



HAL
open science

Recombinant C-type Lectin Receptors production and selective ligand identification: new tools towards immune system tailoring

Silvia Achilli

► **To cite this version:**

Silvia Achilli. Recombinant C-type Lectin Receptors production and selective ligand identification: new tools towards immune system tailoring. Biochemistry [q-bio.BM]. Université Grenoble Alpes, 2018. English. NNT: 2018GREAV014 . tel-01980791

HAL Id: tel-01980791

<https://theses.hal.science/tel-01980791>

Submitted on 14 Jan 2019

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.

THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTE UNIVERSITE
GRENOBLE ALPES**

Spécialité : **Biochimie**

Arrêté ministériel : 25 mai 2016

Présentée par

Silvia ACHILLI

Thèse dirigée par **Pr. Franck FIESCHI**
et codirigée par **Dr. Corinne VIVES**

préparée au sein de l'Institut de Biologie Structurale
dans l'École Doctorale Chimie et Sciences du Vivant

Production recombinante de récepteurs lectines de type C et identification de ligands sélectifs: vers de nouveaux outils pour la modulation du système immunitaire.

Thèse soutenue publiquement le **26 juin 2018**,
devant le jury composé de :

Mme Anne IMBERTY

Dr., Centre de Recherches sur les macromolécules végétales
(Présidente)

Mr Francesco PERI

Pr., Università degli studi di Milano-Bicocca (Rapporteur)

Mr Yann GUERARDEL

Dr., French National Centre for Scientific Research Lille, rôle
(Rapporteur)

Mr Bernd LEPENIES

Pr., University of Veterinary Medicine Hannover (Membre)

Mme Corinne VIVES

Dr., Institut de Biologie Structurale, CEA (Co-directrice de thèse)

Mr Franck FIESCHI

Pr., Université Grenoble-Alpes (Directeur de thèse)



Acknowledgements

First of all and foremost, my profound gratitude goes to Pr. Franck Fieschi and Dr. Corinne Vivès-Deniaud. They welcomed me in their team and gave me extremely valuable guidance and support. Also a big thank for translating the abstract of this thesis in French language.

My immense gratitude goes to Dr. Michel Thépaut. He showed me all the biochemistry tips and helped me immensely with the crystallography section.

I would like to express my deepest gratitude to the reviewers, Prof. Francesco Peri and Dr. Yann Guerardel, who kindly accepted to evaluate my dissertation and examine it.

My immense appreciation goes to Dr. Anne Imberty and Pr. Bernd Lepenies for being member of the jury and contribute to the examination of my work.

I own a great acknowledgement to the Biacore platform, particularly to Dr. Isabelle Bally and Jean Baptiste Reiser, and to the PAOL platform of the institute, specially to Pr. Christine Ebel and Aline le Roy for helping and teaching me to perform experiments.

ITC data were acquired at CIBB institute in Grenoble and Dr. Caroline Mass has my sincere gratitude for her help and advice.

A lot of thanks to Dr. Romain Vives and Pr. Olivier Renaudet for being part of the jury for my CST: they evaluated my work and guided me for two years.

I am indebted to all our collaborators: Dr. Niels Reichardt and Dr. Sonia Serna, who hosted me for two secondments and taught me how to perform glycan array screening, Pr. Bernd Lepenies and (not yet Dr.) Joao Monteiro, who hosted me for one secondment and showed me how to perform binding assays with cells, Dr. Ludovic Landemarre and (not yet Dr.) Blanka Didak for showing me their LECprofile assay, Pr. Anna Bernardi and Dr Laura Medve for all the glycomimetics provided and all the valuable help that you gave me.

A big thank to Dr. Jeroen Codee and his PhD student Tim Hogervorst for their work on the multivalent glycoclusters synthesis and Pr. Yvette van Kooyk and her PhD student Eveline Li for the biological assays.

I am deeply grateful for the IMMUNOSHAPE network for the great workshops and the memorable moments spent together.

Finally, I must acknowledge the financial support given by the Marie Sklodowska-Curie grant agreement No 642870 (ETN-Immunoshape), without which my work would not have been possible.

Ringraziamenti

Molte altre persone hanno contribuito, chi più chi meno, a questo mio successo. Questa parte dei ringraziamenti sarà scritta nella mia lingua madre.

Dedico questo a lavoro a:

I miei genitori e mia nonna Ceci i quali, pur non capendo assolutamente niente di quello che faccio, mi hanno guardato con orgoglio durante tutta la mia difesa.

Mio fratello Stefano, che mi considera più intelligente di quanto effettivamente io non sia.

Gloria, la mia più cara amica e fervida sostenitrice.

Erika, che ha viaggiato per più di 7898 chilometri pur di starmi accanto.

Clara, Adrica e Cinzia, the italian team in Grenoble.

Laura, Charles, Romain, Annelise, François, Muge e Sima, miei colleghi compagni di avventure e pause café.

Dulcis in fundo, **Loreto**. Mi ricorderò sempre di quando, alle 2 di notte, mi hai aiutato a rileggere questo manoscritto al contrario, dalla fine all'inizio, per trovare eventuali errori. Nessuna parola potrà mai esprimere tutta la gratitudine che meriti. Sei il mio porto sicuro nei momenti di sconforto.

Index

Introduction	33
1. The Immune System	35
1.1 Introduction	35
1.2 Innate Immune System	36
1.2.1 Dendritic cells and signals for adaptive immunity activation	37
1.2.2 Pattern recognition receptors	38
2. C-type lectin receptors	41
2.1 Introduction	41
2.2 C type lectin domains and glycan recognition: structural aspects	42
2.2.1 CLR signalling	45
2.2.2 CLR and TLR crosstalk	47
2.3 CLRs considered in the study	49
2.3.1 Blood Dendritic Antigen 2 (BDCA2)	49
2.3.2 DC-SIGN and DC-SIGNR (L-SIGN)	50
2.3.3 DC-Associated C-Type Lectin 1 (dectin-1)	52
2.3.4 Dendritic cell associated C type lectin 2 (Dectin-2)	53
2.3.5 Langerin	54
2.3.6 Liver and lymph node sinusoidal endothelial cell C-type lectin (LSECTin)	56
2.3.7 Macrophages C-type lectin (MCL)	57
2.3.8 Macrophage inducible Ca ²⁺ - dependent lectin (Mincle)	59
3. Glycobiology	65
3.1 Introduction	65
3.2 Design of Mimetic	68
3.2.1 Glycomimetic	69

3.2.2 Non-glycomimetics	70
3.3 Screening technique	71
3.3.1 Lectin microarray	71
3.3.1 Glycan/glycomimetic microarray	72
3.3.2.1 TETRALEC strategy	73
3.4 Multivalent Ligands: Targeting CLR	74
4. Applications	78
4.1 Diagnosis	78
4.2 Imaging	80
4.3 Targeting	82
5. AIM OF THE THESIS	85
Materials & Methods	91
6. Principles	93
6.1 Methods for protein characterization	93
6.1.1 Circular dichroism (CD)	93
6.1.2 Size Exclusion Chromatography Multi Angle Laser Light Scattering SEC-MALLS (PAOL platform)	94
6.2 Methods for characterization of protein-ligand interaction	95
6.2.1 Lectin array (LectPROFILE)	95
6.2.2 Glycan array	96
6.2.3 Surface Plasmon Resonance (SPR)	96
6.2.4 Isothermal Titration Calorimetry (ITC)	101
7. Methods	103
7.1 Production of recombinant C-type lectin constructs	103
7.1.1 Transformation of Ca ₂₊ -competent <i>E.coli</i> strains	103

7.1.2 Cloning of LSEctin His-CRD	103
7.1.3 Preparation of bacterial preculture	105
7.1.4 Preparation of bacterial culture samples for SDS-PAGE	105
7.1.5 DC-SIGN and Langerin over-expression and purification	105
7.1.6 Protein labelling	105
7.1.7 Over-expression and inclusion body preparation of all CLR constructs	105
7.1.7.1 Refolding, purification and labelling of DC-SIGNR ECD	106
7.1.7.2 Refolding, purification and labelling of Dectin-2 Strep-ECD	107
7.1.7.3 Refolding, purification and labelling of Mincle His-ECD	107
7.1.7.4 Refolding and purification of His-GGG-CRD constructs	108
7.2 Biochemical and biophysical protein characterization and ligand analysis	111
7.2.1 Protein samples for SDS-PAGE.	111
7.2.2 Protein concentration determination	111
7.2.3 Sample preparation for circular dichroism	112
7.2.4 Sample preparation for SEC-MALS (PAOL platform)	113
7.2.5 Sample preparation for lectin array analysis	113
7.2.6 Sample preparation for glycan array analysis	113
7.2.7 Sample preparation for flow cytometry	114
7.2.8 Sample preparation for ITC	114
7.2.9 SPR surface preparation for inhibition test	114
7.2.10 SPR surface preparation for direct interaction test	115
Results and Discussion	119
8 Recombinant CLR Production and Functional Test	121
8.1 Cloning	122
8.1.1 LSEctin His-CRD Cloning	125

8.2 Strategies for Recombinant Protein Expression	126
8.2.1 Expression as soluble folded CLR	126
8.2.2 Expression as Inclusion Bodies	131
8.2.2.1 ECD constructs	131
8.2.2.1.1 DC-SIGNR ECD over-expression and purification results	132
8.2.2.1.2 Dectin-2 Strep-ECD over-expression and purification results	133
8.2.2.1.3 Mincle His-ECD over-expression and purification results	135
8.2.2.2 CRD constructs	137
8.2.2.2.1 DC-SIGNR His-CRD	138
8.2.2.2.2 BDCA2 and LSEctin His-CRD	140
8.2.2.2.3 Dectin-1, Dectin-2, Mincle, MCL His-CRD	143
8.3 TETRALEC, Artificial Tetrameric Lectins: a Tool to Screen Ligand and Pathogen Interactions	151
9 Screening: identification of selective ligands towards human CLRs	177
9.1 Glycan array	177
9.1.1 Chemoenzymatic Synthesis of N-glycan Positional Isomers and Evidence for Branch Selective Binding by Monoclonal Antibodies and Human C-type Lectins Receptors	179
9.1.2 Other glycan array screening	195
9.2 Glycomimetic array	197
9.2.1 On-chip screening of a glycomimetic library with C-type lectins reveals structural features responsible for preferential binding of dectin-2 over DC-SIGN/R and langerin	198
9.2.2 Other glycomimetic array screening	215
10 Characterization of new glycomimetics specific to DC-SIGN	217
10.1 Robustness enhancement	219
10.1.1 Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist.	219
10.1.2 Additional data non-presented in the article: enzymatic assay	227
10.1.3 C and Si-Glycosides	228

10.2 Development of glycomimetics selective towards DC-SIGN: the “ammonium binding pocket” strategy	230
10.2.1 Ammonium binding pocket: amino derivatives	232
10.2.2 Ammonium binding pocket: triazole derivatives	234
11 Design of multivalent mannosylated ligands for CLR targeting	247
11.1 Carbohydrate synthesis (Leiden)	248
11.2 ELISA assays (Amsterdam)	249
11.3 SPR Experiments (Grenoble)	251
11.4 FACS Assays (Amsterdam)	255
Conclusions and Perspectives	259
12.1 Remarks on the production of non-commercially available lectins	261
12.2 Exploiting glycan/glycomimetic array screenings	265
12.3 Enhancement and development of new glycomimetics specific to DC-SIGN	267
12.4 Study of mannose clusters targeting CLRs as a future tool to trigger anticancer immune response	271
12.5 Final and general conclusions	273
Bibliography	277
Annexes	291
Annexes Chapter 8	293
Supporting Information Paper 1	325
Annexes Chapter 9	331
Supporting Information Paper 2	333
Supporting Information Paper 3	355
Annexes Chapter 10	397

Supporting Information Paper 4

407

Annexes Chapter 11

413

List of Figures:

Fig.1 The immune system	35
Fig.2 Innate and Adaptive Immunity s	36
Fig.3 Activation of the adaptive immunity	38
Fig.4 PAMPs and DAMPs	39
Fig.5 The four different families of PRRs	39
Fig.6 CLRs and immunity	40
Fig.7 Examples of four animal lectin families	41
Fig.8 Schematic representation of a tetrameric CLR	42
Fig.9 Carbohydrate recognition domain	43
Fig.10 EPN and QPD motifs	43
Fig.11 C-type lectin subfamilies	44
Fig.12 Overview of the four motifs for signalisation	45
Fig.13 Dectin-2 and DCIR signalling	46
Fig.14 Cross talk between CLR and TLR	47
Fig.15 DC-SIGN and TLR4 during Mycobacterium infection	48
Fig.16 Portion of BDCA2 CRD structure	49
Fig.17 Portion of DC-SIGN and DC-SIGNR structure	50
Fig.18 Comparison of the CRD from DC-SIGN and DC-SIGNR,	51
Fig.19 The placenta, DC-SIGN. DC-SIGNR and HIV	52
Fig.20 Murine dectin-1 structure	53
Fig.21 Portion of Dectin-2 structure	54
Fig.22 Langerin structure	55
Fig.23 Birbeck granules	55
Fig.24 Genomic organization of DC-SIGNR, DC-SIGN and LSEctin	56
Fig.25 LSEctin	57
Fig.26 MCL and DC-SIGNR	58
Fig.27 Mincle expression	58
Fig.28 Portion of mincle structure	59
Fig.29 Mincle	60
Fig.30 Graphical overview of the nine CLRs investigated during my PhD	60
Fig.31 Potential information on content of DNA, RNA, protein and glycan	65
Fig.32 N-glycans	66
Fig.33 Linkage points for oligomer formatio	67
Fig.34 Example of calcium ion coordination	67
Fig.35 Cracking the glycode	68
Fig.36 Schematic glycomimetic design	69
Fig.37 Glycomimetics against DC-SIGN developed by Anna Bernardi research group	70

Fig.38 Glycomimetic against langerin	70
Fig.39 Antagonists developed in Laura Kiessling group	71
Fig.40 Glycan array	73
Fig.41 Multivalency	74
Fig.42 Binding mode	75
Fig.43 Multivalent ligands	76
Fig.44 Examples of multivalent inhibitors against DC-SIGN	77
Fig.45 Multivalent ligand as vaccine	77
Fig.46 Exploitation of lectin-carbohydrate interaction in medical biology	78
Fig.47 Application of glycan array.	80
Fig.48 Example of “Smart” probe	81
Fig.49 Example of imaging	81
Fig.50 Targeting	82
Fig.51 TACAs	83
Fig.52 Vaccination using CLR targeting	84
Fig.53 The principal architecture of IMMUNOSHAPE ITN consortium and group involvement.	86
Fig.54 IMMUNOSHAPE network	88
Fig.55 Circular dichroism	93
Fig.56 SEC-MALLS	94
Fig.57 LectPROFILE schematic representation.	95
Fig.58 Glycan array glass slide	96
Fig.59 SPR principle	97
Fig.60 Plasmon resonance effect	97
Fig.61 Analyte interaction and bulk effect	98
Fig.62 Design of competition/inhibition SPR assay	99
Fig.63 SPR competition/inhibition sensorgram and curve	100
Fig.64 Oriented surface approach	101
Fig.65 Calorimetry principle	102
Fig.66 Bacterial cell wall section and location of the periplasm	121
Fig.67 Protein secretion pathway and signal peptides	122
Fig.68 Cloning strategies.	123
Fig.69 LSEctin His-CRD cloning	125
Fig.70 pTUM4 helper plasmid	126
Fig.71 Dectin-2 ompA Strep-ECD expression	127
Fig.72 Dectin-2 Strep-ECD expression	128
Fig.73 Dectin-2 pelB His-ECD expression	128
Fig.74 Dectin-2 pelB His-ECD Western Blot.	129
Fig.75 Dectin-2 pelB His-ECD	129

Fig.76 DC-SIGNR ECD expression	132
Fig.77 DC-SIGNR ECD purification	133
Fig.78 Dectin-2 Strep-ECD expression	133
Fig.79 Dectin-2 Strep-ECD purification	134
Fig.80 Dectin-2 Strep-ECD LECprofile analysis	135
Fig.81 Expression of mincle His-ECD	135
Fig.82 Mincle His-ECD purification	136
Fig.83 GAS recognition by Mincle His-ECD.	137
Fig.84 DC-SIGNR His-CRD expression and purification	138
Fig.85 DC-SIGNR His-CRD His tag cleavage and purification	139
Fig.86 DC-SIGNR CRD biotinylation	139
Fig.87 BDCA2 and LSEctin His-CRD expression	140
Fig.88 BDCA2 biotinylation	141
Fig.89 BDCA2 His-CRD LECtprofile analysis	141
Fig.90 LSEctin His-CRD purification	142
Fig.91 LSEctin His-CRD LectPROFILE analysis	143
Fig.92 Dectin-1, Dectin-2, Mincle and MCL HIS-CRD expressions	143
Fig.93 Dectin-1 purification	144
Fig.94 GAS recognition by dectin-1 His-CRD	144
Fig.95 Mincle His-CRD purification	145
Fig.96 Dectin-2 His-CRD purification	146
Fig.97 Comparison between Dectin-2 His-CRD and Dectin-2 Strep-ECD LectPROFILE analysis	146
Fig.98 Dectin-2 His-CRD CD spectrum.	147
Fig.99 MCL His-CRD purification	147
Fig.100 MCL His-CRD CD spectrum.	148
Fig.101 MCL His-CRD His tag cleavage and purification	148
Fig.102 MCL CRD-biotinylation	149
Fig.103 Example of interaction detection by Agila Scanner	177
Fig.104 Glycan array	178
Fig.105 Glycan microarray incubation	196
Fig.106 Glycomimetic array	197
Fig.107 Other glycomimetic arrays	215
Fig.108 First and second generations of glycomimetics against DC-SIGN	217
Fig.109 6NH ₂ -Man030 structure	218
Fig.110 psDi/DC-SIGN CRD X-ray structure	218
Fig.111 psDi and thio-psDi structures	219
Fig.112 thio-psdi enzymatic reaction	227
Fig.113 List of C-glycosides and Si-glycoside structures	228

Fig.114 ID-246-4, ID-246-5, ID-246-11, ID-246-12 and JCH-423 competition assay against DC-SIGN	229
Fig.115 Best candidates from computational screening	230
Fig.116 The “ammonium binding pocket”	230
Fig.117 Synthetic strategies	231
Fig.118 Synthetic strategy and overall structure of amino derivatives	232
Fig.119 Amino derivatives structures.	232
Fig.120 IC ₅₀ comparison of the amino derivatives	233
Fig.121 Triazole derivatives strategy	234
Fig.122 Man062 competition assay	234
Fig.123 Man065 and Man064 structures	235
Fig.124 Man062, Man065, Man064 and psDI IC ₅₀ comparison.	235
Fig.125 Man066-67-68 structures.	236
Fig.126 Man062 corresponding methylated glycomimetics inhibition assay	236
Fig.127 List of triazole derivatives IC ₅₀ values and corresponding structures.	237
Fig.128 Man069 structure.	238
Fig.129 Man062, Man065 and Man089 competition assay	238
Fig.130 Examples of sensorgrams	239
Fig.131 Titrations of glycomimetic Man069 to DC-SIGN	240
Fig.132 Man069 direct interaction assay	241
Fig.133 One example of obtained crystal and diffraction pattern.	242
Fig.134 Crystal packing and evidence for ligand presence.	242
Fig.135 Man069/DC-SIGN solved X-ray structure	244
Fig.136 Man069 canonical binding	244
Fig.137 Man069 non-canonical binding site	245
Fig.138 Cluster synthesis	248
Fig.139 20 mannosylated compounds	249
Fig.140 ELISA results	250
Fig.141 Inhibition curves	251
Fig.142 Sensorgrams obtained from B6 interaction	253
Fig.143 K _{dapp} values	254
Fig.144 Comparison of the K _{dapp} values	255
Fig.145 Interaction between biotinylated glycoclusters	255
Fig.146 BDCA2, dectin-1, MCL and dectin-2 His-ECD expressions	262
Fig.147 X-ray structure of dectin-2	262
Fig.148 The six asymmetric glycans and the two non-branched corresponding glycans	265
Fig.149 Structure of mannose based glycomimetics that specifically recognise dectin-2	266
Fig.150 Structures and IC ₅₀ comparison	267
Fig.151 Man062 (green) and Man069 (cyan) competition assays	268

List of Tables:

Table.1 Overview of the nine CLRs studied during my PhD	64
Table.2 List of lectins used in cancer biomarker research	79
Table.3 Digest assays of pUC57 LSEctin vector and pET30b vector.	103
Table.4 Preparation of NdeI/HindIII digestion mixtures for the test of positive DNA ligation.	104
Table.5 Refolding concentration, buffer for refolding and dialysis and protein yield/L culture for the His-CRD considered.	109
Table.6 Required volumes for His tag cleavage of BDCA2, DC-SIGNR and MCL CRDs	110
Table.7 Required volumes for sortase reaction on BDCA2, DC-SIGNR and MCL CRDs. mg of TETRALEC complex are also given.	110
Table 8 and 9. Molecular weight, isoelectric points, molar extinction coefficient ϵ and Abs 0.1% values for the considered ECDs and CRDs.	112
Table 10. List of analysed CLRs at the SEC-MALS and their specifications.	113
Table.11 Summary of all the attempts that have been made for soluble expression	130
Table.12 ECD expression and purification yield	149
Table.13 ECD expression, purification yield, functionality assay, CD assay and TETRALEC formation.	150
Table.14 Parameters of the crystal structure.	243
Table.15 Summary of the competition experiments	252

Abbreviations:

AAL : Aleuria Aurantia Lectin	GalT : Galactosyltransferase
APCs : Antigen Presenting Cells	GAS : Group A Streptococcus
ASGPR : Asialoglycoprotein Receptor	GBPs : Glycan Binding Proteins
BDCA2 : Blood Dendritic Antigen 2	GI : Gastrointestinal
BSA-Man : Bovine Serum Albumin Mannosylated	GL : Galectin
CD : Circular Dichroism	GlucNAc : N-acetylglucosamine
cDCs : conventional DCs	gp120 : GlycoProtein 120
CFG : Consortium for Functional Glycomics	GPI : Glycosylphosphatidylinositol
CL : C-type Lectin	hemiITAM : hemi-Immunoreceptor Tyrosine-based Activation Motif
CLRs : C-type Lectin Receptors	HIV : Human Immunodeficiency Virus
CM : Carboxymethyl	HPLC : High Precision Liquid Chromatography
ConA : Concanavalin A	HRP : Horseradish Peroxidase
CRAC : Cholesterol Recognition Amino acid Consensus-like	IFN : type I Interferon
CRD : Carbohydrate Recognition Domain	IgG : Immunoglobulin G
CTLD : C-Type Lectin-like Domain	IL : Interleuchin
CTLs : Cytotoxic T cells	IL : I-type Lectin
DAMPs : Damage-Associated Molecular Patterns	IPTG : Isopropyl 1-thio-D-galactopyranoside
DAP12 : 12-kDa DNAX-Activating Protein	IRAK : Interleukin Receptor-Associated Kinase
DCIR : Dendritic Cell ImmunoReceptor	ITC : Isothermal Titration Calorimetry
DCs : Dendritic Cells	ITIM : Immunoreceptor Tyrosine-based Inhibition Motif
DC-SIGN R : Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin Related	LB : Luria Bertani
DC-SIGN : Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin	LCs : Langerhans Cells
Dectin1/2 : DC-Associated C-Type Lectin 1/2	LDN : GalNAc β 1,4GlcNAc
DLS : Dinamic Light Scattering	LDNF : GalNAc β 1,4[Fuca α 1,3]GlcNAc
DNA : Deoxyribonucleic Acid	LN : N-acetyl lactosamine
DOL : Degree Of Labeling	LPS : Lipopolysaccharide
dsbA : disulphide bond A	LSEctin : Liver and lymph node Sinusoidal Endothelial C-Type lectin
DSF : Differential Scanning Fluorimetry	L-SIGN : Liver/Lymph node-Specific Intercellular adhesion molecule-3-Grabbing Nonintegrin
ECD : Extracellular Domain	mAbs : monoclonal Antibodies
EDC : N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide	MALDI-ToF : Matrix Assisted Laser Desorption Ionisation - Time of Flight
EDTA : Ethylenediaminetetraacetic acid	MALS : Multi Angle Light Scattering
ELISA : Enzyme-Linked ImmunoSorbent Assay	manLAM : Mannosylated Lipoarabinomannan
ER : Endoplasmic Reticulum	MAPK : Mitogen-Activated Protein Kinase
ESI-MS : Electrospray Ionization Mass Spectrometry	MBP : Mannose-Binding Protein
FACS : Fluorescence-Activated Cell Sorting	MCL : Macrophage C-Type Lectin
FCRγ :Fc Receptor γ -chain	MHC : Major Histocompatibility Complex
FucT : Fucosyltransferase	Mincle : Macrophage Inducible Ca ²⁺ -dependent lectin
GalNAc :N-Acetylgalactosamine	MMR : Macrophage-Mannose Receptor
GalNAcT : GalNAc-transferase	moDCs : monocyte-derived DCs

MP : Mannose Posphate
MP3 : Multi step Protein Purification Platform
MW : Molecular Weight
NADPH : Nicotinamide Adenine Dinucleotide Phosphate
Nano ITC LV : Nano Isothermal Titration Calorimetry Low Volume
NF-κB : Nuclear Factor – κB
NHS : N-hydroxysuccinimide
NK : Natural Killer
NLRs : Nucleotide-binding oligomerization domain Like Receptors
NMR : Nuclear Magnetic Resonance
ompA : outer membrane protein A
OVA : Ovoalbumin
PADS : Protecting group Aided Detection and Separation of glycans
PAMPs : Pathogen-Associated Molecular Patterns
PAOL : Protein Analysis On Line
PCR : Polymerase Chain Reaction
pDCs : plasmacytoid DCs
PEG : Polyethylene Glycol
peIB : pectate lyase B
PMN : Polymorphonuclear
PRRs : Pattern Recognition Receptors
RA : Rheumatoid Arthritis
RFU : Relative Fluorescence Units
RFI: Relative Fluorescence Intensity
RLRs : Retinoic acid-inducible gene-I-Like Receptors
RNA : Ribonucleic Acid
ROS : Reactive Oxygen Species
RU : Resonance Unit
SARS : Severe Acute Respiratory Sindrome
SAXS : Small Angle X-ray Scattering
SEC : SecB Dependent
SEC : Size-Exclusion Chromatography
SHP-1/2 : SH2-containing tyrosine Phosphatase-1/2
SLS : Static Light Scattering
SPAAC : Strain-Promoted Azide-Alkyne Cycloaddition
SPR : Surface Plasmon Resonance
SRP : Signal Recognition Particle
SrtA : Sortase A
Syk : spleen tyrosine kinase
TACAs : Tumour-Associated Carbohydrate Antigens
TAT : twin-arginine translocation
TCRs : T Cell Receptors
TDM : trehalose-6,6'-dimycolate
Th : T helper
TLC : Thin Layer Chromatography
TLRs : Toll-Like Receptors
TRAF : TNF Receptor Associated Factor
UPLC-MS : Ultra performance liquid chromatography - tandem mass spectrometer
WC : Whole Cells

Les résumés en Français

Chapitre 1. Le système immunitaire

Le contexte général de ce travail de thèse est le système immunitaire, source vitale de défense contre les agents pathogènes. L'immunité est divisée en deux systèmes communicants, l'immunité innée et l'immunité adaptative. Le paragraphe **1.2** décrit brièvement les acteurs cellulaires les plus importants de l'immunité innée. Ces cellules comprennent les macrophages, les neutrophiles, les basophiles, les éosinophiles, les mastocytes, les cellules tueuses (natural killer cells), et les cellules dendritiques (CD). Toutes ces cellules se développent sous la stimulation de certaines cytokines.

Les CD sont spécialement importantes car elles constituent un lien entre l'immunité innée et adaptative et leur rôle principal est d'induire l'immunité adaptative (paragraphe **1.2.1**). L'identification de micro-organismes étrangers par l'immunité innée est basée sur la reconnaissance de structures moléculaires conservés dans les micro-organismes et absentes chez l'hôte (paragraphe **1.2.2**). Ces structures sont appelées "motifs moléculaires associés à des pathogènes" (PAMPs en anglais). Les récepteurs du système immunitaire inné qui reconnaissent les PAMPs sont appelés Pattern Recognition Receptors (PRR). Les PRRs peuvent être solubles, exprimés à la surface cellulaire ou intra-cellulaires. Les récepteurs lectine de type-C (CLRs) sont l'une des familles de PRRs.

Chapitre 2. Les récepteurs lectine de type-C

Les récepteurs lectine de type-C sont des lectines qui reconnaissent des groupements glucidiques spécifiques présents à la surface de leurs via un domaine structural appelé domaine de reconnaissance de carbohydrates (CRD). Cette reconnaissance est dépendante d'ions Ca^{2+} présents dans le site actif et ceci est expliqué dans le paragraphe **2.2**. Après contact avec le ligand, quatre voies peuvent être exploitées pour moduler le système immunitaire. La signalisation résultante dépend de plusieurs aspects, tel que la typologie du récepteur, son internalisation et la nature des ligands (paragraphe **2.2.1**). Néanmoins, certaines constantes sont conservées parmi les différentes voies de signalisation : le motif basé sur la tyrosine, l'utilisation de SYK et SHP pour la modulation de la transcription et la coopération avec autre récepteur comme le récepteurs de type toll.

Enfin, la dernière partie du chapitre (paragraphe **2.3**) décrit les neuf différentes CLRs étudiées pendant ce travail de thèse : BDCA2, DC-SIGN, DC-SIGNR, dectin-2, dectin-1, langerin, LSECtin, MCL et mincle. Pour chacun d'eux des informations sur la structure, la spécificité de reconnaissance et la voie de signalisation ont été donnés.

Chapitre 3. Glycobiologie

Les lectines décrites dans le chapitre précédent peuvent déchiffrer le glycode de glycanes, porteur d'informations biologiques. Les glycanes sont essentiels dans l'interaction cellule-cellule et la position de leurs groupements hydroxyle est cruciale pour la bio-reconnaissance.

Les sucres sont extrêmement complexes et une raison de cette complexité découle de la grande variété de liaisons possibles pour la formation d'oligomères. Outre leur complexité, les carbohydrates ont faible affinité et spécificité pour leurs partenaires protéiques. Afin de surmonter ces difficultés, des mimétiques de carbohydrates doivent être développés. Trois différentes sections du glycan peuvent être optimisées. Ceci est développé dans des parties du paragraphe 3.2. Une fois que les glycomimétiques ont été synthétisés, leur interaction avec de lectines doit être validé par des expériences biochimique *in vitro* et, notamment, par des techniques de criblage (puce a sucres et puce a lectines, paragraphe 3.3).

L'optimisation et l'étude de ligand monovalent est la première étape vers le développement de inhibiteurs. Pour améliorer la faible affinité d'interaction, des ligands multivalentes sont ensuite envisagés afin d'atteindre l'effet d'avidité. Dans le paragraphe 3.4 différents exemples the ligands multivalents sont donnés.

Chapitre 4. Applications

Ce chapitre surligne la pertinence médicale de l'étude des interactions lectines-carbohydrates. Notamment, un focus a été fait sur le diagnostic (plusieurs maladies étant caractérisées par de changement de motif de glycosylation, sur l'imagerie pour pouvoir observer l'efficacité de l'internalisation au niveau cellulaire et sur le ciblage de lectines, fondamental pour la vaccination et le traitement du cancer.

Chapitre 5. Les objectifs de cette thèse

Le projet s'inscrit dans un contexte international et fait partie du réseau européen IMMUNOSHAPE. Le but du réseau est de combiner l'état de l'art de la synthèse chimique et des technologies de criblage pour développer des molécules immunothérapeutiques multivalentes basée sur des glycanes. Ma contribution au projet a été divisée en trois axes :

- La production de neuf récepteurs lectine de type-C. Différentes approches ont été testées pour optimiser l'expression et la production en bactérie. De plus, pour améliorer artificiellement la multivalence des protéines et l'affinité pour leurs ligands, une nouvelle stratégie visant à

multimeriser le construct CRD a été développée. Le complexe multimeric final comprenant quatre CRD biotinilés a été appelle TETRALEC.

- L'utilisation de méthodes de criblage d'interaction avec les ligands. Les techniques exploitées ont été : LectPROFILE assay, analyse par FACS, puce a sucre et a lectine.
- La caractérisation de l'interaction entre les ligands et les lectines par des études biophysiques. Cet axe est divisé en deux parties. La première est basées sur l'étude d'interaction avec des ligands monovalents et sélectifs pour DC-SIGN, tandis que la deuxième partie étudiée l'interaction avec des ligands multivalents.

Chapitre 6. Principes

Ce chapitre décrit la théorie des techniques utilisées dans ce travail de recherche.

Chapitre 7. Méthodes

Ce chapitre décrit en détail toutes les procédures ainsi que tous les produits chimiques, biologiques et les équipements utilisés dans ce travail de recherche.

Chapitre 8. Préparation de lectines recombinantes et test fonctionnel

Ce chapitre contient les résultats de la production des différentes lectines. En première intention, deux types de constructions ont été réalisées : soit le domaine extracellulaire complet (ECD) soit seulement le domaine de reconnaissance des sucres (CRD). Plusieurs constructions ciblant le periplasm de E.coli afin d'obtenir la protéine directement soluble ont été essayées. Malgré la large gamme de constructions testées, cette approche a conduit soit un faible niveau d'expression soit une expression insoluble. Pour cette raison, la stratégie d'expression periplasmique a été abandonnée. Les résultats d'expression et purification sont présentés au paragraphe **8.2.1**.

Une stratégie alternative est présentée au paragraphe **8.2.2** et consiste en l'expression de protéines sans forme insoluble dans le cytoplasme bactérien, suivi par leur repliement *in vitro*. Trois constructions ECD ont été produites (chapitre **8.2.2.1**). La production de DC-SIGNR-ECD a été réalisée sans obstacle majeur, et de rendements considérables de protéines pures et fonctionnelles ont été obtenus. Les rendements de dectin-

2 Strep-ECD et mincle His-ECD étaient inférieurs. Néanmoins, la fonctionnalité a été confirmée et la spécificité de reconnaissance de dectin-2 a été étudiée plus en détail. La production des constructions CRD a été plus fructueuse (chapitre **8.2.2.2**). DC-SIGNR, BDCA2 et LSEctin His-CRDs ont été produits comme protéines fonctionnelles. MCL His-CRD a été produit et son intégrité structurale a été confirmée mais l'absence de ligands connus a empêché l'évaluation de sa fonctionnalité. Dectin-1 et Dectin-2 His CRD ont été produits mais sont non-fonctionnelles.

Enfin, la stratégie TETRALEC a été exploitée avec succès pour DC-SIGNR et MCL His-CRDs, tandis que des complexes TETRALEC avec biotinylation aléatoire ont été produits pour LSEctin et BDCA2 His-CRDs.

DC-SIGNR TETRALEC a été étudié plus en détail et les résultats sont présentés dans le papier numéro **1** (paragraphe **8.3**). Le complexe a été caractérisé structuralement par SEC-MALS et la validation de sa fonctionnalité a été effectuée sur une puce à sucres dans le laboratoire de Dr. Niels Reichardt en Donostia-San Sebastian (Espagne) et *in cellulo* en utilisant la cytométrie en flux et montrant une interaction avec *Candida albicans* dans le laboratoire de Pr. Bernd Lepenies en Hanover (Allemagne).

Chapitre 9. Criblage : identification de composés sélectifs des CLR humains

Ce chapitre décrit l'utilisation de puces à sucre pour cribler l'interaction entre des panels de glycanes (synthétisés par l'équipe de Dr. Niels Reichardt) et des glycomimétiques (synthétisés par l'équipe de Pr. Anna Bernardi) avec nos lectines marquées par un fluorophore. Dans le paragraphe **9.1.1**, le papier numéro **2** décrit une reconnaissance différentielle par trois lectines (DC-SIGN ECD, DC-SIGNR ECD et LSEctin CRD) de paires de glycans qui sont des isomères de position.

