHz)Synthesis of compound **76b**

The tetraacid **54b** (5.8 mg, 6.4 μmol) and the pseudo-disaccharide **5b** (23.2 mg, 48.9 μmol) were dissolved together in dry DMA (320 μL) under nitrogen atmosphere. The mixture was cooled to 0 $^{\circ}\text{C}$ and HATU (19.6 mg, 51.5 μmol) and DIPEA (18 μL , 103 μmol) were added. The bright yellow solution was heated at 35 $^{\circ}\text{C}$ and stirred overnight monitoring with MS (ESI). The crude was purified with a size-exclusion chromatography using a Sephadex LH-20 column ($\varnothing = 3$ cm, height = 50 cm; eluent: MeOH). Additionally, was performed a flash chromatography (eluent: CHCl_3 - MeOH - H_2O , 7:3:0.5; $R_f = 0.36$) and mixed fractions were further purified by HPLC (Waters Atlantis 21 mm x 10 cm column, gradient from 100% (H_2O + 0.1% HCOOH) to 25% (H_2O + 0.1% HCOOH) - 75% (CH_3CN + 0.1% HCOOH) in 12 min, t_R (product) = 7.6 min). The reunited fractions afforded pure **76b** as a white foam (11.1 mg, 67 %). $[\alpha]_D^{35} = +31.3$ ($c = 0.55$ in MeOH). ^1H NMR (400 MHz, CD_3OD) δ (ppm): 4.93 (d, $J = 1.3$ Hz, 4 H, 1-H), 4.30 (s, 4 H, CH_2O), 4.24 (t, $J = 6.2$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 4.16-4.12 (m, 16 H, CH_2CONH), 3.99-3.96 (m, 4 H, 2- H_D), 3.87-3.82 (m, 8 H, 6-H, 2-H), 3.71-3.49 (m, 32 H, 1- H_D , 3-H, 6- H' , $\text{OCH}_2\text{CH}_2\text{N}$, 5-H, 4-H, $\text{CHH}'\text{N}$), 3.65 (24 H, OCH_3), 3.48-3.33 (14 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$, $\text{OCH}_2\text{CH}_2\text{N}$, $\text{CHH}'\text{N}$), 2.92-2.80 (m, 8 H, 5- H_D , 4- H_D), 2.12-2.01 (m, 8 H, 6- H_{Deq} , 3- H_{Deq}), 1.95 (t, $J = 6.4$ Hz, 2 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 1.82-1.72 (m, 8 H, 3- H_{Dax} , 6- H_{Dax}), 1.35 (s, 3 H, CH_3), 1.18 (s, 6 H, CH_3). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 176.9 (C=O), 176.8 (C=O), 175.1 (C=O), 174.3 (C=O), 172.4 (C=O), 171.7 (C=O), 100.4 (4 C, C1), 75.7 (4 C, C4), 75.4 (4 C, C1 $_D$), 72.6 (4 C, C3), 72.5 (4 C, C2), 72.4 (4 C, C2 $_D$), 71.4 (8 C, CH_2CONH), 68.7 (4 C, C5), 68.2 (4 C, $\text{OCH}_2\text{CH}_2\text{N}$), 67.1 (2 C, CH_2O), 63.8 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 63.1 (4 C, C6), 52.5 (8 C, OCH_3), 49.3 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 49.1 (2 C, CCH_2N), 47.9 (CCH_2O), 42.8 (4 C, CH_2N), 40.3 (8 C, C4 $_D$, C5 $_D$), 40.2 (4 C, $\text{OCH}_2\text{CH}_2\text{N}$), 29.1 ($\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}_3$), 28.9 (4 C, C6 $_D$), 28.3 (4 C, C3 $_D$), 20.0 (2 C, CH_3), 18.1 (CH_3). HRMS (ESI) m/z : calcd for $\text{C}_{106}\text{H}_{167}\text{N}_{11}\text{O}_{62}$ 2587.02851; found 885.33588 [$\text{M}+3\text{Na}$] $^{3+}$, 1316.51256 [$\text{M}+2\text{Na}$] $^{2+}$. MS (MALDI) m/z : calcd for $\text{C}_{110}\text{H}_{168}\text{N}_8\text{O}_{62}$ 2587.0; found 2610.6 [$\text{M}+\text{Na}$] $^{+}$ (matrix DHB).

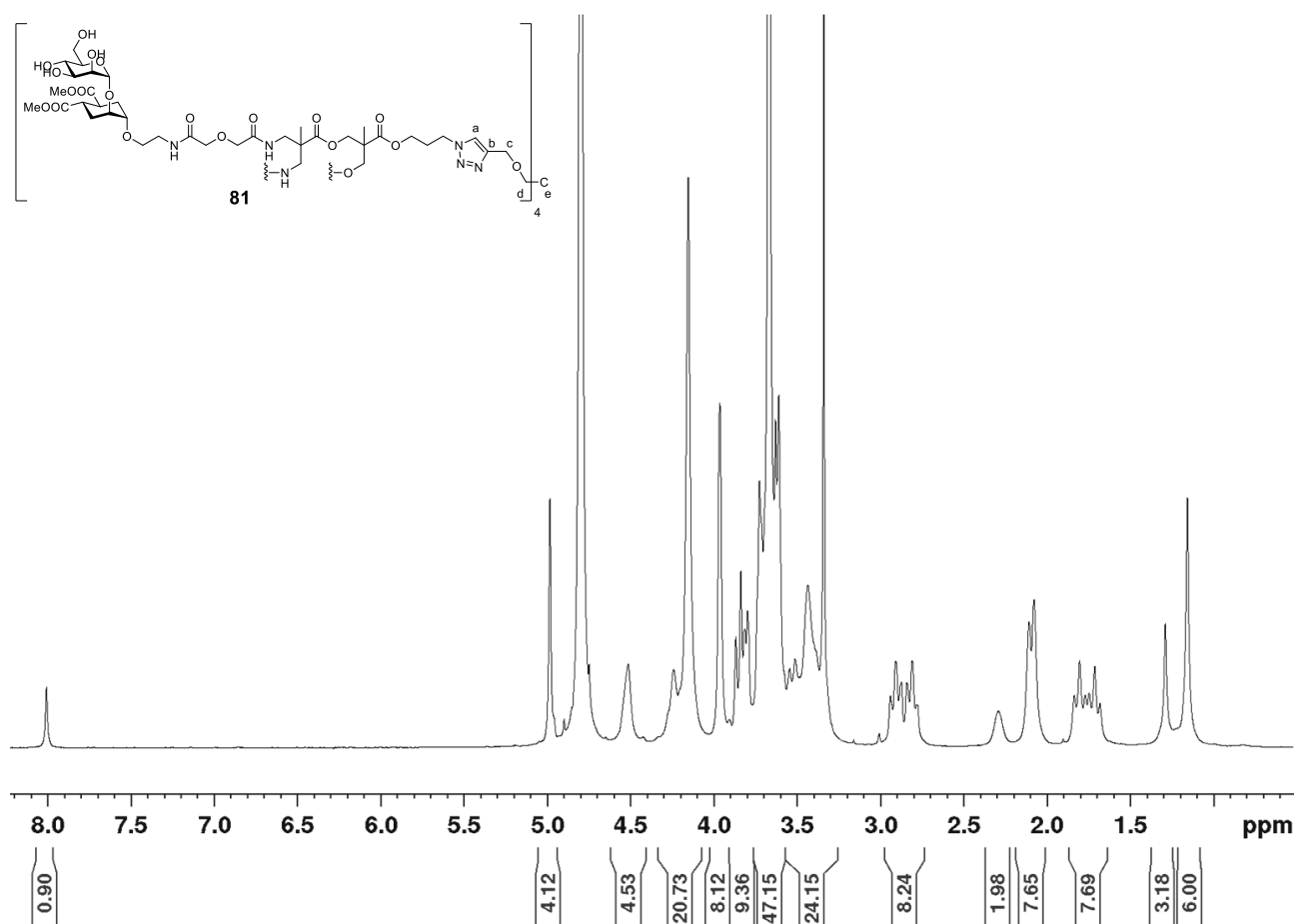
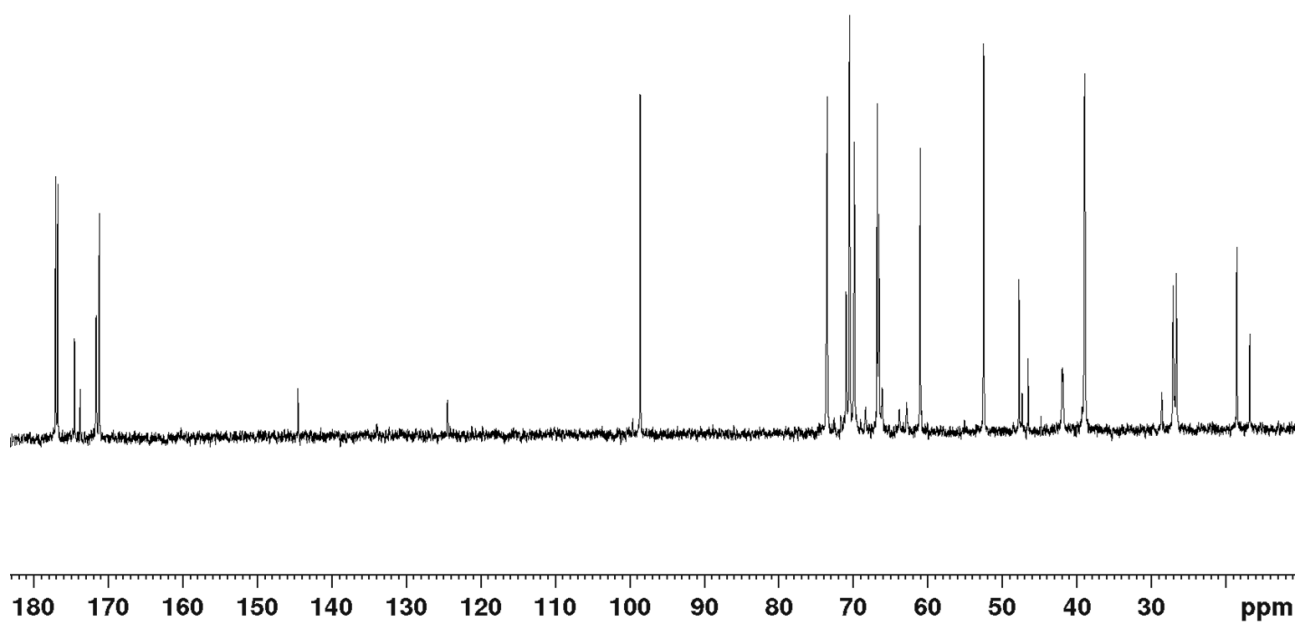
^1H NMR spectrum of 76b in CD_3OD (400 MHz) ^{13}C NMR spectrum of 76b in CD_3OD (100 MHz)

HPLC purification of analytical sample of 76b: HPLC trace. Waters Atlantis 21 mm x 10 cm column, gradient from 100% (H_2O + 0.1% HCOOH) to 25% (H_2O + 0.1% HCOOH) - 75% (CH_3CN + 0.1% HCOOH) in 12 min, $t_R = 7.6$ min.

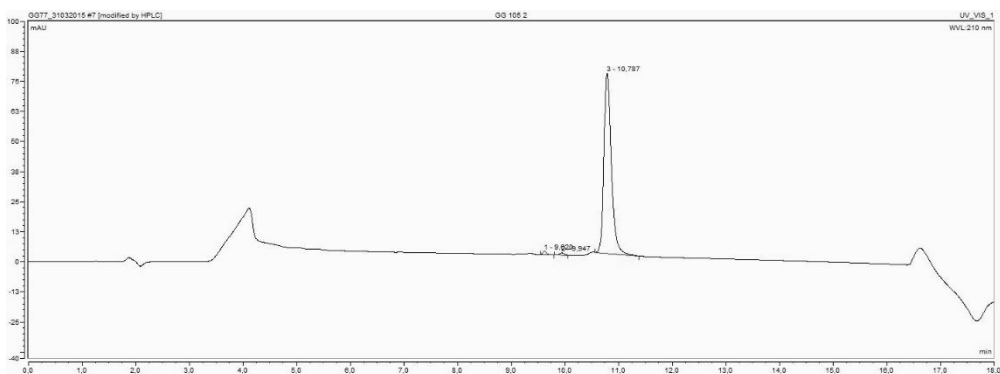


Synthesis of compound **81**

The tetravalent scaffold **80** (740 μg , 2.6 μmol) was dissolved in freshly distilled THF (125 μL) under nitrogen atmosphere. A solution of TBTA (270 μg , 0.5 μmol) in distilled THF (10 μL) was added, followed by 5 μL of a solution of $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ (64 μg , 0.26 μmol) and 10 μL of a solution of sodium ascorbate (200 μg , 1 μmol) both in degassed H_2O . After stirring for 10 min, the mixture changed from colourless to light yellow and finally a solution of dendron **76b** (29.2 mg, 11.3 μmol) in degassed H_2O (100 μL) was added. The mixture was stirred at room temperature, under nitrogen atmosphere, sheltered by light for 15 h. Monitoring with MS (MALDI) (matrix DHB, HCCA) formation of the product along with the di- and tri-substituted intermediates was observed thus **76b** (14.2 mg, 5.5 μmol) and sodium ascorbate (200 μg , 1 μmol) were further added as solution in degassed H_2O . The mixture was stirred as before for additional 15 h, after which MS (MALDI) analysis (matrix DHB, HCCA) confirmed the unique formation of the desired product. The copper scavenger QuadraSil MP was added to the solution which was stirred for 20 min. After filtering, the crude was finally purified by size-exclusion chromatography using a Sephadex LH-20 column ($\text{O} = 3 \text{ cm}$, height = 50 cm; eluent: MeOH) and monitoring by TLC (eluent: $\text{CHCl}_3 - \text{MeOH} - \text{H}_2\text{O}$, 7:3:0.5; $R_f = 0.11$). Dendrimer **81** was recovered as a pure white foam (24.6 mg, 90%). The purity of the compound was assessed by HPLC (Waters Atlantis 21 mm x 10 cm column, gradient from 100% ($\text{H}_2\text{O} + 0.1\% \text{HCOOH}$) to 50% ($\text{H}_2\text{O} + 0.1\% \text{HCOOH}$) - 50% ($\text{CH}_3\text{CN} + 0.1\% \text{HCOOH}$) in 15 min, t_R (product) = 10.8 min). $[\alpha]_D^{27} = +31$ ($c = 0.50$ in MeOH). $^1\text{H NMR}$ (400 MHz, D_2O) δ (ppm): 8.01 (s, 4 H, H_a), 4.99 (s, 16 H, 1-H), 4.32-4.23 (m, 16 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}$, H_c), 4.29-4.13 (m, 88 H, CH_2O , CH_2CONH , $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}$), 3.99-3.94 (m, 32 H, 2-H, 2- H_D), 3.92-3.79 (m, 32 H, 6-H, 3-H), 3.76-3.57 (m, 96 H, 6- H' , 1- H_D , $\text{OCH}_2\text{CH}_2\text{N}$, 5-H, 4-H), 3.66 (s, 48 H, OCH_3), 3.65 (s, 48 H, OCH_3), 3.56-3.32 (m, 72 H, CH_2N , $\text{OCH}_2\text{CH}_2\text{N}$, H_d), 2.91 (td, $J = 12.0, 3.2 \text{ Hz}$, 16 H, 5-H), 2.81 (td, $J = 12.0, 3.2 \text{ Hz}$, 16 H, 4-H), 2.36-2.26 (m, 8 H, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}$), 2.16-2.06 (m, 32 H, 6- H_{Deq} , 3- H_{Deq}), 1.80 (t, $J = 13.1 \text{ Hz}$, 16 H, 3- H_{Dax}), 1.71 (t, $J = 12.9 \text{ Hz}$, 16 H, 6- H_{Dax}), 1.29 (s, 12 H, CH_3), 1.16 (s, 24 H, CH_3). $^{13}\text{C NMR}$ (100 MHz, D_2O) δ (ppm): 177.1 (C=O), 176.9 (C=O), 174.6 (C=O), 173.8 (C=O), 171.7 (C=O), 171.3 (C=O), 144.5 (4C, C_b), 124.4 (4C, C_a), 98.6 (16 C, C_1), 73.5 (16 C, C_{1D}), 73.4 (16 C, C_4), 70.9, 70.5 (48 C, C_2 , C_{2D} , C_3), 69.8 (32 C, CH_2CONH), 68.3 (4 C, C_d), 66.7 (16 C, C_5), 66.5 (16 C, $\text{OCH}_2\text{CH}_2\text{N}$), 66.1 (8 C, CH_2O), 63.8 (4 C, C_c), 62.8 (4 C, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}$), 61.0 (16 C, C_6), 52.5 (16 C, OCH_3), 52.4 (16 C, OCH_3), 47.7 (8 C, CCH_2N), 47.3 (4 C, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}$), 46.5 (4 C, CCH_2O), 44.8 (C_e), 41.9 (16 C, CH_2N), 39.0 (32 C, C_5 , C_4), 38.9 (16 C, $\text{OCH}_2\text{CH}_2\text{N}$), 28.6 (4 C, $\text{OCH}_2\text{CH}_2\text{CH}_2\text{N}$), 27.0 (16 C, C_3), 26.6 (16 C, C_6), 18.6 (8 C, CH_3), 16.8 (4 C, CH_3). HRMS (ESI) m/z : calcd for $\text{C}_{441}\text{H}_{688}\text{N}_{44}\text{O}_{252}$ 10637.25277; found 1795.93886 $[\text{M}+6\text{Na}]^{6+}$, 2150.50807 $[\text{M}+5\text{Na}]^{5+}$, 2682.45649 $[\text{M}+4\text{Na}]^{4+}$.

¹H NMR spectrum of 81 in D₂O (400 MHz)**¹³C NMR spectrum of 81 in D₂O (100 MHz)**

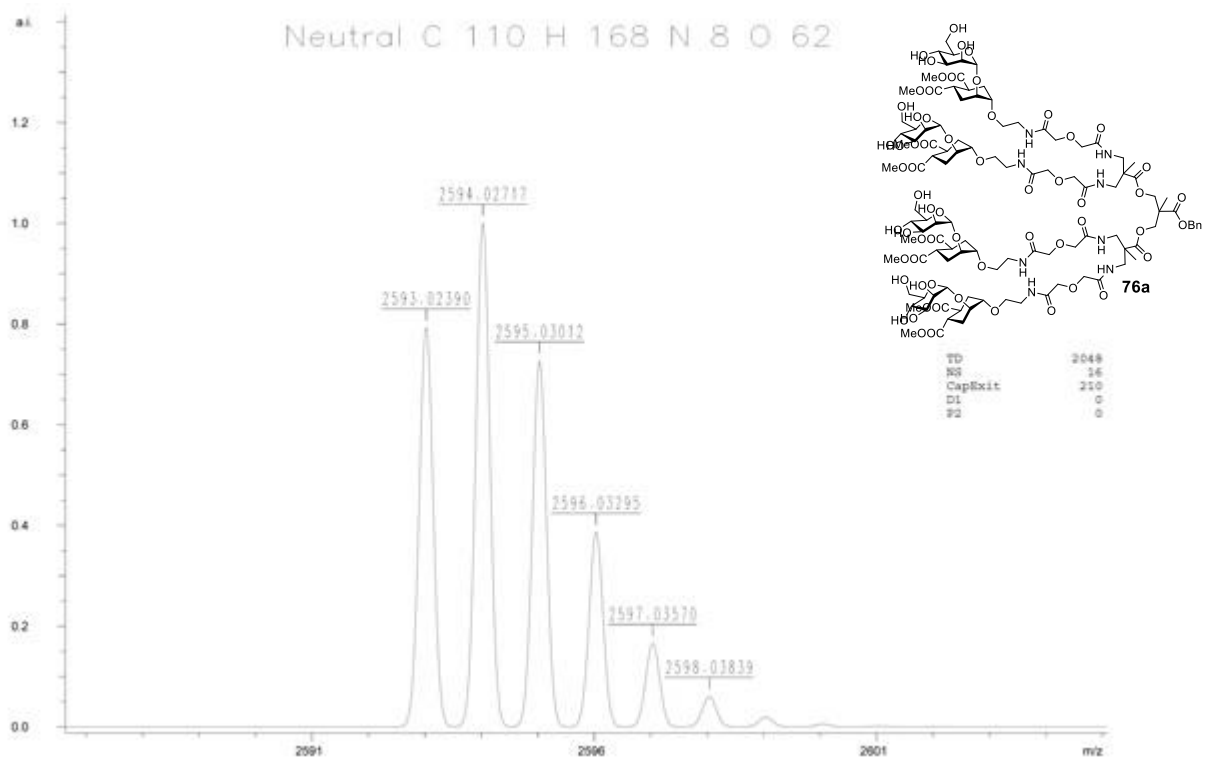
HPLC purification of analytical sample of 81: HPLC trace. Waters Atlantis 21 mm x 10 cm column, gradient from 100% (H₂O + 0.1% HCOOH) to 50% (H₂O + 0.1% HCOOH) - 50% (CH₃CN + 0.1% HCOOH) in 15 min, *t_R* = 10.8 min.



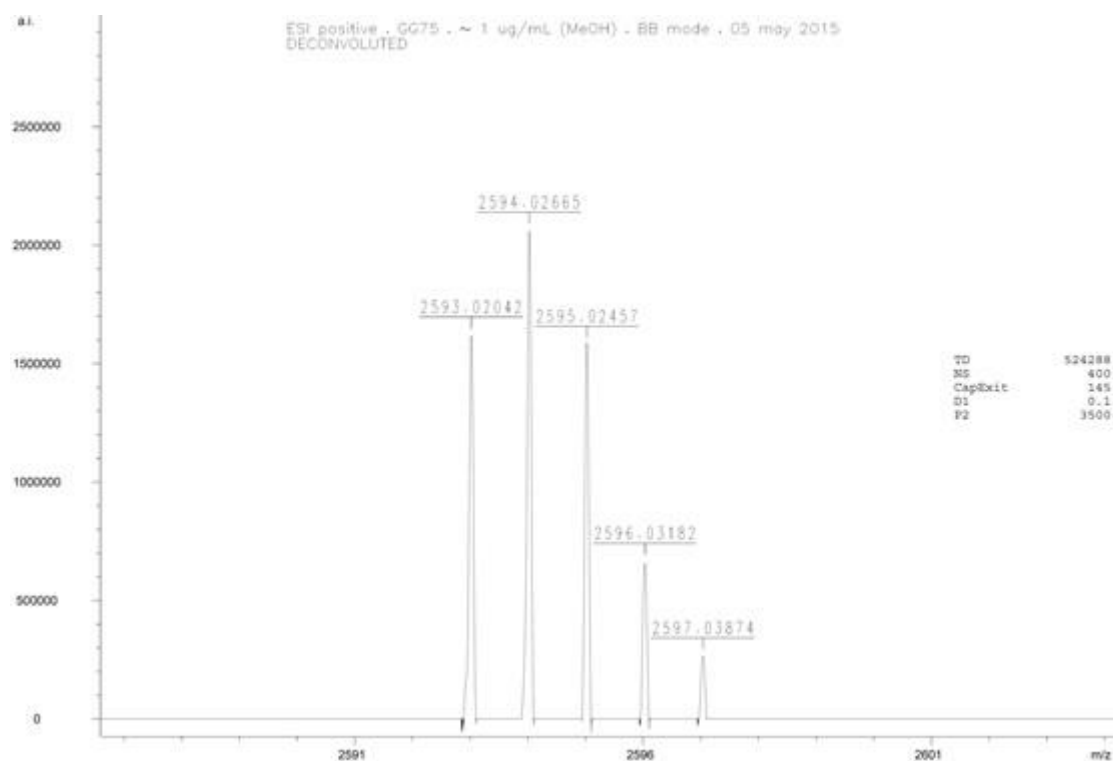
1.6.2.7 Isotopic distribution of dendrons

Calculated (a) and found (b) isotopic distribution for dendron 76a HRMS (ESI):

a) Calculated:



b) Found:

Calcd for neutral $[C_{110}H_{168}N_8O_{62}]$: 2594.02717.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	2593.02390	2593.02042	0.7923	0.7850
2	2594.02717	2594.02665	1.0000	1.0000
3	2595.03012	2595.02457	0.7269	0.7661
4	2596.03295	2596.03182	0.3864	0.3218
5	2597.03570	2597.03874	0.1658	0.1279
6	2598.03839	Not detected	0.0604	Not detected

Calcd for $[C_{110}H_{168}N_8O_{62}Na_2]^{2+}$: 2640.00563.

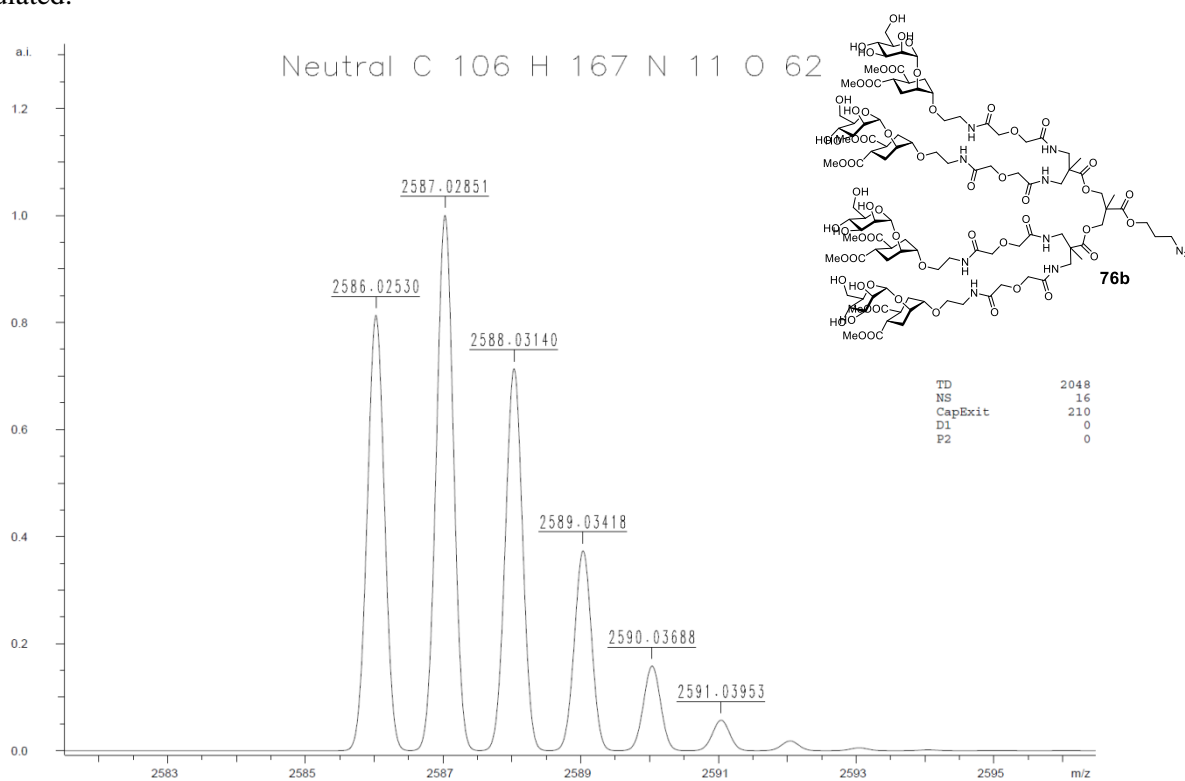
Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1319.50117	1319.49999	0.7923	0.8454
2	1320.00281	1320.00119	1.0000	1.0000
3	1320.50428	1320.50146	0.7269	0.7629
4	1321.00570	1321.00454	0.3864	0.4536
5	1321.50707	1321.50861	0.1658	0.2165
6	1322.00841	Not detected	0.0604	Not detected

Calcd for $[C_{110}H_{168}N_8O_{62}Na_3]^{3+}$: 2662.99485.

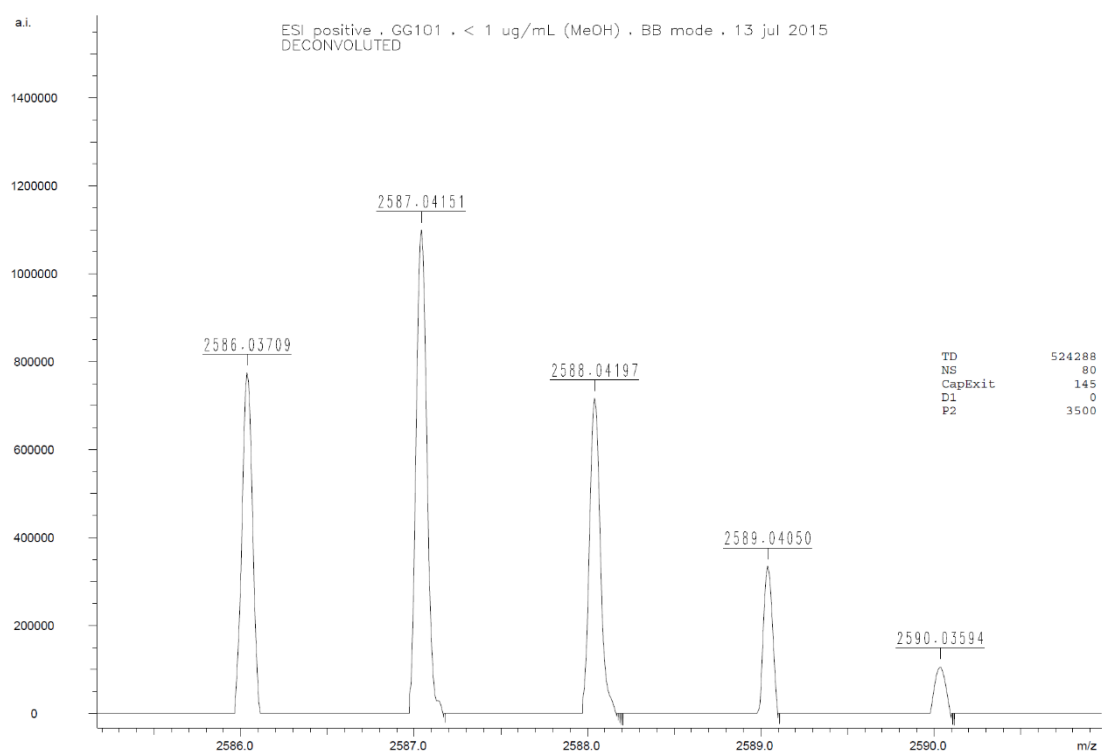
Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	887.33052	887.32894	0.7923	0.6970
2	887.66495	887.66579	1.0000	1.0000
3	887.99926	887.99793	0.7269	0.8485
4	888.33354	Not detected	0.3864	Not detected
5	888.66779	Not detected	0.1658	Not detected
6	889.00202	Not detected	0.0604	Not detected

Calculated (a) and found (b) isotopic distribution for dendron 76b HRMS (ESI):

a) Calculated:



b) Found:



Calcd for neutral [C₁₀₆H₁₆₇N₁₁O₆₂]: 2587.02851.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	2586.02530	2586.03709	0.8131	0.7045
2	2587.02851	2587.04151	1.0000	1.0000
3	2588.03140	2588.04197	0.7134	0.6513
4	2589.03418	2589.04050	0.3732	0.3044
5	2590.03688	2590.03594	0.1578	0.0978
6	2591.03953	Not detected	0.0568	Not detected

Calcd for [C₁₀₆H₁₆₇N₁₁O₆₂Na₂]²⁺: 2633.00697.

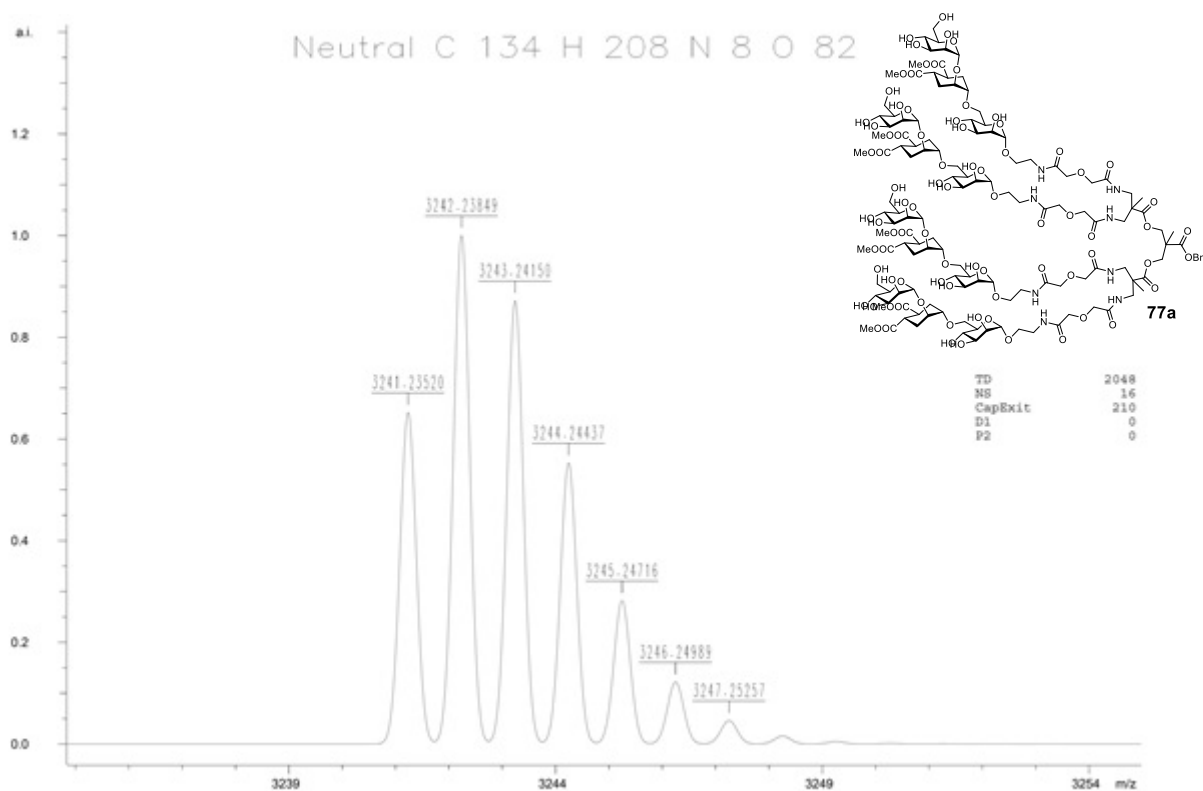
Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1316.00187	1316.00849	0.8131	0.8479
2	1316.50348	1316.51256	1.0000	1.0000
3	1317.00492	1317.01106	0.7133	0.6935
4	1317.50631	1317.51151	0.3732	0.3691
5	1318.00766	Not detected	0.1578	Not detected
6	1318.50899	Not detected	0.0568	Not detected

Calcd for [C₁₀₆H₁₆₇N₁₁O₆₂Na₃]³⁺: 2655.99619.

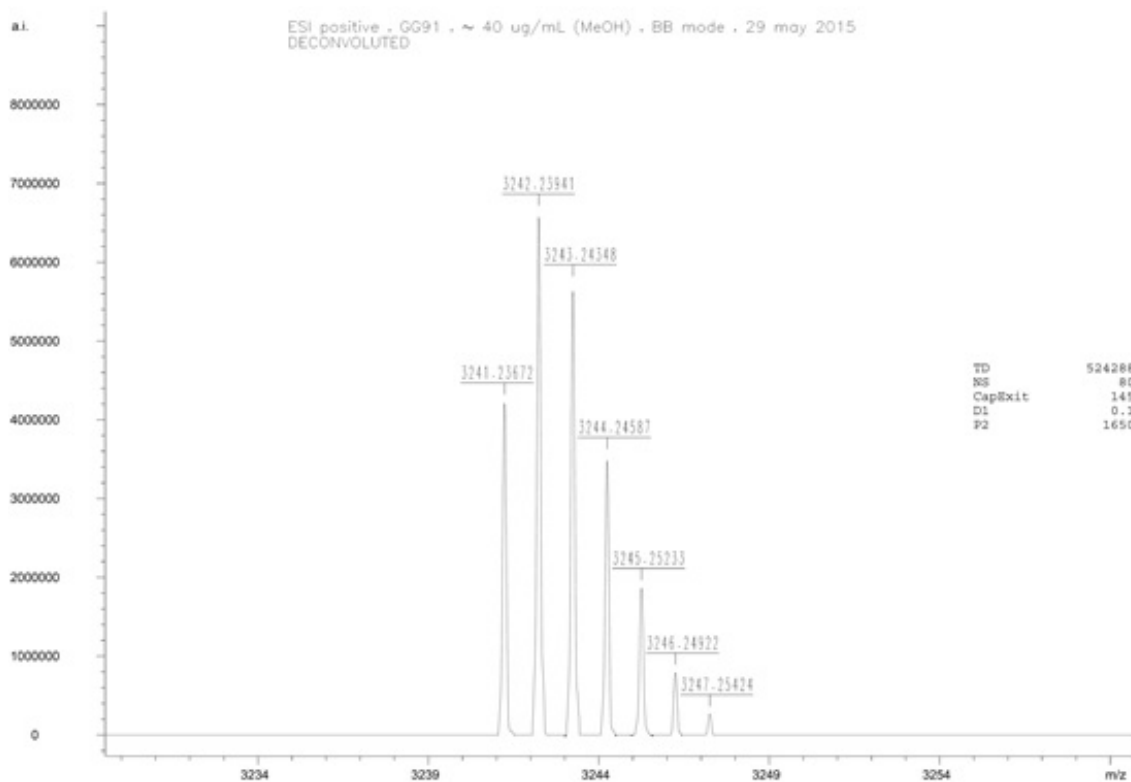
Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	884.99765	885.00161	0.8131	0.5729
2	885.33206	885.33588	1.0000	1.0000
3	885.66635	885.66992	0.7134	0.6432
4	886.00062	886.00203	0.3732	0.3125
5	886.33485	Not detected	0.1579	Not detected
6	886.66906	Not detected	0.0568	Not detected

Calculated (a) and found (b) isotopic distribution for dendron 77a HRMS (ESI):

a) Calculated:



b) Found:



Calcd for neutral $[\text{C}_{134}\text{H}_{208}\text{N}_8\text{O}_{82}]$: 3242.23849.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	3241.23520	3241.23672	0.6519	0.6382
2	3242.23849	3242.23941	1.0000	1.0000
3	3243.24150	3243.24348	0.8717	0.8645
4	3244.24437	3244.24587	0.5528	0.5334
5	3245.24716	3245.25233	0.2819	0.2825
6	3246.24989	3246.24922	0.1219	0.1204
7	3247.25257	3247.25424	0.0461	0.0407

Calcd for $[\text{C}_{134}\text{H}_{208}\text{N}_8\text{O}_{82}\text{Na}_2]^{2+}$: 3288.21694.

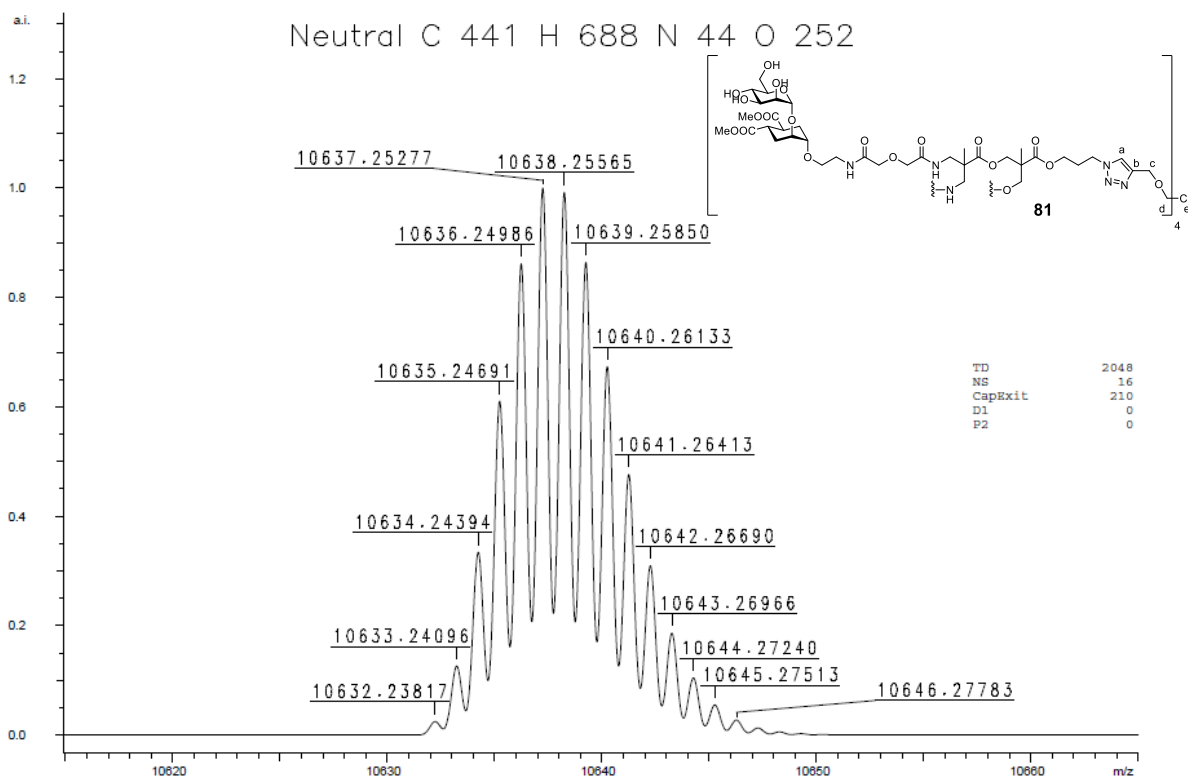
Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1643.60682	1643.60671	0.6519	0.6632
2	1644.10847	1644.10830	1.0000	1.0000
3	1644.60997	1644.61066	0.8717	0.8710
4	1645.11141	1645.11195	0.5527	0.5879
5	1645.61280	1645.61796	0.2819	0.3208
6	1646.11417	Not detected	0.1219	Not detected
7	1646.61551	Not detected	0.0461	Not detected

Calcd for $[\text{C}_{134}\text{H}_{208}\text{N}_8\text{O}_{82}\text{Na}_3]^{3+}$: 3311.20617.

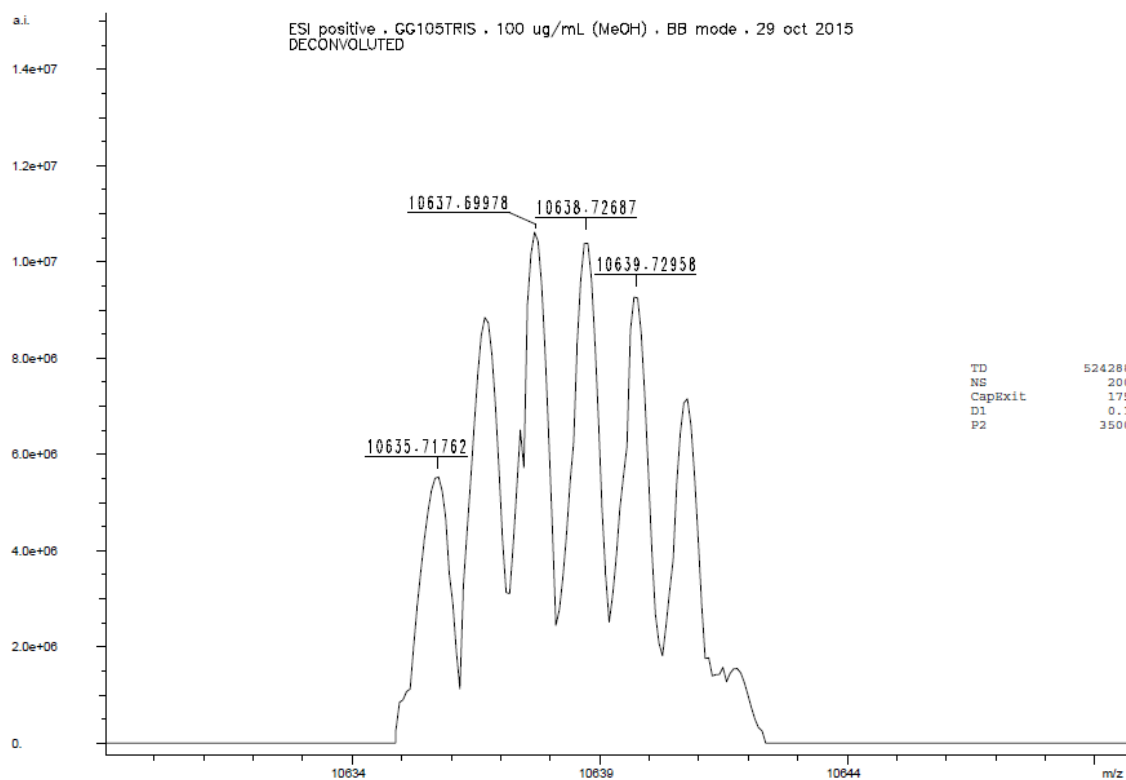
Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1103.40095	1103.40144	0.6519	0.6407
2	1103.73539	1103.73571	1.0000	1.0000
3	1104.06972	1104.07041	0.8717	0.8651
4	1104.40401	1104.40459	0.5528	0.5305
5	1104.73827	1104.73990	0.2819	0.2893
6	1105.07252	1105.07388	0.1218	0.1287
7	1105.40675	1105.40764	0.0461	0.0557

Calculated (a) and found (b) isotopic distribution for dendrimer 81 HRMS (ESI):

a) Calculated:



b) Found:



Calcd for neutral [C₄₄₁H₆₈₈N₄₄O₂₅₂]: 10637.25277.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	10632.23817	Not detected	0.0247	Not detected
2	10633.24096	Not detected	0.1262	Not detected
3	10634.24394	Not detected	0.3343	Not detected
4	10635.24691	10635.71762	0.6104	0.5233
5	10636.24986	10636.69768	0.8615	0.8263
6	10637.25277	10637.69978	1.0000	1.0000
7	10638.25565	10638.72687	0.9923	0.9908
8	10639.25850	10639.72958	0.8643	0.8793
9	10640.26133	10640.74317	0.6736	0.6752
10	10641.26413	Not detected	0.4765	Not detected
11	10642.26690	Not detected	0.3094	Not detected
12	10643.26966	Not detected	0.1861	Not detected
13	10644.27240	Not detected	0.1044	Not detected
14	10645.27513	Not detected	0.0550	Not detected
15	10646.27783	Not detected	0.0274	Not detected

Calcd for [C₄₄₁H₆₈₈N₄₄O₂₅₂Na₄]⁴⁺: 10729.20964.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	2681.04856	Not detected	0.0247	Not detected
2	2681.29937	Not detected	0.1261	Not detected
3	2681.55015	Not detected	0.3343	Not detected
4	2681.80092	Not detected	0.6103	Not detected
5	2682.05167	2682.20456	0.8615	0.8696
6	2682.30241	2682.45649	1.0000	1.0000
7	2682.55314	2682.72037	0.9923	0.9814
8	2682.80386	2682.96842	0.8642	0.9130
9	2683.05458	2683.22542	0.6735	0.7634
10	2683.30528	Not detected	0.4764	Not detected
11	2683.55598	Not detected	0.3094	Not detected
12	2683.80668	Not detected	0.1861	Not detected
13	2684.05737	Not detected	0.1044	Not detected
14	2684.30806	Not detected	0.0550	Not detected
15	2684.55874	Not detected	0.0274	Not detected

Calcd for [C₄₄₁H₆₈₈N₄₄O₂₅₂Na₅]⁵⁺: 10752.20005.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	2149.45359	Not detected	0.0257	Not detected

2	2149.64236	Not detected	0.1274	Not detected
3	2149.84072	Not detected	0.3354	Not detected
4	2150.04018	2150.09093	0.6110	0.5714
5	2150.24001	2150.29852	0.8617	0.8276
6	2150.44001	2150.50807	1.0000	1.0000
7	2150.64010	2150.72202	0.9926	0.9704
8	2150.84025	2150.92523	0.8649	0.9015
9	2151.04043	2151.13063	0.6744	0.7291
10	2151.24063	2151.33624	0.4773	0.4926
11	2151.44085	Not detected	0.3102	Not detected
12	2151.64109	Not detected	0.1867	Not detected
13	2151.84132	Not detected	0.1049	Not detected
14	2152.04157	Not detected	0.0553	Not detected
15	2152.24182	Not detected	0.0275	Not detected
16	2152.44207	Not detected	0.0130	Not detected
17	2152.64232	Not detected	0.0058	Not detected

Calcd for $[C_{441}H_{688}N_{44}O_{252}Na_6]^{6+}$: 10775.18820.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1795.02968	Not detected	0.0248	Not detected
2	1795.19640	1795.29360	0.1263	0.2393
3	1795.36342	1795.45771	0.3344	0.3869
4	1795.53050	1795.61647	0.6105	0.6098
5	1795.69759	1795.77580	0.8615	0.8820
6	1795.86470	1795.93886	1.0000	0.9967
7	1796.03181	1796.10768	0.9924	1.0000
8	1796.19892	1796.27374	0.8644	0.8656
9	1796.36603	1796.44277	0.6737	0.6885
10	1796.53313	1796.60886	0.4765	0.5180
11	1796.70024	1796.77841	0.3095	0.3443
12	1796.86734	1796.94738	0.1861	0.2393
13	1797.03444	Not detected	0.1045	Not detected
14	1797.20154	Not detected	0.0551	Not detected

1.6.2.8 SPR inhibition assay experiments

Dendron affinity was evaluated both for murine MBL-A and MBL-C isoforms contained in murine plasma. Mouse plasma was diluted 100-folds in running buffer (10 mM TRIS buffer containing 150 mM NaCl, 1.2 mM $CaCl_2$, and 0.005% Tween 20, pH 7.4), preincubated for 30 min at 25 °C with or without glycodendrimers at different concentrations, and then flowed simultaneously onto immobilized Man-BSA and BSA. The availability of six parallel flow channels allowed to analyze the effect of four inhibitors concentrations simultaneously, leaving one channel for mouse plasma preincubated with vehicle alone and one channel for MBL-/- mice plasma. Anti-MBL-A antibody (10 μ g/mL, Hycult Biotech, The Netherlands) and anti-MBL-C antibody (10 μ g/mL, Hycult Biotech, The Netherlands) were subsequently injected to specifically recognize

and quantify the amount of free mMBL-A and mMBL-C bound to mannose residues. We considered the SPR signal measured 100 s after antibody's injection. (More details for the assay: ref 52)

The inhibitors effect was calculated as follows:

$$\%INH - SB = \frac{TB - BI}{TB - NSB} * 100$$

where: %INH-SB = percentage of inhibition of either mMBL-A or mMBL-C specific binding to Man-BSA; TB = total binding i.e. maximal mMBL-C binding to Man-BSA (plasma preincubated with vehicle); NSB = non-specific binding signal (plasma from MBL-/- mice); BI = binding signal observed in the presence of inhibitor. Inhibition curves, i.e. %INH-SB versus log[inhibitor], were fitted using the 'one-site competition' equation built into Prism version 6.0 for Windows (Graphpad Software, San Diego, CA). This analysis gives the IC₅₀ value with its 95% confidence intervals.

Diluted plasma from WT mice was preincubated with different concentrations of dendrons, or vehicle, and then injected over MAN-BSA immobilized on the sensor chip. Diluted plasma from MBL-/- mice was injected in parallel. Anti-mMBL-A and anti-mMBL-C antibodies were subsequently injected for 2 min (bars), producing the sensorgrams (i.e. SPR signal in Resonance Units, RU, vs time). The SPR signal is proportional to the amount of MBL bound to mannose, which in turn is proportional to the amount of free MBL in the plasma (i.e. not bound by the inhibitor).

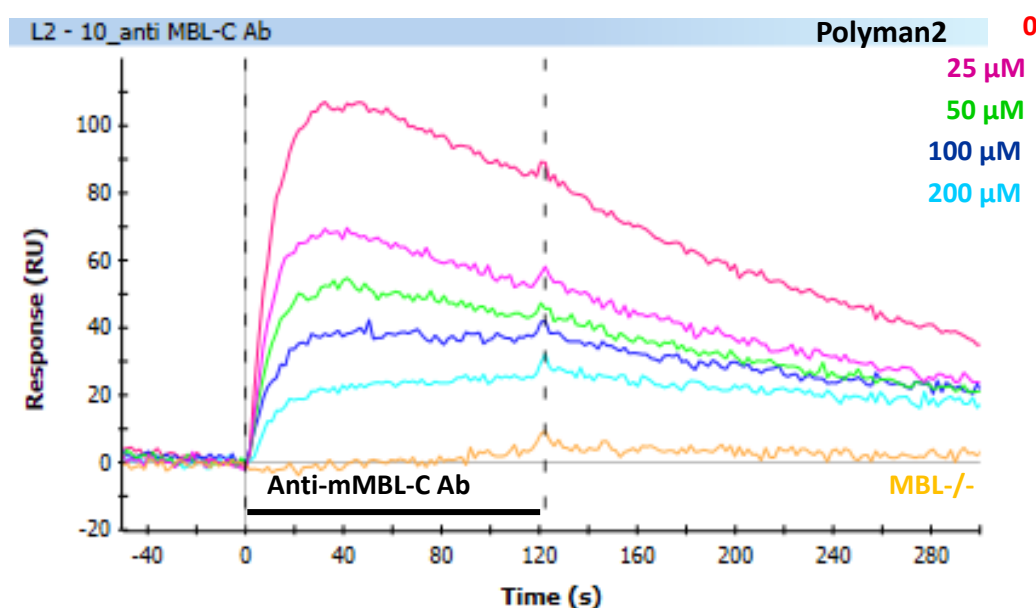


Figure 1.12 - Sensorgram of Polyman2 37 inhibitory effect on mMBL-C. Polyman2 37 dendron was incubated at different concentrations (25 μM, 50 μM, 100 μM, 200 μM) with murine plasma. The solutions were then flowed over a sensor chip and binding of MBL-C to immobilized Man-BSA was revealed flowing anti-MBL-C antibodies. From the concentration-dependent SPR signals an IC₅₀ = 84 μM was calculated.

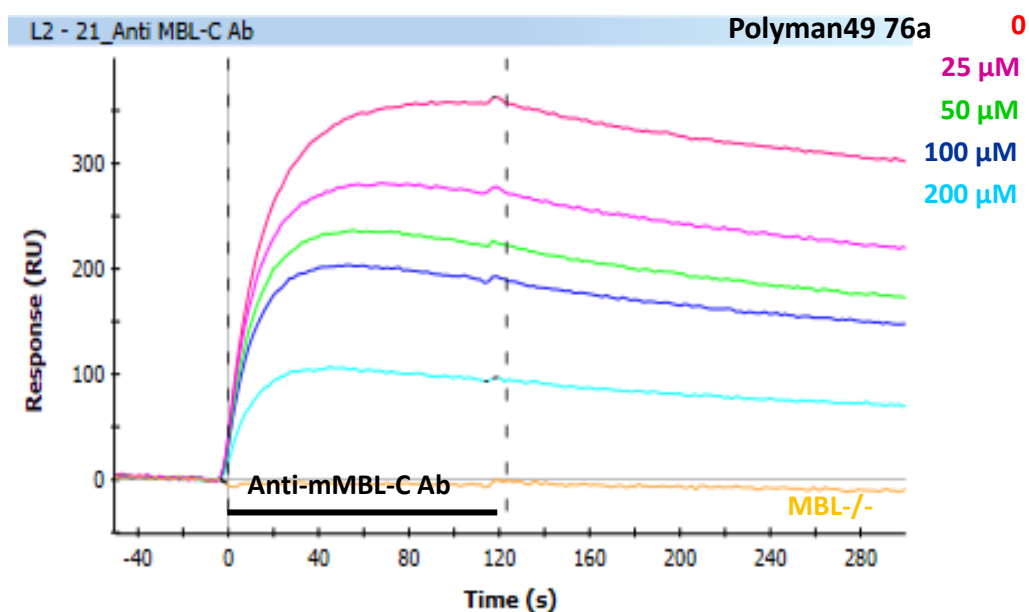


Figure 1.13 - Sensorgram of Polyman49 76a inhibitory effect on mMBL-C. Polyman49 76a dendron was incubated at different concentrations (25 μM , 50 μM , 100 μM , 200 μM) with murine plasma. The solutions were then flowed over a sensor chip and binding of MBL-C to immobilized Man-BSA was revealed flowing anti-MBL-C antibodies. From the concentration-dependent SPR signals an $\text{IC}_{50} = 77 \mu\text{M}$ was calculated.

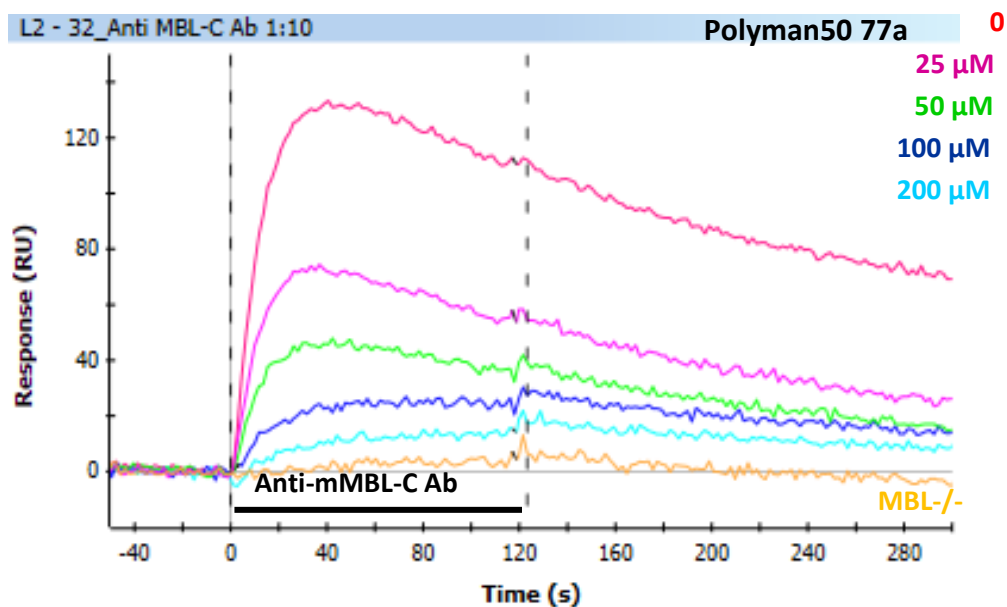


Figure 1.14 - Sensorgram of Polyman50 77a inhibitory effect on mMBL-C. Polyman50 77a dendron was incubated at different concentrations (25 μM , 50 μM , 100 μM , 200 μM) with murine plasma. The solutions were then flowed over a sensor chip and binding of MBL-C to immobilized Man-BSA was revealed flowing anti-MBL-C antibodies. From the concentration-dependent SPR signals an $\text{IC}_{50} = 35 \mu\text{M}$ was calculated.

