



HAL
open science

Chemo-Enzymatic Synthesis of *S. mansoni* O-Glycans and Their Evaluation as Ligands for C-Type Lectin Receptors MGL, DC-SIGN, and DC-SIGNR

Julie Pham, Alvaro Hernandez, Anna Cioce, Silvia Achilli, Giulio Goti, Corinne Vivès, Michel Thepaut, Anna Bernardi, Franck Fieschi, Niels-Christian Reichardt

► To cite this version:

Julie Pham, Alvaro Hernandez, Anna Cioce, Silvia Achilli, Giulio Goti, et al.. Chemo-Enzymatic Synthesis of *S. mansoni* O-Glycans and Their Evaluation as Ligands for C-Type Lectin Receptors MGL, DC-SIGN, and DC-SIGNR. *Chemistry - A European Journal*, 2020, 26 (56), pp.12818-12830. 10.1002/chem.202000291 . hal-02974726

HAL Id: hal-02974726

<https://hal.science/hal-02974726v1>

Submitted on 25 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chemo-enzymatic Synthesis of *S. mansoni* O-glycans and their Evaluation as Ligands for C-type Lectin Receptors MGL, DC-SIGN and DC-SIGNR

Julie Pham¹, Alvaro Hernandez^{1,2}, Anna Cioce¹, Silvia Achilli^{3,4}, Giulio Goti⁵, Corinne Vivès³, Michel Thepaut³, Anna Bernardi⁵, Franck Fieschi³, Niels Reichardt^{*1,6,7}

¹Glycotechnology Group, CIC biomaGUNE, Paseo Miramón 182, 20014 San Sebastian, Spain

²Asparia Glycomics S.L., Mikeletegi 83, 20009 San Sebastian, Spain

³Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, 38100 Grenoble, France

⁴Present address: DCM, UMR 5250, Université Grenoble Alpes, CNRS, 38000 Grenoble (France)

⁵Dipartimento di Chimica, Università degli Studi di Milano, via Golgi 19, 20133 Milano, Italy

⁶CIBER-BBN, Paseo Miramón 182, 20014 San Sebastian, Spain

⁷Basque Research and Technology Alliance (BRTA), Paseo Miramón 182, 20014 San Sebastian, Spain

nreichardt@cicbiomagune.es

Abstract

Due to their interactions with C-type lectin receptors (CLRs), glycans from the helminth *Schistosoma mansoni* represent promising leads for treatment of autoimmune diseases, allergies or cancer. We chemo-enzymatically synthesized nine O-glycans based on the two predominant O-glycan cores observed in the infectious stages of schistosomiasis, the mucin core 2 and the *S. mansoni* core. The O-glycans were fucosylated next to a selection of N-glycans directly on a microarray slide using a recombinant fucosyltransferase and GDP-fucose or GDP-6-azidofucose as donor. Binding assays with fluorescently labelled human CLRs DC-SIGN, DC-SIGNR and MGL revealed the novel O-glycan **O8** as the best ligand for MGL from our panel. Significant binding to DC-SIGN was also found for azido-fucosylated glycans. Contrasting binding specificities were observed between the monovalent carbohydrate recognition domain (CRD) and the tetravalent extracellular domain (ECD) of DC-SIGNR.

Introduction

S. mansoni glycans play an important role in host infection by helminths and mounting evidence suggests that they contribute to the parasite evasion of the host immune system by hijacking C-type lectin receptors (CLRs). CLRs are crucial components in the immune system acting as pattern recognition receptors (PRRs).^[1,2] Their interaction with pathogen associated molecular patterns (PAMPs) at the pathogen surface can lead to a series of signalling events ultimately resulting in a skewed immune response.^[3,4] The glycans on the surface of *S. mansoni* are thought to be responsible for the characteristic T_H2/Treg response observed during chronic schistosomiasis.^[5] Concomitantly, reduced disease related inflammation was observed in autoimmune pathologies such as multiple sclerosis, rheumatoid arthritis, type I diabetes and inflammatory bowel diseases for patients infected with helminths or treated with parasite extracts.^[6]

Although yet to be fully assigned, the *S. mansoni* glycome displays a rich array of immunogenic glycan epitopes such as Le^x and LDNF which are known to interact with host CLRs. Moreover, interesting similarities can be observed between helminth and host

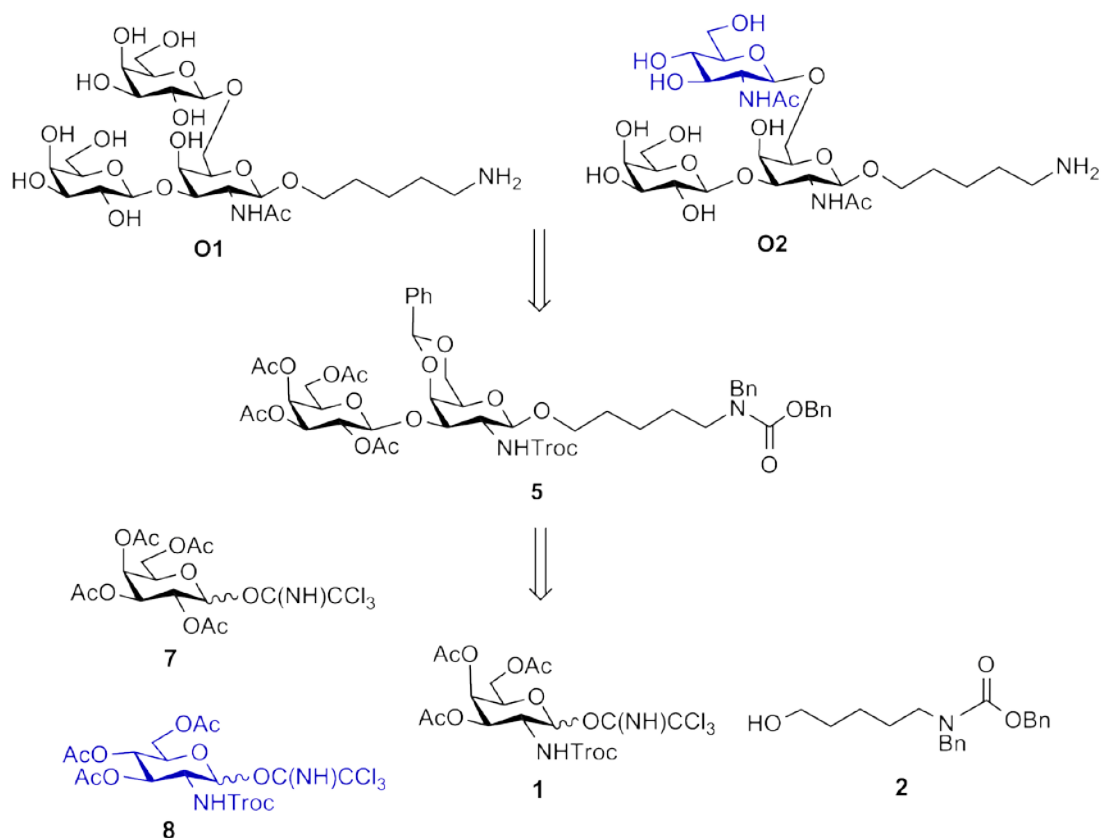
glycans. Indeed, the *S. mansoni* specific O-glycan core, which is predominantly found in the cercariae, closely resembles an O-glycan core found in the intermediate snail host *Biomphalaria glabrata*.^[7,8] On the other hand, the mucin core 2, which is abundantly found in mammalian host mucosal tissues, is predominantly found in the helminth eggs.^[9] Thus, *S. mansoni* is strongly suspected to employ glycan mimicry as a strategy to subvert the host immune system to allow parasitic survival.^[10]

The immunomodulatory properties of *S. mansoni* O-glycans place them as potentially interesting lead compounds for the development of carbohydrate-based drugs to treat immuno-compromised patients with autoimmune diseases, allergies or cancer. However, most studies have focused on parasite N-glycans or the antigenic motifs alone and only few describe the relevance of parasite O-glycans and their interactions with CLRs.^[11,12] This is mainly due to the challenge of isolating pure compounds in sufficient amounts for functional and diagnostic tests.^[13–15]

The aim of our study was to investigate the interaction of *S. mansoni* O-glycans with key CLRs and identify lead structures for improved targeting of immune cells *via* specific CLR-glycan interactions. To this end, we report, here for the first time, the chemoenzymatic synthesis of several *S. mansoni* O-glycans and evaluate their interactions with fluorescently labelled human CLRs DC-SIGN, DC-SIGNR and MGL.

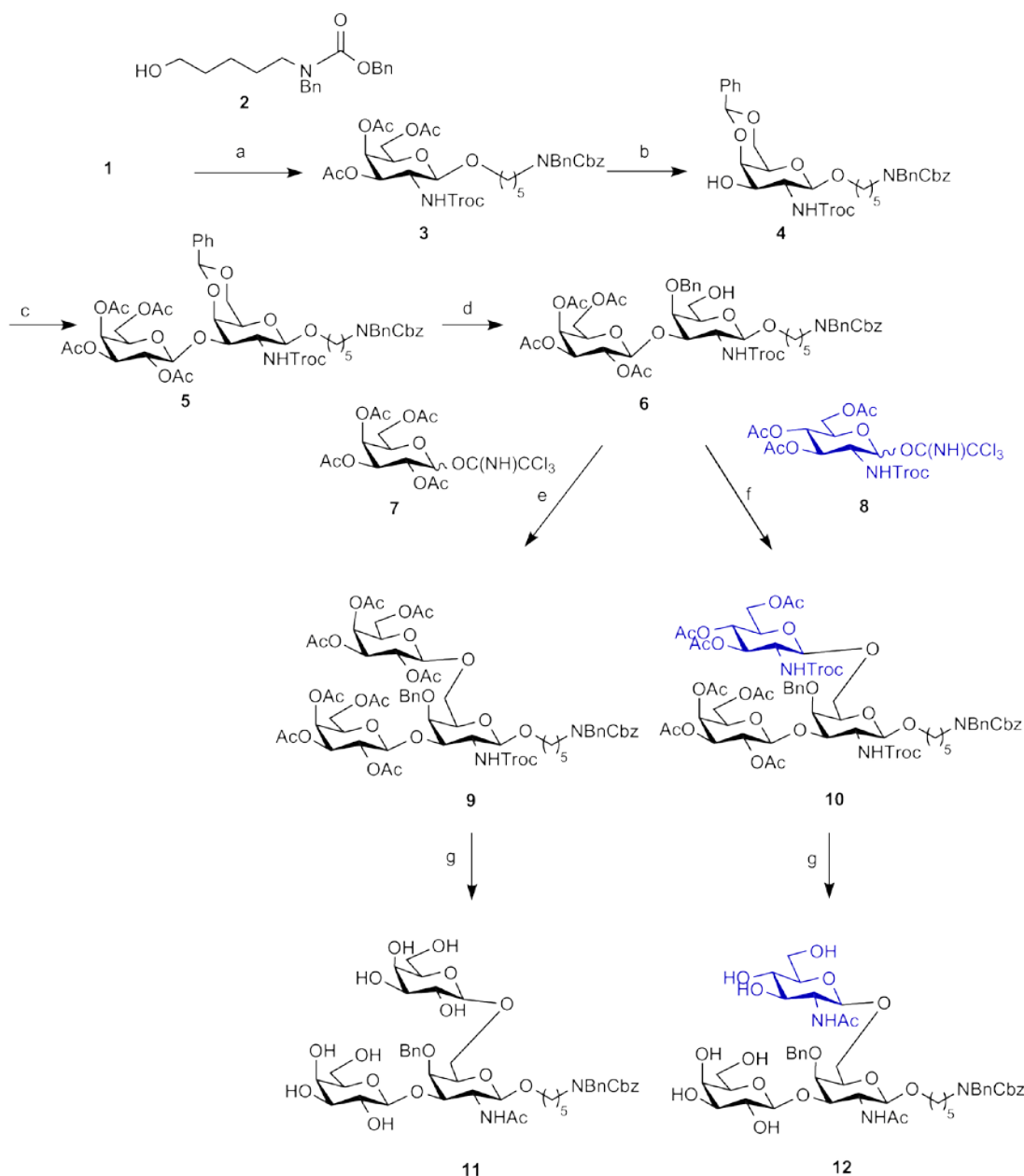
Results and Discussion

The retrosynthetic analysis of the O-glycan cores **O1** and **O2** (Scheme 1) suggested a convergent strategy involving three building blocks. After conjugation of the GalNAc donor **1** with a protected amine linker **2**, glycosylation in O-3 with galactose donor **7** would provide the Gal β 1-3GalNAc disaccharide **5** common to both cores. A second glycosylation in O-6 with galactose (Gal) **7** or N-acetylglucosamine (GlcNAc) donor **8** would furnish the *S. mansoni* core **O1** or mucin 2 core **O2**, respectively.