La deuxième partie du chapitre décrit le criblage de glycomimétiques et les résultats sont présentés dans le papier numéro **3** (paragraphe **9.2.1**). L'interaction obtenue par criblage sur puce a été confirmée aussi par SPR pour DC-SIGN et DC-SIGNR ECDs et de nouveaux glycomimétiques reconnus par dectin-2 ECD ont été identifiés.

Chapitre 10. Caractérisation de nouveaux glycomimétiques spécifiques vers DC-SIGN

Ce chapitre décrit la caractérisation par compétition en SPR de composés glycomimétiques spécifiques de DC-SIGN et est divisée en deux sous-parties principales. L'une décrit la stratégie utilisée pour augmenter la stabilité des glycomimétiques synthétisés et la seconde la recherche de composés encore plus spécifiques de DC-SIGN.

La première partie a été menée via deux collaborations différentes. Le papier numéro **4** (paragraphe **10.1.1**) résulte d'une collaboration avec l'équipe de Pr. Anna Bernardi (Milano, Italie) pour caractériser l'effet du changement de liaisons glycosidiques par des liaisons avec un soufre. La deuxième collaboration a été

meée avec l'équipe du Pr. Jitka Moravcová pour le développement de C-glycosides et les résultats sont présentés au paragraphe **10.1.3**.

La deuxième partie du chapitre (paragraphe **10.2**) regroupe les travaux menés par l'équipe du Pr. Anna Bernardi sur le concept d'une « poche de liaison de ammonium » par des études de modélisations moléculaires. Le criblage par SPR des multiples glycomimétiques développés en ce sens nous ont permis d'identifier le composé Man069 qui présente une forte affinité et spécificité pour DC-SIGN. Les études biophysiques et structurales que nous avons réalisés sur le complexe DC-IGN/Man069 concluent le chapitre.

Chapitre 11. Conception de composés multivalents mannosylés pour le ciblage de CLR

Ce dernier chapitre de résultats est basé sur une collaboration avec deux autres équipes du réseau IMMUNOSHAPÉ : l'équipe de Prof. Yvette van Kooyk (Amsterdam, Pays-Bas) et l'équipe de Prof. Jeroen Codee (Leiden, Pays-Bas). Le projet à la base de cette collaboration était le développement d'une molécule multivalente hautement définie pour la vaccination contre le cancer, un des objectifs finals du réseau IMMUNOSHAPÉ. Notre contribution a consisté en la caractérisation par interaction directe en SPR de différents glycoclusters. Les résultats de cette évaluation sont présentés au paragraphe **11.3**.

Chapitre 12. Conclusions et Perspectives

Tous les résultats sont résumés dans ce chapitre.

- Des remarques sur la production des lectines ont été faites dans le chapitre **9.1**. Les moyens pour améliorer leur production sont aussi suggérés.
- Les résultats de criblage sont résumés et d'autres expériences sont aussi suggérées pour leur valider.
- Les résultats d'interaction pour l'ensemble des glycomimétiques monovalents testés contre DC-SIGN sont résumés et comparés dans le chapitre **9.3**. Un composé optimal a été identifié et son interaction avec DC-SIGN a été de plus étudiée structuralement.
- Les résultats d'interaction pour l'ensemble des composés multivalents testés sont résumés et comparés dans le chapitre **9.4**.

Enfin, dans le chapitre **9.5** des conclusions plus générales sont données, notamment par rapport à l'identification de candidats potentiels permettant d'adresser l'objectif initial à long terme du réseau IMMUNOSHAPÉ : la modulation du système immunitaire.

Part I.
Introduction

1. The Immune System

1.1 Introduction

The immune system is our vital source of defense against infection and damage from external organism and toxin [1]. « Defense » and « immunity » appears as two crucial terms: while « defense » is used in a context of attack, implying defeat or victory, « immunity » is employed as synonym of resistance [2]. Immunity is a very complex network of soluble and cellular factors, with pro- and anti-inflammatory partners, and it is composed, in our body, by two communicating subgroups: the innate immune system and the adaptive immune system (Fig.1) [3].

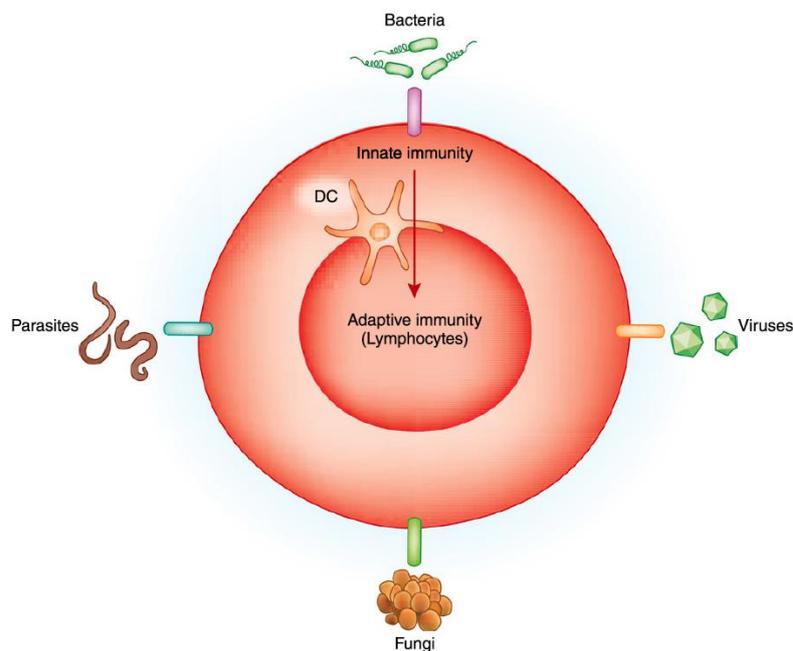


Fig.1 The immune system. By sensing microbes (bacteria, parasites, fungi, and viruses), the innate immune system awakes the adaptive one. Dendritic cells (DCs) link the two components of immunity. DCs are involved in antigen phagocytosis, processing and presenting to lymphocytes (T Cells). Adapted from [2].

The **innate immune system** is passed down from generation to generation [3] and it is characterized by the ability to sort out the self-molecular patterns from non-self or altered elements. The **adaptive immune system**, on the other hand, can develop memory and can adjust itself in response to pathogens. These two defensive systems can communicate through cells that are devoted to antigen processing, called Antigen-Presenting Cells (APCs).

The strong interaction between the innate and adaptive systems is allowed by the production of cytokines and chemokines [4] but this aspect of immune regulation will not be detailed within this manuscript.

1.2 Innate Immune System

Innate immunity, discovered by Elie Metchnikoff in 1916 during phagocytosis studies [5], describes the defense processes by which pathogens are recognized in an immediate yet nonspecific manner. Its typical timescale ranges from seconds to hours after antigen invasion [5]. The first physical defensive barrier is composed by cells of the epithelium, which is impenetrable by most external agents. Cells actively involved in immunity are divided into **lymphocytes** and **phagocytes**. Lymphocytes, which are the actors of the adaptive system, are T cells, B cells and Natural Killer (NK), while phagocytic [6] cells involved in the fast innate response include neutrophils and other types of granulocytes, macrophages, dendritic cells, and mast cells (Fig.2).

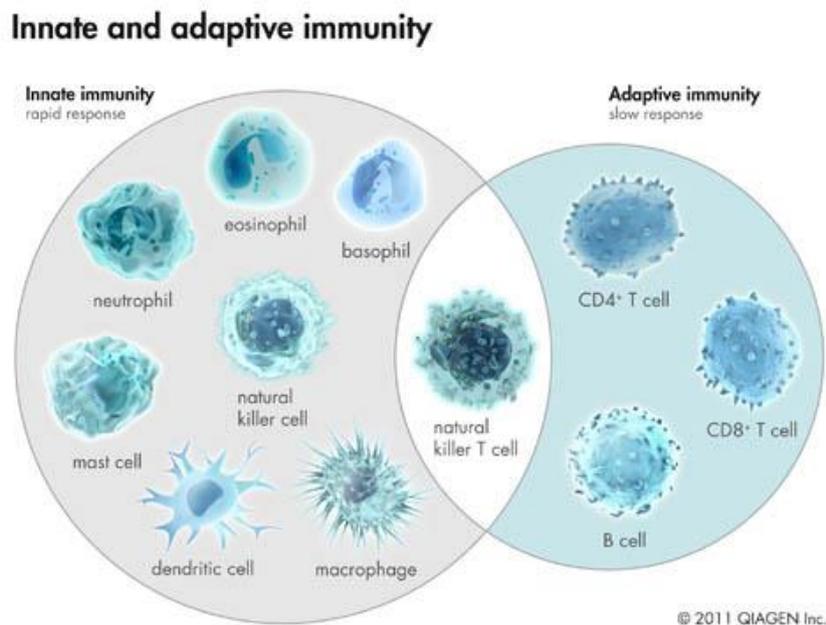


Fig.2 Innate and adaptive immunity. The players (and inter-players) of the innate and adaptive immune systems.

The granulocytes family is composed by neutrophils, eosinophils and basophils, they are all involved in phagocytosis and they can also release granules in the extracellular space upon stimulation [7]. Neutrophils, in particular, can kill the infectious microorganism by using Reactive Oxygen Species (ROS), which are generated by the NADPH oxidase [8].

Macrophages are recruited to the inflammation site and are effective in ingesting microorganism as well as infected neutrophils. Once they have ingested the non-self-agent, macrophages migrate into lymph nodes or die [9].

Dendritic Cells (DCs), together with macrophages and B cells, act as the previously mentioned APCs: they

present antigens to the lymphocytes (T cells) using Major Histocompatibility Complex (MHC) molecules and secrete cytokines that stimulate innate immune cells. Despite the fact that they are considered as part of the innate immunity, macrophages and DCs are executing effectors of the adaptive immune system [10].

1.2.1 Dendritic cells and signals for adaptive immunity activation

How fast the immune system responds to Pathogen-Associated (PAMPs) and Damage-Associated Molecules (DAMPs) is correlated to the availability of innate immune APCs [11]. Dendritic cells are the most potent APCs due to their ability to prime naïve T cells. DCs can be divided into plasmacytoid DCs (pDCs), monocyte-derived DCs (moDCs) and conventional DCs (cDCs) that all share the same hemopoietic cell progenitor. Although scarce [12], **pDCs** have important immunomodulation capabilities in T-cell mediated immune responses during viral infections and in the rapidly release type I interferon (IFN) [13]. **moDCs** are monocytes that have differentiated into DCs under inflammatory conditions and have a back-up role during acute inflammation [14]. **cDCs**, with a stellate morphology, which clearly distinguishes them from macrophages, have been shown to be the major cell type involved in migration into secondary lymphoid tissue and priming of naïve T cells[6]. By leading to an activation of the immune system or to its muffling, DCs play a pivotal and delicate role in balancing homeostasis as a down or over-regulation of the inflammation could, indeed, cause damage and disease [15].

How do DCs communicate with the adaptive immune system? When immature DCs encounter the antigen and receive other immune stimuli, they become mature. Endocytic receptors are then down regulated, CD40, CD80 and CD86 maturation markers are up-regulated, the level of MHC class II is raised and DCs migrate to the lymphoid organs where they communicate with T cells by three different signals (Fig.3a). The first one is based on the interaction between T cell receptors (TCRs) and MHC complexes loaded with antigenic peptide on DCs. The second signal involves the production of co-stimulatory signals and the interaction of the T cell co-stimulatory receptor CD28 with the ligand B7-1 expressed by DCs [16]. The third one is the production of inflammatory cytokines, IL-12 for instance, that helps T cell activation [6]. Without these three proper instructions from DCs, T cells would not be able to acquire effector functions and form memory cells [17]. The nature of the antigen presented and of the cytokine produced during DC maturation influences T cell differentiation. DCs can produce pro-inflammatory (IL-6) or anti-inflammatory (IL-10) cytokines and different chemokines, resulting in the recruitment of different T cell subsets at the infection site. CD4+T cell can differentiate into different T helper type cells Th1, Th2 or Th17.

Immunological synapses form between MHC class I or II from DCs and naïve T cells to allow the transfer of information about pathogen invasion. Several adhesion receptors mediate this cell-cell junction (Fig.3b).

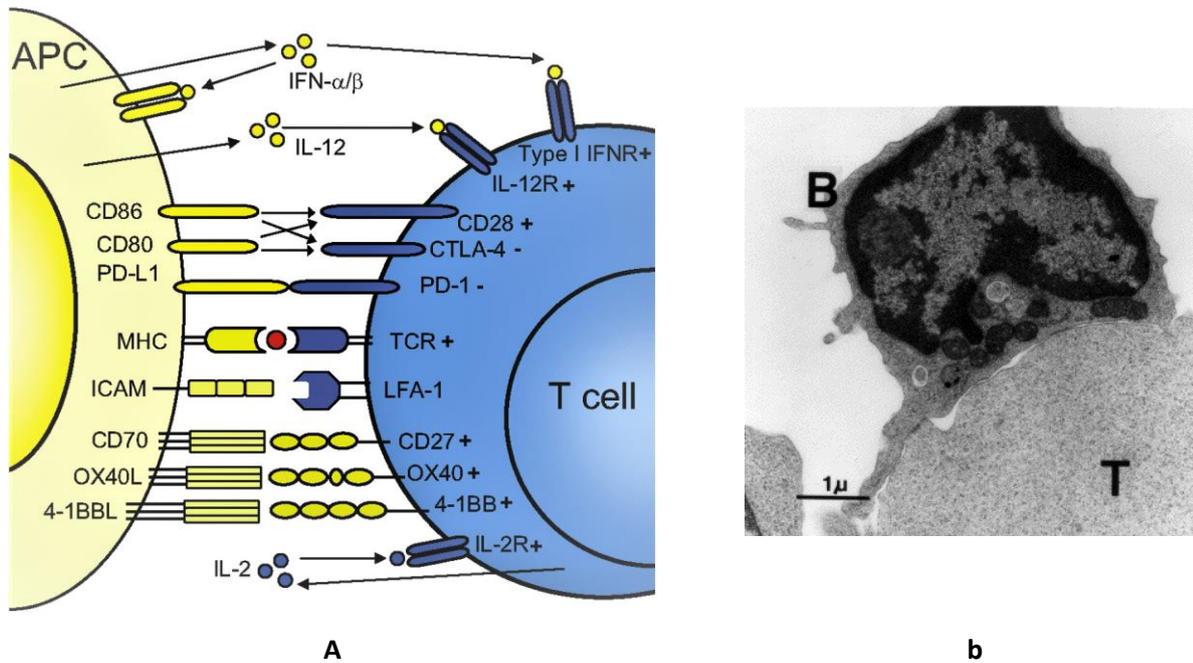


Fig.3 Activation of the adaptive immunity a) APC communication with T cells, adapted from [6], and b) electron micrograph (courtesy of J. W. Uhr) showing a B cell and T cell bound to each other. The bar = 1 μ m.
<http://www.biology-pages.info/I/ImmSynapse.html>.

1.2.2 Pattern recognition receptors

DCs, as professional sentinels of the innate immune cells, screen for pathogens by expressing **Pattern Recognition Receptor** (PRRs). They recognize non-self-elements and, thus, elicit the activation of immune response (inflammatory factor production). Two categories of patterns can be recognized by those receptors: Pathogen-Associated Molecular Patterns (PAMPs) and Damaged-Associated Molecular Patterns (DAMPs). Example of bacterial PAMPs are lipopolysaccharide (LPS), lipoproteins and peptidoglycan, while fungal PAMPs consist in carbohydrates of the cell wall [18]. PAMPs from viral origin are part of the glycoproteic envelope [19]. Once PAMPs are sensed by these receptors, DCs become activated with further production of chemokines and cytokines leading to inflammation (Fig.4). Inflammation could also be caused by damaged cells, e.g. by the above mentioned DAMPs, that results from tissue injury after bacterial infections. It is worth to note that DAMPs are also produced even in non-pathological conditions [20].

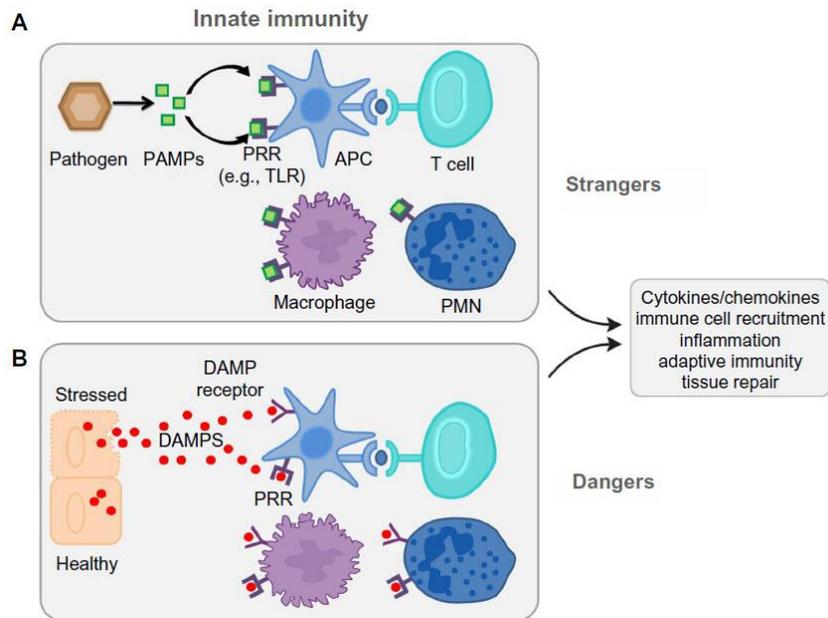


Fig.4 PAMPs and DAMPs. PAMPs and DAMPs interact with PRRs expressed by APCs with an effective cross talk between innate and adaptive immunity (PMN=polymorphonuclear leukocytes, it refers to granulocytes[21]). Adapted from [22].

Those PRRs include the transmembrane **C-type Lectin Receptors (CLRs)** and **Toll-Like Receptors (TLRs)**, together with the cytosolic receptors **Nucleotide-binding oligomerization domain-Like Receptors (NLRs)** and **Retinoic acid-inducible gene-I-Like Receptors (RLRs)** (Fig.5).

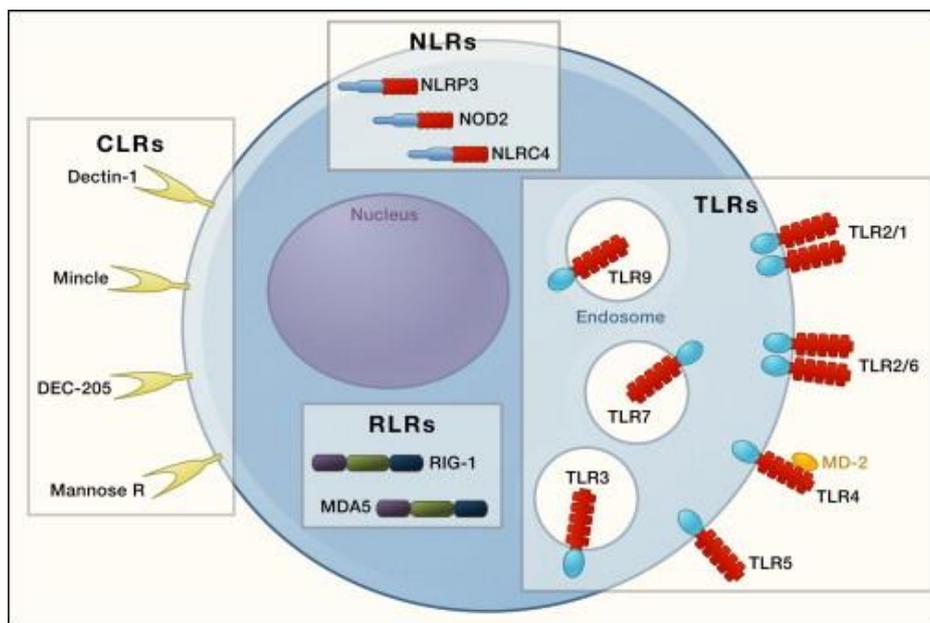


Fig.5 The four different families of PRRs. CLRs, TLRs, NLRs and RLRs. Adapted from [23].

TLRs, the first class of PRRs identified, are localized at the cell surface or at the endosome surface. TLR1,2,4,6 sense lipids, while TLR3,7,8 recognize viral RNA. TLR9 identifies bacterial DNA and it is strongly

involved in pro-inflammatory responses. Its deregulation could contribute to disease progression, e.g. sepsis [24]. Cross talk between PRRs signalling pathways can enhance the specificity of PAMP recognition and a focus on the synergistic activation of TLRs and CLRs will be given in chapter 2.2.2.

NLRs are involved in the regulation of inflammation and apoptosis during bacterial recognition. Finally, RLRs are helicases that sense viral RNA [25].

Depending on the antigen, on the PRR and on the APC involved, the immune system will be shaped towards an initiation of the immune response or towards the maintenance of self-tolerance (Fig.6).

This manuscript will focus on the specific PRR class of C-type lectin receptors, highlighting their structural and functional features and their contribution to the immune system response. C-type lectin receptors expressed by DCs are crucial for tailoring immune responses. They bind pathogens through the recognition of mannose, fucose, galactose and other **carbohydrate** structures. The combination of CLRs on APCs enables the recognition of most classes of human pathogens.

After pathogen uptake, several signalling pathways can induce the expression of specific cytokines and, consequently, trigger T cell differentiation. Some CLRs can directly induce activation of nuclear factor- κ B (NF- κ B), others affect signalling by Toll-like receptors (DC-SIGN) [26]. Therefore, CLRs represent an attractive target for immunotherapeutics.

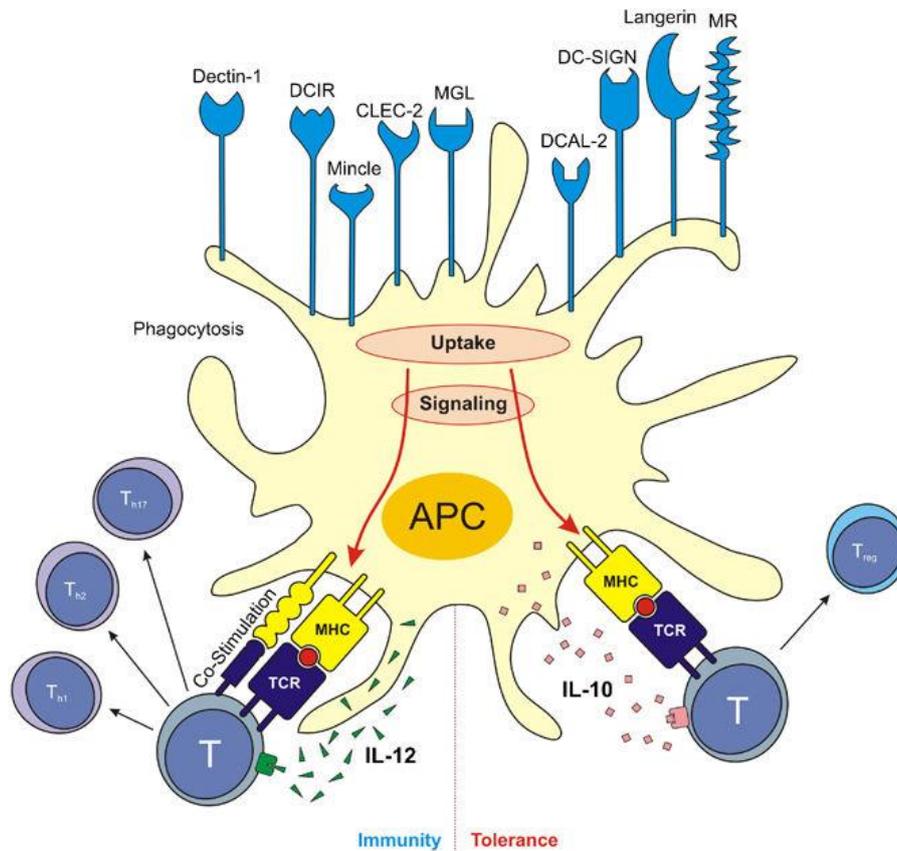


Fig.6 CLRs and immunity. CLRs on DCs are the first step on the activation of the adaptive immune response. Adapted from [27]

2. C-type lectin receptors

2.1 Introduction

C-type lectin receptors are Glycan-Binding Proteins (GBPs) belonging to the large family of lectins. The term lectins derives from the Latin *lectus*, the past principle of *legere*, which means to choose or select [28]. In 1954 W.C. Boyd proposed this name in order to highlight their peculiar specificity [29]. The definition of lectin excludes both antibodies and enzymes, as glycans are not a substrate whose biochemical nature will be altered upon binding to the lectin. Lectins recognize glycan thanks to specific structural domain called Carbohydrate Recognition Domain (**CRD**). To date, 14 different CRD folds have been described and four examples of lectin family involved in different aspects of the immune responses are shown in the following figure (Fig.7).

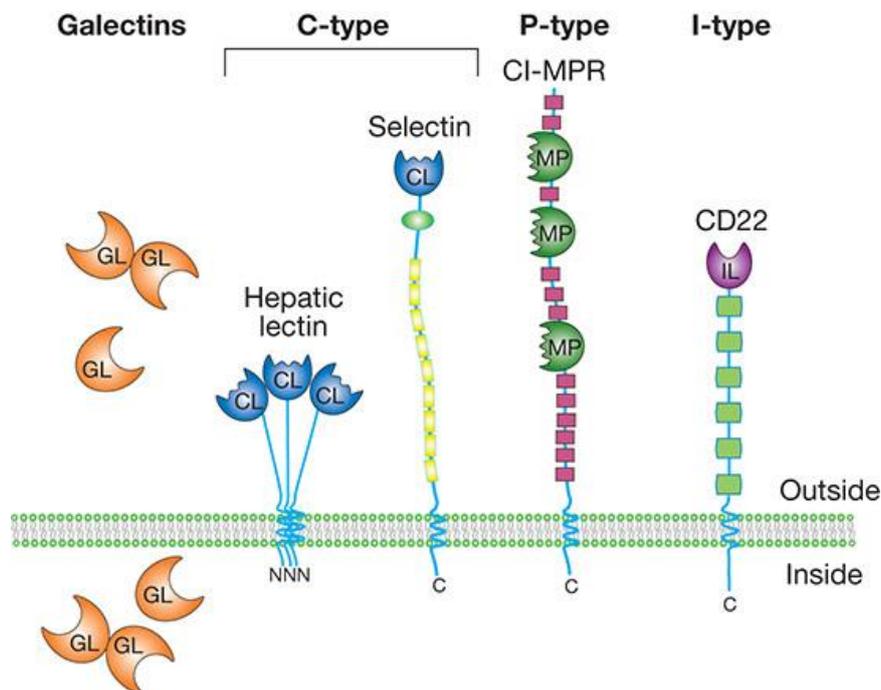


Fig.7 Examples of four animal lectin families. (GL) galectin, (CL) C-type lectin, (MP) P-type lectin, (IL) I-type lectin [30].

The functions mediated by lectins are very diverse. Some lectins mediate interactions between cells and the extracellular matrix, while others are involved in the immune response. For example, L-Selectin is involved in lymphocytes homing [31] whereas serum mannanose binding protein and ficolins activate the complement cascade [32]. Quality control is another important task of lectins. Calnexin and calreticulin in

the Endoplasmic Reticulum (ER) bind glucose on newly synthesized glycoprotein [33] and P-type lectins target lysosomal enzymes to endosomes by recognizing their mannose 6-phosphate [34].

In 1988, the group of Drickamer used for the first time the term « C-type lectin group » [35] to identify Ca^{2+} -dependent lectins.

C-Type Lectins Receptors (CLR) bind to glycolipids and glycoproteins through a well-conserved globular domain CRD.

2.2 C type lectin domains and glycan recognition: structural aspects

CLRs are divided into soluble or transmembrane proteins. In the latter, the CRD is connected to the cellular transmembrane region. In many cases, this connection occurs thanks to a coiled-coiled sequence termed neck domain. This neck domain is composed of a repeated sequence and, depending on the number of repetition and its length, the entire extracellular portion of the CLR, called ExtraCellular Domain (ECD), can oligomerize with a stoichiometry specific for the lectin for one lectin to another. Figure 8 schematises a tetrameric CLR with a neck oligomerization (Fig.8)

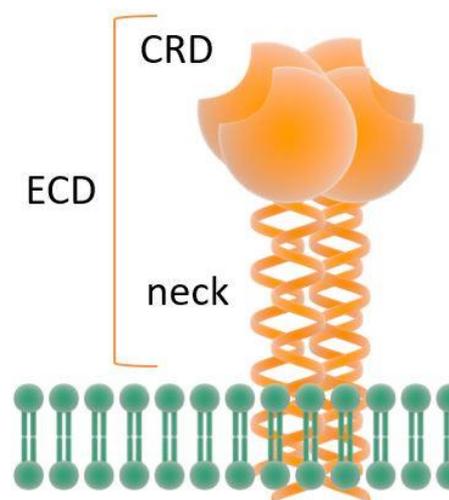


Fig.8 Schematic representation of a tetrameric CLR. The ECD is composed by the neck and the CRD.

To encompass proteins that contains a CRD domain with typical structural fold but that do not bind sugars [36] or lectins that conserve CLR characteristic properties but that are not Ca^{2+} dependent [37], the more general term « C-type Lectin-Like Domains » (CTLD) was introduced afterwards. However, for common usage, in this manuscript the term CRD will be used and not CTLD when referring to the carbohydrate recognition domain.

The CRD is, overall, a loop, with beta-strands at the N- and C-termini connected by two alpha helices and three antiparallel beta sheets (Fig.9).

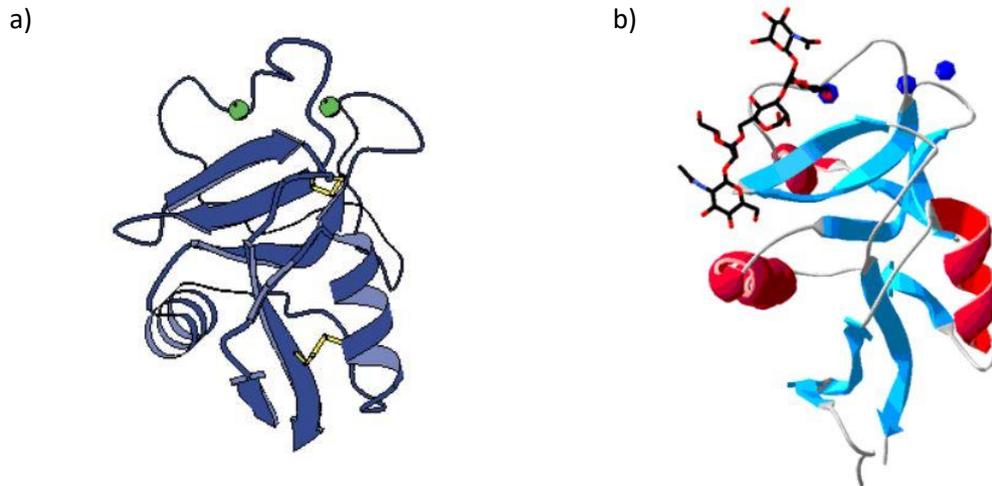


Fig.9 Carnohydrate recognition domain a) Generic CRD structure b) DC-SIGN CRD binding to GlcNAc-Man3-GlcNAc <http://www.imperial.ac.uk/research/animallectins/>

Four cysteines are involved in disulphide bridges, crucial for the correct folding of the protein CRD. Up to four Ca^{2+} binding sites are found, but only one is involved in glycan recognition and is particularly conserved. The other Ca^{2+} binding sites play an important role in the CLR structure stability. The conserved Ca^{2+} binding site is characterized by specific motifs that lead to different specificities of the glycan recognition. The EPN (Gluc-Pro-Asn) motif leads to the interaction with mannose-type ligands (mannose, GlcNAc, glucose), that contains 3-hydroxyl and 4-hydroxyl groups in equatorial position (Fig.10a), while the QPD (Gln-Pro-Asp) motif drives the recognition towards galactose-type glycans (galactose and GalNAc), where the 4-hydroxyl group is axial (Fig.10b)[38]. In both cases the carbonyl side chains are involved in hydrogen bond formation with the specific monosaccharide and also coordinate two bonds with Ca^{2+} . The proline is involved in the backbone conformation.

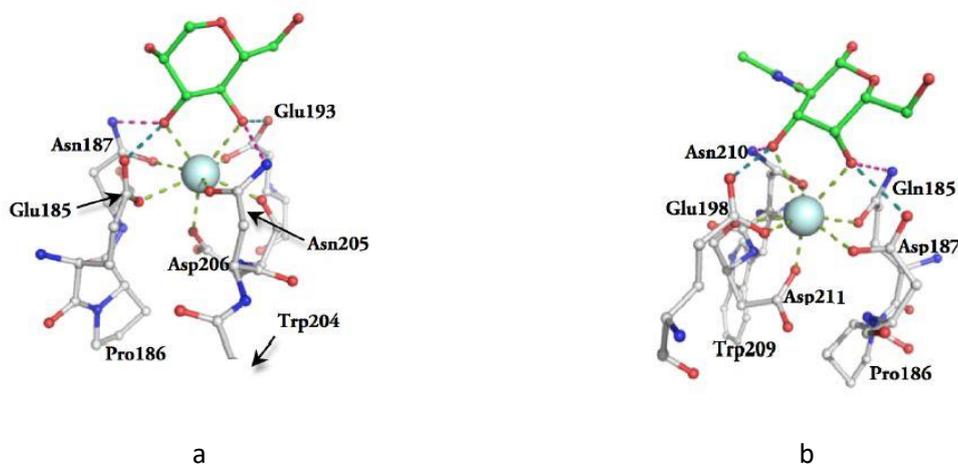


Fig.10 EPN and QPD motif. a) EPN, mannose-type, motif b) QPD, galactose-type, motif. Coordination bond are in green, H-bonds where hydroxyl acts as acceptor or donor are marked in pink or cyan dashed lines, respectively. Adapted from [35].

In addition to the primary binding site, secondary binding sites could be crucial in the glycan recognition and allow a broader interaction with the protein surface [39].

17 different groups of CRDs are described, depending on their domain architecture, phylogeny and function (Fig.11). Group III or Collectin group, for example, includes mannose-binding protein (MBP) with the typical collagen-like triple helical tail. E-Selectin belongs to the group IV and selectins are leukocyte adhesion molecules specific for sialyl-Lewis^x [40]. macrophage mannose receptor (MMR) belongs to the group VI [41],[35].

All the protein studied during my PhD belong are involved in the activation of immunity. For example, Dendritic-cell-specific intercellular adhesion molecule-3-grapping non-integrin (DC-SIGN) and Mincle belong to type II receptors (group II) and they are characterized by a short cytoplasmic tail, a transmembrane domain and an extracellular domain ending with a Ca²⁺-dependent CRD. NK Receptors (Group V) include, as well, transmembrane proteins, e.g. Dectin1, with short cytoplasmic domain but with a CRD often lacking the Ca²⁺ dependency.