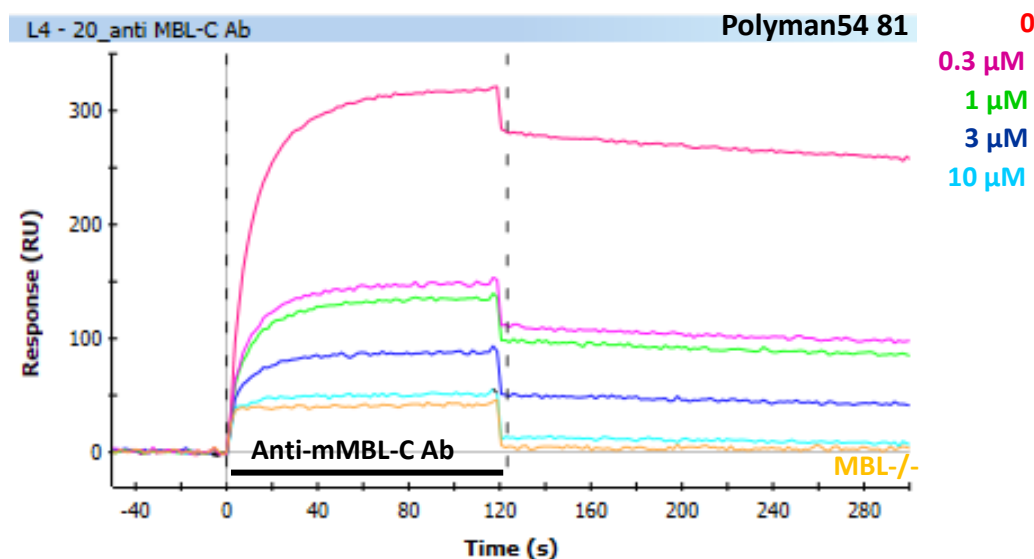


Figure 1.15 - Sensorgram of Polyman54 81 inhibitory effect on mMBL-C. Polyman54 81 dendrimer was incubated at different concentrations (0.3 μM , 1 μM , 3 μM , 10 μM) with murine plasma. The solutions were then flowed over a sensor chip and binding of MBL-C to immobilized Man-BSA was revealed flowing anti-MBL-C antibodies. From the concentration-dependent SPR signals an $\text{IC}_{50} = 1.0 \mu\text{M}$ was calculated.

1.7 References

- ¹ T. Fujita *Nat. Rev. Immunol.* **2002**, 2, 346-353.
- ² W. K. E. Ip, K. Takahashi, R. A. Ezekowitz, L. M. Stuart *Immunol. Rev.* **2009**, 230, 9-21.
- ³ J. Dobó, G. Pál, L. Cervenak, P. Gál *Immunol. Rev.* **2016**, 274, 98-111.
- ⁴ P. Garred, N. Genster, K. Pilely, R. Bayarri-Olmos, A. Rosbjerg, Y. J. Ma, M.-O. Skjoedt *Immunol. Rev.* **2016**, 274, 74-97.
- ⁵ N. V. Dudkina, B. A. Spicer, C. F. Reboul, P. J. Conroy, N. Lukoyanova¹, H. Elmlund, R. H. P. Law, S. M. Ekkel, S. C. Kondos, R. J. A. Goode, G. Ramm, J. C. Whisstock, H. R. Saibil, M. A. Dunstone *Nat. Commun.* **2016**, 7, 10588.
- ⁶ M. Serna, J. L. Giles, B. P. Morgan, D. Bubeck *Nat. Commun.* **2016**, 7, 10587.
- ⁷ T. H. Sharp, A. J. Koster, P. Gros *Cell Rep.* **2016**, 15, 1-8.
- ⁸ P. Garred, S. Thiel, H. O. Madsen, L. P. Ryder, J. C. Jensenius, A. Svejgaard *Clin. Exp. Immunol.* **1992**, 90, 517-521.
- ⁹ S. E. Degn, L. Jensen, A. G. Hansen, D. Duman, M. Tekin, J. C. Jensenius, S. Thiel *J. Immunol.* **2012**, 189, 3957-3969.
- ¹⁰ J. Seyfarth, P. Garred, H. O. Madsen *Hum. Mol. Genet.* **2005**, 14, 2859-2869.
- ¹¹ W. I. Weis, K. Drickamer *Structure* **1994**, 2, 1227-1240.
- ¹² S. Sheriff, C. Y. Y. Chang, R. A. B. Ezekowitz *Nat. Struct. Biol.* **1994**, 1, 789-794.
- ¹³ A. Miller, A. Phillips, J. Gor, R. Wallis, S. J. Perkins *J. Biol. Chem.* **2012**, 287, 3930-3945.
- ¹⁴ P. Garred, F. Larsen, H. O. Madsen, C. Koch *Mol. Immunol.* **2003**, 40, 73-84.
- ¹⁵ T. Mogue, T. Ota, A. I. Tauber, K. N. Sastry *Glycobiology* **1996**, 6, 543-550.
- ¹⁶ R. Wallis, K. Drickamer *J. Biol. Chem.* **1999**, 274, 3580-3589.
- ¹⁷ V. Y. Polotsky, J. T. Belisle, K. Mikusova, R. A. B. Ezekowitz, K. A. J. Joiner *Infect. Dis.* **1997**, 175, 1159-1168.
- ¹⁸ V. Y. Polotsky, W. Fischer, R. A. B. Ezekowitz, K. A. Joiner *Infect. Immun.* **1996**, 64, 380-383.
- ¹⁹ J. Nadesalingam, A. W. Dodds, K. B. M. Reid, N. Palaniyar *J. Immunol.* **2005**, 175, 1785-1794.
- ²⁰ L. Shi, K. Takahashi, J. Dundee, S. Shahroor-Karni, S. Thiel, J. C. Jensenius, F. Gad, M. R. Hamblin, K. N. Sastry, R. A. B. Ezekowitz *J. Exp. Med.* **2004**, 199, 1379-1390.
- ²¹ E. M. Anders, C. A. Hartley, D. C. Jackson *Proc. Nat. Acad. Sci. U.S.A.* **1990**, 87, 4485-4489.
- ²² M. Gadjeva, S. R. Paludan, S. Thiel, V. Slavov, M. Ruseva, K. Eriksson, G.-B. Löwhagen, L. Shi, K. Takahashi, A. Ezekowitz, J. C. Jensenius *Clin. Exp. Immunol.* **2004**, 138, 304-311.
- ²³ M. C. Ghezzi, G. Raponi, S. Angeletti, C. Mancini *J. Infect. Dis.* **1998**, 178, 1743-1749.
- ²⁴ A. L. Laursen, N. Obel, U. Holmskov, J. C. Jensenius, E. M. Aliouat, P. L. Andersen *APMIS* **2003**, 111, 405-415.
- ²⁵ A. J. Nauta, N. Raaschou-Jensen, A. Roos, M. R. Daha, H. O. Madsen, M. C. Borrias-Essers, L. P. Ryder, C. Koch, P. Garred *Eur. J. Immunol.* **2003**, 33, 2853-2863.
- ²⁶ M. C. Walsh, T. Bourcier, K. Takahashi, L. Shi, M. N. Busche, R. P. Rother, S. D. Solomon, R. A. B. Ezekowitz, G. L. Stahl *J. Immunol.* **2005**, 175, 541-546.
- ²⁷ M. Trendelenburg, P. Theroux, A. Stebbins, C. Granger, P. Armstrong, M. Pfisterer *Eur. Heart J.* **2010**, 31, 1181-1187.

- ²⁸ M. N. Busche, V. Pavlov, K. Takahashi, G. L. Stahl *Am. J. Physiol.: Heart Circ. Physiol.* **2009**, *297*, H1853-9.
- ²⁹ J. E. Jordan, M. C. Montalto, G. L. Stahl *Circulation* **2001**, *104*, 1413-1418.
- ³⁰ R. K. Chan, S. I. Ibrahim, K. Takahashi, E. Kwon, M. McCormack, A. Ezekowitz, M. C. Carroll, F. D., Jr Moore, W. G., Jr. Austen *J. Immunol.* **2006**, *177*, 8080-8085.
- ³¹ M. L. Hart, K. A. Ceonzo, L. A. Shaffer, K. Takahashi, R. P. Rother, W. R. Reenstra, J. A. Buras, G. L. Stahl *J. Immunol.* **2005**, *174*, 6373-6380.
- ³² R. A. Matthijsen, J. P. M. Derikx, R. Steffensen, R. M. van Dam, C. H. C. Dejong, W. A. Buurman *Mol. Immunol.* **2009**, *46*, 2244-2248.
- ³³ B. de Vries, S. J. Walter, C. J. Peutz-Kootstra, T. G. A. M. Wolfs, L. W. E. van Heurn, W. A. Buurman *Am. J. Pathol.* **2004**, *165*, 1677-1688.
- ³⁴ R. Gesuete, C. Storini, A. Fantin, M. Stravalaci, E. R. Zanier, F. Orsini, H. Vietsch, M. L. M. Mannesse, B. Ziere, M. Gobbi, M.-G. De Simoni *Ann. Neurol.* **2009**, *66*, 332-342.
- ³⁵ A. Cervera, A. M. Planas, C. Justicia, X. Urrea, J. C. Jensenius, F. Torres, F. Lozano, A. Chamorro *PLoS ONE* **2010**, *2*, e8433.
- ³⁶ For a report from the American Heart Association of the Heart Disease and Stroke Statistics - 2016 Update: D. Mozaffarian, E. J. Benjamin, A. S. Go, D. K. Arnett, M. J. Blaha, M. Cushman, S. R. Das, S. de Ferranti, J.-P. Després, H. J. Fullerton, V. J. Howard, M. D. Huffman, C. R. Isasi, M. C. Jiménez, S. E. Judd, B. M. Kissela, J. H. Lichtman, L. D. Lisabeth, S. Liu, R. H. Mackey, D. J. Magid, D. K. McGuire, E. R. Mohler III, C. S. Moy, P. Muntner, M. E. Mussolino, K. Nasir, R. W. Neumar, G. Nichol, L. Palaniappan, D. K. Pandey, M. J. Reeves, C. J. Rodriguez, W. Rosamond, P. D. Sorlie, J. Stein, A. Towfighi, T. N. Turan, S. S. Virani, D. Woo, R. W. Yeh, M. B. Turner on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee *Circulation* **2016**, *133*, 447-454.
- ³⁷ F. Orsini, P. Villa, S. Parrella, R. Zangari, E. R. Zanier, R. Gesuete, M. Stravalaci, S. Fumagalli, R. Ottria, J. J. Reina, A. Paladini, E. Micotti, R. Ribeiro-Viana, J. Rojo, V. I. Pavlov, G. L. Stahl, A. Bernardi, M. Gobbi, M.-G. De Simoni *Circulation* **2012**, *126*, 1484-1494.
- ³⁸ S. Sattin, A. Daggetti, M. Thépaut, A. Berzi, M. Sánchez-Navarro, G. Tabarani, J. Rojo, F. Fieschi, M. Clerici, A. Bernardi *ACS Chem. Biol.* **2010**, *5*, 301-312.
- ³⁹ W. I. Weis, K. Drickamer, W. A. Hendrickson *Nature* **1992**, *360*, 127-134.
- ⁴⁰ K. K.-S. Ng, K. Drickamer, W. I. Weis *J. Biol. Chem.* **1996**, *271*, 663-674.
- ⁴¹ K. K.-S. Ng, A. R. Kolatkar, S. Park-Snyder, H. Feinberg, D. A. Clark, K. Drickamer, W. I. Weis *J. Biol. Chem.* **2002**, *277*, 16088-16095.
- ⁴² N. Varga, I. Sutkeviciute, R. Ribeiro-Viana, A. Berzi, R. Ramdasi, A. Daggetti, G. Vettoretti, A. Amara, M. Clerici, J. Rojo, F. Fieschi, A. Bernardi *Biomaterials* **2014**, *35*, 4175-4184.
- ⁴³ S. Ordanini, N. Varga, V. Porkolab, M. Thépaut, L. Belvisi, A. Bertaglia, A. Palmioli, A. Berzi, D. Trabattoni, M. Clerici, F. Fieschi, A. Bernardi *Chem. Commun.* **2015**, *51*, 3816-3819.
- ⁴⁴ J. Luczkowiak, S. Sattin, I. Sutkevičiūtė, J. J. Reina, M. Sánchez-Navarro, M. Thépaut, L. Martínez-Prats, A. Daggetti, F. Fieschi, R. Delgado, A. Bernardi, J. Rojo *Bioconjugate Chem.* **2011**, *22*, 1354-1365.
- ⁴⁵ M. Touaibia, A. Wellens, T. C. Shiao, O. Wang, S. Sirois, J. Bouckaert, R. Roy *ChemMedChem* **2007**, *2*, 1190-1201.
- ⁴⁶ The synthesis and characterisation of Polyman31 **42** has never been reported in the literature and is described in the Experimental part of Chapter 1.
- ⁴⁷ C. Colombo, A. Bernardi *Eur. J. Org. Chem.* **2011**, 3911-3919.
- ⁴⁸ H. Ihre, A. Hult, J. M. J. Frechet, I. Gitsov *Macromolecules* **1998**, *31*, 4061-4068.
- ⁴⁹ S. Hernández-Ainsa1, R. Alcalá, J. Barberá, M. Marcos, C. Sánchez, J. L. Serrano *Macromolecules* **2010**, *43*, 2660-2663; yield 51%.
- ⁵⁰ J. S. Moore, S. I. Stupp *Macromolecules* **1990**, *23*, 65-70.
- ⁵¹ Q. H. Sodji, V. Patil, J. R. Kornacki, M. Mrksich, A. K. Oyelere *J. Med. Chem.* **2013**, *56*, 9969-9981.
- ⁵² M. Stravalaci, D. De Blasio, F. Orsini, C. Perego, A. Palmioli, G. Goti, A. Bernardi, M.-G. De Simoni, M. Gobbi *J. Biomol. Screening* **2016**, *21*, 749-757.
- ⁵³ J. Luczkowiak, A. Muñoz, M. Sánchez-Navarro, R. Ribeiro-Viana, A. Ginieis, B. M. Illescas, N. Martín, R. Delgado, J. Rojo *Biomacromolecules* **2013**, *14*, 431-437.
- ⁵⁴ G. Goti, A. Palmioli, M. Stravalaci, S. Sattin, M.-G. De Simoni, M. Gobbi, A. Bernardi *Chem. Eur. J.* **2016**, *22*, 3686-3691.
- ⁵⁵ W. C. Still, M. Kahn, A. Mitra *J. Org. Chem.* **1978**, *43*, 2923-2925.
- ⁵⁶ J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martínez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo, A. Bernardi *ChemMedChem* **2007**, *2*, 1030-1036.
- ⁵⁷ T. Bravman, V. Bronner, K. Lavie, A. Notcovich, G. A. Papalia, D. G. Myszkla *Anal. Biochem.* **2006**, *358*, 281-288.
- ⁵⁸ L. C. Gjelstrup, J. D. Kaspersen, M. A. Behrens, J. S. Pedersen, S. Thiel, P. Kingshott, C. L. Oliveira, N. M. Thielens, T. Vorup-Jensen *J. Immunol.* **2012**, *188*, 1292-1306.
- ⁵⁹ M. Ortega-Munoz, J. Lopez-Jaramillo, F. Hernandez-Mateo, F. Santoyo-Gonzalez *Adv. Synth. Catal.* **2006**, *348*, 2410-2420.
- ⁶⁰ F. Pertici, N. Varga, A. van Duijn, M. Rey-Carrizo, A. Bernardi, R. J. Pieters *Beilstein J. Org. Chem.* **2013**, *9*, 215-222.
- ⁶¹ Described in the Supporting Information of reference 42.

⁶² J. Xie, L. Hu, W. Shi, X. Deng, Z. Cao, Q. Shen *J. Polym. Sci., Part B: Polym. Phys.* **2008**, *46*, 1140-8.

⁶³ B. A. Laurent, S. M. Grayson *J. Am. Chem. Soc.* **2011**, *133*, 13421-13429.

CHAPTER 2

Rational Design of DC-SIGN Multivalent Antagonists

2.1 DC-SIGN in HIV infection: a paradigm in escaping immunity

One of the most known and studied members of the C-type lectin family surely is DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin). DC-SIGN is a transmembrane protein with a structure characterized by a tetrameric arrangement of four homologous subunits (Fig. 2.1). Each subunit presents a cytosolic N-terminus domain, a transmembrane region, a long flexible neck domain and a C-type lectin-like domain. Importantly, the neck domain is formed by seven complete and one incomplete tandem repeats and is responsible for the oligomerization of the protein into the biologically active homo-tetramer.¹

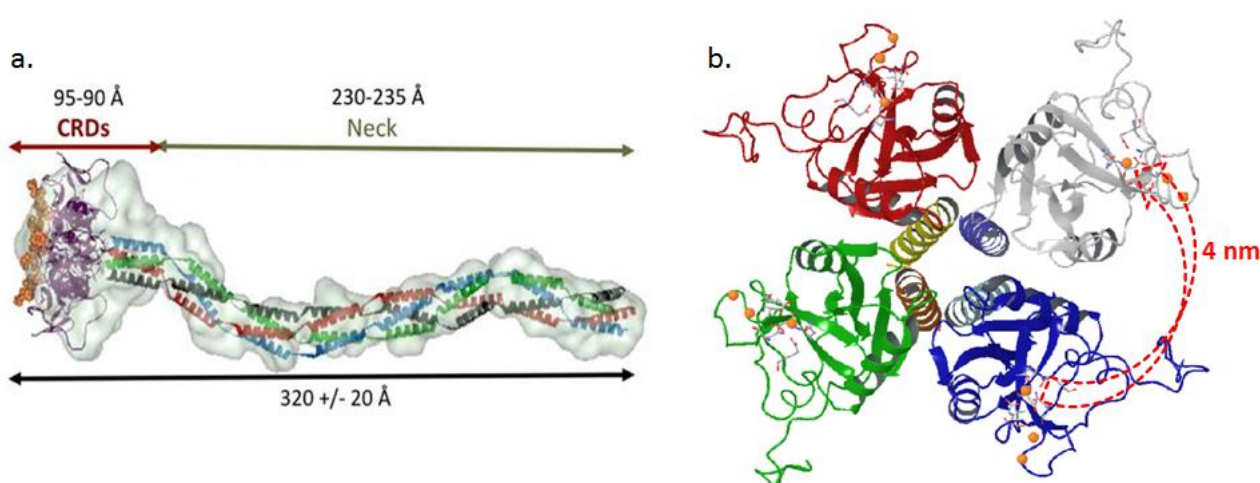


Figure 2.1 – Tetrameric structure of DC-SIGN. **a)** The extracellular domains of DC-SIGN are shown. Four DC-SIGN monomers assemble together to form the active homotetramer. The structure is characterized by the presence of four carbohydrate recognition domains (CRDs) and a long flexible neck formed by seven complete and one incomplete tandem repeats. **b)** The four CRDs of DC-SIGN are directed towards opposite directions in a tetragonal arrangement, with Ca^{2+} ions involved in the binding of glycans spaced around 4 nm between two different CRDs.²

In animals, DC-SIGN is mainly expressed on the surface of dendritic cells (DCs), where it accomplishes two main roles: it acts as a PRR, being able to recognize and bind pathogens, and is also responsible for the adhesion of T lymphocytes (or T cells).^{3,4}

As already described in the Introduction, myeloid DCs are antigen presenting cells (APCs), that are present in an immature state throughout the body, like mucosal and lymphoid tissues, providing important defence mechanism against infections to the organism. Indeed, DCs recognize, and internalize pathogens in order to process and present antigens on major histocompatibility complex (MHC) class I and II molecules to T cells, triggering their activation and differentiation.⁵ More in detail, immature DCs, which show high phagocytic and endocytic capability are responsible for pathogen internalization by several mechanisms including macropinocytosis,⁶ phagocytosis⁷ and adsorptive endocytosis that is mediated by the interaction with PRRs displayed on the myeloid cell surface.⁶ Upon recognition and internalization of antigens, DCs maturation takes place. This process carries out the presentation of processed antigens on their surface and the migration towards secondary lymphoid organs where interactions with T lymphocytes occurs.

Among the PRRs exposed by dendritic cells, a major role is played by the surface C-type lectin DC-SIGN. As a C-type lectin, DC-SIGN is able to bind pathogens selectively recognizing polysaccharide structures

displayed on their surfaces. Moreover, DC-SIGN allows for antigen presentation to T cells, being able to bind to the highly glycosylated intracellular adhesion molecule 3 (ICAM-3) exposed on T lymphocytes.

The affinity and selectivity of DC-SIGN for highly glycosylated structures containing either L-fucose or D-mannose moieties is at the basis of its binding ability towards a series of viruses (HIV, Hepatitis C, Ebola) bacteria (*Mycobacterium tuberculosis*, *Helicobacter pylori*), yeasts (*Candida albicans*), and parasites (*Leishmania* spp, *Schistosoma mansonii*).^{8,9,10} However, pathogens recognized by DC-SIGN may be able to escape from degradation and antigen processing, therefore exploiting the scavenging defensive activity of DC-SIGN as a preferential way to start and spread the infection.^{11,12}

A deeply investigated example of DC-SIGN mediated infection is the HIV-1 (Human Immunodeficiency Virus type 1) infection, which is responsible for AIDS (Acquired Immune Deficiency Syndrome).¹³ The way in which HIV exploits DCs to trigger infection is a controversial topic which is not fully defined. On the basis of experimental evidences, several mechanisms have been proposed; however, multiple mechanisms may be simultaneously operative.¹⁴

What is known is that recognition of the HIV highly mannosylated capsid protein gp120 by DC-SIGN is exploited by the virus to reach the secondary lymphoid organs and to infect CD4⁺ T cells. It has been demonstrated that HIV transmission can occur upon formation of specific complexes between DCs and T cells named infectious-synapses, but this mechanism is still poorly understood.^{15,16,17,18}

Another proposed mechanism arose from the observation that HIV is rapidly internalized by immature DCs into endosomes in a DC-SIGN dependant uptake. HIV containing endosomes are likely targeted to lysosomes for degradation. However, to a certain extent, virus particles have been shown to be released from endosomes in the extracellular environment associated with vesicles known as exosomes. Thus, it has been suggested that infection may originate from exosomes¹⁹ fusion with T cells membrane.²⁰ Moreover, endosomes escape and release of virus particles could represent a mechanism for HIV to reach infectious synapses. Altogether, these mechanisms are referred to as *in trans* infection. On the contrary, several days after the first contact with the virus, T cell infection may be caused by HIV particles originated from *de novo synthesis* in infected DCs. This mechanism is known as *in cis* infection.²¹

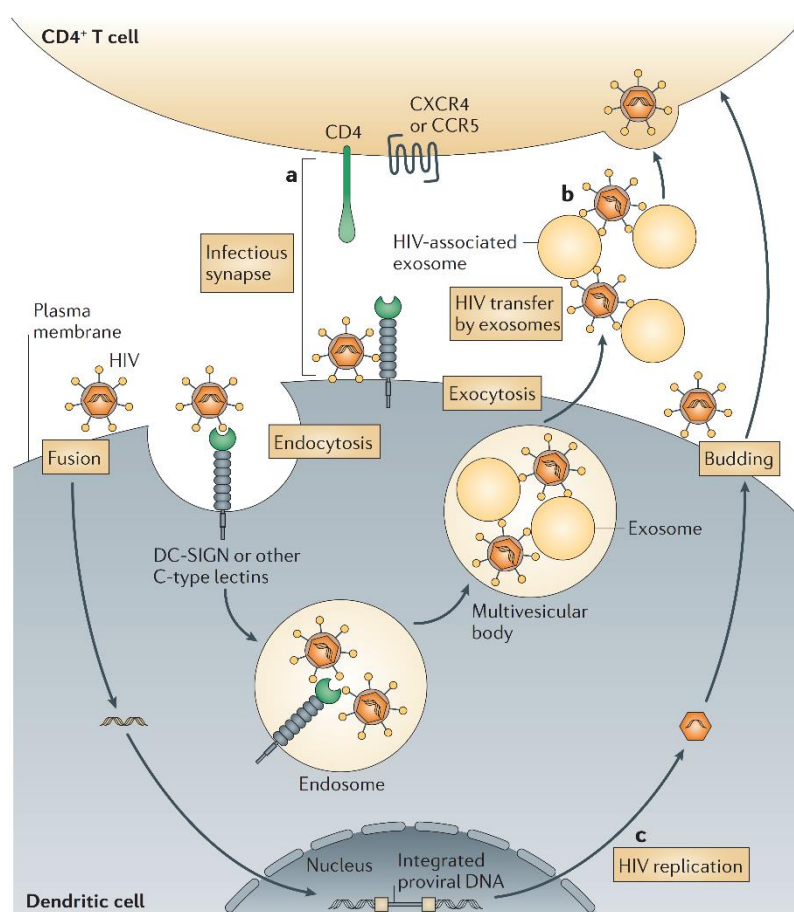


Figure 2.2 - Mechanisms of DC-mediated HIV infection. HIV infection mechanisms can be classified as trans- (pathways a,b) and cis-infection mechanisms (pathway c). **a)** HIV infection may be promoted by the formation of complexes between DCs and T cells named infectious-synapses. **b)** HIV transmission can be mediated by DC-SIGN. DC-SIGN recognizes the highly mannoseylated capsid protein gp120 of HIV, which is internalized into endosomes. However, HIV is able to escape from lysosomal degradation and to infect CD4⁺ T cells. Infection of T cells may arise from HIV release at infectious synapse site or from release of HIV associated to vesicles known as exosomes. Fusion of exosomes with T cells membrane may constitute a mechanism to spread the infection. **c)** Cis-infection is ascribed to HIV virions originated from de novo synthesis in infected DCs.¹⁴

Although a clear picture of the operative mechanisms leading to HIV infection is still missing, the recognition of the virus mediated by DC-SIGN emerges as a key event for the overall process. This fact is furthermore corroborated by the correlation of DC-SIGN expression shutdown with the unproductive formation of infectious synapses and the inhibition of *in trans* infection of CD4⁺ T cells.^{16,22}

2.2 Structure-based synthesis of DC-SIGN ligands

The engagement of DC-SIGN in the first stage of several pathogenic processes offers a convenient way for the development of antiadhesive drugs against many pathogens for which the formulation of vaccine is still a challenge. As a consequence, ligands targeting DC-SIGN constitute the great majority of the C-type lectin antagonists reported so far (see Introduction – Monovalent and Multivalent ligands).

Considering the synthesis of DC-SIGN antagonists as potential HIV drugs, a great challenge is represented by the achievement of selectivity. Indeed, selective targeting of DC-SIGN is required to avoid interfering with other C-type lectins that may supply protecting effects to the host. A clarifying example is given by the protein Langerin. As DC-SIGN, Langerin is a transmembrane C-type lectin showing affinities for mannose and fucose containing glycans, but is characterized by a significantly different structure. Indeed, Langerin quaternary

structure is a homotrimer, in which each subunit, presenting an extracellular region formed by a neck domain and a C-terminus C-type lectin-like domain, assemble together displaying the C-type lectin-like domain in a rigid trigonal fashion spaced around 41.5 Å.²³

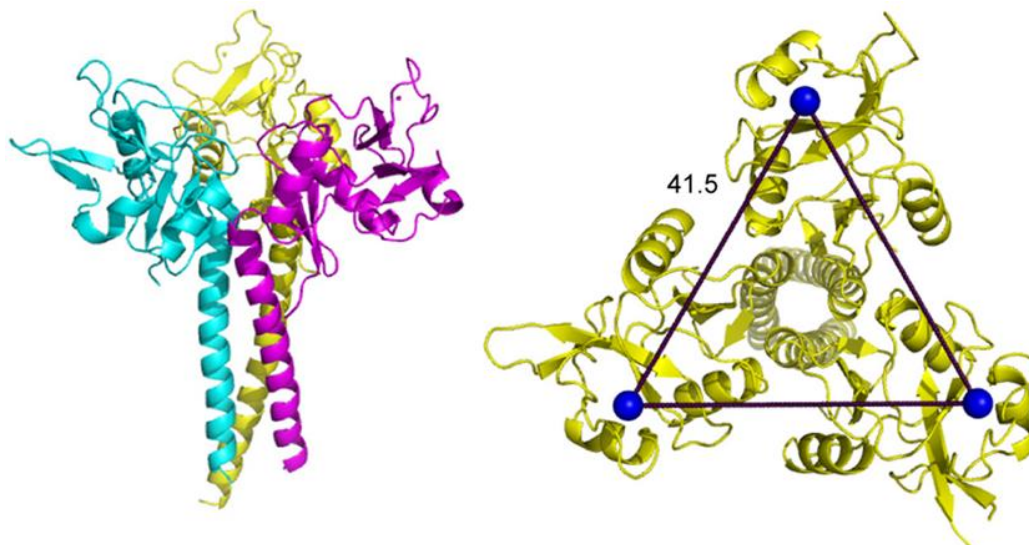


Figure 2.3 – Trimeric structure of Langerin. Extracellular domains of truncated Langerin are shown. The protein is characterized by three subunits that assemble together forming a homotrimer. The three CRDs are arranged in a trigonal disposition spaced about 4 nm from each other.²³

Langerin expression is limited to Langerhans cells, a sub-type of immature dendritic cells located in the epidermis layer and in mucosal stratified epithelia, thus being the first antigen presenting cells encountered by pathogens as HIV.²⁴ As opposed to DC-SIGN, an important protective role is furnished by Langerin which is able to sequester HIV virions into Langerhans cells allowing their degradation.²⁵

In the past few years our group have demonstrated that structure-based design of DC-SIGN antagonists, which combine glycomimetic ligands and optimal multivalent presentation, constitute a successful approach to reach potency and selectivity. In particular, glycomimetics have been loaded onto rigid rod-like spacers giving access to constructs with tuneable length, which were used to bridge two DC-SIGN C-type lectin-like domains (Fig. 2.4).^{26,27}

In the design of chelating agents for DC-SIGN, the length of the compounds is a fundamental parameter since it has to cover at least the distance between two calcium ions at adjacent carbohydrate recognition domains, estimated to be around 4 nm for the tetramer.² Moreover, a good compromise between rigidity and flexibility is required to simultaneously decrease the entropy of the system, thus favouring the binding, while still allowing conformational adaptation of the ligands to the protein.

As monovalent ligands, the pseudo-disaccharide **5** was selected along with the corresponding more potent bis-*p*-hydroxymethylbenzylamide derivative **7** which demonstrated good selectivity against DC-SIGN over Langerin.²⁸ Both monovalent ligands and the corresponding trivalent glycodendrons were used to functionalize the rigid spacers obtaining bidentate and hexavalent compounds with different length and flexibility. The molecular rods were obtained by connection of PEG-containing phenylene-ethynylene units, which guaranteed water solubility of the constructs.²⁹

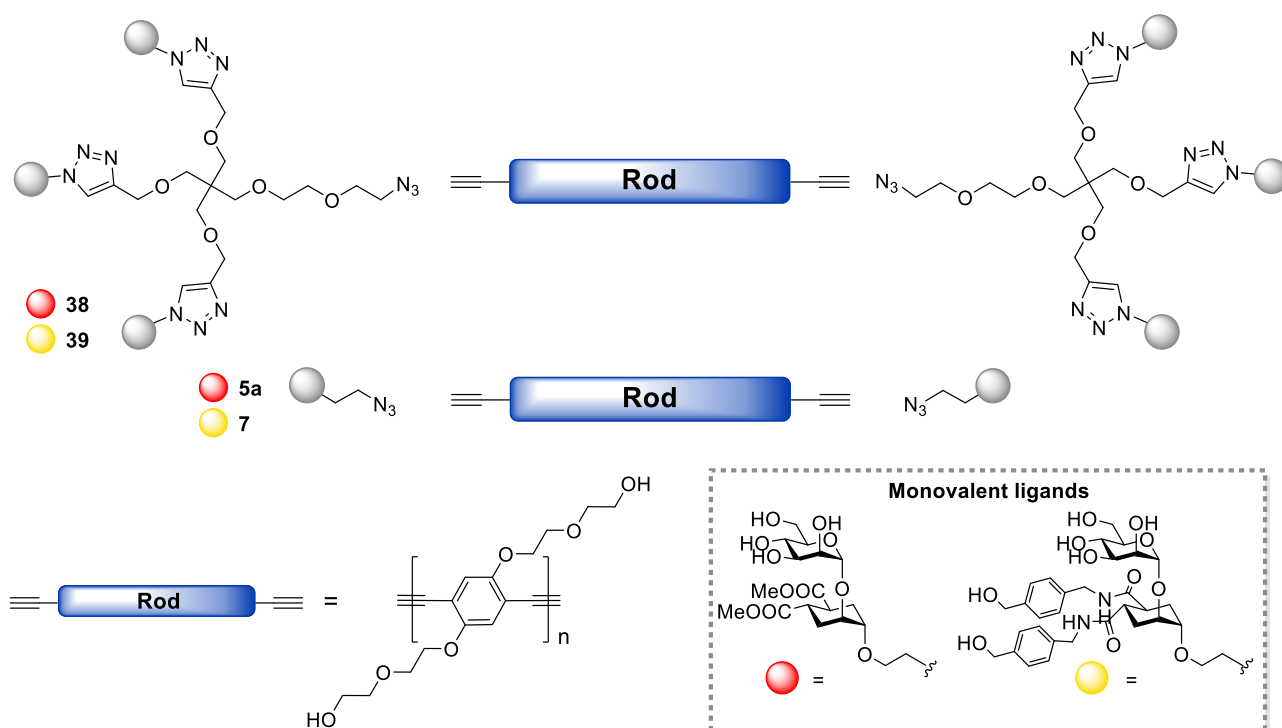


Figure 2.4 – DC-SIGN chelating antagonists. The syntheses of DC-SIGN ligands able to bridge two CRDs have been reported. Hexavalent and divalent ligands have been developed connecting either trivalent glycodendrons **38** and **39** or monovalent glycomimetics **5a** and **7** to a rod like scaffold with variable length. The rod core, constituted by pegylated phenylene-ethynylene units, bestows rigidity and water solubility to the construct.²⁶

Compounds were tested for their binding affinity towards DC-SIGN in a HIV-1 *trans* infection experiment. B-THP-1 cells expressing DC-SIGN were used as a model of DCs and the antagonists potency was evaluated by their ability in inhibiting CD4⁺ T cells infection.

As expected, compounds bearing the more active bis-amide ligand **7** performed as the best. Among these, surprisingly high potency was shown by the divalent **82** ($IC_{50} = 161$ nM), in which just two copies of monovalent **7** were directly connected to a spacer constituted by three phenylene-ethynylene units (Fig. 2.5). Stochastic Dynamics (SD) simulation experiments revealed that the bidentate **82** possesses an optimized pre-organization of the ligands that results in a strong chelation of DC-SIGN binding sites. Indeed, the rigidity imposed by the spacer, limits molecule conformations to extended geometries, with 30% of the sampled structures showing a length higher than 35 Å, thus being able to perform as chelating agents. The length of the structures was calculated as the distance between the two O3 atoms (d_{O3-O3}) of the two mannose residues displayed by the linear **82**, which are known to interact with the Ca²⁺ ions at the carbohydrate binding sites.

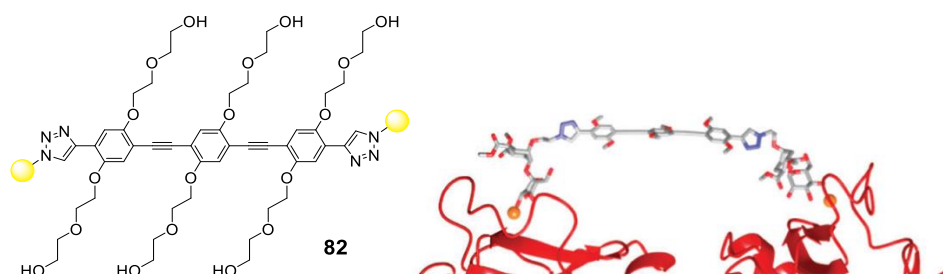


Figure 2.5 – Chelating divalent **82.** A DC-SIGN potent antagonist ($IC_{50} = 161$ nM) was obtained functionalizing a linear scaffold constituted by three phenylene-ethynylene units with two copies of monovalent ligand **7**. Stochastic Dynamics (SD) simulation experiments revealed that construct **82** is characterized by optimal presentation of the glycomimetic moieties, which determines a strong chelating effect.

In a predictable way, increasing the number of monovalent ligands afforded even more potent antagonists with the hexavalent glycodendrimer **43** performing as the most potent ($IC_{50} = 24$ nM) (Fig. 2.6). This compound is characterized by a rod at its core, with the same length of the linear **82**, which was connected to two copies of glycodendrimer **39**. The introduction of two long linkers protruding from the aromatic core bestows a higher flexibility to the molecule. As a consequence, although the maximum length for hexavalent **43** was calculated as 67.5 Å, SD simulations revealed that this construct preferentially lies in folded and compact conformations, with less than 5% of sampled structures having d_{O3-O3} larger than 35 Å. However, the statistical rebinding contribution arising from multiple presentation of glycomimetic moieties represents a trade-off for the unoptimized preorganization of the ligand and, along with chelation, ensures high avidity to the system.

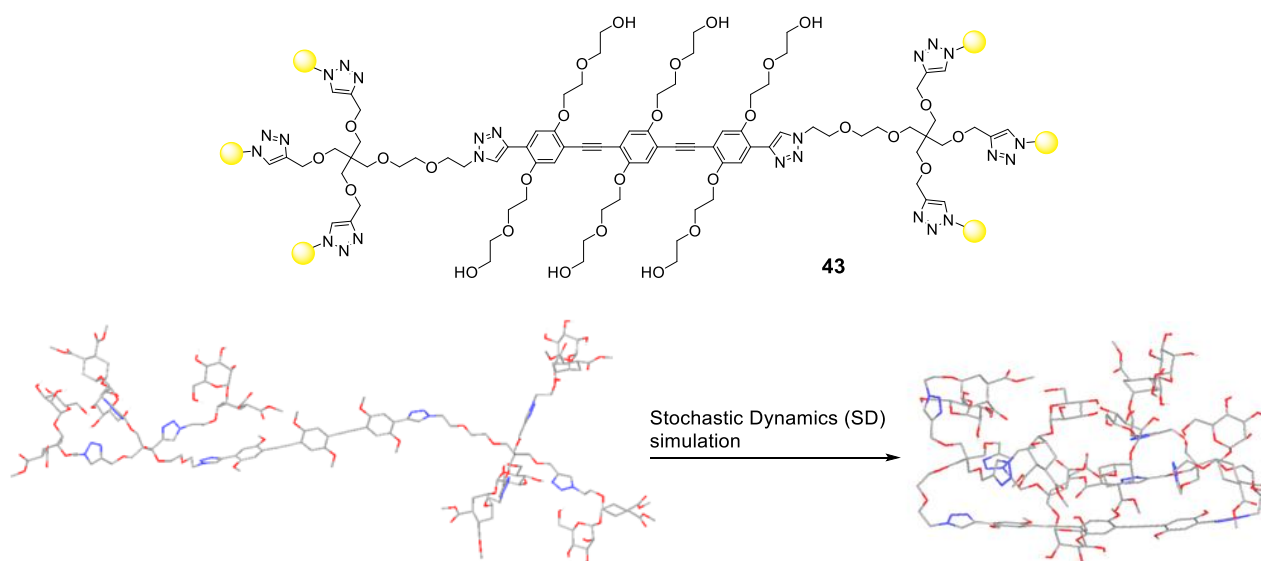


Figure 2.6 – Chelating hexavalent **43.** Glycodendrimer **43** was originated connecting of two copies of glycodendron **38** with a rod like scaffold of three phenylene-ethynylene units. SD simulations have revealed that the flexibility of the construct hamper a proper preorganization of the glycomimetic moieties, with **43** assuming compact conformations. However, statistical rebinding effects and the retained chelating ability allow for strong binding potency ($IC_{50} = 24$ nM).

2.3 Synthesis of tetracoordinating ligands: grabbing DC-SIGN

So far, our group has reported the synthesis of DC-SIGN antagonists able to bridge two carbohydrate binding sites of the protein.²⁶ Since chelation ability proved a strong impact on ligand potency, we reasoned that even stronger antagonists could be obtained by simultaneous binding of the four CRDs of the DC-SIGN tetramer.

We planned to achieve this goal by synthesizing a cross-shaped glycodendrimer **83**, characterized by a tetravalent rigid core functionalized with four copies of trivalent glycodendron **38** (Fig. 2.7). The core is reminiscent of to the rod-like scaffold previously developed for the chelating compounds **82** and **43**. Crucial aspects are the rigidity, which allows preorganization of the ligands directed towards the targeted CRDs, and the presence of PEG (polyethylene glycol) chains, which guarantee the solubility of the construct. As DC-SIGN ligand, we chose to connect the scaffold with the trivalent glycodendron **38**, which displays three copies of the pseudo-dimannoside glycomimetic **5a**. This functionalization allows to extend the core structure, whose diagonal length was calculated to be 22 Å, in order to reach all the four CRDs of DC-SIGN, which are spaced around 55 Å at opposite sides of the tetramer.

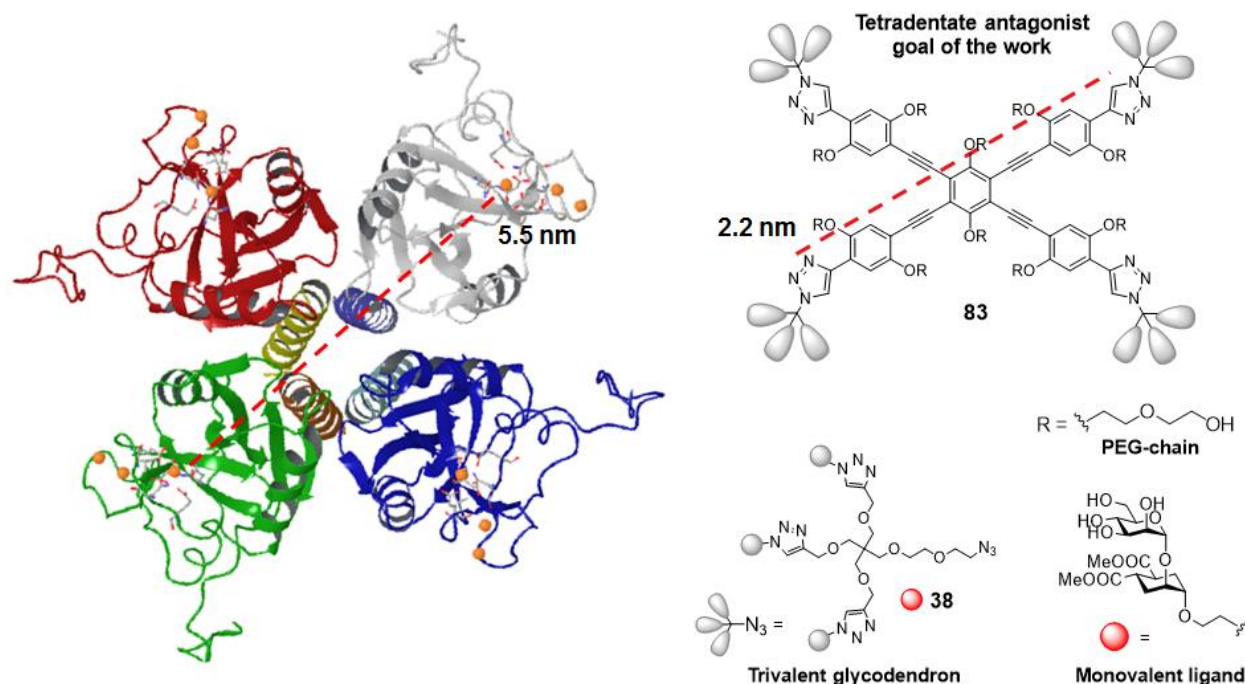
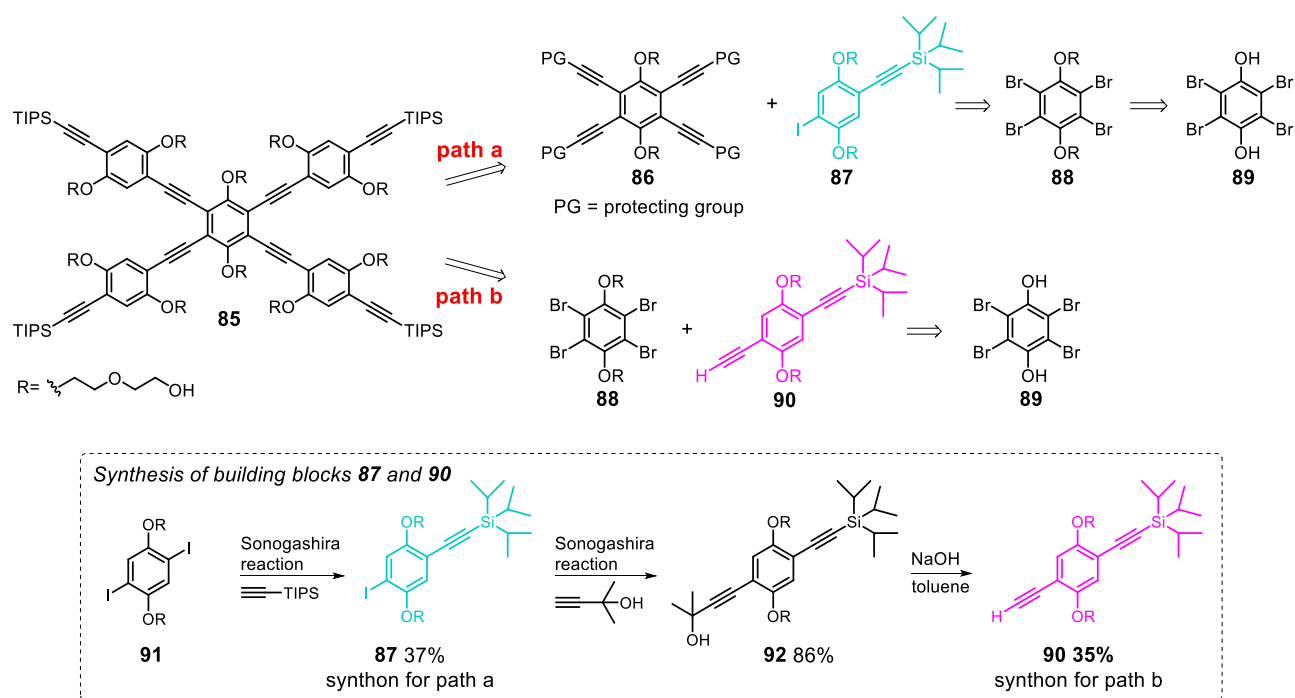


Figure 2.7 – Cross-shaped glycodendrimer 83. In order to allow simultaneous binding of the four CRDs displayed by DC-SIGN, the synthesis of cross shaped **83** was targeted. The functionalization of a tetraivalent core with trivalent dendron **38** is aimed to extend the scaffold length in order to cover the distance between CRDs located at opposite sides of DC-SIGN tetramer (around 5.5 nm).

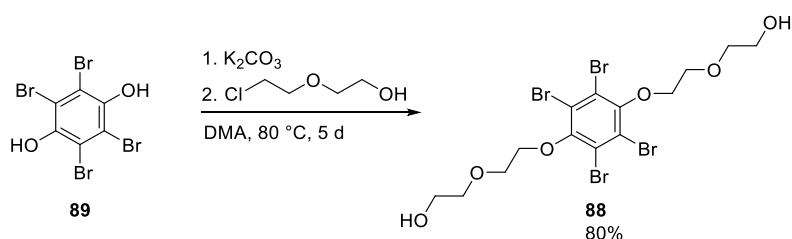
Following a similar strategy for the synthesis of the rod-family compounds, we planned to exploit CuAAC reaction to connect glycodendrone **38** to the rigid tetraivalent core. For this purpose, we identified the TIPS (triisopropyl silyl) protected **85** as a key intermediate (Scheme 2.1). The proposed retrosynthetic analysis shows that **85** can be originated from functionalization of the tetraalkynyl aryl derivative **86** with four copies of the iodide phenylene ethynylene unit **87**. On the other hand, the core unit **86** can arise from Sonogashira reaction of pegylated tetrabromide **88** obtained from alkylation of the corresponding tetrabromo hydroquinone **89** (Scheme 2.1, path a).

A more straightforward pathway envisages the synthesis of key intermediate **85** directly by Sonogashira coupling between the tetrabromo propargyl **88** and the monoprotected bis-alkyne **90** (Fig. 2.1, path b). However, this route was avoided due to the laborious and non-efficient synthesis required for the monoprotected synthon **90**, whose preparation has been reported for the synthesis of rod-like linear scaffolds.³⁰ This building block shares the initial synthetic step with the synthon **87** envisaged in path a. First the bis-iodide derivative **91** is reacted in a Sonogashira coupling to give the mono-functionalized phenylene-ethynylene unit **87** (37%) along with the corresponding bis-alkyne derivative (39%). Since both the products were required in the preparation of rod-like scaffolds, an optimization of the reaction to give **87** as unique product has never been attempted. A further Sonogashira coupling of **87** is then required to install an orthogonally protected alkyne moiety which enables selective deprotection of **92** under basic condition affording synthon **90** in 35% yield.



Scheme 2.1 – Retrosynthetic analysis for scaffold 85. Two pathways have been identified for the synthesis of the core **85**. Path a envisages a Sonogashira coupling between iodide synthon **87** and the tetraalkyne derivative **86**, which can be originated from the tetrabromide **88**, obtained from alkylation reaction of tetrabromohydroquinone **89**. A more straightforward strategy is offered by path b where core **85** is originated from Sonogashira coupling of tetrabromide **88** and the terminal alkyne derivative **90**. However, this route was avoided due to the non-efficient synthesis of **90** (shown in the scheme).

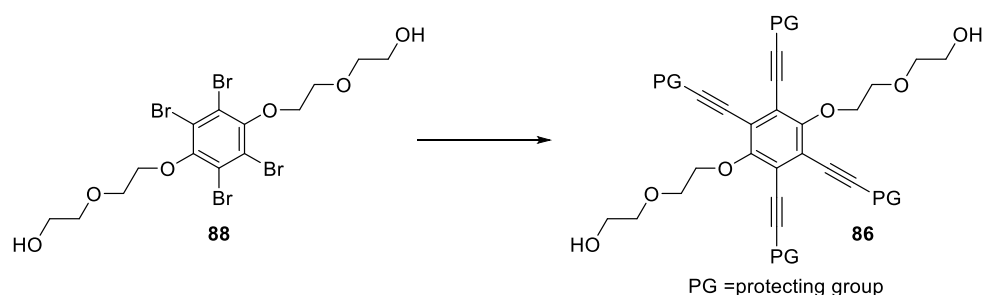
The alkylation of tetrabromo hydroquinone **89** with 2-(2-chloroethoxy)ethanol under basic conditions afforded the pegylated **88**, that constitutes the central unit of the tetravalent rigid core (Scheme 2.2). Almost complete conversion was observed heating for 5 days at 80 °C and monitoring by TLC. Flash chromatography afforded pure product **88** in 80% yield.



Scheme 2.2 - Alkylation reaction of 89. Tetrabromohydroquinone **89** was reacted with K_2CO_3 and 2-(2-chloroethoxy)ethanol affording the pegylated **88** (80%).

In the attempt to synthesize the protected tetraalkynyl **86**, tetrabromo **88** was subjected to Sonogashira coupling. A similar coupling reaction has been reported for 1,4-alkoxy derivatives of tetrabromo hydroquinone using phenylacetylene as the terminal alkyne.³¹ However, in our case the use of variously protected acetylene derivatives was unsuccessful. Reactions were generally performed using $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ as catalyst in combination with CuI in the presence of a base. Low conversion was achieved with TIPS- acetylene performing the reaction at 80 °C, with formation of the mono and disubstituted intermediates (detectable by ESI-MS analysis) (Table 2.1, entry 1). To favour the reactivity, trimethylsilyl- acetylene (TMS- acetylene) was chosen as a less bulky alkyne. Unfortunately, no conversion was observed heating the reaction mixture at 70 °C or irradiating with the microwave at 50 °C (Table 2.1, entry 2 and 3). Increasing the temperature at 100 °C under microwave irradiation led to side-products with higher molecular weight (Table 2.1, entry 4). The use

of 2-methyl-3-butyn-2-ol as an alternative protected acetylene equivalent was also investigated. Formation of up to a trisubstituted intermediate was observed heating in dioxane at 90 °C (Table 2.1, entry 6), while, a tetrasubstituted product was obtained with almost complete conversion irradiating with microwaves at 100 °C in 2 hours (detected by ESI-MS 637.6 [M+Na]⁺) (Table 2.1, entry 7). Surprisingly, a different isomer of the expected **86** was isolated after flash chromatography as revealed by ¹H NMR analysis (Fig. 2.8). Indeed, eight signals for the PEG chains (two being overlapped) were identified, clearly showing the disruption of the molecular symmetry, which is in contrast with the formation of **86**. This constitutional isomer may perhaps arise from undesired Pd catalysed isomerization favoured under the reaction conditions and its full characterization was not pursued.



Entry	Alkyne	Conditions	Result
1	$\equiv\text{Si/Pr}_3$ 12 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ , CuI, Ph ₃ P, Et ₃ N, 80 °C, 7 h	Low conversion in mono- and di- functionalised
2	$\equiv\text{SiMe}_3$ 4.4 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ 3 mol%, CuI 6 mol%, Et ₃ N, DMF, 70 °C, 20 h	No conversion
3	$\equiv\text{SiMe}_3$ 10 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ 20 mol%, CuI 20 mol%, Et ₃ N, DMF, 100 °C, MW, 2 h	High molecular weight side-products
4	$\equiv\text{SiMe}_3$ 4.4 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ 20 mol%, CuI 20 mol%, Et ₃ N, DMF, 50 °C, MW, 2 h	No conversion
5	$\equiv\text{C}(\text{OH})$ 5 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ 10 mol%, CuI 10 mol%, DIPEA, THF, 70 °C, 2 d	No conversion
6	$\equiv\text{C}(\text{OH})$ 5 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ 10 mol%, CuI 20 mol%, DIPEA, 1,4-dioxane, 90 °C, 2.5 d	Mono-, di- and tri-functionalized
7	$\equiv\text{C}(\text{OH})$ 10 mol. eq.	(Ph ₃ P) ₂ PdCl ₂ 20 mol%, CuI 20 mol%, DIPEA, DMF, 100 °C, MW, 2 h	Almost complete but rearrangement to unidentified constitutional isomer

Table 2.1- Alkynylation of 88. The table summarizes the reaction conditions attempted for the Sonogashira coupling of tetrabromide **88** with different sources of protected terminal acetylenes.

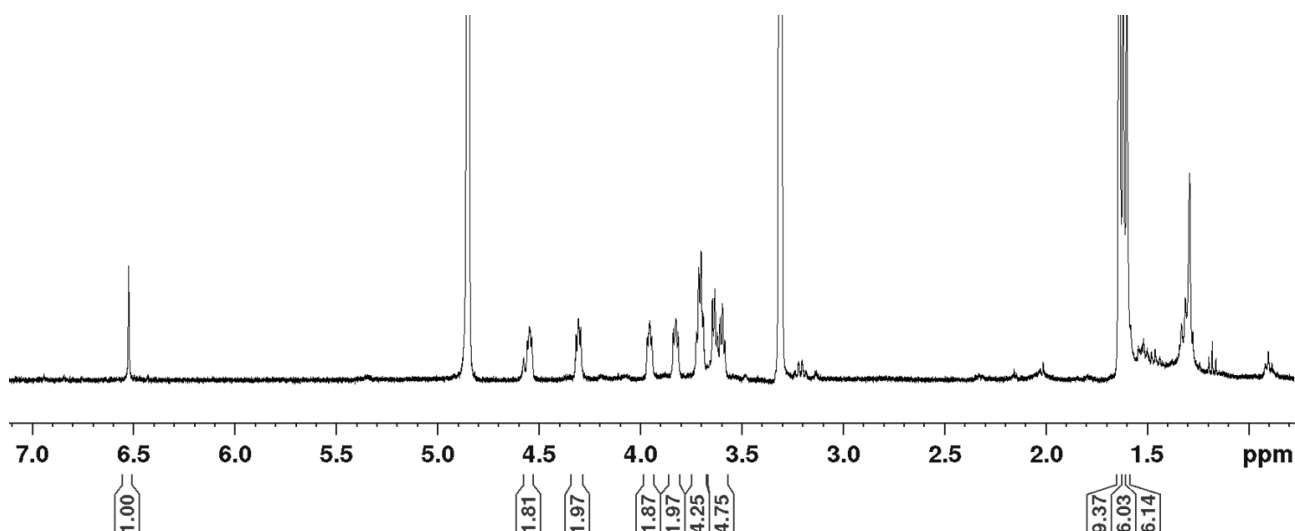
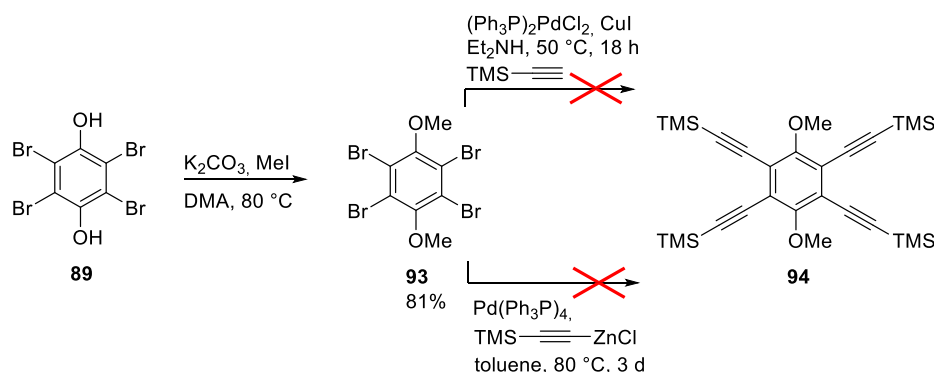


Figure 2.8 - ^1H NMR spectrum of **86** constitutional isomer in CD_3OD (400 MHz). The formation of an isomeric side-product of tetraalkynyl **86** under reaction conditions of (Table 2.1, entry 7) was clearly revealed by ^1H NMR analysis. The presence of a deshielded signal (6.53 ppm), eight different signals for the PEG chains and four distinct signals for the methyl groups can be identified. These data reveal the low degree of symmetry of the molecule, which allows to exclude the formation of the desired highly symmetric **86**. ^1H NMR (400 MHz) δ (ppm): 6.53 (s, 1 H), 4.59-4.53 (m, 2 H, PEG), 4.33-4.28 (m, 2 H, PEG), 3.98-3.93 (m, 2 H, PEG), 3.85-3.80 (m, 2 H, PEG), 3.73-3.68 (m, 4 H, PEG), 3.65-3.57 (m, 4 H, PEG), 1.64 (s, 12 H, CH₃), 1.62 (s, 6 H, CH₃), 1.60 (s, 6 H, CH₃).