Scheme 1. Retrosynthetic analysis for the *S. mansoni* core and core 2 trisaccharides **O1** and **O2**

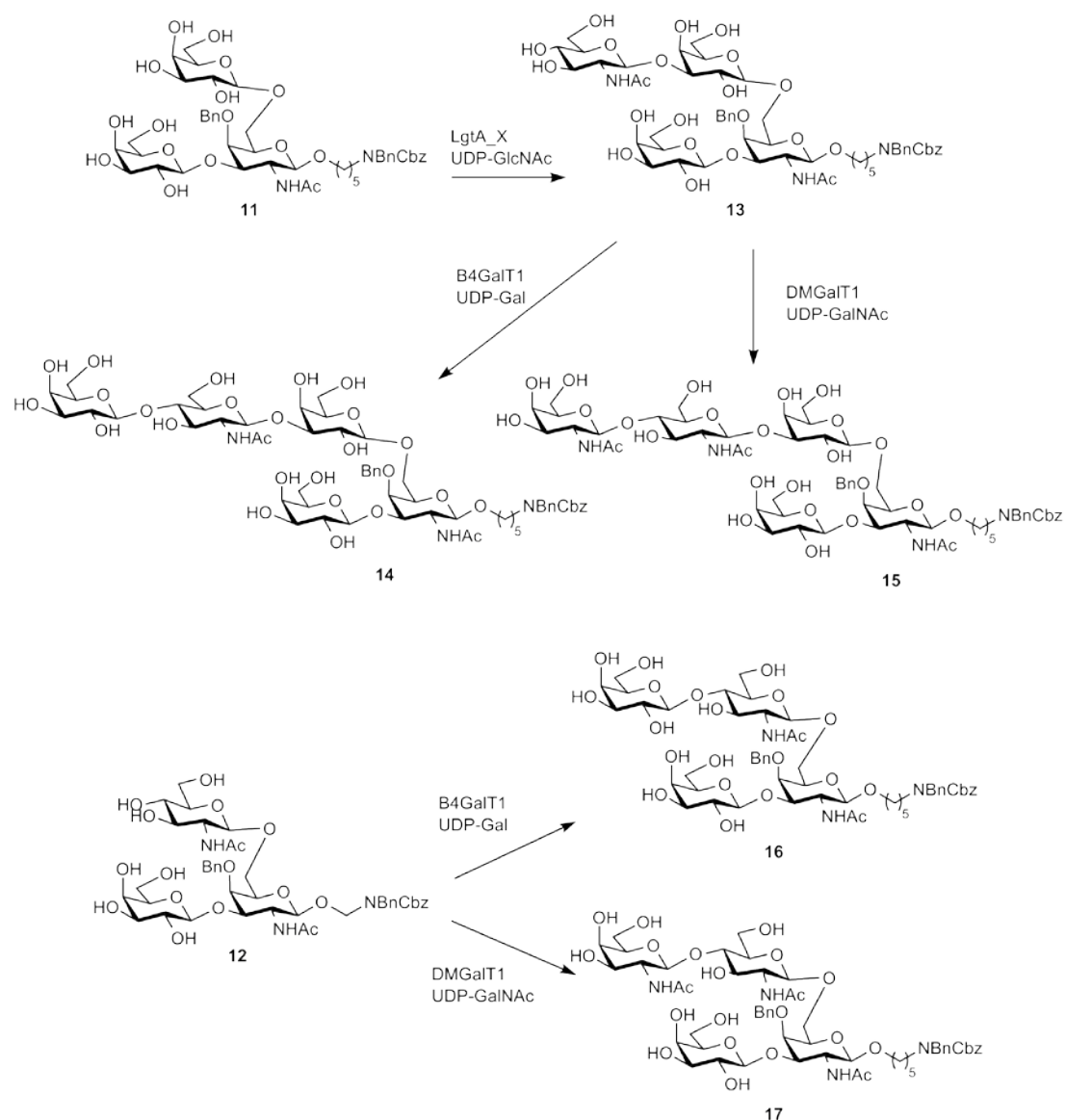
To facilitate the immobilisation of final compounds as ligands on glycan arrays or to carrier proteins, the known galactosamine trichloroacetimidate donor **1**^[16] (Scheme 1) was conjugated with N-benzyl-benzylcarbamate protected aminopentyl linker **2**^[17] to obtain glycoconjugate **3** in 73%. Treatment of **3** with a guanidine/guanidium solution quantitatively hydrolyzed all acetates without affecting the base-sensitive N-Troc group.^[18] Subsequent acetal protection of O-4 and O-6 with benzaldehyde dimethyl acetal under acid catalysis provided **4** in 70%. Glycosylation of the remaining free O-3 with the galactose imidate **7** under TMSOTf catalysis then afforded disaccharide **5** in 62% yield, as a common intermediate in the synthesis of both core structures. Reductive ring opening of the benzylidene acetal in **5** with $\text{BH}_3\cdot\text{THF}$ and catalytic TMSOTf furnished the 4-benzyl ether **6** in a moderate 61% yield. Finally, glycosylation at O-6 either with galactose donor **7** or GlcNAc donor **8** afforded trisaccharides **9** and **10** in 40% and 73% yield, respectively.



Scheme 2. Convergent synthesis of the partially protected *S. mansoni* and mucin 2 cores **11** and **12**. Reagents and conditions: a) Linker **2**, TMSOTf, DCM, -40°C, 73%; b) i) G/GHNO₃, MeOH/DCM; ii) PhCH(OMe)₂, cat. CSA, DCM, 70%; c) Donor, TMSOTf, DCM, -40°C, 62%; d) 1M BH₃.THF, TMSOTf, DCM, 61%; e) Donor **7**, TMSOTf, DCM, -41°C, 40%; f) Donor **8**, TMSOTf, DCM, -40°C, 73%; g) i) 1M TBAF, THF, reflux; ii) Ac₂O, pyridine; iii) 0.5M NaOMe, MeOH, 70-90%

For the enzymatic diversification of the core structures with recombinant glycosyltransferases both O-glycan cores were partially deprotected. The benzyl protection was left untouched at this stage of the synthesis to aid in the separation of products by HPLC after enzymatic modifications. Deprotection of the *N*-Troc functionality under standard conditions with LiOH or zinc/acetic acid amalgam failed to provide clean products but refluxing the compounds **9** and **10** in 1M TBAF in THF afforded the clean *S. mansoni* core **11** in 95% and the mucin core 2 trisaccharide **12** in 78% yield after three steps.^[19–22]

Both core structures were then diversified by enzymatic glycosylations to access the Gal1-4GlcNAc (LN) and GalNAc1-4GlcNAc (LDN) motifs typically observed in the cercarial and egg glycomes of *S. mansoni*. Towards this end, we employed three recombinant glycosyltransferases, a bovine beta-1,4-galactosyltransferase (β 4Gal-T1), a double mutant (Y289LC342T) conferring beta-1,4-N-galactosaminyltransferase (DMGalT1) activity and a *Neisseria meningitidis* beta-1,3-N-acetylglucosaminyltransferase (LgtA_X) construct prepared exclusively for this purpose (scheme 3).^[23–25] The new LgtA_X construct contains two polyhistidine tags and a thioredoxin domain to increase LgtA purity and solubility, respectively.



Scheme 3. Enzymatic diversification of cores structures **11** and **12** yielding glycans **13** (57%), **14** (28%), **15** (49%), **16** (61%) and **17** (59%).

The general substrate and donor specificity of LgtA, including a comparison with the human homologue B3GnT2, has previously been described by Blixt *et al.* Employing N-glycan **G5** (Figure 1) we confirmed the enzymatic activity of LgtA_X to be similar to LgtA in solution, *i.e.* for the elongation of terminal galactose with GlcNAc. To evaluate the usefulness of the

enzyme for diversification of the *S. mansoni* core type trisaccharide we assessed the enzyme selectivity and activity on trisaccharide **11**. Incubation of the O-glycan core with 2 potential acceptor sites afforded 57% of a β 1-6 monosubstituted product **13**, next to only 7% of the bis-substituted product.^[26] Errore. Il segnalibro non è definito.

Before enzymatic introduction of an acid-labile fucose residue we decided to remove remaining benzyl and Cbz protecting groups by hydrogenation. Protected glycans **11-17** were dissolved in a mixture of H₂O/MeOH/AcOH, and hydrogenated under Pd/C(10%) catalysis under atmospheric pressure of H₂(g), which afforded the deprotected final glycans **O1-O7** in 40-100% yield (Figure 1).^[27]

For the synthesis of the proposed Gal β 1-4[Fuca α 1-3]GlcNAc (Le^x)-type and GalNAc β 1-4[Fuca α 1-3]GlcNAc (LDNF) structures we employed a recombinant α 1,3-fucosyltransferase from *Caenorhabditis elegans* (CeFUT6) and a commercial *Helicobacter pylori* α -1,3 fucosyltransferase (HP-FucT). CeFUT6 has a known Lewis-X fucosylation activity which was employed in this study to yield **O8** from **O7** in 73% yield.^[28] HP-FucT has a broad substrate and donor specificity including the C-6 azido-fucose surrogate (FucZ).^[29,30] Introducing an azide function as a biorthogonal handle in the molecule opened the possibility of generating glycomimetics by copper-catalyzed cycloaddition with alkynes to afford a library of novel glycomimetics with potentially improved affinity towards CLRs.^[31-34]

We decided to employ a microarray platform for the rapid preparation of fucosylated and azido-fucosylated compounds. Performing the reactions in parallel and on a microscale decreases the need of expensive azido-donor by several orders of magnitude compared to the solution phase synthesis of individual compounds. The synthetic parasite O-glycans **O1-O8** were printed alongside a set of N-glycans which were chosen as additional fucose acceptors from a glycan library available in our laboratory onto NHS activated indium tin oxide (ITO) slides (Figure 1).^[35,36] The bio-functionality of the printed glycans was confirmed by incubating the array with the plant lectin peanut agglutinin (PNA) ECD recognizing Gal β 1-3GalNAc residues (see SI, Figure S4)^[37,38]

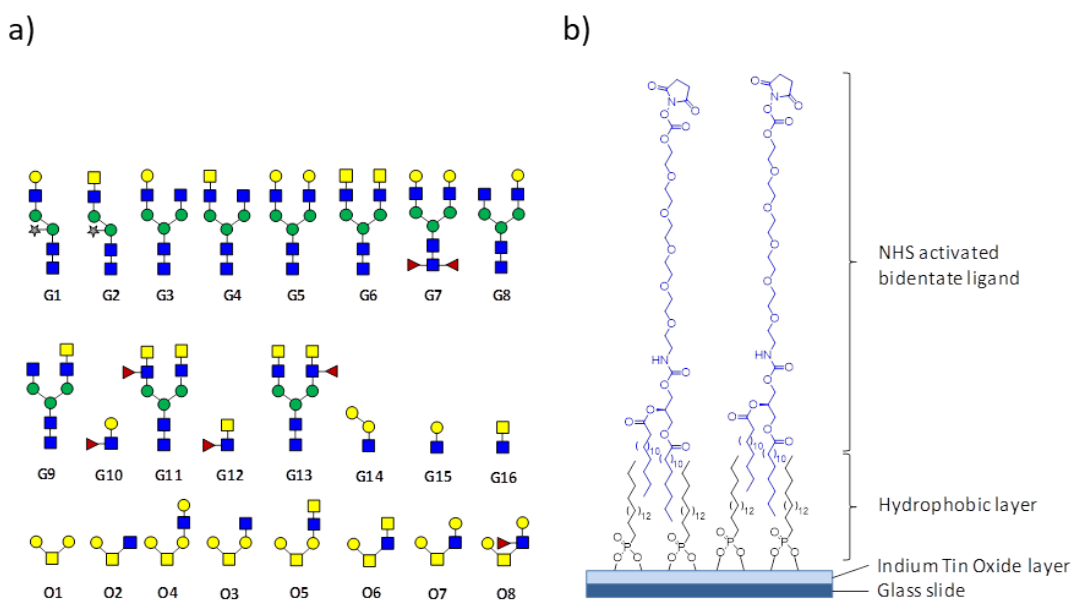


Figure 1. a) Selection of ligands immobilized on the microarray; b) surface functionalization of the ITO slide

Initially, we screened donor and HP-FucT enzyme concentrations to optimize on-chip fucosylation for printed glycans. Yields for the enzymatic on-chip glycosylation were determined by MALDI-TOF MS as ratios of the sum of product signal intensities (glycan#-F for fucosylation and glycan#-Z for azido-fucosylation) to remaining starting material signal intensities (Figure 2, A). Plotting the product ratios for the different glycans showed no increase in product formation beyond an enzyme concentration of 189 mU ml⁻¹ (Figure 2, B) possibly due to enzyme precipitation at higher concentration.