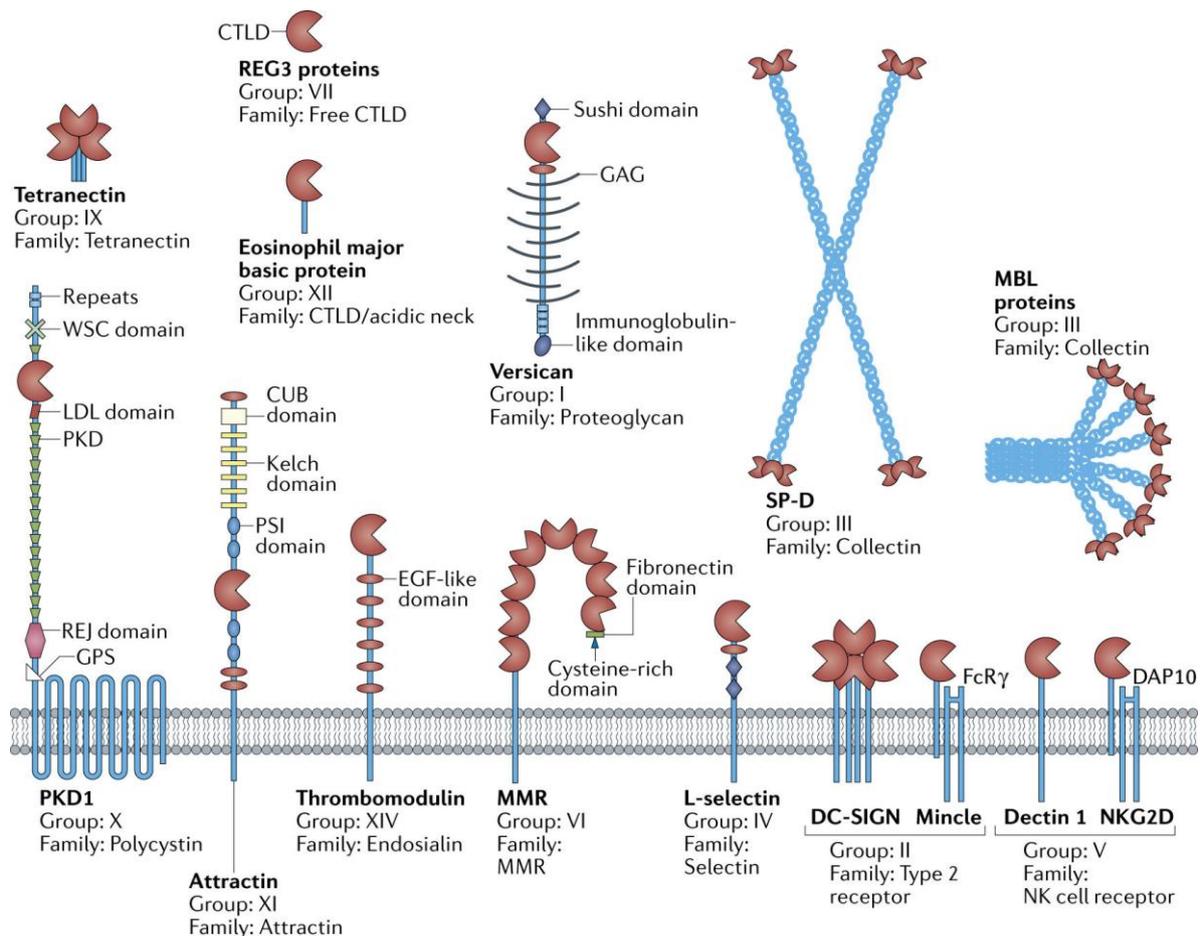


Fig.11 C-type lectin subfamilies [42]

2.2.1 CLR signalling

Once a CLR encounter a pathogenic ligand, **four** different ways could be exploited to warn the immune system (Fig.12).

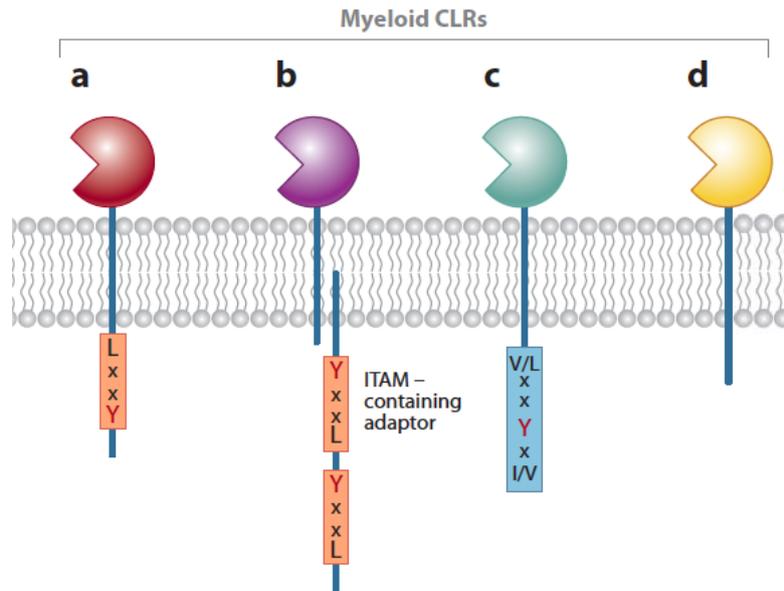


Fig.12 Overview of the four motifs for signalling a) direct CLR coupling to Syk b) Indirect CLR coupling to Syk c) CLR with ITIM domain d) CLR without ITAM or ITIM domain adapted from [43]

- **a) Direct CLR coupling to Syk:** CLR cytoplasmic tail possesses a tyrosine-based motif called hemiITAM (hemi-Immunoreceptor Tyrosine-based Activation Motif). Dimerization of two phosphorylated CLR leads to recruitment of the tyrosine kinase Syk via its tandem SH2 domains. This path leads to myeloid cell activation [43]. An example is dectin-1.
- **b) Indirect CLR coupling to Syk:** CLR cytoplasmic tail possesses a tyrosine-based motif and requires an ITAM containing adaptor molecule, such as Fc Receptor γ -chain (FCR γ). Again the phosphorylation events trigger Syk activation [43]. Examples are dectin-2 (Fig.13a), BDCA2 and mincle.
- **c) CLR with Immunoreceptor Tyrosine-based Inhibition Motif (ITIM) domain** recruits phosphatases leading to a subsequent down regulation of immune response [43]. An example is the Dendritic Cell ImmunoReceptor DCIR.
- **d) CLR without ITAM or ITIM domain:** these CLR can signal after antigen capture of by other signalisation motives and modulate the signalling of other receptors. Examples are DC-SIGN are langerin [41]. DC-SIGN signalisation will be detailed afterwards.

The intracellular signalling pathways activated by dectin-2 and DCIR will be detailed below (Fig.13).

Dectin-2 is a CLR that indirectly couples with Syk and the association with FcR γ via an arginine residue is required for Dectin-2 surface expression. When dectin-2 recognizes fungi, the activation of Syk regulates the production of ROS [44], used as microbicidal agents, and the activation of complex involved in gene transcription regulation (Fig.13a).[43]

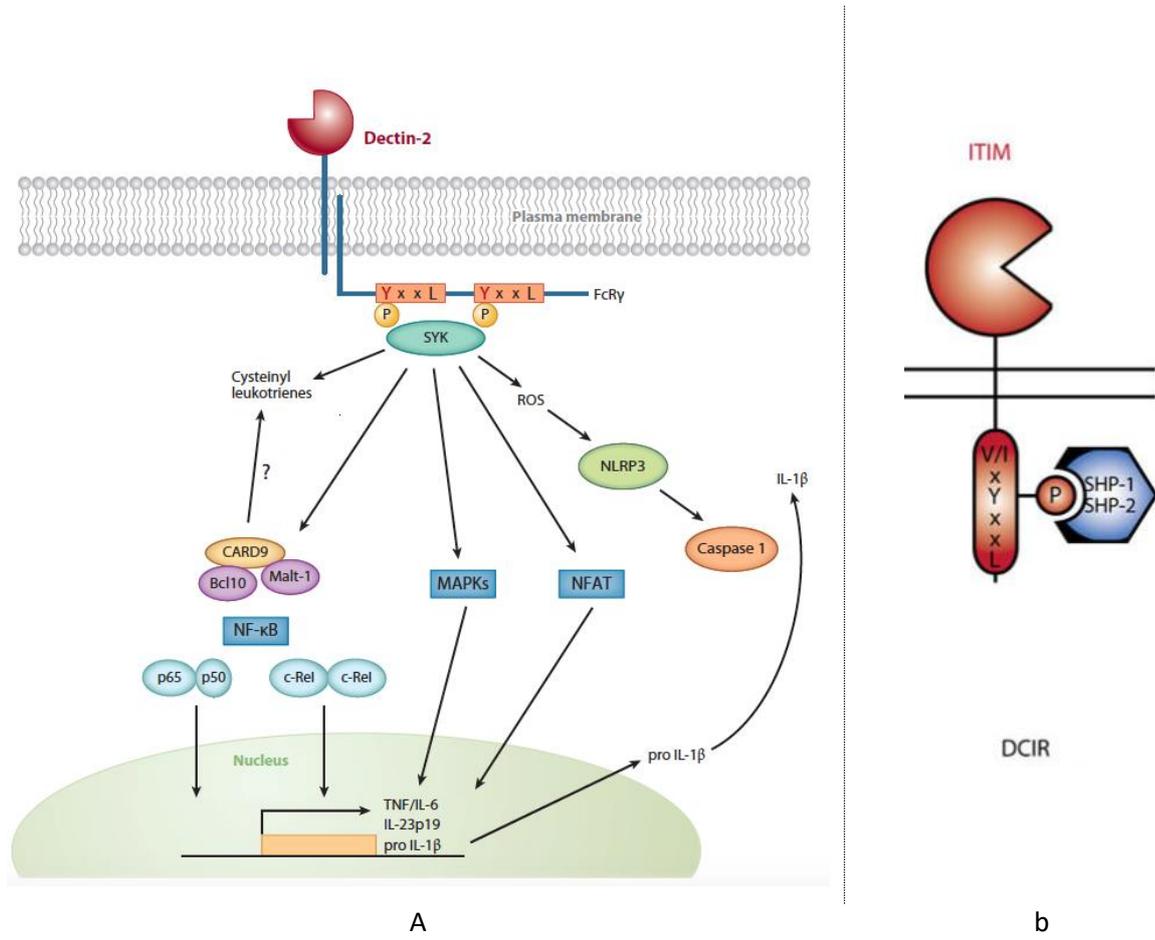


Fig.13 Dectin-2 and DCIR signalling. a) Dectin-2 indirect coupling to SYK and intracellular signalling pathway, adapted from [43] b) DCIR ITIM domain and phosphatase recruitment, adapted from [45].

DCIR, in contrast, bears an ITIM domain believed to mediate inhibitory signals in DCs (fig 13b). Phosphorylated ITIMs mediate recruitment of the SH2-containing tyrosine phosphatase-1 and 2 (SHP-1/2) which negatively controls NF- κ B signalling to antigen response[45].

DC-SIGN do not possess any ITAM or ITIM domain. Nevertheless, it possesses in the cytoplasmic tail a tyrosine motif, necessary for the intracellular signalling [46], and a di-leucine motif involved in endosomal/lysosomal pathway [47]. This motif was shown by *Engering et al* [48] to be involved in the internalization of DC-SIGN-ligand complex.

The outcome of the signalling depends on multiple aspects, such as the type of receptor, the ligand (nature, architecture, density) and the internalization of the receptor involved. Nevertheless, some constants are

kept among the different CLR signalling: the tyrosine-based motif, the use of SYK and SHP for the modulation of gene transcription. Another common point is the ability to cooperate with other receptors to regulate myeloid cell functions [43].

2.2.2 CLR and TLR crosstalk

CLRs expressed at the surface of APCs are not only involved in pathogen recognition but also in their internalisation and a simultaneous activation of CLRs and TLRs can occur to trigger the appropriate immune response. Both PRRs lead to the activation of both nuclear factor-kappa B (NF-κB) and the mitogen-activated protein kinase (MAPK), resulting in an overlapping production of cytokines and chemokines. This net effect strikes, for example, during *Candida* infections, as shown in Figure 14, with a cooperation between TLR4/2 and dectin-1/2. MtD88 adaptor molecule is used by TLRs for the activation of the Interleukin Receptor-Associated Kinases IRAK1, IRAK2 and IRAK4 and subsequent ubiquitination of TNF Receptor Associated Factor TRAF6. Finally, MAPK and NF-κB are activated downstream. CLRs interaction with Syk triggers the recruitment of the Card9/Bcl10/Malt1 protein complex, leading as well to the MAPK and NF-κB activation (Fig.14) [49]. The final outcome is inflammation and the production of cytokines.

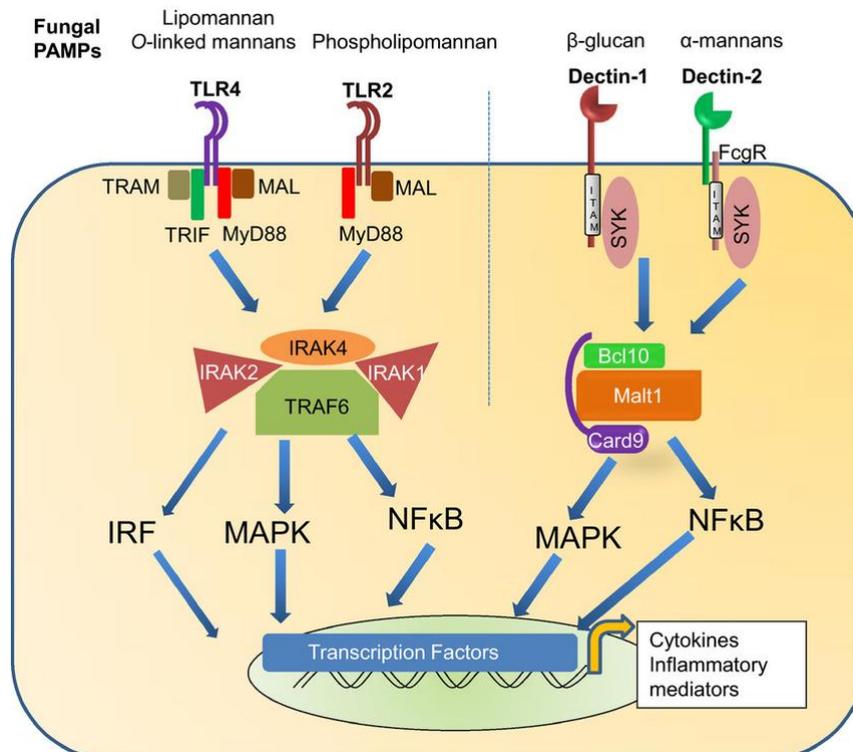


Fig.14 Cross talk between CLR and TLR a) Dectin1/2 and TLR4/2 during *Candida albicans* infection adapted from [49]

In 2018 *S. Gringhuis et al.* [50] have identified the mechanism by which DC-SIGN modulates TLR-dependent responses in human DCs (Fig.15). Mannosyl caps on the terminal D-arabinan (manLAM), found in pathogenic *Mycobacterium*, interact with DC-SIGN, leading to the activation of Raf-1. Activation of Raf-1, in turn, allows the acetylation of p65, the activating subunits of NF- κ B, but only after TLR signalling had activated NF- κ B. Indeed, TLR enables p65 translocation to the nucleus where the latter can then get activated by DC-SIGN downstream pathway. The acetylation of p65 extends the transcriptional activity of NF- κ B and boost the transcription rate of anti-inflammatory IL10 gene.

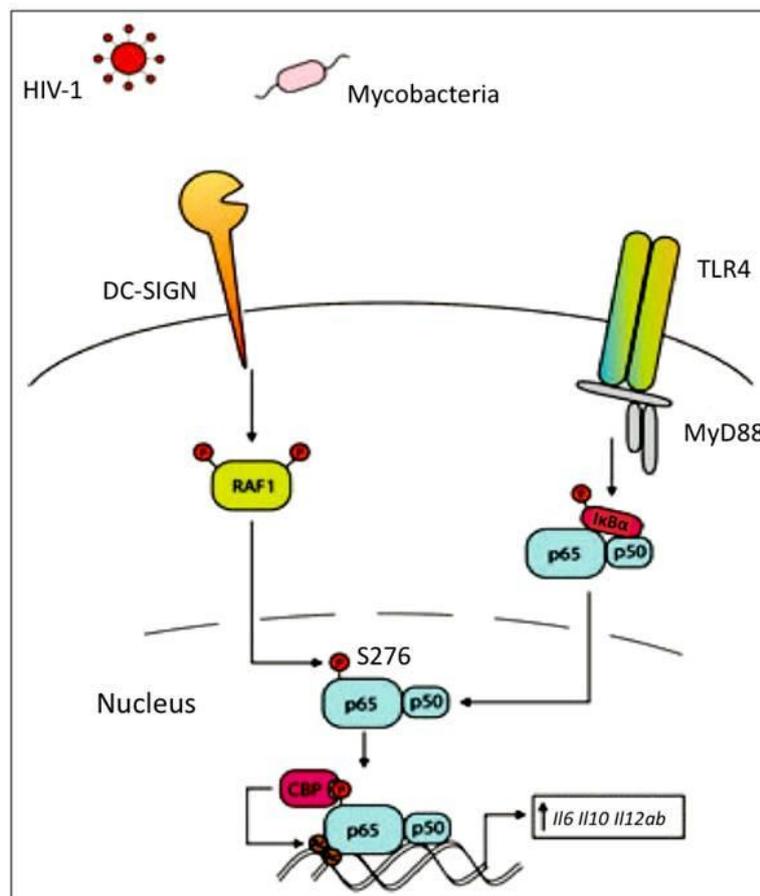


Fig.15 DC-SIGN and TLR4 during Mycobacterium infection. p65 and p50 subunits form NF- κ B complex. Adapted from [51]

2.3 CLR considered in the study

This chapter will focus on the nine different CLR investigated during my PhD: BDCA2, DC-SIGN, DC-SIGNR (L-SIGN), dectin2, dectin1, langerin, LSECtin, MCL and mincle. For each of them some information will be given, from structure features to binding specificities, from their contribution to the signalling cascade to the pathologies in which they are involved.

2.3.1 Blood Dendritic Antigen 2 (BDCA2)

BDCA2 is the only CLR considered in this study exclusively expressed by pDC. BDCA2 CRD has a typical CLR with an EPN sequence at the Ca^{2+} binding site that should lead to interaction with mannose-type glycans. However, Glu178 of the EPN motif is positioned outside the calcium binding site, which is partially occupied by the side chain of Arg179 [52]. A first study [53] in 2011 by glycan array indicated an unusual binding of BDCA2 towards galactose-terminated biantennary glycans. In 2015 *S. Jegouzo et al.* [54] identified that BDCA2 binds in a very selective way glycans containing the epitope $\text{Gal}\beta 1-3/4\text{GlcNAc}\beta 1-2\text{Man}$. Resolution of the CRD structure in the presence of this trisaccharide revealed that the mannose residue interacts with the primary binding site, while the other two sugars contribute to the interaction by occupying « a shallow groove » (Fig.16).

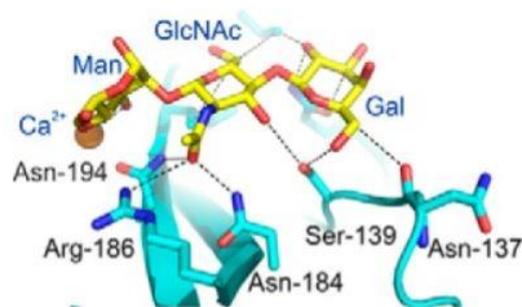


Fig.16 Portion of BDCA2 CRD structure. Complex of BDCA2 CRD with $\text{Gal}\beta 1-4\text{GlcNAc}\beta 1-2\text{Man}$ adapted from [54]

BDCA2 does not have an endocytic activity, being exclusively involved in intracellular signalling pathway. Its cytoplasmic domain does not contain any known signalling motif, therefore it must associate to the transmembrane adaptor FCR γ , which interferes with TLR9-induced activation of pDCs and leads, eventually, to the inhibition of type I IFN secretion [55]. This is an attractive feature to evade type I IFN responses that is used by Hepatitis B virus to facilitate its spreading [56]. Finally, it was also shown that colorectal cancer cells express BDCA2 ligands [55].

2.3.2 DC-SIGN and DC-SIGNR (L-SIGN)

A lot could be said on DC-SIGN (by typing “DC-SIGN” on the research tool of PubMed, one would find 1500 hits). Many thesis in our group have indeed focused on the study and the targeting of DC-SIGN¹. However, for the sake of simplicity, in this chapter I will mainly focus on the similarities and differences between DC-SIGN and its related CLR, DC-SIGNR, both studied during my PhD. For relevant publications based exclusively on DC-SIGN please refer to [57], [58], [59],[60],[61].

The names themselves indicate their importance in the initiation of T cell immunity by interacting with ICAM-3. Dendritic Cell Specific Intracellular adhesion molecule-3 (ICAM-3) Grabbing Nonintegrin (DC-SIGN) and DC-SIGN Related (DC-SIGNR) also termed Liver/Lymph node-Specific Intercellular adhesion molecule-3-Grabbing integrin (L-SIGN) appear to be the product of a gene duplication and share 77% of homology [62]. The EPN motif is present in both proteins and enables mannose binding.

Their binding to mannose derived glycans was studied by *H. Feinberg et al* [63],[64]. As typical feature of mannose binding mode to CLR, equatorial 3- and 4-OHs of the internal sugar form both coordination bonds with the Ca²⁺ (Fig.17). In addition, they form hydrogen bonds with amino acids that also serve as Ca²⁺ ligands. Moreover, the 6-OH forms a water-mediated contact with Asp367 (Asn379 in DC-SIGNR).

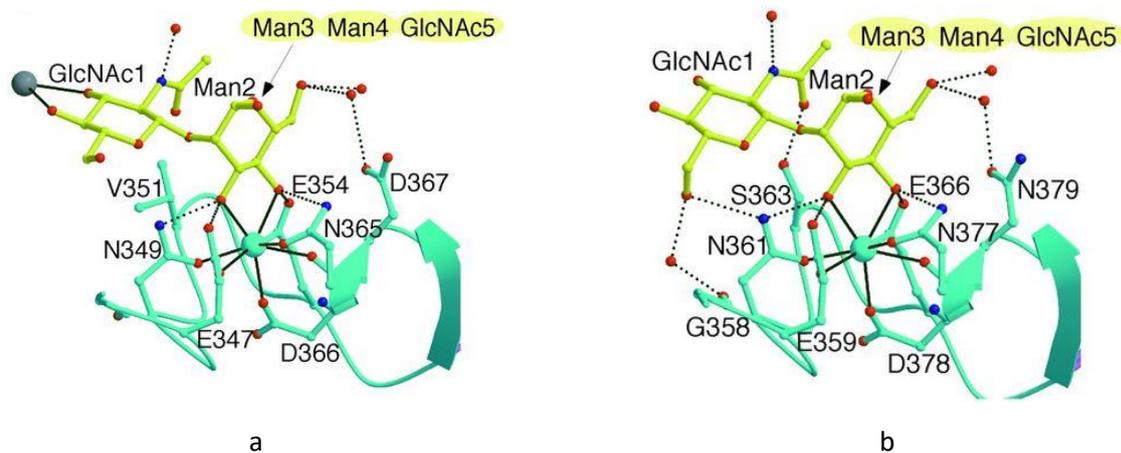


Fig.17 Portion of DC-SIGN and DC-SIGNR CRD structure a) DC-SIGN and b) DC-SIGNR X-ray structure in complex with GlcNAc2- Man3. Adapted from [63].

One primary difference between the two CLRs lies in the binding sites: Val351 of DC-SIGN is replaced by Ser363 in DC-SIGNR. This is the main reason why DC-SIGNR does not interact with fucose-containing Lewis^x antigens, while DC-SIGN enables the van der Waals interactions necessary for fucose interaction [37].

¹ <https://tel.archives-ouvertes.fr/tel-00819832> and <https://tel.archives-ouvertes.fr/tel-01497502>

For both DC-SIGN and DC-SIGNR, the neck repetitions allow protein oligomerization into a tetramer with an affinity increase for glycan recognition. The final coil-coil unit of their neck domain is the most divergent in term of residues and confer different CRD presentations [65],[66]. SAXS analysis performed by Franck Fieschi group [57] revealed different coiled-coil neck arrangements that lead to structural differences : the « closed flower » conformation of DC-SIGN is characterized by the CRD alignment with the neck domain, while the « opened flower » organization of DC-SIGNR prevents DC-SIGNR to have the suitable oligomeric orientation for the recognition of some DC-SIGN ligand (Fig.18).

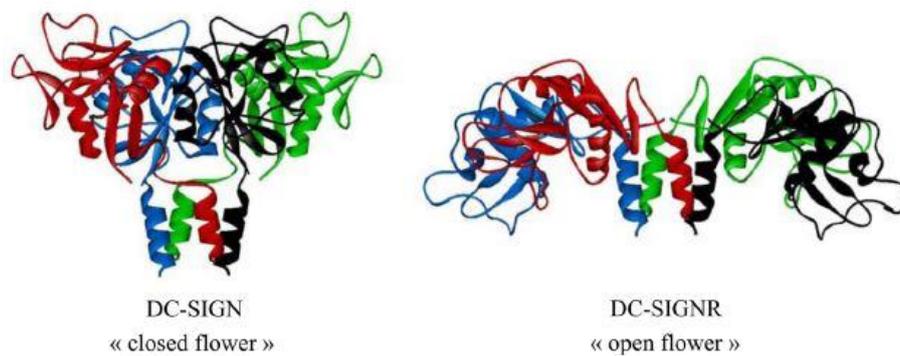


Fig.18 Comparison of the CRD from DC-SIGN and DC-SIGNR. Closed flower for DC-SIGN and open flower for DC-SIGNR, adapted from [57].

The different CRD orientation of the two lectins is responsible for different binding efficiency towards multivalent ligand. This has been recently confirmed by glycan functionalized quantum dots that showed that the spatial orientation of DC-SIGNR does not favour binding to multiple mannose residues [67].

The two receptors are also characterized by different expression patterns. DC-SIGN is highly found at the surface of monocytes and in subsets of immature and mature DCs in dermis, mucosa, spleen and placenta, while the name “L-SIGN” comes from the fact that it is expressed in endothelia cells of lymph nodes and liver but not on DCs. Similarly to DC-SIGN, DC-SIGNR is involved in HIV-1 virus infection by binding to gp120 [60]. Moreover, being found in placenta, it is implicated in the HIV-1 virus mother to foetus transmission (Fig.19) [68].

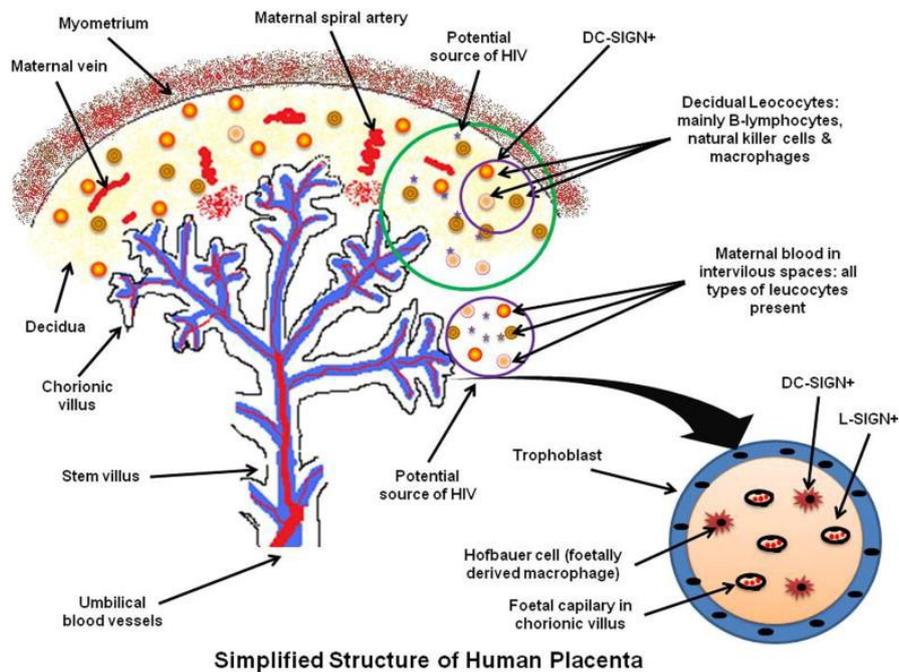


Fig.19 The placenta, DC-SIGN/DC-SIGNR and HIV. The placenta plays an important role in the transmission of HIV-1 infection from mother to foetus. Released viral particles may become adsorbed onto DC-SIGNR on the placental capillary endothelium [68].

DC-SIGNR can establish interactions with ICAM-3-expressing T cells and this may enable activated T cells to recirculate to the liver and to the lymph nodes [69]. DC-SIGNR do not possess ITAM-ITIM domains and shares potential internalization di-leucine motif with DC-SIGN. Nevertheless, it is still under investigation whether it is an endocytic receptor, while DC-SIGN is confirmed to be involved in pathogen uptake.

2.3.3 DC-Associated C-Type Lectin 1 (dectin-1)

Dectin-1 is a natural killer (NK)-cell-receptor-like and it is found on the surface of peripheral blood leukocytes and DCs in muscle, stomach [70] and lung [71]. It is also called β -Glucan receptor because it recognises β -1-3 linked and β -1-6 linked glucan (laminarin and zymosan) from bacteria and fungi (*A. fumigatus*, *C. neoformans*, *C. albicans*). Its presumed binding site is a « shallow groove » [37] defined by Trp221 and His223, and modelling studies suggest that the interaction with laminarin is mostly driven by hydrophobic forces. Unfortunately, all the attempts to crystalize the CRD with long β -glucan chain have failed so far. One peculiarity of this lectin is that it does not require metallic ions for binding. Nevertheless, a Ca^{2+} binding site was found and a structural role was attributed to it. Figure 20 presents the X-ray structure of murine dectin-1 binding to β -glucan chain [37].

Gatner et al. [72] observed that during the recognition of zymosan, both dectin-1 and TLR2 were recruited and their synergistic collaboration was established as essential for the antifungal immunity.

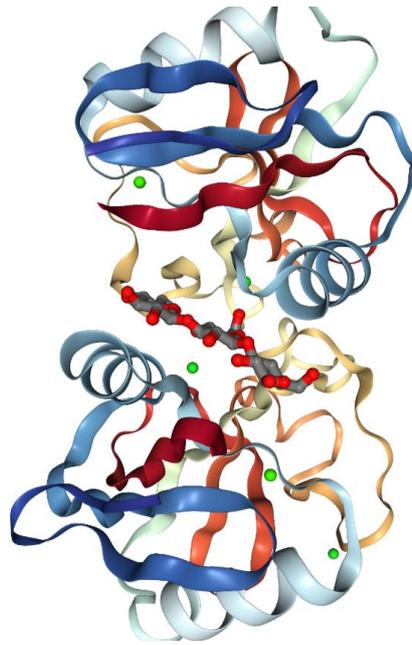


Fig.20 Murine dectin-1 structure. Interaction with β -glucan chain [37]

Dectin-1 possesses an hemiITAM motif that recruits Syk when phosphorylated. Dectin-1 is involved in many immune functions: recognition of fungi induction of respiratory burst [73], cytokine and chemokine production and differentiation of T helper cells [12]. As a consequence, defective surface expression of Dectin-1 results in mucosal fungal infections [74].

2.3.4 Dendritic cell associated C type lectin 2 (Dectin-2)

Dectin-2 belongs to the same family as BDCA2 but, unlike BDCA2, it is expressed by macrophages and dendritic cells in lung, spleen, lymph node and tonsils [75]. Dectin2 was predicted to interact with mannose-type sugars because of the presence of the EPN motif. It was known to recognize *C. albicans*, *S. cerevisiae*, *M. tuberculosis* and *M. audouini* [76] but the specific epitope remained undetermined until 2013. The first hit was given by *T. Ishikawa et al.* [77] who found that the O-linked oligosaccharide from the fungus *Malassezia* recognised by dectin-2 is a α 1-2-linked mannobiose. Finally, in 2017 [78] dectin-2 structure was solved and the binding to $\text{Man}\alpha$ 1-2Man disaccharide was explained. The geometry of the binding site allows binding to terminal disaccharide units but also to $\text{Man}\alpha$ 1-2Man in internal positions in more complex oligosaccharides (Fig.21).

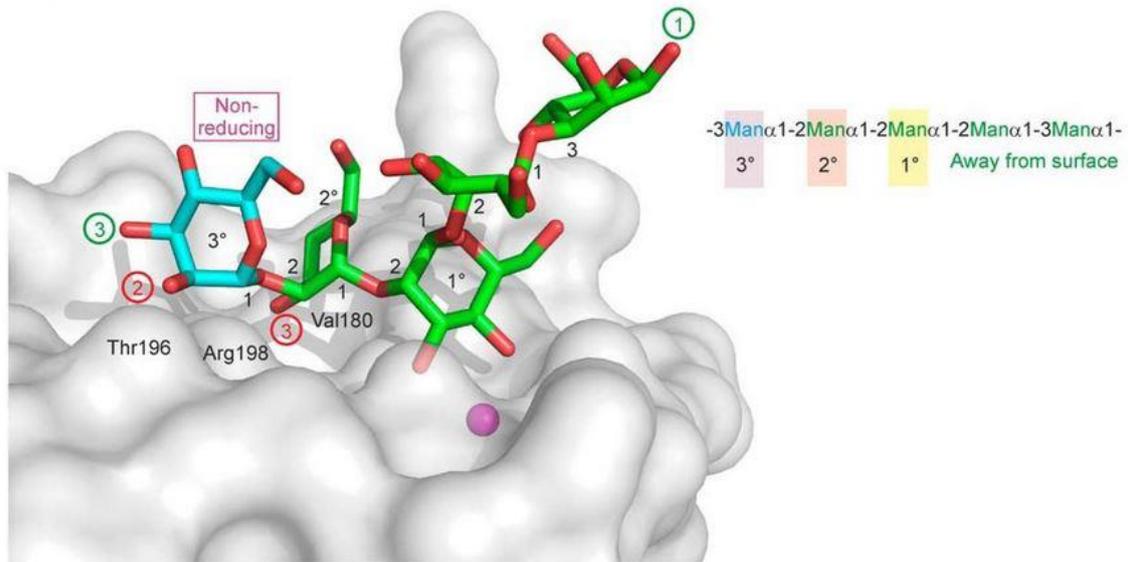


Fig.21 Portion of Dectin-2 structure. Interaction with mannose based oligosaccharide, adapted from [78].

As for BDCA2, the lack of a cytoplasmic consensus motif requires the interaction with the adaptor FCR γ through a positively charged residue of the transmembrane region. Dectin-2 does not have an endocytic activity, being exclusively involved in intracellular signalling pathway. It is involved in the activation of CARD9 upon interaction with the hyphal form of *C. albicans* [79] (see figure 14).

2.3.5 Langerin

Langerin is characterized by a peculiar sugar specificity: although it possesses an EPN motif in its primary binding site, it also interacts with sulphogalactosides *via* electrostatic interactions between Lys299 and Lys313 and the sulphate in position 6 (Fig.22a)[80]. In addition, langerin coil-coiled repetition allows a trimeric oligomerization (Fig.22b)[81] and the interface between the trimer generates the binding sites for calcium independent binding to different sulphated glycans such as glycoaminoglycans (heparin, heparin sulphate, chondroitin sulphate) [82] [80] (Fig.22b and c).