To establish whether the poor reactivity of tetrabromo **88** was determined by the presence of free hydroxyl groups at the end of the PEG chains, the corresponding dimethoxy derivative **93** was synthesized by methylation of tetrabromo hydroquinone **89** and subjected to Sonogashira reaction with trimethylsilylacetylene (Scheme 2.3). The reaction was performed using $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ as catalyst along with CuI , Et_2NH at 50°C for 18 h but no formation of the product was detected by ESI-MS. No conversion was also assessed by analogous reaction of the tetrabromo hydroquinone **89**.

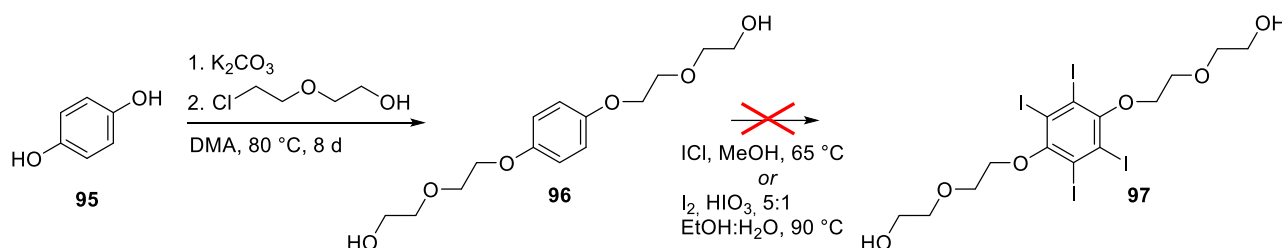
To push the reactivity, an alkynylzinc compound was exploited as a more reactive agent in the transmetalation step envisaged by the catalytic cycle (Scheme 2.3). In particular, a THF solution of trimethylsilylacetylene was treated with $n\text{BuLi}$, followed by the addition of ZnCl_2 to afford a trimethylsilylacetylene zinc chloride complex. This reagent was used in combination with $\text{Pd}(\text{PPh}_3)_4$ at 80°C to convert the tetrabromide **93**. This protocol has been reported as a valuable alternative for substrates that presents poor reactivity under Sonogashira reaction and has been exploited for the synthesis of encumbered hexaalkynyl benzene derivatives.³² Yet, with this methodology we observed no conversion of tetrabromide **93**.



Scheme 2.3 – Alkylation reaction of **93**. Dimethoxy derivative **93** was prepared by methylation of tetrabromohydroquinone **89**. No conversion of **89** was observed either after Sonogashira reaction with trimethylsilylacetylene or using the corresponding alkynylzinc reagent.

Since we did not want to abandon the synthetic strategy adopted, we reasoned that an increased reactivity for the Sonogashira coupling could be displayed by the tetraiodide derivative **97**. Indeed, aryl iodides are known as more reactive species compared to aryl bromides towards Pd oxidative addition. The synthesis of tetraiodo **97** requires as an intermediate the pegylated **96**, which was obtained from alkylation reaction of hydroquinone **95** as a pure compound with no need of chromatographic purification. (Scheme 2.4).

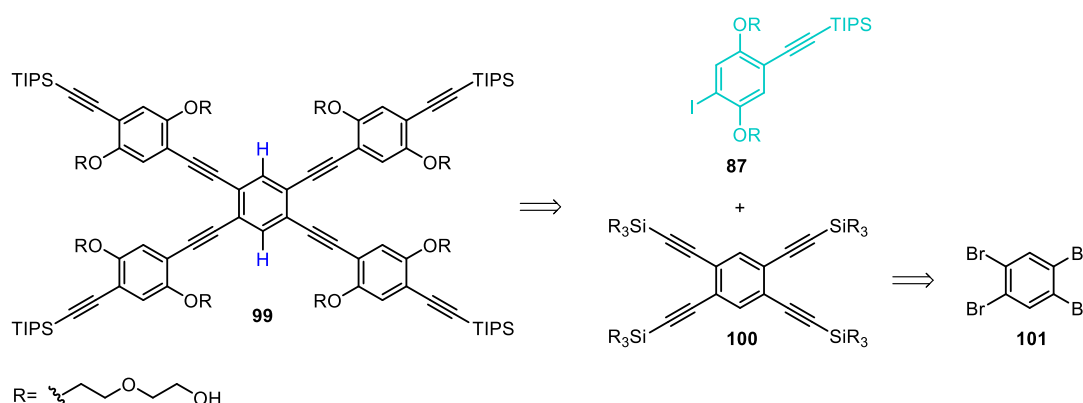
Full iodination of the aromatic ring in **96** turned out to be really challenging. In a first attempt ICl was used as iodinating agent in refluxing methanol.³³ The reaction was monitored by ESI-MS revealing the formation of mono, di and tri iodinated species along with the chlorinated side-products. As an alternative, I₂ was used in combination with HIO₃ at 90 °C, however only the product of disubstitution was observed monitoring by ESI-MS.³⁴ Analogous results were obtained for iodination reaction of the hydroquinone **95**.



Scheme 2.4 – Iodination reaction of **96.** Pegylated **96** was obtained by alkylation reaction of hydroquinone **95**. Iodination of **96** to give tetraiodide **97** failed using either ICl or I₂ in combination with HIO₃.

The failures in the synthesis of the pegylated tetraalkynyl central unit **86** brought us to reconsider our synthetic plan. Therefore, we selected the cross-shaped scaffold **99** as an alternative key intermediate. This core represents a good compromise between synthetic accessibility and similarity with our original goal (Scheme 2.5). The difference with respect to the previous targeted core **85** is given by the central unit which now does not contain the two original PEG chains. Even though this modification may determine a decrease in water solubility and a higher tendency to give pi-stacking interactions, the high content of PEG chains into the core structure may be sufficient to tolerate this variation.

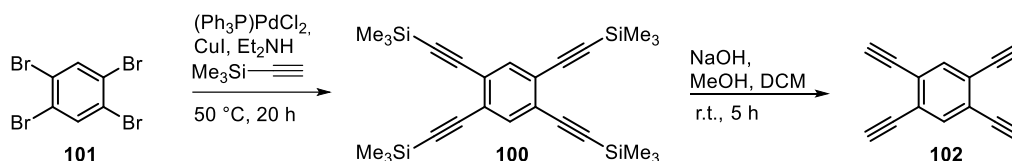
A retrosynthetic analysis for the core **99** identifies the iodide phenylene ethynylene unit **87** and the protected tetraalkynylbenzene **100** as the two key synthons (Scheme 2.5). The synthesis of **100** has already been reported starting from 1,2,4,5-tetrabromobenzene **101**.³⁵



Scheme 2.5 – Retrosynthetic analysis for core **99.** Following the proposed retrosynthesis, the tetravalent core **99** can be obtained from Sonogashira coupling between iodide **87** and the tetraalkynylbenzene **100**. Intermediate **100** can be achieved using tetrabromobenzene **101** as precursor.

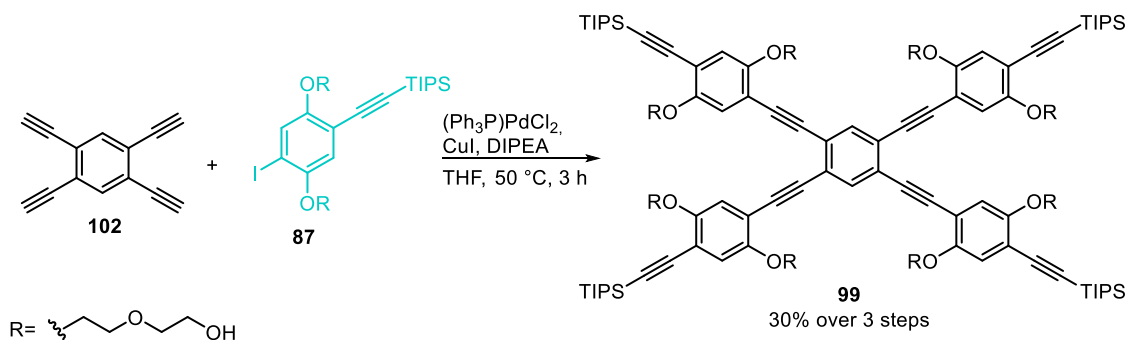
As the first step of the synthesis, 1,2,4,5-tetrabromobenzene **101** was subjected to Sonogashira coupling with trimethylsilylacetylene (Scheme 2.6). The use of $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ and CuI in Et_2NH at 50 °C afforded the desired protected tetraalkynyl **100** as a pure product in 20 h. Remarkably, no conversion was obtained in the previously attempted Sonogashira reaction of pegylated tetrabromo **88** using the same reaction conditions. The great difference in terms of reactivity between substrates **101** and **88** suggests that the presence of alkoxy substituents on the aromatic ring plays detrimental stereoelectronic effects.

The tetravalent **100** was used as a crude in the following deprotection reaction. Removal of the trimethylsilyl groups under basic conditions afforded the tetraalkynyl central unit **102** as a pure product, with no need of further chromatographic purification. Notably, the selective formation of both **100** and **102** was confirmed by ^1H NMR and by electron impact (EI) MS.



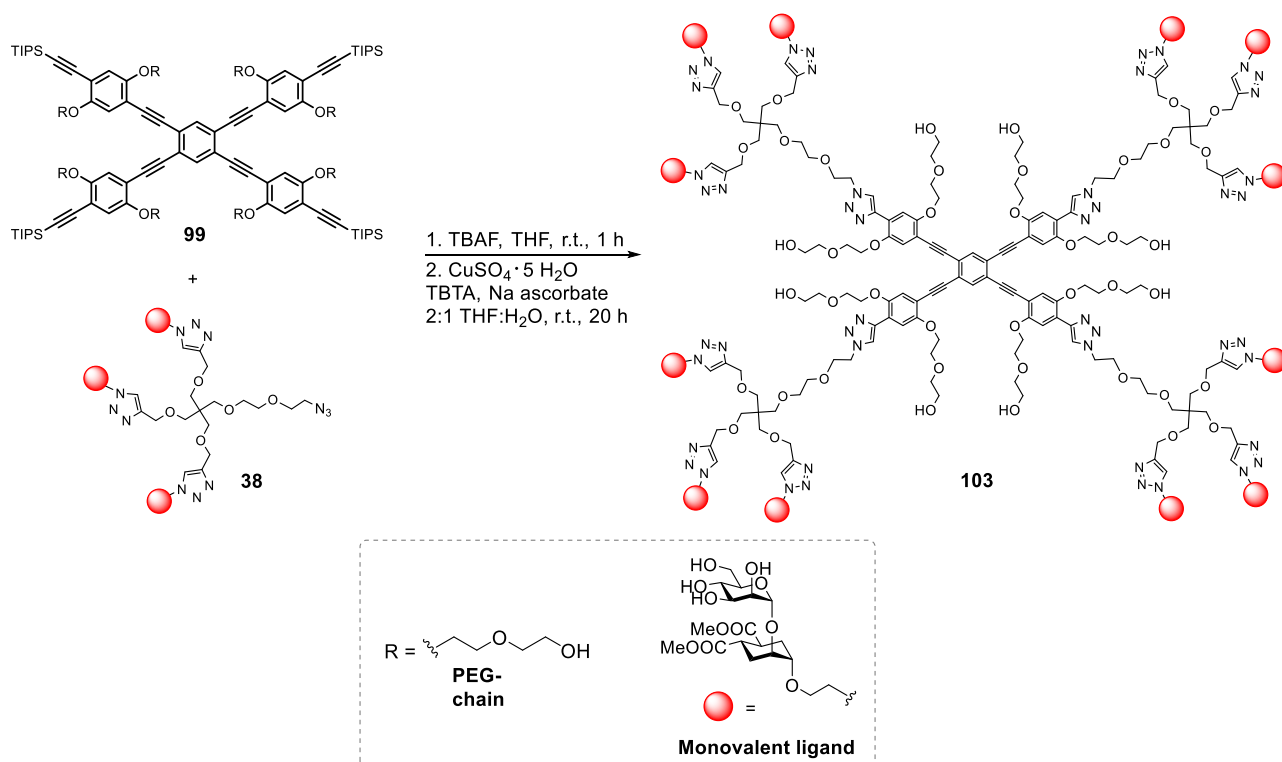
Scheme 2.6 – Synthesis of 1,2,4,5-tetraethynylbenzene 102. Tetrabromobenzene **101** was converted into tetraalkynylbenzene **100** by Sonogashira reaction with trimethylsilylacetylene. Removal of trimethylsilyl protecting groups was achieved treating with NaOH in a 1:1.3 MeOH/DCM mixture. No purification was needed for these synthetic steps.

Finally, the central unit **102** was connected to four copies of iodide synthon **87** by Sonogashira coupling. The temperature is a critical parameter since no conversion was observed performing the reaction at room temperature. However, the desired scaffold **99** was selectively obtained in 3 h heating at 50 °C. The intrinsic fluorescence of the product allows its detection by TLC analysis irradiating at 365 nm. The reaction was also monitored by ESI-MS. After work-up, flash chromatography followed by size-exclusion chromatography (Sephadex LH-20 column) afforded the pure core **99** as an oil in 30% yield over three steps.



Scheme 2.7 – Synthesis of tetravalent core 99. Sonogashira coupling of 1,2,4,5-tetraethynylbenzene **102** with iodide synthon **87** at 50 °C for 3 h afforded the cross-shaped phenylene-ethynylene scaffold **99** in 30% yield (over 3 steps from **101**).

With the phenylene-ethynylene core **99** in hand, the glycodendrimer **103** was finally accessible (Scheme 2.8). Removal of triisopropylsilyl protecting groups was achieved in situ treating with a TBAF solution in THF for 1 h. The reaction was followed by TLC monitoring at 365 nm until full conversion was observed. Then the scaffold was functionalized by CuAAC with four copies of the trivalent glycodendron **38**. The formation of glycodendrimer **103** was assessed by MALDI-TOF MS (DHB matrix) and its isolation was achieved by size-exclusion chromatography (Sephadex LH-20) in very good yield (92%).



Scheme 2.8 – Synthesis of cross-shaped glycodendrimer **103**. The final glycodendrimer **103** was readily obtained by in situ desilylation of the trivalent scaffold **99** and subsequent CuAAC reaction with the azide tagged trivalent pseudo-mannosylated dendron **38**.

2.4 Conclusions

Structure-based design is a successful strategy for the synthesis of effective DC-SIGN antagonists. Potency and selectivity can be jointly achieved by the combination of glycomimetic ligands and multivalent presentation. In particular, our group has recently demonstrated that chelating glycodendrimers able to bridge two DC-SIGN CRDs perform as strong antagonists of HIV *trans* infection.^{26,27} Hexavalent and divalent glycodendrimers of various size were prepared by connecting either trivalent glycodendrons **38**, **39** or monovalent glycomimetic ligands **5a**, **7** to a rigid linear scaffold. Rod-like scaffolds of variable length were obtained from connection of pegylated phenylene-ethynylene units in order to guarantee the water solubility of the constructs. This study highlighted the importance of ligand preorganization and the delicate balance between rigidity and flexibility that have always to be considered in the designing of chelating ligands. Fine-tuned multivalent scaffolds achieve potency with low loading of monovalent glycomimetic ligands (divalent **82** $\text{IC}_{50} = 161 \text{ nM}$), while increasing ligand valency is a valuable strategy for the synthesis of effective ligands with unoptimized scaffolds (hexavalent **43** $\text{IC}_{50} = 24 \text{ nM}$).

Following a similar approach, a cross-shaped tetravalent rigid scaffold was synthesized starting from 1,2,4,5-tetrabromobenzene **101**. Thus, ligands able to simultaneously bind all four CRDs of DC-SIGN were targeted. As a preliminary result the synthesis of the dodecavalent glycodendrimer **103** was achieved. This construct is characterized by an aromatic central unit which was extended with pegylated phenylene-ethynylene moieties. The resulting core was functionalized with four copies of trivalent glycodendrion **38** which increases the length of the construct allowing to cover the distance between two CRDs located at opposite sides of the DC-SIGN tetrameric structure (5.5 nm).

The feasibility of the synthesis that we designed gives now access to a set of cross-shaped ligands. In particular, the functionalization of the core with trivalent glycodendrion **39** may be attempted. This construct is an analogous version of glycodendrion **38** which displays higher affinity for DC-SIGN but also reduced water

solubility. On the other hand, the use of monovalent glycomimetic ligands **5a** and **7** may also be exploited (Fig. 2.9). The reduced length of monovalent ligands, compared to trivalent glycodendrons, demands for the synthesis of antagonists with further extended central core. However, these new constructs may allow to reach potency and selectivity with minimal loading of glycomimetic ligands.

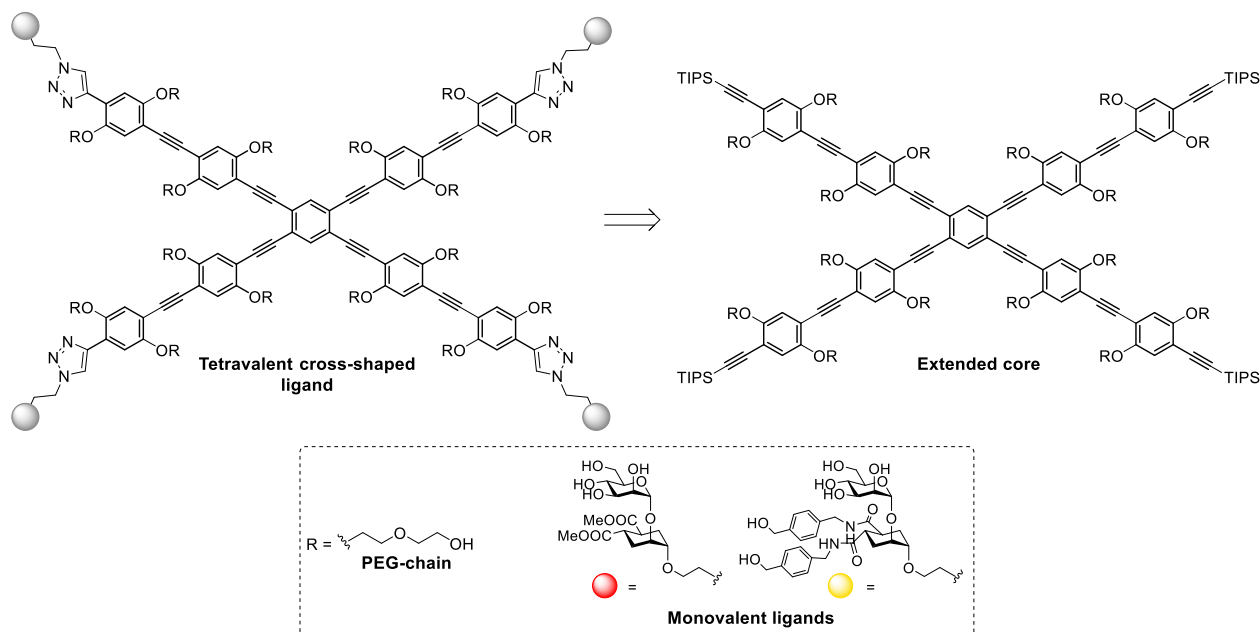


Figure 2.9 – Tetraivalent crossed-shaped ligands. As a future perspective the synthesis of tetraivalent cross-shaped ligands bearing monovalent ligands may be addressed. This goal requires the preparation of antagonists characterized by further extended cross-shaped central core.

2.5 Experimental

2.5.1 General methods and procedures

Chemicals were purchased from commercial sources and used without further purification, unless otherwise indicated. When anhydrous conditions were required, the reactions were performed under nitrogen atmosphere. Anhydrous solvents were purchased from Sigma-Aldrich® with a content of water $\leq 0.005\%$. *N,N'*-Diisopropylethylamine (DIPEA) and triethylamine were dried over calcium hydride, THF and 1,4-dioxane were dried over sodium/benzophenone and freshly distilled before use. Reactions were monitored by analytical thin-layer chromatography (TLC) performed on Silica Gel 60 F₂₅₄ plates (Merck) with UV detection (254 nm and 365 nm) and/or staining with ammonium molybdate acid solution or potassium permanganate alkaline solution. Flash column chromatography was performed according to the method of Still and co-workers³⁶ using silica gel 60 (40–63 μm) (Merck). Automated flash chromatography was performed with Biotage Isolera Prime system, Biotage SNAP KP-Sil cartridges were employed unless otherwise indicated. Size-exclusion chromatography was performed using Sephadex LH-20 from GE Healthcare Life Science. Microwave irradiation was performed by a Biotage Initiator⁺ system. HPLC analyses were performed with an Atlantis T3 5 μm 4.6x100 mm column (Waters) equipped with a Waters 996 Photodiode Array Detector. NMR experiments were recorded either on a Bruker AVANCE-600 MHz, Bruker AVANCE-400 MHz or a Bruker AVANCE-300 MHz instrument at 298 K. Chemical shifts (δ) are reported in ppm. The ¹H and ¹³C NMR resonances of compounds were assigned with the assistance of COSY and HSQC experiments. Multiplicities are assigned as s (singlet), d (doublet), t (triplet), q (quartet), quint (quintet), m (multiplet). EI-MS spectra were collected using a VG AUTOSPEC-M246 spectrometer (double-focusing magnetic sector instrument with EBE geometry) equipped with EI source. Solid samples were introduced via a heated direct insertion probe. ESI-

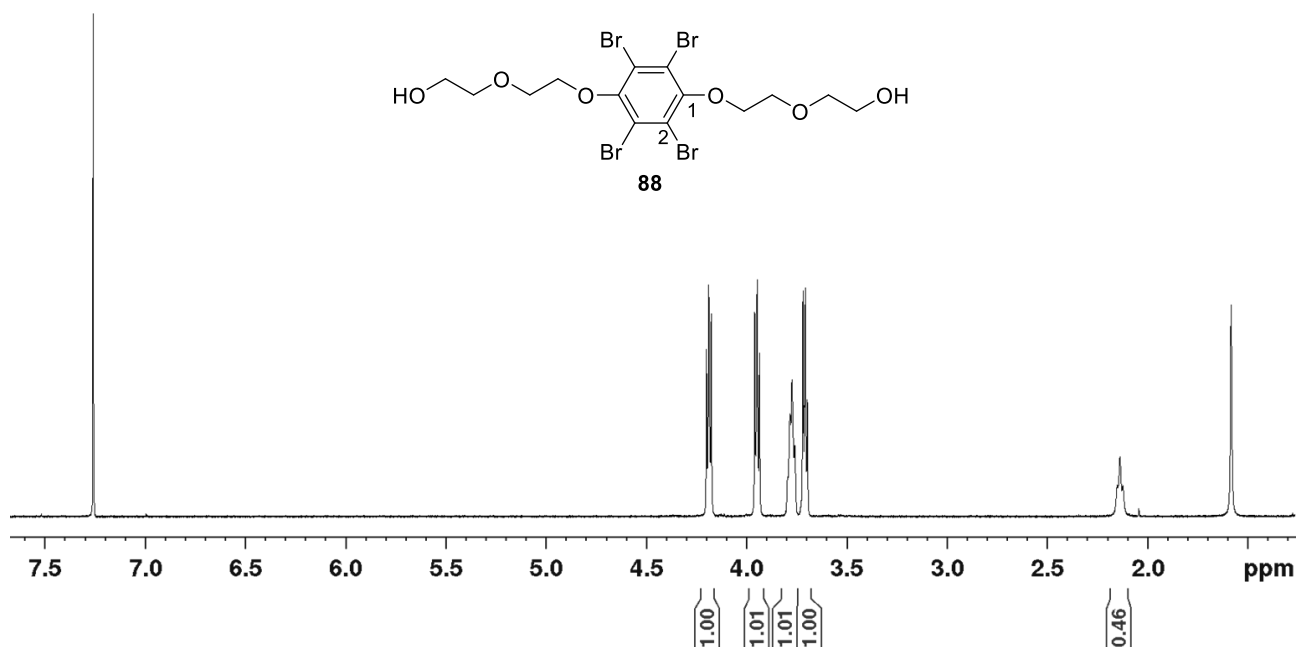
MS spectra were recorded on Waters Micromass Q-TOF (ESI ionization-HRMS). MALDI-TOF MS spectra were recorded on Bruker Daltonics Microflex LT. The following abbreviations are used: CuAAC (copper catalysed azide alkyne cycloaddition), DCM (CH_2Cl_2), DMA (*N,N'*-dimethylacetamide), DMF (*N,N'*-dimethylformamide), DIPEA (*N,N'*-diisopropylethylamine), MW (microwave), TBAF (tetrabutylammonium fluoride), TBTA (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine), THF (tetrahydrofuran). Compounds **38**³⁷ and **87**²⁹ were previously synthesized in our group and already available. Compounds **89**, **95** and **101** are commercially available.

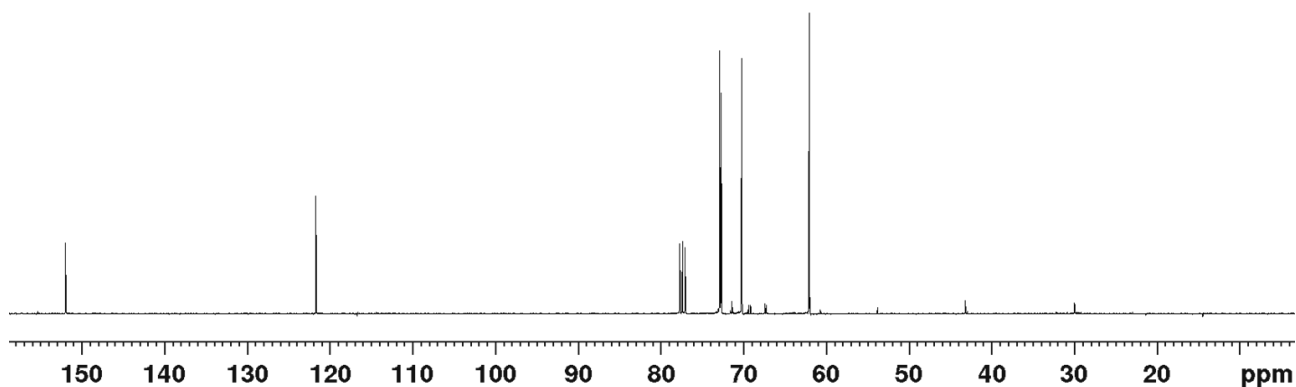
2.5.2 Synthesis of a “cross-shaped” DC-SIGN antagonist

Synthesis of compound **88**

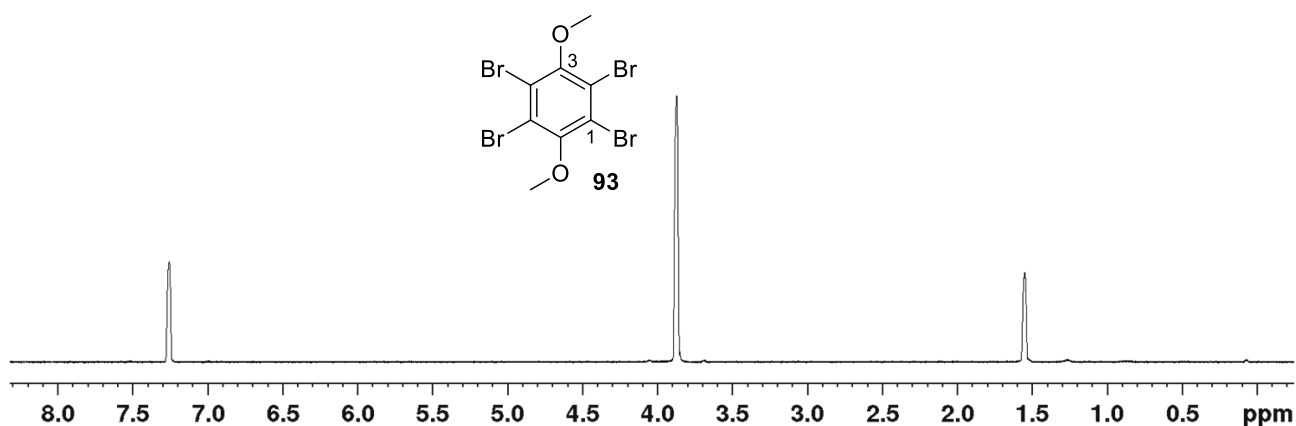
Tetrabromohydroquinone **89** (1 g, 2.35 mmol) was dissolved under nitrogen atmosphere in dry DMA (4.7 mL). K_2CO_3 (1.3 g, 9.41 mmol) was added and the mixture was stirred for 30 min at room temperature turning from a yellow colour to a light orange one. Then, 2(2-chloroethoxy)ethanol (1 mL, 9.47 mmol) was added dropwise and the reaction stirred at 80 °C for 5 d monitoring by TLC (eluent: *n*-hexane - EtOAc, 3:7, $R_f = 0.23$). The reaction was quenched adding a saturated Na_2CO_3 solution (80 mL), the product was extracted with EtOAc (6x60 mL) and reunited organic phases were dried over Na_2SO_4 . After filtration and evaporation of the solvent, the crude was purified by automated flash chromatography (gradient elution: from 40% *n*-hexane - 60% EtOAc to 100% EtOAc) affording pure **88** (1.13 g, 80%) as light pink crystals. ^1H NMR (400 MHz, CDCl_3) δ (ppm): 4.17-4.21 (m, 4 H, $\text{ArOCH}_2\text{CH}_2$), 3.97-3.93 (m, 4 H, $\text{ArOCH}_2\text{CH}_2$), 3.81-3.76 (m, 4 H, $\text{CH}_2\text{CH}_2\text{OH}$), 3.73-3.69 (m, 4 H, $\text{CH}_2\text{CH}_2\text{OH}$), 2.14 (t, $J = 5.7$ Hz, 2 H, OH). ^{13}C NMR (100 MHz, CDCl_3) δ (ppm): 152.0 (2 C, C1), 121.7 (4 C, C2), 72.9 (2 C, $\text{CH}_2\text{CH}_2\text{OH}$), 72.7 (2 C, $\text{ArOCH}_2\text{CH}_2$), 70.2 (2 C, $\text{ArOCH}_2\text{CH}_2$), 62.0 (2 C, $\text{CH}_2\text{CH}_2\text{OH}$). MS (ESI) m/z : calcd for $\text{C}_{14}\text{H}_{18}\text{Br}_4\text{O}_6$ 601.8; found 640.7 $[\text{M}+\text{K}]^+$.

^1H NMR spectrum of **88** in CDCl_3 (400 MHz)



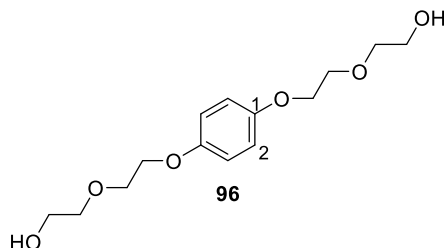
¹³C NMR spectrum of 88 in CDCl₃ (100 MHz)**Synthesis of 1,2,4,5-tetrabromo-3,6-dimethoxybenzene 93**

Tetrabromohydroquinone **89** (500 mg, 1.17 mmol) was dissolved under nitrogen atmosphere in dry DMA (2.3 mL). K₂CO₃ (650 mg, 4.70 mmol) was added and the mixture was stirred for 30 min at room temperature. Then, methyl iodide (290 μL, 4.70 mmol) was added dropwise and the reaction stirred at 80 °C for 4.5 h monitoring by TLC (eluent: *n*-hexane - EtOAc, 10:1, R_f = 0.67). The reaction was quenched adding H₂O (25 mL), the product was extracted with EtOAc (3x25 mL) and reunited organic phases were dried over Na₂SO₄. After filtration and evaporation of the solvent, the crude was purified by automated flash chromatography (gradient elution: from 100% *n*-hexane to 75% *n*-hexane - 25% EtOAc) affording pure **93** (450 mg, 81%). The spectroscopic data are in accordance with those previously reported in the literature.³⁸ ¹H NMR (400 MHz, CDCl₃) δ (ppm): 3.87 (s, OCH₃). ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 153.0 (2 C, C3), 121.4 (4 C, C1), 60.7 (2 C, OCH₃). MS (EI) *m/z*: calcd for C₁₄H₁₈Br₄O₆ 454; found 439 [M-CH₃]⁺, 454 [M]⁺.

¹H NMR spectrum of 93 in CDCl₃ (400 MHz)**Synthesis of compound 96**

Hydroquinone **95** (1 g, 9.1 mmol) was dissolved under nitrogen atmosphere in dry DMA (37.5 mL). K₂CO₃ (10 g, 72.7 mmol) was added and the mixture was stirred for 30 min at room temperature shielding from light. Then, 2(2-chloroethoxy)ethanol (5.3 mL, 36.3 mmol) was added dropwise and the reaction stirred at 80 °C for 8 d monitoring by TLC (eluent: *n*-hexane - EtOAc, 2:8, R_f = 0.18). The reaction was filtered over fritted glass funnel, it was diluted in DCM (50 mL) and washed with H₂O (2x30 mL). The aqueous phases were extracted with DCM (2x30 mL) and reunited organic phases were dried over Na₂SO₄. After filtration and evaporation

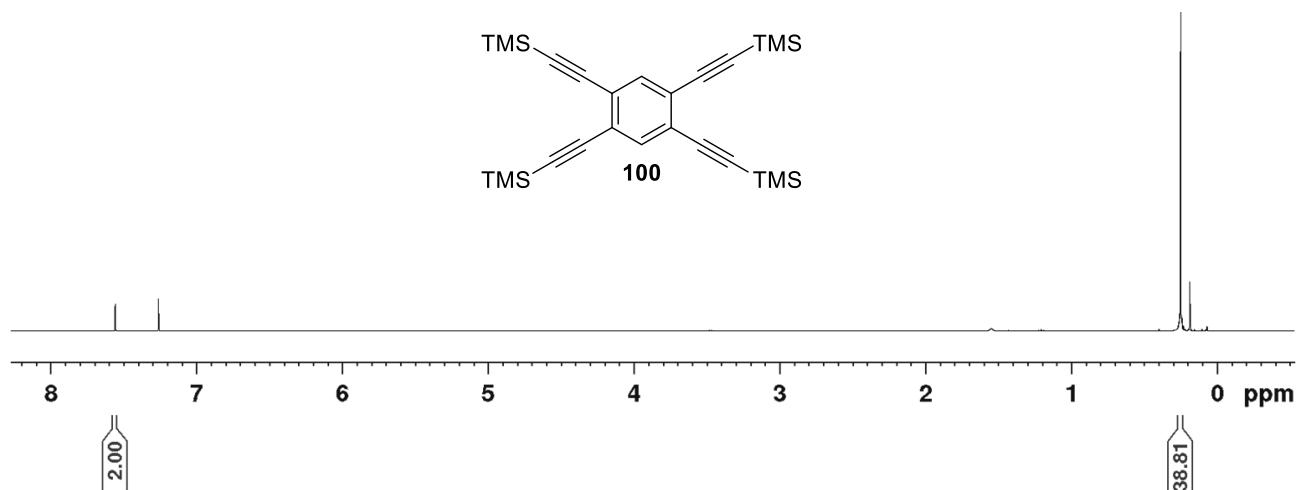
of the solvent, 2.59 g of crude were recovered and judged pure enough by ^1H NMR to be used in the next synthetic step without further purification. ^1H NMR spectroscopic data are in accordance with those previously reported in the literature.³⁹ ^1H NMR (300 MHz, CDCl_3) δ (ppm): 6.85 (s, 4 H, Ar), 4.11-4.06 (m, 4 H, $\text{ArOCH}_2\text{CH}_2$), 3.87-3.82 (m, 4 H, $\text{ArOCH}_2\text{CH}_2$), 3.78-3.73 (m, 4 H, $\text{CH}_2\text{CH}_2\text{OH}$), 3.70-3.64 (m, 4 H, $\text{CH}_2\text{CH}_2\text{OH}$).



Synthesis of 1,2,4,5-tetrakis((trimethylsilyl)ethynyl)benzene **100**³⁵

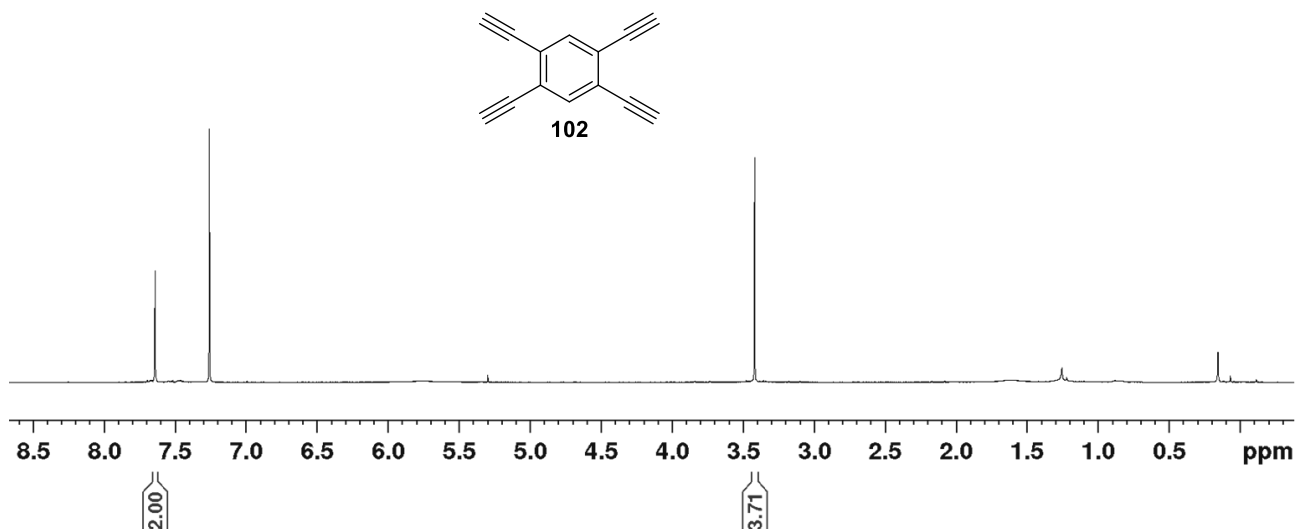
Tetrabromobenzene **101** (158 mg, 0.40 mmol) was dissolved under nitrogen atmosphere with distilled Et_2NH (2 mL) and $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ (7.1 mg, 0.010 mmol), CuI (1.0 mg, 0.005 mmol), ethynyltrimethylsilane (270 μL , 1.92 mmol) were added in the order. The reaction was stirred at 50 $^\circ\text{C}$ for 19 h and TLC analysis showed complete conversion (eluent: *n*-hexane, $R_f = 0.08$). The mixture was filtered over a celite pad and washed with Et_2O . Evaporation of the solvent afforded crude **100** that was pure enough to be used in the next synthetic step without further purification. The spectroscopic data are in accordance with those previously reported in the literature. ^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.56 (s, 2 H, Ar), 0.25 (s, 36 H, $\text{Si}(\text{CH}_3)_3$). MS (ESI) m/z : calcd for $\text{C}_{26}\text{H}_{38}\text{Si}_4$ 462.20; found 485.08 $[\text{M}+\text{Na}]^+$.

^1H NMR spectrum of **100** in CDCl_3 (400 MHz)

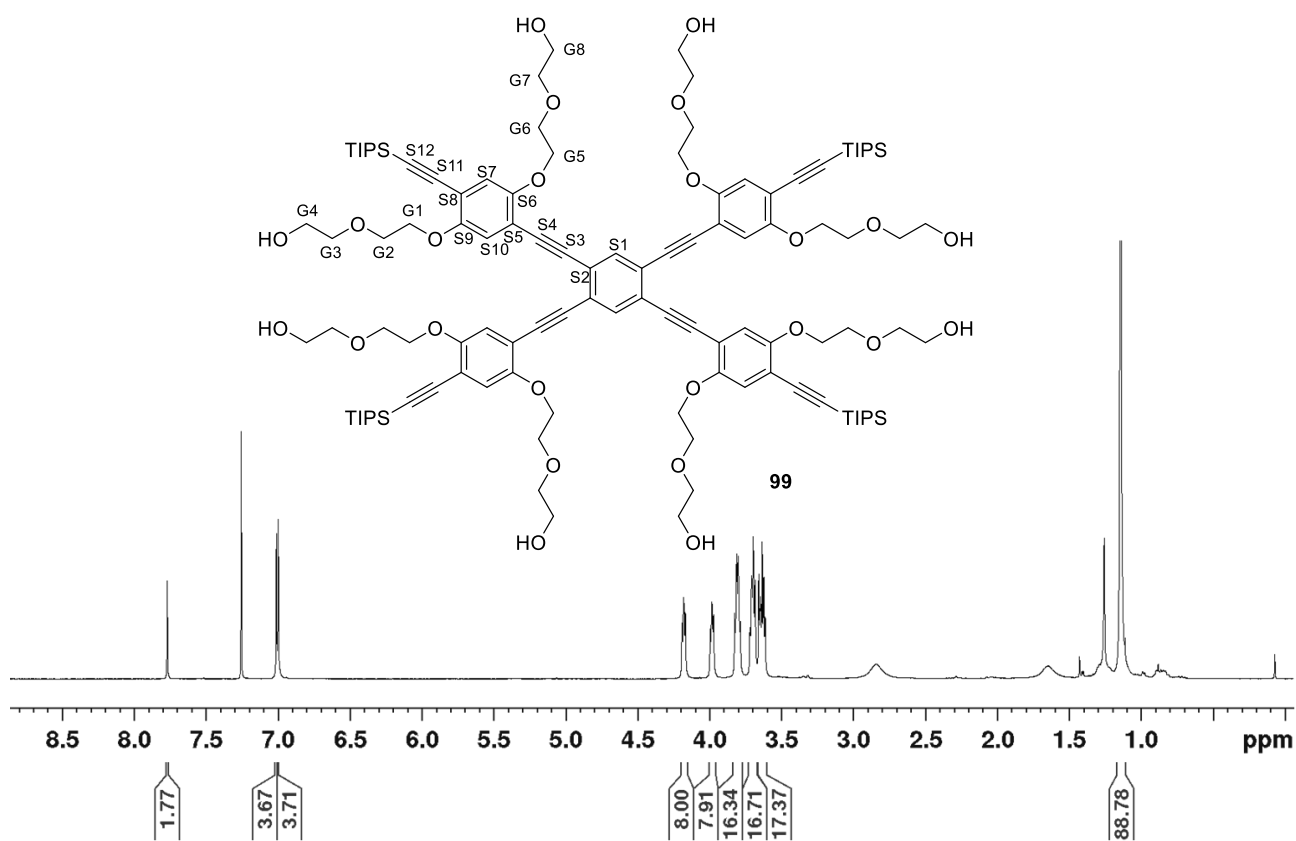
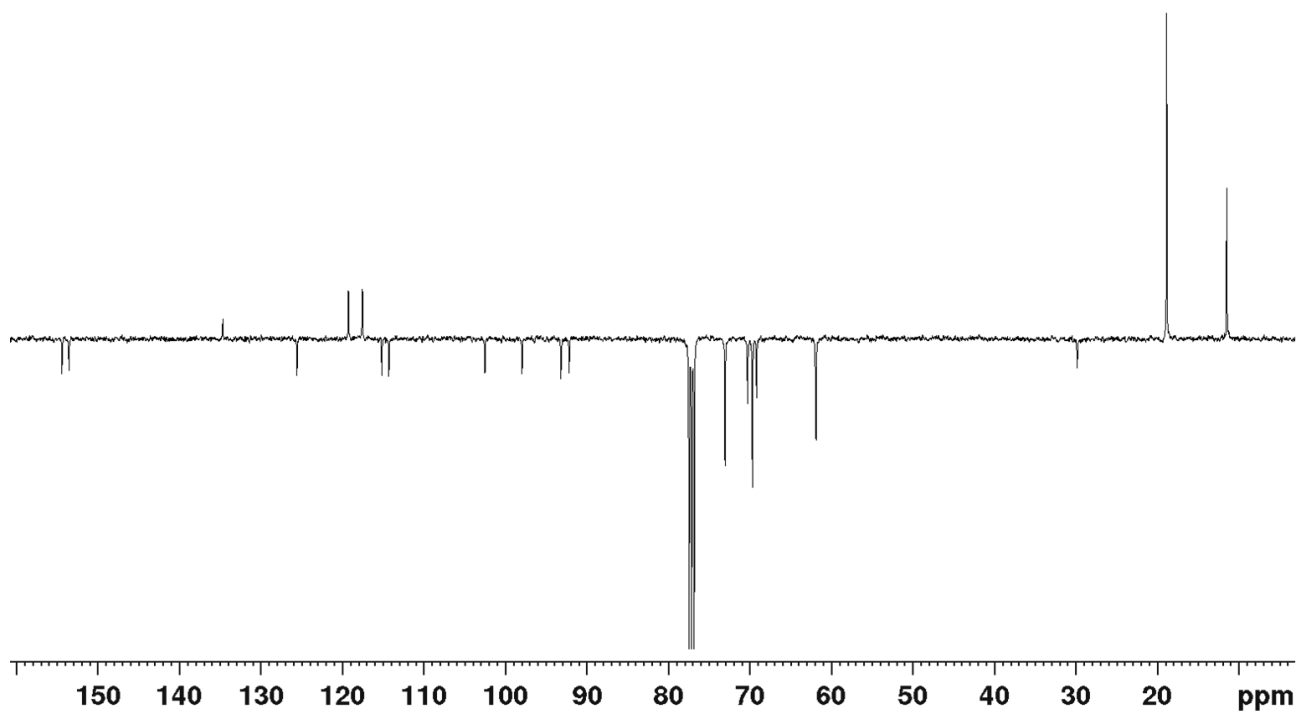


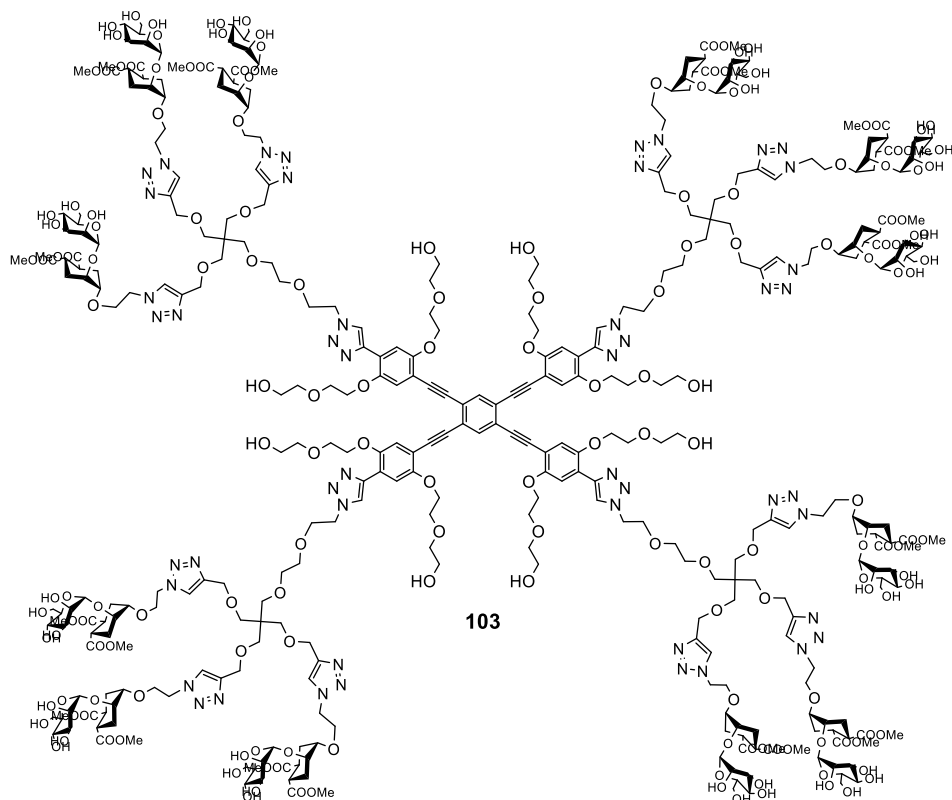
Synthesis of 1,2,4,5-tetraethynyl benzene **102**³⁵

Crude **100** (50.6 mg, 0.109 mmol) was dissolved under nitrogen atmosphere in dry DCM (900 μL). Then a NaOH solution in MeOH (45.2 mg in 700 μL) was added and the reaction was stirred at room temperature for 5 h monitoring by TLC (eluent: *n*-hexane - EtOAc , 20:1, $R_f = 0.33$). The solvent was evaporated, the crude was dissolved in DCM (5 mL) and filtered washing with fresh DCM (5 mL) to remove a white precipitate. The organic phase was washed with brine (2x5 mL) and dried over anhydrous Na_2SO_4 . Evaporation of the solvent afforded crude **102** that was pure enough to be used in the next synthetic step without further purification. The spectroscopic data are in accordance with those previously reported in the literature. ^1H NMR (400 MHz, CDCl_3) δ (ppm): 7.64 (s, 2 H, Ar), 3.42 (s, 4 H, $\text{C}\equiv\text{CH}$). MS (EI) m/z : calcd for C_{14}H_6 174; found 174 $[\text{M}]^+$.

¹H NMR spectrum of **102** in CDCl₃ (400 MHz)Synthesis of compound **99**

Crude **102** (2.7 mg, 0.012 mmol) was dissolved under nitrogen atmosphere in dry THF (70 μ L) and (Ph₃P)₂PdCl₂ (1.3 mg, 0.002 mmol), CuI (1.5 mg, 0.008 mmol), distilled DIPEA (12 μ L, 0.069 mmol) were added in the order. Finally the aryl iodide **87** (40 mg, 0.068) was added as a solution in dry THF (84 μ L). The reaction was stirred at 50 °C for 3 h and complete conversion was assessed by TLC analysis (eluent: DCM - MeOH, 9:1, R_f = 0.61) monitoring at 365 nm. The solvent was evaporated and the product isolated by flash chromatography (eluent: DCM - MeOH, 20:1 for 6 fractions then DCM - MeOH, 15:1). A further purification was performed by size-exclusion chromatography using a Sephadex LH-20 column (\varnothing = 3 cm, height = 50 cm; eluent: MeOH) affording pure **99** (7.4 mg, 30% over three steps from **101**). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.77 (s, 2 H, H_{S1}), 7.02 (s, 4 H, H_{S10}), 7.00 (s, 4 H, H_{S7}), 4.18 (t, *J* = 4.6 Hz, 8 H, H_{G5}), 3.98 (t, *J* = 4.6 Hz, 8 H, H_{G1}), 3.83-3.78 (m, 16 H, H_{G6}, H_{G2}), 3.72-3.68 (m, 16 H, H_{G8}, H_{G4}), 3.66-3.61 (m, 16 H, H_{G7}, H_{G3}), 1.14 (s, 84 H, TIPS). ¹³C NMR (100 MHz, CDCl₃) δ (ppm): 154.4 (4 C, C_{S9}), 153.6 (4 C, C_{S6}), 134.7 (2 C, C_{S1}), 125.6 (4 C, C_{S2}), 119.3 (4 C, C_{S7}), 117.6 (4 C, C_{S10}), 115.2 (4 C, C_{S8}), 114.3 (4 C, C_{S5}), 102.5 (4 C, C_{S11}), 98.0 (4 C, C_{S12}), 93.2 (4 C, C_{S3}), 92.2 (4 C, C_{S4}), 73.1 (8 C, C_{G3}, C_{G7}), 70.3 (4 C, C_{G5}), 69.7 (8 C, C_{G2}, C_{G6}), 69.2 (4 C, C_{G1}), 62.0, 61.9 (8 C, C_{G4}, C_{G8}), 18.9 (24 C, SiCH(CH₃)₂), 11.5 (12 C, SiCH(CH₃)₂). MS (ESI) *m/z*: calcd for C₁₁₄H₁₆₆O₂₄Si₄ 2032.09; found 700.3 [M+3Na]³⁺, 1038.93 [M+2Na]²⁺, 2054.88 [M+Na]⁺. MS (MALDI) *m/z*: calcd for C₁₁₄H₁₆₆O₂₄Si₄ 2032.1; found 2056.1 [M+Na]⁺ (matrix DHB).

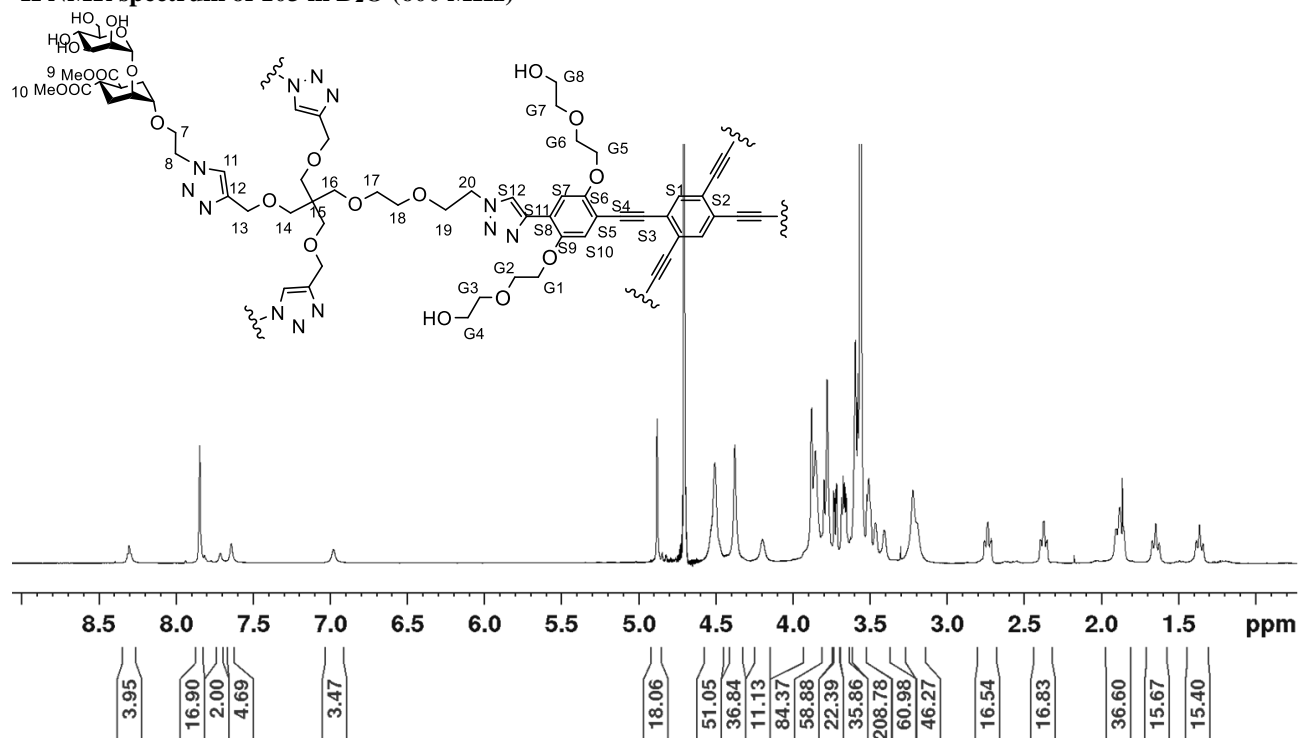
^1H NMR spectrum of **99** in CDCl_3 (400 MHz) ^{13}C NMR spectrum of **99** in CDCl_3 (100 MHz)

Synthesis of compound **103**

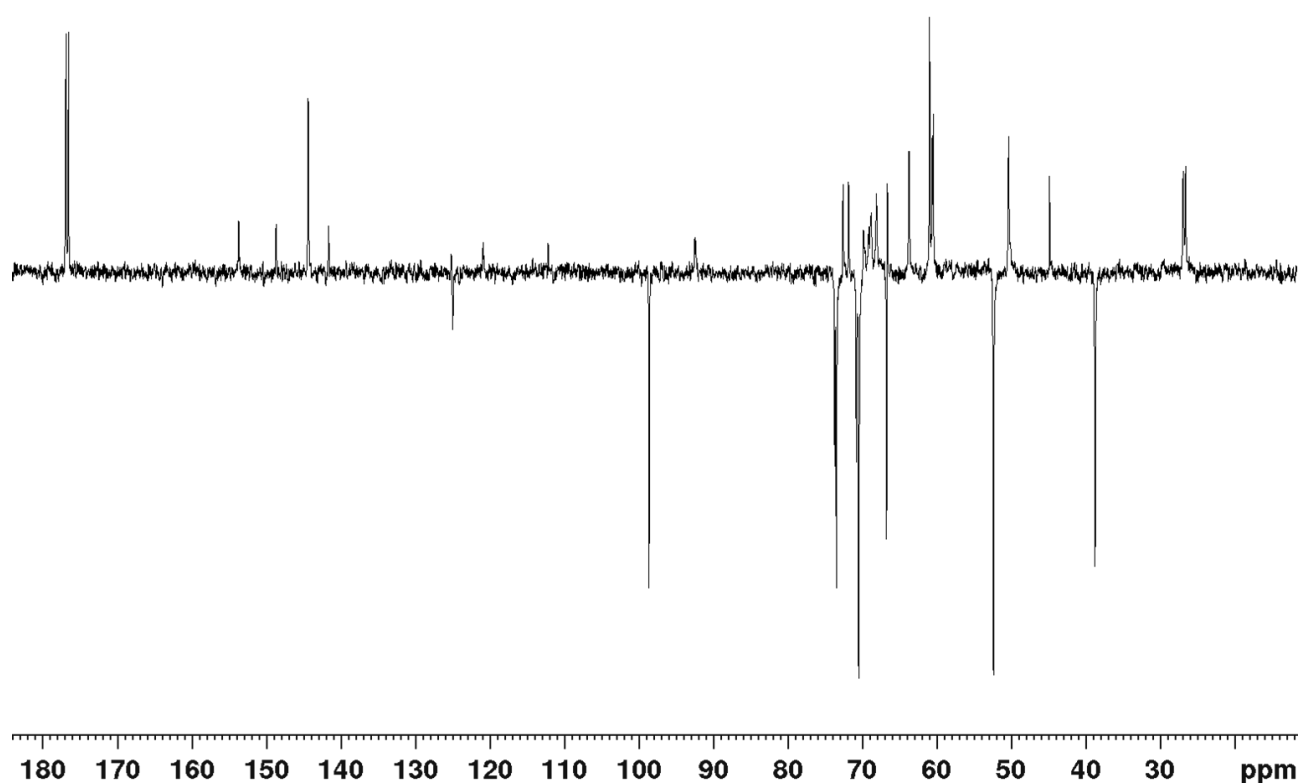
The tetravalent cross-shaped scaffold **99** (5.3 mg, 2.6 μmol) was dissolved in freshly distilled THF (105 μL) under nitrogen atmosphere. TBAF (10 μL) was added as a 1 M solution in THF and the reaction was stirred at room temperature for 1 h. Complete deprotection was assessed by TLC analysis (eluent: DCM - MeOH, 9:1, $R_f = 0.29$) monitoring at 365 nm. A solution of TBTA (280 μg , 0.53 μmol) in freshly distilled THF (38 μL) was added, followed by 13 μL of a solution of $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ (60 μg , 0.24 μmol) and 17 μL of a solution of sodium ascorbate (210 μg , 1.06 μmol) both in degassed H_2O . Finally, dendron **38** (20 mg, 11.4 μmol) was added followed by THF (94 μL) and H_2O (102 μL) to reach a ~ 2:1 THF/ H_2O mixture. The reaction was stirred at room temperature, under nitrogen atmosphere, shielded from light for 15 h. The complete conversion into the desired product was assessed by TLC analysis (eluent: DCM - MeOH, 7:3 + 0.5 H_2O , $R_f = 0.22$) monitoring at 365 nm and by MALDI-TOF MS (matrix DHB, HCCA). The copper scavenger QuadraSil MP was added to the solution which was stirred for 15 min. After filtering, the crude was finally purified by size-exclusion chromatography using a Sephadex LH-20 column ($\varnothing = 3$ cm, height = 50 cm; eluent: MeOH) and monitoring by TLC (eluent: DCM - MeOH, 7:3 + 0.5 H_2O). Dendrimer **103** was recovered as a bright yellow oil (20.3 mg, 92%). The purity was confirmed by HPLC analysis of an analytical sample (Waters Atlantis T3 5 μm 4.6x100 mm column), plateau at 90% ($\text{H}_2\text{O} + 0.1\%$ TFA) – 10% ($\text{CH}_3\text{CN} + 0.1\%$ TFA) for 1 min followed by a gradient to 100% ($\text{CH}_3\text{CN} + 0.1\%$ TFA) in 10 min, t_R (product) = 7.0 min. $[\alpha]_{\text{D}}^{16} = +28.5$ ($c = 0.49$ in MeOH). ^1H NMR (600 MHz, D_2O) δ (ppm): 8.30 (s, 4 H, $\text{H}_{\text{S}12}$), 7.85 (s, 12 H, 11-H), 7.72 (s, 2 H, $\text{H}_{\text{S}1}$), 7.65 (s, 4 H, $\text{H}_{\text{S}7}$ or $\text{H}_{\text{S}10}$), 6.99 (s, 4 H, $\text{H}_{\text{S}7}$ or $\text{H}_{\text{S}10}$), 4.88 (s, 12 H, 1-H), 4.50 (bs, 32 H, 8-H, 20-H), 4.37 (s, 24 H, 13-H), 4.20 (s, 8 H, $\text{H}_{\text{G}5}$), 3.90-3.74 (m, 84 H, 2-H, 7-H, 1- H_D , 6-H, $\text{H}_{\text{G}1}$, $\text{H}_{\text{G}2}$, $\text{H}_{\text{G}6}$), 3.72 (dd, $J = 9.5, 3.3$ Hz, 12 H, 3-H), 3.70-3.38 (m, 176 H, 6- H' , 2- H_D , 4-H, 5-H, 9-H, 10-H, 17-H, 18-H, 19-H, $\text{H}_{\text{G}3}$, $\text{H}_{\text{G}4}$, $\text{H}_{\text{G}7}$, $\text{H}_{\text{G}8}$), 3.26-3.16 (m, 32 H, 14-H, 16-H), 2.73 (td, $J = 12.1, 3.0$ Hz, 12 H, 5- H_D), 2.37 (td, $J = 12.1, 2.7$ Hz, 12 H, 4- H_D), 1.88 (t, $J = 14.0$ Hz, 24 H, 6- H_{Deq} , 3- H_{Deq}), 1.64 (t, $J = 13.2$ Hz, 12 H, 3- H_{Dax}), 1.37 (t, $J = 13.2$ Hz, 12 H, 6- H_{Dax}). ^{13}C NMR (100 MHz, D_2O) δ (ppm): 176.9 (12 C, C=O), 176.6 (12 C, C=O), 153.7, 148.7 (8 C, $\text{C}_{\text{S}9}$, $\text{C}_{\text{S}6}$), 144.4 (12 C, C12), 141.6 (4 C, $\text{C}_{\text{S}11}$), 135.0 (2 C, $\text{C}_{\text{S}1}$), 125.5 (4 C, $\text{C}_{\text{S}12}$), 125.0 (12 C, C11), 120.9, 117.0 (4 C, $\text{C}_{\text{S}7}$ or $\text{C}_{\text{S}10}$), 112.2 (4 C, $\text{C}_{\text{S}7}$ or $\text{C}_{\text{S}10}$), 111.6, 98.6 (12 C, C1), 92.5, 73.7 (12 C, C2 $_D$), 73.5 (12 C, C4), 72.6, 71.8,

70.8 (12 C, C_{1D}), 70.4 (24 C, C₂, C₃), 69.8, 69.4 (8 C, C_{G5}), 68.8, 68.1, 66.7 (12 C, C₅), 66.6 (12 C, C₇), 63.7 (12 C, C₁₃), 61.0 (12 C, C₆), 60.6, 60.5 (8 C, C_{G4}, C_{G8}), 52.4, 52.3 (24 C, C₉, C₁₀), 50.3 (16 C, C₈, C₂₀), 44.9 (4 C, C₁₅), 38.7 (24 C, C_{4D}, C_{5D}), 26.9 (12 C, C_{3D}), 26.6 (12 C, C_{6D}). HRMS (ESI) *m/z*: calcd for C₃₆₆H₅₃₄N₄₈O₁₇₆ 8421.44331; found 1426.56540 [M+6Na]⁶⁺, 1707.28033 [M+5Na]⁵⁺, 1711.67260 [M+6Na]⁵⁺, 2128.36951 [M+4Na]⁴⁺, 8421.46951 by deconvolution.