Repeated glycosylation with freshly prepared enzyme solution and 1mM of GDP-fucose produced higher product conversion than a single exposure with 2mM donor concentration (Figure 2, C). While for glycosylation with the natural GDP-fucose donor, two incubation cycles were sufficient to reach a maximum conversion between 65-90%, similar conversion yields employing the GDP- 6-azido-fucose required 3 cycles.

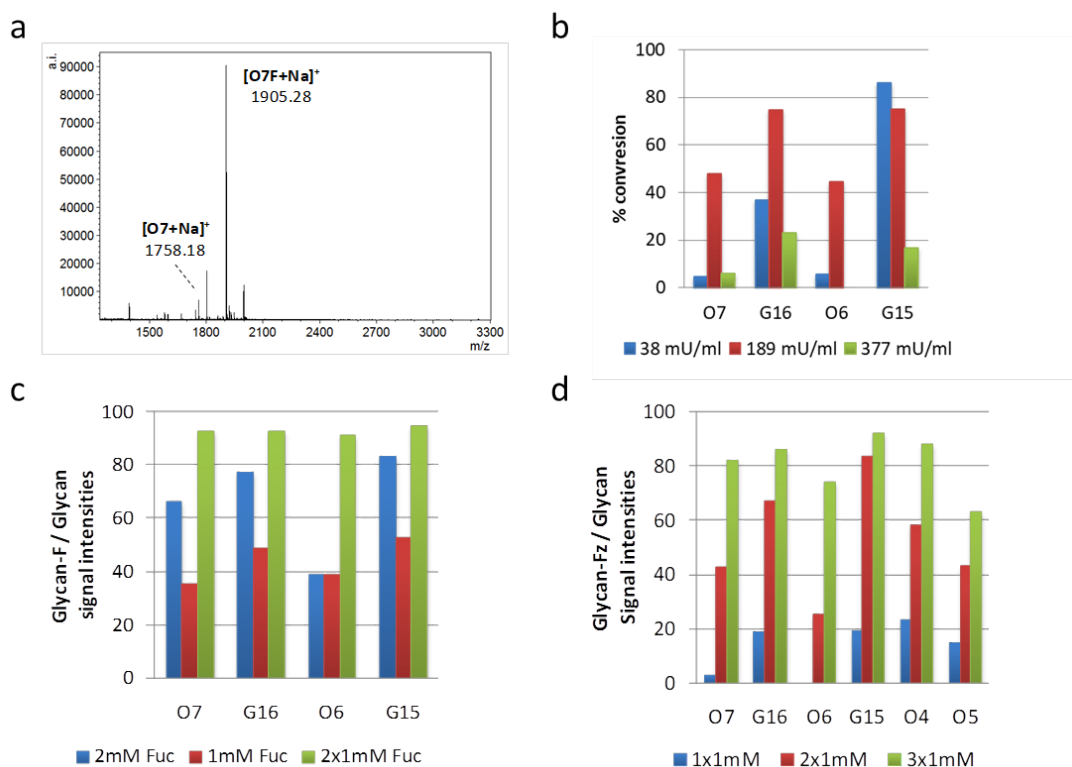


Figure 2. Optimisation of reaction conditions for on-chip fucosylation; a) MALDI-TOF MS of fucosylation of **O7**; b) Effect of enzyme concentration; c) Fucose donor concentration; D) Azido fucose donor concentration on product conversion

Product conversion yields showed a certain degree of structure-dependent variability but were altogether above 80% for the fucose addition and above 75% for the azido-fucose addition (see SI, Figure S5).

As a general note to the reader unfamiliar with glycan array binding experiments, the binding data between different arrays were only compared in relative terms i.e. by comparing ligand binding profiles. Apart from the underlying binding affinity between carbohydrate ligand and lectin, the absolute fluorescence intensities given as RFU values depend on the degree of protein labelling, the incubation time, the washing conditions and image acquisition parameters among others. Some of these parameters are more easily controlled than others and efforts are under way in many labs to improve reproducibility of glycan arrays to allow for a comparison of data between arrays and different immobilisation platforms. Unless an internal reference is used to normalise fluorescence values between different slides or arrays, or extreme measures are put in place to reproduce incubation and analysis conditions for all slides, absolute fluorescence values between different arrays should not be compared.

The effect of the spatial organization of carbohydrate binding domains on glycan interaction was first assessed by comparing the fluorescent binding profiles of the monovalent CRD and the tetrameric ECD of DC-SIGNR on the unmodified array (Figure 3). DC-SIGNR is an endothelial tetrameric C-type lectin receptor that binds to high mannose and complex glycans found on surface glycoproteins of ebola, HIV or hepatitis virus.^[39] It has also been suggested as an entry path for *S. mansoni* larvae that begin their migration towards the intestine and liver.^[40] The glycan binding profile of DC-SIGNR has been previously studied by glycan array and frontal affinity chromatography. Strong binders include high mannose and hybrid structures and to a lesser extent galactosylated complex glycans with terminal GlcNAc residues.^[41–43] Lacking high mannose and hybrid structures on our array, we found strongest binding to the ECD domain for glycans with terminal GlcNAc (**G4**, **O2** and **O3**). The CRD showed preferential binding towards glycans with a *S. mansoni* core and galactose residues **G10**, **G14** and **O4**, while mucin core derivatives **O2** and **O3** which strongly bound the ECD construct did not show interaction with the isolated carbohydrate recognition domain CRD. In addition, and contrary to the ECD, fucosylated glycans (**O8**, **G11**, **G13** and **G7**) displayed enhanced lectin binding compared to their non-fucosylated counterparts for the CRD (**O7** and **G5**).

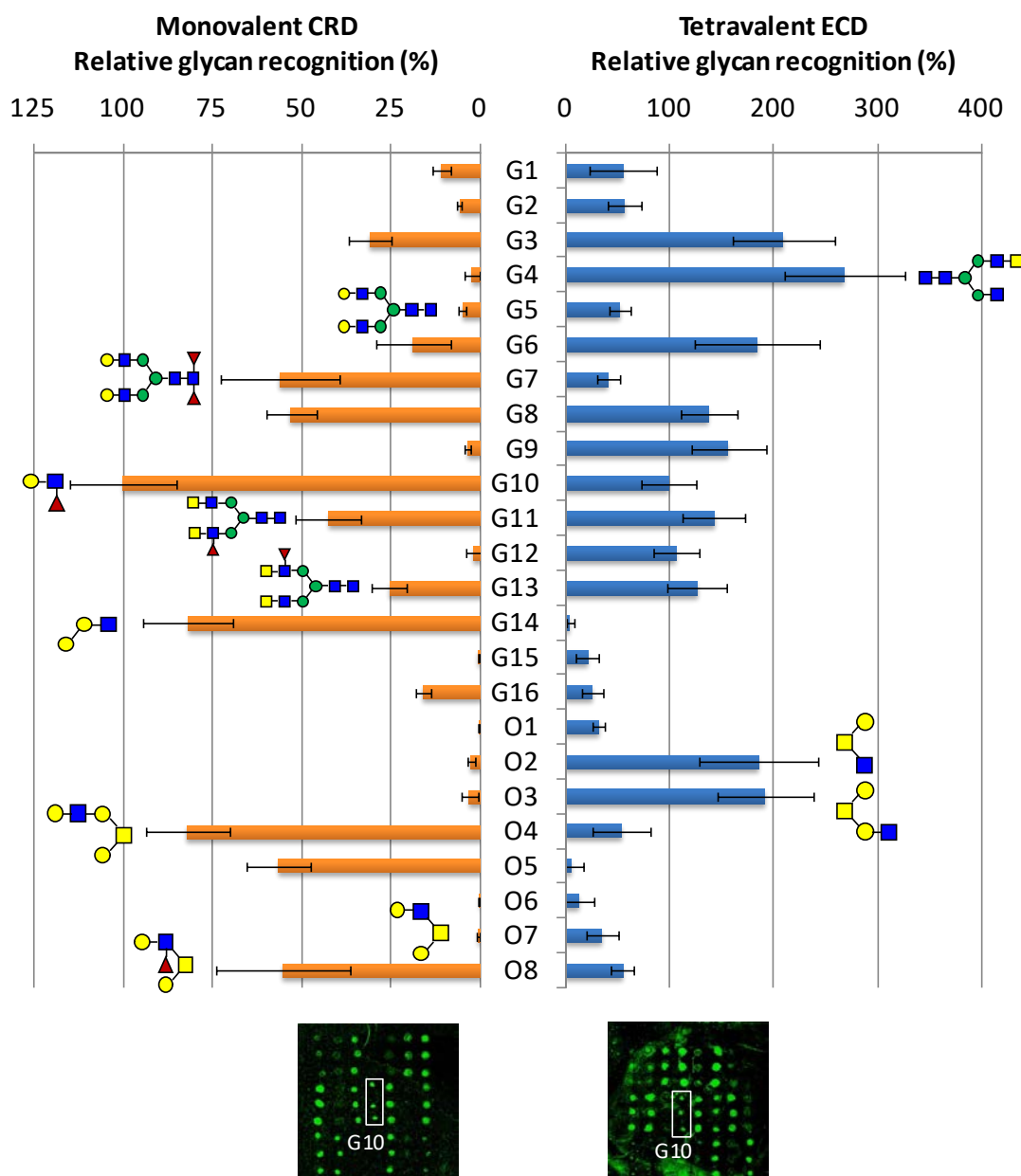


Figure 3. Relative fluorescence intensities and fluorescence image for the unmodified array after incubation with DC-SIGNR CRD (10 μ g/ml, orange bars) and DC-SIGNR ECD (50 μ g/ml, blue bars). Each bar in the histogram represents the average of fluorescence from 3 spots which were normalized to **G10** and the propagation of error is shown as an error bar.