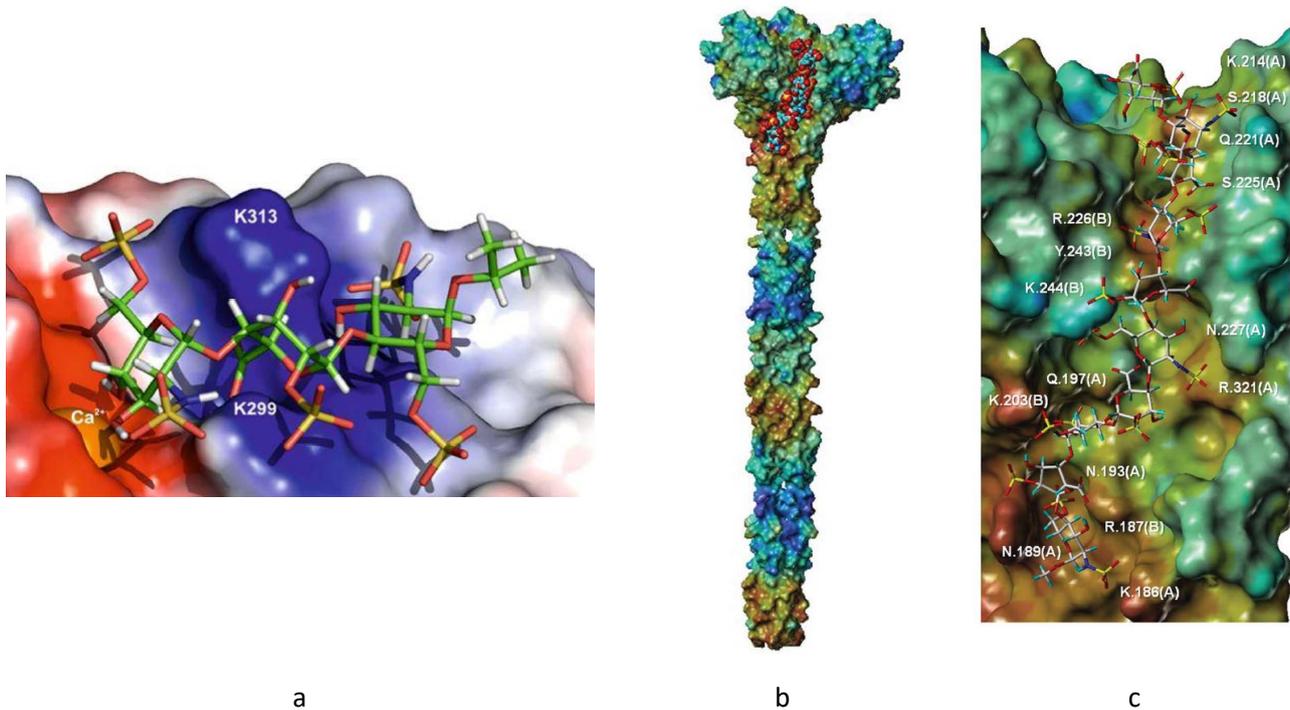


Fig.22 Langerin structure. Langerin interaction with sulphogalactosides, adapted from[80] b) Trimeric structure of langerin and c) calcium independent binding to heparin. Adapted from[82].

Langerin recycles through early endosomal compartments Langerin and DC-SIGN do not have the same behaviour towards HIV-1 infection. While DC-SIGN resides in DCs from the subepithelium, langerin is specifically expressed by DCs of the skin and mucosa called Langerhans cell (LCs) and induces the formation of Birbeck granules, specific cytoplasmic organelles involved in antigen processing (Fig.23). LCs internalization pathway is thought to be involved in the inhibition of HIV-1 transmission upon binding with langerin, whereas DC-SIGN is responsible for the spreading of the virus [83].

Langerin signalling appears to control the LC endosomal trafficking and endocytosis is regulated by a proline-rich motif in the receptor tail. It remains unknown if it regulates DC activation [43].

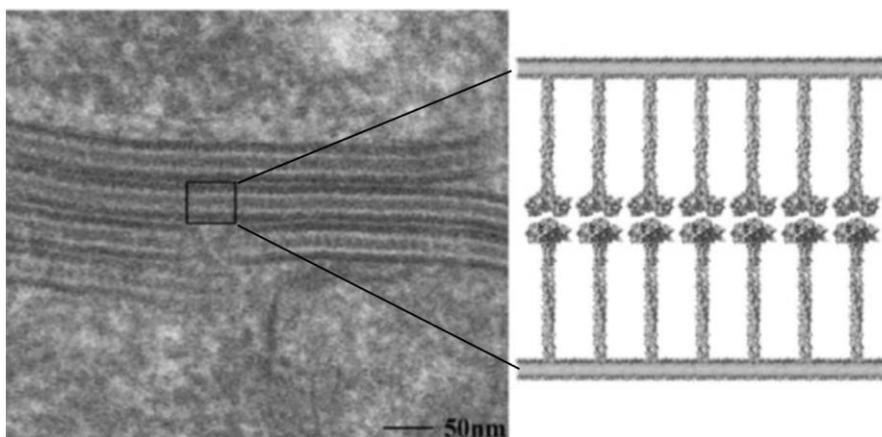


Fig.23 Birbeck granules. Role of langerin in Birbeck granule formation. Adapted from [81].

2.3.6 Liver and lymph node sinusoidal endothelial cell C-type lectin (LSEctin)

LSEctin was first described in 2004 [84] as closely related to DC-SIGN and DC-SIGNR in term of protein domain composition and genomic organization (Fig.24).

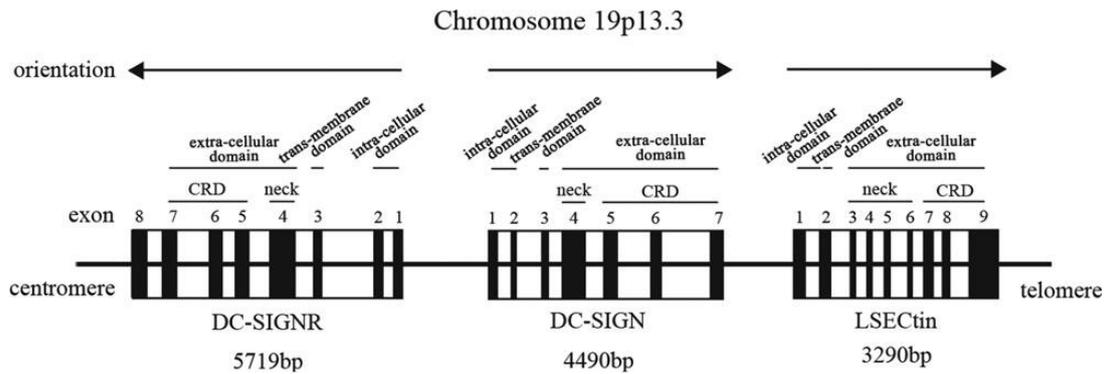


Fig.24 Genomic organization of DC-SIGNR, DC-SIGN and LSEctin [85].

Although reported to be exclusively expressed on liver and lymph node sinusoidal endothelial cells, LSEctin has been later found to be expressed in ex vivo isolated human peripheral blood and thymic DCs [86]. LSEctin is shown to initiate signalling after binding to Ebola *via* association with a 12-kDa DNAX-activating protein (DAP12) and to induce Syk activation [87].

Its neck sequence contains cysteine residues that are potentially involved in inter-chain disulphide bond formation. This may results in a dimerization of LSEctin (Fig.25) [88], although a unique oligomeric structure has not been defined yet.

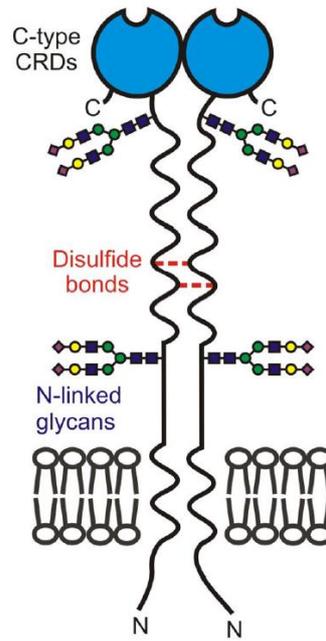


Fig.25 LSECTin. Hypothetical dimeric conformation of LSECTin. Adapted from [88]

LSECTin possesses an EPN motif and it should recognize mannose and related sugars, sharing with DC-SIGN and DC-SIGNR the ability to bind to envelope viral proteins. However, the three lectins bind different viruses: LSECTin recognizes Ebola and Severe acute respiratory syndrome (SARS) but not hepatitis C virus, which is recognized by DC-SIGN and DC-SIGNR [86]. The first extensive investigation on LSECTin binding specificity was performed by *A. Powlesland et al.* [88] on an array of truncated viral glycans. They identified GlcNAc β 1-2Man as the minimum binding epitope. However, LSECTin CRD structure is yet to be solved, in order to get a closer view of its interaction with glycans.

HIV virus is recognised by DC-SIGN and DC-SIGNR, while no virus capture was observed by LSECTin. However, *T. Gramber et al* [89] performed biological study which proved that LSECTin was able to interact with the HIV glycoprotein (gp120) but not with the HIV particle and, therefore, that the interaction with gp120 is not necessarily correlated to the capture of the virus. Moreover, they found that LSECTin, likewise DC-SIGN, is able to have an endocytic activity but with a different intracellular fate [89].

2.3.7 Macrophages C-type lectin (MCL)

Among the CLRs considered during my PhD, MCL is the less characterized. It is expressed by monocytes and macrophages. The consensus Ca²⁺ binding EPN motif is changed into EPD, but remains able to coordinate Ca²⁺ ion[90] (Fig.26).

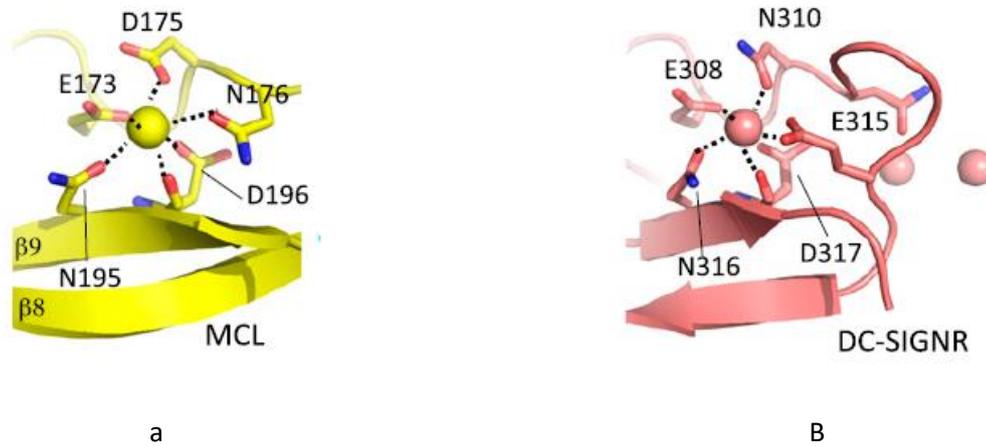


Fig.26 MCL and DC-SIGNR. Comparison of the putative ligand binding sites in MCL and DC-SIGNR. Adapted from [90].

Common carbohydrate structures are not recognized by this unusual lectin, as shown by a microarray screening using MCL-fc construct on the CFG website and by experiments performed by the collaboration between our group and Niels Reichardt (see results chapter). Very little is known about MCL alone; however, it is often studied coupled to Mincle. *Miyake et al.* [91] reported that MCL induces Mincle expression upon stimulation by cord factor (trehalose-6,6'-dimycolate, TDM) as they observed a failure in the activation of the immune system in MCL-deficient mice. Since Mincle is barely detectable on resting cells, it is assumed that MCL recognizes TDM and then induces Mincle expression (Fig.27).

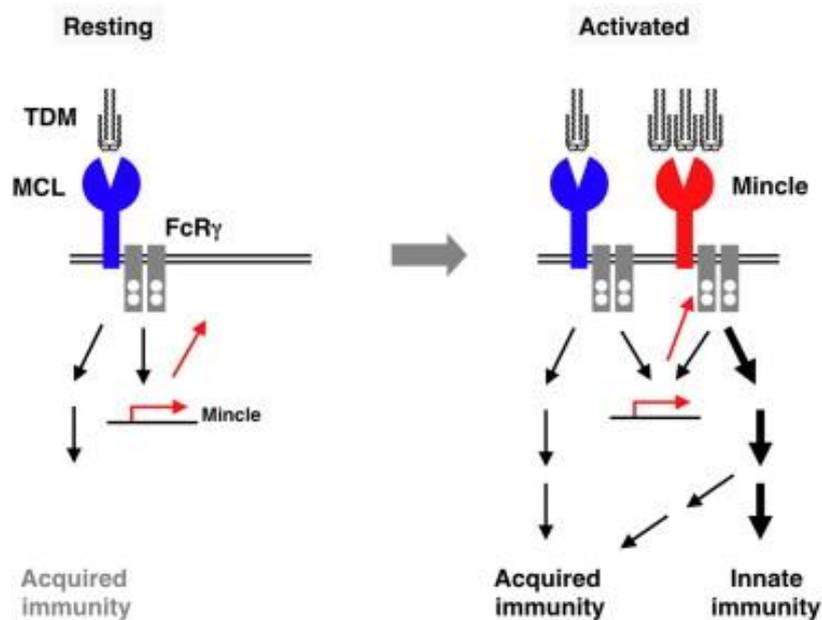


Fig.27 Mincle expression. The supposed expression of Mincle upon MCL interaction with TDM, adapted from [91].

The hypothetical TDM binding site of MCL may be a hydrophobic region that recognizes the acyl group of TDM [90]. MCL interaction with its ligands is an example of weak lectin-glycan binding that are not detected *in vitro*.

2.3.8 Macrophage inducible Ca²⁺- dependent lectin (Mincle)

Mincle is the only CLRs that specifically recognizes glycolipids [92]. It is expressed by macrophages, DCs, B cells, and neutrophils [93]. It strongly interacts TDM which is found in the cell wall of *Mycobacteria*, and it was the first PRR identified to recognize cholesterol crystals [94] through a Cholesterol Recognition Amino acid Consensus-like (CRAC) motif. The crystal structure of the interaction between cow Mincle CRD with TDM has been studied [95]. TDM (Fig) possesses two glucoses that bind to the two primary binding sites of Mincle, forming hydrogen bonds via their OH groups in position 3 and 4. However, a hydrophobic groove interacts with the acyl chains in position 6 of one glucose residues (Fig.28).

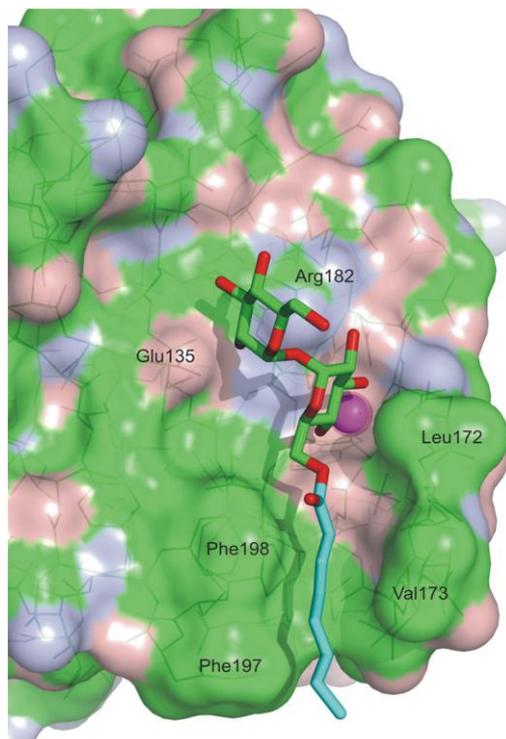


Fig.28 Portion of Mincle structure. Model of the trehalose octanoate conjugate. Adapted from [96].

The hydrophobic interaction is crucial for the strong binding. A solid phase binding competition assay showed that human Mincle binds threalose with 17-folder higher affinity than it binds glucose [95].

Mincle main targets are pathogenic glycans. It does not have high affinity binding towards the mammalian glycan array available at the CFG.

It is selectively associated with an ITAM containing adaptor and it is involved in IL-1 β induction via NF- κ B activation (Fig.29).

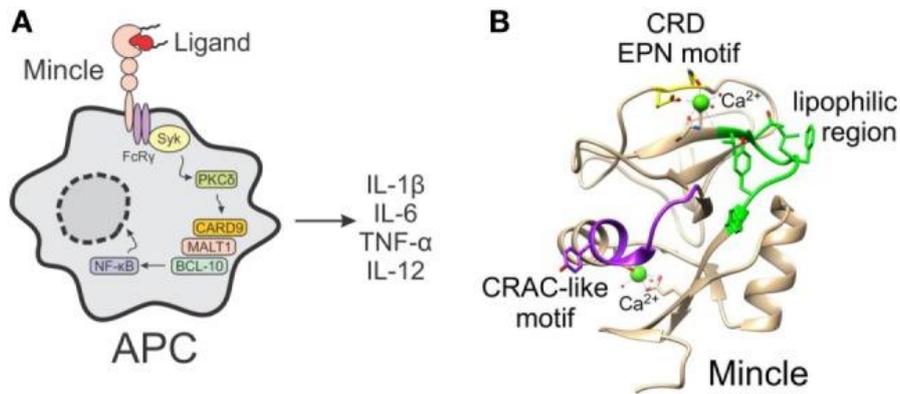


Fig.29 Mincle A) activation of Mincle on APCs. B) Crystal structure with the EPN motif indicated in yellow, the lipophilic region in green, and the cholesterol recognition/interaction amino acid consensus-like (CRAC) motif in purple [97].

CLR figure overview (Fig.30):

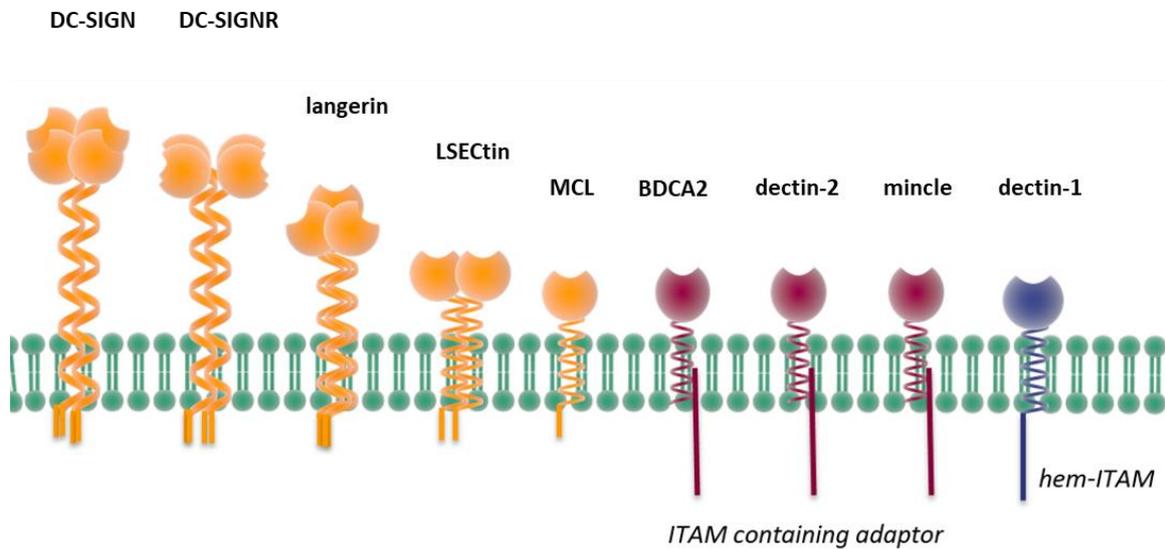


Fig.30 Graphical overview of the nine CLR investigated during my PhD. The CLR different oligomerisation is shown together with the cytoplasmic domain.

CLR table overview (Table.1):

<u>Name</u>	<u>PDB</u>	<u>Tissue/cell</u>	<u>Ligand</u>	<u>Intracellular Signalling Pathway</u>	<u>References</u>
BDCA2 (Blood Dendritic Cell Antigen 2)	3WBP	pDCs	Galβ1-3/4GlcNAcβ1-2Man	ITAM containing adaptor	<p>M. Nagae et al (2014) Crystal structure of CRD of BDCA2, Proteins, 82:1512-1518</p> <p>S. Jegouzo et al (2015) A novel mechanism for binding of galactose terminated glycans by the C-type carbohydrate recognition domain in BDCA2</p>
DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-Grabbing Nonintegrin)	1SL4	Dermis, mucosa, spleen and placenta	<p>Mannose</p> <p>Le^X</p> <p>Le^Y</p> <p>Le^A</p> <p>Le^B</p>	Without ITAM or ITIM domain	<p>(Geijtenbeek et al (2000) Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses, Cell, 100: 575-85).</p> <p>. B. H. Geijtenbeek et al (2001) DC-SIGN, a Dendritic Cell-Specific HIV-1 Receptor Present in Placenta That Infects T Cells In Trans—A Review, Placenta,22</p> <p>JJ. Garcia-Vallejo and Y. van Kooyk The physiological role of DC-SIGN: A tale of mice and men, Trends Immunol., 34:482-6)</p> <p>G. Tabarani et al (2009) DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain, J Biol Chem, 284: 1229-40</p>
DC-SIGNR (DC-SIGN related)/ LSIGN (Liver/Lymph node-Specific ICAM3)	1XAR	Endothelia cells of lymph nodes and liver, placenta	Mannose	Without ITAM or ITIM domain	<p>U.S. Khoo et al (2008) DC-SIGN and L-SIGN: the SIGNs for infection, J Mol Med, 86: 861-74</p> <p>R.C. da Silva et al (2011) Role of DC-</p>

Grabbing Nonintegrin)					<p>SIGN and L-SIGN receptors in HIV-1 vertical transmission, Hum Immunol, 74: 305-11</p> <p>H. Feinberg et al (2001) Structural basis for selective recognition of oligosaccharides by DC-SIG and DC-SIGNR, Science, 294: 2163</p> <p>Y. Guo et al (2017) Dissecting multivalent lectin-carbohydrate recognition using polyvalent multifunctional glycan-quantum dots, J Am Chem Soc, 139:11833-11844</p>
Dectin-1 (DC-associated C-type lectin 1)/ Beta-Glucan Receptor	2CL8 (murine)	Peripheral blood leukocytes and DCs in muscle, stomach and lung.	B-glucan (e.g. zymosan and laminarin)	hemi-ITAM	<p>J. Brown et al (2007) structure of the fungal beta glucan binding immune receptor decti-1 : implications for function, Protein Science, 16 : 1042-1052</p> <p>B.N. Gantner et al (2004) Collaborative induction of inflammatory responses by Dectin-1 and Toll-like Receptor 2, J Exp Med 179 : 1107-17</p>
Dectin-2 (DC-associated C-type lectin 2)	5VYB	Macrophages and dendritic cells in lung, spleen, lymph node and tonsils	α -mannan (Man α 1-2Man)	ITAM containing adaptor	<p>L. Graham and D. Brown (2009) The Dectin2 family of C type lectins in immunity and homeostasis, Cytokine, 48: 148-155</p> <p>H. Feinberg et al (2017) mechanism of pathogen recognition by human dectin 2, Journal of Biological chemistry 292 (32): 13402-13414</p>
Langerin	3P7F	LCs in epithelia (Birbeck granules)	Mannose, sulphated glycosaminoglycan (GAG)	Without ITAM or ITIM domain	<p>L. de Witte et al (2007) Langerin is a natural barrier to HIV-1 transmission by Langerhans cells, Nature medicine, 13: 387-371).</p> <p>M. Thepaut et al</p>

					<p>(2009) Structural studies of Langerin and Birbeck granule: a macromolecular organization model, Biochemistry, 48:2684-98</p> <p>J.C. Munoz-Garcia et al (2015) Langerin-Heparin interaction: two binding sites for small and large ligands as revealed by a combination of NMR spectroscopy and cross-linking mapping experiments, J. Am. Chem.Soc., 137:4100-4110</p>
<p>LSEctin (Liver and lymph node Sinusoidal Endothelial cell C-type lectin)</p>	—	<p>Liver and lymph node sinusoidal endothelial cells, isolated human peripheral blood and thymic DCs.</p>	GlcNAc β 1-2Man	<p>Without ITAM or ITIM domain</p>	<p>Powlesland et al (2008) A novel mechanism for LSEctin binding to Ebola virus surface glycoprotein through truncated glycans, Journal of biological chemistry, 283:593-602</p> <p>T. Gramber et al (2008) Interactions of LSEctin and DC-SIGN/DC-SIGNR with viral ligands: differential pH dependence, internalization and virion binding, Virology, 373:189-201)</p>
<p>MCL (Macrophage C-type Lectin)</p>	3WHD	<p>Monocytes and macrophages</p>	<p>Unknown (Trehalose-6,6'-dimycolate?)</p>	<p>ITAM containing adaptor</p>	<p>Furujawa et al (2013) Structural analysis for glycolipid recognition by C type lectins Mincle and MCL, PNAS, 110:</p> <p>Miyake et al (2013) C-type lectin MCL is an FcγR-coupled receptor that mediates the adjuvanticity of mycobacterial cord factor, Immunity, 38:1050-1062</p>

<p>Mincle (Macrophage inducible Ca²⁺-dependent (C-type) lectin)</p>	<p>3WH2</p>	<p>Macrophage, DCs, B cells, and neutrophils</p>	<p>Cholesterol crystals Glycosyl diacylglycerols, e.g-trehalose-6,6'-dimycolate</p>	<p>ITAM containing adaptor</p>	<p>S.J. Williams (2017) Sensing lipids with Mincle : structure and function, Frontiers in Immunology, 8 :1662</p> <p>Rambaruth et al (2015) Mouse Mincle: characterization as a model for human Mincle and evolutionary implications, Molecules 20:6670-6682</p>
---	-------------	--	---	--------------------------------	--

Table.1 Overview of the nine CLRs studied during my PhD. The following information are summarised: PDB number, expression, ligand recognition, x, signalisation motif and relevant literature.

3. Glycobiology

3.1 Introduction

Protein-carbohydrate interactions serve multiple functions in the immune system, as anticipated in chapter 2. This new chapter will focus on the world of glycans.

Glycans are essential in cell-cell interactions and almost all cell-surface and secreted proteins are glycoproteins [98]. Gabriel A. Rabinovich qualified carbohydrates as “cousins” of nucleic acid, since DNA and RNA are essentially composed of polysaccharides, and “roommates” for proteins [99] since the glycosylation represents the major class of post-translational modifications. Glycosylation, which is the formation of a glycosidic linkage between the reducing end of β -GlcNAc chain and asparagine (N-linked) or the hydroxyl group of serine/threonine (O-linked), dramatically enhances the functional diversity of protein (Fig.31).

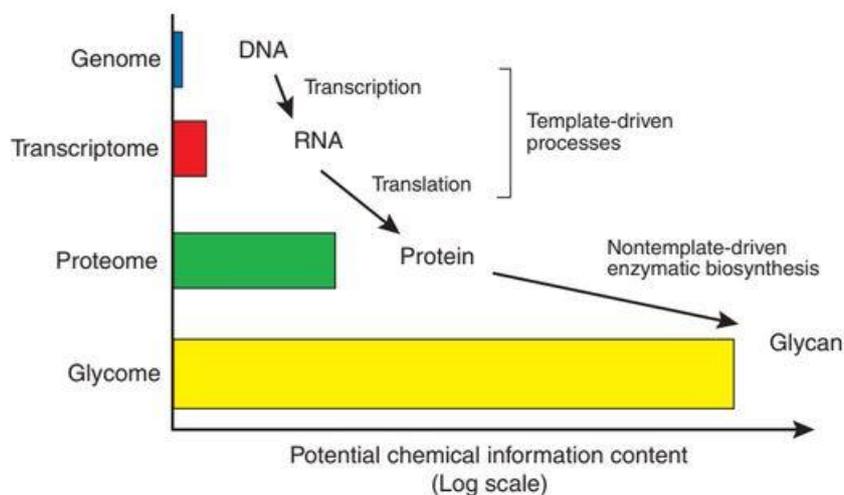


Fig.31 Potential information content of DNA, RNA, protein and glycan. Protein expression is based on a genetically encoded template but post-translational modifications dramatically enhance their functional diversity. The biosynthesis of glycans is not encoded via a template-driven system. Adapted from [100].

While DNA and peptide are linear molecules with no branches, glycans are incredibly complex and this peculiar complexity comes from the necessity of a multitude of signals at the cellular level. For DNA, there are 4096 possible ways to build a 6-mer, for protein 64 million, for sugars 193 million possible configuration. This complexity explains the 35 years delay between the development of tools for synthesizing sugars compared to those for DNA and peptides [100]. Even more intriguing is the ability of sugars to cover an immense diversity of biological functions with a non-template synthesis, as carbohydrates are not encoded by the genome. The genome codes, *de facto*, for enzymes like

glycosyltransferases and glycosidases that will subsequently determine the glycosylation patterns of glycolipids and glycoprotein [101].

N-glycosylation is found in all domains of life with some differences and three different types of N-glycans can be found in eukaryotes (Fig.32): high Mannose, Complex and Hybrid glycans. They share a common core structure that includes the first two N-acetylglucosamine residues and the first three mannose residues. [102] [103]

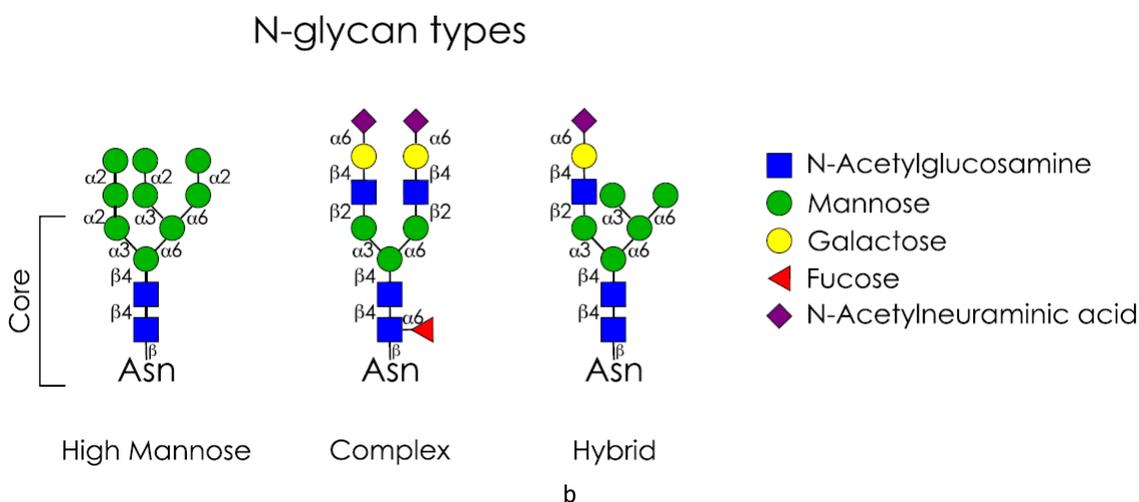


Fig.32 N-glycans. Three types of N-glycans

There are several reasons for glycan complexity. Glycans can have α or β configurations, they can form bond *via* different linkage positions (1 \rightarrow 1,2,3,4,6 for hexopyranose) (Fig.32), the ring size could be pyranosidic or furanosidic. They are usually branched to have their termini accessible in a high local density surface. Moreover, some sites undertake modifications such as acetylation or phosphorylation, for instance. The number of possible linkages points in Fig. 33 clearly underlines the high complexity that results from carbohydrate oligomer formation compared to nucleotides and amino acids [104].

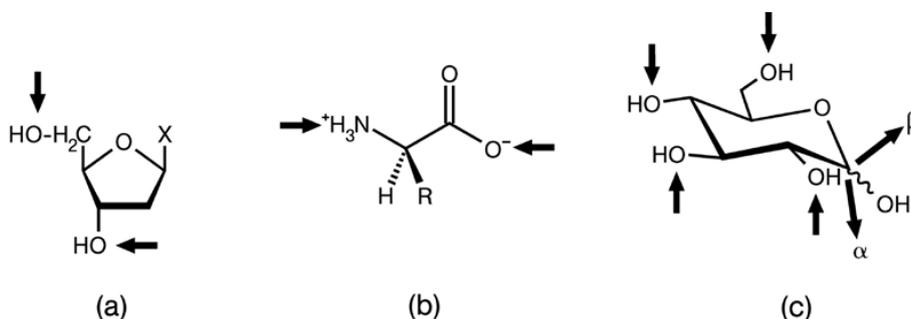


Fig.33 Linkage points for oligomer formation from a) nucleotides, b) amino acids and c) carbohydrates, adapted from [104].

Carbohydrate hydroxyl groups are used for the coordination of Ca^{2+} ions and they participate in H-bond formation, while C-H bonds are involved in van der Waals interactions or π -interactions with Trp or Tyr. The position of one or two hydroxyl groups can be crucial for bio recognition. When considering the example of

mannose and galactose in CLR interactions, the axial or equatorial hydroxyl group in position 4 conditions the recognition by the EPN or QPD motif of the lectin. Mannose binding protein C-type lectin, for example, involves a Ca^{2+} ion to probe the presence of the equatorial 4-hydroxyl group together with two H bonds [105], as already described in chapter 2.2. (Fig.34)

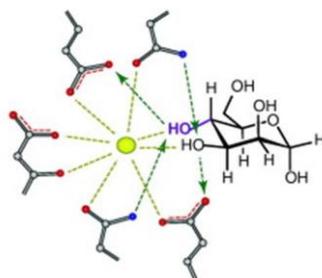


Fig.34 Example of calcium ion coordination. Glycan binding to a mannose binding protein. Adapted from [105].

Glycomics, with the analysis of the structure and function of glycans, gave to Glycobiology a dimension in biomedicine [100]. Although the chemico-physical properties of simple carbohydrates are known, the same thing does not apply to complex ones. In 2001 the Consortium for Functional Glycomics (CFG, <http://www.functionalglycomics.org/>) was created by James Paulson to help the investigation of carbohydrate roles in biological systems. The CFG is a fundamental resource of glycan microarrays with covalently attached hundreds of different carbohydrates.

A discussion about carbohydrate necessarily involves references to lectins; while sugars are able to carry the biological information, lectins can read and decipher their glycode [106]. Glycobiology is a relatively new field and methodologies to study the glycans, the lectins and their binding properties are currently evolving. The following figure gives an overview of different technologies exploited to study the glycan world (Fig.35).

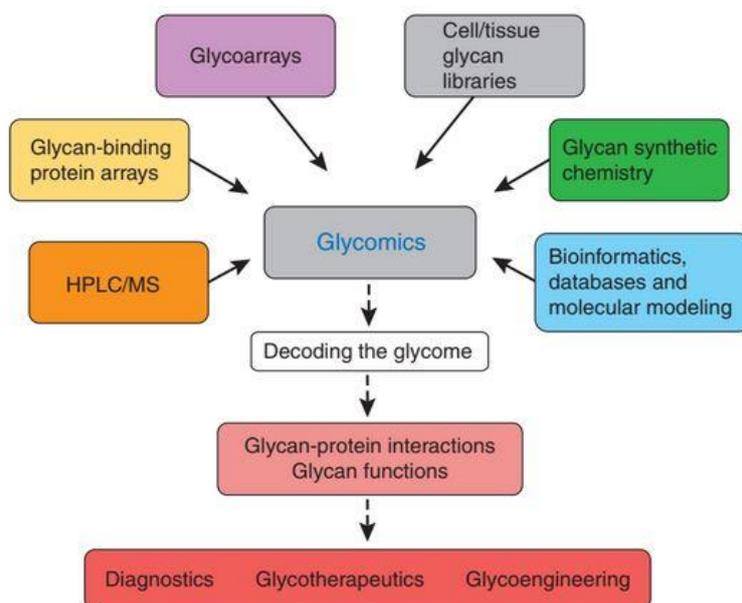


Fig.35 Cracking the glycode. Emerging technologies to explore structure-function relationships of the glycome. Adapted from [100].