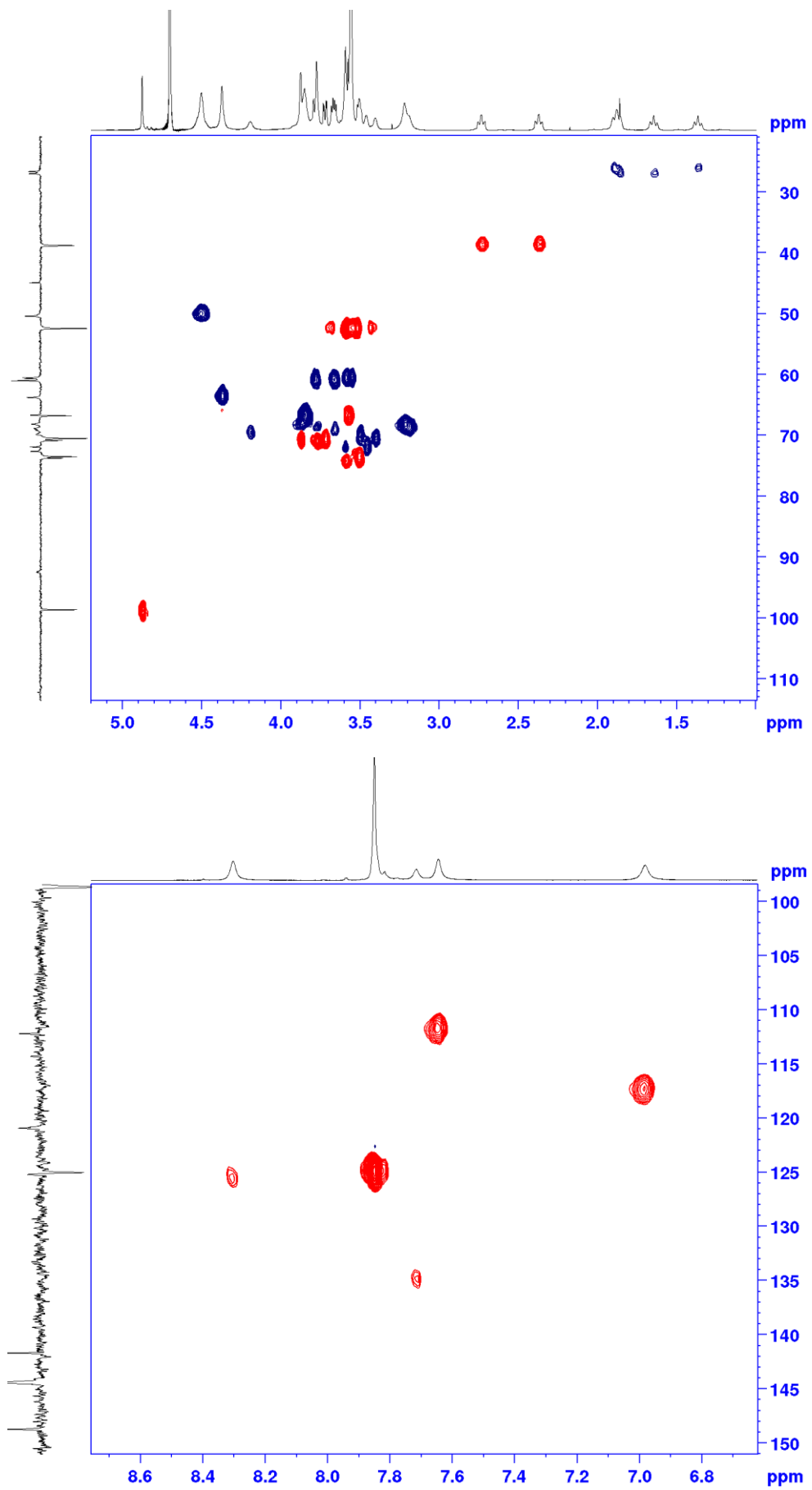
¹H NMR spectrum of 103 in D₂O (600 MHz)



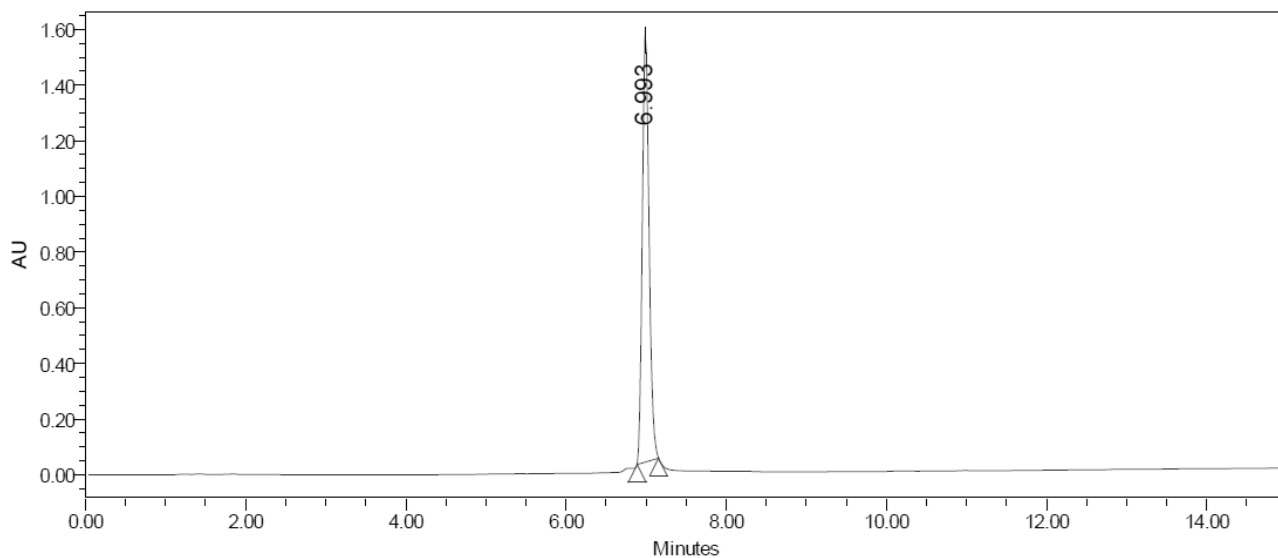
¹³C NMR spectrum of 103 in D₂O (100MHz)



HSQC NMR spectrum of 103 in D₂O (600 MHz)

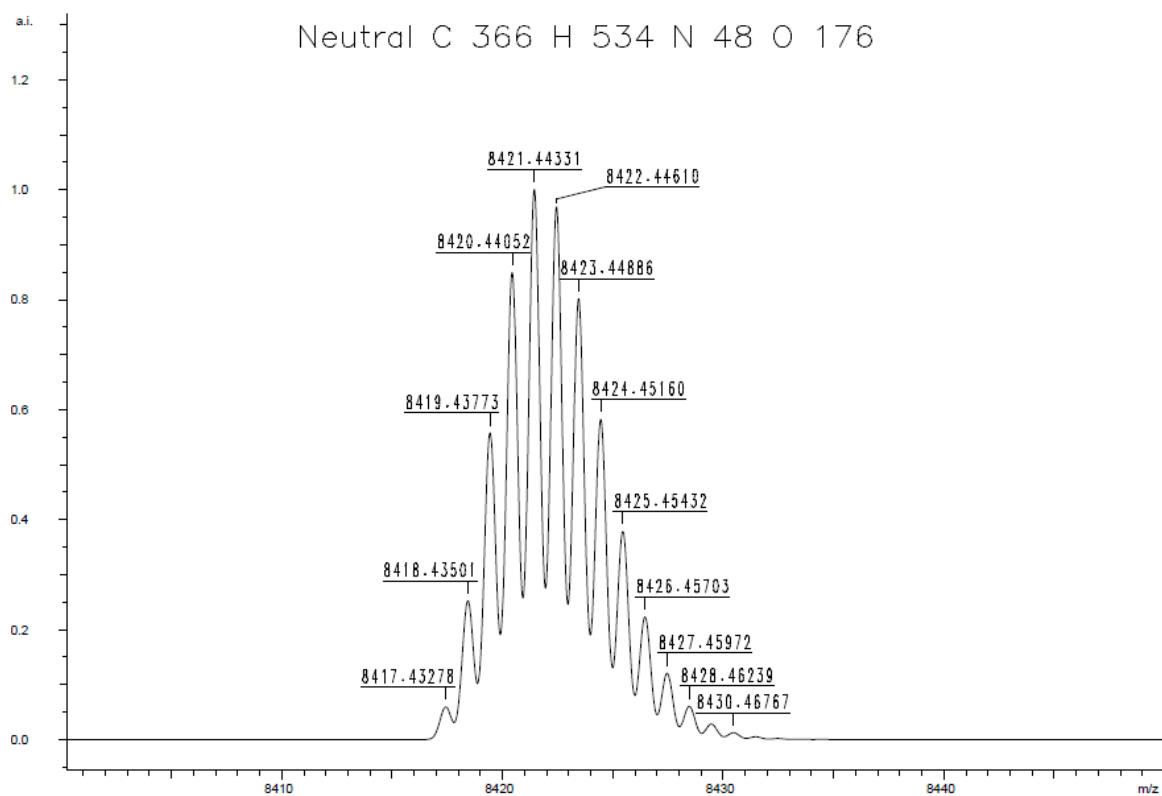


HPLC of analytical sample of 103: HPLC trace. Waters Atlantis T3 5 μm 4.6x100 mm column, a plateau at 90% (H_2O + 0.1% TFA) – 10% (CH_3CN +0.1%) for 1 min was followed by a gradient from 90% (H_2O + 0.1% TFA) – 10% (CH_3CN +0.1%) to 100% (CH_3CN +0.1%) in 10 min followed by a plateau for 1 min.



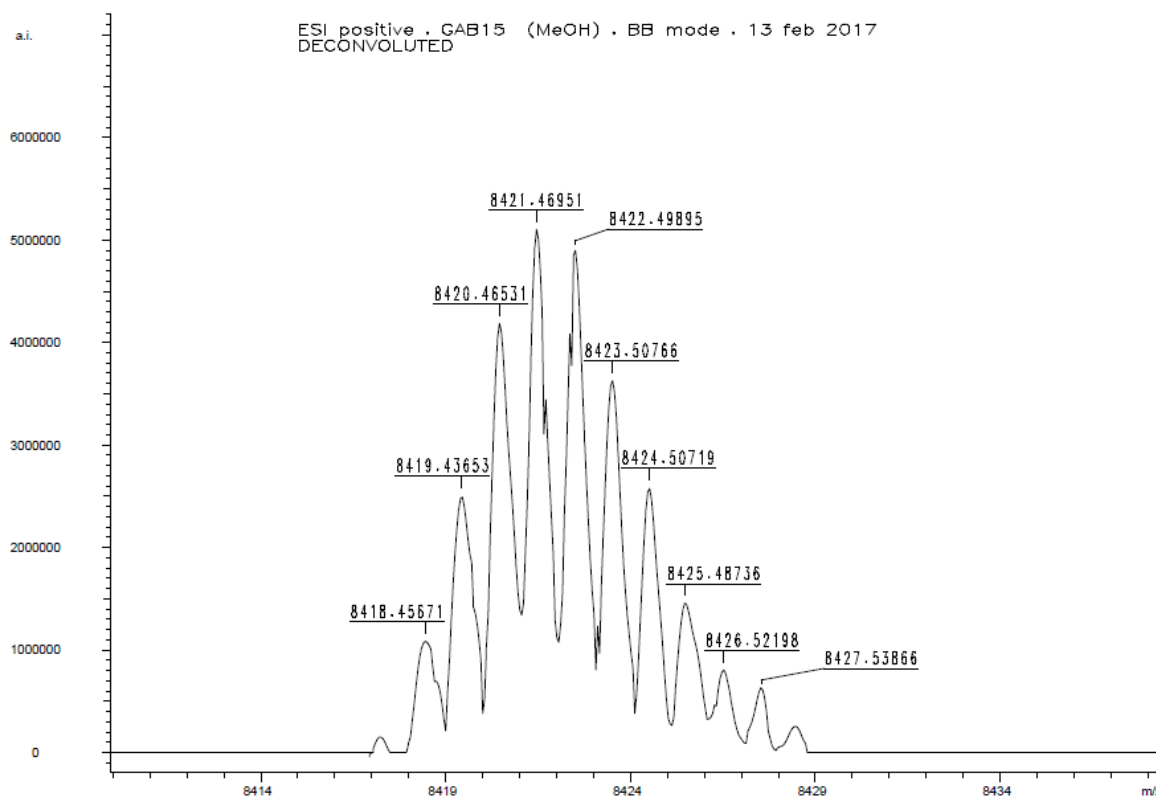
Calculated (a) and found (b) isotopic distribution for dendron 103 HRMS (ESI):

a) Calculated:



Chapter 2

b) Found:



Calcd for neutral $[C_{366}H_{534}N_{48}O_{176}]$: 8421.44331.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	8417.43278	Not detected	0.0592	Not detected
2	8418.43501	8418.45671	0.2520	0.2125
3	8419.43773	8419.43653	0.5573	0.4919
4	8420.44052	8420.46531	0.8492	0.8168
5	8421.44331	8421.46951	1.0000	1.0000
6	8422.44610	8422.49895	0.9682	0.9638
7	8423.44886	8423.50766	0.8013	0.7119
8	8424.45160	8424.50719	0.5820	0.5067
9	8425.45432	8425.48736	0.3782	0.2857
10	8426.45703	8426.52198	0.2230	0.1578
11	8427.45972	8427.53866	0.1207	0.1238
12	8428.46239	Not detected	0.0605	Not detected

Calcd for $[C_{366}H_{534}N_{48}O_{176}Na_4]^{4+}$: 8513.40007.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	2127.34749	Not detected	0.0592	Not detected
2	2127.59801	Not detected	0.2520	Not detected
3	2127.84867	2127.83845	0.5573	0.5362

4	2128.09936	2128.11542	0.8492	0.8809
5	2128.35005	2128.36951	1.0000	1.0000
6	2128.60074	2128.63333	0.9682	0.9660
7	2128.85143	2128.88062	0.8013	0.6766
8	2129.10211	2129.14678	0.5820	0.4681
9	2129.35278	2129.40902	0.3782	0.3660
10	2129.60346	Not detected	0.2230	Not detected
11	2129.85413	Not detected	0.1207	Not detected
12	2130.10479	Not detected	0.0605	Not detected

Calcd for $[\text{C}_{366}\text{H}_{533}\text{N}_{48}\text{O}_{176}\text{Na}_6]^{5+}$: 8558.37124.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1710.87226	Not detected	0.0592	Not detected
2	1711.07266	Not detected	0.2521	Not detected
3	1711.27318	Not detected	0.5574	Not detected
4	1711.47372	1711.48194	0.8493	0.7530
5	1711.67428	1711.67260	1.0000	1.0000
6	1711.87483	1711.87500	0.9683	0.8300
7	1712.07537	1712.07173	0.8014	0.6275
8	1712.27591	1712.28010	0.5821	0.4980
9	1712.47645	1712.50227	0.3782	0.3806
10	1712.67699	Not detected	0.2230	Not detected
11	1712.87752	Not detected	0.1207	Not detected
12	1713.07805	Not detected	0.0605	Not detected

Calcd for $[\text{C}_{366}\text{H}_{534}\text{N}_{48}\text{O}_{176}\text{Na}_5]^{5+}$: 8536.38929.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1706.47585	Not detected	0.0592	Not detected
2	1706.67626	Not detected	0.2520	Not detected
3	1706.87679	Not detected	0.5573	Not detected
4	1707.07733	1707.08130	0.8492	0.7853
5	1707.27789	1707.28033	1.0000	1.0000
6	1707.47844	1707.48293	0.9683	0.9633
7	1707.67898	1707.69000	0.8014	0.7827
8	1707.87953	1707.89063	0.5821	0.5576
9	1708.08007	Not detected	0.3782	Not detected
10	1708.28061	Not detected	0.2230	Not detected
11	1708.48114	Not detected	0.1207	Not detected
12	1708.68167	Not detected	0.0605	Not detected
13	1708.88220	Not detected	0.0283	Not detected

Calcd for $[C_{366}H_{534}N_{48}O_{176}Na_6]^{6+}$: 8559.37851.

Peak	Mass		Relative intensity	
	Calculated	Found	Calculated	Found
1	1425.89476	Not detected	0.0592	Not detected
2	1426.06176	Not detected	0.2521	Not detected
3	1426.22886	1426.23787	0.5573	0.5478
4	1426.39598	1426.39920	0.8492	0.8391
5	1426.56311	1426.56540	1.0000	1.0000
6	1426.73023	1426.73855	0.9683	0.9739
7	1426.89736	1426.90216	0.8014	0.7000
8	1427.06447	1427.06811	0.5820	0.6130
9	1427.23159	Not detected	0.3782	Not detected
10	1427.39870	Not detected	0.2230	Not detected
11	1427.56582	Not detected	0.1207	Not detected
12	1427.73292	Not detected	0.0605	Not detected
13	1427.90003	Not detected	0.0283	Not detected
14	1428.06713	Not detected	0.0124	Not detected
15	1428.23423	Not detected	0.0051	Not detected

2.6 References

- ¹ D. A. Mitchell, A. J. Fadden, K. Drickamer *J. Biol. Chem.* **2001**, 276, 28939-28945.
- ² G. Tabarani, M. Thépaut, D. Stroebel, C. Ebel, C. Vivès, P. Vachette, D. Durand, F. Fieschi *J. Biol. Chem.* **2009**, 284, 21229-21240.
- ³ R. M. Steinman *Cell* **2000**, 100, 491-494.
- ⁴ A. C. Cambi, C. G. Figdor *Curr. Opin. Cell Biol.* **2003**, 15, 539-546.
- ⁵ J. Banchemereau, R. M. Steinman *Nature* **1998**, 392, 245-252.
- ⁶ F. Sallusto, M. Cella, C. Danieli, A. Lanzavecchia *J. Exp. Med.* **1995**, 182, 389-400.
- ⁷ K. Inaba, M. Inaba, M. Naito, R. M. Steinman *J. Exp. Med.* **1993**, 178, 479-488.
- ⁸ Y. van Kooyk, B. Appelmek, T. B. H. Geijtenbeek *Trends Mol. Med.* **2003**, 9, 153-159.
- ⁹ Y. van Kooyk, T. B. H. Geijtenbeek *Nat. Rev. Immunol.* **2003**, 3, 697-709.
- ¹⁰ B. J. Appelmek *J. Immunol.* **2003**, 170, 1635-1639.
- ¹¹ T. B. H. Geijtenbeek, D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. F. van Duijnhoven, J. Middel, I. L. M. H. A. Cornelissen, H. S. L. M. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor, Y. van Kooyk *Cell* **2000**, 100, 587-597.
- ¹² Y. van Kooyk, A. Engering, A. N. Lekkerkerker, I. S. Ludwig, T. B. H. Geijtenbeek *Curr. Opin. Immunol.* **2004**, 16, 488-493.
- ¹³ R. C. Gallo *Retrovirology* **2006**, 3, 72.
- ¹⁴ L. Wu, V.N. KewalRamani *Nat. Rev. Immunol.* **2006**, 6, 859-868.
- ¹⁵ D. McDonald, L. Wu, S. M. Bohks, V. N. KewalRamani, D. Unutmaz, T. J. Hope *Science* **2003**, 300, 1295-1297.
- ¹⁶ J.-F. Arrighi, M. Pion, E. Garcia, J.-M. Escola, Y. van Kooyk, T. B. H. Geijtenbeek, V. Piguet *J. Exp. Med.* **2004**, 200, 1279-1288.
- ¹⁷ S. G. Turville, J. J. Santos, I. Frank, P. U. Cameron, J. Wilkinson, M. Miranda-Saksena, J. Dable, H. Stössel, N. Romani, M. Piatak Jr, J. D. Lifson, M. Pope, A. L. Cunningham *Blood* **2004**, 103, 2170-2179.
- ¹⁸ E. Garcia, M. Pion, A. Pelchen-Matthews, L. Collinson, J.-F. Arrighi, G. Blot, F. Leuba, J.-M. Escola, N. Demaurex, M. Marsh, V. Piguet *Traffic* **2005**, 6, 488-501.
- ¹⁹ C. Thery, L. Zitvogel, S. Amigorena *Nat. Rev. Immunol.* **2002**, 2, 569-579.
- ²⁰ R. D. Wiley, S. Gummuluru *Proc. Nat. Acad. Sci. U.S.A.* **2006**, 103, 738-743.
- ²¹ L. de Witte, A. Nabatov, T. B. H. Geijtenbeek *Trends Mol. Med.* **2008**, 14, 12-19.

-
- ²² J.-F. Arrighi, M. Pion, M. Wiznerowicz, T. B. H. Geijtenbeek, E. Garcia, S. Abraham, F. Leuba, V. Dutoit, O. Ducrey-Rundquist, Y. van Kooyk, D. Trono, V. Piguet *J. Virol.* **2004**, *78*, 10848-10855.
- ²³ H. Feinberg, A. S. Powlesland, M. E. Taylor, W. I. Weis *J. Biol. Chem.* **2010**, *285*, 13285-13293.
- ²⁴ J. Valladeau, O. Ravel, C. Dezutter-Dambuyant, K. Moore, M. Kleijmeer, Y. Liu, V. Duvert-Frances, C. Vincent, D. Schmitt, J. Davoust, C. Caux, S. Lebecque, S. Saeland *Immunity* **2000**, *12*, 71-81.
- ²⁵ L. de Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. W. P. de Jong, T. de Gruijl, V. Piguet, Y. van Kooyk, T. B. H. Geijtenbeek *Nat. Med.* **2007**, *13*, 367-371.
- ²⁶ S. Ordanini, N. Varga, V. Porkolab, M. Thépaut, L. Belvisi, A. Bertaglia, A. Palmioli, A. Berzi, D. Trabattoni, M. Clerici, F. Fieschi, A. Bernardi *Chem. Commun.* **2015**, *51*, 3816-3819.
- ²⁷ A. Berzi, S. Ordanini, B. Joosten, D. Trabattoni, A. Cambi, A. Bernardi, M. Clerici *Sci. Rep.* **2016**, *6*, 35373.
- ²⁸ N. Varga, I. Sutkeviciute, C. Guzzi, J. McGeagh, I. Petit-Haertlein, S. Gugliotta, J. Weiser, J. Angulo, F. Fieschi, A. Bernardi *Chem. Eur. J.* **2013**, *19*, 4786-4797.
- ²⁹ F. Pertici, N. Varga, A. van Duijn, M. Rey-Carrizo, A. Bernardi, R. J. Pieters *Beilstein J. Org. Chem.* **2013**, *9*, 215-222.
- ³⁰ Stefania Ordanini, 2014, PhD thesis, University of Milan, R09678.
- ³¹ D. Janietz, S. Katholy, B. Dietzel, L. Brehmer *Langmuir* **1997**, *13*, 7217-7222.
- ³² M. Sonoda, A. Inaba, K. Itahashi, Y. Tobe *Org. Lett.* **2001**, *3*, 2419-2421.
- ³³ R. G. Brisbois, R. A. Wanke, K. A. Stubbs, R. V. Stick, Iodine monochloride *e-EROS Encyclopedia of Reagents for Organic Synthesis*, **2001**.
- ³⁴ F. Buonsanti, L. Lattuada, G. Leonardi, F. Uggeri, E. Vignale, M. Visigalli *EP 2394984A1*.
- ³⁵ S. Yuan, B. Dorney, D. White, S. Kirklin, P. Zapol, L. Yu, D.-J. Liu *Chem. Commun.* **2010**, *46*, 4547-4549.
- ³⁶ W. C. Still, M. Kahn, A. Mitra *J. Org. Chem.* **1978**, *43*, 2923-2925.
- ³⁷ N. Varga, I. Sutkeviciute, R. Ribeiro-Viana, A. Berzi, R. Ramdasi, A. Daggetti, G. Vettoretti, A. Amara, M. Clerici, J. Rojo, F. Fieschi, A. Bernardi *Biomaterials* **2014**, *35*, 4175-4184.
- ³⁸ M. Rack, B. Hauschel, M. Hanack *Chem. Ber.* **1996**, *129*, 237-242.
- ³⁹ C. Wang, H. Tsai, H. Shih, S. Jeon, Z. Xu, D. Williams, S. S. Iyer, T. C. Sanchez, L. Wang, M. Cotlet, H. Wang *ACS Appl. Mater. Interfaces* **2010**, *2*, 738-747.

CHAPTER 3

Glycomimetic Lead Structures Identification:
Exploiting MALDI-TOF Detectable Glycan Microarrays**3.1 Glycan Microarrays: general features and applications**

An important contribution in the study of protein-carbohydrate interactions is offered by glycan microarray technology. A microarray consists of a library of compounds of interest, which are displayed in a controlled fashion over a suitable surface. This technique dates back to the mid '90s, when it was first applied to DNA¹ and proteins^{2,3} and was readily adapted to glycans. In particular, the first glycan microarrays were described in 2002 by several research groups, who achieved printing of glycans by automated arrayers^{4,5} or immobilized them with various methodologies over printed slides or multiwell plates.^{6,7,8,9,10} Since then, the application of this technique has literally exploded and different methodologies concerning preparation of the arrays and their analysis have been developed.¹¹

The success of glycan microarrays can be understood looking at the general features of the experiment (Fig. 3.1). In a typical set-up, selected glycans are immobilized, usually using an automatic spotter, over properly functionalized or coated slides. Following immobilization, target lectins are incubated with the array to study their interactions with printed glycans, while unbound proteins are washed away. In most of the cases, the extent of lectin affinity for each glycan of the library is detected by fluorescence. The intensity of fluorescence is a direct measurement of the bound lectin, which can be obtained either using already labelled proteins or indirectly, performing a second incubation with a labelled tag able to recognize and bind to the lectin, e.g. an antibody or fluorescently labelled streptavidin in the case of biotinylated lectins. Therefore, this process allows displaying in a well defined and tunable arrangement a large number of compounds over a relatively small surface. Moreover, the entire library can be screened by fast analysis requiring tiny amounts of both glycans (usually pmol) and lectins. Finally, the high density of monovalent ligands immobilized on the array is able to mimic carbohydrate-protein interactions as they occur in Nature, with lectins showing avidity and selectively binding to multivalent counterparts.

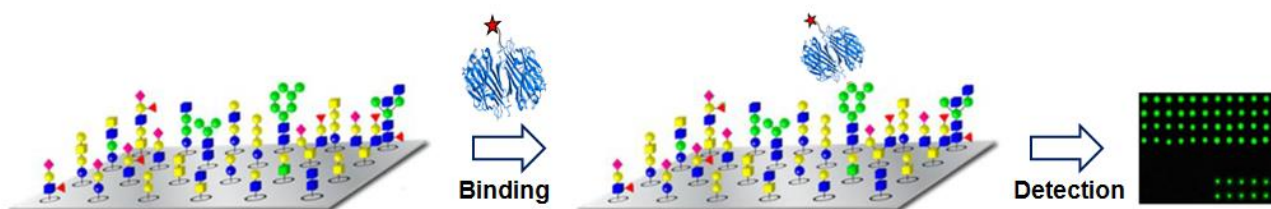


Figure 3.1 – Glycan microarray screening. In a typical experiment involving glycan microarrays, several different glycan structures are immobilized on a surface and screened for interactions with a target lectin (either labelled or not). The extent of lectin binding is usually revealed by fluorescence spectroscopy.

The density of spotted ligands is an important parameter that has to be considered in microarray analysis. Indeed, if glycan concentration is too high, saturation of the binding signal may occur with the lectin showing comparable binding affinity both for high affinity and low affinity ligands. On the contrary, printing glycans at reduced concentration allows highlighting the preferential binding of lectins for different carbohydrate structures.¹² In recent years, the influence of ligand orientation has also been investigated. Various studies have pointed out that carbohydrate-protein interactions of biological systems are only roughly mimicked by bidimensional microarrays and that tridimensional presentation of glycan ligands may play a crucial role in directing binding selectivity.¹³ These effects have been investigated developing three-dimensional microarrays,¹⁴ in which printed glycans have been grafted on various synthetic scaffolds, comprising neoglycoproteins,¹⁵ glycodendrimers¹⁶ or glycopolymers.^{17,18} However, a high degree of control over glycan

surface density and presentation is hard to achieve mainly due to difficulties in the characterisation of the device surface.

Fast screening of large glycan libraries is one of the most attractive properties of microarray analysis that has been exploited to assess binding specificities of newly discovered lectins. Among these, screening campaigns have been addressed also towards several C-type lectins, including DC-SIGN, langerin, Dectin-1, MGL and SIGN-R1-7.^{19,20,21,22,23} In this manner, naturally occurring ligands can be identified opening the doors to more accurate structural studies and allowing rational design of glycomimetic analogues.

Glycan microarrays have also been addressed to the identification of synthetic lectin antagonists. In a recent work, Paulson and coworkers have described the identification of a highly potent Siglec-7 ligand through the screening of a sialoside library that was directly synthesised on the array.²⁴ Remarkably, a library of 1140 sialoside analogues was readily obtained by printing a set of 12 different sialosides bearing an alkyne moiety followed by CuAAC reaction with 94 different azides.

Microarrays represent a promising tool for diagnostic purposes.²⁵ This technique combines fast screening and low detection limits. Moreover, bacteria and viruses can be directly incubated on the array surface to study their interactions and selectivities towards the printed glycans.²⁶ The Seeberger group has demonstrated that *E. coli*, which can bind to surfaces exploiting the adhesive lectin FimH, selectively binds to mannose in carbohydrate microarrays.²⁷ This result suggests the possible application of glycan arrays as a test for the detection of pathogens. Indeed, characterizing different bacterial species on the basis of their binding specificity towards different glycan structures would allow, in theory, to assess the presence and to discriminate between different species of pathogens in strains and biological fluids.

3.2 Immobilization chemistry

The immobilization of glycans is the key step for the generation of carbohydrate microarrays. Depending on the type of chemistry on which they rely, immobilization strategies can be classified into covalent immobilization methodologies and non-covalent immobilization methodologies. In any case, a proper functionalization of the glycans is necessary to match the chemical nature of the array surface.

3.2.1 Covalent immobilization

This technique requires the formation of a new covalent bond between the array surface and the glycans (Fig. 3.2). Covalent immobilization and particularly the robust amine chemistry is probably the most exploited strategy for the synthesis of glycan microarrays. Reported studies employing this methodology comprise the immobilization of amino tethered glycans with a surface activated as *N*-hydroxysuccinimide (NHS)-ester²⁶ or containing reactive epoxide groups.²⁸ Condensation reactions between aldehydes and amines, hydrazides and *O*-substituted hydroxylamines have also been applied; notably this strategy includes the use of non functionalised reducing glycans as the aldehyde counterpart, which however determines the loss of the reducing end residue of the glycan in its native form.

Along with the amine chemistry, thiol chemistry allowed for robust immobilization technique. Thiol-ene reactivity between a thiol group and a maleimide moiety is the most commonly adopted combination,^{29,30} however disulphide exchange reaction has also been applied allowing the synthesis of an SPR detectable array.³¹

Cycloaddition reactions have been successfully used for glycan immobilization, in particular Diels Alder⁹ and CuAAC^{8,18} reaction have been applied to array libraries of monosaccharides and oligosaccharides.

Finally, a convenient method that allows glycan immobilization using non derivatized sugars is the photoimmobilization. Activating the array surface with aryl-trifluoromethyl diazirine or phthalimido moieties the photochemical immobilization of bacterial polysaccharides, dextrans, xyloglucans and glycoproteins has

been achieved.^{32,33,34,35} A reverse strategy has also been established, in which perfluoroarylazide functionalized mono and disaccharides were photochemically linked onto a non activated PEG-ylated surface.³⁶

Covalent immobilization

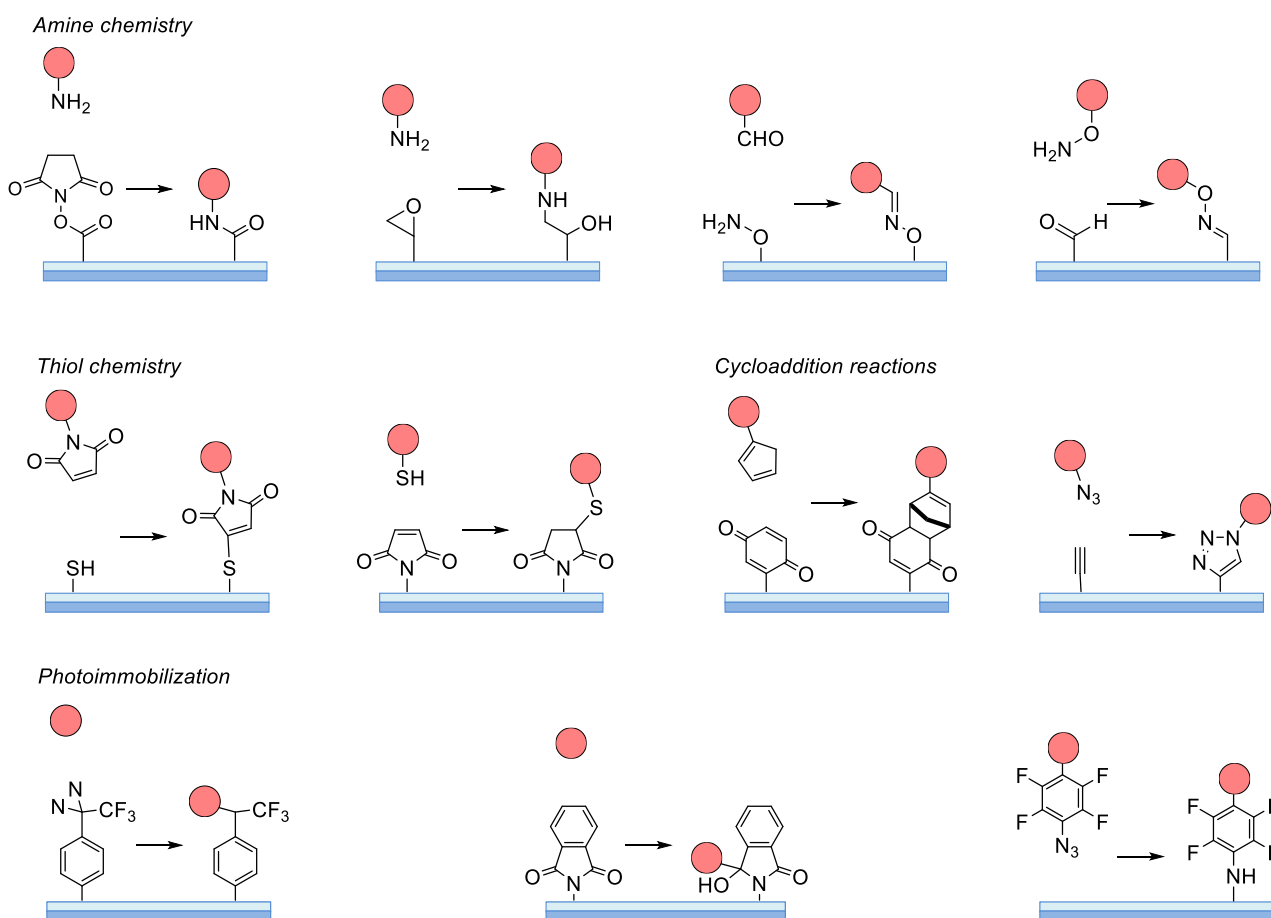


Figure 3.2 – Covalent immobilization techniques. Most of the described methodologies for the fabrication of microarrays rely on covalent immobilization, where a covalent bond is directly formed between glycans and the array surface. Robust amine chemistry is the most developed, but also thiol chemistry along with cycloaddition reactions and photoimmobilization processes have been reported.

3.2.2 Non-covalent immobilization

This strategy envisages the adsorption of selected glycans over the array surface through non covalent interactions (Fig. 3.3). Non-covalent methods were reported since the early development of glycan microarrays, when it was observed that polysaccharides, glycoproteins and proteoglycans spontaneously adsorb on nitrocellulose and oxidized polystyrene surfaces.^{5,6} A clear disadvantage of this methodology is the uncontrolled, random orientation of the immobilized glycans.

With a similar approach, electrostatic interactions were exploited to immobilize naturally charged glycosaminoglycans or properly modified negatively charged dextrans over a positively charged amino tethered surface.^{37,38}

The two methodologies described so far are both limited to high molecular weight glycans. On the contrary, non-covalent immobilization of mono up to pentasaccharides has been reported using glycolipids. Glycolipids are obtained tethering glycans with a long alkyl chain, thus allowing hydrophobic interactions and consequent adsorption over a polystyrene surface.⁷

A similar principle was used by the Pohl group who adopted a perfluorinated surface (C_8F_{17}) to immobilize monosaccharides bearing the same perfluorinated (C_8F_{17}) chain.³⁹

Finally, DNA⁴⁰ and biotin-streptavidin⁴¹ based methods have been also described, even though the latter has never been applied for immobilization of monosaccharides.

Non-covalent immobilization

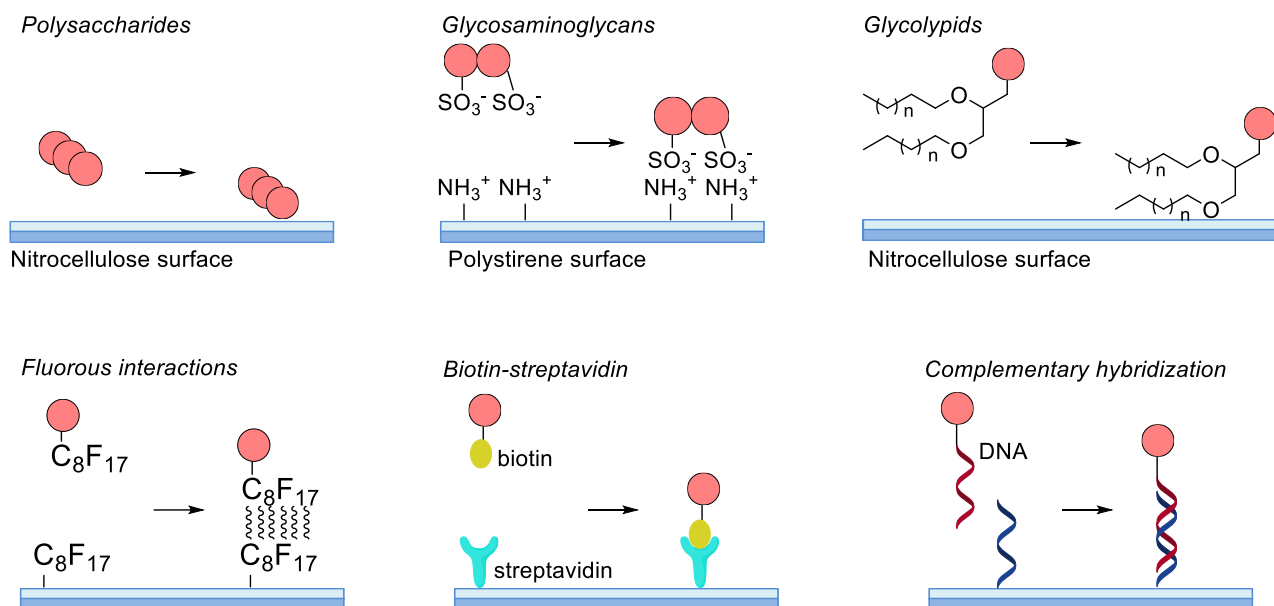
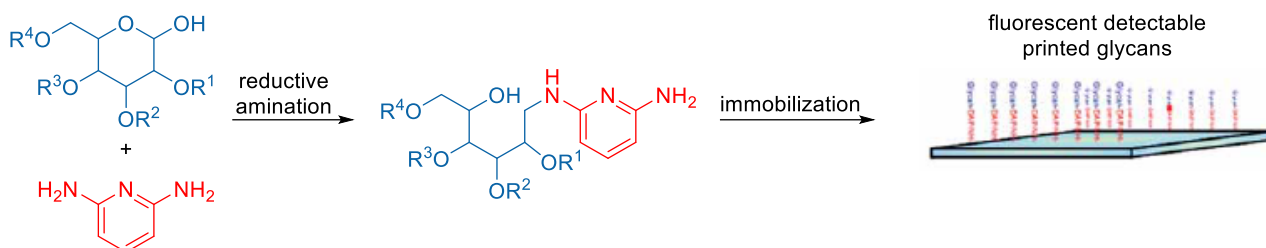


Figure 3.3 – Non-covalent immobilization techniques. Non-covalent immobilization envisages the absorption of conveniently functionalized glycans on a surface, usually by hydrophobic, fluoruous or electrostatic interactions. Polysaccharides have shown to spontaneously absorb on nitrocellulose surfaces, while other methods exploit biotin/streptavidin interaction or ssDNA pairing.

3.3 Detection techniques and surface characterization

Among the possible applications of glycan microarrays, the investigation of carbohydrate-lectin interactions is certainly one of the most attractive. The most common technique to measure the relative affinities between lectins and printed saccharides is fluorescence spectroscopy, even though SPR detection has been successfully applied.^{31,42,43,44}

One of the most challenging aspects of glycan microarrays fabrication is represented by surface characterization. With the aim of quantifying the extent of glycan immobilization Cummings and co-workers developed a method using fluorescent labelled glycans.⁴⁵ The selected carbohydrates were functionalized with a fluorescent diaminopyridyl moiety prior to immobilization. Then the reaction over epoxide activated slides was monitored and quantified measuring the fluorescence intensity (Scheme 3.1). The limitation of this application resides in the reductive amination conjugation step that determines loss of structural complexity of the reducing glycan moiety.



Scheme 3.1 – Immobilization of fluorescent glycans. Reductive amination of glycan reducing end with diaminopyridine was exploited to immobilize fluorescent detectable carbohydrates. This method allowed to characterize array surface in terms of glycan density by fluorescent spectroscopy.⁴⁵

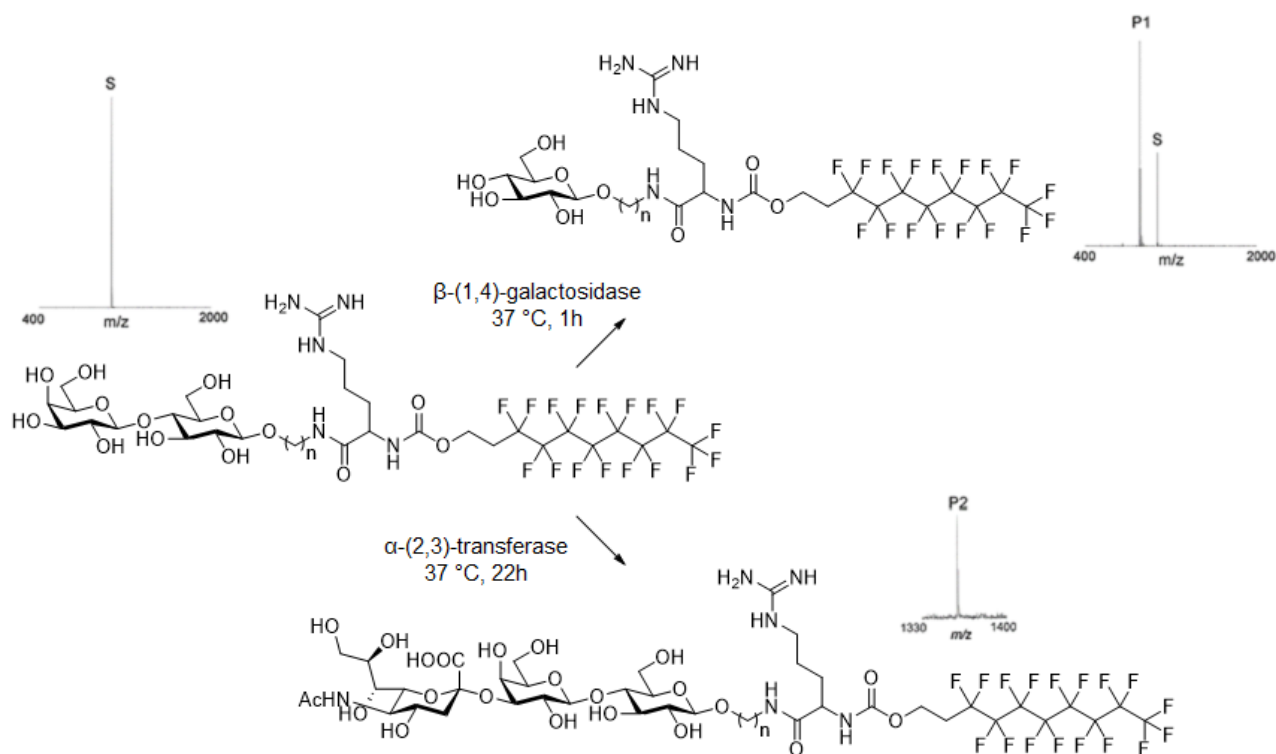
Printing of a fluorescent dye on a given array was established by Wong and co-workers⁴⁶ as a general method to estimate glycan density. In particular, solutions of fluorescein cadaverine were printed at different concentrations onto *N*-hydroxysuccinimide activated slides and fluorescence was detected by an array scanner. After the immobilization time, the slides were washed with a PBS buffer solution and analysed again by fluorescence. Plotting the average fluorescence intensity obtained from quenched slides with the logarithm of the concentration allowed to establish the concentration at which saturation of the surface is reached. Moreover, knowing the solution concentrations (*C*) and the volumes (*V*) employed during printing, the density of spotted compound could be calculated as follows:

$$\delta = \frac{Np}{\pi \cdot r^2} \quad \text{where} \quad Np = \frac{C \cdot V \cdot N_A \cdot Q_{post}}{Q_{pre}}$$

where *Np* is the number of spotted molecules, *r* the radius of the spots, *N_A* the Avogadro's constant, *Q_{post}* is the fluorescence intensity post-quench and *Q_{pre}* is the fluorescence intensity pre-quench. The limit of this approach is the use of fluorescein as a dye which presents very different morphology and chemophysical properties when compared to glycans. Indeed, steric bulk and topology are likely to affect be important parameters affecting glycan density on arrays, which may vary significantly depending on the complexity and architecture of printed glycans.

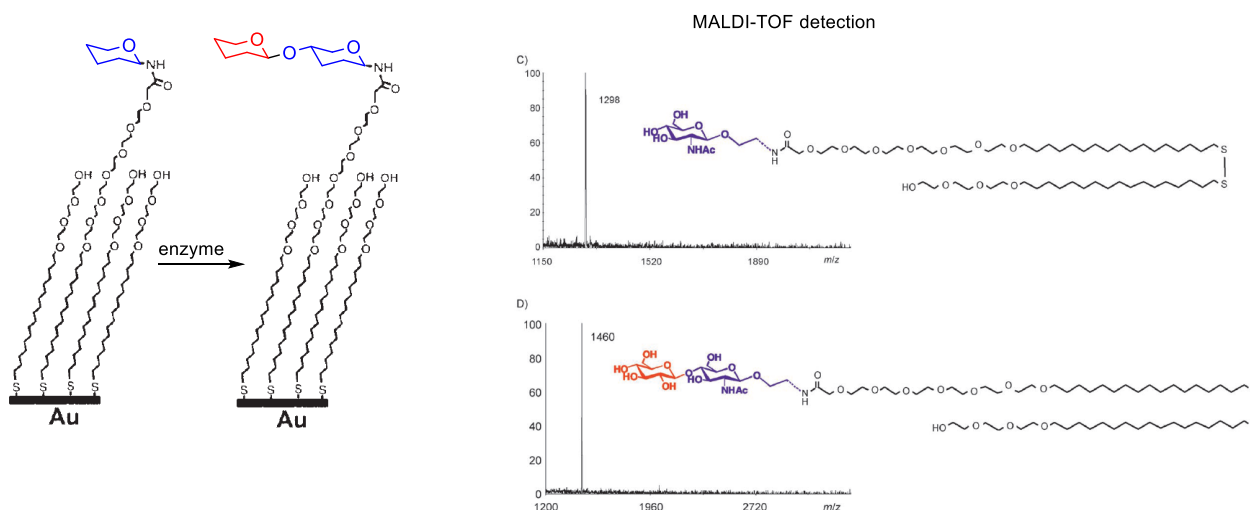
An alternative way to characterize arrays surface is provided by mass spectroscopy. In a pioneering work, oligosaccharides were conveniently functionalised with a photocleavable linker and covalently immobilized on porous silicon based slides. Photolytic cleavage and analysis of the array was possible using desorption/ionization on silicon mass spectroscopy (DIOS-MS) technique. Unfortunately this method is affected by a low signal to noise ratio (S/N), which is mostly due to incomplete cleavage of the linker.⁴⁷

Surface characterization by mass spectroscopy is a powerful technique that has been also exploited to monitor on-chip enzymatic reactions. The on-chip transformation of the lactose derivative **S** using either a β-1,4-galactosidase or a α-2,3-sialyltransferase has been reported (Scheme 3.2).⁴⁸ The extent of the enzymatic reactions was assessed by the exotic Nanostructure-Initiator Mass Spectrometry (NIMS) technique. Particularly, **S** was tethered with a fluorour-tagged linker that enabled immobilization through non-covalent interactions on a perfluorinated surface. Upon laser irradiation, desorption and ionization take place and carbohydrates can be detected. Notably, the presence a guanidine moiety in the saccharide linker favored the ionization process providing a methodology with a low detection limit (S/N ~ 20 at 500 amol).



Scheme 3.2 – On-chip reactions are detectable by MS spectrometry. Glycan arrays based on compounds immobilized through fluororous interactions are suitable to be analysed by NIMS (Nanostructure-Initiator Mass Spectrometry). In particular, glycan *S* was elaborated using either a β -1,4-galactosidase or a α -2,3-sialyltransferase. The formation of the corresponding products *P1* and *P2* was assessed by NIMS.⁴⁸

In a similar manner, arrays based on self-assembled monolayers (SAMs) on a gold surface have been described and exploited to monitor on-chip enzymatic modifications of printed saccharides (Scheme 3.3). This platform offers great versatility in terms of detection, with SPR analysis and MALDI-TOF spectrometry being particularly efficient.⁴⁹



Scheme 3.3 – SAMs based glycan arrays. Glycosylation reactions of glycans immobilized to self-assembled monolayers on a gold chip were successfully monitored by MALDI-TOF MS. From the two spectra we can appreciate the shifting of the mass signal towards higher molecular weight upon glycosylation reaction.⁴⁹

In a recent work, Reichardt and coworkers demonstrated that hydrophobically immobilized glycan arrays are a suitable device for studying lectin-carbohydrate interactions by MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometry (MS).⁵⁰ Particularly, a small library of *N*-glycans tethered with an amino group was printed over an indium-tin oxide (ITO) surface coated with an NHS activated lipidic double layer. The resulting glycan array (Fig. 3.4) proved to be a versatile device allowing both fluorescent and MALDI-TOF MS detection. As a proof of principle experiment, concanavalin A (Con A) was incubated with the spotted glycans and binding interactions were detected by MALDI-TOF MS. Moreover, partial on-chip trypsin digestion and MALDI-TOF/TOF analysis allowed peptide sequencing and assignment of fragment ions, pointing out that this method is a powerful technology for the identification of unknown lectins binding to already characterized glycan molecular patterns.

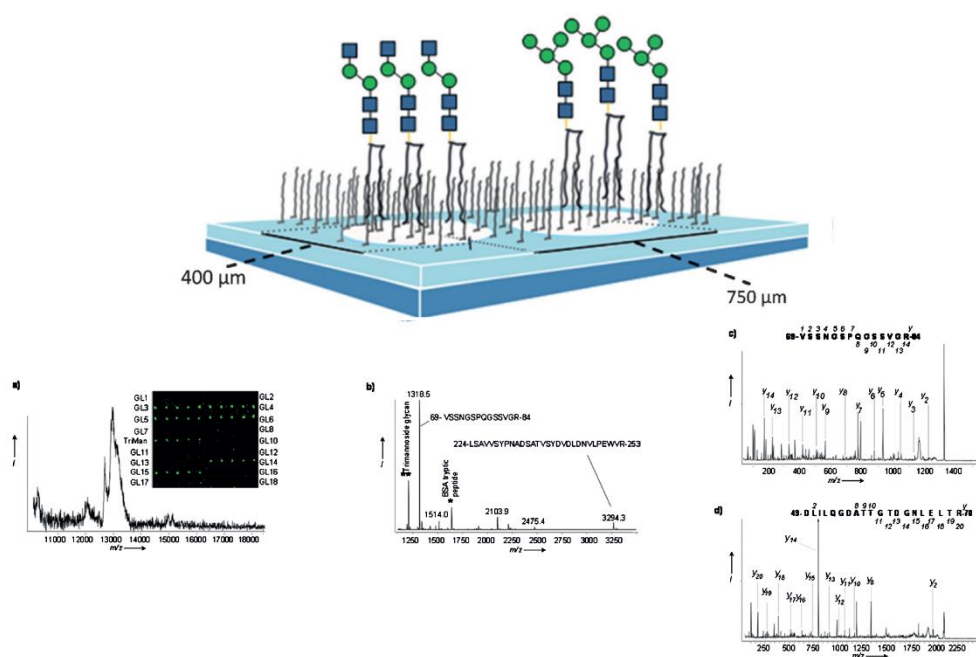


Figure 3.4 – Hydrophobically immobilized glycan arrays detectable by MALDI-TOF MS. Glycan-lectin interactions are detectable by MALDI-TOF MS using hydrophobically immobilized glycan arrays. **a)** The signal of concanavalin A (Con A) was detected for glycans able to interact with this protein. Control experiment was performed using fluorescently labelled Con A. **b)** This technique would also allow for determination of structure of unknown proteins. As a proof of concept, on chip trypsin digestion of Con A was performed leading to fragments that were detected by MALDI-TOF MS. **c)** and **d)** Sequencing of the protein main fragments was achieved by on-chip MALDI-TOF/TOF analysis.⁵⁰

3.4 On-chip glycomimetic library expansion

MALDI-TOF detectable glycan arrays combine fast screening with ease of surface characterization, which make them an attractive platform for the study of carbohydrate-protein interactions and, in particular, for the identification of novel lectin ligands. Indeed, new lead structures can be selected among hundreds of compounds by fast screening of wide glycan libraries. Moreover, analysis by MALDI-TOF spectrometry furnishes a characterisation of the surface, glycans can be easily identified, their on-chip elaborations can be properly monitored and interactions with lectins can be successfully detected.