The observed differences in binding profiles towards monovalent CRD and tetraivalent ECD receptor organization are striking and have been observed previously for other systems.^[44] Based on our array binding data we can only speculate on the underlying molecular basis determining the observed change in specificity. Rapid rebinding to multiple equivalent binding sites in close proximity could improve binding to some structures, while others could face additional binding penalties due to steric clashes in the tetrameric lectin organization. In addition, chelation could favor binding to compounds which are capable of interacting with more than one binding site. This observation could have practical implications for lectin binding studies as for simplicity and practical reasons these are often

carried out on the monomeric CRD, which is easier to express and purify. The natural presentation of DC-SIGNR, however, is determined by the organization via the stem region into defined tetramers. Any development of multivalent inhibitors should therefore take the natural receptor organization into account. DC-SIGNR, unlike DC-SIGN, binds to neoglycoconjugates presenting all Lewis antigens except Lewis X epitopes.^[42] We therefore turned our attention to DC-SIGN. Recently, we reported complementary binding selectivity towards one of the isomeric N glycans **G3/G8** and **G4/G9** by DC-SIGN and DC-SIGNR.^[43]

DC-SIGN is found on dendritic cells and macrophages as a tetrameric lectin and displays specificity for mannose, fucose and GlcNAc residues. N-glycans bearing Le^x and LDNF epitopes in the soluble egg antigens have been shown to bind to DC-SIGN and inhibit DC activation.^[45,46] Incubation of the unmodified initial array with fluorescently labeled DC-SIGN ECD showed strongest binding to glycans presenting Le^x (**G10, O8**) or LDNF (**G11-G13**) glycan determinants but not to the bis-core fucosylated N-glycan **G7** (**Figure 4**) possibly due to the impact of core fucosylation on glycan conformation and hence lectin recognition.^[47] Glycans **G3, G4** and **O3** presenting a free GlcNAc residue on the 6-arm barely showed interaction with the lectin. However, in a previous screen on a glycan array with different surface architecture they have been shown to bind to DC-SIGN with similar strength as LDNF structures.^[43] Among the strongest binders in our previously reported screen were bi-antennary glycans with terminal GlcNAc residues and high mannose structures, not included in the present array.

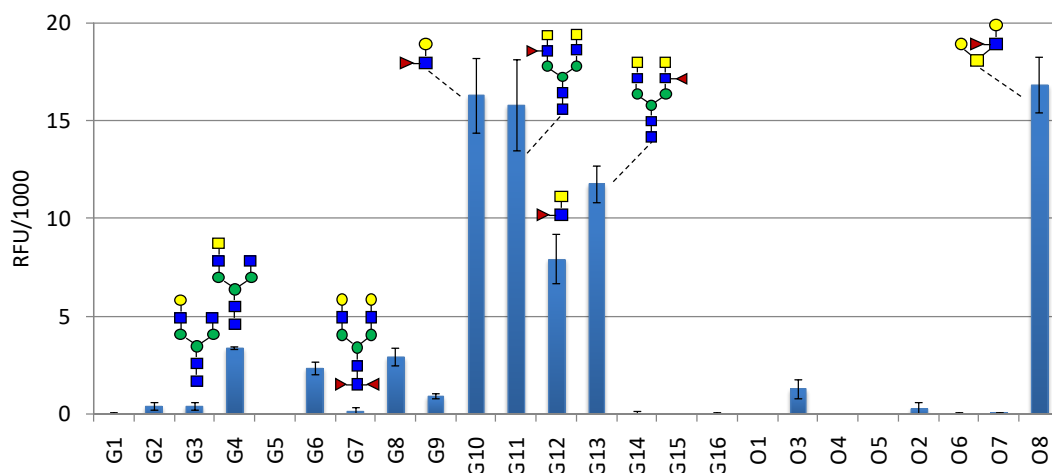


Figure 4. Binding profile of the array immobilized glycans with DC-SIGN ECD (10µg/ml). Histogram bars represent the average fluorescence from 3 spots and error bars represent the standard deviation.

After on-chip fucosylation with HP-FucT, DC-SIGN bound to all enzymatically extended structures, though the spot-to-spot variability increased considerably, possibly due to partial removal of the non-covalently attached glycoconjugates during prolonged washing procedures (**Figure 5**).

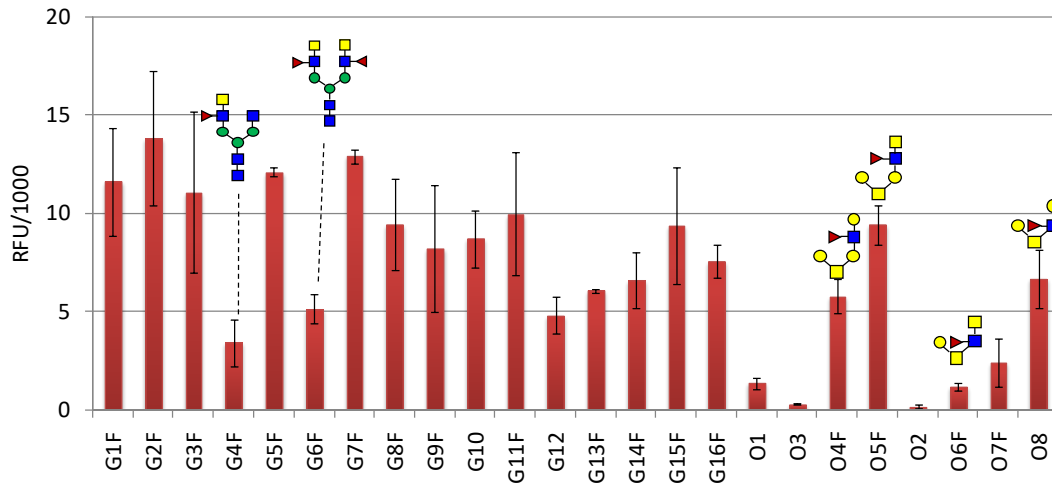


Figure 5. Glycan binding of glycan array with tetraivalent DC-SIGN ECD (10µg/ml) after on-chip fucosylation with HP-FucT. Each bar in the histogram represents the average of fluorescence from 3 spots and the standard deviation is given as an error bar.

Nevertheless, the population of fucosylated O-glycans in our array bound with intensity similar to the fucosylated N-glycans, making them equally as valuable in the analyses of CLR-glycan interaction, as previously stated by Van Diepen *et al.* [48]

Lewis X binding to DC-SIGN CRD is thought to be favoured by Van der Waals interactions between the fucose C-2 and the neighbouring V351.^[49] However, given the limitations in glycan-lectin interaction predictions, it was difficult to predict whether the azide modification at C-6 would affect glycan binding.^[50,51] Despite also showing an increased degree of variability, DC-SIGN bound to all azido-fucosylated structures of our array, indicating that the C-6 modification does not significantly contribute to glycan-CLR interaction (Figure 6). Additionally, it was interesting to note that glycans **G11Z** and **G13Z** containing an LDNF and azido-LDNF epitope were observed to be the highest binders.

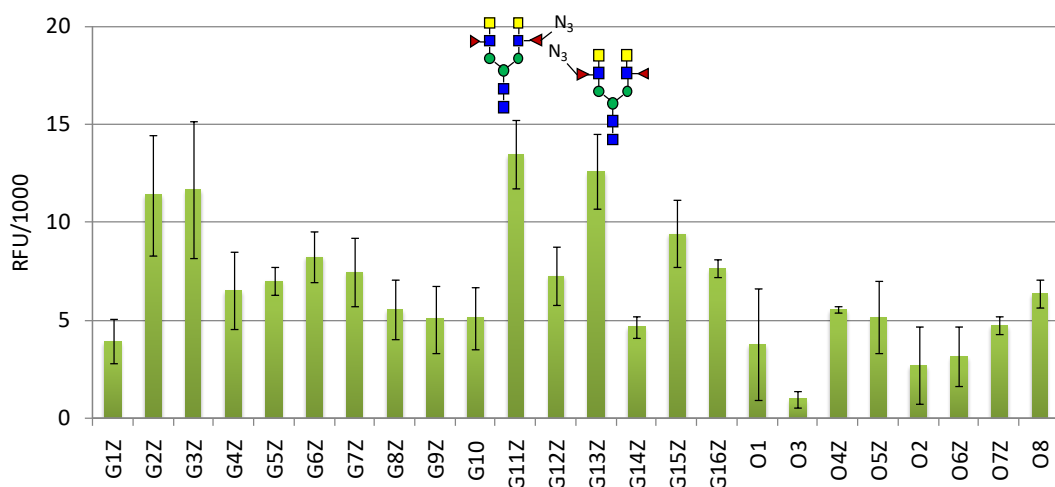


Figure 6. Glycan binding profile for DC-SIGN ECD (10µg/ml) after modification of the glycan array with the C6-azido-fucose. Each bar in the histogram represents the average fluorescence for 3 replicate spots and the standard deviations are given as an error bar.

Finally, we looked at macrophage galactose lectin (MGL) binding to our collection of immobilised O- and N-glycans. MGL is a PRR expressed at the cell surface of macrophages and dendritic cells that recognizes predominantly terminal GalNAc residues found on Tn antigen glycosylated mucins overexpressed on cancer cells or the GalNAc β 1-4GlcNAc (LDN) disaccharide motif of the parasite *S.mansoni*.^[52] Targeting MGL with high affinity glycans or glycomimetics has been suggested as a strategy for macrophage targeting in cancer immune therapy and drug delivery.^[53,54] In line with its reported binding specificity, we found strong interactions of MGL with all glycans presenting terminal GalNAc residues of the unmodified array (Figure 7). We observed a low degree of oligosaccharide context dependent alteration of the interaction of GalNAc residues with MGL. A similar binding strength was seen for the isolated LDN epitope **G16** and glycans presenting the LDN motif on one or both antennae.^[43] Glycan **O5** with an asymmetrically extended *S. mansoni* core, presenting a single GalNAc residue, was the strongest binding ligand on the array.

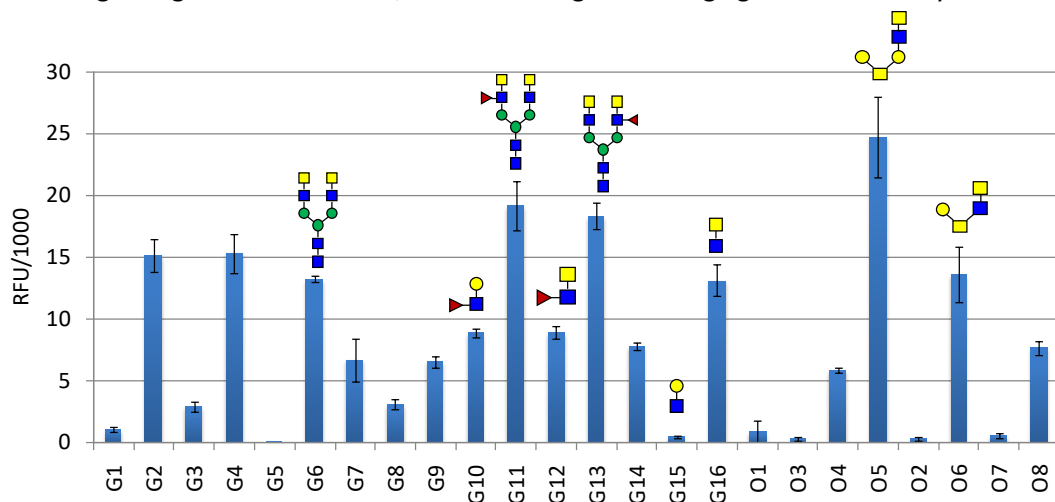


Figure 7. Glycan binding profile with MGL ECD (10 μ g/ml). Histogram bars represent the average fluorescence from 3 spots and the standard deviation is given as an error bar.

Molecular modelling studies suggest that the presence of a α 1,3 fucose residue would disrupt the chelation of the GalNAc 3-OH with calcium at the binding site but our MGL binding profile results does not support this generalized rule.^[55] Although lower binding was observed for the LDNF glycan **G12** than for the isolated LDN epitope **G16**, both monofucosylated bi-antennary glycans **G11** and **G13** showed higher fluorescence intensities than the non-fucosylated compound **G6**. In addition, we were surprised to observe binding of the Le^x epitope **G10** with a similar binding strength as for the LDNF epitope **G12**, neither of which was anticipated to bind MGL as fucosylation should inhibit binding. These results suggest that fucosylation may indeed, induce alternate binding modes for non-GalNAc ligands in MGL CRD.

On-chip enzymatic fucosylation led to a generally reduced binding to MGL with the exception of **O5F** and **G2F** which maintained high affinities (**Figure 8, a**).