The very first bottleneck in this field resides in the glycan synthesis and analysis. Often, multistep synthesis is required and starts from the conversion of an available compound into the desired product using known reactions. Protective groups are used to drive the intermediate formation towards the suitable final product. Nevertheless, other undesired intermediate compounds could form, leading to low yield after purification step. Moreover, the manipulation of protective group itself is difficult and often requires multiple steps. The most advanced synthetic approach is an automated solid-phase synthesis strategy for glycan assembly, which can create 50-mers within days [107]. However, this approach is far from being commonly used by all the laboratories and, moreover, the production of each single building block required for the automated synthesis demands specific optimization.

For more detailed information about technologies exploited during my PhD to study the relationship between glycan and CLRs, please refer to chapter 6 and 7.

3.2 Design of Mimetic

Besides their complexity and the difficulty of their synthesis, carbohydrates are characterized by low affinity interaction and overlapping specificity for their proteic partners. Natural sugars could bind to multiple lectins and are thus inappropriate to specifically target one single CLR. Moreover, their pharmacokinetic properties do not encourage the use of glycans as drugs: carbohydrates, in fact, are rapidly digested or they cannot passively diffuse through the intestine.

To overcome all the above-mentioned difficulties, compounds able to selectively target carbohydrate binding sites with drug-like properties have to be developed [108].

3.2.1 Glycomimetic

One approach to reach that goal is to synthesise compounds mimicking the natural sugar. Such glycomimetics are non-carbohydrate players that could be used as alternative because they often retain the geometry of natural molecules and modulate their biological activity. In the context of CLRs, the approach has been rationalized by defining three sections that should be taken into consideration during the glycomimetic design (Fig.36) [108].

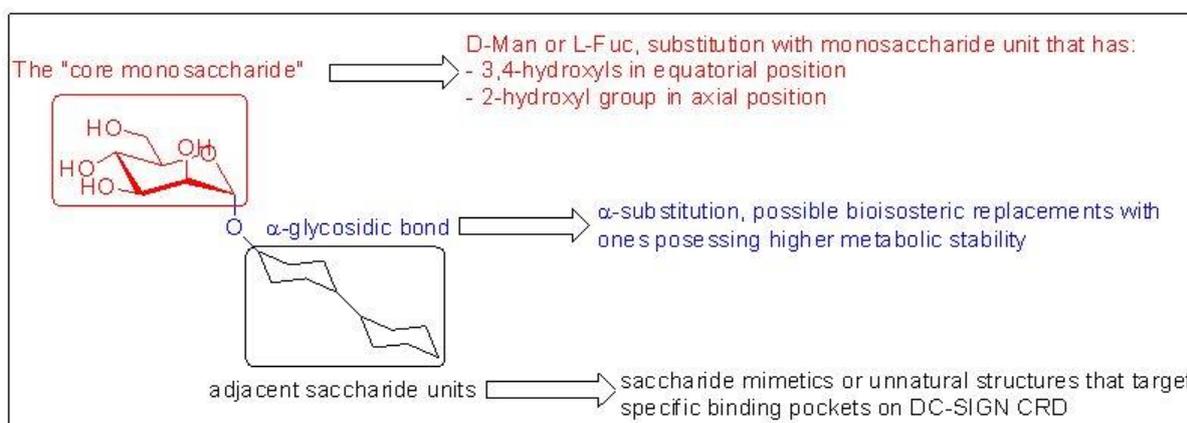


Fig.36 Schematic glycomimetic design. Schematic representation of the three different section in the glycomimetic design for DC-SIGN. Adapted from [108].

This schematic representation applicable to different CLRs was particularly exploited for DC-SIGN. DC-SIGN was initially identified as a receptor of mannose in HIV infection [109], and the epitope necessary for the ligand recognition is Di-Mannose. Example of the "core monosaccharide" modification was recently corroborated by a collaboration between our group and the group of Anna Bernardi (*Fieschi et al*, unpublished), with a modification of the "core" mannose in position C-2 in order to reach Phe313 of DC-SIGN. Details about the project will be provided in chapter 8.3.

Glycosidic bonds are subject to glycosidase hydrolysis and effort should be directed towards other surrogated bonds, which is the second point addressed in the above scheme. Recently, a collaboration between our group and the group of Anna Bernardi [110] showed that a pseudo-1,2-thiol-mannobioside (thiol-psDi) has the same binding affinity for DC-SIGN than pseudo-mannobioside (psDi) but with enhanced stability towards glycosidases (Fig.37a and b). A better analysis will be given in chapter 8.3.

Last but not least, adjacent monosaccharide units also influence the binding specificity. Recently, the question regarding the overlapping affinity between DC-SIGN and langerin has been addressed and a

rational design approach was used to selectively target DC-SIGN in disfavour of langerin [111]. Figure (Fig.37c) shows the structure of 6-amino-Man030 compound for which affinity and specificity enhancement towards DC-SIGN has been achieved by reaching a secondary binding site.

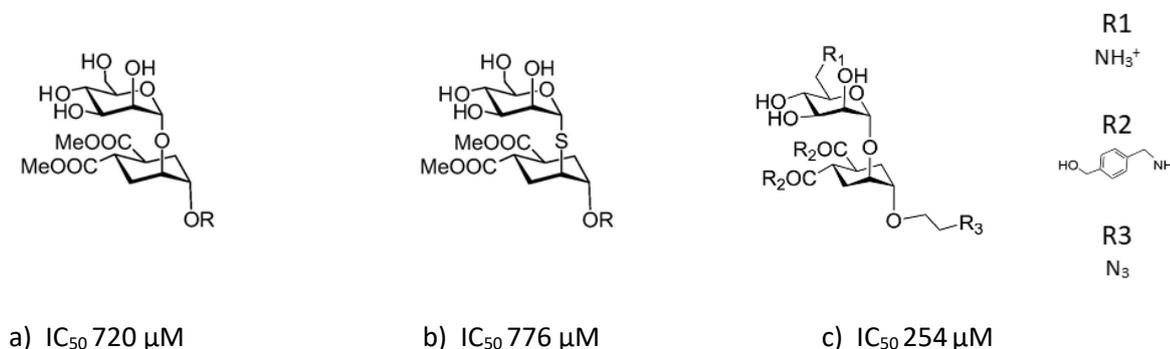


Fig.37 Glycomimetics against DC-SIGN developed by Anna Bernardi research group (a) psDi, b) thiol-PsDi and c) 6-amino-Man030. Adapted from [112], [110] and [111]. R=H

The design of antagonists against DC-SIGN is engaging and the group of Rademacher brought its contribution by targeting lectin secondary druggable binding sites [39]. Efforts start to be made also on the development of mimics towards other lectins. For example, Rademacher group [113] studied ligand design for langerin. By combining in silico studies and ^{19}F R2-filtered NMR, a 2-deoxy-2-carboxamido- α -mannoside analogue was identified with a K_d of 4.3 mM, a 5-fold affinity increase over natural monosaccharide ManNAc (22 mM) (Fig.38).

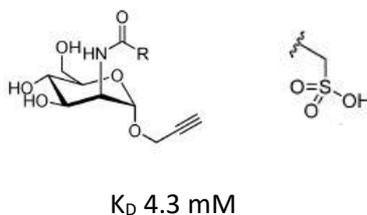


Fig.38 Glycomimetic against langerin. 2-deoxy-2-carboxamido- α -mannoside analogue against langerin, from [113].

3.2.2 Non-glycomimetics

The development of compounds that mimic carbohydrate has not been the only investigated approach. The research group led by Laura Kiessling contributed to the development of glyco and especially non-glycomimetics against DC-SIGN. By high-throughput fluorescence-based competition assay from two commercial libraries, 36000 non glycocompounds were screened against DC-SIGN as antagonist and one of them showed an inhibitory power of 1.6 ± 0.5 μ M [114] (Fig.39a). Two series of DC-SIGN inhibitors were then

developed: a shikimic acid-derived glycomimetic scaffold series (Fig.39b), with an IC_{50} of 3.2 ± 0.6 mM (N-acetylmannosamine $IC_{50} = 11.2 \pm 0.7$ mM) [115] and a quinoxalinones based collection that allowed to reach an IC_{50} of $0.31 \mu\text{M} \pm 0.13$ (Fig.39c) [116].

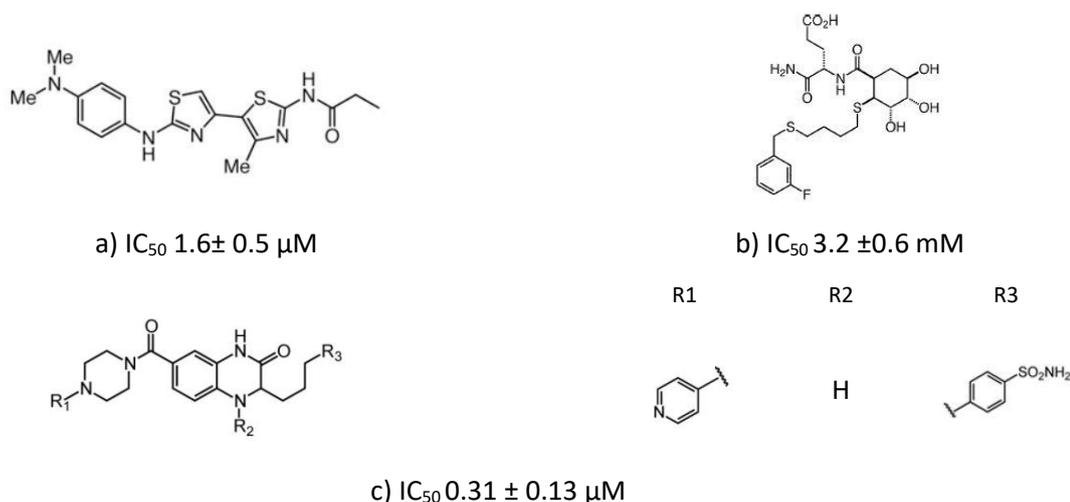


Fig.39 Antagonists developed in Laura Kiessling group a) b) shikimic acid derivative c) quinoxalinone derivative. Adapted from [114],[115] and [116].

The synthesis of mimetic compounds targeting CLRs has to be validated by functional *in vitro* biochemical tests. Screening techniques are used to assess whether the un-natural carbohydrate derivatives interact with lectins. The following paragraphs will give an overview on two main screening techniques exploited during my PhD: lectin microarray and glycan microarray. Of course, the putative candidates will subsequently be analysed with biophysical technique, such as Surface Plasmon Resonance and Isothermal Titration Calorimetry (chapter 6), to quantify and characterize the interaction with the lectin. The following paragraphs will give an overview on two main screening techniques exploited during my PhD: lectin microarray and glycan microarray.

3.3 Screening technique

Libraries of glycans and glycomimetics can be analysed by analytical tools such as lectin microarrays and glycomimetic/glycan arrays. In both techniques, series of molecules are attached onto a supporting material and used as a platform for biological sample screening [28].

3.3.1 Lectin microarray

Lectin microarrays are a fast tool to monitor carbohydrate and glycomimetic hypothetical interaction with the target lectin. The basic format consists in the immobilization of the lectin onto a surface (chip or plate) and the binding partner is added in solution. Lectins either adsorb on the surface or are covalently attached to it ; alternatively, immobilization can be achieved via a biotin moiety. The detection of the interaction can be direct or indirect. The direct one implies that the carbohydrate partner is labelled with a fluorophore, for example, and the fluorescence is measured immediately after binding. For indirect revelation, on the other hand, the partner is not labelled but is, for instance, recognized by a fluorescent secondary antibody. The revelation, in this second situation, occurs in two different steps [106].

During my PhD I had the possibility to work with a French company located in Orléans called GLYcoDiag (<http://www.glycodiag.com/>). They have developed LectPROFILE Plates to detect the interaction between adsorbed lectins and glycoproteins in solution by indirect revelation. The detailed technique is explained in chapters 6 and 7.

3.3.2 Glycan/glycomimetic microarrays

Glycan microarrays are based on the glycan attachment onto a surface. The glycan/glycomimetic can be a) covalently immobilized on gold or glass b) absorbed on nitrocellulose c) absorbed on polystyrenes through hydrophobic interactions d) immobilized by a biotin moiety. Again, fluorescence is used for detection using lectins labelled prior screening. The selection of the surface and the method of functionalization can be crucial for successful detection. The following figure contains a schematic representation of the necessary steps for glycoprofiling (Fig.40). [106]

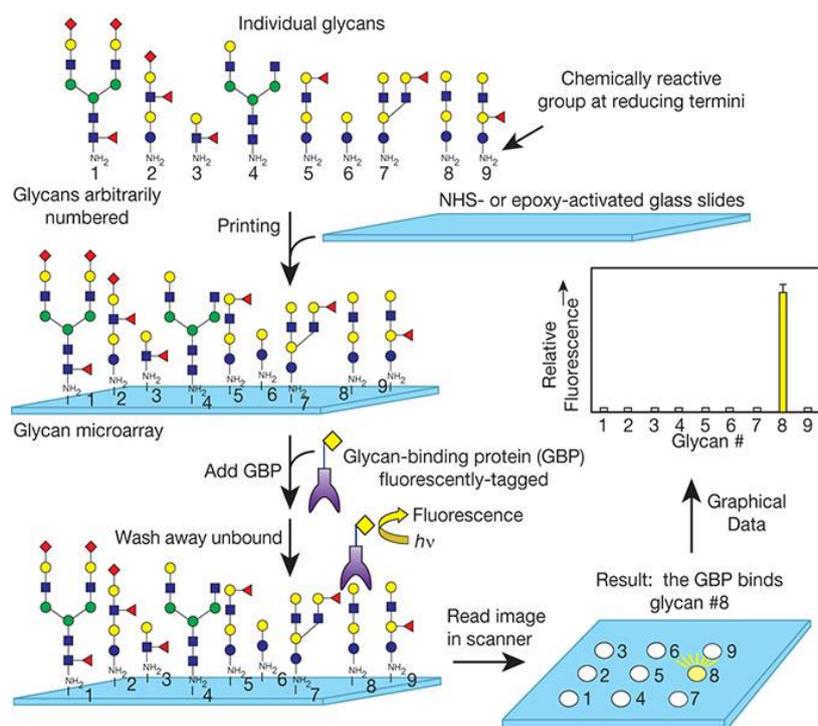


Fig.40 Glycan array. Schematic representation of the glycan array technique, adapted from [117].

During my PhD, I had the possibility to work with Niels Reichardt laboratory located in Donostia-San Sebastian (Spain) (<https://glycotechnology.net/>) on their glycan microarray platform. Their platform is used to detect the interaction between immobilized glycans/glycomimetic and fluorescently labelled lectins. The detailed technique is explained in sections 6 and 7.

3.3.2.1 TETRALEC strategy

We have already mentioned that CLR have low affinity for their saccharidic partners and that feature can be problematic for ligand screening. Besides, for some lectins, e.g. MCL, the entire protein interacts with the endogenous ligand, while the binding of the single CRD to the isolated glycan is undetectable by screening technique (chapter 2.3.7).

To encompass this difficulty, the group of Drickamer [88] developed the first example of artificial CLR multivalency to enhance binding affinity. To study LSEctin binding to a panel of ligands, they created an artificial tetrameric protein exploiting the strong biotin /tetrameric streptavidin interaction. They produced recombinant LSEctin directly functionalized with biotin that formed with streptavidin a tetramer of CLRs, enabling multivalency at protein level.

With the same purpose, we developed an artificial oligomeric protein displaying four CRDs in order to enable in vitro ligand binding biochemical assay. This multivalent tetramer is obtained by a site specific

biotin labeling of the CRD with the use of the bacterial enzyme Sortase A (SrtA), following the method improved by H. Antos [118]. In the Gram-positive *Staphylococcus aureus*, the SrtA cleaves the LPXTG motif present in target proteins and links them to the amino terminal group of five glycine of the peptidoglycan [119]. The enzymatic activity of SrtA has been widely used to link several compounds, protein to protein [120] or even PEG to surface [121]. Here, SrtA is used to enzymatically couple the N-Terminus of the protein to a biotinylated peptide. The resulting biotinylated CRD is then complexed to NeutrAvidin to obtain a final molecule exposing four glycan binding sites, named hereafter TETRALEC.

Technical details about the strategy and some preliminary results will be given in chapter 8.1.

3.4 Multivalent Ligands: Targeting CLRs

Optimization of monovalent antagonists is the first step on the long road towards the development of inhibitors. Monosaccharides or small oligosaccharides in isolation, as already mentioned, tend to be low-affinity ligands for lectins, often with dissociation constants in the millimolar range. Nature deals with the « low affinity » issue by exploiting multiple binding events. The accumulation of weak affinity leads to an apparent strong interaction, an effect called avidity. While affinity refers to the direct interaction of a single CRD with a monovalent ligand, avidity refers to the overall strength of multivalent interactions [117]. Multivalent binding plays a crucial role in the cell-surface recognition. Such multivalent binding could be achieved by different means and, in order to facilitate it and accommodate different targets, CLR glycan-binding sites have to cluster. The result is either a clustering of single CLR in micro domains, or multiple CRDs in a single polypeptide chain or polypeptide oligomers each containing a single CRD (Fig.41) [122].

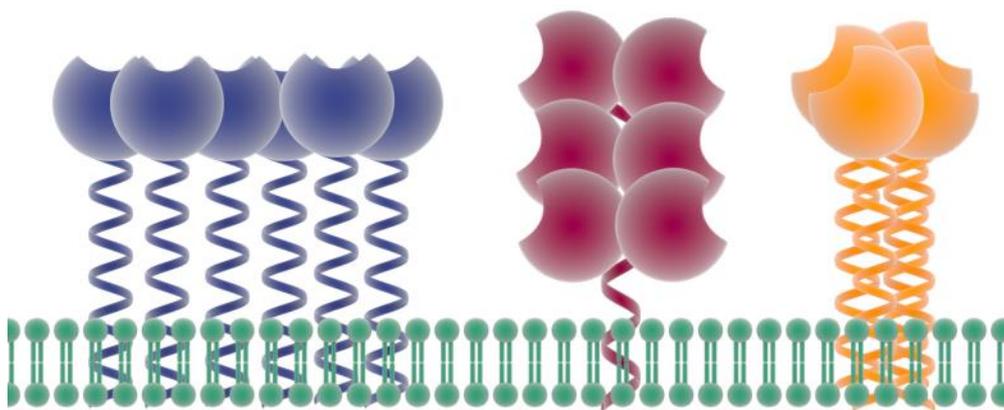


Fig.41 Multivalency. Single CLR in micro domains (blues), or multiple CRDs in a single polypeptide chain (red) or polypeptide oligomers each containing a single CRD (orange).

Multivalent ligands participate as well to high-avidity binding and contribute to the “glycan cluster effect” [123] or « velcro effect » that occurs when multivalency is reached on both protein and glycan side.

There are four main mechanisms by which multivalent ligands can interact with their receptors (Fig.42). Oligomeric CLRs could have their binding sites occupied by multiple binding elements (chelate effect) (Fig.42a), while monomeric receptors could clusterize to accommodate multivalent partners (clustering binding) (Fig.42b). As mentioned in chapter 2.2, CLRs could also possess secondary binding sites exploitable for binding (Fig.42c). Finally, the rebinding or statistical association phenomena (Fig.42d) occurs when multivalent ligands could display high epitopes concentrations leading to an increasing of apparent binding affinity [124].

Physiological multivalent ligands are too structurally heterogeneous or complex to be used in the identification of relevant molecular mechanism. Therefore, multivalent synthetic glycans or glycomimetics/non glycomimetics are created to identify high affinity ligand. The ultimate ligand should combine high monovalent affinity ligand with an appropriate multivalent presentation. Chemists can play around different parameters of the multivalent compounds: the scaffold structures, the nature and the number of binding elements and the density of binding elements. The scaffold, or main core, carries linkers or spacers to the terminal binding units. Multivalent ligands include both low valency compounds, such as short polymers, glycoclusters or peptide conjugates and high valency compounds, such as dendrimers, liposomes or nanoparticles (Fig43). The glycoclusters, hence, could expose several but controlled number of sugars/derivatives and the different way of presentation could help to rationalise the best presentation mode.

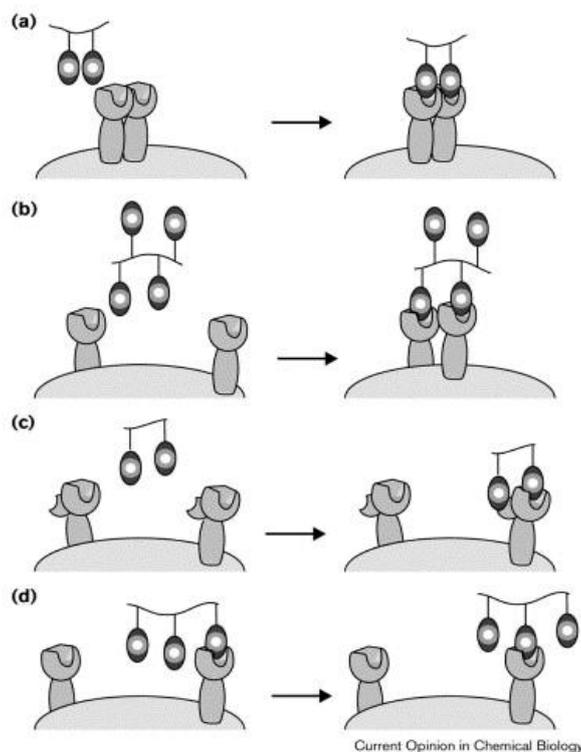


Fig.42 Binding mode. Chelate binding (a), clustering binding (b), secondary binding (c) and rebinding (d). Adapted from [124].

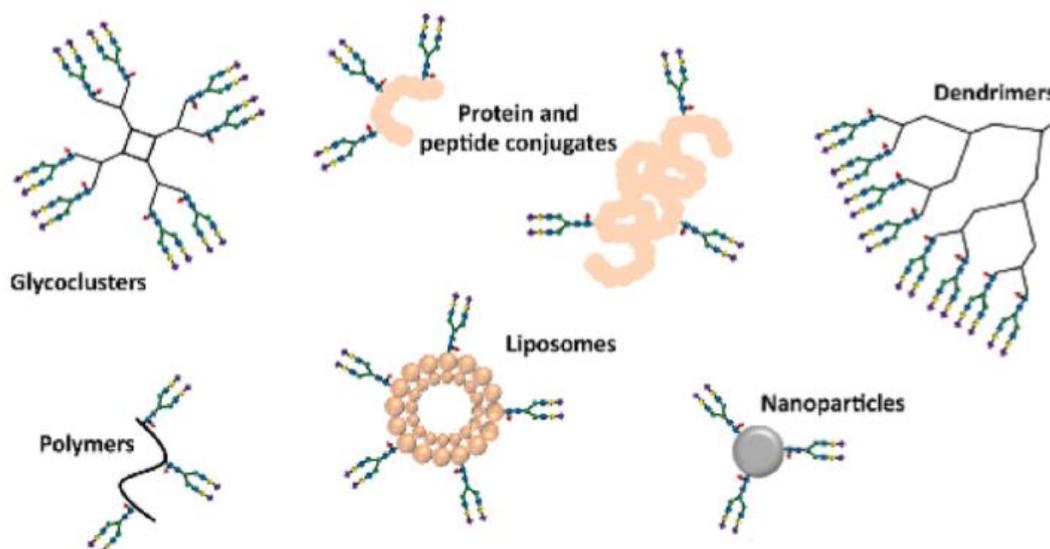


Fig.43 Multivalent ligands: glycoclusters, protein and peptide conjugates, dendrimers, polymers, liposomes and nanoparticles. Adapted from [125].

Again, DC-SIGN was an attractive target and there are several examples of multivalent ligands that have been optimized towards this CLR. Firstly, **dendrimers** have been extensively exploited for CLR targeting purpose since they are non-toxic macromolecules with defined globular shape. The very first polyvalent DC-SIGN antagonist is a polydisperse Boltor-type dendritic polyester platform functionalized with 16/32 copies of mannose (Fig.44a). The dendrimer bearing 32 mannose units inhibit gp120 binding with an IC_{50} in the μM range [126]. The group of Anna Bernardi has recently developed a rod-like dendrimer named Polyman 26 (PM26) that was shown to inhibit DC-SIGN-mediated HIV infection in nanomolar concentrations and to be internalized by DCs into the endolysosomal compartment [127](Fig.44b). **Neo-glycoproteins** and **glycopolymers** can also be used as multivalent compound against DC-SIGN. For example, BSA bearing 30 mannoses was showed to have a KD of 31 ± 7 nM when interacting with DC-SIGN [64] while star-shaped glycopolymers have affinities in the picomolar range [128].

Glycofullerenes compounds, which facilitates spherical presentation of glyco-based epitopes, were synthesized by the group of Javier Rojo to block DC-SIGN interaction with Ebola pseudo type viral particles. (Fig.44c) [129]. Finally, the world at nano scale was as well explore, intrigued by the possibility to mimic the natural saccharidic presentation coupled to the peculiar optical properties of the metallic core. **Gold nanoparticle** (AuNPs) are an outstanding example of nanosystem that could be used in a biological context and, for instance, *D.Arosio et al* [130] functionalized AuNPs with α -fucosyl- β -alanyl amide in order to compete with HIV glycoprotein gp120 on DC-SIGN expressing cells (Fig.44d).

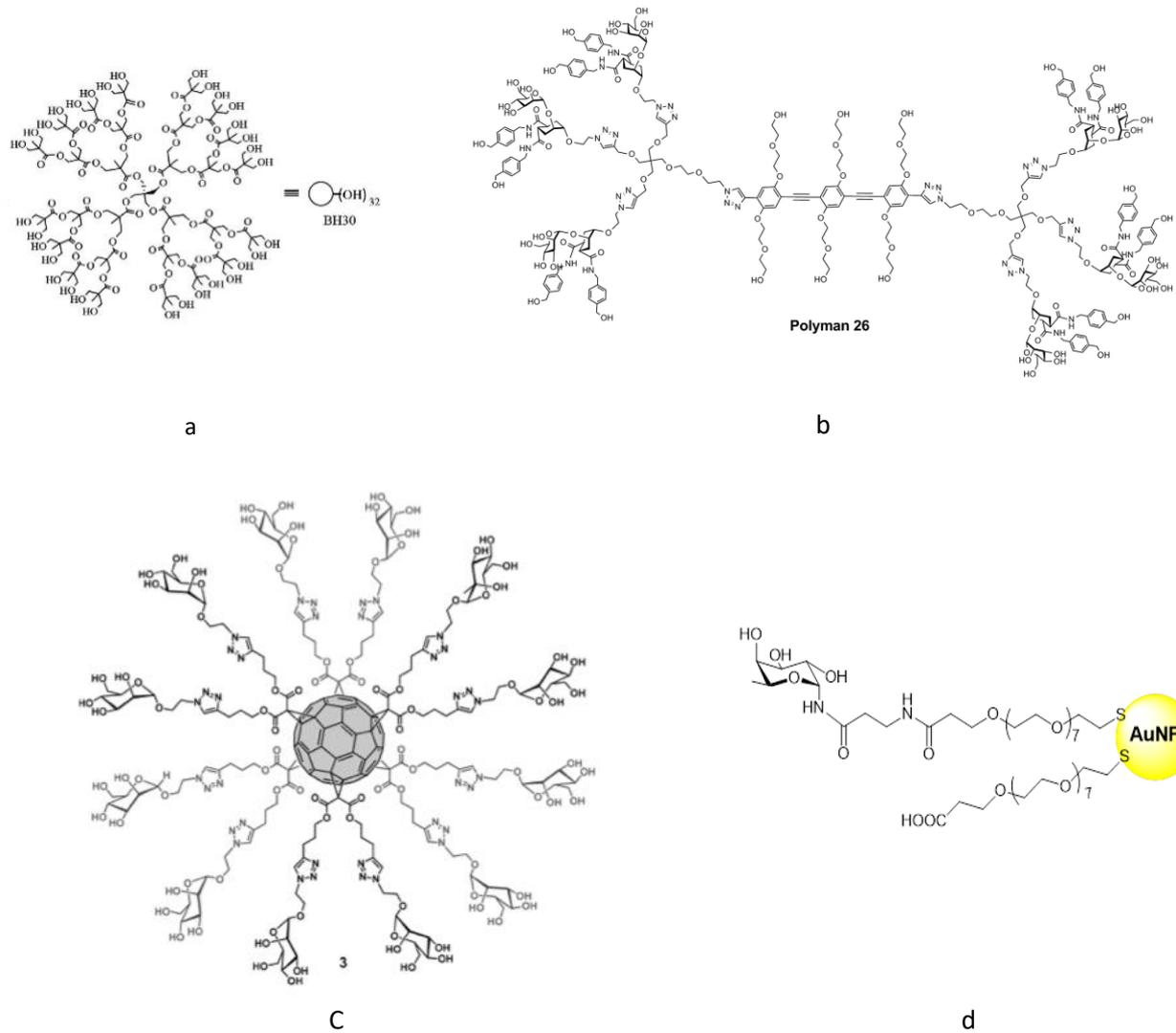


Fig.44 Examples of multivalent inhibitors against DC-SIGN a) Boltom-type b) rod-like dendrimer PM26 c) Le^b-conjugated PAMAM d) AuNPs functionalized with α -fucosyl- β -alanyl amide.

Multivalent ligands can be synthesized for two different purposes: to act as inhibitors or effectors. The inhibitors interfere with receptor-ligand binding, as above showed for DC-SIGN, while effectors could help to understand and manipulate a cellular response. For instance, a glycocluster carbohydrate-based vaccine against *Streptococcus pneumoniae* was created by coating gold nanoparticle with the repeating units of the bacterial capsular polysaccharide together with a T-helper peptide to evoke specific immune response (Fig.45) [131].

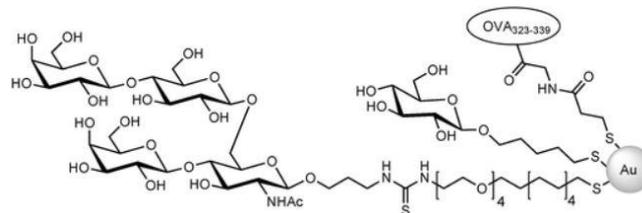


Fig.45 Multivalent ligand as vaccine. *Streptococcus pneumoniae* carbohydrate-based vaccine. Adapted from [131]

A lot of the effort has been made towards the study of glyco-conjugates for possible *in vitro* and *in vivo* applications. The next chapter will show some examples of practical use of carbohydrate-based constructs.

4. Applications

In the previous chapter, the complexity of carbohydrates and some strategies exploited to study and enhance their interaction with CLRs were discussed. This part is meant to highlight how protein-carbohydrate interactions have medical relevance for the treatment of immune - and inflammation-mediated diseases. A focus will be made on the use of lectin-glycan interactions as tools for diagnosis, imaging and targeting (Fig.46).

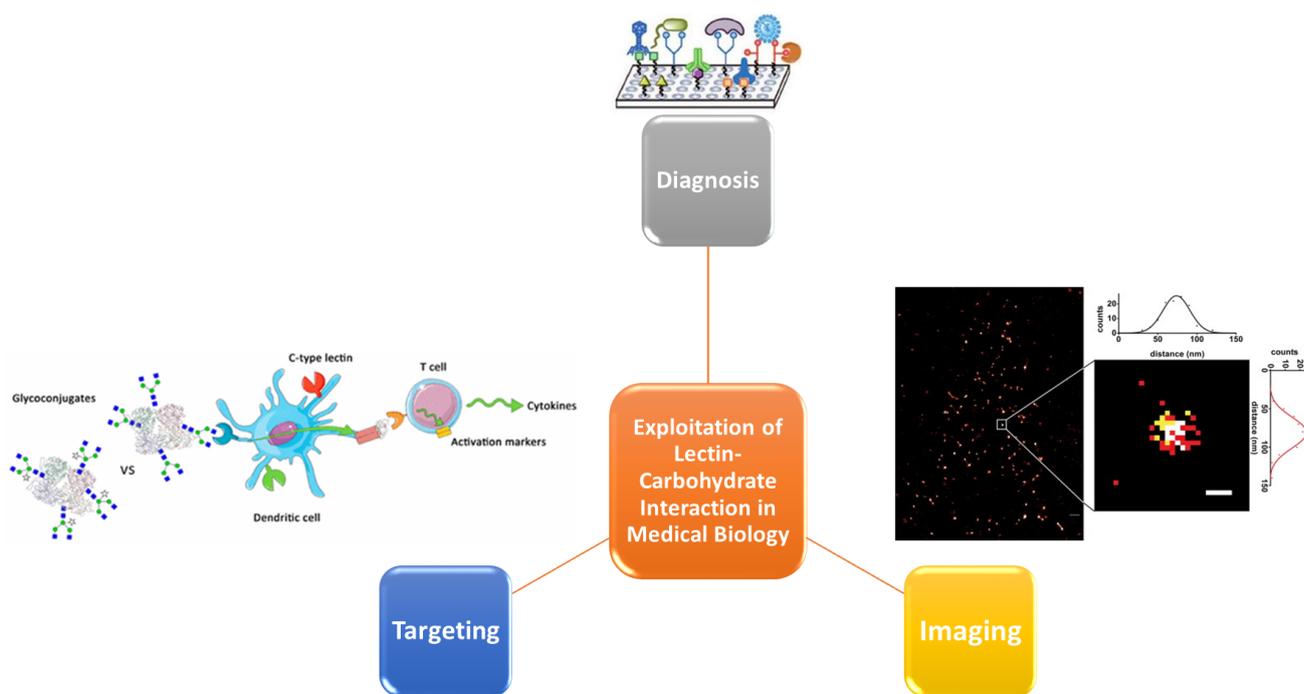


Fig.46 Exploitation of lectin-carbohydrate interaction in medical biology. Flowchart of possible applications of protein-carbohydrate interactions. Adapted from [132], [133], [134].