In collaboration with the group of Niels C. Reichardt at CICbiomaGUNE in San Sebastian (Spain), we aimed to exploit the features of MALDI TOF detectable glycan arrays to perform the on-chip chemoenzymatic synthesis of a glycomimetic library (Fig. 3.5). We thought that the ease of surface characterization of these devices would allow to synthesize and expand chemically or enzymatically the library of glycans directly on

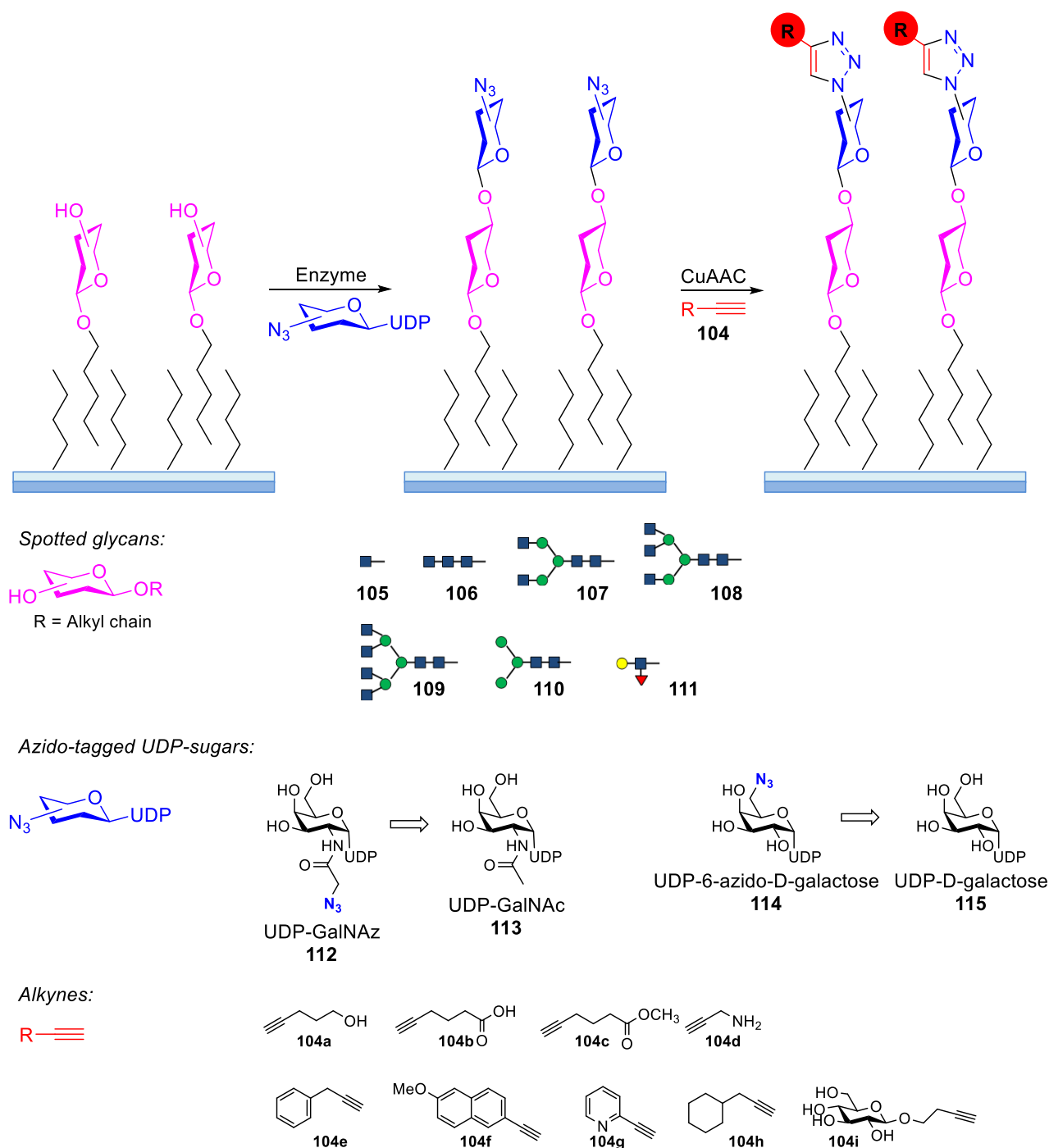
the array. The strategy that we adopted envisages the use of glycans functionalized with alkyl chains embedded onto hydrophobic surface from which they can be desorbed after laser ionization. Elaboration of the glycans was investigated by enzymatic reactions, which constitute a convenient method for the stereoselective transformation of glycans, both in solution and directly on micro-arrays. In particular, our goal was to enzymatically elongate arrayed glycans using “clickable” NDP-sugar donors, allowing further library diversification by CuAAC reaction. Following this methodology, large libraries of glycomimetic able to explore different chemical spaces can be rapidly generated and tested as potential ligand in interactions with C-type lectins. Remarkably, each elaboration step and the extent of the reactions performed on the array can be monitored by MALDI-TOF MS analysis.



Figure 3.5 – General strategy for on-chip library expansion. A small library of glycans will be properly functionalized directly on the array allowing further reactivity and diversification of glycan library. This will allow fast generation of different glycan structures. Moreover, the extent of reactions will be detectable by MALDI-TOF MS.

3.4.1 On-array chemoenzymatic reactions: optimization studies

To assess the feasibility of the approach that we devised, both on-chip enzymatic glycosylation and CuAAC reactions were investigated and optimized in a model experiment (Scheme 3.4). In detail, glycans ranging from a monosaccharide to a nonasaccharide **105-111** were selected to build up a glycan array. ITO slides were chosen as a suitable surface for both MS and fluorescence detection and properly functionalized by self-assembly to form a hydrophobic monolayer on which glycans tethering long alkyl chains got adsorbed. Then, glycan elongation by enzymatic glycosylation was tested using two glycoside donors: the commercially available UDP-GalNAz **112** and the 6-azido-UDP-Gal **114**, which was previously synthesized in our group⁵¹ following a known procedure.⁵² The introduction of azido tagged glycoside units allowed for further diversification and library expansion by CuAAC reaction, which was optimized and applied to a set of nine alkynes **104a-i**.

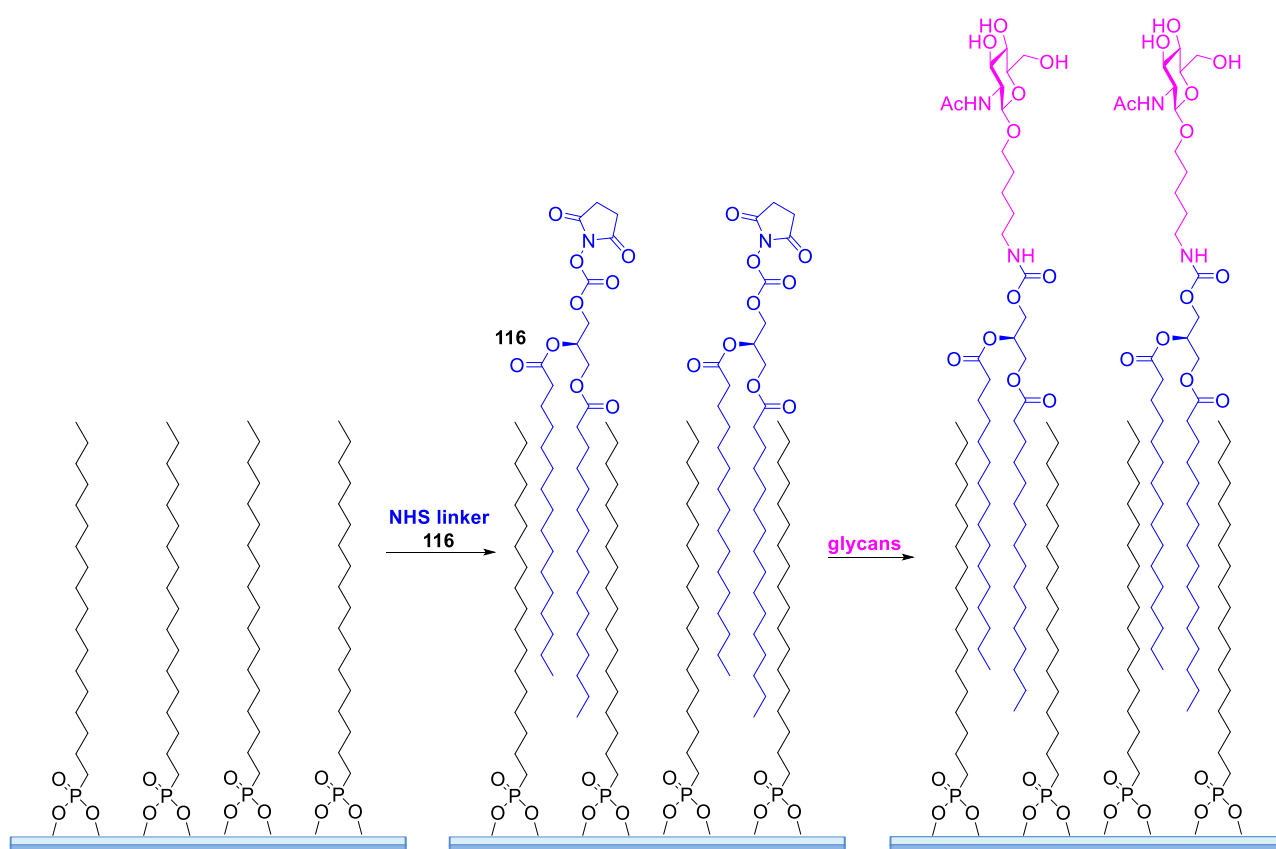


Scheme 3.4 – On-chip glycosylation and CuAAC reactions. The methodology we designed envisages the immobilization of a small library of glycans **105-111** that will be expanded directly on the array. We aimed to introduce an azide tag by glycosylation reaction using the labelled UDP-GalNAz **112** and UDP-6-azido-D-galactose **114**, respectively mimic structure of the natural **113** and **115**. The azide functionalities will allow diversification through CuAAC with a small library of alkynes **104a-i**.

3.4.1.1 Glycan printing

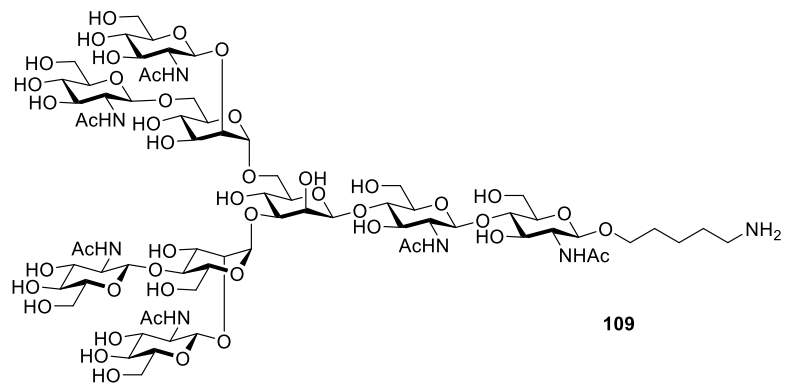
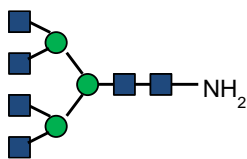
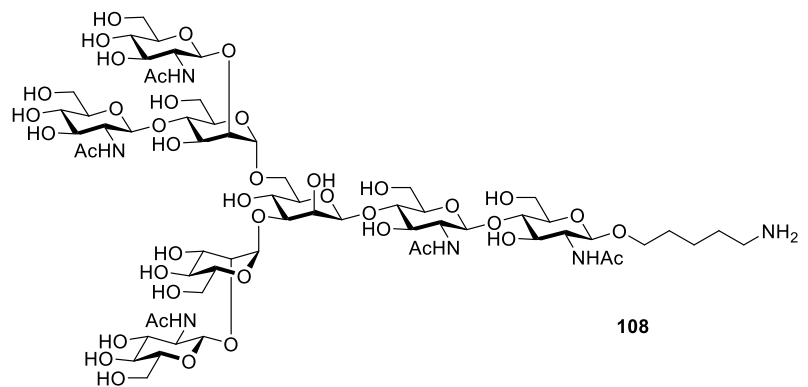
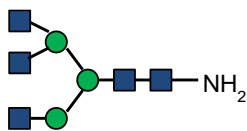
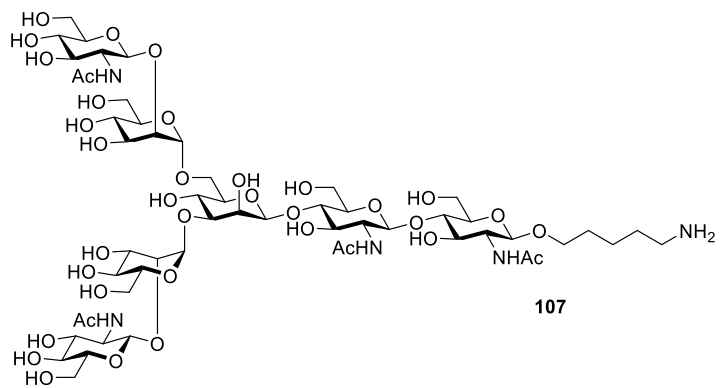
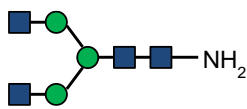
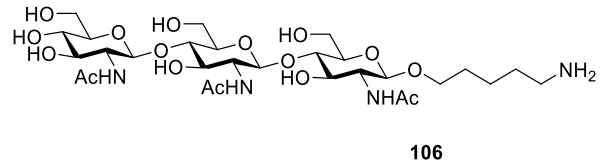
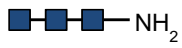
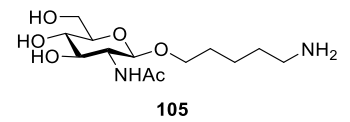
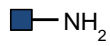
In the construction of glycan microarrays a careful selection of the nature of the surface which has to match with the immobilization method is essential. For our purposes, we chose commercially available glass slides coated with an indium tin oxide (ITO) layer as a transparent and conductive platform, which ensures both fluorescence and MALDI-TOF MS detection.

In order to allow the hydrophobic immobilization of glycans, ITO slides had to be properly functionalized. Therefore, ITO slides were initially treated with a basic piranha solution ($\text{H}_2\text{O}:\text{NH}_3:\text{H}_2\text{O}_2 = 6:1:1$), which allowed to maximize the hydroxyl groups on the surface and to remove eventual adsorbed organic traces. Incubation with a solution of octadecyl phosphonic acid in THF (3 h, r.t.) and finally curing the slides at 140 °C overnight allowed to get a self-assembled hydrophobic monolayer. The use of phosphonic acids is known as a preferential method for the generation of self-assembled monolayers on ITO slides.⁵³ Then, a CHCl_3 solution of the *N*-hydroxy succinimide (NHS) activated glycerol derived linker **116**, which was previously synthesized in the Reichardt group at CIC biomaGUNE,⁵⁰ was incubated with the hydrophobic slides, leading to a reactive lipidic double layer where glycans could finally be immobilized (Scheme 3.5). In particular, glycans were conveniently functionalized with a C-5 amino linker and robotically printed as solutions (50 μM in phosphate buffer pH 8.5) on the slides where they got bound as carbamates upon reaction with the activated NHS carbonates. Finally, remaining NHS reactive groups were quenched treating the slides with an ethanolamine solution (50 mM in borate buffer pH 9.0).



Scheme 3.5 – Immobilization of the glycans. ITO slides were conveniently functionalized with octadecyl phosphonate to get a hydrophobic surface, which was incubated with a CHCl_3 solution of NHS activated linker **116**. Amino tethered glycans were then printed over the reactive lipidic double layer getting immobilized through the formation of carbamate bonds.

As a first attempt, *N*-glycans presenting a C5 amino linker and characterised by different degrees of complexity were selected and printed on the ITO slides (Fig. 3.6). All glycans **105-109** display a terminal GlcNAc residue, except for the pentasaccharide **110** and the Lewis^X trisaccharide **111**. (Once immobilized on the slide, we will refer to glycans **105-111** as **105i-111i**).



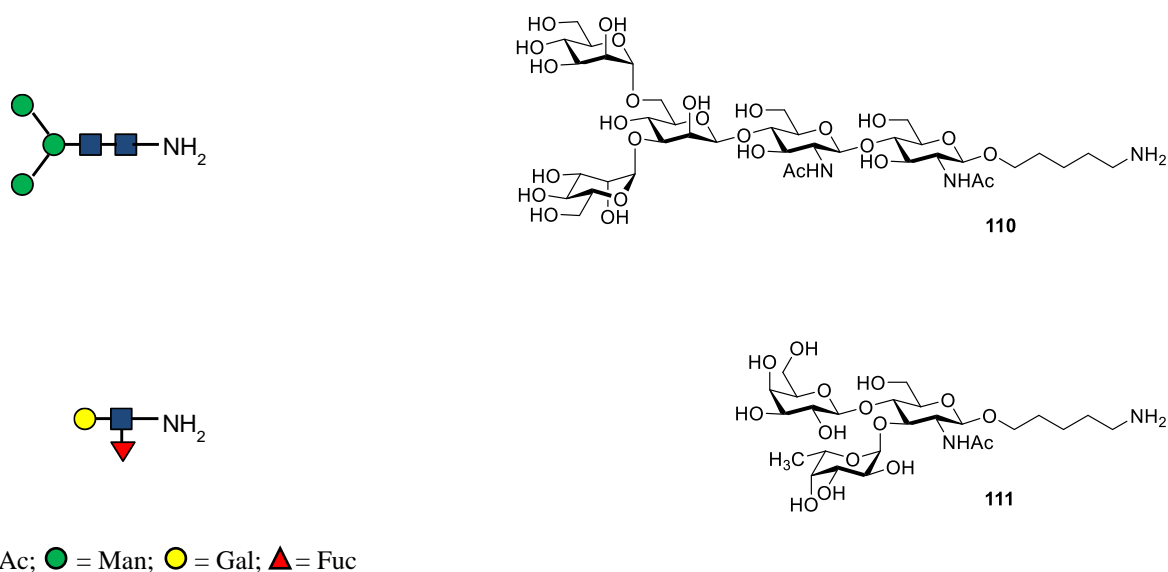


Figure 3.6 – Amino tethered glycans printed on the array. Glycans **105-111** were selected to be printed on the array. The extended chemical structures are reported along with their schematic representation.

3.4.1.2 Enzyme activity assay: clickable NDP sugars as glycoside donors

To establish the methodology, glycosylation reaction with the azido labelled UDP sugars **112** and **114** as glycoside donors were first studied in solution with two related enzymes. The ability of the enzymes in transferring the unnatural sugar nucleotides **112** and **114** was evaluated and compared with glycosylation reactions using the respective naturally occurring UDP-*N*-acetyl-D-galactosamine **113** and UDP-D-galactose **115** (Scheme 3.4).

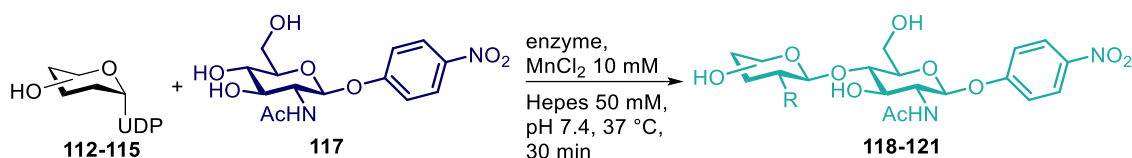
The glycosylation reactions were performed using β -1,4-galactosyl transferase I (GalT-1) enzyme, which selectively transfers UDP activated galactosyl donors to *N*-acetyl glucosamine residues forming a β -1 \rightarrow 4 glycosidic bond. In particular, two versions of the GalT-1 enzyme were exploited, the native enzyme and the double mutant C342T&Y289L. While the C342T mutation allows for better and easier folding during the protein expression and purification process,⁵⁴ the additional Y289L mutation allows modulating enzyme selectivity. Indeed, the native GalT-1 enzyme can use UDP-Gal as a donor; it shows only a minimal activity (0.1%) in transferring GalNAc from the related UDP-GalNAc nucleoside. On the contrary, it was demonstrated that GalNAc residues are successfully accepted by the GalT-1 Y289L mutant, in which the Tyr289 at the glycoside donor binding site is substituted with the non coordinating Leu289.⁵⁵ As already reported, the mimic structure of the 6-azido-UDP-Gal **114** is a substrate of the truncated bovine galactosyl transferase GalT-1,⁵² while the commercially available GalNAz **112** can be transferred using the double mutant version of the protein, the GalT-1-DM.

Both GalT-1 and GalT-1-DM were overexpressed in transformed *E. coli*⁵⁶ and recovered after cell lysis from inclusion bodies. The enzymes were folded and purified by His-Trap column eluting with a 500 mM imidazole Tris buffer solution pH 8.5. The fractions were analysed by SDS-page and dialysed affording GalT-1 (0.50 mg/mL) and GalT-1-DM (0.62 mg/mL) solutions.

In order to evaluate the activities of the expressed enzymes towards the natural glycoside donors **113**, **115** and the respective azido derivatives **112** and **114**, glycosylation reaction were performed in solution using the UV visible 1-*p*-nitrophenyl-*N*-acetyl glucosamine **117** as acceptor (Scheme 2). The reactions were generally performed mixing an equimolar ratio of the glycoside donor and acceptor with the enzyme solution (3.3 μ L) in HEPES buffer (50 mM, pH 7.4) containing MnCl₂ and shaking at 37 °C. The reactions were stopped after 30 minutes quenching with acetonitrile and analytical samples were analysed by UPLC-MS detecting at 298 nm. From the resulting chromatograms (Fig. 3.7), the conversions were estimated by area integration of the

absorbance peaks of the acceptor **117** (dark blue peak) and of the disaccharide products **118-121** (light blue peaks). The corresponding enzyme units for the solutions of GalT-1 and GalT-1-DM were calculated (1U = the amount of enzyme that transforms 1 μmol of a given substrate in 1 min).

The GalT-1-DM enzyme was able to transfer both the native UDP-N-acetyl-D-galactosamine **113** and the unnatural counterpart **112** (26% and 46% conversion respectively in 30 minutes) (Fig. 3.7a,b); on the other hand, the GalT-1 enzyme was found to be active only with UDP-galactose **115** as donor (46% conversion in 30 minutes) (Fig. 3.7c), showing to be completely unable to accept **114** as a substrate. Attempts to force the reaction using a higher amount of the enzyme (0.76 mU taking reaction of UDP-D-galactose **115** with GalT-1 as a reference) and an excess of donor (2.5 and 5.0 eq) failed. Only using 13 μL = 1.75 mU of enzyme at room temperature for 14 h a low conversion (2%) was observed (Fig. 3.7d). In a further trial, alkaline phosphatase was added to favour the reaction thermodynamics⁵⁷ but was found not to affect the reaction outcome. Analogous results were obtained in trials employing a commercially available GalT-1 supplied by Sigma-Aldrich, thus suggesting a possible decomposition of the activated UDP-6-azido-D-galactose **115**. However, since ¹H and ³¹P NMR and ESI-MS spectra of this compound were all in agreement with those reported in literature,⁵² it is reasonable to argue that such a low conversion is likely to be ascribed to the low affinity of GalT-1 enzyme for this substrate. As a consequence, on-chip glycosylation reactions with UDP-6-azido-D-galactose **115** as a donor seem not feasible and we focused our attention on the use of UDP-GalNAz **112**.



Scheme 3.6 – Enzyme activity assay. The activity of expressed enzymes GalT-1 and GalT-1-DM were evaluated by in solution glycosylation reactions using both natural glycoside donors **113** and **115**, as well as the azido derivatives **112** and **114**. The *p*-nitrophenyl GlcNAc **117** was employed as glycoside acceptor.

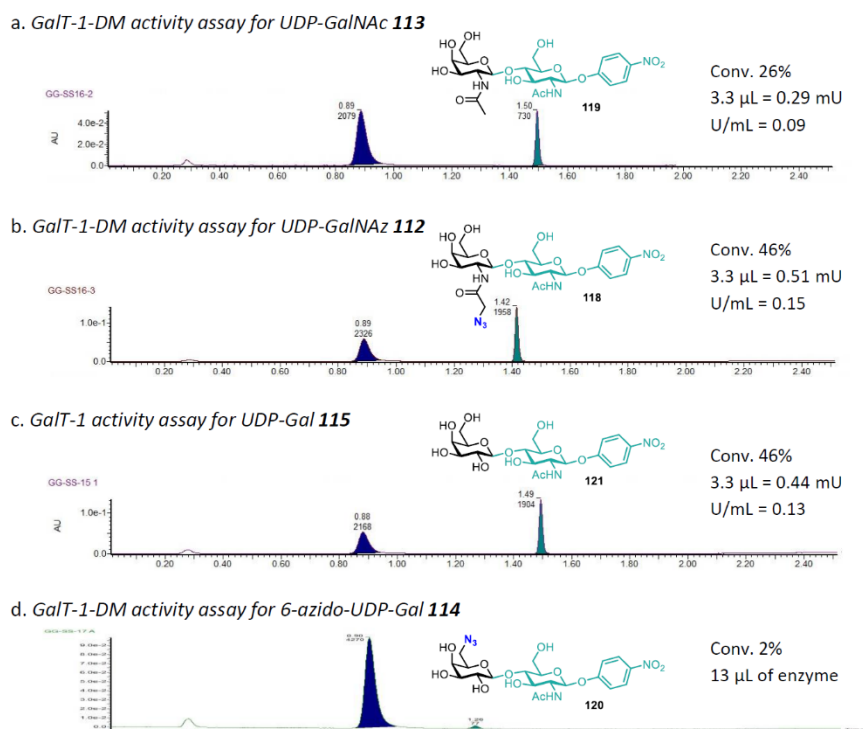


Figure 3.7 – UPLC-MS chromatograms for enzymatic glycosylation reactions. Reaction samples were analysed by UPLC-MS. The *p*-nitrophenyl labelled acceptor **117** and the disaccharide products were both UV-visible monitoring at 289 nm and conversions were estimated upon area integration of absorbance peaks. PNP-GlcNAc **117** $t_R = 0.89$ min. **a**) GalNAc- β -1,4-PNP-GlcNAc **119** $t_R = 1.50$ min. **b**) GalNAz- β -1,4-PNP-GlcNAc **118** $t_R = 1.42$ min. **c**) Gal- β -1,4-PNP-GlcNAc **121** $t_R = 1.49$ min. **d**) 6-azido-Gal- β -1,4-PNP-GlcNAc **120** $t_R = 1.26$ min. From chromatograms **a,b,c** the enzyme units for the corresponding reaction were calculated.

3.4.1.3 On-chip glycosylation

The enzymatic glycosylation of arrayed glycans was investigated using GalT-1-DM enzyme in order to transfer either the natural substrate UDP-*N*-acetyl galactosamine **113** or the labelled UDP-GalNAz **112**. The reactions were performed in similar conditions as previously established in solution. The ITO slides were incubated overnight at 37 °C with a 500 μ L solution of both the enzyme and the glycoside donor (glycoside donor 1 mM, MnCl₂ 5 mM, Hepes 50 mM, GalT-1-DM enzyme 125 μ L of a 0.62 mg/mL solution, BSA IgG free 2 μ L of a saturated solution). The extent of the reactions was evaluated by robotically printing a DHB matrix solution onto the slides and analysing by MALDI-TOF spectrometry. Almost complete conversion was assessed for the monosaccharide **105i** with both the GalNAc donor **113**, yielding the GalNAc- β -1,4-GlcNAc disaccharide **122i** (Fig. 3.8b; 1126.996 [M+Na]⁺, 1142.984 [M+K]⁺), and the azido derivative **112**, yielding the corresponding azido labelled disaccharide **123i** (Fig. 3.8c 1168.079 [M+Na]⁺, 1184.059 [M+K]⁺). Notably, the MS peaks for the azido labelled disaccharide **123i** were always associated with signals at lower molecular weight (-28) due to decomposition of the azido group upon ionization (Fig. 3.8c 1142.070 [M-N₂+H₂+Na]⁺, 1158.046 [M-N₂+H₂+K]⁺). Complete conversion was obtained also for glycosylation of the linear trisaccharide **106i** with UDP-*N*-acetylgalactosamine **113**, leading to the tetrasaccharide **124i** (Fig. 3.9b; 1533.198 [M+Na]⁺, 1549.201 [M+K]⁺). On the contrary, a decreased activity was shown by GalT-1-DM in the glycosylation of **106i** with UDP-GalNAz **112** as the donor and only partial conversion was achieved (Fig. 3.9c; 1548.317 [M-N₂+H₂+Na]⁺, 1562.289 [M-N₂+H₂+K]⁺, 1574.320 [M+Na]⁺, 1590.319 [M+K]⁺).

The glycosylation of multiantennary glycans was less satisfactory: elongation of **109-111** with UDP-GalNAc **113** always afforded the desired products along with partially galactosylated intermediates, likely in regioisomeric mixtures (respectively Fig.3.10b, 3.11b, 3.12b). On the contrary, using UDP-GalNAz **3.9** as glycoside donor no signals ascribed either to the starting material or to the products and the corresponding intermediates could be detected (respectively Fig.3.10c, 3.11c, 3.12c). As expected, no glycosylation product was observed in reactions involving the pentasaccharide **110i** and the LewisX **111i**, which were included in the screening as negative controls.

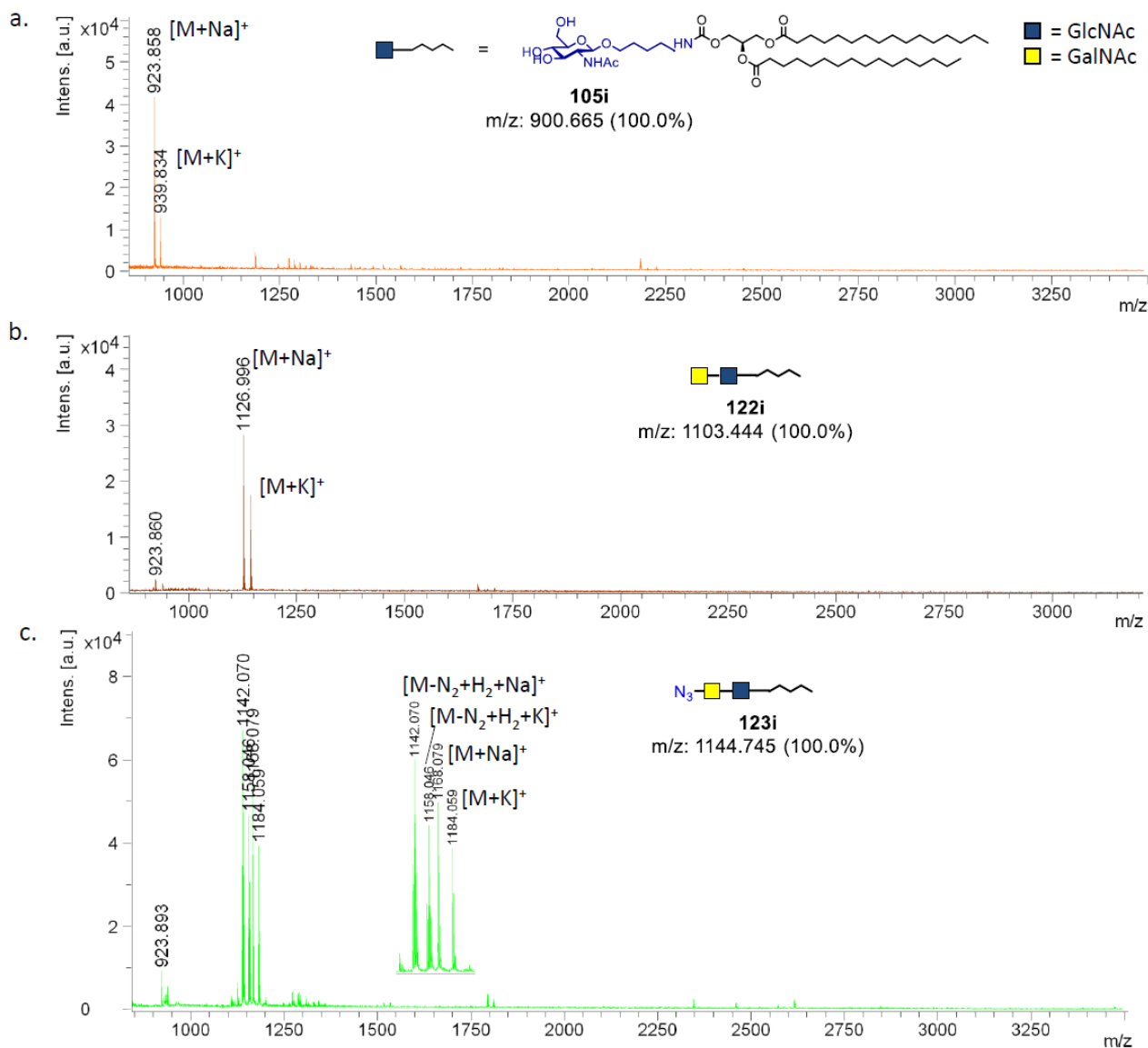


Figure 3.8 – MALDI-TOF MS spectra of on-chip enzymatic elongations of monosaccharide 105i. a) MS spectrum of immobilized glucosamine derivative **105i**: m/z calcd 900.665; found 923.858 [M+Na]⁺, 939.834 [M+K]⁺. b) MS spectrum acquired after enzymatic glycosylation of **105i** with UDP-GalNAc **113**. Almost complete conversion to disaccharide **122i** was observed: m/z calcd 1103.444; found 1126.996 [M+Na]⁺, 1141.866 [M+K]⁺. c) MS spectrum acquired after enzymatic glycosylation of **105i** with UDP-GalNAz **112**. Almost complete conversion to azido tagged disaccharide **123i** was observed: m/z calcd 1144.745; found 1168.079 [M+Na]⁺, 1184.059 [M+K]⁺. Peaks at lower molecular weight (-28) deriving from decomposition of azide group due to ionization were also detected: m/z 1142.070 [M-N₂+H₂+Na]⁺, 1158.046 [M-N₂+H₂+K]⁺.

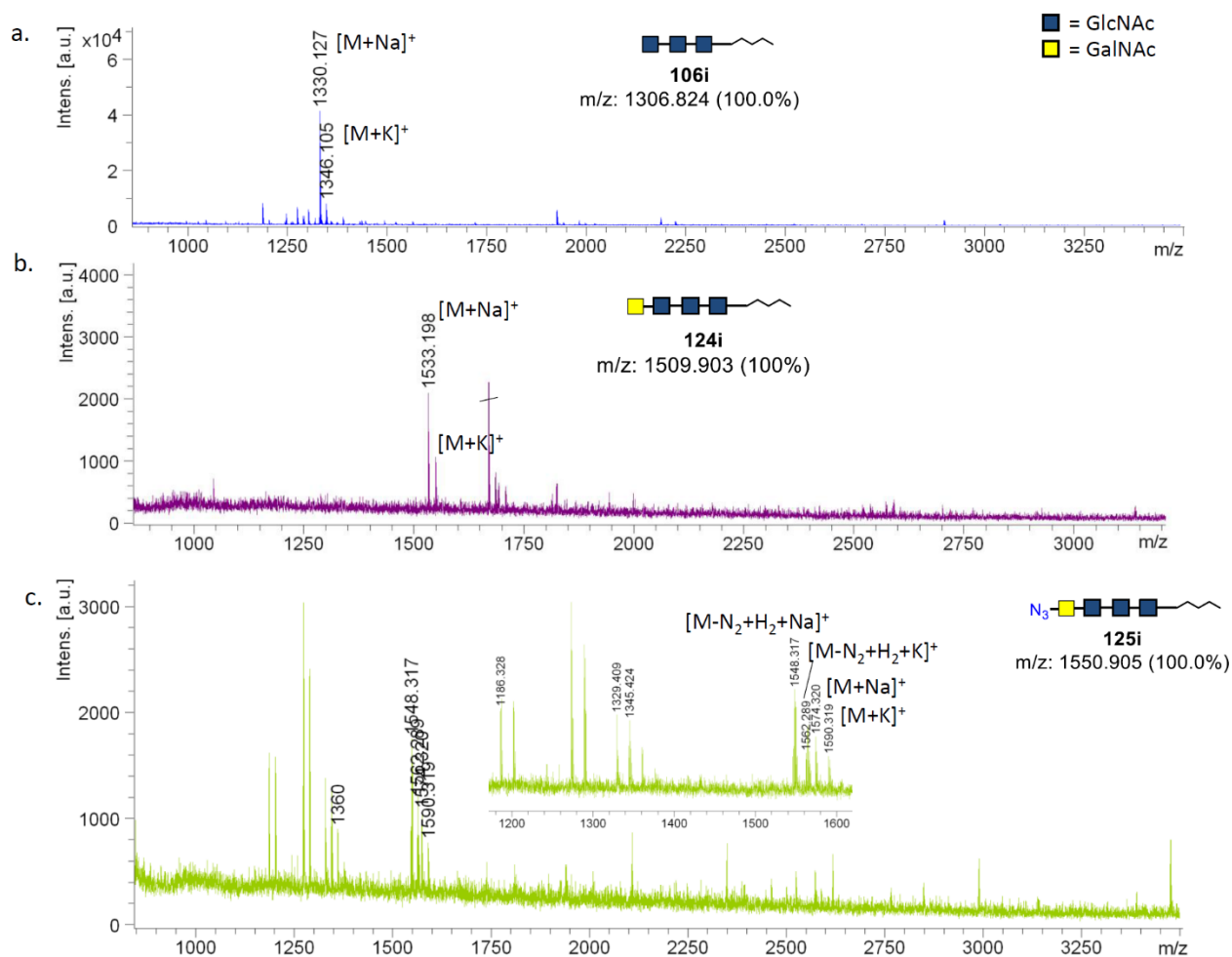


Figure 3.9 - MALDI-TOF MS spectra of on-chip enzymatic elongations of trisaccharide 106i. a) MS spectrum of immobilized trisaccharide **106i**: m/z calcd 1306.824; found 1330.127 [$M+Na$]⁺, 1346.105 [$M+K$]⁺. b) MS spectrum acquired after enzymatic glycosylation of **106i** with UDP-GalNAc **113**. Complete conversion to tetrasaccharide **124i** was observed: m/z calcd 1509.903; found 1533.198 [$M+Na$]⁺, 1549.021 [$M+K$]⁺. c) MS spectrum acquired after enzymatic glycosylation of **106i** with UDP-GalNAz **112**. Partial conversion to azido labelled tetrasaccharide **125i** was observed: m/z calcd 1550.905; found 1574.320 [$M+Na$]⁺, 1590.319 [$M+K$]⁺. Peaks at lower molecular weight (-28) deriving from decomposition of azide group due to ionization were also detected: m/z 1548.317 [$M-N_2+H_2+Na$]⁺, 1562.289 [$M-N_2+H_2+K$]⁺.

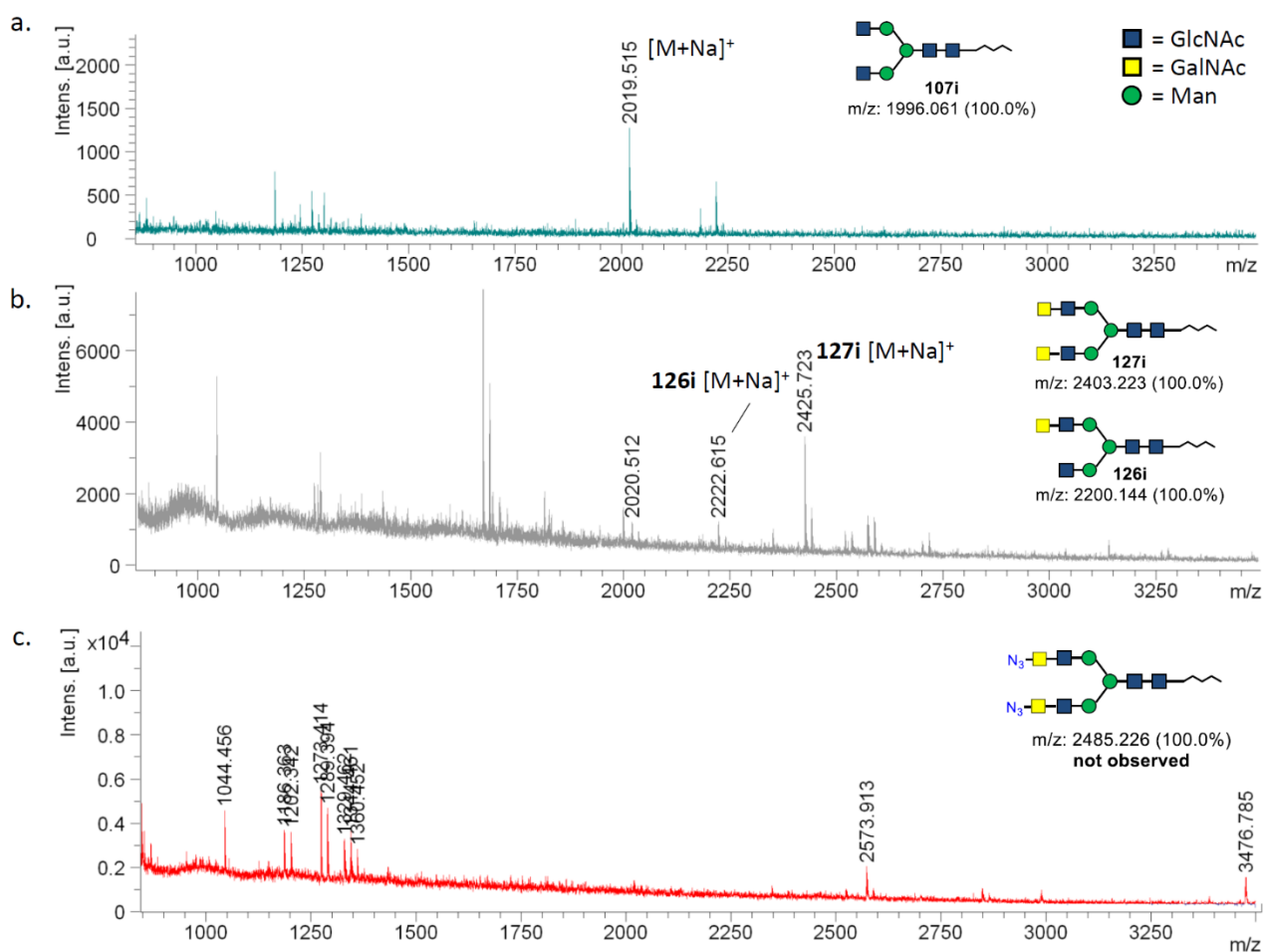


Figure 3.10 - MALDI-TOF MS spectra of on-chip enzymatic elongations of biantennary **107i.** a) MS spectrum of immobilized biantennary **107i**: m/z calcd 1996.061; found 2019.515 $[M+Na]^+$. b) MS spectrum acquired after enzymatic glycosylation of **107i** with UDP-GalNAc **113**. Almost complete conversion to di-GalNAc containing glycan **127i** was achieved: m/z calcd 2403.223; found 2425.723 $[M+Na]^+$. Partially glycosylated **126i** was also observed at minor extent: m/z calcd 2200.144; found 2222.615 $[M+Na]^+$. c) Enzymatic glycosylation of **107i** with UDP-GalNAz **112** was unsuccessful and no peaks ascribed to the product were detected.

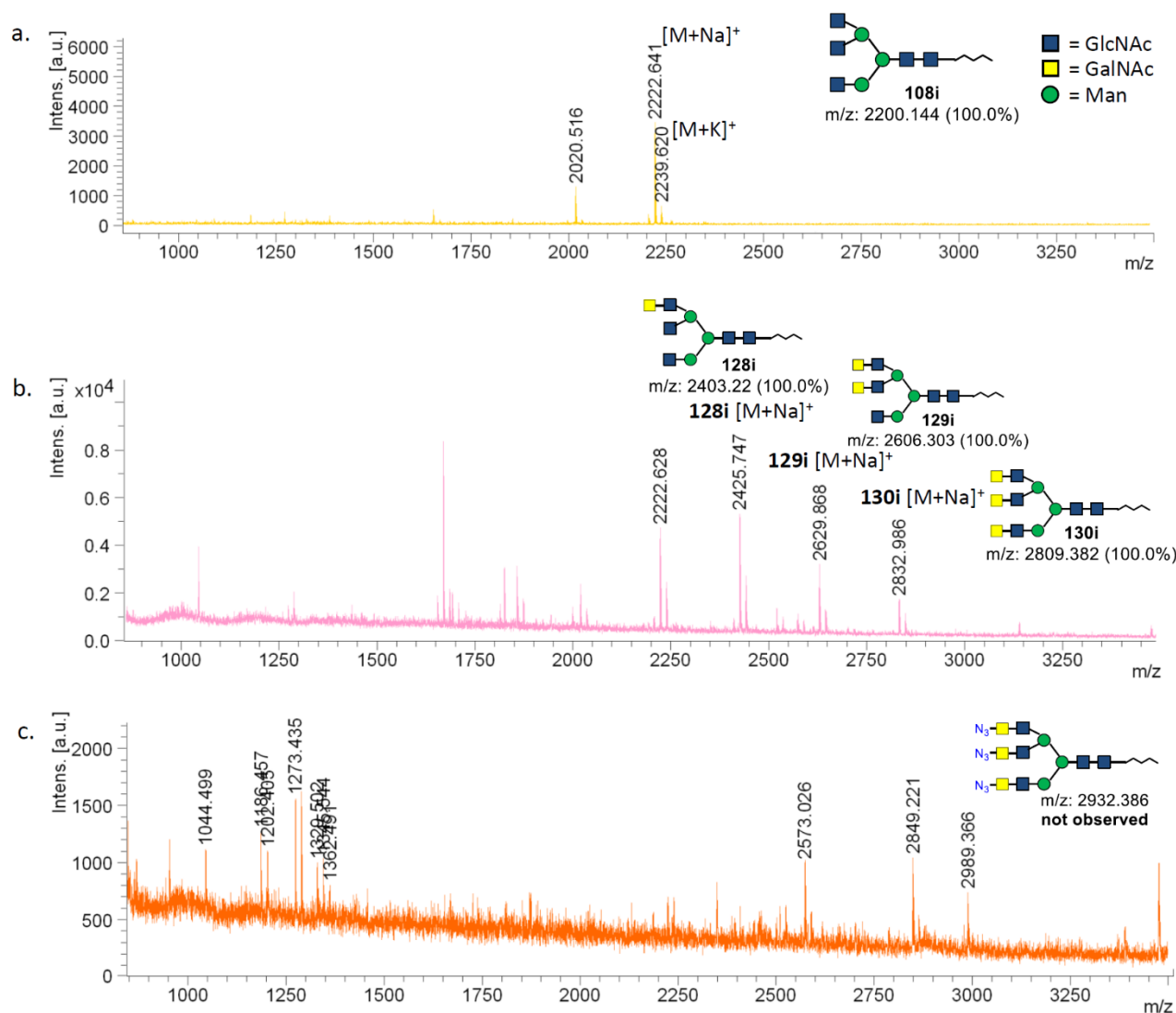


Figure 3.11 - MALDI-TOF MS spectra of on-chip enzymatic elongations of triantennary **108i.** a) MS spectrum of immobilized triantennary **108i**: m/z calcd 2200.144; found 2222.641 [M+Na]⁺, 2239.620 [M+K]⁺. b) MS spectrum acquired after enzymatic glycosylation of **108i** with UDP-GalNAc **113**. Incomplete conversion to give the triglycosylated product **130i** was observed: m/z calcd 2809.382; found 2832.986 [M+Na]⁺. Partially glycosylated intermediates **128i** and **129i** were detected. Diglycosylated intermediate **129i**: m/z calcd 2606.303; found 2629.868 [M+Na]⁺. Monoglycosylated intermediate **128i**: m/z calcd 2403.22; found 2425.747 [M+Na]⁺. c) Enzymatic glycosylation of **108i** with UDP-GalNAz **112** was unsuccessful and no peaks ascribed to the product were detected.

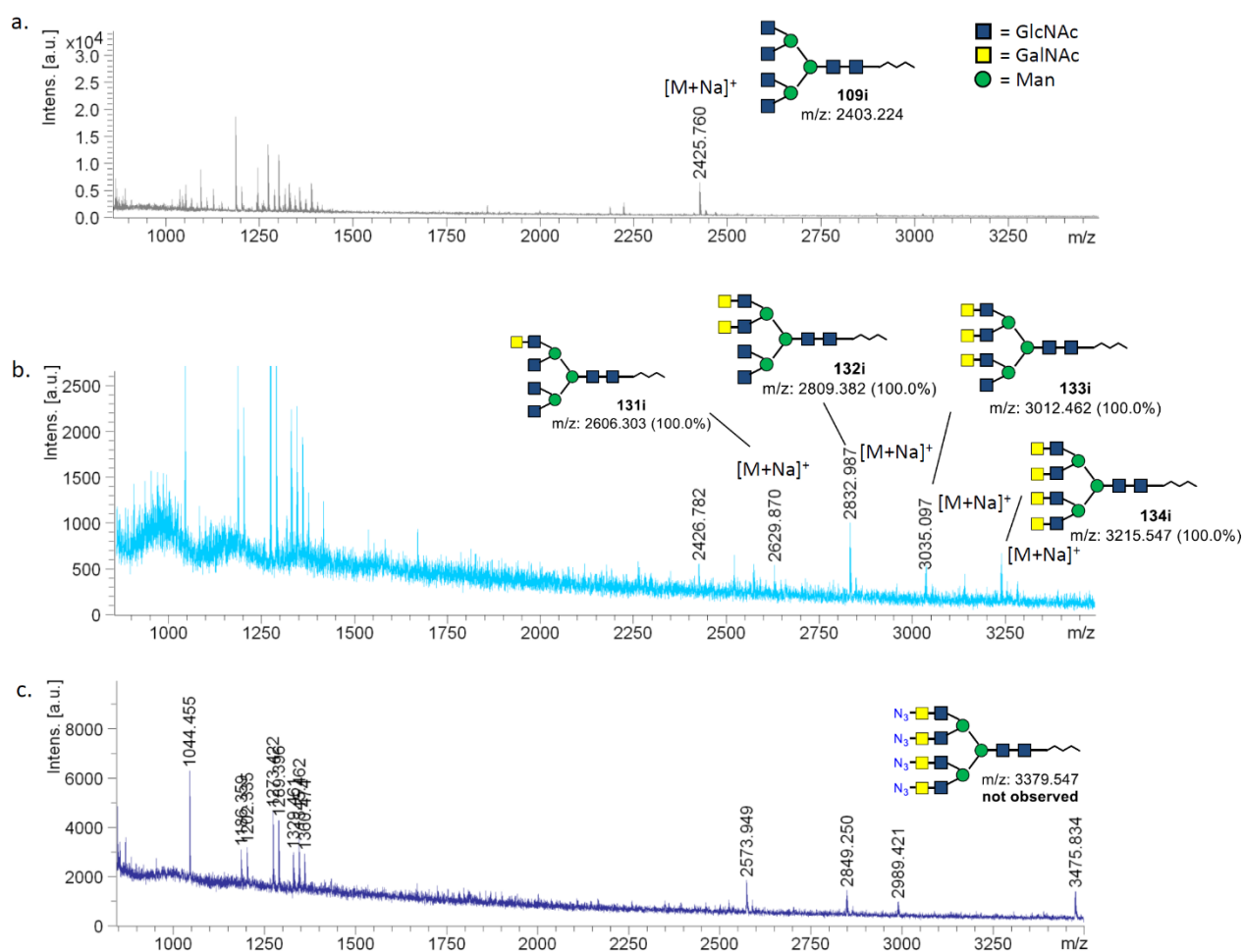
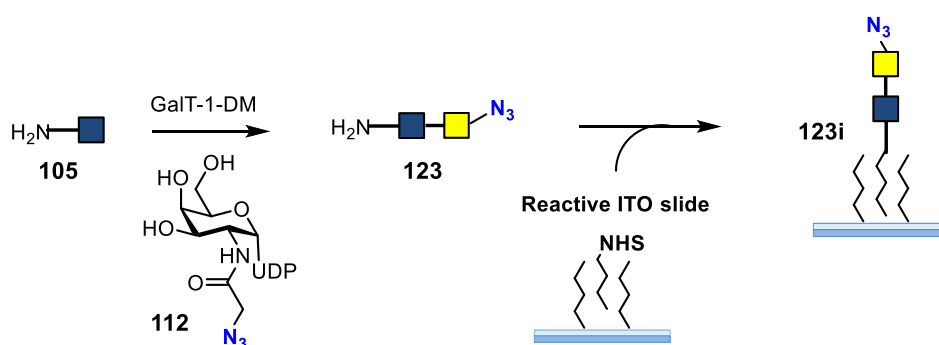


Figure 3.12 - MALDI-TOF MS spectra of on-chip enzymatic elongations of tetraantennary **109i.** a) MS spectrum of immobilized tetraantennary **109i**: m/z calcd 2403.224; found 2425.760 [M+Na]⁺. b) MS spectrum acquired after enzymatic glycosylation of **109i** with UDP-GalNAc **113**. Incomplete conversion to give the tetraglycosylated product **134i** was observed: m/z calcd 3215.547; found 3243.898 [M+Na]⁺. Partially glycosylated intermediates **131i**, **132i** and **133i** were also detected. Triglycosylated intermediate **133i**: m/z calcd 3012.462; found 3035.097 [M+Na]⁺. Diglycosylated intermediate **132i**: m/z calcd 2809.382; found 2832.987 [M+Na]⁺. Monoglycosylated intermediate **131i**: m/z calcd 2606.303; found 2629.870 [M+Na]⁺. c) Enzymatic glycosylation of **109i** with UDP-GalNAz **112** was unsuccessful and no peak ascribed to the product was detected.

3.4.1.4 Library diversification: CuAAC of arrayed glycans

The introduction of the azide tag into the arrayed glycan structure allowed for the on-chip library diversification by CuAAC reaction. Since the on-chip enzymatic glycosylation of multiantennary glycans with tagged unnatural glycosides could not be optimized, monoantennary **123i** and **125i** were selected for setting up the general conditions for the CuAAC reaction on the array. Moreover, when printed on ITO slides, these glycans showed a good S/N ratio analysing by MALDI-TOF spectrometry.

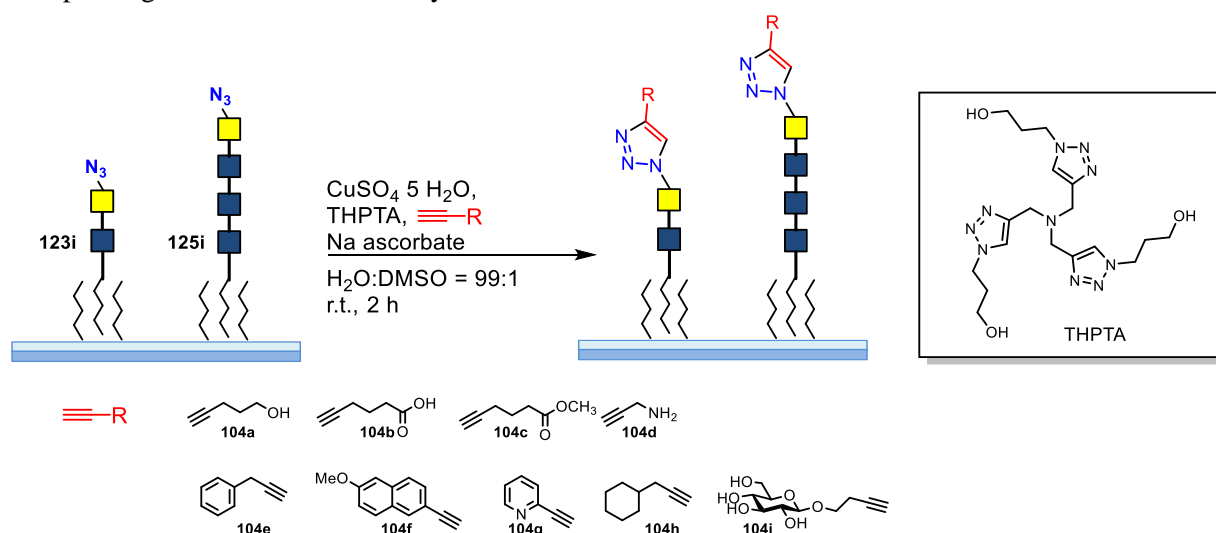
Generally, experiments were conducted with the azido **123i** and **125i** obtained by on-chip glycosylation, however to limit the use of the synthetically precious UDP-GalNAz **112**, in solution glycosylation was also exploited. In particular, in solution enzymatic glycosylation of the GlcNAc derivative **105** using GalT-1-DM enzyme and UDP-GalNAz **112** afforded the corresponding amino tethered disaccharide **123** that was readily purified by a graphitized carbon cartridge (eluting with H₂O - CH₃CN, 6:4) and printed over the ITO slides leading to the immobilized **123i** (Scheme 3.7).



Scheme 3.7 – In solution preparation of azido tagged glycans. Following an alternative strategy, the immobilized glycan **123i** could be prepared by in solution glycosylation of **105** with UDP-GalNAz **112** followed by printing on NHS activated ITO slides.

The monoantennary **123i** and **125i** were subjected to cycloaddition reactions with variously functionalized terminal alkynes (Scheme 3.8). All the selected alkynes were commercially available, except for the methyl hexynoate **104c** which was readily obtained by esterification of the corresponding acid **104b**⁵⁸ and the homopropargyl glucose **104i** that was obtained by glycosylation reaction 3-butyn-1-ol with β -D-glucose.

Importantly, to prevent disruption of the hydrophobic interactions between the arrayed glycans and the hydrophobic ITO slides, all reactions on the array must be performed in water solutions and reaction conditions have to be carefully tuned. For this reason, cycloaddition of glycans **123i** and **125i** were generally performed in a 99:1 $\text{H}_2\text{O}/\text{DMSO}$ mixture as a good compromise between low content of DMSO and alkyne solubility. Cu(I) catalysed azide alkyne cycloadditions were performed following the methodology reported by Sharpless and co-workers,^{59,60} thus ITO slides with the printed azido tagged glycans were incubated with reaction mixtures containing terminal alkyne, $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$, Na ascorbate and the THPTA (tris(3-hydroxypropyltriazolylmethyl)amine) ligand (5 mol. eq. relative to the Cu salt). The THPTA ligand⁶¹ is a hydrophilic version of the more common TBTA (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine) ligand, which is employed as an additive in CuAAC reaction to stabilize the copper complex in its catalytically active Cu(I) oxidation state. Reactions were conducted for 2 hours, shielding from light and finally analysed by DHB matrix printing and MALDI-TOF analysis.