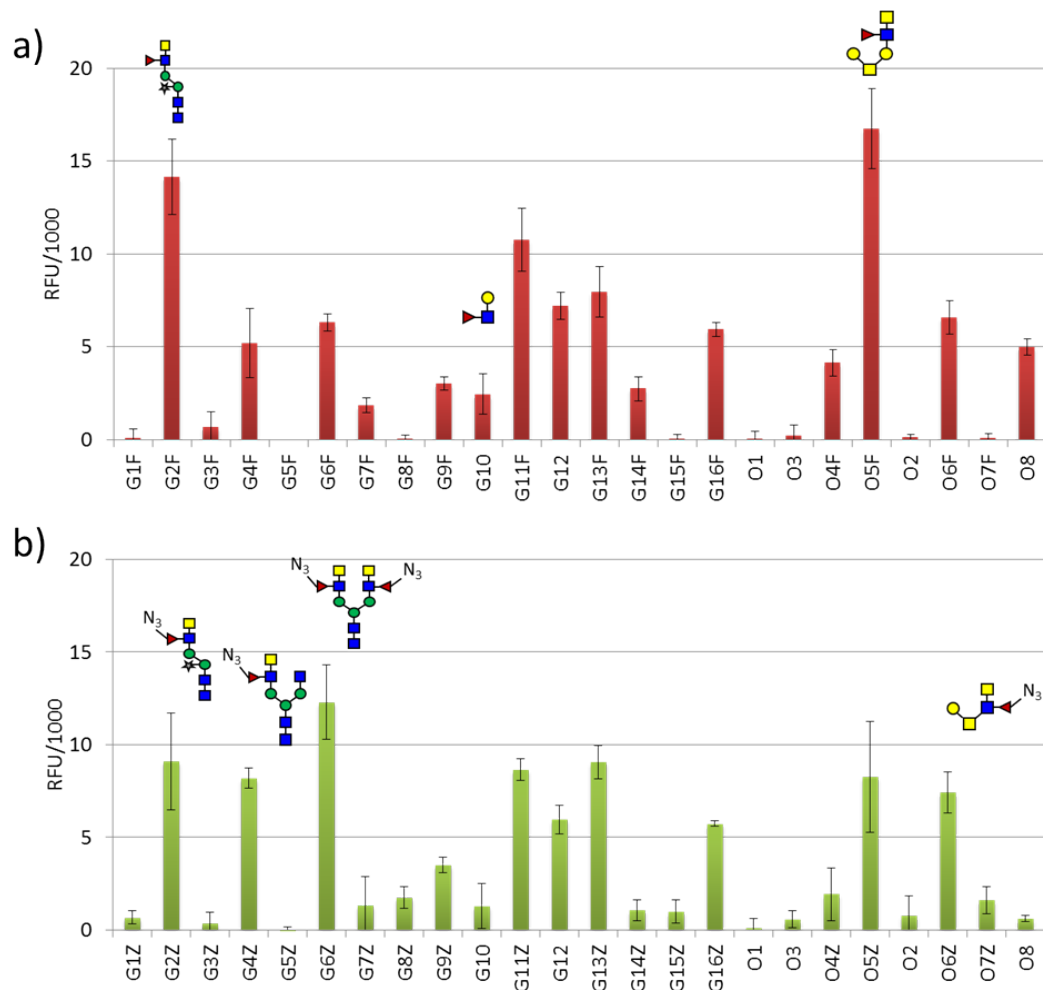


Figure 8. Fluorescence intensities for a) fucosylated γ and b) the azido-fucosylated glycan array after incubation with MGL ECD (10 μ g/ml). Histogram bars represent the average fluorescence from 3 spots and the standard deviation is given as an error bar.

Despite the spot-to-spot variability increasing considerably here too, the comparison of relative binding profiles showed that the elongation with 6-azido-fucose decrease binding to most of the immobilized glycans with the exception of **G2Z**, **G4Z** and **G6Z** which remained relatively unchanged (**Figure 8, b**).

Conclusions

We have chemo-enzymatically synthesized eight novel O-glycans based on the *S. mansoni* specific core and the mammalian mucin core 2. The unexpected regioselectivity for the 6-arm of the *S. mansoni* core, observed for the recombinant glycosyltransferase LgtA_X, provided a simple entry into products with enzymatic elongations on the 6-arm. The eight synthesized O-glycans were printed alongside a number of N-glycans and were enzymatically modified in good to excellent yields with fucose and non-natural 6-azido-fucose. Glycan array based preliminary binding studies with the human C-type lectin receptors MGL, DC-SIGNR and DC-SIGN showed interesting novel insights into the fine specificity of CLR-glycan binding. This study highlights the potential of employing glycan arrays to identify leads for the further development of glycomimetics with improved CLR

targeting properties. In addition, our comparison of the binding profiles of CRD and ECD domains of DC-SIGNR underscores the importance of multivalent receptor presentation for inducing binding with an additional level of glycan selectivity.

Experimental Part

Chemical Synthesis

5-(benzyl(benzyloxycarbonyl)amino)pentyl3,4,6-tri-O-acetyl-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino) β -D-galactopyranoside **3**

To a solution of **1**^[16] (11.8 g, 18.9 mmol) in dry DCM (30ml) on activated molecular sieves and under argon was added a solution of N-benzyl-N-(5-hydroxypentyl)carbamate (7.42 g, 22.7 mmol) **2** in dry DCM (200 ml). The solution was placed at 0°C before TMSOTf (0.690 ml, 3.8 mmol) was added dropwise. The reaction was allowed to warm to RT and stirred for 2 hours after which TLC showed full consumption of starting material. The reaction mixture was quenched with Et₃N and filtered through celite. The crude sample was purified by flash column chromatography (15→40% EtOAc:Tol) to yield **3** as a white gum, 11 g, 73%. ¹H NMR (500 MHz, CDCl₃) δ 7.41 – 7.23 (m, 9H, Ar), 7.17 (d, *J* = 8.1 Hz, 1H, Ar), 5.36 (d, *J* = 3.3 Hz, 1H, H-4), 5.25 – 5.09 (m, 3H, H-3, CH₂Bn), 4.79 – 4.59 (m, 2H, CH₂Troc), 4.59 – 4.40 (m, 3H, H-1, CH₂Bn), 4.14 (qd, *J* = 11.2, 6.7 Hz, 2H, 2H-6), 3.94 – 3.70 (m, 3H, H-5, H-2, CH_{linker}), 3.50 – 3.25 (m, 1H, CH_{linker}), 3.23 – 3.12 (m, 2H, CH₂), 2.14 (s, 3H, OCH₃), 2.05 (s, 3H, OCH₃), 1.99 (s, 3H, OCH₃), 1.59 – 1.43 (m, 4H, 2CH_{2linker}), 1.39 – 1.18 (m, 2H, CH_{2linker}). ¹³C NMR (126 MHz, CDCl₃) δ 170.45, 170.37, 156.77, 156.30, 154.38, 137.79, 136.74, 128.57, 128.44, 127.92, 127.81, 127.73, 127.35, 127.20, 101.25, 95.67, 74.30, 70.50, 69.86, 69.73, 67.21, 66.76, 61.48, 52.73, 50.48, 50.37, 47.28, 46.10, 29.04, 28.66, 27.76, 26.99, 22.96, 20.69, 20.65. HRMS (MALDI-ToF): *m/z* calcd. for C₃₅H₄₃Cl₃N₂O₁₂ [M+Na]⁺ : 811.1778, found: 811.1728. [α]_D²⁰ = -11.8° (c=1, CHCl₃)

5-(benzyl(benzyloxycarbonyl)amino)pentyl4,6-O-benzylidene-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino) β -D-galactopyranoside **4**

Compound **3** (10.2 g, 12.9 mmol) was dissolved in G/GHNO₃^[18] (700 ml) and the solution was stirred at RT for 15 mins until full consumption of starting material was observed by TLC. The reaction mixture was concentrated *in vacuo* then washed thoroughly with DCM and filtered through celite. The filtrate was concentrated to yield a pale pink residue. This was dissolved in dry acetonitrile (100 ml) under argon and benzaldehyde dimethyl acetal (3.4 ml, 2.25 mmol) and camphor sulfonic acid (600 mg, 2.6 mmol) were added. The reaction was stirred at RT overnight after which successful conversion was observed by TLC. Et₃N was added and the reaction mixture was concentrated *in vacuo*. The resulting residue was purified by flash column chromatography (3→10% MeOH:DCM) to yield **4** as a white solid, 6.78 g, 70%. ¹H NMR (500 MHz, CDCl₃) δ 7.51 (d, *J* = 5.2 Hz, 2H, Ar), 7.42 – 7.22 (m, 11H, Ar), 7.17 (d, *J* = 7.0 Hz, 2H, Ar), 5.57 (s, 1H, CH_{Ph}), 5.23 – 5.13 (m, 2H, CH₂Ph), 4.72 (d, *J* = 11.7 Hz, 2H, CH₂Troc), 4.50 (m, 3H, H-1, CH₂Bn), 4.32 (d, *J* = 12.4 Hz, 1H, H-6_a), 4.20 (d, *J* = 3.6 Hz, 1H, H-4), 4.08 (dd, *J* = 12.4, 1.9 Hz, 1H, H-6_b), 4.03 – 3.82 (m, 2H, H-3, CH₂), 3.65 (m, 1H, H-2), 3.5 (s, 1H, H-5), 3.43-3.25 (m, 1H, CH_{linker}), 3.25-3.15 (m, 2H, CH₂), 1.62 – 1.41 (m, 2H, 2CH_{2linker}), 1.39 – 1.16 (m, 2H, CH_{2linker}). ¹³C NMR (126 MHz, CDCl₃) δ 152.34, 151.88, 150.76, 150.49, 134.39, 134.36, 134.34, 134.15, 133.47, 133.32, 126.34, 125.72, 125.69, 125.60,

125.43, 125.16, 125.10, 124.98, 124.88, 124.57, 124.54, 124.39, 123.71, 99.78, 99.53, 99.33, 94.60, 75.14, 75.07, 74.49, 70.66, 70.37, 69.68, 69.62, 69.52, 67.75, 67.72, 66.93, 56.43, 51.84, 51.71, 48.87, 47.78, 31.72, 31.33, 30.52, 29.75, 26.02, 25.90. HRMS (MALDI-Tof) m/z calcd. for $C_{36}H_{41}Cl_3N_2O_9$ $[M+Na]^+$: 773.1774, found: 773.1772. $[\alpha]_D^{20} = -3.5^\circ$ ($c=1$, $CHCl_3$)

5-(benzyl (benzyloxycarbonyl)amino) pentyl 2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-4,6-O-benzylidene-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino) β -D-galactopyranoside 5

To a solution of **4** (6.26 g, 8.3mmol) and **7**^[56](5.0 g, 10.1 mmol) in dry DCM (120 ml) on activated molecular sieves at $-40^\circ C$ and under argon was added TMSOTf (0.301 ml, 1.7 mmol) and the reaction was stirred at -20 for 1hr as monitored by TLC. Triethylamine was added and the reaction was filtered through celite then concentrated to be purified by FCC (20 \rightarrow 70% EtOAc:Hex) yielding **5** as a white foam, 5.6 g, 62%. 1H NMR (500 MHz, $CDCl_3$) δ 7.55 (d, $J = 7.6$ Hz, 2H, Ar), 7.40 – 7.20 (m, 12H, Ar), 7.16 (d, $J = 7.3$ Hz, 1H, Ar), 5.57 (s, 1H, CH_{Ph}), 5.36 (d, $J = 3.4, 1.2$ Hz, 1H, H-4'), 5.21 (dd, 1H, H-2'), 5.16 (s, 2H, CO_2CH_{2Bn}), 4.97 (dd, $J = 10.4, 3.5$ Hz, 1H, H-3'), 4.90 – 4.80 (m, 2H, H-1, $CHCl_3$), 4.77 (d, $J = 8.0$ Hz, 1H, H-1'), 4.65 – 4.55 (m, 1H, $CHCl_3$), 4.47 (s, 3H, H-3, NCH_{2Bn}), 4.34 – 4.28 (m, 2H, H-6_a, H-4), 4.15 (ddd, 2H, 2H-6'), 4.06 (dd, $J = 12.4, 1.8$ Hz, 1H, H-6_b), 3.92 (s, 1H, CH_{linker}), 3.87 (td, 1H, H-5'), 3.53 (s, 1H, H-2), 3.44 (s, 1H, H-5), 3.39 (s, 1H, CH_{linker}), 3.23 (s, 2H, $CH_{2linker}$), 2.16 (s, 3H, Ac), 2.04 (d, $J = 5.7$ Hz, 6H, 2 CH_3), 1.97 (s, 3H, CH_3), 1.52 (d, $J = 28.7$ Hz, 4H, $CH_{2linker}$), 1.36 – 1.23 (m, 2H, $CH_{2linker}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.31, 170.08, 169.34, 156.77, 156.22, 154.05, 137.86, 136.86, 136.74, 128.84, 128.55, 128.47, 128.11, 127.93, 127.80, 127.30, 127.19, 126.26, 101.86, 100.69, 99.74, 95.61, 76.01, 74.23, 70.87, 70.82, 69.69, 69.41, 69.19, 68.87, 67.17, 67.08, 66.45, 61.56, 53.82, 50.35, 47.26, 45.95, 29.68, 29.08, 28.81, 27.77, 27.27, 23.31, 23.06, 20.71, 20.56. HRMS $[M+Na]^+$ (MALDI-Tof) m/z calcd. for $C_{50}H_{63}Cl_3N_2O_{18}$ $[M+Na]^+$: 1103.2724 found: 1103.2716. $[\alpha]_D^{20} = +13.9^\circ$ ($c=1$, $CHCl_3$)