4.1 Diagnosis

As already commented in section 3, glycosylation is the most common post-translational modification and glycans are involved in multiple important biological functions. Several diseases are characterized by changes in the cell glycosylation pattern, features that can be exploited in the context of diagnosis. Being

able to diagnose disease at early stage of their development constitutes an important challenge, as it could be crucial for the success of treatment. With that goal, lectins can be used to detect glycan modifications [135], as shown in Table 2 that lists lectins used in cancer biomarker research:

Lectin	Abbreviation	Specificity	Glycan linkage	References
African legume (<i>Griffonia (Bandeiraea) simplicifolia</i>) lectin-I	GSLI (BSLI)	α -Gal; α -GalNAc	O-linked	<i>Lescar et al. (2002)</i>
Asparagus pea (<i>Lotus tetragonolobus</i>) lectin	LTL	Fuc α 1-3(Gal β 1-4)GlcNAc, Fuc α 1-2Gal β 1-4GlcNAc	N-linked	<i>Pereira & Kabat (1974), Yan et al. (1997)</i>
Koji (<i>Aspergillus oryzae</i>) lectin	AOL	α 1,6-fucosylated	N-linked	<i>Matsumura et al. (2007)</i>
Castorbean (<i>Ricinus communis</i>) agglutinin	RCA	Gal β 1-4GlcNAc; terminal β -D-Gal	N-linked	<i>Harley & Beevers (1986), Wang et al. (2011)</i>
Champedak (<i>Artocarpus integer</i>) galactose binding lectin	CGB	Gal; GalNAc	O-linked	<i>Hashim et al. (1991), Gabrielsen et al. (2014)</i>
Champedak (<i>Artocarpus integer</i>) mannose binding lectin	CMB	Man	N-linked	<i>Lim, Chua & Hashim (1997), Gabrielsen et al. (2014)</i>
Daffodil (<i>Narcissus pseudonarcissus</i>) lectin	NPL	α -Man, prefers polymannose structures containing α -1,6 linkages	N-linked	<i>Kaku et al. (1990), Lopez et al. (2002)</i>
Elderberry (<i>Sambucus nigra</i>) agglutinin	SNA	Neu5Ac α 2-6Gal(NAc)-R	N- and O-linked	<i>Shibuya et al. (1987), Silva, Gomes & Garcia (2017)</i>
Gorse or furze (<i>Ulex europaeus</i>) seed agglutinin-I	UEA-I	Fuc α 1-2Gal-R	N- and O-linked	<i>Holthofer et al. (1982), Rudrappan & Veeran (2016)</i>
Jackbean (<i>Canavalia ensiformis</i>) lectin	ConA	α -Man; α -Glc	N-linked	<i>Percin et al. (2012)</i>
Jackfruit (<i>Artocarpus heterophyllus</i>) lectin	Jacalin	Gal; GalNAc	O-linked	<i>Kabir (1995), Jagtap & Bapat (2010)</i>
Lentil (<i>Lens culinaris</i>) hemagglutinin	LcH	Man; Glc (Affinity enhanced with α -Fuc attached to N-acetylchitobiose)	N-linked	<i>Howard et al. (1971), Chan et al. (2015)</i>
Amur maackia (<i>Maackia amurensis</i>) lectin II	MAL II	Sia α 2-3Gal β 1-4GlcNAc; Sia α 2-3Gal β 1-3GalNAc	N- and O-linked	<i>Konami et al. (1994), Geisler & Jarvis (2011)</i>
Orange peel fungus (<i>Aleuria aurantia</i>) lectin	AAL	Fuc α 1-6GlcNAc; Fuc α 1-3LacNAc	N- and O-linked	<i>Hassan et al. (2015)</i>
Peanut (<i>Arachis hypogaea</i>) agglutinin	PNA	Gal β 1-3GalNAc; Gal	O-linked	<i>Chacko & Appukuttan (2001), Vijayan (2007)</i>
Chinese green dragon (<i>Pinellia pedatisecta</i>) agglutinin	PPA	Man	N-linked	<i>Li et al. (2014)</i>
Poke weed (<i>Phytolacca americana</i>) mitogen lectin	PWM	GlcNAc oligomers	N-linked	<i>Kino et al. (1995), Ahmad et al. (2009)</i>
Red kidney bean (<i>Phaseolus vulgaris</i>) lectin	PHA-L	Bisecting GlcNAc	N-linked	<i>Kaneda et al. (2002), Movafagh et al. (2013)</i>
Thorn-apple (<i>Datura stramonium</i>) lectin	DSL	(GlcNAc β 4) _n	N-linked	<i>Yamashita et al. (1987), Abbott et al. (2010)</i>
Wheat germ (<i>Triticum vulgaris</i>) agglutinin	WGA	GlcNAc β 1-4GlcNAc β 1-4GlcNAc; Neu5Ac	N-linked	<i>Nagata & Burger (1972), Parasuraman et al. (2014)</i>
White button mushroom (<i>Agaricus bisporus</i>) lectin	ABL	GalNAc; Gal β 1,3GalNAc (T antigen); sialyl-Gal β	O-linked	<i>Nakamura-Tsuruta et al. (2006), Hassan et al. (2015)</i>

Table.2 List of lectins used in cancer biomarker research. Adapted from [135].

Lectin arrays, as explained in section x, is one of the possible techniques for the detection of glycan-lectin interactions. For instance, an array of 45 different lectins was used to detect the binding towards glycoproteins from colorectal cancer cells [136]. In particular, one lectin from *Agaricus bisporus* was identified as a potential new predictive biomarker.

Another example of carbohydrate-protein interactions exploited for diagnosis concerns Rheumatoid arthritis (RA), a chronic autoimmune inflammatory disease in which most of the patients produces auto-antibodies directed against immunoglobulin G (IgG). Modification of IgG glycosylation has been associated with RA and a recent study by ultra-performance liquid chromatography showed that IgG glycan aberrancy

can be detected years before the beginning of the disease and considered as a novel risk factor for RA [137].

Glycan arrays have been considered in the context of diagnosis. They have been utilised to screen patient serum from blood to facilitate diagnosis and identification of bacterial and viral markers (Fig.47) [132].

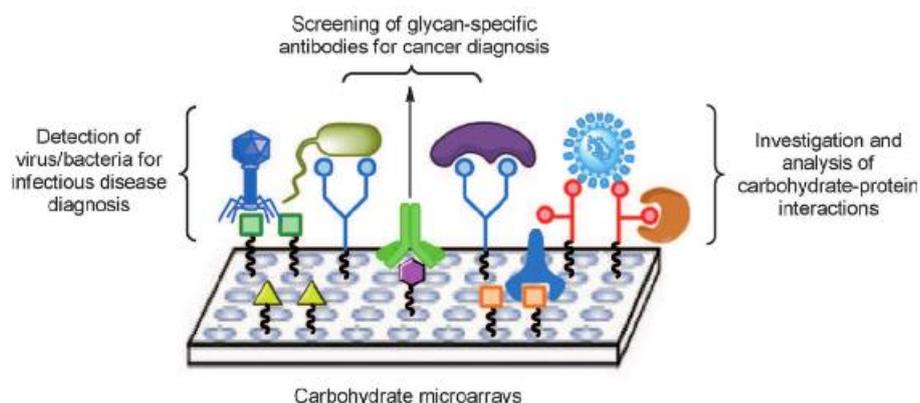


Fig.47 Application of glycan array. Glycan array is used to screen bacterial, viral and cancer markers. Adapted [132]

Microorganisms, such as *Burkholderia pseudomallei*, *Bacillus anthracis*, and *Francisella tularensis*, are indeed known to expose a wide variety of glycan-binding ligands on their cell surface, detectable by glycan array[138].

Glycan arrays, composed of pathogen derived carbohydrates, have been exploited as well to detect in the sera antibody against malaria [139] and hemagglutinin proteins from human influenza virus were studied on carbohydrate array to define their binding epitope [140].

4.2 Imaging

As mentioned in chapter “Glycobiology”, natural sugar mimics are developed to target, for example, specific CLRs expressed by APCs. Qualitative and quantitative characterization of the compound-lectin interaction is a prerequisite before moving towards *in vitro* and *in vivo* studies. When performing experiments at the cellular level, being able to visualise if the designed compound has been internalized in the cell informs about its efficacy. **Fluorescent conjugates** are useful probes to track the internalization by microscopy. *Zoran Arsov et al* [141] designed a probe against DC-SIGN containing a D-mannose moiety for CLR targeting and a pH-sensitive rhodamine dye (Fig.48).

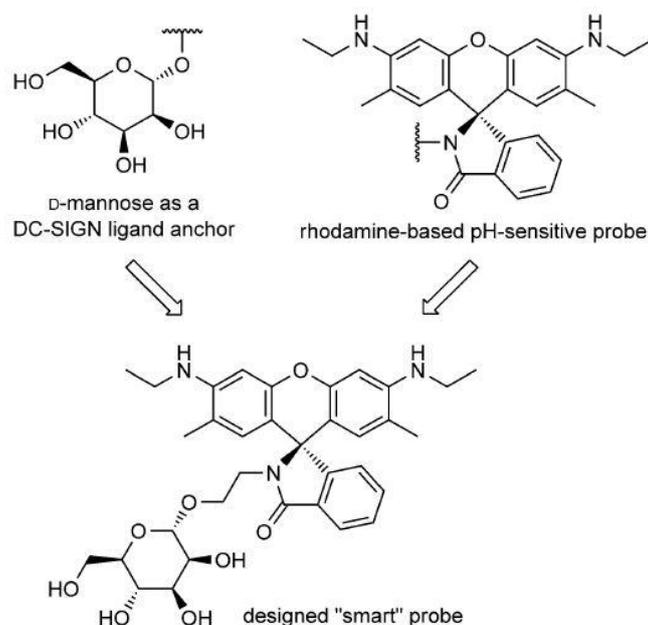


Fig.48 Example of smart probe. pH-sensitive “Smart” probe for DC-SIGN targeting. Adapted from [141].

This rhodamine-based probe becomes fluorescent in acidic conditions and, therefore, it enables after binding to DC-SIGN, to follow by microscopy internalization into endosomes.

The already mentioned dendrimer PM26 (3.4) is another example of Fluorescent conjugates. Its rod-like spacer is naturally fluorescent dendrimer. Confocal microscopy was used to assess its internalisation by DCs into the endolysosomal compartment. Figure 49 shows in red human immature Monocyte Derived Dendritic Cells and in green the internalised DC-SIGN and PM26 complex.

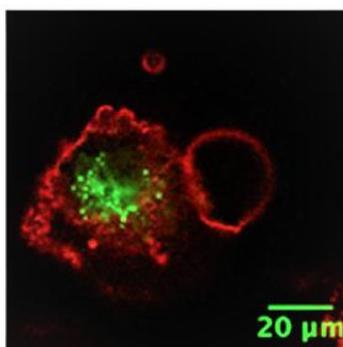


Fig.49 Example of imaging. DC-SIGN and PM26 complex internalisation by human immature Monocyte Derived Dendritic Cells.

PM26 was therefore shown to be internalized *via* DC-SIGN-mediated endocytosis and it ended up in lysosomes. Moreover, further results of chemokines production indicated that it modulated diverse innate responses and that it should be further developed for immunomodulatory approaches.

4.3 Targeting

Glycans are highly specifically recognised by lectins which makes them suitable carrier molecules to target drugs to different cells and tissues [142]. The use of targeted drug delivery reduces the side effects of drugs on healthy tissues by enhancing drug exposure on affected sites [101].

One example is the asialoglycoprotein receptor (ASGPR), a hepatic C-type animal lectin involved in the uptake of desialylated glycoproteins. It binds, in a Ca^{2+} dependent manner, to Gal and GalNAc residues and has been exploited as a site for drug targeting using Gal exposing carriers. For instance, liposomes decorated with asialofetuin can be used as drug carriers for intracellular delivery to liver cells [143].

While ASGPR targeting on hepatocytes is a mean for liver-specific drug delivery, targeting **myeloid CLR**s of the innate immunity has the additional advantage to shape immune responses. *K. Brzezicka et al* [134] showed that small structural glycan modifications can impact the uptake by DCs and subsequent T cell activation. In particular, they found that nonxylosylated ovalbumin (OVA) glycoconjugate increased binding to SIGNR3, one murine isoforms of DC-SIGN, when compared to xylosylated-OVA and unconjugated OVA. The Internalisation of nonxylosylated OVA led to T cell differentiation in Th1 and a production of IL-2 and IFN- γ (Fig.50).

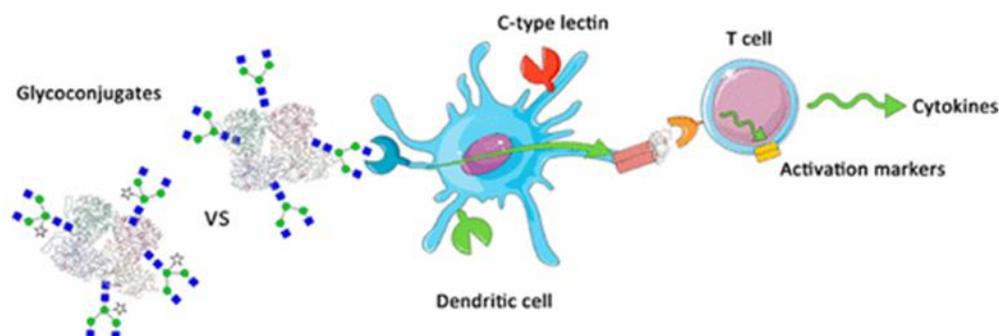


Fig.50 Targeting. Nonxylosylated and xylosylated ovalbumin glycoconjugate SIGN3 targeting and T cell activation. Adapted from [134].

CLR targeting may be fundamental for vaccination and tumour therapy. The next chapter presents some examples of strategy for cancer vaccination.

4.3.1 Vaccines against cancer

The glycan “fingerprints”/ structure change with the development of cancer and inflammation. There is, indeed, a link between glycan structure and the progression of disease. The alteration could be either an under- or over-expression of naturally occurring glycans or the neo-expression of glycans normally restricted to embryonic tissues. Features typical of malignancy development are the increase of branching

of N-glycans, sialylation, shortening of O-glycans chain and appearance of Lewis type antigens in glycoproteins. An extensive list of tumour-associated carbohydrate antigens (TACAs) is presented in the following picture (Fig.51) [132]:

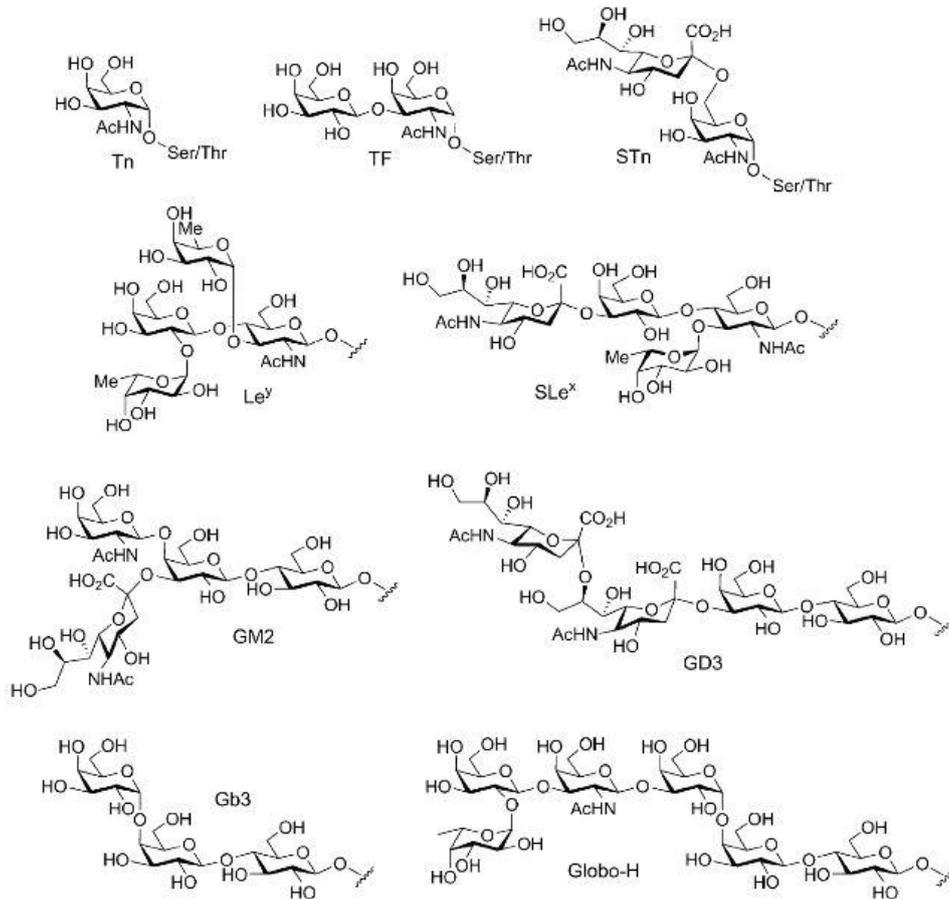


Fig.51 TACAs. Tumour-associated carbohydrate antigens. Adapted from [132]

Overexpression of glycoproteins such as mucin can be considered as tumour feature and used as diagnostic markers of cancer [144]. The expression of tumour-specific glycans is a consequence of decreased expression of normal epithelial glycans like disialyl-Lewis a, sialyl 6-sulfoLewis X and 3'-sulfated Lewis a epitopes [145].

Many tumour-associated glycans are poorly expressed in normal tissue and can be recognized as "self" if the immune system machinery has not been further activated. Moreover, similarly to pathogens and virus, tumour carbohydrates can escape the immune surveillance. For example, the glycoprotein MUC1 is a mucin highly expressed in breast cancer and is able to prevent monocyte differentiation to DCs.

However, the immune system can be guided to target cancer cells through their glycosylation pattern. One way to **trigger the immune system** against cancer cells is based on the induction of cytotoxic T cells (CTLs) against tumour specific peptides by targeting DCs, as already mentioned in the previous chapter. Since DCs

are the professional APCs for the induction of immune responses, they are vital in vaccination strategies (Fig.52).

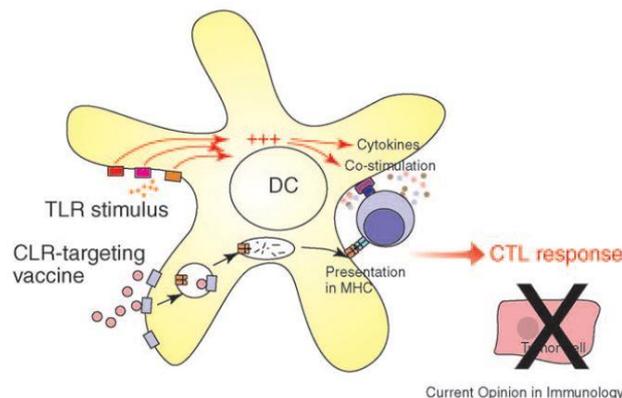


Fig.52 Vaccination using CLRs targeting. DC targeting vaccination strategy. Adapted from [145]

Glycans or glycomimetics specific for each CLR can be coupled to the antigenic peptide. It helps the targeting and the antigen internalization, ultimately improving tumour specific T cell response [145].

To date, this DC-based immunotherapy is mainly based on *ex vivo* generated autologous DCs loaded with tumour antigens [146]. DCs used in this strategy are derived from monocytes isolated from the patient blood. Cells are cultured in the presence of various cytokines to produce immature DCs and loaded with antigen before or following maturation. DCs can be targeted either with antibodies, against DC-SIGN for instance [147], or with carbohydrate selectively recognized by the desired CLR.

The design of glycan-based ligands specific to each CLR would be highly valuable to shape the immune response towards the desired immunological response. This is precisely the goal of the ITN **IMMUNOSHAPE** network to which my PhD project belongs.

5. AIM OF THE THESIS

The research work based on C-type lectin receptors and the development of their antagonist/agonist has started a long time before my arrival to the laboratory of Prof. Franck Fieschi. A tight collaboration established with the laboratory of Prof. Anna Bernardi (Università degli Studi di Milano) allowed the development of mannose-based glycomimetics selective towards DC-SIGN, while collaboration with Dr. Javier Rojo (Instituto de Investigaciones Químicas, CSIC, University of Sevilla) focused on the development of multivalent scaffolds (e.g. FP 7 ITN Carmusys). The pursuit aim was to optimise ligand affinity towards a single lectin without, however, looking at the potential cross reactivity between different CLRs

A conceptually different approach was then undertaken to address the selectivity problem. A large number of glycans and glycomimetics had to be investigated to cover major immunomodulatory glycan structures. A large number of CLR receptors, as well, were targeted to interact with their counterpart by using screening technology. The objective was to take advantage of existing arrays and available compound libraries and focus on their derivatisation to improve selectivity and binding to specific CLRs. The CLR selective binding would allow specific APC targeting that could be used as powerful strategy to shape the immune response towards the desired immunological response.

This project was part of a successful European ITN **IMMUNOSHAPE** consortium (<http://immunoshape.eu/>) which seeks to combine state of the art of synthesis and screening technology to develop lead structures for highly selective glycan based multivalent immunotherapeutics. 14 European academic and industrial partners with high-level expertise in different fields lead the IMMUNOSHAPE consortium, including **biochemists** from Grenoble (our group, Institut de Biologie Structurale) and Hannover (*Prof. Bernd Lepenies* group, Infection Immunology) **carbohydrate chemists** from Berlin (*Pr. Peter Seeberger* group, Max Planck Institute of Colloids and Interfaces), San Sebastian-Donostia (*Dr. Niels Reichardt* group, Glycotechnology Group), Milano (*Pr. Anna Bernardi* group, Dipartimento di Chimica), Sevilla (*Dr. Javier Rojo* group, Glycosystems Laboratory), Manchester (Prof. Sabine Flitsch group, The Manchester Institute of Biotechnology) and Leiden (*Pr. Jeroen Codee* group Leiden Institute of Chemistry), **cell biologists and immunologist** from Amsterdam (*Pr. Yvette van Kooyk* group, VU University Medical Centre, DC4U), Hannover (*Prof. Bernd Lepenies* group, Infection Immunology), Heidelberg (*Dr. Reinhard S. Albiez* and *Dr. Frank Momburg* group, Translational Immunology), experts of **multivalent platform** in Bilbao (Midatech Pharma, R&D/Nanoparticle Development Department) and in Orléans (GLYcoDIAG).

I have joined the network in June 2015, i.e. three months after it has started, and its workflow consists of several stages (Fig.53):

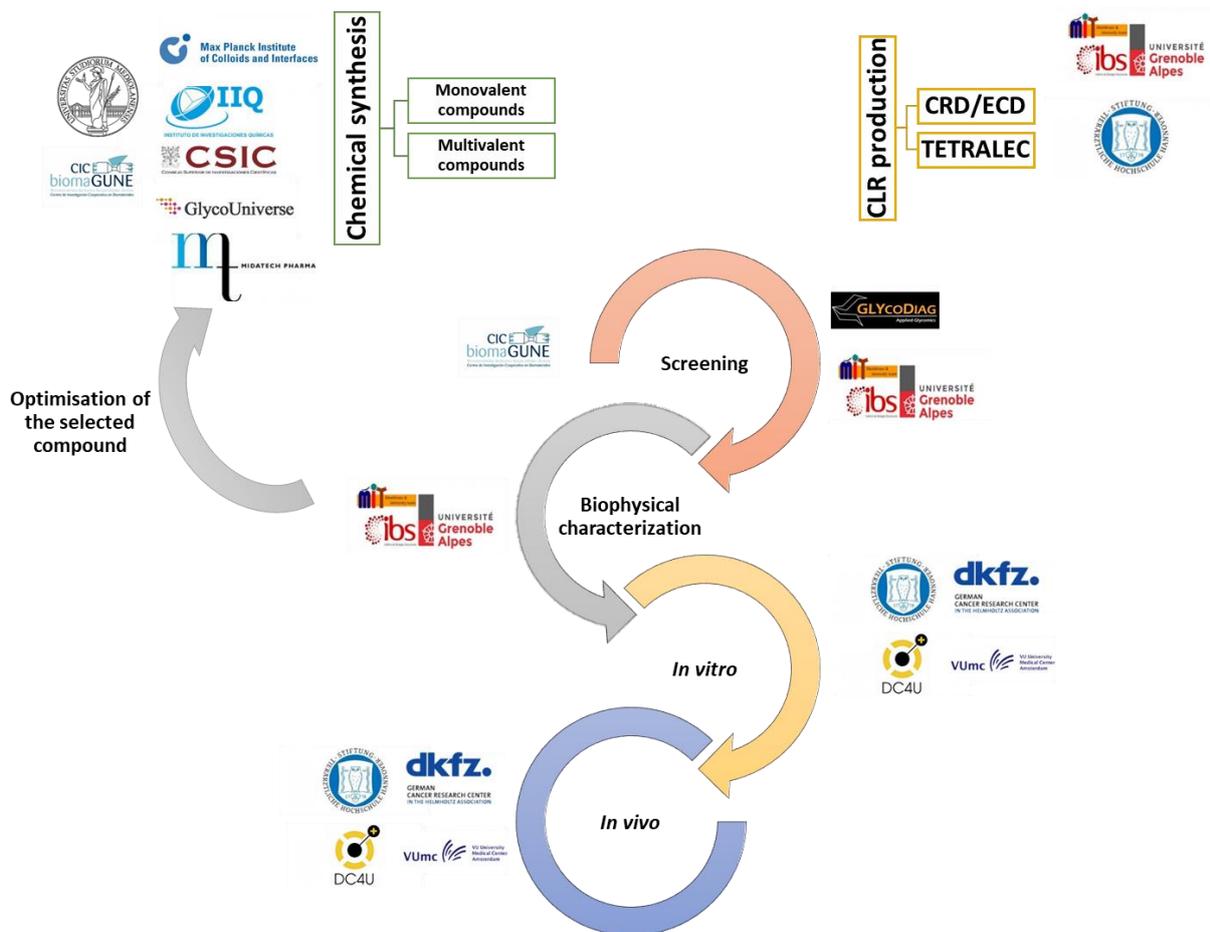


Fig.53 The principal architecture of IMMUNOSHAPE ITN consortium and group involvement.

The final long-term goal of this multidisciplinary network was to tailor the immune system response and bringing together synthetic chemistry, biochemistry, structural biology, microarray technology and immunology will help to develop novel glycan based immunomodulators.

The work performed during my PhD contributed to this large multidisciplinary research project by three different axes:

1. CLR production

Our main support to the project revolves around the production of DC-SIGN and langerin, for which the expression systems have already been set up in the Prof. Franck Fieschi laboratory, and the development of recombinant expression and purification protocols for seven other not commercially available human CLRs, namely BDCA2, DCSIGNR, Dectin-1, Dectin-2, LSECtin, MCL and Mincle. For each CLRs, two types of constructs were generated: the CRD, on one hand, and the full extracellular domain ECD, on the other. Multiple disulphide bridges (S-S) ensure the structural integrity of the carbohydrate binding domain but the proper S-S formation is a difficult task to achieve *via* bacterial recombinant expression. Thus, several

approach in parallel has been explored for CLR production in this work:

- The soluble and functional production of CLR in the bacterial periplasm, with its oxidative medium, through the combinatorial use of different strategies. Briefly, the signal peptide of proteins capable of crossing the periplasmic membrane were fused to the N-terminus of the CLR constructs. Expression was attempted in various bacterial strains.
- The high yield expression in bacterial cytoplasm as inclusion bodies followed by refolding steps. Indeed, our team had already established expertise in this approach of protein expression in *E.coli* inclusion bodies (IBs).

Finally, in the case of CRD production, monovalent binding can lead to poor ligand affinity impairing the screening in the first step. Thus, in order to artificially enhance the protein multivalency, and the affinity towards ligands, a strategy aiming to multimerize CRD constructs were developed, where specific biotin labelling of CRDs was coupled to the association to a tetramer of NeutrAvidin®. This final multimeric complex exposing four CRDs has been called the TETRALEC strategy.

2. The use of ligand screening methods

Two screening methods were exploited for functionality control assays:

- During Blanka Didak (PhD student of the IMMUNOSHAPE ITN consortium) secondment in Grenoble (04/07/2016-15/07/2016), LECprofile assay from GLYcoDIAG Company (Ludovic Lademarre), was used to investigate protein functionality and selectivity.
- During my secondment in the laboratory of Prof. Bernd Lepenies (11/12/2017-15/12/2017), human CLR were screened against Group A Streptococcus using FACS analysis.

Finally, glycan and glycomimetic microarrays were screened for their interaction with our library of recombinant human CLR during my two secondments in the laboratory of Dr. Niels Reichardt (02/05/2016-13/05/2016 and 03/04/2017-07/04/2017).

3. Biophysical studies of CLR ligand interaction and specificity.

The screening of the interaction between CLR and ligands provided qualitative information about the binding. Nevertheless, quantitative information (IC_{50} and K_d) regarding the specificity and the strength of the interaction were needed to guide final selection of some glycomimetics. In addition, quantitative binding assays allow to follow the optimisation of other previously identified glycomimetics. Diverse biophysical studies were performed to evaluate the interaction in the context of mainly two different collaborations within the IMMUNOSHAPE network:

-Monovalent glycomimetics were designed by our chemistry partner in Milano (laboratory of Prof. Anna Bernardi) and screened for their ability to selectively bind specific CLRs. This work has been performed by our team using SPR (competition assay) and ITC techniques.

-Simultaneously to the monovalent ligand development, the team in Leiden led by Jeroen Codee has synthesized multivalent compounds. Natural short chain glycans were loaded on a peptidic backbone with different multivalency and tested by SPR by using direct interaction assay. The candidates were also tested using ELISA and FACS analysis by the laboratory of Yvette van Kooyk. This last project was the first step on the design of a highly defined molecule for cancer vaccination by targeting CLRs.

Below, a map representation of the IMMUNOSHAPE groups (Fig.54). In red are highlighted our closed collaborators.



Fig.54 IMMUNOSHAPE network. In red CICbiomaGUNE (Niel Reichardt), GLYcoDiag (Ludovic Lademarre) , Leiden University (Jeroen Codee), VUmc (Yvette van Kooyk), Hannover Veterinary Universityt (Bernd Lепенies) and Università degli Studi di Milano (Anna Bernardi)

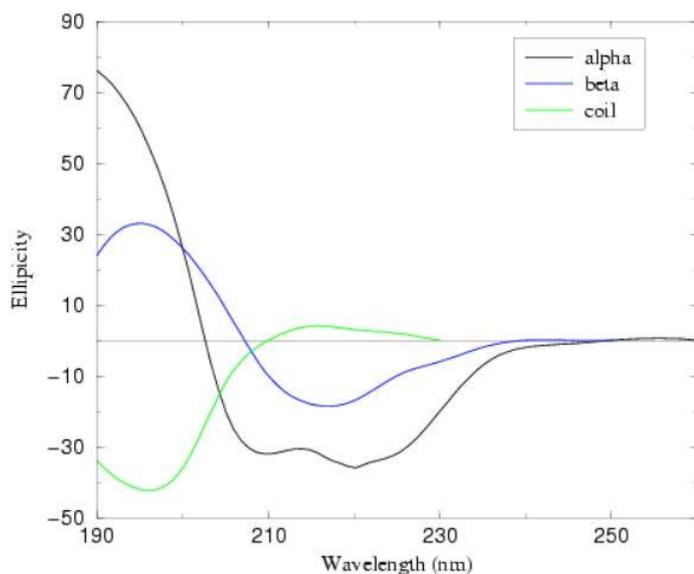
Part II.
Materials & Methods

6. Principles

6.1 Methods for protein characterization

6.1.1 Circular dichroism (CD)

The CD is a technique used to study the secondary structure of proteins (α -helices and β -sheets) by exploiting the different absorption of left-handed polarized light versus right-handed polarized light of optical active structure. α -helices and β -sheets have typical spectral signatures (Fig.55a). Thus, CD spectroscopy technique can be used as a sample quality analysis in order to control the correct folding of a given protein.



$$\theta(\text{deg.cm}^2.\text{dmol}^{-1}) = \frac{\text{Ellipticity}(\text{mdeg}).10^6}{\text{Pathlength}(\text{mm}).[\text{Protein}](\mu\text{M}).n}$$

a

b

Fig.55 Circular dichroism a) Typical far UV spectra signature for α -helices, β -sheet and random coil (www.isbg.fr/biophysics-characterisation/circular-dichroism). b) Equation used to calculate the molar ellipticity. mdeg is the CD signal, M the molecular weight in g/mL, C the concentration in g/L and L the path of the cell (cm).

The spectropolarimeter JASCO J-810 was used and, once the blank subtraction was done, the CD signal, expressed mdeg, was converted in molar ellipticity (Fig55b). Please refer to **7.2.3** for the detailed preparation of sample.

6.1.2 Size Exclusion Chromatography Multi Angle Laser Light Scattering SEC-MALLS (PAOL platform)

Protein Analysis On Line (PAOL) is a platform that combines the separation of macromolecules in solution and their characterization in term of mass, composition and size (<http://www.isbg.fr/biophysics-characterisation/protein-analysis-on-line-paol/>). Along with a Size-Exclusion Chromatography (SEC), it allows to perform simultaneously static and dynamic light scattering (MALS), as well as measurements of refractive index and absorbance. In general, MALS techniques measure the light scattered from the particles within a sample, at defined angles. A block scheme of the instrument is depicted in Figure 56:

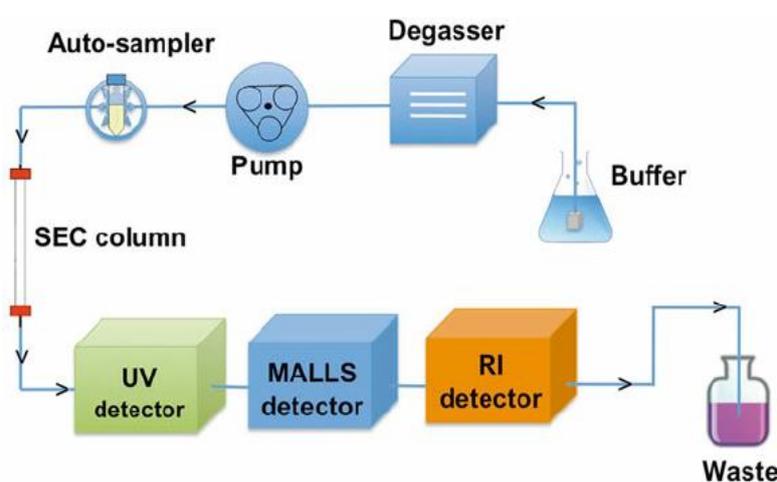


Fig.56 SEC MALLS. General block scheme of SEC-MALLS instrument [148].

Dynamic light scattering (DLS) and Static Light Scattering (SLS) are both exploited from MALS measurements. At a given angle (90°), the intensity of fluctuation of the scattered light measured in DLS will be directly linked to the dynamical properties of the sample particles in solution. It will be possible to calculate their diffusion coefficient, thus their hydrodynamic radius, by using the following equation:

$$D = \frac{K_B T}{6\pi\eta R_H}$$

Where D is the diffusion coefficient, K_B the Boltzmann constant, T is the temperature (K), η the solvent viscosity and R_H the hydrodynamic radius. DLS measurements are analysed in terms of particle size distribution, so that the presence of aggregates is easily detected.

SLS measurements are obtained at multi angles and collect the time averaged scattered light as function of the scattering angle. For large particles > 20 nm, it allows to obtain indications of the radius of gyration, which is another estimation of the particle size. From the intensity of the scattered light and the knowledge from refractive index measurements, the molecular weight of the particles in solution can be determined:

$$I = kcMW \left(\frac{\partial n}{\partial c} \right)^2 \quad k = \frac{4\pi n_0^2}{\lambda^4 N_A}$$

Where I is intensity of the scattered light, c the concentration in g/mL, (dn/dc) represent the variation of refractive index of the solution with the concentration, n₀ the refractive index of the solvent and N_A the Avogadro number. Absorbance and refractive index measurements are combined to determine stoichiometry complexes.

6.2 Methods for characterization of protein-ligand interaction

6.2.1 Lectin array (LectPROFILE)

During my PhD, I had the opportunity to use the LectPROFILE platform developed by the GLYCoDIAG Company (<http://www.glycodiag.com>). Briefly, the lectins were absorbed on a 96/well black plate (**Biomat**) and then screened with glycoproteins (asialofetuin, thyroglobulin, etc.) and neo glycoproteins. The latter were as well provided by our industrial partner and were composed by a molecule of BSA functionalized with a controlled number of glycans (mannosylated BSA, galactosylated BSA, etc.). Both glycoproteins and neo glycoproteins possessed a molecule of biotin that helped the revelation of the binding by the use of a fluorescently labelled streptavidin. A schematic representation of the approach is given in Fig.57.

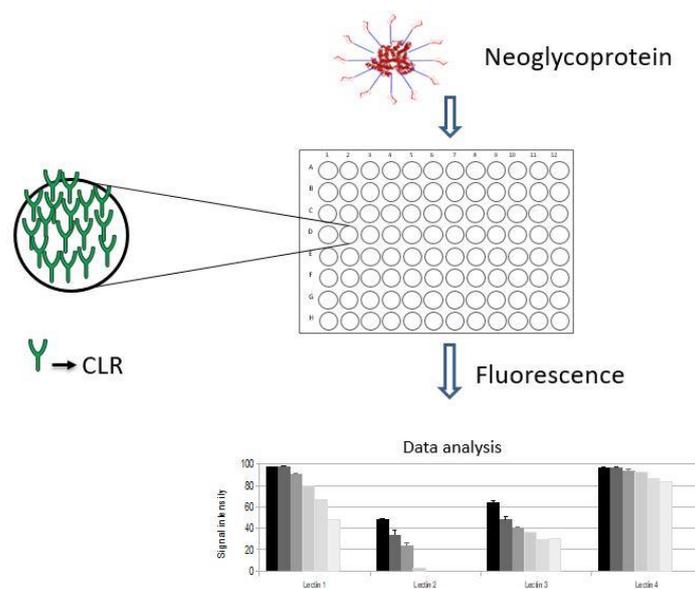


Fig.57 LectPROFILE schematic representation.