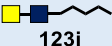
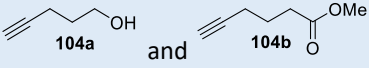

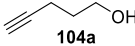
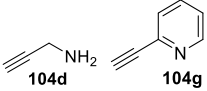


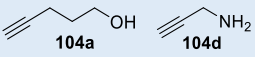


Scheme 3.8 - Cu catalyzed cycloaddition of 123i and 125i with terminal alkynes 104a-i. The reactions were performed in a 99:1 $\text{H}_2\text{O}/\text{DMSO}$ mixture to ensure stability of the array and enough solubility of the alkynes. The water soluble Cu(I) stabilizing ligand THPTA was involved.

In a first attempt, glycan **123i** was reacted with pentynol **104a** and hexynoic acid **104b**. The slides were incubated with 100 μL of reactive solutions with a 4 mM concentration of the alkynes, 2.5 mM of Na ascorbate and 100 μM CuSO_4 , using a 99:1 $\text{H}_2\text{O}/\text{DMSO}$ mixture as solvent (Table 3.1, entry 1). No signals of the products could be detected by MALDI-TOF MS. However no signal of unreacted disaccharide **123i** was observed either, indicating a probable disruption of the hydrophobic bilayer upon reaction.

More encouraging results were obtained using reactive solutions with a lower concentration (2 orders of magnitude) of the catalytic system partners (CuSO_4 , THPTA, sodium ascorbate) and higher concentrations (14 fold) of terminal alkyne employed. (Table 3.1, entry 2). Almost complete conversion was achieved in the cycloaddition reaction of azido labelled **123i** with 4-pentyn-1-ol **104a** using a 99:1 $\text{H}_2\text{O}/\text{DMSO}$ mixture as the solvent. Increasing the content of DMSO was detrimental for both the reaction outcome and the array stability; indeed partial conversion and slightly decreased S/N ratio were observed with a 95:5 $\text{H}_2\text{O}/\text{DMSO}$ mixture (not shown in Table 3.1).

The conditions of (Table 3.1, entry 2) were extended to the cycloaddition of **123i** with the selected set of alkynes obtaining partial conversions using propargylamine **104d** and 2-ethynyl pyridine **104g**. Similar results were obtained for cycloaddition of azido tetrasaccharide **125i** (Table 3.1, entry 3). Low conversions were achieved with pentynol **104a** and propargylamine **104d**, while complete conversion was obtained with 2-ethynyl pyridine **104g**. Decreasing the alkyne concentration to a comparable range with that of the catalyst and the additives led to a reduced reactivity and cycloaddition of azido disaccharide **123i** was achieved only with propargylamine **104d** and 2-ethynylpyridine **104g**, respectively with low and partial conversion (Table 3.1, entry 4). Performing the reaction in a more oxygen free environment to stabilize the catalyst in its active Cu(I) oxidation state was attempted using degassed H_2O (bubbled with N_2), which however did not affect the reaction outcome. Finally, homogeneously increasing the concentrations by one order of magnitude as reported in (Table 3.1, entry 5) allowed reaching a better balance between array stability and reactivity. Indeed, CuAAC reaction of azido **123i** with the terminal alkynes **104a-i** was achieved with complete conversion using propargylamine **104d** and almost complete conversions were observed with pentynol **104a**, methyl hexynoate **104c**, the methoxynaphthalene derivative **104f** and 2-ethynylpyridine **104g**.

Entry	Substrate	CuSO_4	THPTA	Alkyne	Na Ascorbate	Conversion
1 ^a		100 μM	500 μM	4 mM	2.5 mM	Not observed with 
2		1.15 μM	5.75 μM	57.5 mM	11.5 μM	Almost complete:  Partial: 
3		1.15 μM	5.75 μM	57.5 mM	11.5 μM	Complete:  Low: 


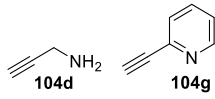

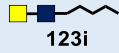
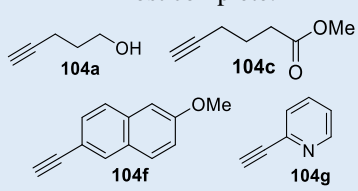
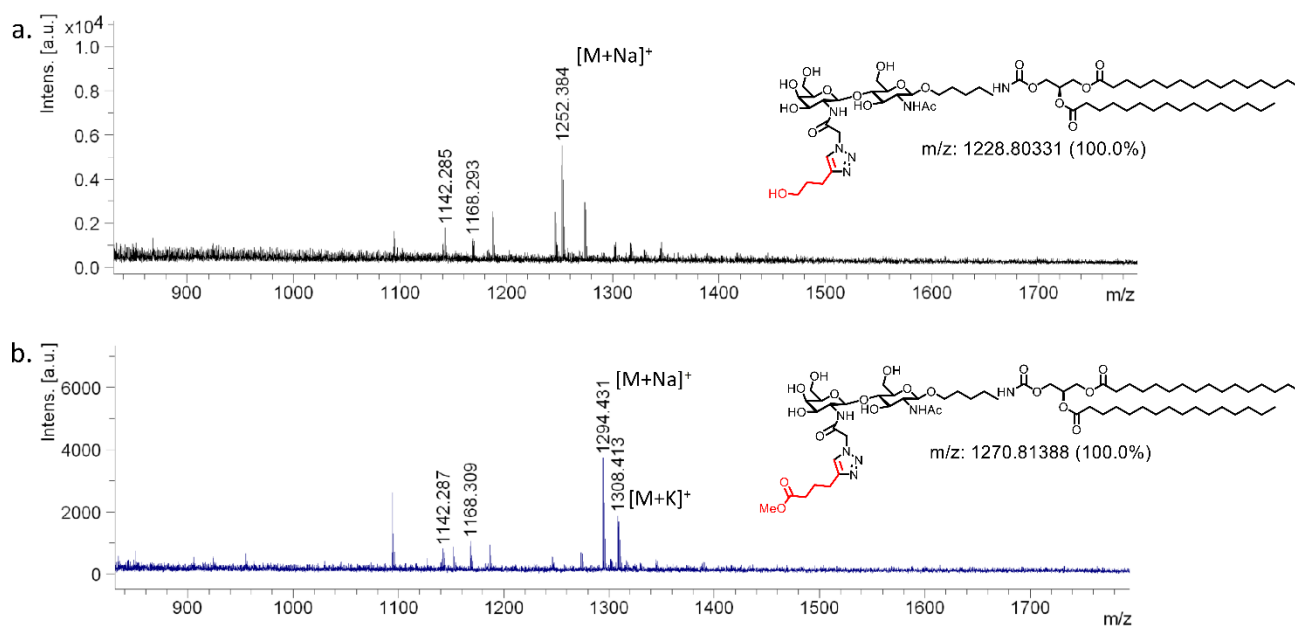
4		1.15 μM	5.75 μM	57.5 μM	11.5 μM	Partial:  Complete: 
5		11.5 μM	57.5 μM	575 μM	115 μM	Almost complete: 

Table 3.1 – CuAAC reaction conditions. Cycloaddition reactions were performed incubating arrayed glycans with reactive solutions (100 μL) containing different alkynes in 99:1 $\text{H}_2\text{O}/\text{DMSO}$ mixture. The concentrations of the reactants are reported. ^a Reaction performed using either pentynol **104a** or hexynoic acid **104b**.



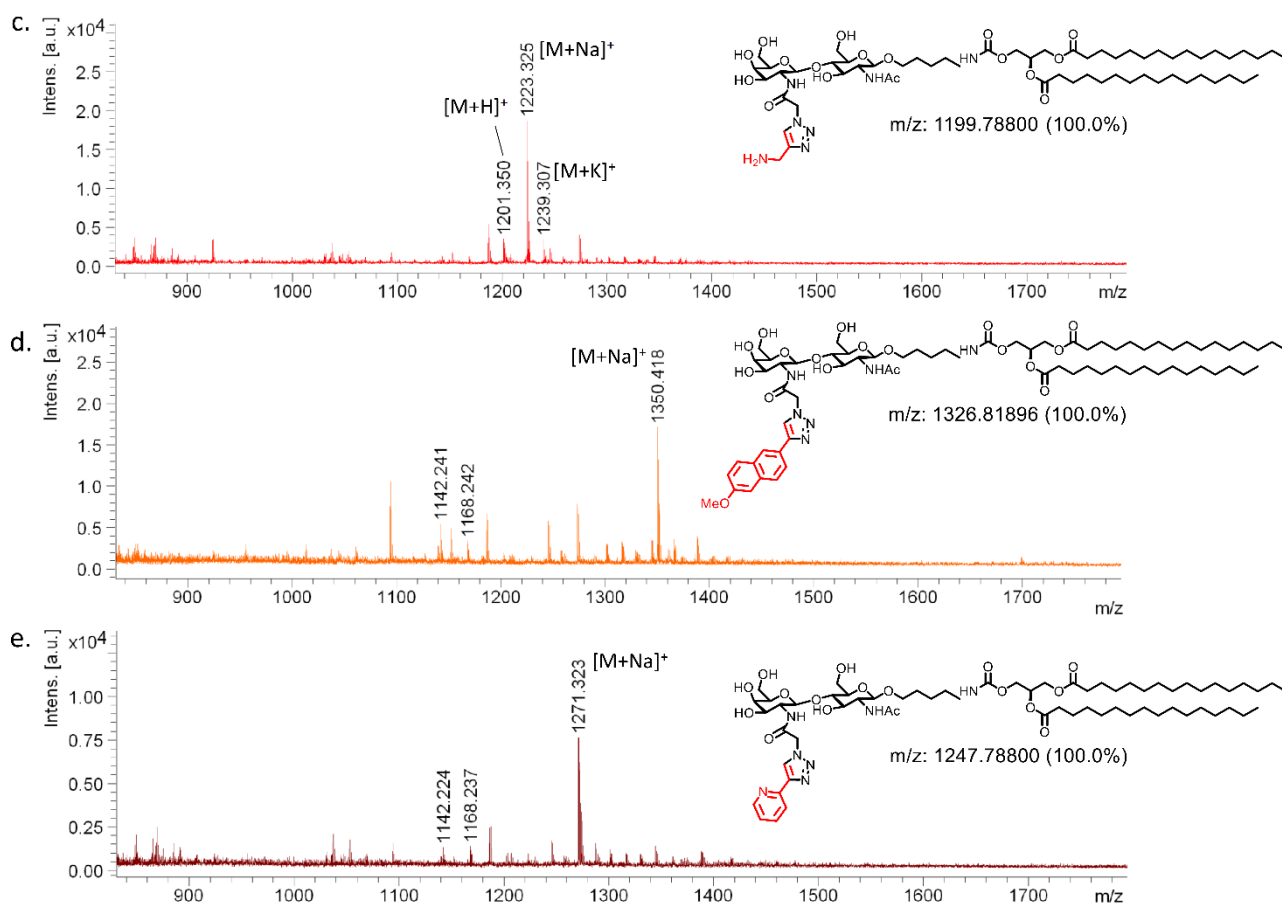


Figure 3.13 - MALDI-TOF MS spectra of on-chip CuAAC reactions. The spectra of glycan **123i** successfully functionalized by CuAAC reaction with terminal alkynes are reported. Except for reaction with propargylamine **104d**, traces of unreacted glycan were observed ($m/z \sim 1142$ [M+Na]⁺, 1168 [M+K]⁺). **a)** Almost complete conversion was achieved with pentynol **104a**: m/z calcd 1228.803; found 1252.384 [M+Na]⁺. **b)** Almost complete conversion was obtained with methylhexynoate **104c**: m/z calcd 1270.814; found 1294.431 [M+Na]⁺, 1308.413 [M+K]⁺. **c)** Complete conversion was assessed for propargylamine **104d**: m/z calcd 1199.788; found 1201.350 [M+H]⁺, 1223.325 [M+Na]⁺, 1239.307 [M+K]⁺. Almost complete conversion was observed for the naphthalene derivative **104f**: m/z calcd 1326.819; found 1350.418 [M+Na]⁺. Almost complete conversion was achieved for 2-ethynylpyridine **104g**: m/z calcd 1247.788; found 1271.323 [M+Na]⁺.

These encouraging results suggest that the overall strategy of on-chip library expansion by glycosylation and CuAAC reactions is feasible. However, since array stability turned to be a crucial and delicate parameter, a convenient alternative may be to perform the enzymatic glycosylation in solution. Indeed, glycosylation with tagged monosaccharide donors of amino tethered acceptors leads to labelled glycans that, after simple purification with graphite cartridges, can be spotted on the array for on-chip CuAAC library expansion. This approach simultaneously limits the amount of synthetically precious sugar nucleotides needed for enzymatic reactions and minimizes array decay resulting in MALDI-TOF spectra with better S/N ratio. Moreover, this method enables for glycosylation of multiantennary compounds, allowing their printing and enriching the complexity of glycans displayed on the array.

Further studies to identify more efficient conditions for on-chip CuAAC are currently being performed at CIC biomaGUNE. The main goal is to broaden the scope of alkynes while improving conversions. Indeed the reaction failure observed using some members of the terminal alkyne set (5-hexynoic acid **104b**, 3-phenylpropyne **104e**, 3-cyclohexylpropyne **104h**, 3-butynyl- β -D-glucopyranoside **104i**) points out that further optimization is required. Interestingly, the absence of conversion using the water soluble 3-butynyl- β -D-

glucopyranoside **104i** suggests that, even though solubility may be an issue for some alkynes, additional parameters have to be fine-tuned. In this regard, the amount of catalyst used may represent a crucial point.

3.4.1.5 Array stability: fluorescent analysis

Since desorption and leaching of printed glycans is possibly occurring during physical treatments and chemical transformations on the array, we aimed to study the stability of hydrophobically coated ITO slides by fluorescence studies. For our purpose, fluorescein cadaverine was printed on the slides, which were then subjected to the same conditions used during glycosylation and CuAAC reactions. The stability of the array was evaluated measuring and quantifying the loss of fluorescent signal.

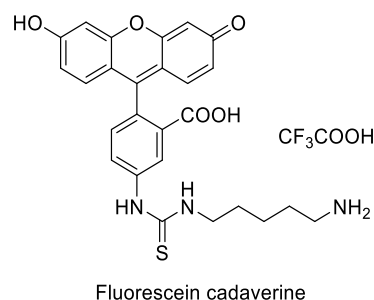


Figure 3.13 – Fluorescein cadaverine structure. The amino tethered fluorescein cadaverine allowed to be immobilized on ITO slides to assess array stability by fluorescence studies.

To evaluate the extent of glycan loss, different concentrations of fluorescein cadaverine solutions (ranging from 5 mM to 1 fM) were printed on activated ITO slides. The slides were quenched with ethanolamine solution (50 mM in borate buffer pH = 9.3) and treated as during enzymatic glycosylation (incubation overnight at 37 °C with aqueous solution containing Hepes buffer 50 mM and MnCl₂ 10 mM) and CuAAC reactions (incubation for 2 h at r.t. with a 95:5 H₂O/DMSO mixture). Fluorescence intensity was measured after each step and leaching of glycans was evaluated as the loss of intensity after simulation of the CuAAC reaction compared to fluorescence measured after quenching (Table 3.2). Fluorescence was detected by array scanner irradiating at 532 nm.

Unexpectedly, poor linear correlation was observed between fluorescein concentration and fluorescent signal after quenching of the slides. The loss of glycans at 50 μM, the concentration at which our glycans are spotted on the ITO slides, was found to be around 14% which is an acceptable value for our experiments. However, fluorescence measurements were sometimes affected by the quality of the slides possibly due to unspecific binding. An improvement may perhaps be obtained using a less hydrophobic and bulky fluorescent dye such as a compound belonging to the cyanine family.

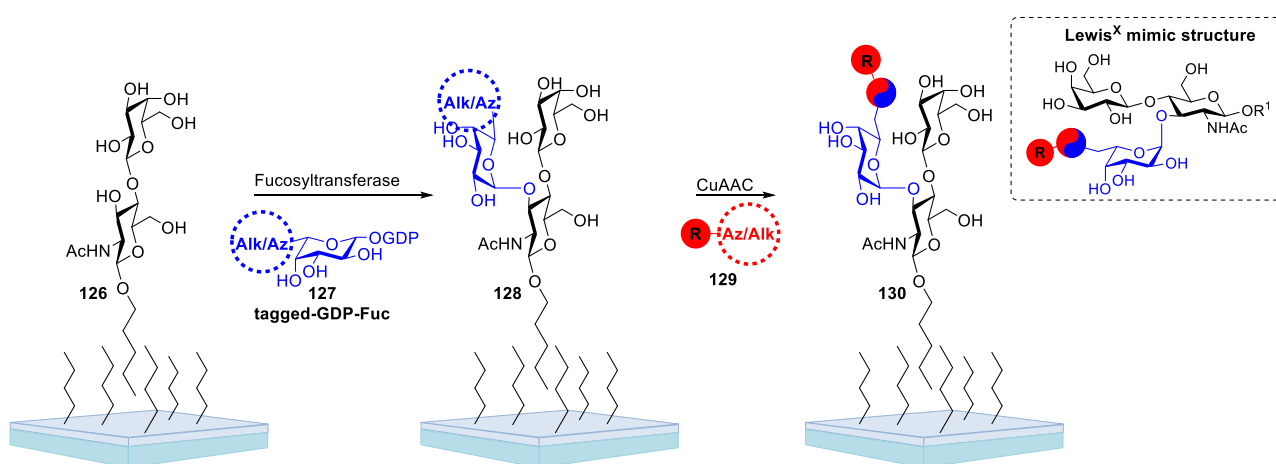
Concentration	After quench	After glycosylation	After CuAAC	Fluorescence intensity loss (compared to fluorescence after quenching)
5 mM	1517	426	756	34%
1mM	1065	851	850	11%
500 μM	1072	1072	702	21%
100 μM	964	564	928	2%
50 μM	801	602	602	14%
10 μM	371	371	151	42%
5 μM	291	193	152	31%

Table 3.2 – On-array fluorescent studies. The Table reports the median fluorescence intensities measured after printing of fluorescein cadaverine on ITO slides at different concentrations. In particular, the slides were treated as during

glycosylation and CuAAC reaction. Fluorescence intensities were detected after quenching of the slides with ethanolamine and after glycosylation reaction and CuAAC reaction. The loss of glycans was estimated comparing fluorescence intensity after CuAAC reaction with initial fluorescence (after quenching of the slides).

3.4.2 Synthesis of Lewis^X mimics array: looking for lead structures

Once established the feasibility of the methodology, we decided to apply this strategy for the on-chip synthesis of a library of Lewis^X mimics. The natural occurring Lewis^X moiety is a fucose containing trisaccharide, Gal β (1,4)[Fuc α (1,3)]GlcNAc, which has shown to perform as a ligand for DC-SIGN.¹⁹ Considering the relevance of DC-SIGN as potential pharmaceutical target, we reasoned that the fast generation of a Lewis^X mimics library would constitute a very appealing and efficient way for the identification of novel and potent lead structures. Hence, following a similar approach as discussed above, Lewis^X epitopes could be introduced by enzymatic glycosylation of arrayed lactosamine moieties **126** using conveniently tagged GDP-fucose glycoside donors **127** and allowing for library generation by CuAAC reaction (Scheme 3.9).

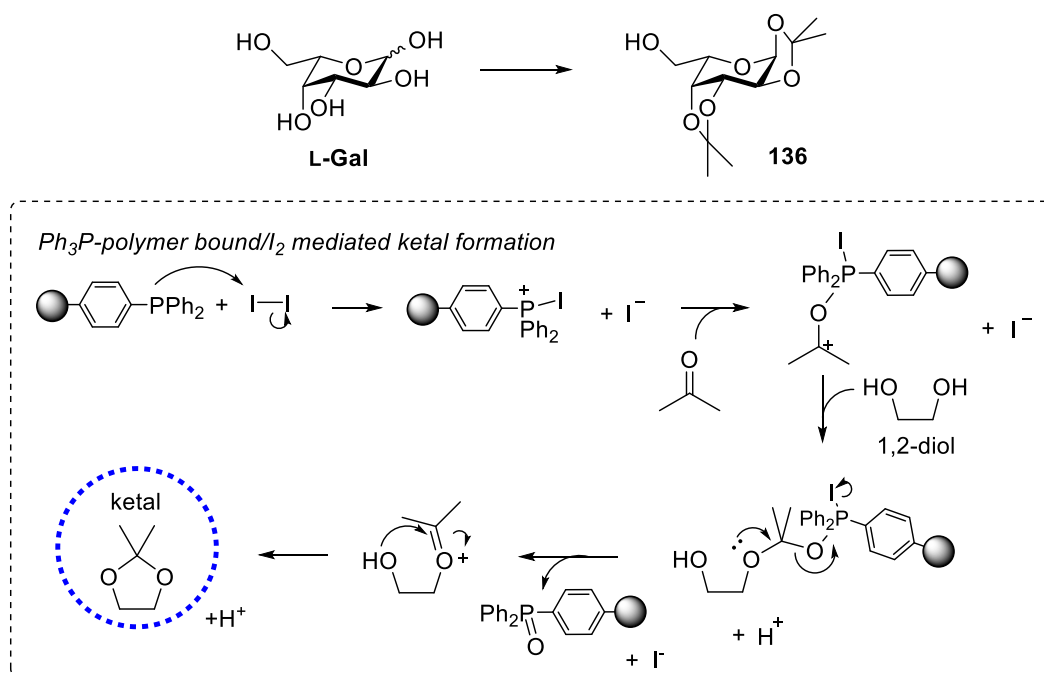


Scheme 3.9 – On-array Lewis^X mimics library generation. On-chip synthesis of Lewis^X mimics can be achieved by enzymatic fucosylation of printed lactosamine moieties **126**. The use of labelled GDP-fucose derivatives allows the introduction of a tag (azide or alkyne function) that will be exploited to perform CuAAC reactions on the array.

Azide or alkyne-tagged GDP-fucose mimics are not commercially available and we decided to synthesize them on our own. In particular, the 6-azido-GDP-galactose **127** and the 5-ethynyl-GDP-fucose **128** (Scheme 3.10) were selected as known substrates of the α -1,3-fucosyltransferase from *Helicobacter pylori*.⁶²

Two main strategies have been reported for the synthesis of sugar nucleotide diphosphates (Scheme 3.10).⁶³ Sugar phosphate derivatives such as **132** or **133** can be employed as nucleophilic counterpart in reactions with activated nucleotide monophosphates X-GMP (Scheme 3.10, path a). Most of the developed methodologies rely on this synthetic pathway and various activating groups X have been reported.⁶⁴ Alternatively, carbohydrate derivatives can be activated by leaving groups at the anomeric position, thus acting as electrophiles in nucleophilic substitution reactions with nucleotide diphosphates (Scheme 3.10, path b). Notably, this route was exploited by Timmons and co-workers for the synthesis of GDP-L-fucose.⁶⁵ Both pathways were considered to design a retrosynthetic analysis of tagged glycoside donors **127** and **128**. Unfortunately, both strategies are usually characterized by overall low yields, which is particularly true for the synthesis of GDP-L-fucose derivatives. However, fewer synthetic steps are required in pathway b, which was therefore selected as the more convenient route. Therefore, the synthesis of tagged GDP-fucose **127** and **128** was achieved coupling GDP with fucosyl bromide derivatives **134**, **135**, which were obtained by elaboration of L-galactose.

fast and clean reaction was obtained treating L-galactose with a polymer-bound triphenylphosphine and iodine in acetone at room temperature.⁶⁹ This reaction gave complete conversion in 30 minutes (monitoring by TLC analysis) (Table 3.3, entry 3). Moreover, a simple filtration allows to get rid of triphenylphosphine oxide side-product immobilized on the polymeric resin, which may otherwise complicate the purification step. Despite no work-up has been reported for this methodology, we found that quenching of the hydrogen iodide developed during the reaction was essential to avoid decomposition of bis-acetonide **136**. Thus, the reaction was quenched directly adding Na₂CO₃ as a solid which was later on filtered away along with the polymer-bound triphenylphosphine oxide. Removal of iodine was achieved by chromatographic purification. Surprisingly, low yields have always been observed after purification by automated flash chromatography (37-42% by direct phase), which is possibly due to hydrolysis of **136** under acidic catalysis during dry load. On the contrary, flash column chromatography afforded **136** in 75% yield.



Scheme 3.12 – L-Gal protection as bis-acetonide derivative **136**. L-Gal was protected as bis-acetonide by reaction with polymer-bound triphenylphosphine and iodine in acetone. The reaction involves the formation of a phosphonium ion intermediate which leads to the activation of acetone towards nucleophilic attack by a diol. Finally, the intermediate evolves towards the formation of an oxonium ion which allows the formation of the ketal.

Entry	Conditions	Yield
1	CuSO ₄ , H ₂ SO ₄ (30 mol%), acetone	Partial conversion
2	2,2-dimethoxypropane, <i>p</i> TsOH (10 mol%), 30 °C	50% after flash chromatography
3	1. Ph ₃ P polymer-bound, I ₂ , acetone, r.t., 30 min 2. Na ₂ CO ₃	Pure product by ¹ H NMR 75% after flash chromatography

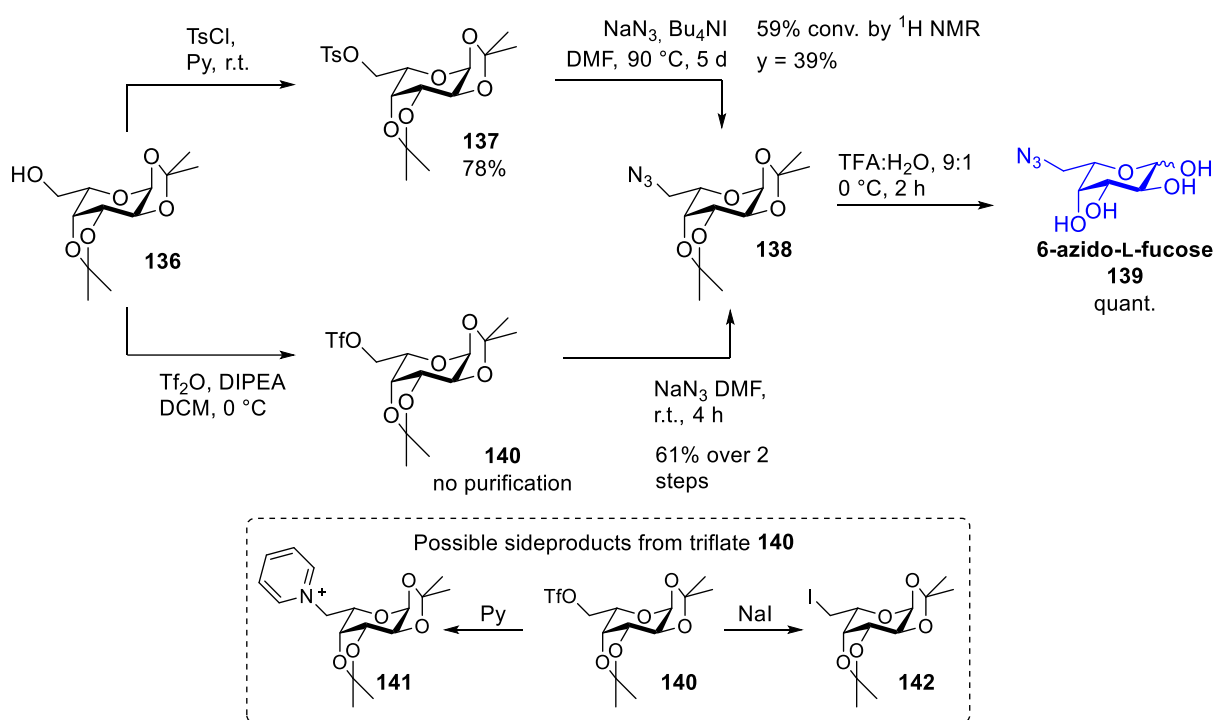
Table 3.3 – Synthesis of L-Gal bis-acetonide. The table reports the conditions attempted for the protection of L-Gal as bis-acetonide.

The bis-acetonide **136** gave access to the 6-azido-L-fucose **139**, a key intermediate in the synthesis of the corresponding 6-azido-GDP-L-fucose **127** (Scheme 3.13). Following a procedure that was adopted in our group for the synthesis of the enantiomeric 6-azido-D-galactose,⁵¹ the free hydroxyl group of **136** was activated

to give the corresponding tosylate **137** in good yield, which was subjected to nucleophilic substitution by sodium azide in the presence of a catalytic amount of tetrabutylammonium iodide. Despite the prolonged reaction times at high temperature (5 days at 90 °C) only partial conversion was achieved (59% estimating by ^1H NMR) and pure azide **138** was obtained after flash chromatography only in 39% yield.

The low conversion with tosylate **137** brought us to explore an alternative route envisaging the more reactive triflate intermediate **140** (Scheme 3.13).⁷⁰ Complete conversion of bis-acetonide **136** was achieved upon reaction with triflic anhydride in the presence of DIPEA at low temperature and the resulting triflate **140** was directly submitted to nucleophilic substitution with sodium azide. Complete conversion was achieved in only four hours and pure azide **138** was isolated in 61% yield over two steps. Notably, in the synthesis of triflate **140**, DIPEA was preferred to the usually employed pyridine in order to avoid the massive formation of side product **141** (clearly identified both by ^1H NMR and ESI-MS), which proved to react slowly and incompletely with sodium azide in the following substitution step. The purity of bis-acetonide **136** is also a critical point. In particular, contamination from iodine should be avoided since the formation of iodide **142** has been observed. However, this side product could be completely converted into the desired azide **138** upon reaction with NaN_3 for 7 days at 80 °C.

Finally, removal of the acetonide groups from protected azide **138** with a 9:1 TFA/ H_2O mixture quantitatively afforded 6-azido-L-fucose **139** without purification needed (Scheme 3.13).⁷¹



Scheme 3.13 – Synthesis of 6-azido-L-fucose **139.** The azido labelled **139** was obtained following two alternative routes starting from the protected bis-acetonide **136**. Particularly, the synthetic strategy involving triflate **140** as intermediate turned to be the most advantageous. The use of pyridine as a base in the formation of triflate **140** was avoided to prevent the formation of little reactive **141**. Iodide contamination of bis-acetonide **136** can lead to the formation of **142** as side-product.

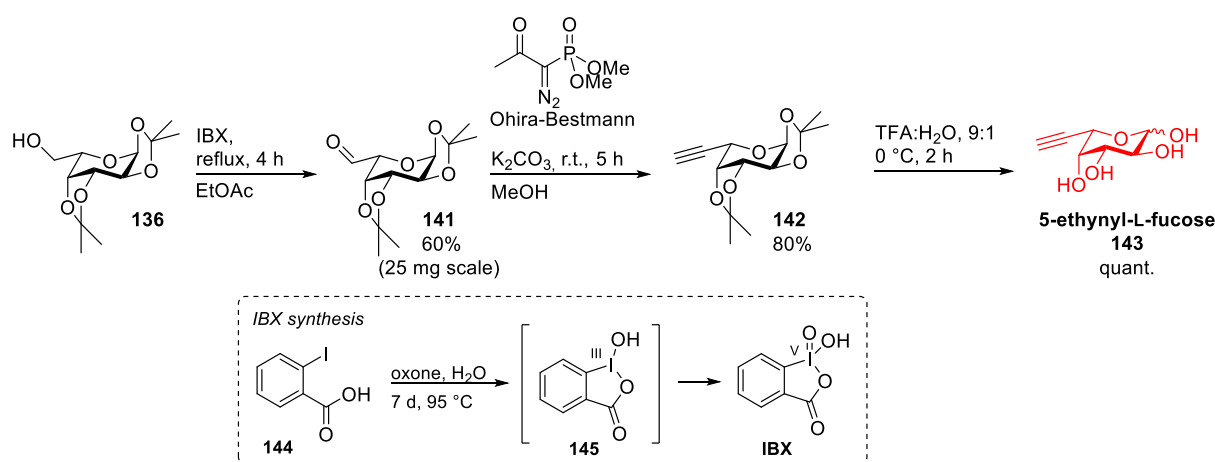
As mentioned above, the bis-acetonide **136** was also employed for the preparation of 5-ethynyl-GDP-L-fucose **128** which required the synthesis of 5-ethynyl-L-fucose as key intermediate (Scheme 3.14).⁷¹

Oxidation of the free hydroxyl group of **136** was achieved using IBX in refluxing ethyl acetate. The aldehyde **141** was isolated by flash chromatography in 60% yield. The required hypervalent iodine reagent IBX was synthesized from 2-iodobenzoic acid **144** with oxone as highly oxidizing agent.⁷² The reaction was

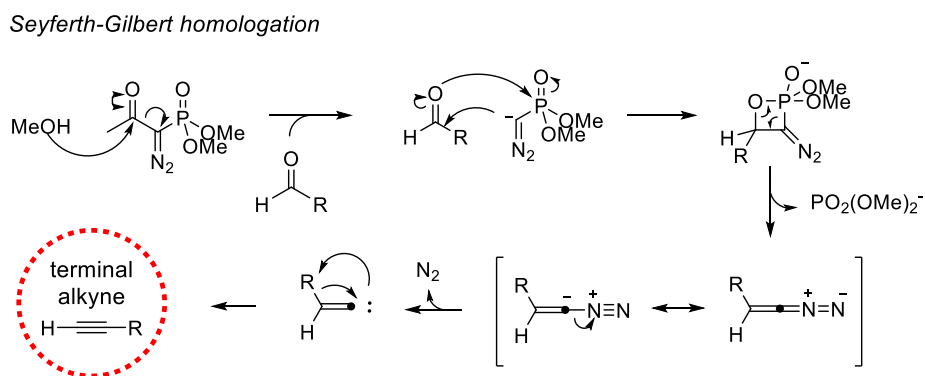
monitored by ^1H NMR, which allowed to assess the complete conversion of the partially oxidized I(III) intermediate **145**. Unfortunately, the yield of aldehyde **141** literally dropped as the reaction was scaled up (15% with ~ 1.3 g of **136**). A better outcome may be obtained optimizing the reaction conditions or using a different oxidation methodology with alternative oxidants.⁶⁶

The alkyne **142** was obtained in good yield (80%) from **141** by Seyferth-Gilbert homologation reaction (for the mechanism see Scheme 3.15). The reaction was performed using the Ohira-Bestmann diazo compound reagent.⁷³ This relatively expensive reagent allows for clean and fast reactions and in-situ generation using more easily available and cheaper precursors have also been reported.⁷⁴

Finally acetonide removal from **142** in 9:1 TFA/ H_2O mixture quantitatively afforded the 5-ethynyl-L-fucose **143** with no purification needed.



Scheme 3.14 – Synthesis of 5-ethynyl-L-fucose **143**. Alkyne sugar **143** was generated by oxidation of bis-acetonide **136** in presence of IBX and Seyferth-Gilbert homologation using the Ohira-Bestmann reagent. The synthesis of IBX from 2-iodobenzoic acid **144** is also shown.



Scheme 3.15 – Seyferth-Gilbert homologation mechanism. The Ohira-Bestmann reagent decomposes in situ to give a stabilized anion that attacks the aldehyde with the formation of a four membered ring. Rearrangement and elimination of nitrogen lead to the formation of a carbene that, upon transposition of the R group, affords the desired terminal alkyne.

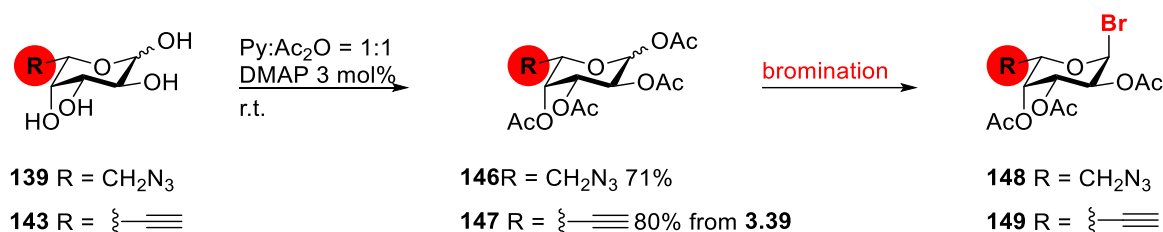
To obtain the targeted GDP derivatives **127** and **128**, both the tagged 6-azido-L-fucose **139** and 5-ethynyl-L-fucose **143** had to be activated as anomeric bromides (Scheme 3.16). These intermediates were first peracetylated with acetic anhydride in pyridine. The reaction afforded the tetra-*O*-acetates **146** and **147** in good yield (71% and 80% over 2 steps, respectively starting from bis-acetonide intermediates **138** and **142**). ^1H NMR analysis revealed that the products were obtained as a α/β anomeric mixture of pyranosidic forms (1.6:1 β/α ratio for tetra-*O*-acetate azido fucose **146**; 1.2:1 β/α ratio for tetra-*O*-acetate ethynylfucose **147**) along with the corresponding α/β mixture of furanose isomers (11%, 1:1 β/α ratio for tetra-*O*-acetates azido fucose; 43%,

1:1.3 β/α ratio for tetra-*O*-acetates ethynylfucose **147**).⁷⁰ The desired pyranose tetra-*O*-acetates **146** and **147** were not isolated, but the mixtures of isomers were directly subjected to bromination, which selectively afforded the pyranose α -bromides **148** and **149** (see below). The furanose forms of tetra-*O*-acetates azido **146** were not converted during bromination and were still observable in the reaction crude by ¹H NMR.

The selective introduction of a bromine atom at the anomeric position of both peracetylated **146** and **147** was not trivial. This is particularly true for the preparation of the 6-azido bromide **148**, for which different methodologies were explored. The use of an HBr solution 33% in acetic acid as brominating agent, which was exploited in our group for the synthesis of the analogous (OAc)₃- α -D-fucopyranosyl bromide,⁷⁵ failed (Table 3.4, entry 1). Indeed, no formation of the desired bromide **148** was observed, while a major side-product, probably deriving from unwanted substitution of the azide function, was clearly detectable by ¹H NMR. Reactions involving either PBr₃ or TiBr₄ left the azide group untouched,^{70,76} but low conversion was always assessed monitoring by TLC even after long reaction times (Table 3.4, entry 2,3). This problem was overcome by performing the reaction with TiBr₄ at 40 °C in a microwave assisted mode, which successfully achieved almost complete conversion to bromide **148** in 1.5 h (Table 3.4, entry 4).⁷⁷ Most of the unreacted peracetylated **146** was identified by ¹H NMR as the α -OAc anomer, suggesting a lower reactivity towards bromination compared to the β -OAc anomer.

As opposed to the bromination of **146**, the tetraacetylated alkyne derivative **147** was successfully activated as the corresponding bromide **149** using a 33% HBr solution in acetic acid in 1 h (Table 3.4, entry 5).

Fucosyl bromides **148** and **149** were not isolated and could be employed directly in the next synthetic step. Notably, both compounds were selectively obtained as α -anomers due to the strong anomeric effect of the bromine substituent.



Scheme 3.16 – Synthesis of sugar bromides 148 and 149. The azide and alkyne fucose mimics **139** and **143** were peracetylated and selectively transformed into the α -sugar bromides **148** and **149**, respectively.

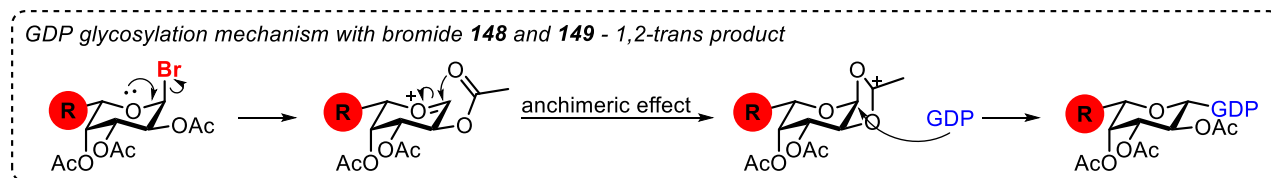
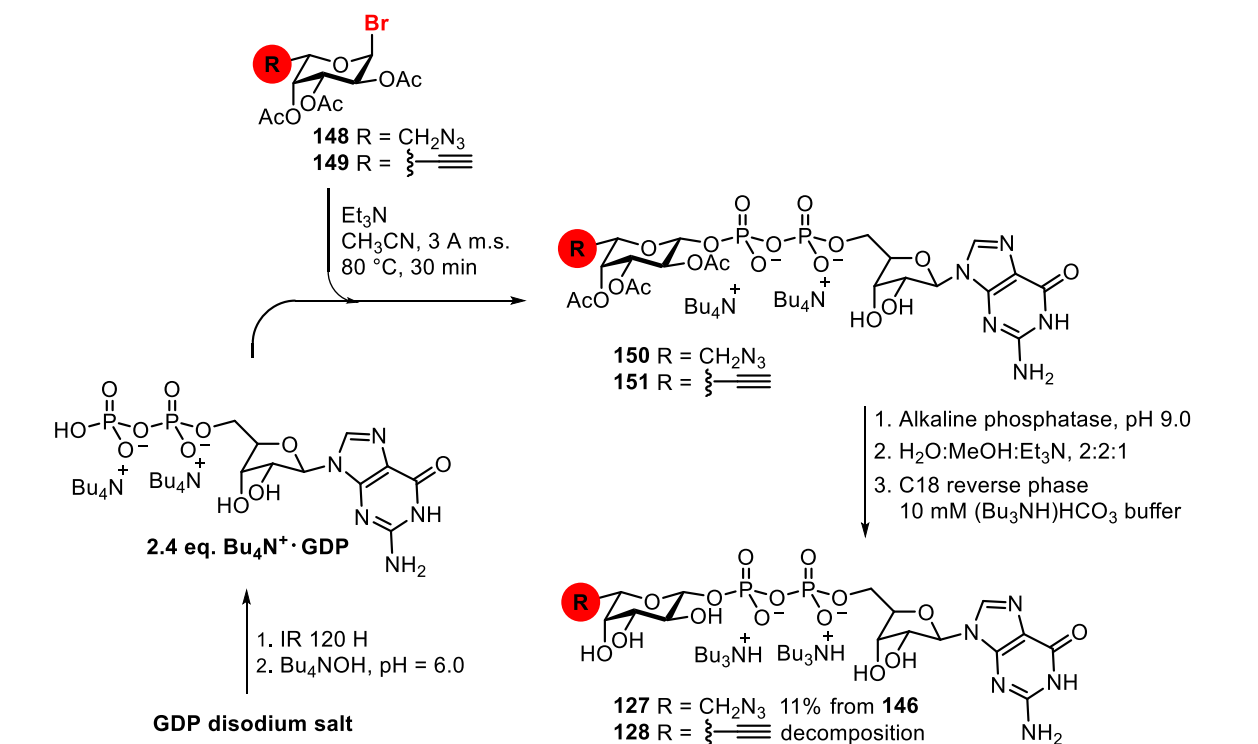
Substrate	Entry	Conditions	Results
	1	33% HBr in AcOH, DCM, r.t., 50 min	Undesired product, loss of N ₃
	2	PBr ₃ , H ₂ O, DCM, r.t., 3 h	148 Low conversion by TLC
	3	TiBr ₄ , 9:1 DCM/EtOAc, r.t., 2 d	148 Low conversion by TLC
	4	TiBr ₄ , 9:1 DCM/EtOAc, 40 °C, MW, 1.5 h	148 Almost complete conversion by ¹ H NMR
	5	33% HBr in AcOH, DCM, r.t., 1 h	149 Almost complete conversion by ¹ H NMR

Table 3.4 – Bromination conditions. The table reports the reaction conditions explored for bromination of the azido tagged **146** (entry 1-4) and the alkyne labelled **147** (entry 5).

With the activated bromide sugars **148** and **149** in hands, the coupling reaction with GDP was finally possible (Scheme 3.17).⁶⁵ GDP, purchased as the sodium salt, was protonated using an ion exchange acidic resin (IR 120 H) and finally titrated with a solution of Bu₄NOH to pH 6. This is a critical point since reactions with GDP Bu₄N salts titrated at pH 5 or 7 gave poor results.

The reactions with fucosyl bromide **148** and **149** were performed under anhydrous conditions in CH₃CN at 80 °C. After 30 minutes TLC analysis revealed the complete consumption of the glycoside donors while the formation of the glycosyl-*O*-acetates GDP derivatives **150**, **151** was assessed by analytical HPLC analysis (Fig. 3.15a and 3.16a. All the analysis were performed monitoring at 254 nm eluting at 1 mL/min with a gradient from 90% (aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, pH 4)) – 10% (CH₃CN) to 40% (aqueous buffer) – 60% (CH₃CN) in 2 min followed by a plateau for 2 min and a second gradient until 100% (CH₃CN) in 2 min. **150** *t*_R 8.3 min, **151** *t*_R 8.1 min). Notably, a lower formation was observed for the 5-ethynyl-(*O*-Ac)₃-GDP-L-fucose **151**. To simplify isolation of the products, the crude reaction mixtures were treated with alkaline phosphatase enzyme at controlled pH 9 to hydrolyze the unreacted GDP into the corresponding pyrophosphate and guanosine nucleoside (Fig. 3.15b and 3.16b; guanosine *t*_R 1.4 min). After enzymatic digestion, the crudes were treated with a 2:2:1 H₂O/MeOH/Et₃N mixture for 16 hours to afford the desired **127** and **128**, as revealed by HPLC analysis (Fig. 3.15c and 3.16c; **127** *t*_R 6.9 min; **128** *t*_R 6.5 min). Finally, reverse phase chromatography with 10 mM (Bu₃NH)HCO₃ aqueous buffer was used in order to isolate the products as tributylammonium salts. Satisfyingly, the 6-azido-GDP-L-fucose **127** was selectively obtained as the β anomer (for glycosylation mechanism see Scheme 3.17) in 11% overall yield starting from the 6-azido-(*O*Ac)₄-L-fucose **146**.

On the contrary, the 5-ethynyl-GDP-L-fucose **128** totally decomposed during this final purification. Hence, an optimization of deprotection and purification step is likely required.



Scheme 3.17 – Synthesis of GDP-fucose mimics **127 and **128**.** Sugar bromides **148** and **149** were reacted in presence of GDP and subsequently deacetylated to give the targeted GDP-derivatives **127** and **128**. The glycosylation reaction selectively afforded the GDP-sugars as β -anomers due to the anchimeric effect of the participating acetyl group.

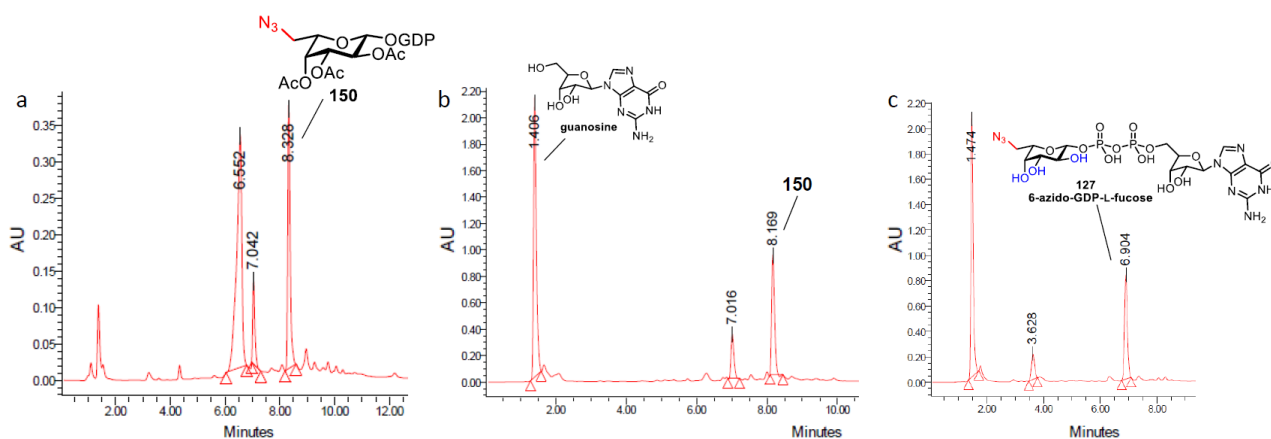


Figure 3.15 – HPLC traces of reaction crudes for the synthesis of 6-azido-GDP-L-fucose **127 starting from GDP and fucosyl bromide **148**.** All the analysis were performed monitoring at 254 nm eluting at 1 mL/min with a gradient from 90% (aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, pH 4)) – 10% (CH₃CN) to 40% (aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, pH 4)) – 60% (CH₃CN) in 8 min followed by a plateau for 2 min and a second gradient until 100% (CH₃CN) in 2 min. **a)** HPLC trace for the crude of GDP glycosylation using fucosyl bromide **148**. Glycosylation products **150** t_R = 8.3 min; unreacted GDP t_R = 6.5 min. **b)** HPLC trace after enzymatic digestion of the crude with alkaline phosphatase. The disappearance of GDP is associated with the appearance of a second

signal which was ascribed to guanosine, $t_R = 1.4$ min. **c**) HPLC trace for the crude of solvolysis reaction in 2:2:1 H₂O/MeOH/Et₃N. Removal of acetyl groups afforded the desired product **127** $t_R = 6.9$ min.

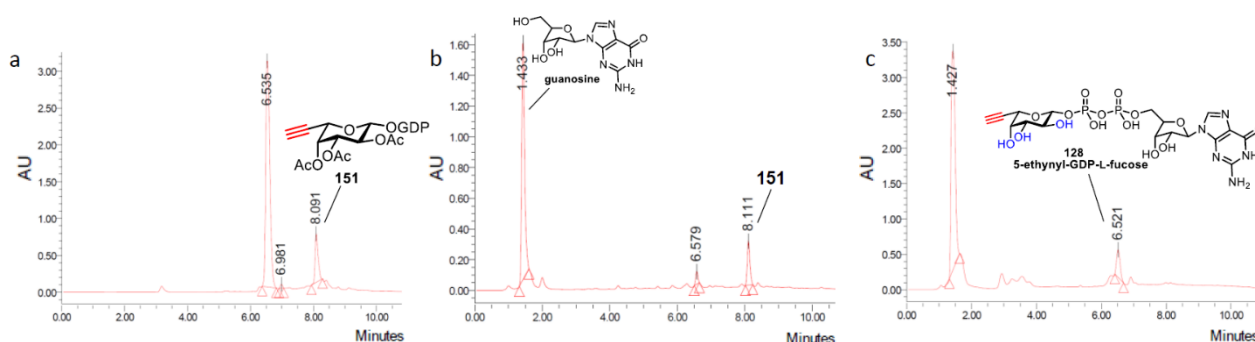
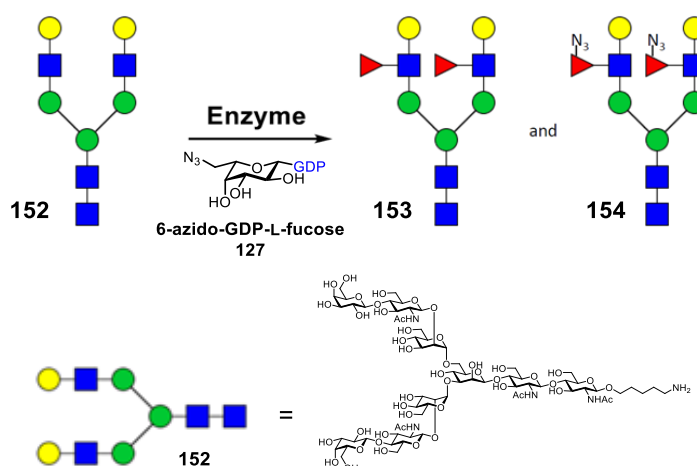


Figure 3.16 – HPLC traces of reaction crudes for the synthesis of 5-alkynyl-GDP-L-fucose 128 starting from GDP and fucosyl bromide 149. All the analysis were performed monitoring at 254 nm eluting at 1mL/min with a gradient from 90% (aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, pH 4)) – 10% (CH₃CN) to 40% (aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, pH 4)) – 60% (CH₃CN) in 8 min followed by a plateau for 2 min and a second gradient until 100% (CH₃CN) in 2 min. **a**) HPLC trace for the crude of GDP glycosylation using fucosyl bromide **149**. Glycosylation product **151** $t_R = 8.1$ min; unreacted GDP $t_R = 6.5$ min. **b**) HPLC trace after enzymatic digestion of the crude with alkaline phosphatase. The disappearance of GDP signal is associated with the appearance of a second signal which was ascribed to guanosine, $t_R = 1.4$ min. **c**) HPLC trace for the crude of solvolysis reaction in 2:2:1 H₂O/MeOH/Et₃N. Removal of acetyl groups afforded the desired product **128** $t_R = 6.5$ min.

In order to introduce Lewis^X epitopes on arrayed glycans, different enzymes were screened in solution for glycosylation reactions with biantennary **152** as the model glycoside acceptor. The enzyme ability to transfer the unnatural 6-azido-GDP-L-fucose **127** was compared with the corresponding natural glycoside donor GDP-L-fucose. Fucosylation with CeFUT6 enzyme proceeded completely using GDP-L-fucose leading to glycan **153**, as it was assessed by MALDI-TOF MS. On the contrary, the CeFUT6 enzyme⁷⁸ completely failed in reaction with 6-azido-GDP-L-fucose **127** as glycoside donor and no formation of **154** was observed. However, full conversion of **152** glycan with both the natural and unnatural glycoside donors was obtained using the α -1,3-fucosyltransferase from *H. pylori* (Fig. 3.17). This experiment demonstrates that α -1,3-fucosyltransferase can be used for direct glycosylation of glycan arrays with the azido tagged **127** allowing the on-chip synthesis of Lewis^X mimics library. Studies to achieve this goal are currently on going at CIC biomaGUNE.



Scheme 3.18 – In solution fucosylation tests. The activity of CeFUT6 and α -1,3-fucosyltransferase from *H. pylori* in transferring the azido-GDP-L-fucose **127** were evaluated by in solution reactions and compared with corresponding transformation using the natural GDP-fucose.

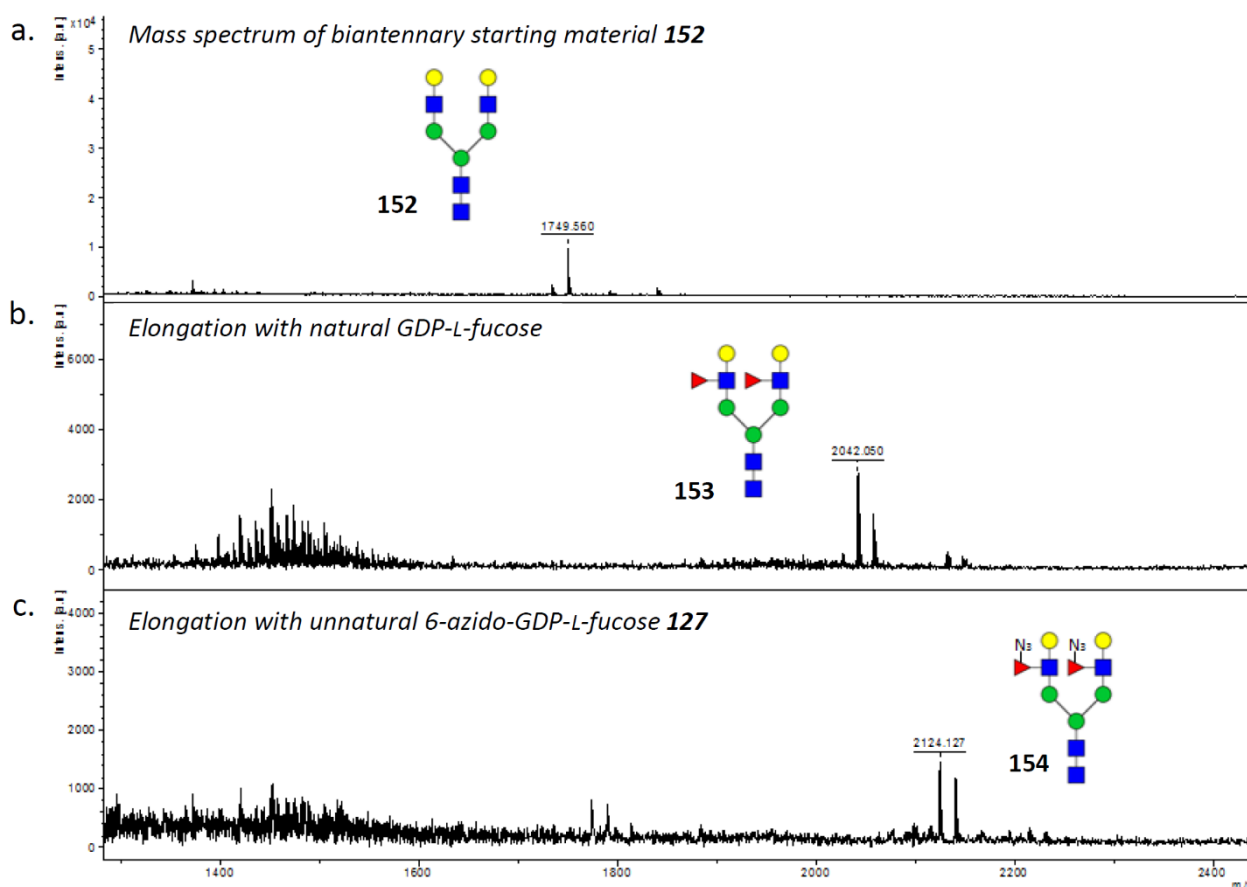


Figure 3.17 – MALDI-TOF MS spectra for enzymatic fucosylation of biantennary 152. Biantennary **152** was successfully fucosylated using α -1,3-fucosyltransferase from *H. pylori*. This enzyme was able to accept and transfer both the natural GDP-L-fucose and the 6-azido-GDP-L-fucose **127**. **a)** MALDI-TOF spectrum of biantennary **152**, m/z : calcd 1725.681; found 1749.560 [M+Na]⁺. **b)** Biantennary **153** obtained by fucosylation reaction of **152** with GDP-fucose, m/z : calcd 2017.797; found 2042.050 [M+Na]⁺. **c)** Azido tagged **154** obtained by fucosylation reaction of **152** with unnatural 6-azido-GDP-L-fucose **127**, m/z : calcd 2099.800; found 2124.127 [M+Na]⁺.

3.5 Conclusions

Glycan microarrays are a useful tool in the identification of new lead structures enabling fast analysis of large glycan libraries. We decided to exploit this technology to develop new glycomimetic structures able to perform as potent lectin antagonists. In particular, we established a methodology for on-chip generation and expansion of glycan libraries by enzymatic glycosylation and CuAAC reaction. Glycosylation with unnatural azido tagged glycoside donors (UDP-GalNAz and 6-azido-UDP-D-Gal) was investigated, showing that an optimization is required for the complete functionalization of multiantennary glycans. For this reason, monoantennary glycans were selected as the most suitable structures for preliminary CuAAC reaction experiments. Promising results were obtained for a small library of alkynes, most of which reacted almost completely with the arrayed tagged glycans.

This piece of work provided proof of concept that a glycan library non-covalently immobilized on ITO microarrays using the Reichardt's approach can be expanded and diversified using two subsequent transformations, i.e. an enzymatic glycosylation with a nucleotide donor of a tagged monosaccharide, followed by CuAAC reactions. However, from a practical point of view, a much more robust system with improved S/N

ratio can be obtained if the enzymatic transformation is performed in solution and only the final CuAAC modification is actually carried out on the array.