5-(benzyl (benzyloxycarbonyl)amino) pentyl 2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-4-O-benzyl-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino) β -D-galactopyranoside 6

To a solution of **5** (5.6 g, 5.2 mmol) in dry DCM (35ml) under Ar at $0^\circ C$ was added 1M $BH_3.THF$ (20.7 ml, 20.7 mmol) dropwise. The solution was allowed to cool before adding TMSOTf (0.467 ml, 2.59 mmol) dropwise and the reaction was stirred at $0^\circ C$ under Ar for 1.5 hours after which full conversion was observed by TLC. The ice bath was removed and the solution was quenched with a solution of MeOH:Et₃N (10:1) until effervescence ceased. The reaction mixture was concentrated *in vacuo* and purified by FCC (50 \rightarrow 100% EtOAc:Hex) to yield **6** as a white solid, 3.4 g, 61%. 1H NMR (500 MHz, $CDCl_3$) δ 7.43 (d, $J = 6.8$ Hz, 2H, Ar), 7.39 – 7.13 (m, 13H, Ar), 5.43 – 5.38 (s, 1H, H-4'), 5.27 (dd, $J = 10.5, 8.0$ Hz, 1H, H-2'), 5.17 (d, $J = 10.9$ Hz, 2H, OCH_{2Bn}), 4.99 (dd, $J = 10.5, 3.4$ Hz, 1H, H-3'), 4.92 (d, $J = 11.8$ Hz, 1H, 4- OCH_{2Bn}), 4.84 – 4.64 (m, 5H, CH_{2Troc} , 4- OCH_{2Bn} , H-1, H-1'), 4.47 (d, $J = 8.8$ Hz, 3H, H-3, NCH_{2Bn}), 4.17 (ddq, $J = 15.9, 11.1, 6.1, 4.7$ Hz, 2H, H-6'), 3.98 – 3.65 (m, 4H, H5', H-4, CH_{linker} , H-6_a), 3.53 – 3.14 (m, 6H, H-6_b, H-2, H-5, CH_{linker} , $CH_{2linker}$), 2.14 (s, 3H, CH_3), 2.10 (s, 3H, CH_3), 2.02 (s, 3H, CH_3), 1.99 (s, 3H, CH_3), 1.53 (d, $J = 18.6$ Hz, 4H, 2 $CH_{2linker}$), 1.31 (s, 2H, $CH_{2linker}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.41, 170.15, 170.07, 169.51, 138.31, 137.82, 129.09,

128.55 , 128.46 , 128.35 , 128.03 , 127.94 , 127.79 , 127.30 , 127.19 , 102.45 , 99.95 , 95.60 , 78.73 , 74.53 , 74.42 , 74.22 , 73.97 , 73.85 , 70.74 , 70.68 , 69.88 , 69.53 , 69.04 , 67.20 , 67.09 , 61.79 , 61.27 , 54.89 , 50.37 , 47.24 , 45.97 , 29.12 , 28.75 , 27.78 , 27.19 , 23.28 , 23.07 , 20.72 , 20.63 , 20.56 . HRMS (MALDI-Tof) m/z calcd. for $C_{50}H_{61}Cl_3N_2O_{18}$ $[M+Na]^+$: 1105.2881, found: 1105.2839. $[\alpha]_D^{20} = -18.4^\circ$ ($c=1$, $CHCl_3$)

5-(benzyl (benzyloxycarbonyl)amino) pentyl 2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 6)]-4-O-benzyl-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino) β -D-galactopyranoside 9

To a solution of **6** (470 mg, 0.434 mmol) and **7** (256 mg, 0.520 mmol) in dry DCM (5ml) on activated molecular sieves, under argon and at $-60^\circ C$ was added TMSOTf (12 μ l, 65.1 μ mol) dropwise. The reaction was stirred between $-40^\circ C$ and $-20^\circ C$ for 1hr as monitored by TLC. Triethylamine was added and the reaction was filtered through celite then concentrated *in vacuo* before purification by FCC (40 \rightarrow 100% EtOAc:Hex) to yield **9** as a white foam, 252 mg, 41%. 1H NMR (500 MHz, $CDCl_3$) δ 7.43 – 7.21 (m, 14H, Ar), 7.16 (d, $J = 7.4$ Hz, 1H, Ar), 5.99 (s, 0.5H, NH_{Troc}), 5.42 – 5.35 (m, 2H, H-4', H-4''), 5.31 (d, $J = 10.4$ Hz, 0.5H, NH_{Troc}), 5.24 (dd, $J = 10.5, 7.9$ Hz, 1H, H-2'), 5.21 – 5.11 (m, 3H, H-2', OCH_{2Bn}), 5.01 – 4.89 (m, 3H, H-3, H-3', CH_{Troc}), 4.83 – 4.62 (m, 6H, H-1, H-1', 4- OCH_{2Bn} , CH_{Troc}), 4.58 – 4.40 (m, 4H, H-1'', H-3, NCH_{2Bn}), 4.21 – 4.09 (m, 4H, H-6', H-6''), 3.96 – 3.61 (m, 6H, 2H-6, H-4, H-5', H-5'', CH_{linker}), 3.60 – 3.54 (m, 1H, H-5), 3.52 – 3.14 (m, 4H, H-2, 3 CH_{linker}), 2.15 – 1.95 (m, 24H, 8 CH_3), 1.52 (d, $J = 26.0$ Hz, 4H, 2 $CH_{2linker}$), 1.36 – 1.26 (m, 2H, $CH_{2linker}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.49, 170.43, 170.34, 170.27, 170.22, 170.19, 169.57, 169.30, 156.35, 154.23, 138.35, 137.96, 129.07, 128.66, 128.58, 128.28, 128.07, 127.97, 127.90, 127.42, 127.30, 102.46, 101.35, 99.88, 78.38, 75.15, 74.36, 70.95, 70.82, 70.78, 70.75, 69.14, 69.06, 67.29, 67.15, 67.03, 61.32, 61.08, 55.02, 50.50, 47.45, 29.81, 29.24, 28.84, 27.94, 27.42, 23.60, 23.28, 20.94, 20.85, 20.79, 20.76, 20.68. HRMS (MALDI-Tof) m/z calcd. for $C_{64}H_{79}Cl_3N_2O_{27}$ $[M+Na]^+$: 1435.383, found: 1435.3923. $[\alpha]_D^{20} = -11.3^\circ$ ($c=1$, $CHCl_3$)

5-(benzyl (benzyloxycarbonyl)amino) pentyl 2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-[3,4,6-tri-O-acetyl-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino)- β -D-glucopyranosyl-(1 \rightarrow 6)]-4-O-benzyl-2-deoxy-2-((2,2,2-trichloroethoxycarbonylamino) β -D-galactopyranoside 10

To a solution of **6** (420 mg, 0.387 mmol) and with the imidate **8** (291 mg, 0.465 mmol) in dry DCM (5ml) on activated molecular sieves, under argon and at $-60^\circ C$ was added TMSOTf (10.5 μ L, 58.1 μ mol). The reaction was stirred between $-40^\circ C$ and $-20^\circ C$ for 1hr until full reaction was observed as monitored by TLC. Triethylamine was added and the reaction was filtered through celite. The reaction crude was then concentrated *in vacuo* and purified by FCC (40 \rightarrow 60% EtOAc:Hex) yielding **10** as a white solid, 438 mg, 73%. 1H NMR (500 MHz, $CDCl_3$) δ 7.42 – 7.13 (m, 15H, Ar), 5.98 (s, 0.5H, NH_{Troc}), 5.45 (s, 0.5H, NH_{Troc}), 5.39 (s, 1H, H-4'), 5.30 – 5.13 (m, 4H, H-2', H-3'', OCH_{2Bn}), 5.06 (t, $J = 9.6$ Hz, 1H, H-4''), 4.98 (dd, $J = 10.5, 3.5$ Hz, 1H, H-3'), 4.91 (d, $J = 11.3$ Hz, 1H, CH_{Troc}), 4.84 – 4.59 (m, 6H, CH_{Troc} , 4- OCH_{2Bn} , H-1, H-1', H-1''), 4.46 (dt, $J = 22.7, 12.8$ Hz, 3H, NCH_{2Bn} , H-3), 4.27 (dd, $J = 12.3, 4.7$ Hz, 2H, H-6'_a, H-6''_a), 4.12 (ddd, $J = 26.1, 11.8, 4.5$ Hz, 2H, H-6'_b, H-6''_b), 3.96 – 3.25 (m, 10H, H-4, H-5, H-5', H-5'', 2H-6, $CH_{2linker}$, H-2, H-2''), 3.19 (s, 2H, $CH_{2linker}$), 2.16 – 1.95 (m, 21H, 7 CH_3), 1.65 – 1.44 (m, 4H, 2 $CH_{2linker}$), 1.38 – 1.28 (m, 2H, $CH_{2linker}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.73, 170.58,

170.22, 170.12, 169.52, 156.32, 154.11, 138.38, 137.83, 136.67, 128.90, 128.58, 128.49, 128.18, 127.97, 127.79, 127.34, 127.23, 102.31, 101.03, 100.07, 95.63, 78.22, 75.09, 74.40, 74.23, 72.18, 71.82, 70.79, 70.70, 69.00, 68.50, 67.24, 67.09, 62.03, 61.18, 56.38, 54.80, 50.39, 47.35, 28.83, 27.81, 27.20, 23.37, 20.75, 20.68, 20.65, 20.59. HRMS (MALDI-Tof) m/z calcd. for $C_{65}H_{79}Cl_6N_3O_{27}$ $[M+Na]^+$: 1566.2928, found: 1566.2886. $[\alpha]_D^{20} = -6.2^\circ$ (c=1, $CHCl_3$)

5-(benzyl(benzyloxycarbonyl)amino)pentyl β -D-galactopyranosyl-(1 \rightarrow 3)-[β -D-galactopyranosyl-(1 \rightarrow 6)]-4-O-benzyl-2-deoxy-2-acetamido- β -D-galactopyranoside **11**