The fluorescence intensity was measured by using the CLARIOstar® plate reader (**BMG LABTECH**).

6.2.2 Glycan array

During my PhD, I had the opportunity to perform two secondments in Dr Niels Reichardt's laboratory (CICbioMAGUNE, Donostia-San Sebastian, Spain) and use their microarray platform. 134 synthetic N-glycans (mostly parasite and plant structures) and glycomimetics from Anna Bernardi, containing a reactive primary amines at their reducing termini were printed onto *N*-hydroxysuccinimide NHS ester activated glass microarray slides Nexterion® H (**Schott AG**) (Fig.58).

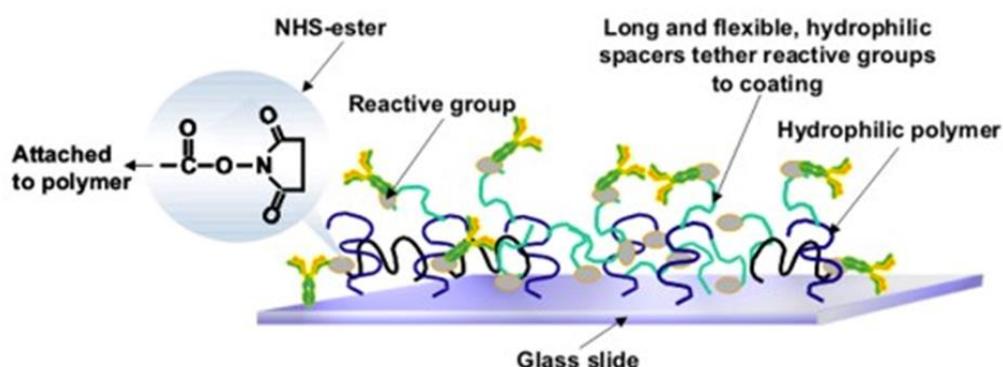


Fig.58 Glycan array glass slide. NHS ester glass microarray used during the two secondments [149].

Four spots for the same glycan and BSA as positive control were used. The fluorescence measurements after binding to labelled lectins were performed in an Agilent G265BA microarray scanner system (**Agilent Technologies**) and the quantification was achieved by ProScanArray® Express software (**Perkin Elmer**).

6.2.3 Surface Plasmon Resonance (SPR)

SPR is a biophysical technique that can study the biomolecular interaction of label-free compounds in real-time. The phenomena can be understood from the electromagnetic theory about the light reflection, transmission, and absorption for the multi-layer medium. You can find all the detailed SPR principles in the book written by H. Raether [150].

Surface Plasmon Resonance is a phenomenon that occurs when polarized light hits a metal film (silver or gold) at the interface of media with different refractive indices (glass and air for instance). When the polarized light is addressed through a prism on a sensor chip with a gold film, the free electrons of the gold, named plasmons, will swing and act as a mirror by reflecting the light (Fig.59) [151].

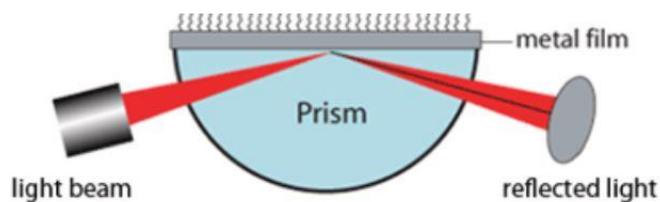


Fig.59 SPR principle. The light is reflected by the plasmons of the gold film
<http://biosensingusa.com/technologies/surface-plasmon-resonance/surface-plasmon-resonance-work/>

At a specific incident angle called resonance angle, the plasmons absorb light, reducing the reflected light intensity and creating a dark line in the reflected beam (Fig.60). When a molecular binding event takes place near the gold film, a shift of the dark line is observed and, by monitoring this shift vs. time, we can study molecular binding events and binding kinetics.

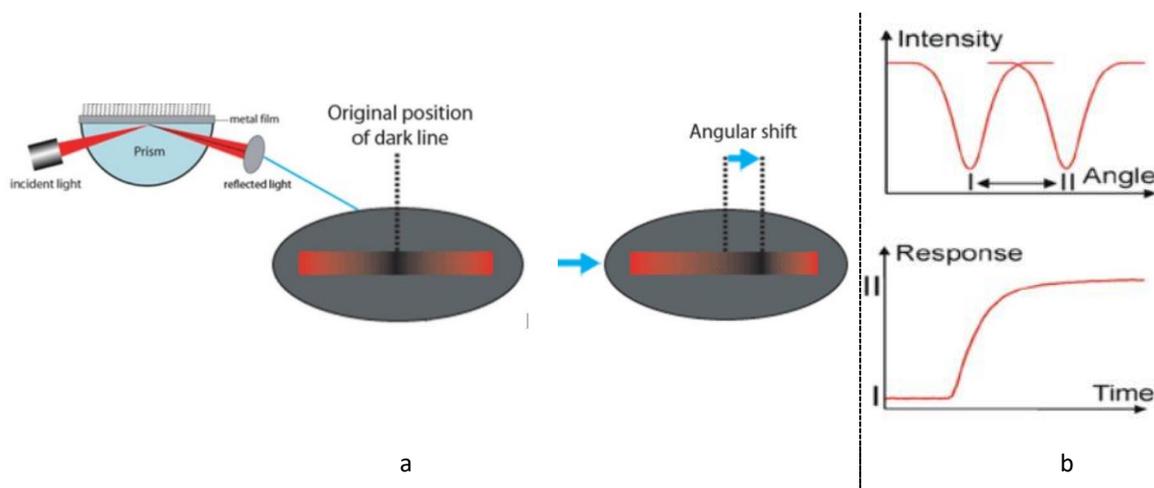


Fig.60 Plasmon resonance effect. a) At a specific angle, the plasmons absorb light and this can be seen by a dark line in the reflected beam. Adapted from <http://biosensingusa.com/technologies/surface-plasmon-resonance/surface-plasmon-resonance-work/>. b) The angle shift is seen in the sensorgram as a change of response during time, from [152].

Above the gold film, a dextran surface, in most cases, could be used for functionalization purposes. The compound immobilized on the latter is called ligand, while the partner in solution is the analyte. When these two components interact, the index of refraction at the metal surface changes and so does the resonance angle. This angle modification is detected by the BIAcore system and recorded by the BIAeval software as SPR signal or resonance units (RU).

The graphical result is called sensorgram, where the RU are expressed as function of time, and it consists of an association, steady state and dissociation phase (Fig.61a). In the association phase, the analyte is injected and the obtained shift is originated from the differences between the flow buffer and the sample buffer. When the number of association events is equal to the number of dissociation events, the plateau

or equilibrium or steady state has been reached, and the response at this stage is called Req. Finally, during the dissociation phase, the analyte is not injected anymore over the surface and the number of dissociation events get the better over association events. When the affinity with the ligand is strong, harsh condition of regeneration of the surface could be used to go back to the value of the baseline at the beginning of the interaction.

The analyte is injected at different concentration to obtain a sensorgram where signal is function of time. Information about the kinetics of reaction (k_{on} or k_{off}) for the association and dissociation respectively, the affinity (K_d) or, in the case of competition experiment, about the half maximal inhibitory concentration (IC_{50}) are obtained. However, care must be used during the design of a SPR experiment: the analytes should have sufficient mass in order to generate a significant change in refractive index and also an overall affinity for the ligand comprises between K_d values of $1 \cdot 10^{-4}$ - $1 \cdot 10^{-12}$ in order to measure reliable binding. In the case of weak affinity between an analyte and the ligand, multivalent analyte compounds might be needed to improve the binding. Moreover, the mass transport limitation must be avoid, which is mainly caused by the analyte and surface sites being located physically at different point (Fig.61b) [153].

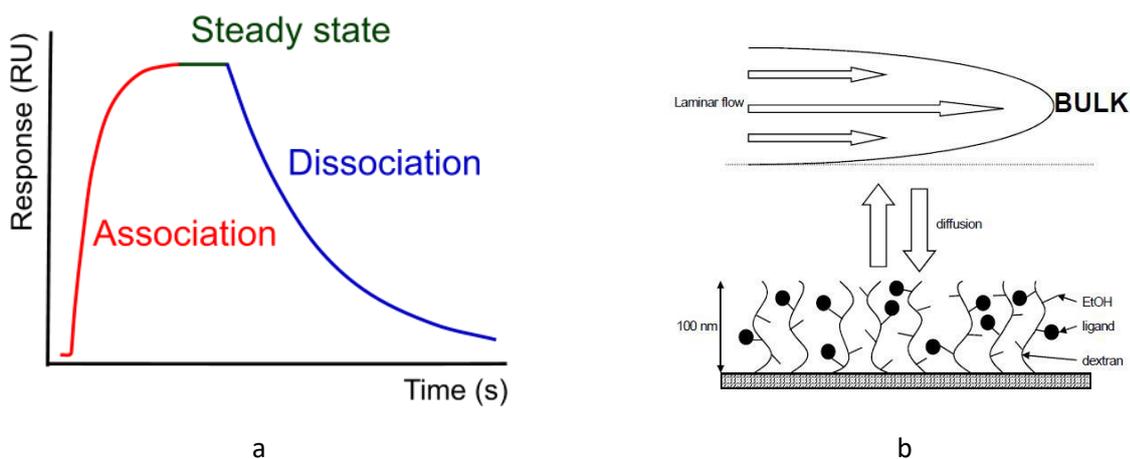


Fig.61 Analyte interaction and bulk effect a) Typical sensorgram of the analyte interacting with the ligand. The steady state corresponds to the equilibrium phase, while the dissociation happens when the analyte is not injected anymore over the surface b) bulk effect visual explanation (<https://www.sprpages.nl>).

The analyte needs to transfer from the bulk solution to the immobilized ligand on the 3D dextran surface and this includes the diffusion of the analyte within the dextran layer. The mass transport limitation occurs when the rate of mass transport is equal or slower than the association rate constant.

During my PhD, I had the possibility to use two different model to analyse the data: *4-parameter logistic* and *steady state affinity*.

4-parameter logistic model for IC₅₀ determination

Small saccharide or monovalent glycomimetic have low MW and low affinity for the analyte and their binding to the ligand is therefore difficult to detect. An indirect approach by competition test has therefore been developed by our group and gives quantitative information on the inhibitory power of the small molecule (IC₅₀). The sensor chip is functionalized with a glycoprotein, namely BSA mannosylated (BSA-Man), and the receptor, for which we want to evaluate the affinity of the tested compound, is co-injected over the surface at fixed concentration with increasing compound concentration (Fig.60). Surface and sample preparations are detailed in 7.2.9. The binding of the compound to the lectin-receptor will impair receptor binding to the ligand on the surface and will thus lowers the binding signal. Therefore, in this set up, the compound as the status of “inhibitors “of the analyte–ligand interaction (lectin/BASMan interaction in that example) (Fig.62). Indeed, as a function of the compound concentration used, an IC₅₀ value (Inhibitory Concentration allowing 50% of signal inhibition) can be obtained and allows evaluating the affinity of the compound for the analyte in the conditions tested.

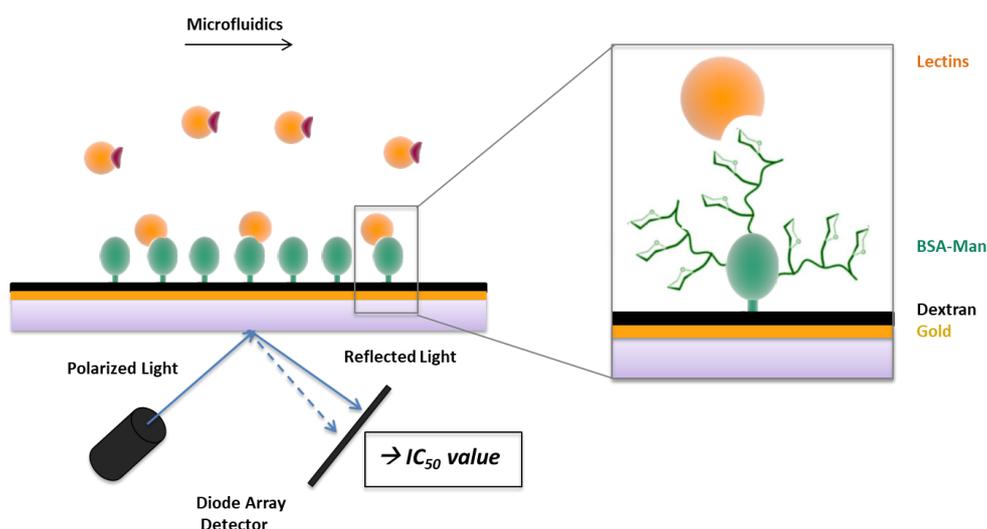


Fig.62 Design of competition/inhibition SPR assay. Inspired by [154].

The Req maximal values are used to calculate the IC₅₀ value through the following equation:

$$IC_{50} = A_1 \left(\frac{R_{hi} - R_{lo}}{R_{hi} - 50} - 1 \right)^{\frac{1}{A_2}}$$

Where R_{hi} is the maximal y-axis value of the curve, R_{lo} is the minimal y axis value of the curve, A₁ the x axis value corresponding to the middle linear part of the curve and A₂ the slope. Moreover, Req maximal values were converted in percentage of protein activity by considering 100% of protein activity the Req maximal obtained in the condition without inhibitor (Fig.63)

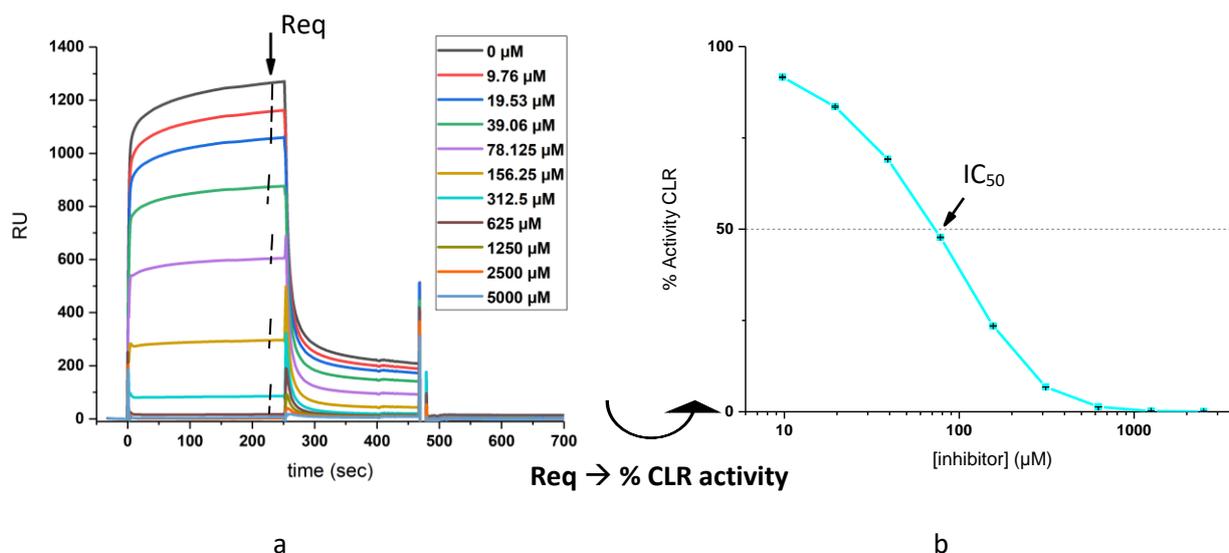


Fig.63 SPR competition/inhibition assay sensorgram and curve. a) Example of sensorgram with increasing concentration of inhibitors and corresponding Req maximal values b) Conversion of Req maximal values in % of protein activity and IC₅₀ graphical representation.

This approach, however, has an important limitation: being primarily based on the affinity between the protein and the surface, the system can detect the interaction of the inhibitor with the protein only if it is less strong than the protein-surface interaction. For example, DC-SIGN has an affinity (K_d) for the mannosylated surface of 5 µM and, therefore, inhibition in the range of nM cannot be viewed. For interaction in the nM range, direct interaction assay is recommended, considering that strong affinity interactions increase the probability to observe a reliable signal even with small molecular weight compound.

State affinity model for K_d determination

In the direct interaction approach, the lectin is immobilized on the surface in an oriented way. Surface and sample preparations are detailed in **7.2.10**. The lectin possesses a StrepTagII N-terminal sequence that is recognized by the StrepTactine. Once the StrepTactine is randomly immobilized on the sensor chip, it can recognize the tag and lead to an oriented immobilization of the protein by its N-terminus (Fig.64).

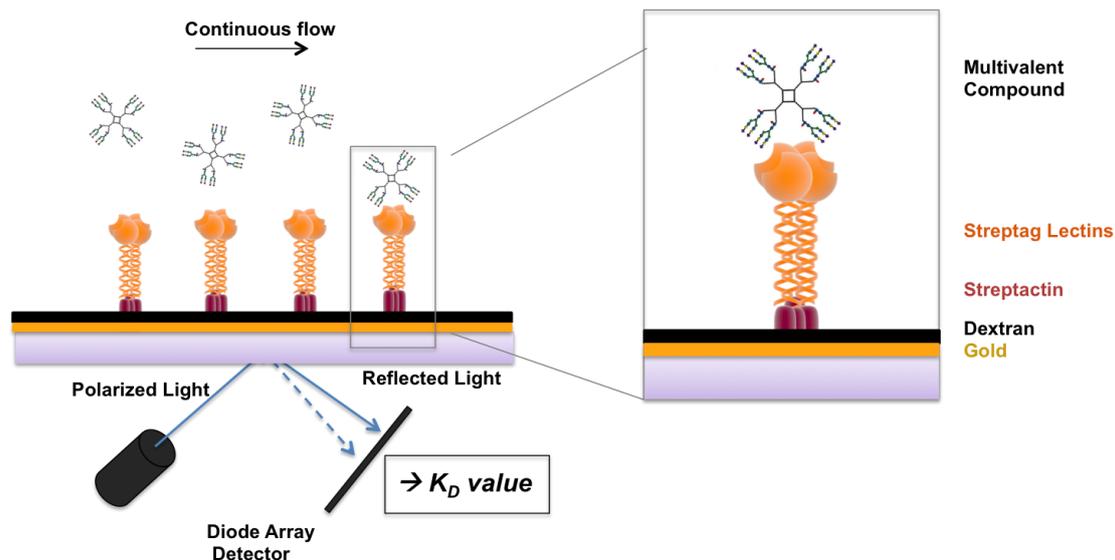


Fig.64 Oriented surface approach. Inspired by [154] and [125].

The multivalent ligand is injected at increasing concentration. The K_d is obtained with the steady state model by fitting a plot of R_{eq} against the concentration:

$$R_{eq} = \frac{CR_{max}}{K_D + C}$$

R_{eq} is the RU at the equilibrium at a given concentration, R_{max} the RU at the equilibrium at the highest concentration and C the concentration of analyte. Steady state affinity measurements are not affected by mass transportation limitations since they are based on report points taken at equilibrium

When using multivalent compound, it is important to keep in mind that the found K_d reflects the affinity for the surface and not for an individual receptor. However, this multisite interaction onto a surface may be close to the real interaction mode at the cellular surface level (with respect to comparable density of the ligand onto the surface). In chapters 8.3 and 8.4 the term “apparent” K_d (or K_{dapp}) will be used.

6.2.4 Isothermal Titration Calorimetry (ITC)

ITC is another biophysical technique that can be used to study the biomolecular interaction of label-free compounds in real-time in solution. By measuring heat transfer during binding between ligand (glycan) and receptor (CLR), it determines binding stoichiometry (n), binding constants (K_d), and thermodynamics parameters (variations of enthalpy (ΔH) and entropy (ΔS)). The glycan is added at several intervals at increasing concentration and the heat released upon the binding is measured with respect to a reference cell. The output given by the iTC200 instrumentation is the heat absorbed or evolved during the binding

and data are afterwards integrated to obtain a titration curve (kcal/mole of injectant versus the molar ratio) that could be used for modelling and K_d calculation (Fig.65).

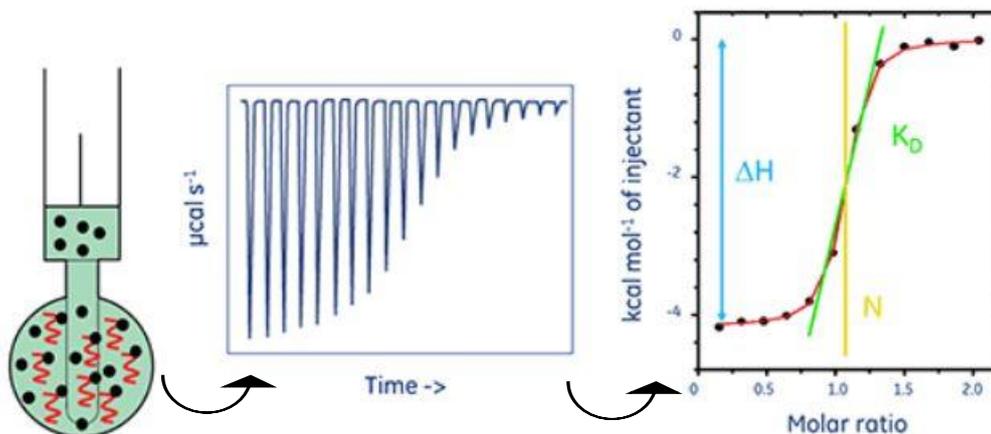


Fig.65 Calorimetry principle

(<http://www.malvern.com/en/products/technology/isothermal-titration-calorimetry/default.aspx>).

The technique does not require any labelling. However, it requires some caution: the buffer composition must be identical for the ligand and receptor, nevertheless a titration ligand against buffer should be performed to remove the buffer contribution. Moreover, it requires a substantial quantity of protein/ligand that can be calculated using the following equation:

$$c = \frac{[R]}{K_D} * n$$

Where [R] is the concentration of the receptor in mol.L⁻¹, K_d is the dissociation constant and n the hypothetical binding stoichiometry. The c value should be $10 < c < 1000$ in the experiment set up to allow the determination of K_d value [155].

7. Methods

7.1 Production of recombinant C-type lectin constructs

7.1.1 Transformation of Ca²⁺-competent *E.coli* strains

Four Ca²⁺-competent *E.coli* strains were used in this work, namely TOP10, BL21(DE3), Tuner (DE3) and Shuffle(DE3). The protocol of Ca²⁺-competent bacterial cell transformation consisted of the following steps: 1 µl of plasmid DNA was added to 25 µl of commercially Ca²⁺-competent cells and gently shaken. After 30 min of incubation on ice, heat-shock was performed for 45 s in 42°C water. 500 µl of Luria Bertani (LB) broth were added to the reaction and incubation at 37 °C 180 rpm for 1 hour was performed. The sample was centrifuged at 4500 rpm for 10 min, the supernatant discarded and the pellet resuspended in a residual volume of supernatant. The re-suspended cells were plated on petri dishes with LB-Agar containing the respective antibiotic and incubated overnight at 37°C.

7.1.2 Cloning of LSEctin His-CRD

pUC57 plasmid containing synthetic human genes encoding human LSEctin CRD (amino acids 162-292) codon optimised for the efficient production in *E. coli* were manufactured by GeneCust Europe (Luxembourg) Additional sequences coding for a 6His tag, a factor Xa cleavage site and 3 Glycines were including between the initiation codon and the first encoding codon of the synthetic gene. pUC57 LSEctin vector and pET30b expression vector (from **Novagen**) were consecutively digested with NdeI and HindIII restriction endonucleases (from **Fermentas-Thermo Fisher Scientific Inc**). Reaction mixtures were prepared as shown in Table 3.

Reaction mixtures	For pUC57 LSEctin vector (µL)	For pET30b vector (µL)
DNA	7 (1 µg)	7.5 (1 µg)
NdeI	2	2
HindIII	2	2
Buffer green	4	4
H ₂ O	25	24.5

Table.3 Digest assays of pUC57 LSEctin vector and pET30b vector.

The reaction was conducted for 1h at 37°C. Digested samples were analysed on a 1% agarose (w/v) gel prepared in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA). 10 µL of each sample were loaded onto the gel and “Mass Ruler DNA ladder Mix ready-to-use” (from **Fermentas-Thermo Fisher Scientific Inc**) was used as a size marker.

The digested pET30b vector and the LSEctin insert were purified via the E.Z.N.A. Gel extraction kit (from **OMEGA**). 990 ng of pET30b vector was dephosphorylated using FastAP Thermosensitive Alkaline phosphatase following the manufacturer conditions. LSEctin fragment was ligated to the dephosphorylated pET30b vector using Rapid DNA Ligation Kit (from **Roche**). The reaction mixture was prepared as follows: 1 µL digested pET30b vector was mixed with 7 µL of LSEctin insert and 2 µL of dilution buffer, then 10 µL of ligation buffer was added and mixed, finally 1 µL of ligase was added into the reaction . A negative control was performed by replacing LSEctin insert by water. The ligation reactions were carried out at room temperature for 15 min. Subsequently, 25 µl of commercial TOP10 *E.coli* cells were transformed with 5 µl of ligation reaction as described in **7.1.1**. The cells were plated on Petri dish with LB-Agar in the presence of 50 µg/mL kanamycin and grown at 37°C overnight.

Four clones were picked up from the LB-Agar plate and grown in 5 ml LB containing of 50 µg/mL kanamycin overnight at 37°C. Minipreps were prepared from 3 ml overnight culture with QIAprep Spin kit (from **QIAGEN**) according to manufacturer’s instructions, and the concentration of plasmid DNA was determined by measuring A_{260nm} of the sample and knowing that 50 µg/mL double-stranded DNA has a $A_{260nm} = 1$.

The presence of LSEctin in the plasmids was tested by NdeI/HindIII restriction digest. Empty pET30b and pUC57 LSEctin His-CRD were used as controls.

The reaction mix is described in Table 4.

Enzyme mixtures for 6 samples	Reaction mixtures
7 µL green buffer	5 µL (330 ng) plasmid
2 µL NdeI	5 µL enzyme mix
2 µL HindIII	positive control: pUC57 LSEctin
24 µL of H ₂ O	negative control empty pET30b

Table.4 Preparation of NdeI/HindIII digestion mixtures for the test of positive DNA ligation.

The reaction was conducted for 1h at 37°C. Digested sample were analysed in 1% agarose gel as described above.

7.1.3 Preparation of bacterial preculture

From the transformation plate, one clone was selected and inoculated into 50 mL of LB both with the respective antibiotic and cultivated overnight at 37°C. The prepared pre-culture was used to start the culture for protein over-expression (in case of BL21(DE3), C41(DE3), Tuner(DE3)- for SHuffle(DE3) strain, a temperature of 30°C was recommended) or to purify the vector from the cells (in case of TOP10 cells).

7.1.4 Preparation of bacterial culture samples for SDS-PAGE

For each constructs, a 5 mL sample of the culture before induction and a sample of culture at the end of the culture were centrifuged at 4500 rpm for 15 min. The supernatants were discarded. To the pellets obtained, 0,5 mL of 6X SDS-PAGE sample loading buffer (with or without β -mercaptoethanol) was added. After re-suspending, the cells were disrupted by sonication for 50 s at 10% amplitude.

7.1.5 DC-SIGN and Langerin over-expression and purification

Both protein over-expression and purification have been optimized before my arrival in the laboratory. For details, please refer to [57] and [81] respectively.

7.1.6 Protein labelling

Lectins were dialysed against 25 mM HEPES pH 8, 150 mM NaCl, 4 mM CaCl_2 and 1 μL of 10 mg/mL Cy3-NHS ester (**GeneCopoeia**) in DMSO was added each 0.3 mg of protein. The reactions were gently shaken at RT for 2 h and then at 4°C for 4 h. Excess dye was removed by repeated dialysis dialyzed (Spectra/Por® dialysis membrane, cut off 3.5 kDa, 9.3 mL/cm from **Spectrum laboratories**) of 3 h against 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl_2 (3hx2). The amount of attached Cy3 was estimated spectrophotometrically based on the dye epsilon (ϵ_{550} 150 000 $\text{cm}^{-1}\text{M}^{-1}$) and protein epsilon.

7.1.7 Over-expression and inclusion body preparation of all CLR constructs

E.coli strain BL21(DE3) was transformed with pETt30b/CLR plasmid according to protocol in subsection 7.1.1. 50 mL of overnight preculture was used to inoculate to 1L of LB broth with 50 $\mu\text{g}/\text{mL}$ kanamycin. The cells were grown at 37°C for 2/3h, then the protein over-expression was induced by addition of IPTG to final concentration of 1 mM and the culture was continued at 37°C for 3h. Centrifugation at 5000xg for 20

min at 4°C allowed cells harvesting, followed by the pellet resuspension in 30 mL of 25 mM Tris-HCl pH 8, 150 mM NaCl buffer and 1 tablet of “complete EDTA-free” (**Roche**) protease inhibitor cocktail. Cells were disrupted by sonication (sonicator BRANSON digital Sonifier® from **Emerson Electric Co**) for two rounds at 90% amplitude for a total time of 12 min for each round, using 2 s sonication and 10 s pauses in between and keeping cells in ice. Inclusion bodies together with the cell debris were collected by centrifugation at 100 000xg for 30 min at 4°C (ultracentrifuge with Ti45 rotor are from **Beckman Coulter**). To isolate inclusion bodies from bacterial cell debris, the pellet was resuspended using Potter-Elvehjem homogenizer in 30 mL of buffer containing 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, 2 M urea and 1% Triton-X100 (**Anatrace**®), centrifuged at 100 000xg for 30 min at 4°C. A washing step was performed by resuspending the pellet in 30 mL of 25 mM Tris-HCl pH 8, 150 mM NaCl buffer with Potter-Elvehjem homogenizer and centrifuged at 100 000xg for 30 min at 4°C.

7.1.7.1 Refolding, purification and labelling of DC-SIGNR ECD

Refolding. Inclusion bodies were solubilized with Potter-Elvehjem homogenizer in 30 mL of buffer containing 6 M Gdn-HCl (**Euromedex**), 150 mM NaCl, 25 mM Tris-HCl pH 8 and 0.01% β-mercaptoethanol and centrifuged at 100 000xg for 30 min at 4°C to eliminate the insoluble fraction. Solubilized DC-SIGNR ECD (typically 30 mL at around 5 mg/mL) was adjusted, before refolding, at a concentration of 2 mg/mL in the refolding buffer detailed above. Then from this starting concentration of 2mg/mL the protein has been refolded by a 5x flash dilution into a buffer containing 1,25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM CaCl₂ at 4 °C. The resulting protein solution was overnight against 4,59 L of 25 mM Tris-HCl pH 8 to remove Gdn-HCl. Two additional dialysis of 3h at 4°C against 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ were performed. Precipitates were eliminated by several cycle of ultracentrifugation.

Purification. Refolded DC-SIGNR ECD was purified in two steps at 4°C: an affinity chromatography followed by an exclusion chromatography. Firstly, refolded DC-SIGNR ECD solution was loaded onto a 20 mL Mannan-agarose column (**Sigma**) previously equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. After column washing, the elution step was performed using 25 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA buffer. 2.5 mL/min flow rate was maintained during the purification. The use of Multi Step Protein Purification (MP3) platform (<http://www.isbg.fr/samples-preparation/multistep-protein-purification/article/multistep-protein-purification>) allowed us to perform an exclusion chromatography automatically and immediately afterwards the affinity chromatography (Xpress FPLC system). A Superose-12 column (from x) was equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer and eluted in the same buffer with a flow rate of 1 mL/min. Eluted fractions were pooled and concentrated to 8 mg with VIVASPIN 20 (**Sartorius**, 5 kDa cut off) and frozen in liquid nitrogen for storage at -80°C

Protein labelling (section 7.1.5). Two different Degree of Labelling (DOL) were obtained for DC-SIGNR ECD. From the 2 mg dialyzed, 1 mg labelled with 1 μ L and the other 1 mg with 3 μ L of Cy3. The obtained DOLs were 0.4 and 0.95, respectively.

7.1.7.2 Refolding, purification and labelling of Dectin-2 S-ECD

Refolding. Inclusion bodies were solubilized with Potter-Elvehjem homogenizer in 30 mL of buffer containing 6 M Gdn-HCl, 150 mM NaCl, 25 mM Tris-HCl pH 8 and 0.01% β -mercaptoethanol and centrifuged at 100 000xg for 30 min at 4°C to eliminate insoluble part. Solubilized dectin-2 ECD (1,7 mg/mL, 30 mL) was refolded with a starting concentration of 1 mg/mL by flash dilution into buffer containing 1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl₂ at 4 °C. The resulting protein solution was dialyzed overnight against 2.96 L of water to remove Gdn-HCl. Two additional dialysis of 3h at 4°C against 4.5L of 25 mM Tris-HCl pH, 150 mM NaCl, 4 mM CaCl₂ were performed. Precipitates were eliminated by several cycle of ultracentrifugation.

Purification. Refolded Dectin-2 S-ECD was purified at 4°C through an affinity chromatography using AktaFPLC system (**GE Healthcare**). Firstly, Refolded Dectin-2 ECD solution was loaded to 1 mL StrepTactin column (**GE Healthcare**) previously equilibrated in 25 mM Tris-HCl pH, 150 mM NaCl, 4 mM CaCl₂ buffer. After column washing, the elution step was performed using 25 mM Tris-HCl pH, 150 mM NaCl, 4 mM CaCl₂ , 2.5 mM desthiobiotin (**IBA**) buffer. 1 mL/min flow rate was maintained during the purification. Eluted fractions were pooled, dialysed against 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ and concentrated to 0.8 mg/mL, 1 mL and frozen in liquid nitrogen for storage at -80°C.

Protein labelling (section 7.1.5). 250 μ g of Dectin-2 ECD were dialyzed and labelled with 3 μ L of Cy3, giving a DOL of 0.42.

7.1.7.3 Refolding, purification and labelling of Mincle His-ECD

Refolding. Inclusion bodies were solubilized with Potter-Elvehjem homogenizer in 30 mL of buffer containing 6 M Gdn-HCl, 150 mM NaCl, 25 mM Tris-HCl pH 8 and 0.01% β -mercaptoethanol and centrifuged at 100 000xg for 30 min at 4°C to eliminate the insoluble part. Solubilized Mincle His-ECD (3,346 mg/mL, 30 mL) was refolded at a starting concentration of 1 mg/mL by a 5x flash dilution into buffer containing 1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl₂ at 4 °C. The resulting protein solution was dialyzed overnight against 3 L of water to remove Gdn-HCl. Two additional dialysis of 3h at 4°C against 4.5L of 25 mM Tris-HCl

pH 8, 150 mM NaCl, 4 mM CaCl₂ were prepared. The precipitates were eliminated by several cycle of ultracentrifugation.

Purification. Refolded Mincle His-ECD was purified using AktaFPLC system in two steps at 4°C: an affinity chromatography followed by exclusion chromatography. Firstly, refolded Mincle ECD solution was loaded to 1 mL HisTrap column (**GE Healthcare**) previously equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. After column washing, the elution step was performed using 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, 0.5 M Imidazole buffer. 1 mL/min flow rate of buffer was maintained during the purification. The eluted protein was dialysed against 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ to eliminate imidazole and concentrated up to 2 mL. Mincle His-ECD was then injected into a methacrylate-modified polymer containing Toyopearl® column (120 mL, 0.5-80 kDa, **Tosoh Bioscience**) previously equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. 1 mL/min low rate of buffer was maintained during the purification. Eluted fractions were pooled into two different aliquots and concentrated to 0,3 mg/mL. 1,3 mL (pool n°1, hypothetical aggregation) and 1,587 mg/mL, 0.8 mL (pool n°2, hypothetical dimer), were both frozen in liquid nitrogen for storage at -80°C.