Following this innovative approach, we aimed to synthesize a library of Lewis^X mimics for the identification of new potent antagonist of the DC-SIGN lectin. The strategy we designed envisages the use of labelled GDP-fucose derivatives; consequently the synthesis of the 6-azido-GDP-L-fucose **127** and 5-ethynyl-GDP-L-fucose **128** was undertaken. Unfortunately, the alkyne derivative **128** decomposed during the last purification stage and its synthesis is still under study. However, the azido-GDP-fucose derivative **127** was successfully prepared and proved to be readily accepted as glycoside donor by the α -1,3 fucosyltransferase from *Helicobacter pylori*. Thus, this substrate gives access to the enzymatic synthesis of azido labelled Lewis^X epitopes, which will be exploited for the on-array synthesis of Lewis^X mimics library at CIC biomaGUNE.

3.6 Experimental

3.6.1 General methods and procedures

Microarrays were printed employing a robotic piezoelectric non-contact spotter SciFLEXARRAYER S11 (Scienion, Berlin, Germany). Indium tin oxide (ITO) coated glass slides (75 mm x 25 mm) were obtained from Hudson Surface Technology, Inc. (Fort Lee, NJ). The slides have a nominal transmittance of >78 % and an ITO thickness of 130 nm. Modified surfaces were stored under vacuum conditions until use. Enzymatic and CuAAC reactions on the slides were performed using 16 wells Fast Frame® incubation chambers from Whatman (Kent, UK). All aqueous solutions used for array fabrication and protein expression were prepared from nanopure water produced with a Diamond UV water purification system (Branstead International). pET30a plasmid bearing the bovine β -1,4-galactosyltransferase-1 gene was obtained from Dr. Peter Both, University of Manchester, UK. Bovine β -1,4-galactosyltransferase-1 double mutant (C342T and Y289L) was prepared as previously described.⁵⁵ OD_{280nm} of protein solutions and OD_{600nm} of bacterial cultures were measured in a Nanodrop®ND-1000 spectrophotometer (Thermo Scientific, Wilmington, USA). Chromatography was performed in an Acquity UPLC system equipped with photodiode array detector (PDA) using an Acquity BEH Amide column (50x 2.1 mm, 1.7 μ m) from Waters (Milford, MA, USA). The mass spectrometry detection was carried out using a time-of-flight mass spectrometer (ESI-TOF) LCT Premier XE from Waters (Milford, MA, USA) with an electrospray ionization source, working in positive /W mode. The MS range acquired was between m/z 100-1000. Masslynx v4.1 software was used to analyze chromatograms and spectra (Waters, Milford, MA, USA). MALDI-TOF mass analyses were performed on an Ultraflex extreme III time-of-flight mass spectrometer equipped with a pulsed Nd:YAG laser (355 nm) and controlled by FlexControl 3.3 (Bruker Daltonics, Bremen, Germany). The acquired data were processed using the Bruker software FlexAnalysis 3.3. Fluorescence measurements were performed in an Agilent G265BA microarray scanner system (Agilent Technologies, Santa Clara, USA). Quantification was performed with ProScanArray® Express software (Perkin Elmer, Shelton, USA).

Chemicals were purchased from commercial sources and used without further purification, unless otherwise specified. When anhydrous conditions were required, the reactions were performed under nitrogen or argon atmosphere. Anhydrous solvents were purchased from Sigma-Aldrich® with a content of water \leq 0.005 %. *N,N*-diisopropylethylamine (DIPEA), triethylamine, acetonitrile were dried over calcium hydride and freshly distilled. Reactions were monitored by analytical thin-layer chromatography (TLC) performed on Silica Gel 60 F₂₅₄ plates or aluminium sheets (Merck) with UV detection (254 nm and 365 nm) and/or staining with ammonium molybdate acid solution, potassium permanganate alkaline solution or vanillin solution. Titrations and reactions at controlled pH were performed monitoring with a PH 211 Microprocessor pH-meter (HANNA instruments) equipped with a HI 1331B electrode (HANNA instruments). Flash column chromatography was performed according to the method of Still and co-workers⁷⁹ using silica gel 60 (40-63 μ m) (Merck). Automated flash chromatography was performed with Biotage Isolera Prime system, Biotage SNAP KP-Sil

cartridges were employed unless otherwise indicated. Microwave irradiation was performed by a Biotage Initiator⁺ system. HPLC analyses were performed with an Atlantis T3 5 μm 4.6x100 mm column (Waters) equipped with a Waters 996 Photodiode Array Detector. HPLC purifications were performed on Dionex Ultimate 3000 equipped with Dionex RS Variable Wavelength Detector (column: Atlantis Prep T3 OBDTM 5 μm 19 x 100 mm; flow 15 mL/min). NMR experiments were recorded on a Bruker AVANCE-400 MHz and a Bruker UltraShield 500 MHz instrument at 298 K. Chemical shifts (δ) are reported in ppm. The ¹H and ¹³C NMR resonances of compounds were assigned with the assistance of COSY and HSQC experiments. Multiplicities are assigned as s (singlet), d (doublet), t (triplet), q (quartet), quint (quintet), m (multiplet). ESI mass spectra were recorded on Waters Micromass Q-TOF (ESI ionization-HRMS), ThermoFisher LCQ apparatus (ESI ionization). The following abbreviations are used: BSA (bovine serum albumin), DCM (CH₂Cl₂), DHB (2,3-dihydroxybenzoic acid), DIPEA (*N,N*-diisopropylethylamine), DMAP (4-dimethylaminopyridine), DMF (*N,N*-dimethylformamide), DMSO (dimethylsulfoxide), GDP (guanosine diphosphate), IBX (2-iodoxybenzoic acid), MW (microwave), NHS (*N*-hydroxysuccinimide), ODPa (octadecylphosphonic acid), THPTA (tris(3-hydroxypropyl)triazolylmethyl)amine, TFA (trifluoroacetic acid), UDP (uridine diphosphate).

Compounds **105**,⁸⁰ **106**,⁸¹ **107**,⁸² **108**,⁸² **110**,⁸² **109**,⁸³ **111**,⁸⁴ **152**,⁸³ **116**,⁵⁰ **117**, **104i** were previously synthesized and available in the Reichardt group. Compound **114** was previously prepared in the Bernardi group⁵¹ adapting a reported procedure.⁵² Compounds **104a**, **104b**, **104d-h**, **112**, **113**, **115**, L-galactono-1,4-lactone, L-Gal, polymer-bound triphenylphosphine, dimethyl(1-diazo-2-oxopropyl)phosphonate, GDP sodium salt, guanosine 5'-diphospho- β -L-fucose sodium salt (GDP-Fuc), fluorescein cadaverine are commercially available. Alkaline phosphatase from bovine intestinal mucosa was purchased from Sigma-Aldrich, α -1,3/4 Fucosyltransferase from *Helicobacter pylori* was purchased from Chemily Glycoscience (Atlanta, USA) and α -1,3 fucosyltransferase from *C. elegans* (CeFUT6) was prepared as previously described.⁷⁸

3.6.2 Array fabrication and characterization

3.6.2.1 Synthesis of self-assembled hydrophobic monolayer on ITO slides

Indium tin oxide (ITO) slides were washed with a 6:1:1 H₂O/NH₃ (25% aqueous solution)/H₂O₂ (30% w/w aqueous solution) basic piranha solution at 80 °C for 1 h. The slides were washed with nanopure H₂O and dried under argon flow. Then slides were incubated with 1 mM octadecylphosphonic acid (ODPA) solution in THF at room temperature for 3 h, dried under argon flow and annealed at 140 °C for 20 h. Finally, slides were sonicated in MeOH for 5 min and kept in MeOH for additional 30 min to remove any non covalently bound ODPA. Slides were dried under stream of argon and kept under high vacuum until use.

3.6.2.2 Reactive double lipidic layer formation

Hydrophobic ITO slides were incubated in Petri dishes with a 1 mM linker **116** solution in CHCl₃ at room temperature overnight. Slides were dried under stream of argon.

3.6.2.3 General procedure for glycan printing and detection

Glycan solutions (50 μM in phosphate buffer 300 mM, pH 8.7) were prepared and stored in a 384 multi-well plate. Glycan solutions were robotically printed (\approx 15 nL, \approx 0.7 pmol) onto hydrophobic ITO slides bearing adsorbed linkers **116**. Immobilization reaction was performed overnight at 17 °C and controlled humidity \sim 50%. The unreacted NHS groups were quenched by incubation with a 50 mM ethanolamine solution in 50 mM borate buffer pH 9.3 for 40 min. Slides were washed with nanopure H₂O and finally dried under argon flow. For glycan detection, a DHB solution (10 mg/mL in H₂O + 0.1 % TFA + 5% CH₃CN) was robotically printed over spotted glycans and the slides were analysed by MALDI-TOF MS.

3.6.2.4 GalT-1 and GalT-1-DM expression and purification

Glycerol stocks of either *E. coli* GalT-1 or GalT-1-C342T&Y289L mutants were plated on Luria Bertani (LB) medium (1% NaCl, 1% peptone, 0.5% yeast extract) supplemented with agar (15-20 g/L) and kanamycin (50 µg/mL). Plates were incubated at 37 °C overnight to allow formation of *E. coli* colonies.

A single colony was inoculated in 4 mL of LB medium containing kanamycin (50 µg/mL) and incubated at 37 °C shaking at 250 rpm overnight. This culture was used to inoculate 250 mL of LB medium containing kanamycin (50 µg/mL). Cells were grown at 37 °C, 250 rpm until the culture reached an OD_{600nm} value of 0.6. Then, protein expression was induced by the addition of IPTG (1 mM final concentration) for 4 h at 37 °C, 250 rpm. The culture was harvested by centrifugation (4500 rpm, 4 °C, 10 min), washed with Tris 50 mM, NaCl 150 mM, pH=8.5 aqueous solution and the pellet was stored overnight in the freezer. The pellet was then resuspended in 10 mL of lysis buffer (Tris 50 mM, NaCl 150 mM, 1% Triton X-100, EDTA 10 mM, PMSF, pH=8.5) and lysis was performed by sonication followed by centrifugation. Resuspension of the pellet, sonication and centrifugation were repeated a second time. The inclusion bodies were washed with water (1x) and stored at -20 °C until use.

Inclusion bodies were solubilized with a guanidine-HCl 5 M aqueous solution until OD_{280nm} value of 1.9 to 2.0. Then the protein solution was filtered through a 0.22 µm filter and diluted ten times in folding buffer containing Tris-HCl 50 mM, pH=8.5, NaCl 10.56 mM, KCl 0.44 mM, MgCl₂ 2.2 mM, CaCl₂ 2.2 mM, guanidine-HCl 0.5 M, cysteamine 8 mM, cystamine 4 mM, L-arginine 0.55 M.⁸⁵ The protein was allowed to renature for 48 h at 4 °C and then dialyzed against Tris 25 mM, 150 mM NaCl, pH 8.5 H₂O solution. Precipitated proteins were removed by centrifugation (4500 rpm, 4 °C, 15 min). Imidazole 10 mM was added to the supernatant which was purified by HisTrap™ HP 5 mL column controlled by an ÄKTA™ protein purifier and equilibrated with binding buffer (Tris 25 mM, NaCl 300 mM, 10 mM imidazole pH =8.5). The column was washed with 5 CV of binding buffer and the protein is eluted with Tris 25 mM, NaCl 300 mM, imidazole 500 mM, pH =8.5. Eluted fractions were analyzed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE), 12% (vol/vol) acrylamide gels in a Bio-Rad Mini Protean system. Gels were stained with Coomassie Blue G-250 (Fig. 3.18). Protein containing fractions were pooled and dialyzed against Tris 25 mM, NaCl 150 mM, pH 7.5. Precipitated proteins were eliminated by centrifugation (4500 rpm, 4 °C, 15 min) and the protein solution was concentrated in Vivaspin 20 devices (10000 MWCO) reaching proper OD_{280nm} values: GalT-1-C342T obtained as 5.75 mL solution with a 0.50 mg/mL concentration and GalT-1-C342T&Y289L mutant as 4.75 mL solution with a 0.62 mg/mL concentration. Finally the proteins were freeze dried and stored at -20 °C until use.

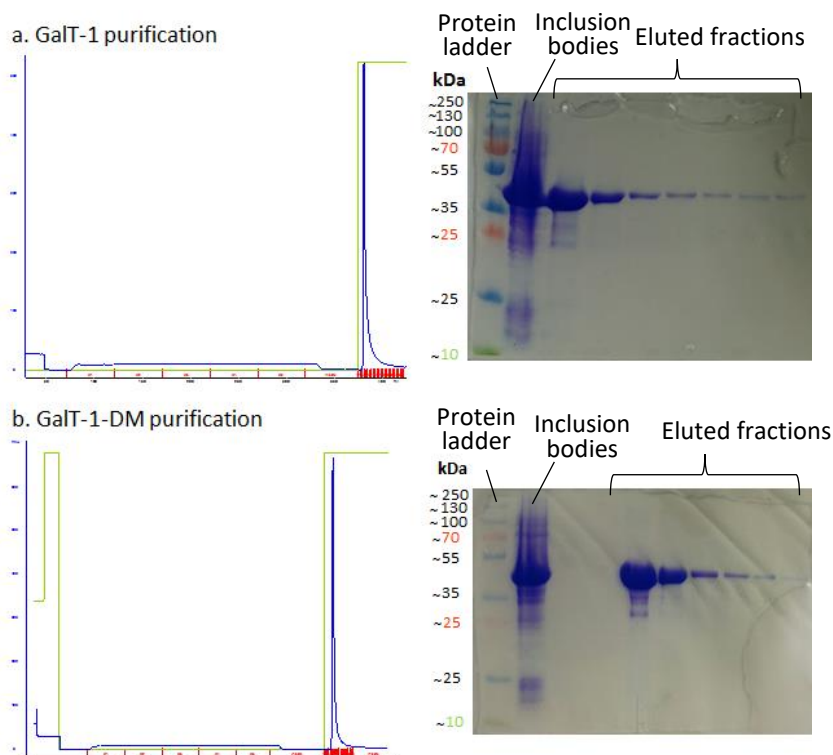
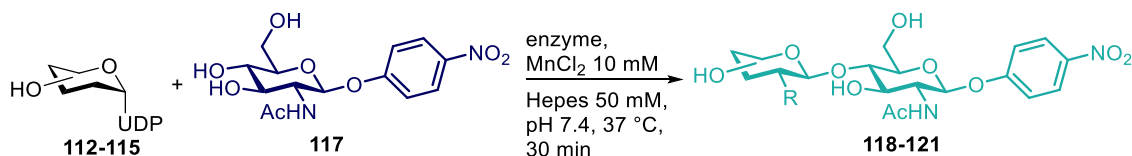


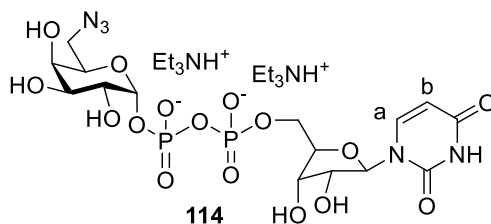
Figure 3.18 – GalT-1 and GalT-1-DM purification by HisTrap™ HP column. The chromatograms of purification for both GalT-1 (a.) and GalT-1-DM (b.) enzymes are reported. Fractions were analyzed by SDS-PAGE and stained by Coomassie blue G-250. The corresponding acrylamide gels are shown. Samples of enzymes prior purification were also run as a reference (second well from the left).

3.6.2.5 General procedure for GalT-1 and GalT-1-DM enzyme activity assay



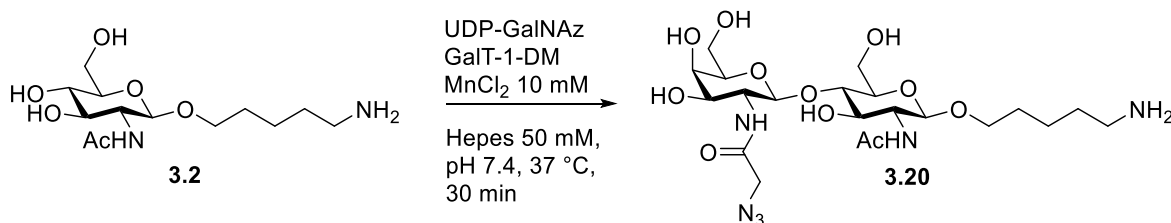
Reactions were generally performed mixing a PNP-GlcNAc **117** 20 mM solution (5 μ L, 100 nmol), a UDP-glycoside donor **112-115** 20 mM solution (5 μ L, 100 nmol) and 10 μ L of enzyme solution in 30 μ L of Hepes buffer (50 mM) pH=7.4 containing $MnCl_2$ 10 mM. Reactions were incubated at 37 $^{\circ}C$ for 30 min. The reactions were quenched adding 1 mL of CH_3CN and analyzed by UPLC-MS. The gradient elution buffers were A (water + 100 mM of ammonium formate) and B (acetonitrile). The gradient method was as follows: 0-0.5 min at 5% A, 0.5-2 min to 60% A, 2-2.5 min at 60% A, 2.5-3 min to 5% A, 3-5 min at 5% A. Total run time was 5 min and the flow rate was set at 500 μ L/min. Chromatograms were analysed for UV absorbance at 298 nm measuring the peak areas of both starting material **117** and the products **118-121**.

3.6.2.6 Spectroscopic data of 6-azido-UDP-D-galactose **114**



The integrity of **114** was assessed by NMR and MS spectra, which were in agreement with data from the literature.⁵² ¹H NMR (500 MHz, D₂O) δ (ppm): 7.9 (d, 1 H, $J = 8$ Hz, Ha), 5.92-5.88 (m, 2 H, Hb, 1rib-H), 5.55 (dd, 1 H, $J = 7.2, 3.6$ Hz, 1-H), 4.31-4.22 (m, 2 H, 2rib-H, 3rib-H), 4.21-4.14 (m, 4 H, 4rib-H, 5-H, 5rib-H, 5'rib-H), 3.91 (d, 1 H, $J = 2.8$ Hz, 4-H), 3.83 (dd, 1 H, $J = 10.4, 3.2$ Hz, 3-H), 3.72 (dt, 1 H, 2-H), 3.49 (dd, 1 H, $J = 12.8, 7.2$ Hz, 6-H), 3.38 (dd, 1 H, $J = 12.8, 6.0$ Hz, 6'-H). ³¹P NMR (202.4 MHz, D₂O) δ (ppm): -11.16 (d, $J = 20.2$ Hz), 12.90 (d, $J = 20.2$ Hz). MS (ESI negative mode) m/z : calcd for C₁₅H₂₃N₅O₁₆P₂ 591.1; found 590.7 [M-H]⁻.

3.6.2.7 Enzymatic glycosylation with UDP-GalNAz. Synthesis of azido tagged **123**



A solution (1 mL) containing **105** (400 nmol), UDP-GalNAz H₂O solution (400 nmol) and GalT-1-DM enzyme (250 μ L of a 0.62 mg/mL solution) in 500 mM Hepes solution pH 7.4 with 10 mM MnCl₂ was incubated at 37° C overnight. Reaction progress was monitored by MALDI-TOF MS. The disaccharide **123** was purified by a SampliQ high performance graphitized carbon cartridge (1 mL) from Agilent Technologies eluting with H₂O (2x 1 mL) followed by a 6:4 H₂O/CH₃CN mixture (1 mL). The fractions were analysed by MALDI-TOF MS (m/z : calcd for C₂₁H₃₈N₆O₁₁ 550.26; found 550.68 [M+H]⁺, 572.62 [M+Na]⁺) and collected obtaining pure **123** (176.0 μ g, 80%).

3.6.2.8 General procedure for on-chip glycosylation

Using incubation chambers, slides with arrayed glycans were compartmentalized and incubated with reaction solution (100 μ L per well) of UDP-glycoside donor 1 mM in Hepes buffer 50 mM (pH 7.4) containing MnCl₂ 5 mM and 50 μ L of enzyme solution. 2 μ L of a BSA IgG-free H₂O saturated solution were added. Reactions were performed overnight at 37 °C. Slides were washed with H₂O sonicating for 30 sec and dried under a stream of argon. The reactions were monitored by printing on top of immobilized glycans with DHB solution (10 mg/mL in H₂O + 0.1 % TFA + 5% CH₃CN) and analysed.

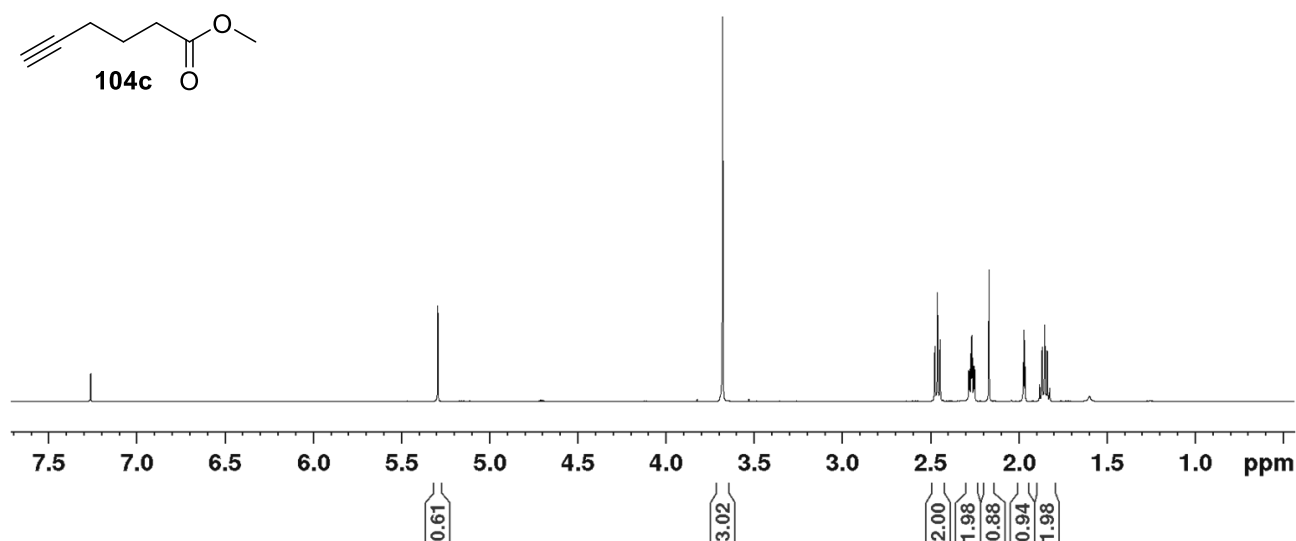
3.6.2.9 General procedure for on-chip CuAAC

Slides containing azido tagged glycans were incubated (140 μ L per well of incubation chambers) with reaction solutions prepared by adding CuSO₄·5 H₂O, THPTA, a terminal alkyne and Na ascorbate in the order to a 99:1 H₂O/DMSO mixture. Reactions were performed for 2 h at room temperature, shielding from light. Slides were washed twice with a 99:1 H₂O/DMSO mixture and finally with H₂O. The reactions were monitored printing the slides with a DHB solution (10 mg/mL in H₂O + 0.1 % TFA + 5% CH₃CN) and analysing by MALDI-TOF MS.

3.6.2.10 Synthesis of methyl hexynoate **104c**⁵⁸

To a solution of 5-hexynoic acid (200 mg, 1.78 mmol) in dry DCM (1.8 mL) under argon atmosphere camphorsulfonic acid (4 mg, 0.014 mmol) and dry methanol (600 μ L, 14.8 mmol) were added. The reaction was stirred at 50 °C for 17 h monitoring by TLC (R_f = 0.35; *n*-hexane - EtOAc, 4:1). The mixture was diluted in DCM (5 mL) and the organic phase washed with a NaHCO₃ saturated aqueous solution (2x20 mL). The aqueous phases were extracted with DCM (2x5 mL). Reunited organic phases were dried over MgSO₄ and, after filtration and evaporation of the solvent, a white precipitate was removed by centrifugation (2x5 min at 10 000 rpm) giving pure **104c** as a yellow liquid (128.4 mg, 64%). ¹H NMR (500 MHz, CDCl₃) δ (ppm): 3.70 (s, 3 H, OCH₃), 2.48 (t, J = 7.4 Hz, 2 H, 2-H), 2.29 (td, J = 6.9, 2.6 Hz, 2 H, 4-H), 1.99 (t, J = 2.6 Hz, 1 H, 6-H), 1.87 (quint, J = 7.2 Hz, 2 H, 3-H).

¹H NMR spectrum of **104c** in CDCl₃ (500 MHz)



3.6.2.11 Fluorescein printing and fluorescence detection

5-((5-aminopentyl)thioureidyl)fluorescein TFA salt solutions at different concentrations (5 mM, 1 mM, 500 μ M, 100 μ M, 50 μ M, 10 μ M, 5 μ M, 1 μ M, 500 nM, 100 nM, 50 nM, 10 nM, 5 nM, 1 nM, 500 pM, 100 pM, 50 pM, 10 pM, 5 pM, 1 pM, 500 fM, 100 fM, 50 fM, 10 fM, 5 fM, 1 fM) in phosphate buffer 300 mM, pH 8.7 were prepared and robotically printed (35 droplets, \approx 12 nL) onto hydrophobic ITO slides bearing adsorbed linkers **116**. After printing, slides were allowed to react overnight at 17 °C and controlled humidity \sim 50%. The slides were quenched by incubation with a 50 mM ethanolamine solution in 50 mM borate buffer pH 9.3 for 40 min. Slides were washed with nanopure H₂O and finally dried under argon flow. Slides were analysed by microarray scanner system (λ_{ex} = 532 nm, λ_{em} = 550-610 nm).

3.6.2.12 General procedure for enzymatic fucosylation of biantennary **152**

Solutions (15 μ L) containing biantennary **152** (1 nmol), either GDP-L-fucose or 6-azido-GDP-L-fucose **127** (20 nmol), in 80 mM Mes buffer pH 6.5 containing 20 mM MnCl₂ for CeFUT6 (6.5 μ L of 0.2 mg/mL) or in 50 mM Tris buffer pH 7.5 containing 10 mM MnCl₂ for α -1,3-fucosyltransferase from *H. pylori* solution (6.5 μ L of 0.2 mg/mL) were incubated overnight at room temperature and conversions were monitored by MALDI-TOF MS.

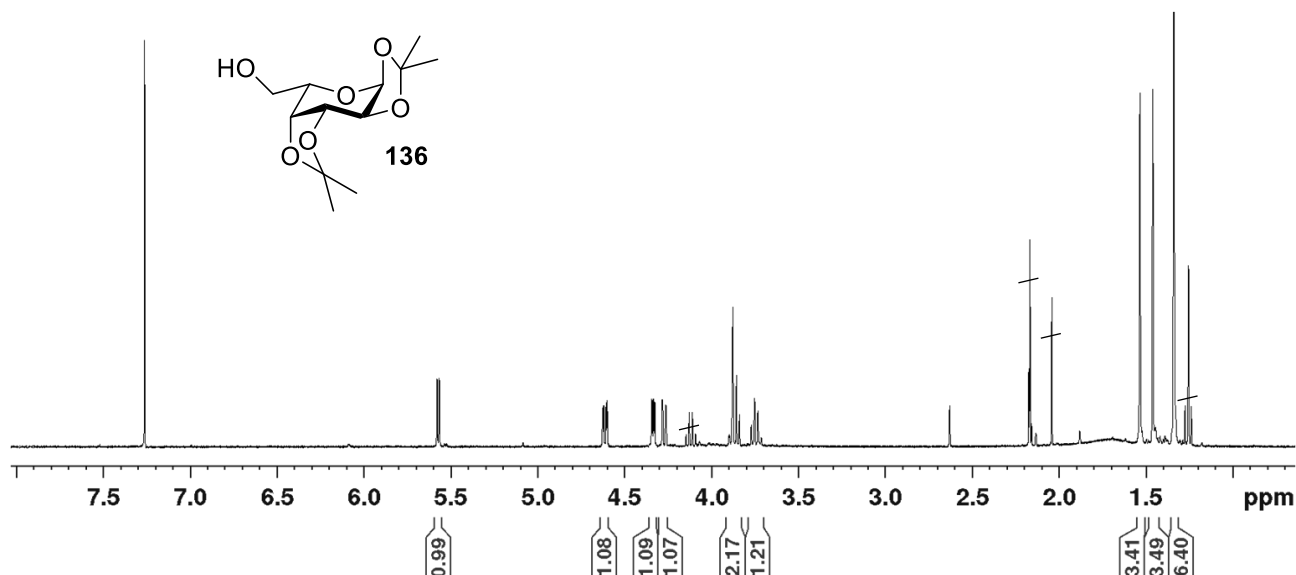
3.6.3 Synthesis of tagged GDP-L-fucose glycoside donors

Synthesis of 1,2:3,4-di-*O*-isopropylidene- α -L-galactopyranose **136**

To a suspension of polymer-bound triphenylphosphine \sim 3 mmol/g (1.39 g, 4.16 mmol) in dry acetone (60 mL, dried with 3 Å molecular sieves) a solution of iodine (1.05 g, 4.16 mmol) in dry acetone (60 mL dried with 3

Å molecular sieves) was added. The mixture was stirred under nitrogen atmosphere for 15 min at room temperature. Then, L-galactose (750 mg, 2.08 mmol) was added as a solid and the reaction was stirred under nitrogen atmosphere at room temperature for 30 min. Complete conversion was assessed by TLC (eluent: *n*-hexane - EtOAc, 2:1) and the reaction was quenched adding Na₂CO₃ (440 mg, 4.2 mmol) to the mixture, which was stirred at room temperature for 15 min. The reaction mixture was filtered by a fritted funnel, concentrated by evaporation of the solvent and purified by flash chromatography (eluent: *n*-hexane - EtOAc, 2:1; R_f = 0.21) affording pure **136** as a colourless syrup (820 mg, 75%). ¹H NMR spectroscopic data were in agreement with those previously reported in the literature.⁷⁰ ¹H NMR (400 MHz, CDCl₃) δ (ppm): 5.57 (d, *J* = 5.0 Hz, 1 H, 1-H), 4.62 (dd, *J* = 8.0, 2.3 Hz, 1 H, 3-H), 4.34 (dd, *J* = 5.0, 2.4 Hz, 1 H, 2-H), 4.23 (dd, *J* = 8.2, 1.0 Hz, 1 H, 4-H), 3.91-3.83 (m, 2 H, 5-H, 6-H), 3.78-3.711 (m, 1 H, 6-H'), 1.54 (s, 3 H, OCCH₃), 1.46 (s, 3 H, OCCH₃), 1.34 (s, 6 H, OC(CH₃)₂).

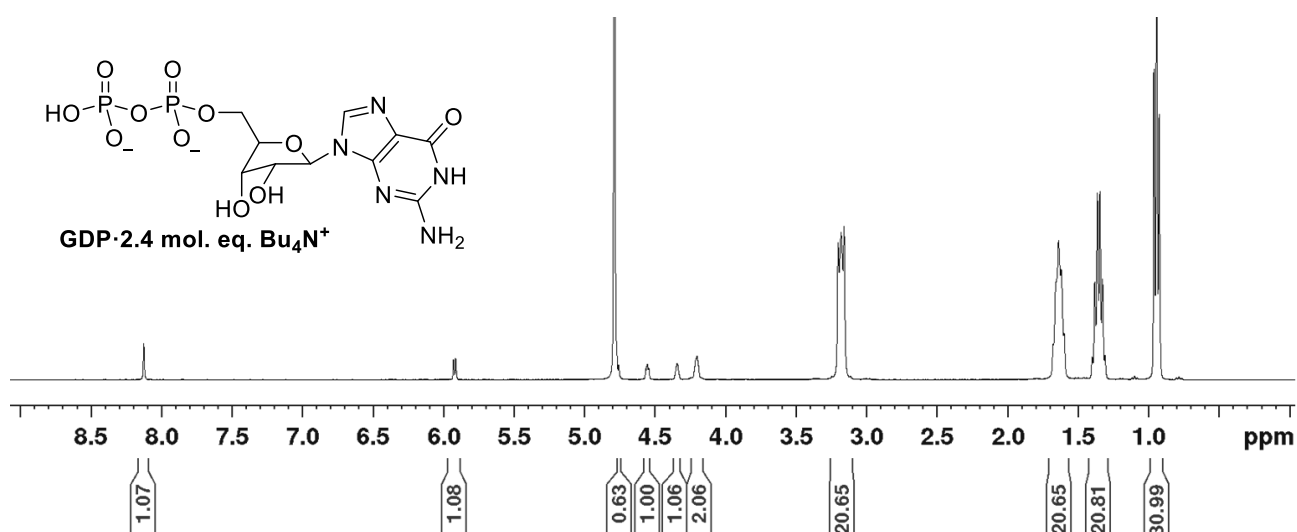
¹H NMR spectrum of **136** in CDCl₃ (400 MHz) Minor impurities ascribed to EtOAc are visible in the reported spectrum.



Titration of GDP sodium salt

GDP sodium salt (150 mg, ~0.322 mmol) was dissolved in the minimum amount of water and loaded onto a column packed with 35 mL of activated ion exchange acidic resin (Amberlite IR 120 H, 1.8 meq/mL). The compound was eluted with H₂O and the eluted fractions were analysed by deposition on TLC monitoring with UV at 254 nm. Fractions containing GDP as free acid were reunited and titrated to pH 6 with a Bu₄NOH aqueous solution monitoring with a pH-meter. First 100 mL of a 40% w/v Bu₄NOH solution were added, followed by few drops of a 100 fold more concentrated Bu₄NOH solution. The resulting solution was finally freeze dried affording GDP·2.4 mol. eq. Bu₄N⁺ as a white solid. ¹H NMR (400 MHz, D₂O) δ (ppm): 8.13 (s, 1 H, Ha), 5.92 (d, *J* = 6.1 Hz, 1 H, 1-H), 4.76 (m, 1 H, 2-H), 4.55 (t, *J* = 4.1 Hz, 1 H, 3-H), 4.34 (bs, 1 H, 4-H), 4.22 (bs, 2 H, 5-H, 5-H'), 3.22-3.14 (m, Bu₄N⁺), 1.69-1.58 (m, Bu₄N⁺), 1.36 (sextet, *J* = 7.2 Hz, Bu₄N⁺), 0.94 (t, *J* = 7.2 Hz, Bu₄N⁺).

^1H NMR spectrum of GDP·2.4 mol. eq. Bu_4N^+ in D_2O (400 MHz)

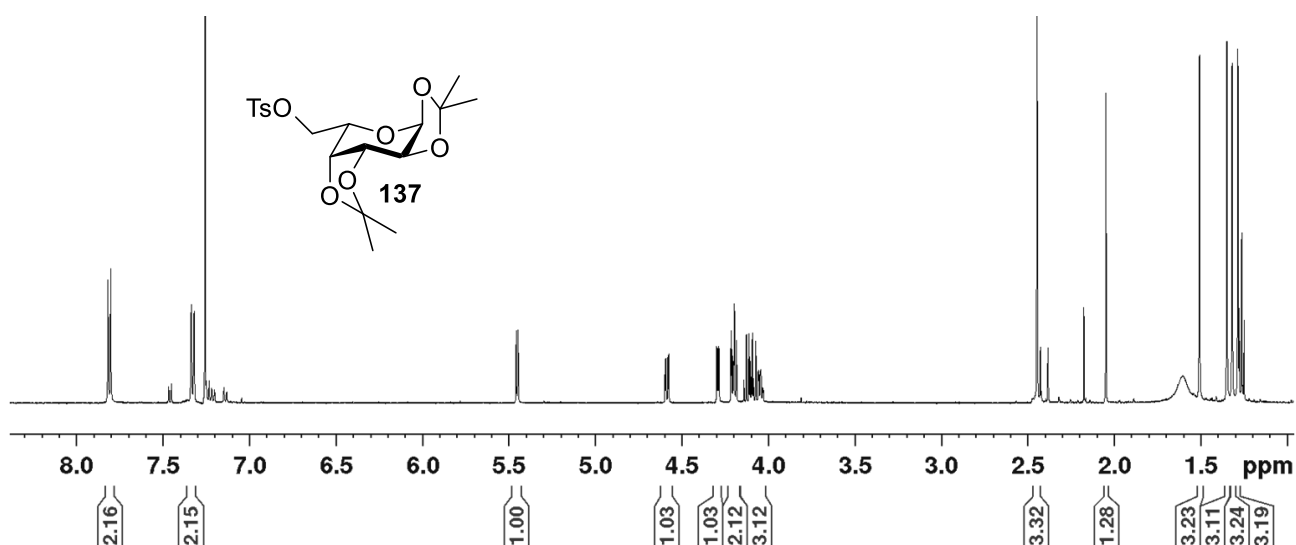


3.6.3.1 Preparation of 6-azido-GDP- β -L-fucose 127

Synthesis of 1,2:3,4-di-*O*-isopropylidene-6-*O*-(*p*-toluenesulfonyl)- α -L-galactopyranose 137

Bis-acetonide **136** (300 mg, 1.16 mmol) was dissolved in dry pyridine (2.3 mL) under argon atmosphere and cooled to 0 °C. Then, tosyl chloride (550 mg, 2.9 mmol) was added and the reaction stirred at room temperature for 4 h monitoring by TLC (eluent: *n*-hexane - EtOAc, 3:1). The reaction was diluted with EtOAc (20 mL) and the organic phase washed with a saturated solution of CuSO_4 (2x20 mL), H_2O (2x10 mL) and brine (2x10 mL). The organic phase was dried on anhydrous MgSO_4 and, after filtration and evaporation of the solvent, purified by flash chromatography (eluent: *n*-hexane - EtOAc, 8:2; $R_f = 0.44$) affording **137** as a colourless syrup (370 mg, 78%). ^1H NMR spectroscopic data were in agreement with those previously reported in the literature. ^1H NMR (500 MHz, CDCl_3) δ (ppm): 7.81 (d, $J = 7.9$ Hz, 2 H, Ar), 7.33 (d, $J = 7.9$ Hz, 2 H, Ar), 5.45 (d, $J = 5.0$ Hz, 1 H, 1-H), 4.59 (dd, $J = 7.9, 2.6$ Hz, 1 H, 3-H), 4.30 (dd, $J = 4.9, 2.5$ Hz, 1 H, 2-H), 4.22-4.18 (m, 2 H, 4-H, 6-H), 4.11-4.03 (m, 2 H, 5-H, 6-H'), 1.50 (s, 3 H, OCCH_3), 1.34 (s, 3 H, OCCH_3), 1.31 (s, 3 H, OCCH_3), 1.28 (s, 3 H, OCCH_3).

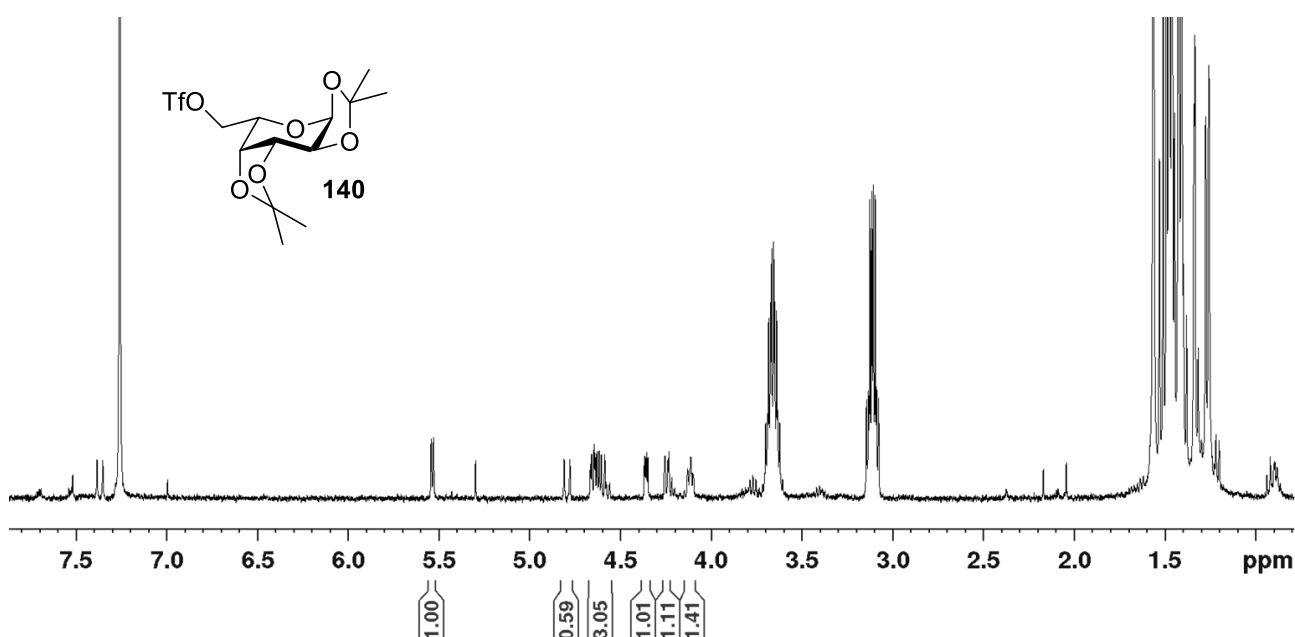
^1H NMR spectrum of **137** in CDCl_3 (400 MHz). Minor impurities ascribed to TsCl are visible in the reported spectrum.



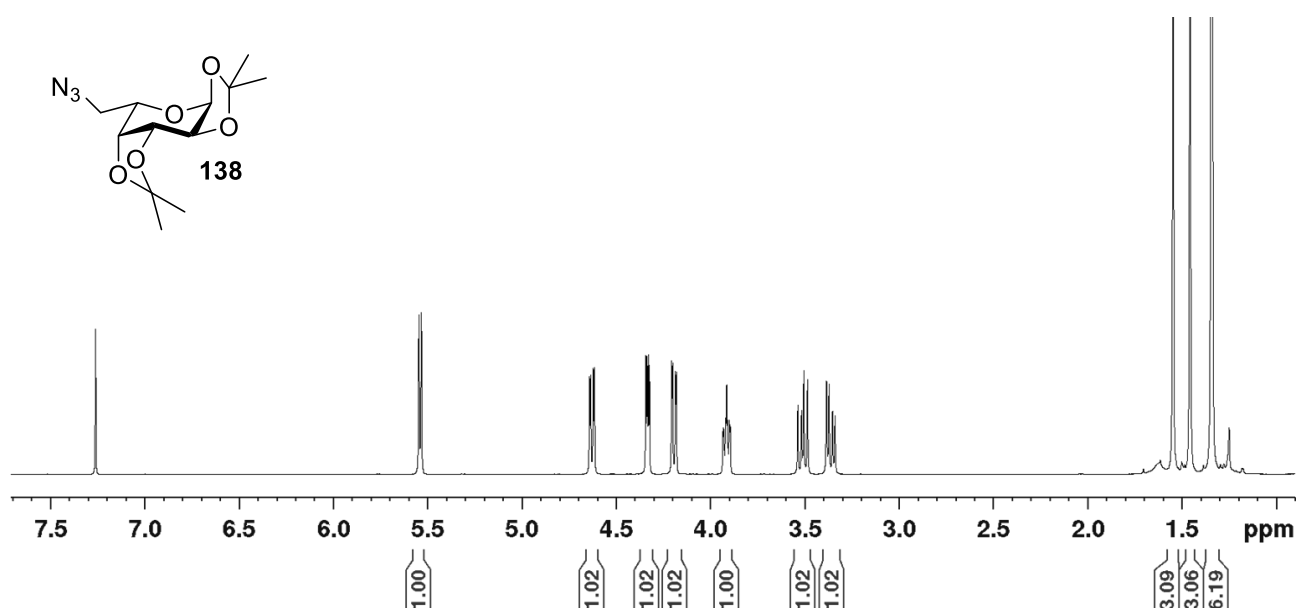
Synthesis of 1,2:3,4-di-*O*-isopropylidene-6-*O*-(trifluoromethanesulfonyl)- α -L-galactopyranose **140⁸⁶**

To a stirred solution of bis-acetonide **136** (360 mg, 1.38 mmol) in dry DCM (6.6 mL) under nitrogen atmosphere distilled DIPEA (1.3 mL) was added. The reaction was cooled to 0 °C and Tf₂O (470 μ L, 2.77 mmol) was slowly added. The reaction was stirred at 0 °C for 15 min monitoring by TLC (eluent: *n*-hexane - EtOAc, 7:3, R_f = 0.63), which showed partial conversion. Complete conversion was achieved by a second addition of Tf₂O (235 μ L, 1.38 mmol) stirring the reaction at 0 °C for 15 min. The reaction was quenched with cold H₂O (25 mL, 0 °C) and the product was extracted with cold DCM (2x40 mL, 0 °C). Reunited organic phases were dried over anhydrous Na₂SO₄. Filtration and evaporation of the solvent afforded crude **140**, which was pure enough to be employed in the next synthetic step without further purification. ¹H NMR (400 MHz, CDCl₃) δ (ppm): 5.54 (d, *J* = 5.2 Hz, 1 H, 1-H), 4.65 (dd, *J* = 8.1, 2.9 Hz, 1 H), 4.67-4.58 (m, 3 H), 4.36 (dd, *J* = 4.8, 2.6 Hz, 1 H), 4.25 (dd, *J* = 7.6, 1.8 Hz, 1 H), 4.14-4.09 (m, 1 H), 1.48 (s, 3 H, OCCH₃), 1.46 (s, 3 H, OCCH₃), 1.43 (s, 3 H, OCCH₃), 1.41 (s, 3 H, OCCH₃).

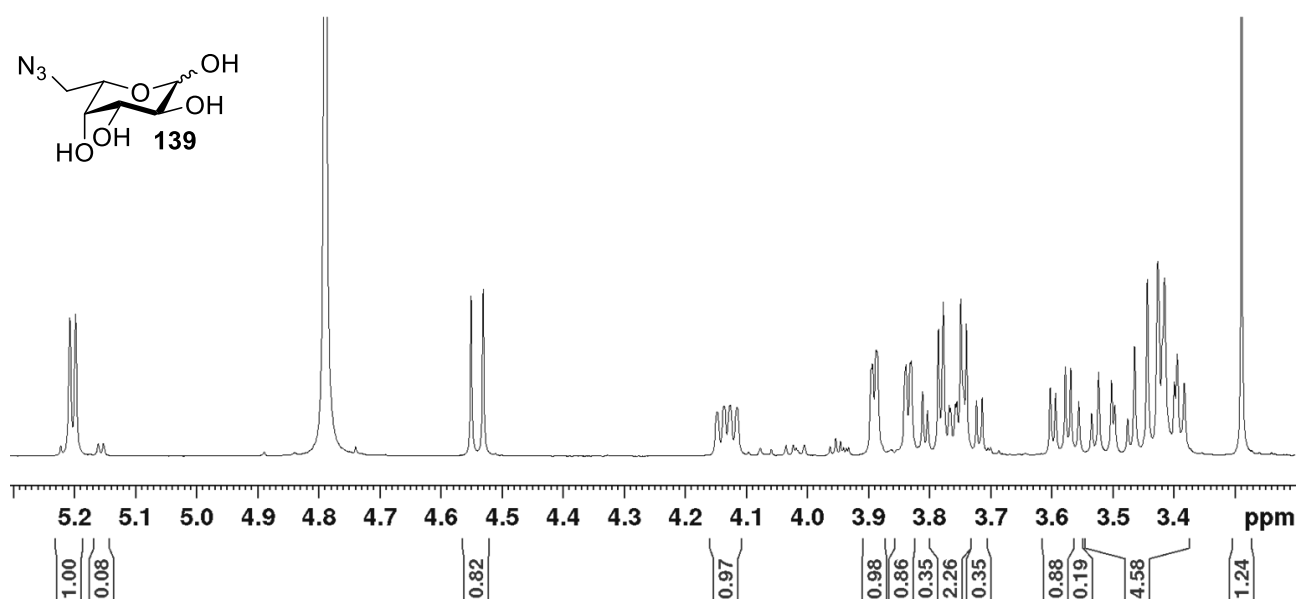
¹H NMR spectrum of crude **140** in CDCl₃ (400 MHz)

**Synthesis of 1,2:3,4-di-*O*-isopropylidene-6-azido- α -L-fucopyranose **138****

Crude **140** (42 mg, 0.11 mmol) was dissolved in dry DMF (400 μ L) under nitrogen atmosphere, NaN₃ (34.5 mg, 0.53 mmol) was added and the reaction stirred for 4 h at room temperature. Complete conversion was assessed by TLC (eluent: *n*-hexane - EtOAc, 10:1). H₂O (6 mL) was added and the product extracted with EtOAc (3x4 mL). The organic phase was dried over Na₂SO₄. After filtration and evaporation of the solvent the crude was purified by flash chromatography (eluent: *n*-hexane - EtOAc, 15:1, R_f = 0.16) affording pure **138** as a colourless syrup (18.5 mg, 61% over two steps starting from **136**). ¹H NMR spectroscopic data were in agreement with those previously reported in the literature.⁷⁰ ¹H NMR (400 MHz, CDCl₃) δ (ppm): 5.54 (d, *J* = 5.0 Hz, 1 H, 1-H), 4.63 (dd, *J* = 8.0, 2.4 Hz, 1 H, 3-H), 4.33 (dd, *J* = 5.0, 2.4 Hz, 1 H, 2-H), 4.19 (dd, *J* = 8.0, 2.0 Hz, 1 H, 4-H), 3.91 (ddd, *J* = 8.0, 5.5, 1.9 Hz, 1 H, 5-H), 3.51 (dd, *J* = 12.8, 8.0 Hz, 1 H, 6-H), 3.36 (dd, *J* = 12.5, 5.3 Hz, 1 H, 6-H'), 1.54 (s, 3 H, OCCH₃), 1.45 (s, 3 H, OCCH₃), 1.34 (s, 3 H, OCCH₃), 1.33 (s, 3 H, OCCH₃).

¹H NMR spectrum of 138 in CDCl₃ (400 MHz)**Synthesis of 6-azido-L-fucose 139**

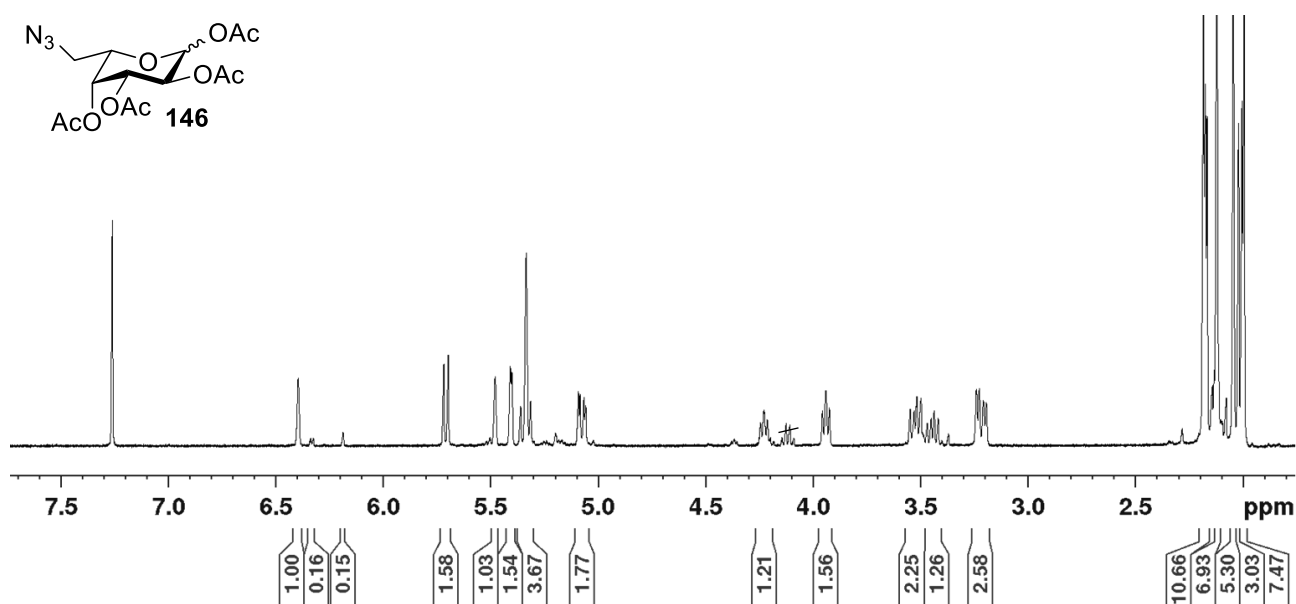
Compound **138** (64.2 mg, 0.219 mmol) was dissolved at 0 °C in a 9:1 TFA:H₂O solution (2.2 mL), which was stirred at 0 °C for 2 h monitoring by TLC (eluent: DCM - MeOH, 8:2, $R_f = 0.48$). The reaction mixture was diluted and co-evaporated with toluene (3x3 mL) and finally co-evaporated with Et₂O (3x3 mL) affording crude **139**, which was used in the following step without further purification. ¹H NMR spectroscopic data were in agreement with those previously reported in the literature.⁷⁰ ¹H NMR (400 MHz, D₂O) δ (ppm): 5.20 (d, $J = 3.9$ Hz, 1 α -H), 4.54 (d, $J = 7.9$ Hz, 1 β -H), 4.13 (dd, $J = 8.7, 4.7$ Hz), 3.89 (dd, $J = 3.2, 0.7$ Hz), 3.83 (dd, $J = 3.3, 0.7$ Hz), 3.81 (d, $J = 3.3$ Hz), 3.78 (d, $J = 3.3$ Hz), 3.76 (dd, $J = 3.8, 0.9$ Hz), 3.74 (d, $J = 3.6$ Hz), 3.71 (d, $J = 4.0$ Hz), 3.60-3.37 (m).

¹H NMR spectrum of 139 in D₂O (400 MHz)**Synthesis of 1,2,3,4-tetra-O-acetyl-6-azido-L-fucopyranose 146**

6-Azido-L-fucose **139** (41 mg, 0.20 mmol) was dissolved in dry pyridine (370 μ L) followed by addition of Ac₂O (370 μ L). A catalytic amount of DMAP was added (0.8 mg, 3 mol%) and the reaction was stirred under

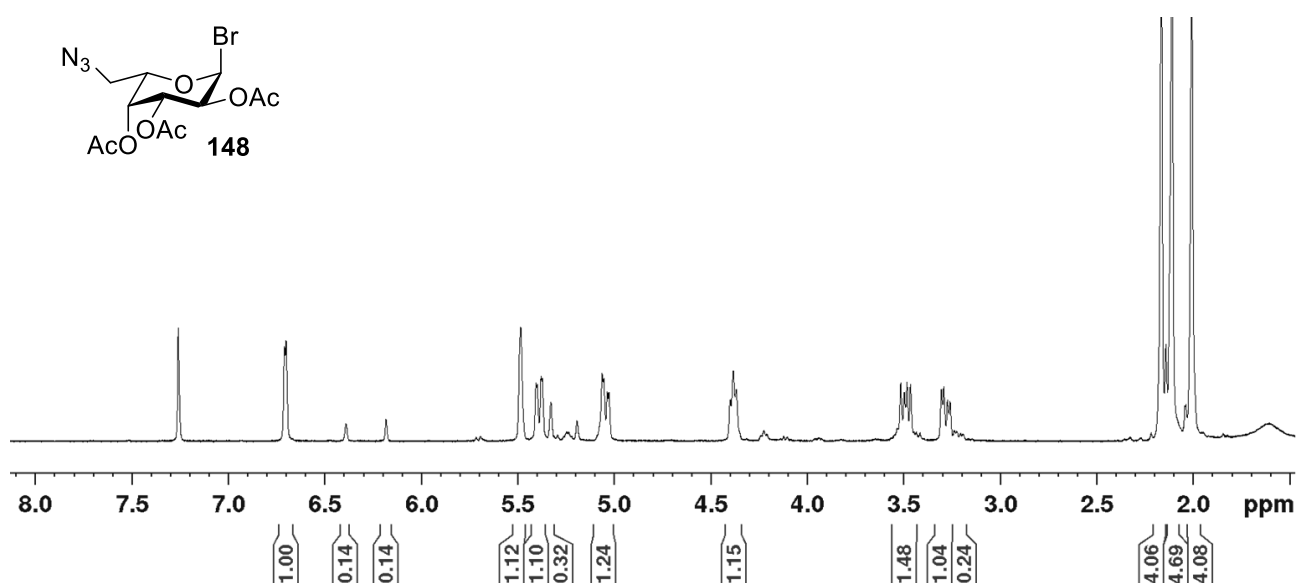
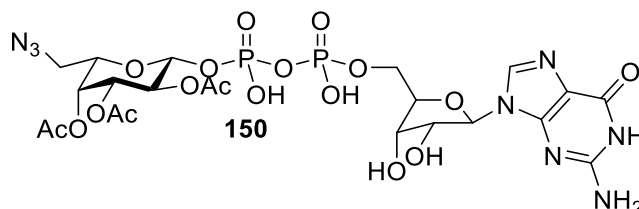
nitrogen atmosphere at room temperature for 18 h monitoring by TLC (eluent: *n*-hexane - EtOAc, 6:4). The mixture was diluted with EtOAc (10 mL) and washed with 1 M HCl aqueous solution (2x20 mL), saturated NaHCO₃ solution (2x20 mL), brine (10 mL). Aqueous phases were extracted with EtOAc (2x10 mL each) and reunited organic phases dried over anhydrous Na₂SO₄. Purification by flash chromatography (eluent: *n*-hexane - EtOAc, 7:3, *R_f* = 0.33) afforded pure **146** (52.8 mg, 71% over two steps from **138**) as a α/β anomeric mixture of pyranosidic forms (89%, 1.6:1 β/α ratio) along with the corresponding α,β mixture of furanose isomers (11%, 1:1 β/α ratio). ¹H NMR spectroscopic data were in agreement with those previously reported in the literature.⁷⁰ ¹H NMR (400 MHz, CDCl₃) δ (ppm): 6.40 (s, 1 α -HPyr), 6.33 (d, *J* = 4.6 Hz, 1 β -HFur), 6.19 (s, 1 α -HFur), 5.71 (d, *J* = 8.0 Hz, 1 β -HPyr), 5.48 (s), 5.40 (d, *J* = 3.5 Hz), 5.38-5.30 (m), 5.08 (dd, *J* = 10.4, 3.4 Hz), 4.23 (t, *J* = 6.7 Hz), 3.94 (t, *J* = 6.7 Hz), 3.57-3.41 (m), 3.22 (dd, *J* = 12.8, 5.3 Hz), 2.18, 2.17, 2.12, 2.05, 2.04, 2.02, 2.00, 1.99 (s, CH₃CO).

¹H NMR spectrum of **146** in CDCl₃ (400 MHz) Minor impurities ascribed to EtOAc are visible in the reported spectrum.