To a solution of **9** (154 mg, 0.109 mmol) in THF (3ml) was added TBAF (1M in THF, 0.163 ml, 0.163 mmol) dropwise. The yellow reaction mixture was refluxed for 1.5 hours after which it was placed on ice and quenched with MeOH. The reaction mixture was concentrated to dryness *in vacuo* before being redissolved in anhydrous pyridine (1 ml). Acetic anhydride (0.3 ml, 3.27 mmol) was added and the reaction was left to stir overnight. MeOH was added and the reaction mixture was concentrated *in vacuo*. The crude was purified by FCC (50 \rightarrow 100% EtOAc:Hex) to yield 5- (benzyl (benzyloxycarbonyl)amino) pentyl 2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-[2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 6)]-4-O-benzyl-2-deoxy-2- acetamido- β -D-galactopyranoside a pale yellow solid, 125 mg, 90%. 1H NMR (500 MHz, Chloroform-d) δ 7.43 – 7.13 (m, 15H,Ar), 5.38 (dd, J = 12.7, 3.5, 1.1 Hz, 2H, H-4',H-4''), 5.24 (t, J = 9.2 Hz, 1H,H-2'), 5.19 – 5.11 (m, 3H, OCH_{2Bn} , H-2''), 5.02 (dd, J = 10.4, 3.4 Hz, 1H, H-3'), 4.95 (dd, J = 10.5, 3.4 Hz, 1H, H-3''), 4.88 (dd, 2H, H-1, 4- OCH_{Bn}), 4.75 – 4.63 (m, 3H, H-1', H-3, 4- OCH_{Bn}), 4.60 – 4.38 (m, 3H, H-1'', NCH_{2Bn}), 4.22 – 4.09 (m, 4H, 2H-6', 2H-6''), 3.96 – 3.55 (m, 7H, H-4, H-5, H-5', H-5'', 2H-6, CH_{linker}), 3.43 – 3.28 (m, 2H, H-2, CH_{linker}), 3.16 (s, 2H, $CH_{2linker}$), 2.15 – 1.93 (m, 27H, 9 CH_3), 1.54 (s, 4H, 2 $CH_{2linker}$), 1.37 – 1.21 (m, 2H, $CH_{2linker}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.37, 170.31, 170.25, 170.17, 170.10, 170.07, 169.18, 138.26, 137.74, 101.95, 101.18, 99.26, 77.88, 75.12, 74.13, 70.83, 70.58, 69.88, 69.21, 68.91, 67.18, 67.06, 67.02, 61.26, 60.98, 55.21, 50.30, 47.29, 45.91, 29.63, 29.10, 28.71, 27.41, 23.59, 23.12, 20.81, 20.64, 20.61, 20.54. HRMS (MALDI-Tof) m/z calcd. for $C_{63}H_{80}N_2O_{26}$ $[M+Na]^+$: 1303.4893, found: 1303.5106, $[\alpha]_D^{20} = -20.5$ (c=1 $CHCl_3$). The product (120 mg, 94 μ mol) was redissolved in dry MeOH (4 ml) and 0.5M NaOMe (1.5 ml, 750 μ mol) was added. The RM was stirred at RT for 1 hour after which it was quenched with Amberlite[®]IR 120 (H). The filtrate was concentrated, purified by Sephadex LH-20 (D=3.5 cm, H=45 cm, MeOH) and lyophilized to yield **11** as a white powder, 87.5 mg, 99%. 1H NMR (500 MHz, Methanol-d4) δ 7.46 – 7.16 (m, 15H,Ar), 5.15 (d, J = 16.3 Hz, 2H, OCH_{2Bn}), 4.99 (d, J = 11.6 Hz, 1H, 4- OCH_{Bn}), 4.70 (d, J = 11.6 Hz, 1H, 4- OCH_{Bn}), 4.50 (s, 2H, NCH_{2Bn}), 4.47 (s, 1H, H-1), 4.30 (d, J = 7.6 Hz, 1H, H-1'), 4.24 (d, J = 7.5 Hz, 1H, H-1''), 4.08 (dd, J = 15.4, 6.2 Hz, 2H, H-2, H-4), 3.90 (dd, J = 10.9, 2.9 Hz, 1H, H-3), 3.87 – 3.69 (m, 10H, H-4', H-4'', 2H-6, 2H-6', 2H-6'', H-5, CH_{linker}), 3.58 (dd, J = 9.7, 7.6 Hz, 1H, H-2'), 3.48 (tdd, J = 20.9, 10.2, 7.3 Hz, 7H, H-2'', H-3', H-3'', H-5',H-5'',H-2'', CH_{linker}), 3.22 (dd, J = 14.3, 6.9 Hz, 2H, $CH_{2linker}$), 1.97 – 1.89 (m, 3H, CH_3), 1.57 – 1.42 (m, 4H, 2 $CH_{2linker}$), 1.28 – 1.21 (m, 2H, $CH_{2linker}$). ^{13}C NMR (126 MHz, MeOD) δ 174.30, 140.51, 129.59, 129.55, 129.05, 128.36, 107.18, 105.20, 102.59, 81.67, 77.15, 76.96, 76.54, 75.63, 75.13, 74.97, 74.56, 72.63, 72.54, 70.28, 70.09, 69.83, 68.47, 62.85, 62.26, 53.74, 49.85, 30.24, 24.24, 23.27. HRMS (MALDI-Tof) m/z calcd. for $C_{47}H_{64}N_2O_{23}$ $[M+Na]^+$: 967.4052, found: 967.408, $[\alpha]_D^{20} = -5.7$ (c= 1, MeOH)

5-(benzyl(benzyloxycarbonyl)amino)pentyl β -D-galactopyranosyl-(1 \rightarrow 3)-[2-deoxy-2-acetamido- β -D-glucopyranosyl-(1 \rightarrow 6)]-4-O-benzyl-2-deoxy-2-acetamido- β -D-galactopyranoside **12**

To a solution of **10** (390 mg, 0.253 mmol) in THF (6ml) was added TBAF (1M in THF, 0.607 ml, 0.607 mmol) dropwise. The yellow reaction mixture was refluxed for 1.5 hours after which it was placed on ice and quenched with MeOH. The reaction mixture was concentrated to dryness *in vacuo* before being redissolved in anhydrous pyridine (3ml). Acetic anhydride (0.6 ml, 6.32 mmol) was added and the reaction was left to stir overnight. MeOH was added and the reaction mixture was concentrated *in vacuo*. The crude was purified by FCC (0 \rightarrow 5% MeOH:DCM) to yield 5- (benzyl (benzyloxycarbonyl)amino) pentyl 2,3,4,6-tetra-O-acetyl- β -D-galactopyranosyl-(1 \rightarrow 3)-[3,4,6-tri-O-acetyl-2-deoxy-2-acetamido- β -D-glucopyranosyl-(1 \rightarrow 6)]-4-O-benzyl-2-deoxy-2-acetamido- β -D-galactopyranoside as a brown solid, 226 mg, 70%. ^1H NMR (500 MHz, CDCl_3) δ 7.43 – 7.13 (m, 15H, Ar), 5.39 (d, J = 3.5 Hz, 1H, H-4'), 5.26 – 5.10 (m, 4H, H-2', H-3'', $\text{OCH}_{2\text{Bn}}$), 5.08 – 4.98 (m, 2H, H-4'', H-3'), 4.87 (d, J = 11.4 Hz, 1H, 4- OCH_{Bn}), 4.82 (d, J = 8.3 Hz, 1H, H-1), 4.64 (dt, J = 27.6, 9.0 Hz, 4H, 4- OC_{HBn} , H-1', H-1'', H-3), 4.49 (m, J = 17.7, 16.6 Hz, 2H, $\text{NCH}_{2\text{Bn}}$), 4.30 – 4.03 (m, 4H, 2H-6', 2H-6''), 3.97 – 3.71 (m, 5H, H-2'', H6_a, H-4, H-5', $\text{CH}_{\text{linker}}$), 3.70 – 3.57 (m, 3H, H6_b, H-5'', H-5), 3.52 – 3.25 (m, 3H, H-2, $\text{CH}_{2\text{linker}}$), 3.24 – 3.11 (m, 2H, $\text{CH}_{2\text{linker}}$), 2.15 – 1.84 (m, 24H), 1.64 – 1.42 (m, 4H, $\text{CH}_{2\text{linker}}$), 1.36 – 1.26 (m, 2H, $\text{CH}_{2\text{linker}}$). ^{13}C NMR (126 MHz, CDCl_3) δ 170.88, 170.75, 170.22, 170.12, 169.44, 169.23, 138.52, 137.76, 128.86, 128.60, 128.51, 128.12, 127.72, 127.68, 127.38, 127.27, 101.87, 100.95, 99.53, 75.34, 74.33, 73.63, 72.62, 71.76, 70.88, 70.77, 69.66, 69.26, 68.96, 68.53, 67.23, 67.18, 62.08, 61.17, 54.88, 50.35, 47.32, 28.82, 27.34, 23.66, 23.30, 20.85, 20.79, 20.75, 20.71, 20.68, 20.65, 20.59. HRMS (MALDI-Tof) m/z calcd. for $\text{C}_{63}\text{H}_{81}\text{N}_3\text{O}_{25}$ [$\text{M}+\text{Na}$] $^+$: 1302.5057, found: 1302.5106, $[\alpha]_{\text{D}}^{20} = -18.2^\circ$ ($c = 1$, CHCl_3). The product (218 mg, 0.170 mmol) was redissolved in dry MeOH (9 ml) and 0.5M NaOMe (2.5 ml, 1.24 mmol) was added. The RM was stirred at RT for 1 hour after which it was quenched with Amberlite[®]IR 120 (H). The filtrate was concentrated, purified by Sephadex LH-20 (D=3.5cm, H=45cm, MeOH) and lyophilized to yield **12** as a white powder, 153 mg, 100%. ^1H NMR (500 MHz, Methanol- d_4) δ 7.45 – 7.40 (m, 2H, Ar), 7.40 – 7.20 (m, 12H, Ar), 7.18 (s, 1H, Ar), 5.15 (d, J = 15.8 Hz, 2H, $\text{OCH}_{2\text{Bn}}$), 4.98 (d, J = 11.6 Hz, 1H, 4- OCH_{Bn}), 4.67 (d, J = 11.5 Hz, 1H, 4- OCH_{Bn}), 4.50 (s, 2H, $\text{NCH}_{2\text{Bn}}$), 4.41 (brs, 1H, H-1), 4.36 (d, J = 8.4 Hz, 1H, H-1''), 4.30 (d, J = 7.6 Hz, 1H, H-1'), 4.06 (brs, 1H, H-2), 4.02 (d, J = 3.1 Hz, 1H, H-4), 3.91 – 3.73 (m, 6H, H6'_a, H6''_b, H6_c, H-3, H-4'), 3.71 – 3.54 (m, 5H, H6'_a, H6''_b, H6_c, H-5, H-2', H-2'', $\text{CH}_{\text{linker}}$), 3.54 – 3.50 (m, 1H, H-5'), 3.49 – 3.33 (m, 3H, H-3', H-3'', $\text{CH}_{\text{linker}}$), 3.24 (m, 3H, H-5', $\text{CH}_{2\text{linker}}$), 1.92 (d, 6H, 2CH), 1.52 (s, 4H, 2 $\text{CH}_{2\text{linker}}$), 1.29 (s, 2H, $\text{CH}_{2\text{linker}}$). ^{13}C NMR (126 MHz, MeOD) δ 174.24, 173.63, 158.36, 157.84, 140.36, 139.09, 137.95, 129.55, 129.48, 129.03, 128.88, 128.63, 128.36, 128.28, 107.00, 102.61, 102.42, 81.45, 77.79, 77.30, 76.93, 75.98, 75.65, 74.75, 74.47, 72.57, 71.99, 70.24, 70.12, 69.74, 68.29, 62.82, 62.67, 57.18, 53.64, 51.42, 51.22, 49.85, 47.44, 30.14, 28.85, 28.35, 24.19, 23.28, 23.20. HRMS (MALDI-Tof) m/z calcd. for $\text{C}_{49}\text{H}_{67}\text{N}_3\text{O}_{18}$ [$\text{M}+\text{Na}$] $^+$: 1008.4317, found: 1008.437, $[\alpha]_{\text{D}}^{20} = -6.3$ ($c=0,1$, MeOH)

Acknowledgements

J.P., S.A., A.C., N.C.R, and F.F. were supported by the EU Horizon 2020 Research and Innovation Program (Marie Skłodowska-Curie Grant 642870, ETN-Immunoshape). N.C.R additionally acknowledges funding from the Ministry of Science and Education (MINECO) Grant No. CTQ2017-90039-R and RTC-2017-6126-1 and the Maria de Maeztu Units of Excellence Program from the Spanish State Research Agency-Grant No.MDM-2017-0720.F.F.acknowledge also the French Agence Nationale de la Recherche (ANR) PIA for Glyco@Alps (ANR- 15-IDEX-02).The Multistep Protein Purification Platform (MP3) was exploited for human DC-SIGN/R and MGL ECD production with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology.

References

- [1] J. C. Paulson, O. Blixt, B. E. Collins, *Nat. Chem. Biol.* **2006**, *2*, 238–248.
- [2] J. C. Hoving, G. J. Wilson, G. D. Brown, *Cell. Microbiol.* **2014**, *16*, 185–194.
- [3] B. Miguel, M. Celeste, M. Teresa, *Protein Kinases* **2012**, DOI 10.5772/37771.
- [4] E. J. Pearce, A. S. MacDonald, *Nat. Rev. Immunol.* **2002**, *2*, 499–511.
- [5] R. R. White, K. Artavanis-Tsakonas, *Virulence* **2012**, *3*, 668–677.
- [6] L. M. Kuijk, I. Van Die, *IUBMB Life* **2010**, *62*, 303–312.
- [7] H. Stepan, M. Pabst, F. Altmann, H. Geyer, R. Geyer, E. Staudacher, *Glycoconj. J.* **2012**, *29*, 189–198.
- [8] E. Staudacher, *Molecules* **2015**, *20*, 10622–10640.
- [9] C. H. Smit, A. Van Diepen, D. L. Nguyen, M. Wuhrer, K. F. Hoffmann, A. M. Deelder, C. H. Hokke, *Mol. Cell. Proteomics* **2015**, *14*, 1750–1769.
- [10] T. P. Yoshino, X. J. Wu, L. A. Gonzalez, C. H. Hokke, *Exp. Parasitol.* **2013**, *133*, 28–36.
- [11] A. Van Diepen, C. H. Smit, L. Van Egmond, N. B. Kabatereine, A. P. De, D. W. Dunne, C. H. Hokke, *PLoS Negl. Trop. Dis.* **2012**, *6*, 1–10.
- [12] N. S. Prasanphanich, A. E. Luyai, X. Song, J. Heimburg-Molinari, M. Mandalasi, M. Mickum, D. F. Smith, A. K. Nyame, R. D. Cummings, *Glycobiology* **2014**, *24*, 619–637.
- [13] P. H. Jensen, D. Kolarich, N. H. Packer, *FEBS J.* **2010**, *277*, 81–94.
- [14] S. Mulagapati, V. Koppolu, T. S. Raju, *Biochemistry* **2017**, *56*, 1218–1226.
- [15] K. Yamada, S. Hyodo, M. Kinoshita, T. Hayakawa, K. Takehi, *Anal. Chem.* **2010**, *82*, 7436–7443.
- [16] B. Sun, A. V. Pukin, G. M. Visser, H. Zuilhof, *Tetrahedron Lett.* **2006**, *47*, 7371–7374.
- [17] P. Czechura, N. Guedes, S. Kopitzki, N. Vazquez, M. Martin-Lomas, N. C. Reichardt,

- Chem. Commun.* **2011**, *47*, 2390–2392.
- [18] U. Ellervik, G. Magnusson, *Tetrahedron Lett.* **1997**, *38*, 1627–1628.
- [19] D. Benito-Alifonso, R. A. Jones, A. T. Tran, H. Woodward, N. Smith, M. C. Galan, *Beilstein J. Org. Chem.* **2013**, *9*, 1867–1872.
- [20] T. B. Windholz, D. B. R. Johnston, *Tetrahedron Lett.* **1967**, *8*, 2555–2557.
- [21] C. Huang, N. Wang, K. Fujiki, Y. Otsuka, M. Akamatsu, Y. Fujimoto, K. Fukase, *J. Carbohydr. Chem.* **2010**, *29*, 289–298.
- [22] H. Liu, Y. Zhang, R. Wei, G. Andolina, X. Li, *J. Am. Chem. Soc.* **2017**, *139*, 13420–13428.
- [23] E. E. Boeggeman, B. Ramakrishnan, P. K. Qasba, *Protein Expr. Purif.* **2003**, *30*, 219–229.
- [24] Z. S. Kawar, I. Van Die, R. D. Cummings, *J. Biol. Chem.* **2002**, *277*, 34924–34932.
- [25] O. Blixt, N. Razi, *Methods Enzymol.* **2006**, *415*, 137–53.
- [26] K. Naruchi, T. Hamamoto, M. Kurogochi, H. Hinou, H. Shimizu, T. Matsushita, N. Fujitani, H. Kondo, S. I. Nishimura, *J. Org. Chem.* **2006**, *71*, 9609–9621.
- [27] M. A. Oberli, M. L. Hecht, P. Bindschädler, A. Adibekian, T. Adam, P. H. Seeberger, *Chem. Biol.* **2011**, *18*, 580–588.
- [28] S. Yan, S. Serna, N. C. Reichardt, K. Paschinger, I. B. H. Wilson, *J. Biol. Chem.* **2013**, *288*, 21015–21028.
- [29] S. Yan, S. Serna, N.-C. Reichardt, K. Paschinger, I. B. H. Wilson, *J. Biol. Chem.* **2013**, *288*, DOI 10.1074/jbc.M113.479147.
- [30] D. Soriano del Amo, W. Wang, C. Besanceney, T. Zheng, Y. He, B. Gerwe, R. D. Seidel, P. Wu, *Carbohydr. Res.* **2010**, *345*, 1107–1113.
- [31] C. D. Rillahan, E. Schwartz, R. McBride, V. V. Fokin, J. C. Paulson, *Angew. Chem. Int. Ed.* **2012**, *51*, 11014–11018.
- [32] C. D. Rillahan, E. Schwartz, C. Rademacher, R. McBride, J. Rangarajan, V. V. Fokin, J. C. Paulson, *ACS Chem. Biol.* **2013**, *8*, 1417–1422.
- [33] J. Tejler, F. Skogman, H. Leffler, U. J. Nilsson, *Carbohydr. Res.* **2007**, *342*, 1869–1875.
- [34] T. Zheng, H. Jiang, M. Gros, D. Soriano del Amo, S. Sundaram, G. Lauvau, F. Marlow, Y. Liu, P. Stanley, P. Wu, *Angew. Chem. Int. Ed.* **2011**, *50*, 4113–4118.
- [35] K. Brzezicka, U. Vogel, S. Serna, T. Johannssen, B. Lepenies, N. C. Reichardt, *ACS Chem. Biol.* **2016**, *11*, 2347–2356.
- [36] 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.
- [37] B. Ultrastructure, *J. Biol. Chem.* **1975**, *250*, 8518–8523.
- [38] M. J. Swamy, D. Gupta, S. K. Mahanta, A. Surolia, *Carbohydr. Res.* **1991**, *213*, 59–67.
- [39] W. Van Breedam, S. Pöhlmann, H. W. Favoreel, R. J. de Groot, H. J. Nauwynck, *FEMS*

Microbiol. Rev. **2014**, *38*, 598–632.

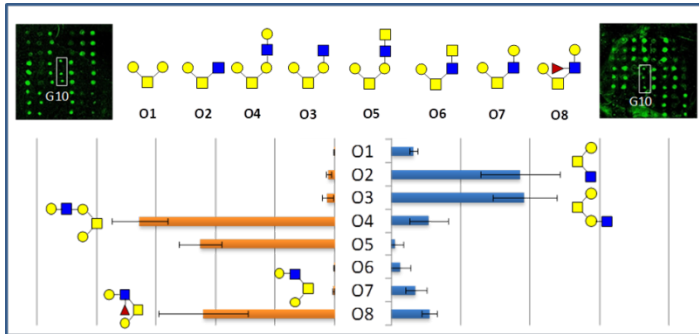
- [40] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell*, Garland Science, New York, **2002**.
- [41] R. Yabe, H. Tateno, J. Hirabayashi, *FEBS J.* **2010**, *3*, 4010–4026.
- [42] Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis, K. Drickamer, *Nat. Struct. Mol. Biol.* **2004**, *11*, 591–598.
- [43] B. Echeverria, F. Fieschi, J. Pham, S. Achilli, C. H. Hokke, N.-C. Reichardt, M. Thépaut, C. Vivès, S. Serna, *ACS Chem. Biol.* **2018**, *13*, 2269–2279.
- [44] P. J. Coombs, R. Harrison, S. Pemberton, A. Quintero-Martinez, S. Parry, S. M. Haslam, A. Dell, M. E. Taylor, K. Drickamer, *J. Mol. Biol.* **2010**, *396*, 685–696.
- [45] M. H. J. Meevissen, N. N. Driessen, H. H. Smits, R. Versteegh, S. J. van Vliet, Y. van Kooyk, G. Schramm, A. M. Deelder, H. Haas, M. Yazdanbakhsh, et al., *Int. J. Parasitol.* **2012**, *42*, 269–277.
- [46] E. Van Liempt, S. J. Van Vliet, A. Engering, J. J. Garc, C. M. C. Bank, M. Sanchez-hernandez, Y. Van Kooyk, I. Van Die, *Mol. Immunol.* **2007**, *44*, 2605–2615.
- [47] C. Unverzagt, S. André, J. Seifert, S. Kojima, C. Fink, G. Srikrishna, H. Freeze, K. Kayser, H.-J. Gabius, *J. Med. Chem.* **2002**, *45*, 478–91.
- [48] A. van Diepen, A. J. van der Plas, R. P. Kozak, L. Royle, D. W. Dunne, C. H. Hokke, *Int. J. Parasitol.* **2015**, DOI 10.1016/j.ijpara.2015.02.008.
- [49] E. Van Liempt, A. Imberty, C. M. C. Bank, S. J. Van Vliet, Y. Van Kooyk, T. B. H. Geijtenbeek, I. Van Die, **2004**, *279*, 33161–33167.
- [50] M. Thépaut, C. Guzzi, I. Sutkeviciute, S. Sattin, R. Ribeiro-Viana, N. Varga, E. Chabrol, J. Rojo, A. Bernardi, J. Angulo, et al., *J. Am. Chem. Soc.* **2013**, *135*, 2518–2529.
- [51] C. Guzzi, F. Doro, J. Reina, M. Th, A. Bernardi, J. Rojo, P. M. Nieto, **2011**, *2*, 7705–7712.
- [52] S. J. van Vliet, E. van Liempt, E. Saeland, C. A. Aarnoudse, B. Appelmelk, T. Irimura, T. B. H. Geijtenbeek, O. Blixt, R. Alvarez, I. van Die, et al., *Int. Immunol.* **2005**, *17*, 661–669.
- [53] B. Lepenies, J. Lee, S. Sonkaria, *Adv. Drug Deliv. Rev.* **2013**, DOI 10.1016/j.addr.2013.05.007.
- [54] L. L. Eggink, K. F. Roby, R. Cote, J. Kenneth Hooper, *J. Immunother. Cancer* **2018**, *6*, 1–16.
- [55] S. A. Jégouzo, A. Quintero-Martínez, X. Ouyang, Á. Dos Santos, M. E. Taylor, K. Drickamer, *Glycobiology* **2013**, *23*, 853–864.
- [56] H. Yu, X. Chen, *Org. Lett.* **2006**, *8*, 2393–2396.

TOC text : A series of parasite O-glycans structures have been prepared by chemoenzymatic synthesis and their affinity towards several C-type lectins screened using glycan arrays

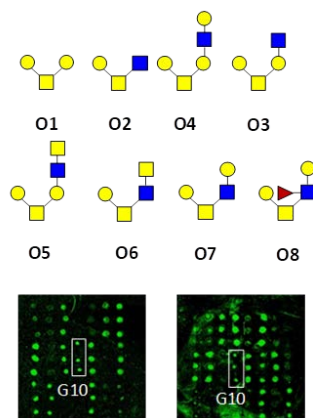
TOC graphic

Please choose most suitable:

Toc 1



Toc 2



Keywords. Carbohydrates, enzymes, glycoconjugates, microarrays, lectins