Synthesis of 2,3,4-tri-*O*-acetyl-6-azido- α -L-fucopyranosyl bromide **148**

Compound **146** (19.7 mg, 0.054 mmol) was dissolved in a 9:1 DCM/EtOAc mixture (600 μ L). DCM was previously distilled from CaH₂ and EtOAc was dried with 3 Å molecular sieves. The solution was cooled to 0 °C under nitrogen atmosphere and TiBr₄ was added (63 mg, 0.171 mmol). The reaction was stirred at 40 °C with microwave irradiation for 1.5 h. Almost complete conversion was assessed by TLC (eluent: *n*-hexane - EtOAc, 8:2, *R_f* = 0.26) and the reaction was quenched adding AcONa (61.9 mg, 0.755 mmol) and stirring for 15 min at room temperature. Cold H₂O (15 mL, 0 °C) was added and the product extracted with cold DCM (2x10 mL, 0 °C). Organic phases were then extensively washed with cold H₂O (5x10 mL, 0 °C) and dried over anhydrous Na₂SO₄. Filtration and evaporation of the solvent afforded crude **148** that was employed in the following synthetic step without further purification. ¹H NMR spectroscopic data were in agreement with those previously reported in the literature.⁷⁷ ¹H NMR (400 MHz, CDCl₃) δ (ppm): 6.71 (d, *J* = 3.7 Hz, 1 H, 1-H), 5.49 (s, 1 H, 4-H), 5.39 (dd, *J* = 10.5, 2.6 Hz, 1 H, 3-H), 5.05 (dd, *J* = 10.5, 3.7 Hz, 1 H, 2-H), 4.39 (t, *J* = 6.3 Hz, 1 H, 5-H), 3.49 (dd, *J* = 12.9, 7.5 Hz, 1 H, 6-H), 3.28 (dd, *J* = 12.7, 5.1 Hz, 1 H, 6-H'), 2.17 (s, 3 H, CH₃CO), 2.11 (s, 3 H, CH₃CO), 2.01 (s, 6 H, CH₃CO).

¹H NMR spectrum of 148 in CDCl₃ (400 MHz)**Synthesis of 2,3,4-tri-*O*-acetyl-6-azido-GDP-β-L-fucose 150**

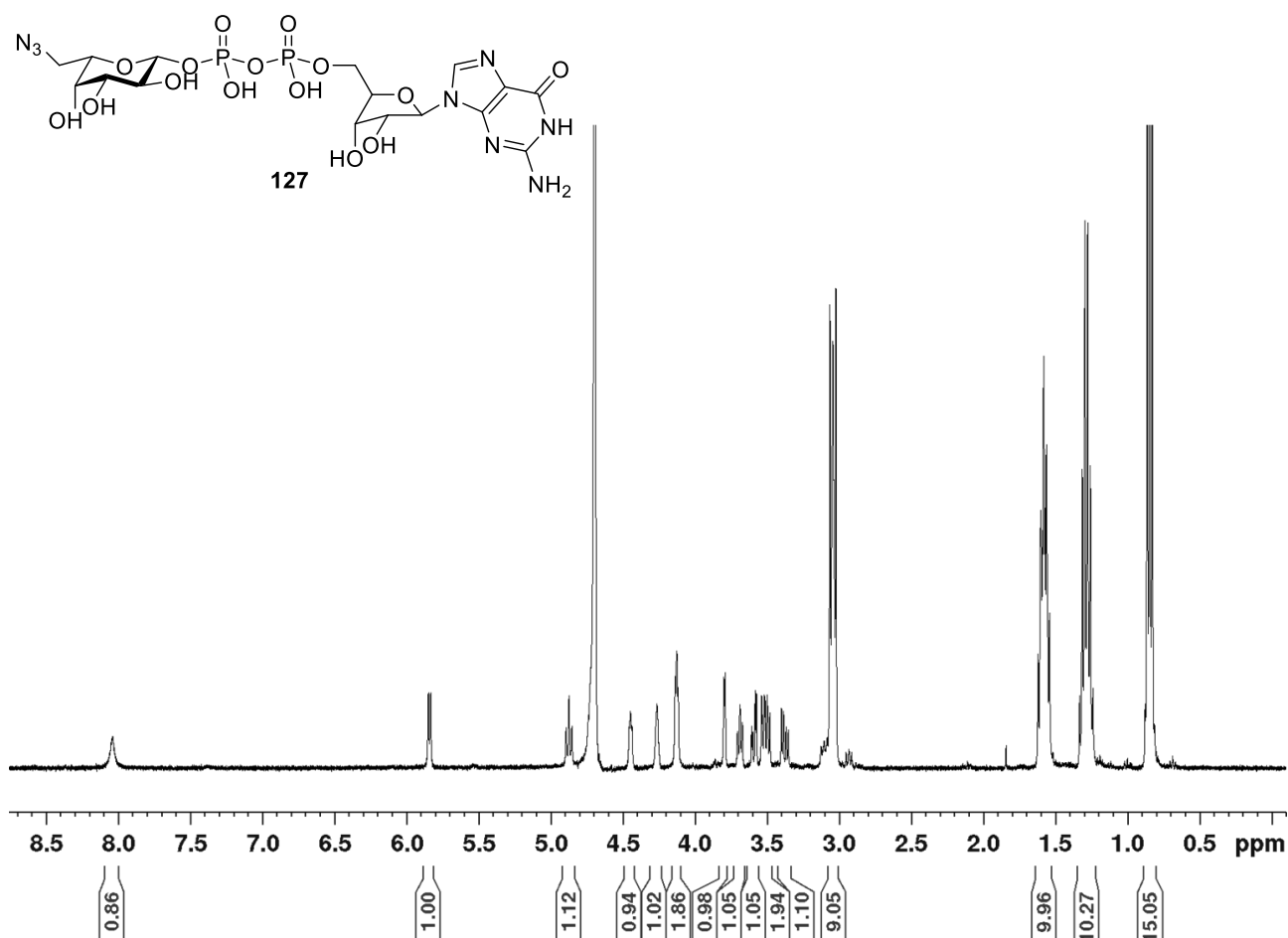
Titrated GDP·2.4 mol. eq. Bu₄N⁺ (46 mg, 0.045 mmol) was dissolved in distilled CH₃CN (1.0 mL) in a Schlenk tube under nitrogen atmosphere. Distilled Et₃N (7 μL, 0.05 mmol) and powdered 3 Å molecular sieves were added and the mixture was stirred at room temperature for 15 min. A solution of fucosyl bromide **148** (19.7 mg, 0.05 mmol) in dry CH₃CN (1.5 mL) was then cannulated into the Schlenk tube and the reaction stirred at 80 °C for 30 min. Complete conversion of fucosyl bromide **148** was assessed by TLC (eluent: *n*-hexane – EtOAc, 7:3, R_f = 0.36). Molecular sieves were removed by filtration over a celite pad and evaporation of the solvent afforded crude **150** that was analysed by HPLC (Waters Atlantis T3 5 μm 4.6x100 mm column; eluents A = aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, adjusted to pH 4 with H₃PO₄ solution), B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-10 min at 40% A, 10-12 min to 0% A, eluting at 1 mL/min, monitoring at 254 nm; t_R = 8.3 min). The corresponding HPLC trace is reported in the discussion of Chapter 3 at page 159. MS (ESI) *m/z*: calcd for C₂₂H₃₀N₈O₁₈P₂ 756.47; found 754.97 [M-H].

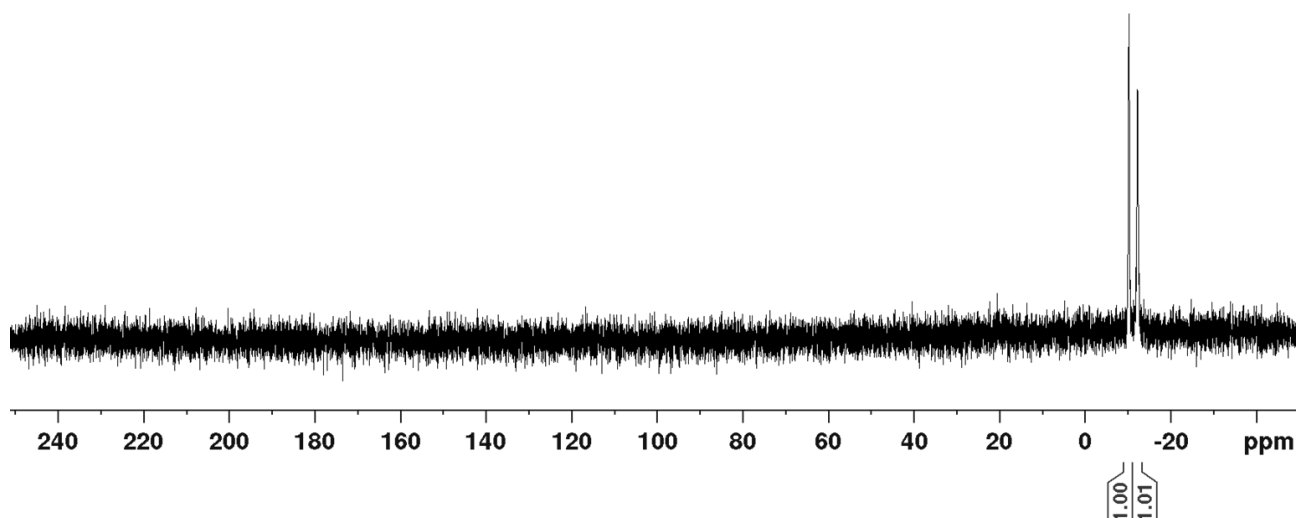
Synthesis of 6-azido-GDP-β-L-fucose 127

Crude **150** (0.045 mmol, assuming a 100% yield for the step above) was dissolved in H₂O (1.5 mL). Distilled Et₃N (16 μL) was added reaching pH 9.6 followed by addition of alkaline phosphatase (50 μL of 1 U/μL solution). The reaction was performed at 28 °C for 20 h and enzymatic hydrolysis of GDP to guanosine was monitored by HPLC (Waters Atlantis T3 5 μm 4.6x100 mm column; eluents A = aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, adjusted to pH 4 with H₃PO₄ solution), B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-10 min at 40% A, 10-12 min to 0% A, eluting at 1 mL/min, monitoring at 254 nm; t_R (GDP) = 6.5 min, t_R (guanosine) = 1.4 min). The solvent was evaporated and the crude dissolved in a 2:2:1 H₂O/MeOH/Et₃N (distilled) mixture (1.5 mL) and stirred at room temperature for 16 h monitoring by HPLC (Waters Atlantis T3 5 μm 4.6x100 mm column; eluents A = aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, adjusted to pH 4 with H₃PO₄ solution), B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-

10 min at 40% A, 10-12 min to 0% A, eluting at 1 mL/min, monitoring at 254 nm; $t_R = 6.9$ min). The solvent was evaporated and the crude purified by automated reverse phase chromatography (Biotage SNAP C18 12g cartridge, gradient elution: from 100% H₂O + 10 mM (Bu₃NH)HCO₃ to 40% H₂O + 10 mM (Bu₃NH)HCO₃ - 60% CH₃CN). The 10 mM (Bu₃NH)HCO₃ aqueous buffer was prepared adding Bu₃N (2.38 mL, 10 mmol) to H₂O (1 L), stirring at 0 °C and bubbling CO₂ for 5 h until pH ~ 5 was reached. Eluted fractions were analysed by HPLC with the conditions reported above. The fractions containing the peak at $t_R = 6.9$ min were collected and the desired product was obtained as a tributylammonium salt, 6-azido-GDP-L-fucose · 1.63 eq. Bu₃NH⁺ **127** (5.62 mg, 11% from **146**). The purity of the compound was assessed by HPLC analysis. HPLC traces are reported in the discussion of Chapter 3 at page 159. Spectroscopic data were in agreement with those previously reported in the literature.⁷⁷ ¹H NMR (400 MHz, D₂O) δ (ppm): 8.13 (s, 1 H, Ha), 5.93 (d, $J = 6.4$ Hz, 1 H, 1-H_R), 4.97 (t, $J = 7.7$ Hz, 1 H, 1-H_F), 4.82 (1 H, 2-H_R), 4.55-4.53 (m, 1 H, 3-H_R), 4.35 (bs, 1 H, 4-H_R), 4.23-4.21 (m, 2 H, 5-H_R, 5-H_F), 3.89 (d, $J = 3.1$ Hz, 1 H, 4-H_F), 3.78 (t, $J = 6.9$ Hz, 1 H, 5-H_F), 3.68 (dd, $J = 10.1, 3.3$ Hz, 1 H, 3-H_F), 3.65-3.57 (m, 2 H, 2-H_F, 6-H_F), 3.47 (dd, $J = 12.8, 5.8$ Hz, 1 H, 6-H_F), 3.18-3.11 (m, Bu₃NH⁺), 1.73-1.63 (m, Bu₃NH⁺), 1.33 (sextet, $J = 7.4$ Hz, Bu₃NH⁺), 0.94 (t, $J = 7.4$ Hz, Bu₃NH⁺). ³¹P NMR (162 MHz, D₂O) δ (ppm): -10.16, -12.09. MS (ESI) m/z : calcd for C₁₆H₂₄N₈O₁₅P₂ 630.08; found 628.86 [M-H]⁻, 650.83 [M-2H+Na]⁻.

¹H NMR spectrum of 6-azido-GDP- β -L-fucose **127** in D₂O (400 MHz)

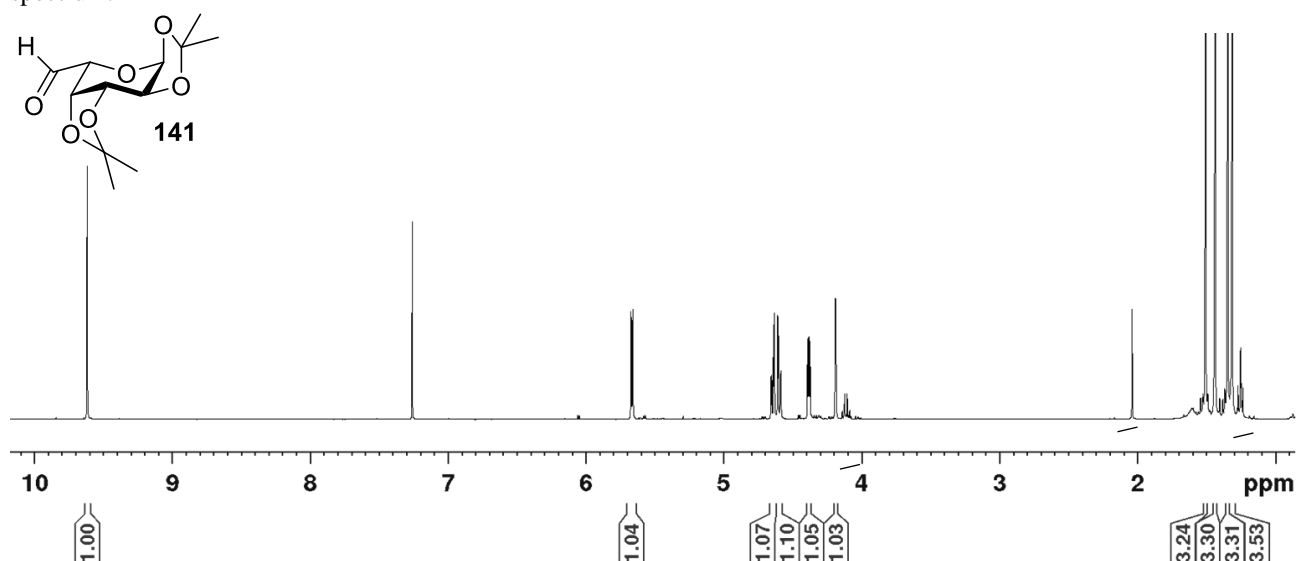


³¹P NMR spectrum of 6-azido-GDP-β-L-fucose 127 in D₂O (162 MHz)**3.6.3.2 Preparation of 5-ethynyl-GDP-L-fucose 128****Synthesis of 1,2:3,4-di-O-isopropylidene-α-L-galacto-hexadialdo-1,5-pyranose 141**

Bis-acetonide **136** (24.7 mg, 0.095 mmol) was dissolved under nitrogen atmosphere in EtOAc (1 mL) dried with 4 Å molecular sieves. Then, freshly prepared IBX (79.7 mg, 0.285 mmol) was added and the mixture stirred at refluxing temperature for 5 h. Partial conversion was observed by TLC (eluent: *n*-hexane - EtOAc, 3:1), thus IBX was further added (39.8 mg, 0.142 mmol) and the mixture stirred for 2 h reaching complete conversion. A saturated NaHCO₃ solution (4 mL) was added and the product extracted with EtOAc (3x6 mL). Organic phases were dried over Na₂SO₄. Filtration and evaporation of the solvent afforded a crude that was purified by flash chromatography (eluent: *n*-hexane - EtOAc, 3:1, R_f = 0.25) to obtain pure **141** (14.8 mg, 60 %) as a colourless oil. Spectroscopic data were in agreement with those previously reported in the literature.⁷¹

¹H NMR (400 MHz, CDCl₃) δ (ppm): 9.62 (s, 1 H, CHO), 5.67 (d, *J* = 5.0 Hz, 1 H, 1-H), 4.64 (dd, *J* = 7.9, 2.5 Hz, 1 H, 3-H), 4.59 (dd, *J* = 7.8, 2.2 Hz, 1 H, 4-H), 4.38 (dd, *J* = 5.0, 2.5 Hz, 1 H, 2-H), 4.19 (d, *J* = 2.1 Hz, 1 H, 5-H), 1.51 (s, 3 H, OCCH₃), 1.44 (s, 3 H, OCCH₃), 1.35 (s, 3 H, OCCH₃), 1.32 (s, 3 H, OCCH₃).

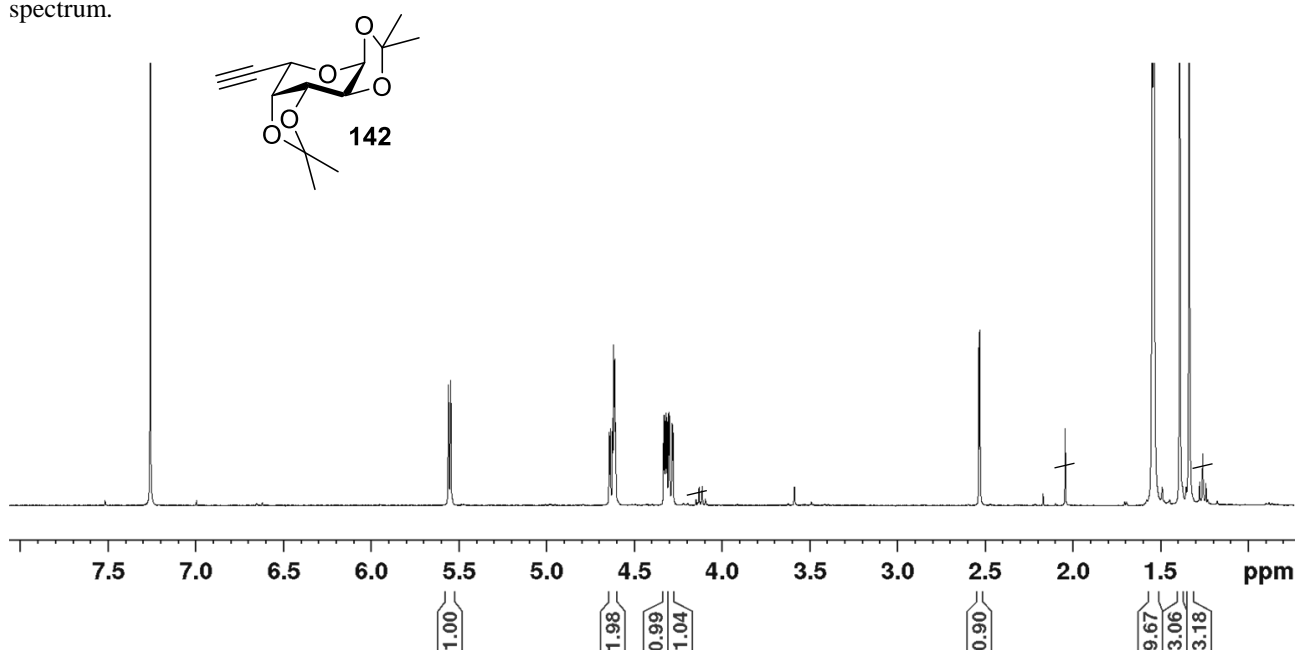
¹H NMR spectrum of **141** in CDCl₃ (400 MHz) Minor impurities ascribed to EtOAc are visible in the reported spectrum.



Synthesis of 1,2:3,4-di-*O*-isopropylidene-5-ethynyl- α -L-fucopyranose **142**

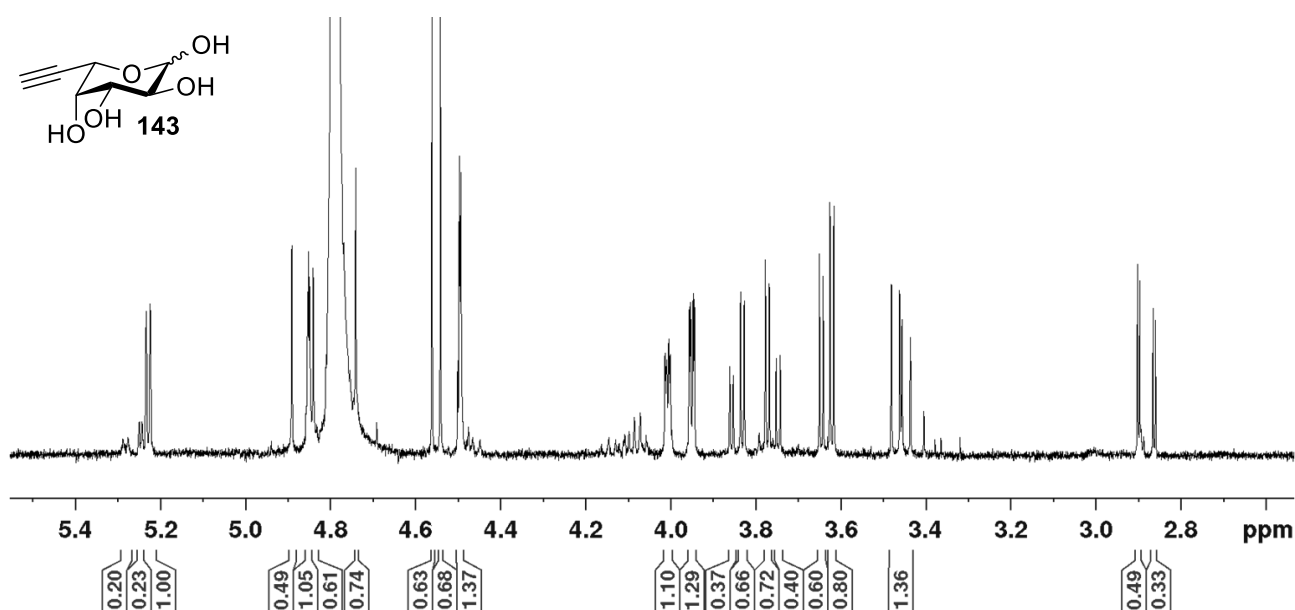
To a solution of aldehyde **141** (183 mg, 0.708 mmol) under nitrogen atmosphere in dry MeOH (4.7 mL) K_2CO_3 (196 mg, 1.415 mmol) was added. The mixture was cooled to 0 °C and dimethyl(1-diazo-2-oxopropyl)phosphonate (127 μ L, 0.849 mmol) was added. The reaction was stirred at room temperature for 4.5 h monitoring by TLC (eluent: *n*-hexane - EtOAc, 4:1). H_2O (40 mL) was added, the product was extracted with EtOAc (3x60 mL) and organic phases were dried over Na_2SO_4 . After filtration and evaporation of the solvent, the crude was purified by flash chromatography (eluent: *n*-hexane - EtOAc, 4:1, R_f = 0.50) affording pure **142** (143.5 mg, 80%) as a colourless oil. 1H NMR spectroscopic data were in agreement with those previously reported in the literature.⁷¹ 1H NMR (400 MHz, $CDCl_3$) δ (ppm): 5.55 (d, J = 4.9 Hz, 1 H, 1-H), 4.63 (dd, J = 7.8, 2.6 Hz, 1 H, 3-H), 4.61 (t, J = 2.2 Hz, 1 H, 5-H), 4.32 (dd, J = 5.1, 2.5 Hz, 1 H, 2-H), 4.29 (dd, J = 7.8, 2.0 Hz, 1 H, 4-H), 2.54 (d, J = 2.4 Hz, 1 H, $H-C\equiv C$), 1.54 (s, 3 H, $OCCH_3$), 1.53 (s, 3 H, $OCCH_3$), 1.39 (s, 3 H, $OCCH_3$), 1.33 (s, 3 H, $OCCH_3$).

1H NMR spectrum of **142** in $CDCl_3$ (400 MHz) Minor impurities ascribed to EtOAc are visible in the reported spectrum.



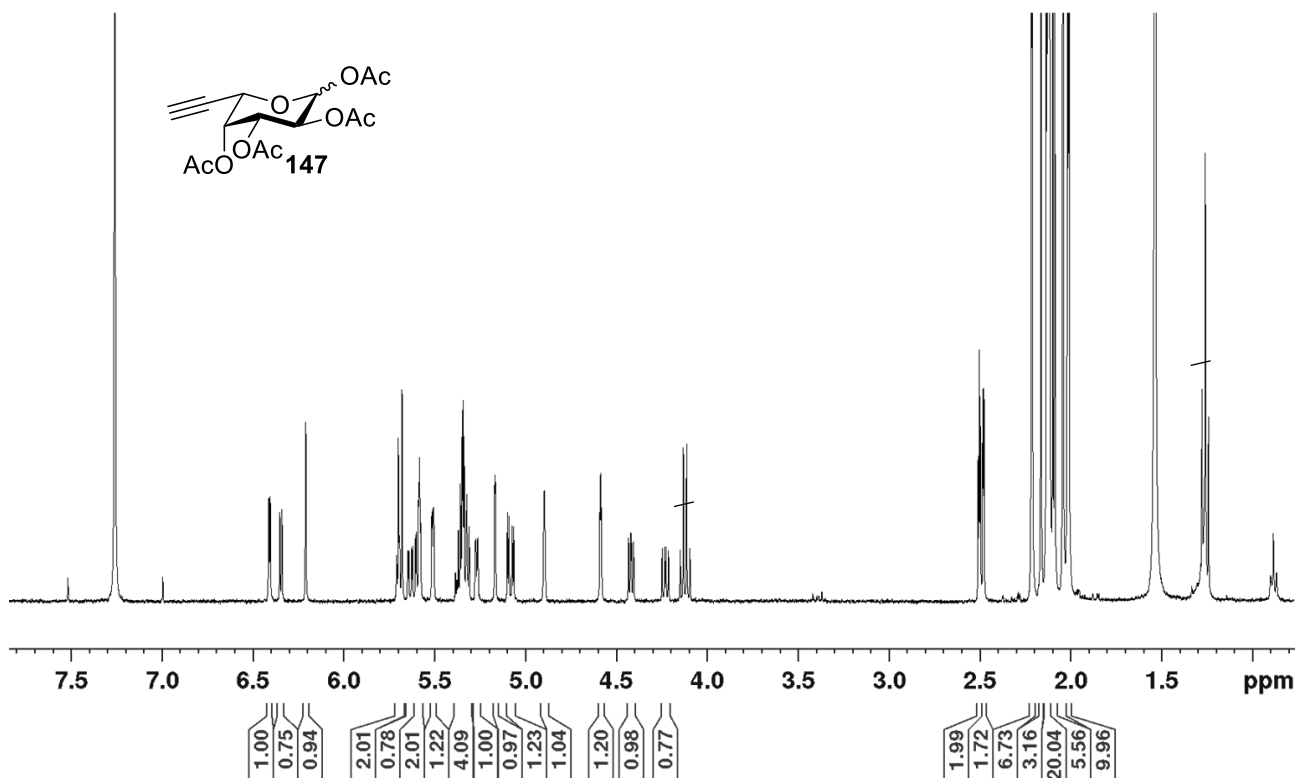
Synthesis of 5-ethynyl-L-fucose **143**

Compound **142** (142.7 mg, 0.56 mmol) was dissolved at 0 °C in a 9:1 TFA: H_2O solution (6.2 mL), which was stirred at 0 °C for 1.5 h monitoring by TLC (eluent: DCM - MeOH, 8:2, R_f = 0.60). The reaction mixture was diluted and co-evaporated with toluene (3x4 mL) and finally co-evaporated with Et_2O (3x4 mL) affording crude **143**, which was used in following step without further purification. Spectroscopic data were in agreement with those previously reported in the literature.⁷¹ 1H NMR (400 MHz, D_2O) δ (ppm): 5.28 (d, J = 5.1 Hz, 1 β -HFur), 5.25 (d, J = 2.9 Hz, 1 α -HFur), 5.22 (d, J = 3.8 Hz, 1 α -HPyr), 4.85 (dd, J = 2.2, 1.4 Hz, 5 α -HPyr), 4.55 (d, J = 7.9 Hz, 1 β -HPyr), 4.49 (dd, J = 2.2, 1.4 Hz, 5 β -HPyr), 4.01 (dd, J = 3.4, 1.4 Hz, 4 α -HPyr), 3.95 (dd, J = 3.5, 1.3 Hz, 4 β -HPyr), 3.84 (dd, J = 10.3, 3.4 Hz, 3 α -HPyr), 3.76 (dd, J = 10.3, 3.8 Hz, 2 α -HPyr), 3.63 (dd, J = 10.0, 3.5 Hz, 3 β -HPyr), 3.45 (dd, J = 10.0, 8.0 Hz, 2 β -HPyr), 2.90 (d, J = 2.3 Hz, $H-C\equiv C$), 2.86 (d, J = 2.3 Hz, $H-C\equiv C$).

¹H NMR spectrum of **143** in D₂O (400 MHz)Synthesis of 1,2,3,4-tetra-*O*-acetyl-5-ethynyl-L-fucopyranose **147**

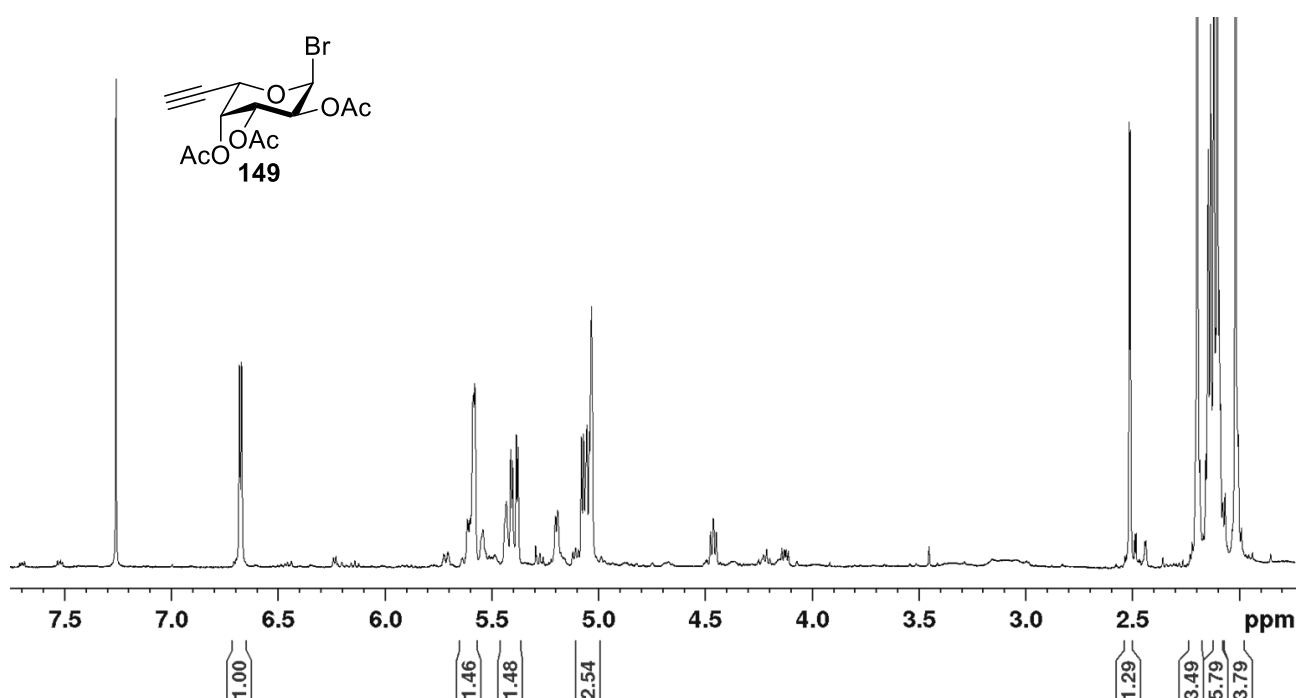
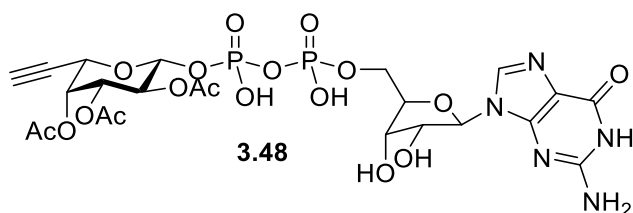
5-ethynyl-L-fucose **143** (97.5 mg, 0.56 mmol, assuming 100% yield in the previous step) was dissolved in dry pyridine (950 μ L) followed by addition of Ac₂O (950 μ L). A catalytic amount of DMAP was added (4.9 mg, 7 mol%) and the reaction was stirred under nitrogen atmosphere at room temperature for 18 h monitoring by TLC (eluent: *n*-hexane - EtOAc, 7:3). The mixture was diluted with EtOAc (25 mL) and washed with 1 M HCl aqueous solution (2x20 mL), saturated NaHCO₃ solution (2x20 mL), brine (20 mL). Aqueous phases were extracted with EtOAc (2x15 mL each) and reunited organic phases dried over anhydrous Na₂SO₄. Purification by flash chromatography (eluent: *n*-hexane - EtOAc, 7:3, R_f = 0.28) afforded pure **147** (153.8 mg, 80% over 2 steps from **142**) as a α,β anomeric mixture of pyranosidic forms (56%, 1.2:1 β/α ratio) along with the corresponding α,β mixture of furanose isomers (44%, 1:1.3 β/α ratio). ¹H NMR spectroscopic data were in agreement with those previously reported in the literature.⁸⁷ ¹H NMR (400 MHz, CDCl₃) δ (ppm): 6.41 (d, J = 3.2 Hz, 1 α -HPyr), 6.35 (d, J = 4.7 Hz, 1 β -HFur), 6.21 (s, 1 α -HFur), 5.70-5.69 (m, 3 β -HFur), 5.69 (d, J = 8.3 Hz, 1 β -HPyr), 5.64 (dd, J = 8.6, 2.4 Hz, 5 β -HFur), 5.61-5.57 (m, 5 α -HFur, 4 α -HPyr), 5.51 (dd, J = 3.4, 1.2 Hz, 4 β -HPyr), 5.40-5.33 (2 β -HPyr, 2 α -HPyr, 2 β -HFur, 3 α -HPyr), 5.27 (dd, J = 4.4, 1.2 Hz, 3 α -HFur), 5.17 (d, J = 1.2 Hz, 2 α -HFur), 5.08 (dd, J = 10.2, 3.5 Hz, 3 β -HPyr), 4.90 (t, J = 1.5 Hz, 5 α -HPyr), 4.59 (t, J = 1.8 Hz, 5 β -HPyr), 4.42 (dd, J = 6.3, 4.7 Hz, 4 α -HFur), 4.23 (dd, J = 8.4, 5.6 Hz, 4 β -HFur), 2.51 (d, J = 2.5 Hz, H_{α} -C \equiv C-HFur), 2.50 (d, J = 2.5 Hz, H_{β} -C \equiv C-HPyr), 2.48 (d, J = 2.3 Hz, H_{β} -C \equiv C-HFur, H_{α} -C \equiv C-HPyr), 2.22, 2.21, 2.14-2.12, 2.10, 2.09, 2.04, 2.02, 2.01, 2.01 (CH₃CO).

^1H NMR spectrum of **147** in CDCl_3 (400 MHz) Minor impurities ascribed to EtOAc are visible in the reported spectrum.

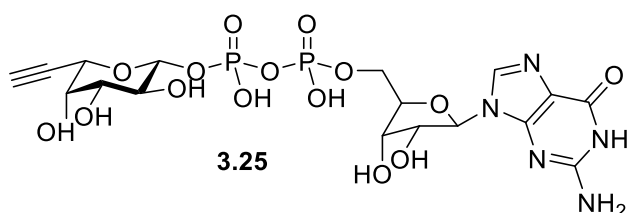


Synthesis of 2,3,4-tetra-*O*-acetyl-5-ethynyl- α -L-fucopyranosyl bromide **149**

Compound **147** (30.2 mg, 0.081 mmol) was dissolved at 0 °C under nitrogen atmosphere in a 33% HBr solution in AcOH (140 μL , 0.81 mmol). The mixture was stirred for 15 min at room temperature, then dry DCM (80 μL) was added and the reaction stirred for additional 20 min. The reaction was monitored by TLC (eluent: *n*-hexane – EtOAc, 7:3, $R_f = 0.53$) and quenched by addition of cold H_2O (12 mL, 0 °C). The product was extracted with cold DCM (2x8 mL, 0 °C) and reunited organic phases were washed with cold H_2O (2x12 mL, 0 °C) and with cold saturated NaHCO_3 solution (3x12 mL, 0 °C). The organic phase was dried over anhydrous Na_2SO_4 and filtration and evaporation of the solvent afforded crude **149** which was directly use without further purification in the next synthetic step. ^1H NMR (400 MHz, CDCl_3) δ (ppm): 6.68 (d, $J = 3.8$ Hz, 1 H, 1-H), 5.59 (d, $J = 3.7$ Hz, 1 H, 4-H or 5-H), 5.39 (dd, $J = 10.7, 3.4$ Hz, 1 H, 3-H), 5.06 (dd, $J = 8.2, 1.0$ Hz, 1 H, 2-H), 5.05-5.03 (m, 1 H, 4-H or 5-H), 2.51 (d, $J = 2.0$ Hz, 1 H, 7-H), 2.20 (s, 3 H, CH_3CO), 2.10 (s, 3 H, CH_3CO), 2.02 (s, 3 H, CH_3CO).

¹H NMR spectrum of **149** in CDCl₃ (400 MHz)Synthesis of 2,3,4-tri-*O*-acetyl-5-ethynyl-GDP- β -L-fucose **151**

Titrated GDP·2.4 mol. eq. Bu₄N⁺ (65.5 mg, 0.064 mmol) was dissolved in distilled CH₃CN (1.2 mL) in a Schlenk tube under nitrogen atmosphere. Distilled Et₃N (9 μ L, 0.064 mmol) and powdered 3 Å molecular sieves were added and the mixture was stirred at room temperature for 15 min. A solution of fucosyl bromide **149** (23.4 mg, 0.064 mmol) in distilled CH₃CN (2.0 mL) was then cannulated into the Schlenk tube and the reaction stirred at 80 °C for 30 min. Complete conversion of fucosyl bromide **148** was assessed by TLC (eluent: *n*-hexane - EtOAc, 7:3). Molecular sieves were removed by filtration over a celite pad and evaporation of the solvent afforded crude **151** that was analysed by HPLC (Waters Atlantis T3 5 μ m 4.6x100 mm column; eluents A = aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, adjusted to pH 4 with H₃PO₄ solution), B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-10 min at 40% A, 10-12 min to 0% A, eluting at 1 mL/min, monitoring at 254 nm; *t*_R = 8.1 min). The corresponding HPLC trace is reported in the discussion of Chapter 3 at page 160. MS (ESI negative mode) *m/z*: calcd for C₁₅H₂₃N₅O₁₆P₂ 599.07; found 598.53 [M-H]⁻.

Synthesis of 5-ethynyl-GDP-L-fucose **128**

Crude **151** (theoretically 0.064 mmol) was dissolved in H₂O (1.8 mL). Distilled Et₃N (10 μ L) was added reaching pH 8.9 followed by addition of alkaline phosphatase (59 μ L of 1 U/ μ L solution). The reaction was

performed at 30 °C for 20 h and enzymatic hydrolysis of GDP to guanosine was monitored by HPLC (Waters Atlantis T3 5 µm 4.6x100 mm column; eluents A = aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, adjusted to pH 4 with H₃PO₄ solution), B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-10 min at 40% A, 10-12 min to 0% A, eluting at 1 mL/min, monitoring at 254 nm; *t_R* (GDP) = 6.5 min, *t_R* (guanosine) = 1.4 min). The solvent was evaporated and the crude dissolved at 0 °C in a 2:2:1 H₂O/MeOH/Et₃N (distilled) mixture (2.0 mL) and stirred at room temperature for 16 h monitoring by HPLC (Waters Atlantis T3 5 µm 4.6x100 mm column; eluents A = aqueous buffer (12 mM Bu₄NBr, 10 mM KH₂PO₄, 5% CH₃CN, adjusted to pH 4 with H₃PO₄ solution), B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-10 min at 40% A, 10-12 min to 0% A, eluting at 1 mL/min, monitoring at 254 nm; *t_R* = 6.5 min). The solvent was evaporated and the crude purified by HPLC (Waters Atlantis Prep T3 OBDTM 5 µm 19 x 100 mm column; eluents A = H₂O + 10 mM (Bu₃NH)HCO₃, B = CH₃CN; gradient 0-8 min from 90% A to 40% A, 8-10 min at 40% A, 10-12 min to 0% A, eluting at 15 mL/min, monitoring at 254 nm; *t_R* = 5.7. The 10 mM (Bu₃NH)HCO₃ aqueous buffer was prepared adding Bu₃N (2.38 mL, 10 mmol) to H₂O (1 L), stirring at 0 °C and bubbling CO₂ for 5 h until pH ~ 5 was reached. Eluted fractions were analysed by HPLC which revealed complete decomposition of the product. HPLC traces are reported in the discussion of Chapter 3 at page 160.

3.7 References

- ¹ M. Schena, D. Shalon, R. W. Davis, P. O. Brown *Science* **1995**, *270*, 467-470.
- ² A. Lueking, M. Horn, H. Eickhoff, K. Bussow, H. Lehrach, G. Walter *Anal. Biochem.* **1999**, *270*, 103-111.
- ³ G. MacBeath, S. L. Schreiber *Science* **2000**, *289*, 1760-1763.
- ⁴ S. Fukui, T. Feizi, C. Galustian, A. M. Lawson, W. Chai *Nat. Biotechnol.* **2002**, *20*, 1011-1017.
- ⁵ D. Wang, S. Liu, B. J. Trummer, C. Deng, A. Wang *Nat. Biotechnol.* **2002**, *20*, 275-281.
- ⁶ W. G. T. Willats, S. E. Rasmussen, T. Kristensen, J. D. Mikkelsen, J. P. Knox *Proteomics* **2002**, *2*, 1666-1671.
- ⁷ M. C. Bryan, O. Plettenburg, P. Sears, D. Rabuka, S. Wacowich-Sgarbi, C.-H. Wong *Chem. Biol.* **2002**, *9*, 713-720.
- ⁸ F. Fazio, M. C. Bryan, O. Blixt, J. C. Paulson, C.-H. Wong *J. Am. Chem. Soc.* **2002**, *124*, 14397-14402.
- ⁹ B. T. Houseman, M. Mrksich *Chem. Biol.* **2002**, *9*, 443-454.
- ¹⁰ S. Park, I. Shin *Angew. Chem. Int. Ed.* **2002**, *41*, 3180-3182.
- ¹¹ For selected reviews: a) A. Geissner, P. H. Seeberger *Annu. Rev. Anal. Chem.* **2016**, *9*, 223-247; b) C. D. Rillahan, J. C. Paulson *Annu. Rev. Biochem.* **2011**, *80*, 797-823; c) C.-Y. Wu, P.-H. Liang, C.-H. Wong *Org. Biomol. Chem.* **2009**, *7*, 2247-2254; d) T. Horlacher, P. H. Seeberger *Chem. Soc. Rev.* **2008**, *37*, 1414-1422.
- ¹² Y. Zhang, Q. Li, L. G. Rodriguez, J. C. Gildersleeve *J. Am. Chem. Soc.* **2010**, *132*, 9653-9662.
- ¹³ L. Wang, R. D. Cummings, D. F. Smith, M. Huflejt, C. T. Campbell, J. C. Gildersleeve, J. Q. Gerlach, M. Kilcoyne, L. Joshi, S. Serna, N.-C. Reichardt, N. Parera Pera, R. J. Pieters, W. Eng, L. K. Mahal *Glycobiology* **2014**, *24*, 507-517.
- ¹⁴ K. Godula, C. R. Bertozzi *J. Am. Chem. Soc.* **2012**, *134*, 15732-15742.
- ¹⁵ O. Oyelaran, Q. Li, D. Farnsworth, J. C. Gildersleeve, *J. Proteome Res.* **2009**, *8*, 3529-3538.
- ¹⁶ X. Zhou, C. Turchi, D. Wang *J. Proteome Res.* **2009**, *8*, 5031-5040.
- ¹⁷ K. Godula, D. Rabuka, K. T. Nam, C. R. Bertozzi *Angew. Chem. Int. Ed.* **2009**, *48*, 4973-4976.
- ¹⁸ C. Zilio, A. Bernardi, A. Palmioli, M. Salina, G. Tagliabue, M. Buscaglia, R. Consonni, M. Chiari *Sens. Actuators, B* **2015**, *215*, 412-420.
- ¹⁹ Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis, K. Drickamer *Nat. Struct. Mol. Biol.* **2004**, *11*, 591-598.
- ²⁰ C. Galustian, C. G. Park, W. Chai, M. Kiso, S. A. Bruening, Y.-S. Kang, R. M. Steinman, T. Feizi *Int. Immunol.* **2004**, *16*, 853-866.
- ²¹ S. K. Singh, I. Streng-Ouwehand, M. Litjens, D. R. Weelij, J. J. Garcia-Vallejo, S. J. van Vliet, E. Saeland, Y. van Kooyk *Mol. Immunol.* **2009**, *46*, 1240-1249.
- ²² E. van Liempt, C. M. C. Bank, P. Mehta, J. J. Garcia-Vallejo, Z. S. Kavar, R. Geyer, R. A. Alvarez, R. D. Cummings, Y. van Kooyk, I. van Die *FEBS Lett.* **2006**, *580*, 6123-6131.
- ²³ A. S. Palma, T. Feizi, Y. Zhang, M. S. Stoll, A. M. Lawson, E. Díaz-Rodríguez, M. A. Campanero-Rhodes, J. Costa, S. Gordon, G. D. Brown, W. Chai *J. Biol. Chem.* **2006**, *281*, 5771-5779.
- ²⁴ C. D. Rillahan, E. Schwartz, C. Rademacher, R. McBride, J. Rangarajan, V. V. Fokin, J. C. Paulson *ACS Chem. Biol.* **2013**, *8*, 1417-1422.
- ²⁵ P.-H. Liang, C.-Y. Wu, W. A. Greenberg, C.-H. Wong *Curr. Opin. Chem. Biol.* **2008**, *12*, 86-92.

- ²⁶ O. Blixt, S. Head, T. Mondala, C. Scanlan, M. E. Huflejt, R. Alvarez, M. C. Bryan, F. Fazio, D. Calarese, J. Stevens, N. Razi, D. J. Stevens, J. J. Skehel, I. van Die, D. R. Burton, I. A. Wilson, R. Cummings, N. Bovin, C.-H. Wong, J. C. Paulson *Proc. Nat. Acad. Sci. U.S.A.* **2004**, *101*, 17033.
- ²⁷ M. D. Disney, P. H. Seeberger *Chem. Biol.* **2004**, *11*, 1701-1707.
- ²⁸ X. Song, B. Xia, Y. Lasanajak, D. F. Smith, R. D. Cummings *Glycoconjugate J.* **2008**, *25*, 15-25.
- ²⁹ S. Park, M.-R. Lee, S.-J. Pyo, I. Shin *J. Am. Chem. Soc.* **2004**, *126*, 4812-4819.
- ³⁰ B. T. Houseman, E. S. Gawalt, M. Mrksich *Langmuir* **2003**, *19*, 1522-1531.
- ³¹ E. A. Smith, W. D. Thomas, L. L. Kiessling, R. M. Corn *J. Am. Chem. Soc.* **2003**, *125*, 6140-6148.
- ³² D. Wang, G. T. Carroll, N. J. Turro, J. T. Koberstein, P. Kovac, R. Saksena, R. Adamo, L. A. Herzenberg, L. A. Herzenberg, L. Steinman *Proteomics* **2007**, *7*, 180-184.
- ³³ M. Shipp, Matthew; R. Nadella, H. Gao, V. Farkas, H. Sigrist, A. Faik *Glycoconjugate J.* **2008**, *25*, 49-58.
- ³⁴ S. Angeloni, J. L. Ridet, N. Kusy, H. Gao, F. Crevoisier, S. Guinchard, S. Kochhar, H. Sigrist, N. Sprenger *Glycobiology* **2005**, *15*, 31-41.
- ³⁵ G. T. Carroll, D. Wang, N. J. Turro, J. T. Koberstein *Langmuir* **2006**, *22*, 2899-2905.
- ³⁶ Z. Pei, H. Yu, M. Theurer, A. Walden, P. Nilsson, M. Yan, O. Ramstrom *ChemBioChem* **2007**, *8*, 166-168.
- ³⁷ E. L. Shipp, L. C. Hsieh-Wilson *Chem. Biol.* **2007**, *14*, 195-208.
- ³⁸ O. Carion, J. Lefebvre, G. Dubreucq, L. Dahri-Correia, J. Correia, O. Melnyk *ChemBioChem* **2006**, *7*, 817-826.
- ³⁹ K.-S. Ko, F. A. Jaipuri, N. L. Pohl *J. Am. Chem. Soc.* **2005**, *127*, 13162-13163.
- ⁴⁰ Y. Chevolut, C. Bouillon, S. Vidal, F. Morvan, A. Meyer, J.-P. Cloarec, A. Jochum, J.-P. Praly, J.-J. Vasseur, E. Souteyrand *Angew. Chem. Int. Ed.* **2007**, *46*, 2398-2402.
- ⁴¹ K. Godula, C. R. Bertozzi *J. Am. Chem. Soc.* **2010**, *132*, 9963-9965.
- ⁴² D. M. Ratner, E. W. Adams, J. Su, B. R. O'Keefe, M. Mrksich, P. H. Seeberger *ChemBioChem* **2004**, *5*, 379-382.
- ⁴³ B. S. Bochner, R. A. Alvarez, P. Mehta, N. V. Bovin, O. Blixt, J. R. White, R. L. Schnaar *J. Biol. Chem.* **2005**, *280*, 4307-4312.
- ⁴⁴ R. Karamanska, J. Clarke, O. Blixt, J. I. MacRae, J. Q. Zhang, P. R. Crocker, N. Laurent, A. Wright, S. L. Flitsch, D. A. Russell, R. A. Field *Glycoconjugate J.* **2008**, *25*, 69-74.
- ⁴⁵ B. Xia, Z. S. Kawar, T. Ju, R. A. Alvarez, G. P. Sachdev, R. D. Cummings *Nat. Methods* **2005**, *2*, 845-850.
- ⁴⁶ P.-H. Liang, S.-K. Wang, C.-H. Wong *J. Am. Chem. Soc.* **2007**, *129*, 11177-11184.
- ⁴⁷ J.-C. Lee, C.-Y. Wu, J. V. Apon, G. Siuzdak, C.-H. Wong *Angew. Chem. Int. Ed.* **2006**, *45*, 2753-2757.
- ⁴⁸ T. R. Northen, J.-C. Lee, L. Hoang, J. Raymond, D.-R. Hwang, S. M. Yannone, C.-H. Wong, G. Siuzdak *Proc. Nat. Acad. Sci. U.S.A.* **2008**, *105*, 3678-3683.
- ⁴⁹ Z.-L. Zhi, N. Laurent, A. K. Powell, R. Karamanska, M. Fais, J. Voglmeir, A. Wright, J. M. Blackburn, P. R. Crocker, D. A. Russell, S. Flitsch, R. A. Field, J. E. Turnbull *ChemBioChem* **2008**, *9*, 1568-1575.
- ⁵⁰ A. Beloqui, J. Calvo, S. Serna, S. Yan, I. B. H. Wilson, M. Martin-Lomas, N.-C. Reichardt *Angew. Chem. Int. Ed.* **2013**, *52*, 7477-7481.
- ⁵¹ Caterina Zilio, 2014, PhD thesis, University of Milan, R09676.
- ⁵² M. Bosco, S. Le Gall, C. Rihouey, S. Couve-Bonnaire, M. Bardor, P. Lerouge, X. Pannecoucke *Tetrahedron Lett.* **2008**, *49*, 2294-2297.
- ⁵³ S. P. Pujari, L. Scheres, A. T. M. Marcelis, H. Zuillhof *Angew. Chem. Int. Ed.* **2014**, *53*, 6322-6356.
- ⁵⁴ B. Ramakrishnan, P. S. Shah, P. K. Qasba *J. Biol. Chem.* **2001**, *276*, 37665-37671.
- ⁵⁵ B. Ramakrishnan, P. K. Qasba *J. Biol. Chem.* **2002**, *277*, 20833-20839.
- ⁵⁶ R. Šardžik, A. P. Green, N. Laurent, P. Both, C. Fontana, J. Voglmeir, M. J. Weissenborn, R. Haddoub, P. Grassi, S. M. Haslam, G. Widmalm, S. L. Flitsch *J. Am. Chem. Soc.* **2012**, *134*, 4521-4524.
- ⁵⁷ C. Unverzagt, H. Kunz, J. C. Paulson *J. Am. Chem. Soc.* **1990**, *112*, 9308-9309.
- ⁵⁸ Adapted procedure from O. Dasse, A. Mahadevan, L. Han, B. R. Martin, V. D. Marzo, R. K. Razdan *Tetrahedron* **2000**, *56*, 9195-9202.
- ⁵⁹ C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057-3064.
- ⁶⁰ V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-2599.
- ⁶¹ T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin *Org. Lett.* **2004**, *6*, 2853-2855.
- ⁶² H.-Y. Lee, C.-Y. Chen, T.-I. Tsai, S.-T. Li, K.-H. Lin, Y.-Y. Cheng, C.-T. Ren, T.-J. R. Cheng, C.-YJ Wu, C.-H. Wong *J. Am. Chem. Soc.* **2014**, *136*, 16844-16853.
- ⁶³ For a review: G. K. Wagner, T. Pesnota, R. A. Field *Nat. Prod. Rep.* **2009**, *26*, 1172-1194.
- ⁶⁴ a) V. Wittmann, C.-H. Wong *J. Org. Chem.* **1997**, *62*, 2144-2147; b) G. Baisch, R. Ohrlein *Biorg. Med. Chem.* **1997**, *5*, 383-391; c) S. Wendicke, S. Warnecke, C. Meier *Angew. Chem. Int. Ed.* **2008**, *47*, 1500-1502; d) T.-C. Lin, J.-M. Fang *Tetrahedron Lett.* **2011**, *52*, 2232-2234; e) H. Tanaka, Y. Yoshimura, M. R. Jørgensen, J. A. Cuesta-Seijo, O. Hindsgaul *Angew. Chem. Int. Ed.* **2012**, *51*, 11531-11534; f) S. Mohamady, A. Desoky, S. D. Taylor *Org. Lett.* **2012**, *14*, 402-405.
- ⁶⁵ S. C. Timmons, D. L. Jakeman *Org. Lett.* **2007**, *9*, 1227-1230.
- ⁶⁶ M. Sawa, T.-L. Hsu, T. Itoh, M. Sugiyama, S. R. Hanson, P. K. Vogt, C.-H. Wong *Proc. Nat. Acad. Sci. U.S.A.* **2006**, *103*, 12371-12376.
- ⁶⁷ K. Bock, I. Lundt, C. Pedersen *Carbohydr. Res.* **1981**, *90*, 7-16.
- ⁶⁸ F. Legrand, C. Breyton, P. Guillet, C. Ebel, G. Durand *J. Org. Chem.* **2016**, *81*, 681-688.

-
- ⁶⁹ S. Pedatella, A. Guaragna, D. D'Alonzo, M. De Nisco, G. Palumbo *Synthesis* **2006**, 2, 305-308.
- ⁷⁰ G. Srivastava, K. J. Kaur, O. Hindsgaul, M. M. Palcic *J. Biol. Chem.* **1992**, 267, 22356-22361.
- ⁷¹ W. Yi, X. Liu, Y. Li, J. Li, C. Xia, G. Zhou, W. Zhang, W. Zhao, X. Chen, P. G. Wang *Proc. Nat. Acad. Sci. U.S.A.* **2009**, 106, 4207-4212.
- ⁷² M. Frigerio, M. Santagostino, S. Sputore *J. Org. Chem.* **1999**, 64, 4537-4538.
- ⁷³ S. Mueller, B. Liepold, G. J. Roth, H. J. Bestmann *Synlett* **1996**, 521.
- ⁷⁴ T. H. Jepsen, J. L. Kristensen *J. Org. Chem* **2014**, 79, 9423-9426.
- ⁷⁵ Matteo Panza - Synthesis of substituted 2-phenyl-benzofurans as HSP90 modulators. Master Degree Thesis AY2013/2014.
- ⁷⁶ S. Tokutake, N. Yamaji, M. Kato *Chem. Pharm. Bull.* **1990**, 38, 13-18.
- ⁷⁷ N. Lunau, K. Seelhorst, S. Kahl, K. Tscherch, C. Stacke, S. Rohn, J. Thiem, U. Hahn, C. Meier *Chem. Eur. J.* **2013**, 19, 17379-17390.
- ⁷⁸ S. Yan, S. Serna, N.-C. Reichardt, K. Paschinger, I. B. H. Wilson *J. Biol. Chem.* **2013**, 288, 21015-21028.
- ⁷⁹ W. C. Still, M. Kahn, A. Mitra, *J. Org. Chem.* **1978**, 43, 2923-2925.
- ⁸⁰ B. T. Houseman, M. Mrksich *Chem. Biol.* **2002**, 9, 443-454.
- ⁸¹ L. P. Calle, B. Echeverria, A. Franconetti, S. Serna, M. C. Fernandez-Alonso, T. Diercks, F. J. Canada, A. Arda, N.-C. Reichardt, J. Jimenez-Barbero *Chem. Eur. J.* **2015**, 21, 11408-11416.
- ⁸² S. Serna, J. Etxebarria, N. Ruiz, M. Martin-Lomas, N.-C. Reichardt *Chem. Eur. J.* **2010**, 16, 13163-13175.
- ⁸³ S. Serna, S. Yan, M. Martin-Lomas, I. B. H. Wilson, N.-C. Reichardt *J. Am. Chem. Soc.* **2011**, 133, 16495-16502.
- ⁸⁴ J. Etxebarria, S. Serna, A. Beloqui, M. Martin-Lomas, N.-C. Reichardt *Chem. Eur. J.* **2013**, 19, 4776-478.
- ⁸⁵ E. E. Boeggeman, B. Ramakrishnan, P. K. Qasba *Protein Expression Purif.* **2003**, 30, 219-229.
- ⁸⁶ Adapted from reference 70.
- ⁸⁷ T.-L. Hsu, S. R. Hanson, K. Kishikawa, S.-K. Wang, M. Sawa, C.-H. Wong *Proc. Nat. Acad. Sci. U.S.A.* **2007**, 104, 2614-2619.