Protein labelling (section 7.1.5). 150 µg from pool n°1 and 300 µg from pool n°2 were dialyzed and labelled with 1 µL and 2 µL of Cy3, respectively. The obtained DOLs were 0.73 and 0.79, respectively.

7.1.7.4 Refolding and purification of His-GGG-CRD constructs

Refolding. Inclusion bodies were solubilized with Potter-Elvehjem homogenizer in 30 mL of buffer containing 6 M Gdn-HCl, 150 mM NaCl, 25 mM Tris-HCl pH 8 and 0.01% β-mercaptoethanol and centrifuged at 100 000xg for 30 min at 4°C to eliminate the insoluble part. Solubilized His-GGG-CRD was refolded by 5x flash dilution into refolding buffer at 4 °C. The resulting protein solution was dialyzed overnight to remove Gdn-HCl. Two additional dialysis of 3h at 4°C against 4.5L of 25 mM Tris-HCl pH, 150 mM NaCl, 4 mM CaCl₂ were performed. The precipitates were eliminated by several cycle of ultracentrifugation.

Purification. Refolded His-GGG-CRD was purified using AktaFPLC system at 4°C: through an affinity chromatography. Firstly, refolded His-GGG-CRD solution was loaded to 1 mL HisTrap column previously equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. After column washing, the elution step was performed using 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, 0.5 M Imidazole buffer. 1 mL/min flow rate of was maintained during the purification. The eluted protein was dialysed against 25 mM Tris-HCl pH, 150 mM NaCl, 4 mM CaCl₂ to eliminate imidazole and concentrated.

Values specific for each proteins are showed in the following table (Table 5):

<i>His-CRD</i>	<i>Refolding C (mg/mL)</i>	<i>Refolding buffer</i>	<i>Dialysis buffer</i>	<i>Protein yield/L culture (mg)</i>
BDCA2	1	1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl ₂	H ₂ O	3
DC-SIGNR	2	1,25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM CaCl ₂	25 mM Tris-HCl pH 8	20
Dectin-1	1	1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl ₂	H ₂ O	0.15
Dectin-2	1	1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl ₂	H ₂ O	2,2
LSEctin	1	1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl ₂	H ₂ O	0.7
MCL	2	1,25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM CaCl ₂	25 mM Tris-HCl pH 8	7.6
Mincle	1	1,25 M NaCl, 200 mM Tris-HCl pH 8, 25 mM CaCl ₂	H ₂ O	x

Table.5 Refolding concentration, buffer for refolding and dialysis and protein yield/L culture for the His-CRD considered.

His tag cleavage. His-GGG-CRD was cleaved using the factor Xa (**Thermo Fischer**) following the ratio recommended by the company: 1 µg of factor Xa per 50 µg of His-GGG-CRD protein at 1 mg/mL. The reaction was performed overnight at RT under agitation and then injected into Toyopearl® exclusion chromatography column previously equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. 1 mL/min flow rate of was maintained during the purification. Eluted fractions were pooled and concentrated up to 1 mg.

Values specific for each proteins are showed in the following table (Table 6):

<i>Sample</i>	<i>Volume (µL)</i>	<i>Volume (µL) factor Xa</i>	<i>Volume (µL) buffer</i>
BDCA2	2500 (2 mg/mL)	96	2204
DC-SIGNR	483 (8.3 mg/mL)	80	3436
MCL	420 (7,2 mg/mL)	60	2520

Table.6 Required volumes for His tag cleavage of BDCA2, DC-SIGNR and MCL CRDs.

Sortase-directed biotinylation and TETRALEC formation. The protocol from [156] was used for the biotinylation of GGG-CRD. The protein exposing three glycines at the N-terminus (1 equivalent) was mixed with the peptide biotin-LPRT-Ome (MW= 725.9 Da, **Covalab**) (5 eq.) and His-tag Sortase A (SrtA) (0.3 eq) from *S.aureus*, recombinantly produced in the lab, in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. The reaction was incubated at 37°C for 6h under agitation. The kinetic of reaction was followed by ESI-MS (electrospray ionization mass spectrometry) : 10 µL of reaction was analysed at 0h, 2h, 4h, 6h,8h and overnight. When the reaction was completed, the solution was loaded onto a 1 mL HisTrap column previously equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ buffer. After column washing, the elution step was performed using 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, 0.5 M Imidazole buffer. 1 mL/min flow rate of buffer was maintained during the purification. The His tagged sortase was retained by the HisTrap column where the untagged biotin-CRD was eluted during the washing step and was pooled and dialysed against 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ to eliminate un-reacted biotin.

Finally, **NeutrAvidin**® (MW = 14.5 kDa, **Thermo Fisher**) sample previously labelled with Cy3-fluorophore (2.9 mg/mL, DOL=0.5) was mixed to Biotin-CRD with a molar ratio of 1:1 and the reaction was incubated overnight at 4°C under agitation. The obtained CRD-TETRALEC was frozen in liquid nitrogen for storage at -80°C.

Values specific for each protein are showed in the following table (Table 7):

<i>Sample</i>	<i>Volume (µL) 1 eq</i>	<i>Volume (µL) Biotin-LPRT-Ome 5 eq</i>	<i>Volume (µL) Sortase 0,3 eq</i>	<i>TETRALEC (mg)</i>
BDCA2	600 (140 µM)	210.2	25.2	unachieved
DC-SIGNR	506.5 (120 µM)	151.96	18	0.4
MCL	500 (120 µM)	150	18	0.62

Table.7 Required volumes for sortase reaction on BDCA2, DC-SIGNR and MCL CRDs. mg of TETRALEC complex are also given.

Random biotinylation and TETRALEC formation. BDCA2 CRD (0,385 mg) was dialyzed against 25 mM HEPES pH 8, 150 mM NaCl and 4 mM CaCl₂ and 69,8 µL (25 eq) of EZ-Link Sulfo-NHS-Biotin (**Thermo Scientific**) was used for random labeling on lectin primary amines. The reaction was conducted at RT for 1h, under agitation, and was followed by two dialysis against 25 mM Tris-HCl pH 8, 150 mM NaCl and 4 mM CaCl₂. The Degree of Biotin (DOB) was determined using the Pierce Biotin Quantification kit (**Thermo Scientific**) and it revealed to be 2.18.

Finally, NeutrAvidin sample previously labelled with Cy3-fluorophore (2.9 mg/mL, DOL=0.5) was coupled to random-BDCA2-Biotin-CRD with a molar ratio of 1:1 and the reaction was incubated overnight at 4°C under agitation. The obtained 0,789 mg of random TETRALEC was frozen in liquid nitrogen for storage at -80°C.

7.2 Biochemical and biophysical protein characterization and ligand analysis

7.2.1 Protein samples for SDS-PAGE.

Gel SDS-PAGE

Bis-acrylamide (30%), SDS (20%) and TEMED solution are from **Euromedex**. Acetic acid, persulphate acid and Brilliant blue R from **Carlo Erba**. The ladders Page ruler Unstained and Prestained proteins are from **Thermo Fisher Scientific**.

1:1 volumes of protein solution and SDS-PAGE loading buffer 1X with or without β-mercaptoethanol (**Carl Roth**) were loaded on the gel. Generally, 10 µL of sample is loaded on 15% or 12% Tris. The gel migration undergoes for 45 minutes at 220 V with **Bio-Rad system**.

Western-Blot

After gel migration for 45 minutes at 220 V, proteins are transferred from the SDS-PAGE gel to a PVDF membrane (**Bio-Rad**) at 300 mA for 50 minutes using a buffer of transfer (25 mM Tris, 0.2 M glycine). 5% of milk diluted in PVS-tween20 was used to block for 1 hour. Anti-polyHistine-HRP antibodies (**Sigma**) from mouse were incubated for 1 hour, followed by three washes of 20 min with 50 mL of PBS-Tween20 buffer. The revelation was performed with the kit Sigma Fast 3-3' Diamino benzidine tablet (**Sigma**) following the protocol recommended by the company

7.2.2 Protein concentration determination

Protein concentration was determined by measuring the absorbance at 280 nM through NanoDrop 2000c spectrophotometer (**Thermo Scientific**). Molar extinction coefficient ϵ and Abs 0.1% values (OD for 1

mg/mL of protein) obtained by the amino acid sequence in ExPASy server (ProtParam tool) are presented in the following Table 8 and 9.

ECDs	MM (Daltons)	Isoelectric Point	ϵ (L*mol⁻¹*cm⁻¹)	Abs 0.1%
<u>DC-SIGN</u>	38845	5.1	70180	1.812
<u>Langerin</u>	28359	7.7	56000	1.972
<u>DC-SIGNR</u>	37167	4.92	60890	1.638
<u>Dectin-2 Strep</u>	21404	5.79	65930	3.080
<u>Mincle His</u>	22349	5.96	42565	1.905

His-CRDs	MM (Daltons)	Isoelectric Point	ϵ (L*mol⁻¹*cm⁻¹)	Abs 0.1%
<u>BDCA2</u>	17111	7.14	44835	2.620
<u>DC-SIGNR</u>	17202	5.85	48845	2.839
<u>Dectin-1</u>	16833	6.35	44835	2.663
<u>Dectin-2</u>	17009	5.95	50335	2.959
<u>LSEctin</u>	16429	6.28	48845	2.973
<u>MCL</u>	17322	6.43	43345	2.502
<u>Mincle</u>	17911	5.74	39335/38960	2.196

Table 8 and 9. Molecular weight, isoelectric points, molar extinction coefficient ϵ and Abs 0.1% values for the considered ECDs and CRDs.

7.2.3 Sample preparation for circular dichroism

Dectin-2 GGG-His-CRD and MCL GGG-His-CRD samples were analysed at a concentration of 1,46 mg/mL and 7,681 mg/mL, respectively, by using a cuvette of 100 microns. The experiences were conducted at 20 °C. The reference was performed with the buffer 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂.

7.2.4 Sample preparation for SEC-MALS (PAOL platform)

A KW 802,5 column was used and equilibrated with the filtered at 0,1 μm running buffer 25 mM Tris-HCl pH 7,5, 150 mM NaCl, 4 mM CaCl_2 at a 0.5 mL/min of flow rate. The calibration was done using 95 μL of BSA 2 mg/mL in the running buffer. The following Table 10 lists the sample analysed on the column after 20800xg centrifugation for 15 min at 4°C

<i>Proteins</i>	<i>C (mg/mL)</i>	<i>vol (μL)</i>	<i>MW (KDaltons)</i>	<i>E ($\text{mL}\cdot\text{g}^{-1}\cdot\text{cm}^{-1}$)</i>	<i>Dn/dc (mL/g)</i>
BSA	2	95	66	667	0.186
DC-SIGNR ECD	5	50	37,2	1936	0,191
DC-SIGNR Fc	1,25	100	39,7	1625	0,188
DC-SIGNR TETRALEC	2,21	50	31	2311	0,191
MCL TETRALEC	0,8	95	16,6	2607	0,196
NeutrAvidin	2,9	20	14,48	1662	0,185

Table 10. List of analysed CLR's at the SEC-MALS and their specifications.

7.2.5 Sample preparation for lectin array analysis

The protocol is under GLYcoDIAG Company property.

7.2.6 Sample preparation for glycan array analysis

Cy3 labeled C-type lectins were diluted in incubation buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl_2 , pH=7.5 containing 0.5% bovine serum albumin (BSA) and 0.005% Tween[®]-20). C-type lectin solutions (200 μL per array) were used to incubate individual wells on a glycan array slide at 4 °C for 18 hours. Arrays were washed with incubation buffer without BSA, H_2O and dried in a slide spinner.

Fluorescence measurements were performed on a microarray scanner (Agilent G2565BA, **Agilent Technologies**) at 10 μm resolution. Quantification of fluorescence was performed by ProScanArray[®] Express software (Perkin Elmer) employing an adaptive circle quantification method from 50 μm (minimum spot diameter) to 300 μm (maximum spot diameter). Average RFU values with local background subtraction

of four spots and standard deviation of the mean were reported using Microsoft Excel and GraphPad Prism® software.

7.2.7 Sample preparation for FACS

Heat-killed *Candida albicans* (**InvivoGen**) and GAS (**ATCC cell lines**) were stained for 15 min with 1 μ M of DNA-staining dye Syto61 (**ThermoFisher Scientific**) at 4°C. The samples were subsequently washed two times with 1x PBS. Then, samples were incubated for 1h either with 250 ng of the respective CLR-hFc fusion proteins in lectin-binding buffer (50 mM HEPES, 5 mM MgCl₂, 5 mM CaCl₂, pH 7.4) or with different concentrations of human CLR constructs in its respective lectin-binding buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, pH 8.0). After washing once with the lectin-binding buffer, the pellet was suspended in a 1:200 PE-conjugated goat anti-human Fc antibody (Dianova) and incubated for 20 min at 4°C, for detection of the bound CLR-hFc fusion proteins. Finally, cells were washed two times and flow-cytometric analysis was performed using an Attune NxT Flow Cytometer (**ThermoFisher Scientific**). The gating strategy applied was a first gate in the *Candida albicans* and GAS population, followed by a single cell population gating for doublets exclusion. In the single cell population gate, Syto61 positive cells were selected and further analyzed for CLR binding. The same gating strategy was performed for all experimental conditions within one experiment. Flow cytometry data were analyzed using the FlowJo version 10 software (**FlowJo**).

7.2.8 Sample preparation for ITC

ITC experiments were performed at 25 °C using a TA Instrument Nano Isothermal Titration Calorimeter Low Volume (Nano ITC LV) with a 190 μ L cell volume. Compound Man069 and DC-SIGN ECD were prepared in 25 mM Tris-HCl at pH 8, 150 mM NaCl, 4 mM CaCl₂ and 4% DMSO. The compound was stepwise injected (1.03 μ L) into DC-SIGN solution using 5 min intervals between injections. Then, 100 μ M of DC-SIGN ECD and 2.5 mM compound concentrations were used. The blank titrations (compounds to buffer) were done for subtraction of dilution heat from the integrated data. A one-site binding model was fit to the data (nanoAnalyse 2.20 TA), yielding association constants (K_A) and binding enthalpies (ΔH).

7.2.9 SPR surface preparation for inhibition test

Biacore T200 was used for measurements on CM (carboxymethyl) 3 sensor chip S series (GE-Healthcare) and the system was equilibrated in HBS-P buffer (0.01 M HEPES pH7.4, 0.15 M NaCl, 0.05% P20). Flow cells

1, 2 and 3 were activated by a 1:1 ratio of 0.2 M *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide (ECD) (**GE-Healthcare**) and 0.05 M *N*-hydroxysuccinimide (NHS). Flow cell 1 was used as reference and functionalized with BSA, while flow cells 2 and 3 were functionalized with mannan BSA. BSA and mannan BSA were used at a concentration of 60 µg/mL in 10 mM C₂H₃NaO₂ (NaOAc) pH 4. Once the functionalization has occurred, 30 µL of 1M ethanolamine-HCl pH 8 (**GE-Healthcare**) were used to deactivate the un-functionalized reactive groups. Two consecutive washings were then applied: 5 µL of 10 mM HCl and 5 µL of 50 mM EDTA. 2500-3500 RU of functionalization were considered adequate for the subsequent experiments. All the previous steps were performed using a flow of 5 µL/min

For all the inhibition tests, the sensor chip was equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, 0.05% P20 running buffer and ECD constructs were tested.

Surface stability control: 20 µM of DC-SIGN is injected 3 times with a flow rate of 5 µL/min. Each single injection is followed by a surface regeneration with EDTA 50 M pH 8. The control of the surface stability could be ideally done after every 10 cycles of experiment to assess stability of the interaction.

Lectin ECD titration and K_{Dapp} determination: titrations with DC-SIGN DC-SIGNR and langerin were performed at increasing concentration. DC-SIGN: 0,195 µM; 0.39 µM; 0.78 µM; 1,5 µM; 3,125 µM; 6,25 µM; 12.5 µM; 25 µM; 50 µM. DC-SIGNR: the same concentration plus 80 µM.

Sample preparation: lyophilized inhibitors were resuspended in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂ at the highest reachable concentration (6.5 mM-20 mM). 4% DMSO was eventually added to overcome water insolubility problem. All compound stock solutions were centrifuged at 13000 rpm at 4°C for 10 min to remove aggregates.

All inhibition assays have been performed using serial dilution of the inhibitor by factor 2. The starting sample was made in a total volume of 120 µl with a concentration of 20 µM DC-SIGN/langerin or 30 µM DC-SIGNR and 5000 µM (or 2500 µM) of the selected antagonist.

Results reproducibility: in order to have a statistical reproducibility, flow cells 2 and 3 were used to test the inhibitors. During all the run, an appropriate control compound was used.

7.2.10 SPR surface preparation for direct interaction test

Biacore T200 was used for measurements on CM3 sensor chip S series and the system was equilibrated in HBS-P buffer (0.01 M HEPES pH7.4, 0.15 M NaCl, 0.05% P20). This surface preparation for direct interaction test is characterized by two different stages of functionalization. Flow cells 1, 2, 3 and 4 were activated by a

80 μL of 1:1 ratio 0.2 M *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide (EDC) and 0.05 M *N*-hydroxysuccinimide (NHS). The activation was followed by the functionalization with 170 $\mu\text{g}/\text{mL}$ of StrepTactin (IBA, 1 mg/ml) in 10 mM NaOAc pH4 buffer using 5 $\mu\text{L}/\text{min}$ of flow. Once this first functionalization has occurred, 80 μL of 1M ethanolamine-HCL pH 8 were used to deactivate the un-functionalized reactive groups. Two consecutive washings were then applied at 100 $\mu\text{L}/\text{min}$: 100 μL of 10 mM HCl and 100 μL of 1 M NaCl/ 50 mM NaOH. The second functionalization is performed by exploiting the Strep tag of DC-SIGN S. 49 $\mu\text{g}/\text{mL}$ of DC-SIGN S is injected over the surface at flow of 5 $\mu\text{L}/\text{min}$. For more detail about the production of DC-SIGN S-ECD please refer to the thesis of Ieva Sutkeviciute². Again, 80 μL of 1M ethanolamine-HCL pH 8 were used to deactivate the un-functionalized reactive groups. Two consecutive washings were then applied at 100 $\mu\text{L}/\text{min}$: 100 μL of 10 mM HCl and 100 μL of 1 M NaCl/ 50 mM NaOH.

The same procedure was applied for Langerin (55.9 $\mu\text{g}/\text{mL}$). For more detail about the production of langerin S-ECD please refer to the thesis of Eric Chabrol³.

For all direct interaction tests, the sensor chip has to be equilibrated in the following running buffer: 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl_2 , 0.05% P20.

Surface stability control: 27 nM of Man-BSA was injected 3 times with a flow rate of 5 $\mu\text{L}/\text{min}$. Each single injection is followed by a surface regeneration with EDTA 50 M pH 8. The control of the surface stability could be ideally done after every 10 cycles of experiment to assess the continuity of the functionalization level.

Man-BSA titration and K_{dapp} determination: a titration with Man-BSA was performed at increasing concentration.

The regeneration was performed with 100 $\mu\text{L}/\text{min}$ of 50 mM Gly-NaOH pH 12, 0.15% Triton X100, 25 mM EDTA pH8.

Sample preparation: lyophilised ligands were re-solubilized in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl_2 at the highest reachable concentration (0.25-6.5 mM). 4% DMSO was eventually added to overcome water insolubility problem. All compound stock solutions were centrifuged at 13000 rpm at 4°C for 10 min to remove aggregates.

All direct interaction assays were performed using serial dilution of the compound by factor 2 and the starting sample was made in a total volume of 120 μL . The regeneration was performed with 100 $\mu\text{L}/\text{min}$ of 50 mM Gly-NaOH pH 12, 0.15% Triton X100, 25 mM EDTA pH8.

² <https://tel.archives-ouvertes.fr/tel-00819832>

³ <https://tel.archives-ouvertes.fr/tel-00743636>

Results reproducibility: in order to have a statistical reproducibility, flow cells 2 and 3 were used to test the inhibitors. During all the run, an appropriate control compound was used.

Software used for the analysis: Biacore T200 Evaluation software 3.1 and OriginPro 2017

Part III.
Results and Discussions

8. Recombinant CLR Production and Functional Test

Proper folding of CLR CRDs requires the appropriate formation of disulphide bridges. This step is difficult to achieve *via* recombinant bacterial expression as *E.coli* cytoplasm constitutes a reductive environment. So far, most of the lectins produced in Pr. Fieschi laboratory are over-expressed as inclusion bodies and subsequently refolded *in vitro*. Such a strategy enables, for instance, high yield production of DC-SIGN ECD (50 mg/L of culture). However, this approach has some drawbacks: each refolding protocol is protein dependent and has to be individually optimised. Moreover, the functionality of the refolded protein has to be subsequently assessed, which could be problematic when no ligand has been identified or when the ligand is not commercially accessible. Therefore, for the seven non-commercially available considered CLRs, in addition to constructs that enable insoluble production, an alternative strategy was also investigated. The latter addresses proteins towards *E.coli* periplasm, which provides an oxidative environment suitable for disulphide bridges formation and protein functional production (Fig.66).

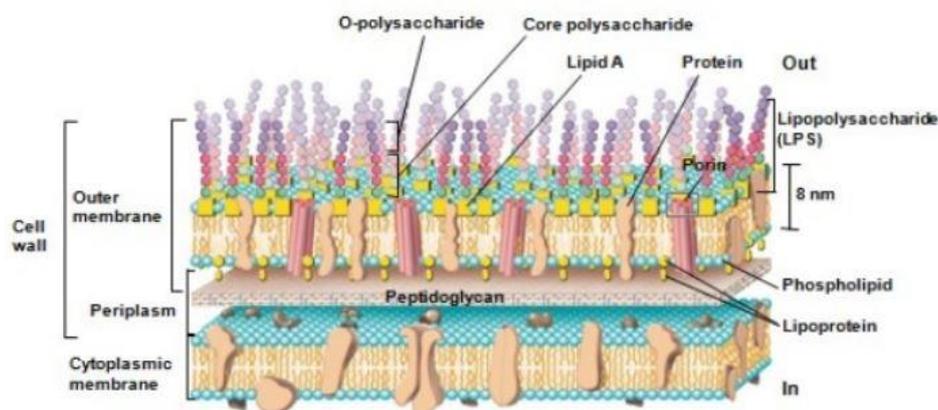


Fig.66 bacterial cell wall section and location of the periplasm (3rd chapter of Microbiology 8th edition).

Periplasmic targeting was attempted using the signal peptides of the proteins ompA (outer membrane protein)[157] and pelB (pectate lyase B)[158] both enabling inner membrane crossing *via* the SEC (SecB-dependent) bacterial post translational machinery (figxa). DsbA signal peptide (disulphide bond A)[159] was also exploited to attempt co-translational membrane crossing *via* SRP (signal recognition particle) pathway (Fig.67a). The three signal peptides share the same overall amino acid organization (Fig.67b): the N-terminus is charged, the central region is composed by hydrophobic amino acids and the C-terminus comprises polar amino acids.

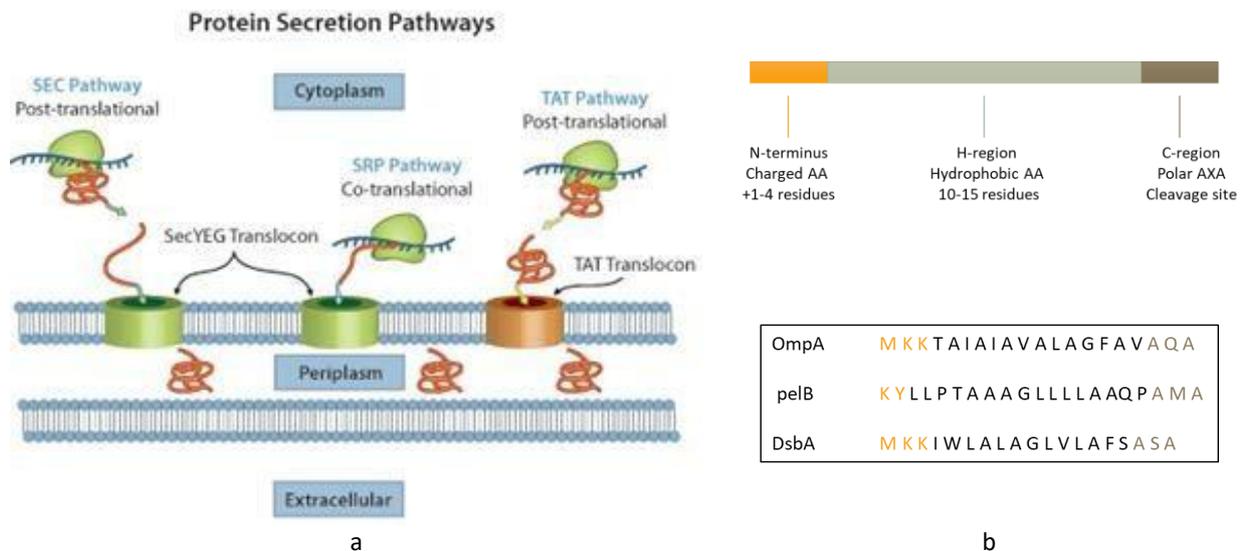


Fig.67 Protein secretion pathway and signal peptides a) SEC post-translational pathway and SRP co-translational pathway (the TAT pathway was not explored). www.athenaes.com b) Overall amino acid composition of three signal peptides.

However, this strategy presents an important disadvantage: the yield of production is often very low (in the μg range). The CRD of langerin is produced properly folded in the *E.coli* periplasm with a similar approach in our laboratory [160].

8.1 Cloning

A large range of constructs have been prepared for all the lectins, either before my arrival or during the beginning of my PhD. For clarity sake, and also because I have not been directly involved in their production, the details of their preparation will not be described, except for pET30b LSEctin His-CRD for which I have handled the cloning step. A schematic flowchart of the construct strategy is found in Fig.68.

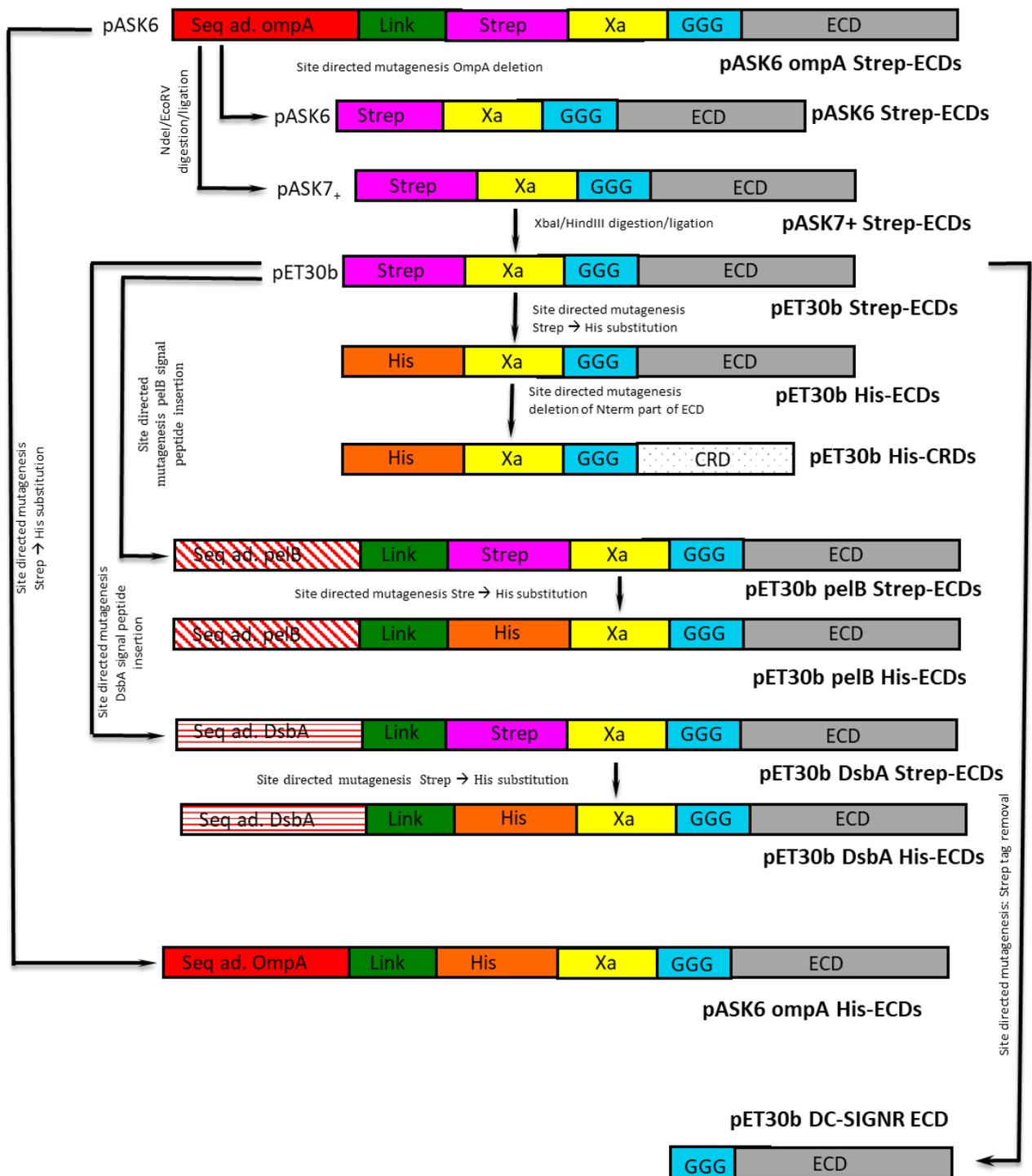


Fig.68 Cloning strategies.

All the original human genes have been manufactured by **GeneCust Europe** (Luxembourg). They consisted in sequences optimised for *E.coli* expression to circumvent rare codon obstacles. The synthetic genes contained the coding sequence of the CLR ECD (DC-SIGNR: 80-399; dectin-2: 46-209; mincle: 40-219) preceded by three glycines for subsequent sortase functionalisation.

ECD constructs

pASK constructs for soluble expression:

The first series of constructs were made by the ligation of the lectin synthetic sequences into the pASK6 periplasmic expression vector (**IBA Lifesciences**) between the BsaI cloning sites, which allow an orientated insertion. The resulting vectors encode for proteins presenting from N-terminus to C-terminus: the signal peptide of ompA (for periplasmic targeting), a Strep-tag II affinity tag, a Xa factor cleavage site, three glycines and the lectin ECD. This series of construct will be termed pASK6 ompA Strep-ECDs.

Another series of constructs called pASK6 ompA His-ECDs was prepared from the former by replacing the Strep-tag II by a His Tag via site directed mutagenesis (QuikChange Lightning Multi Site-Directed Mutagenesis Kit from **Agilent**), a strategy that has in the past optimised the level of expression for other lectins studied in the laboratory.

Constructs for insoluble expression:

Four series of constructs were made to undertake insoluble expression trials in three different vectors

- The first one was obtained by a deletion site-directed mutagenesis on pASK6 ompA Strep-ECD constructs by removing the ompA signal peptide. This series of construct will be termed pASK6 Strep-ECDs.
- The second one was obtained from the above series by NdeI/EcoRV restriction enzyme digestion and ligation into the pASK7+ vector (**IBA Lifesciences**) designed for cytoplasmic expression. This series of construct will be termed pASK7+ Strep-ECDs.
- In pASK vectors, genes are under the control of a Tet promotor inducible with anhydrotetracycline. We decided to test a stronger promoter: the T7P of the pET30b vector (**Novagen**). Therefore, the third series was generated from the latter by XbaI/HindIII digestion and ligation into pET30b vector. This series of construct will be termed pET30b Strep-ECDs.

For the reason mentioned above and because His Tag seems more prone to refolding approaches than Strep-tag II, the Strep-tag II of the pET30b Strep-ECDs series has been replaced by a His Tag (pET30b His-ECDs vectors) by site-directed mutagenesis .

In addition, a construct of DC-SIGNR ECD in pET30b without any tag was prepared as previously done for DC-SIGN in the laboratory.

pET30b constructs for soluble expression:

From the pET30b Strep-ECD constructs, two additional sets of vectors for periplasmic soluble expression were created by the insertion of pelB and DsbA signal peptides by site-directed mutagenesis. This series of construct will be termed pET30 pelB Strep-ECD and pET30 DsbA Strep-ECD, respectively.

Once again, on the two series of constructs mentioned above the His counterparts were prepared by replacing the Strep-tag II by a His Tag via site directed mutagenesis. This series of construct will be termed pET30 pelB His-ECD and pET30 DsbA His-ECD.

CRD constructs

All the CRD constructs were generated by site directed mutagenesis from the pET30b His-ECD vectors. The strategy consisted in a deletion to remove the N-terminal part of the ECD. The CRD sequences corresponded to amino acids 79-213 for BDCA2, 264-399 for DC-SIGNR, 111-247 for dectin-1, 75-209 for dectin-2, 162-292 for LSEctin, 81-215 for MCL and 77-219 for Mincl.

For some of the CLRs (C80 for BDCA2, C76 for dectin2, C82 for MCL and C78 for mincl) the presence of a N terminal cysteine not involved in disulphide bridge formation was awaited to be problematic during the refolding procedure and was therefore substituted by a serine.

8.1.1. LSEctin His-CRD Cloning

LSEctin His-CRD synthetic sequence was transferred from the pUC57 vector in which it was originally cloned by GENECUST Europe to the pET30b expression vector. Both constructs were digested by NdeI and HindIII restriction enzymes (Fig.69a) and the LSEctin His-CRD fragment (indicated by the red arrow) was ligated into NdeI/HindIII digested pET30b. Insertion of LSEctin His-CRD was assessed by NdeI/HindIII digest on several clones obtained after ligation and transformation (Fig.69b)

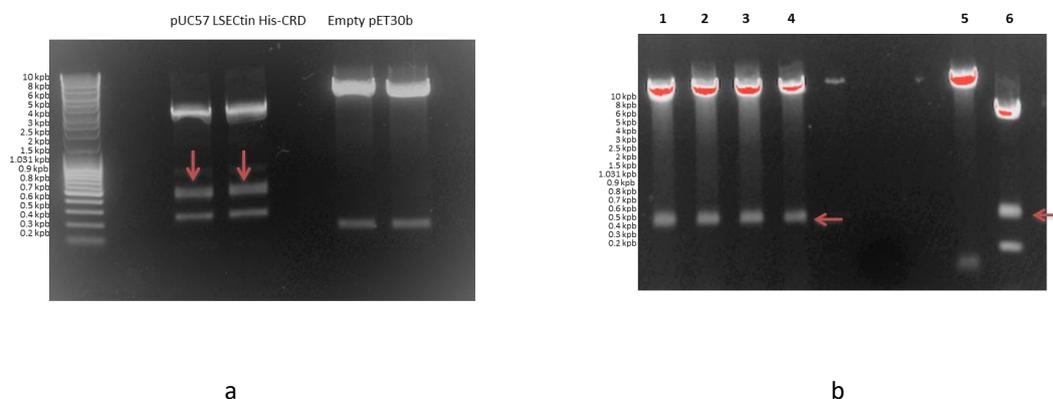


Fig.69 LSEctin His-CRD cloning. a) Results of pUC57 LSEctin His-CRD and pET30b vector NdeI/HindIII digest. The band at 450 bp corresponds to LSEctin His-CRD. b) Screening of pET30b-LSEctin His-CRD clones by NdeI/HindIII digest. 1-2-3-4 = pET30b LSEctin His-CRD minipreps, 5 = negative control: empty pET30b, 6 =

positive control: pUC57 LSEctin His-CRD.

The positive samples, characterised by presence of the 450 bp band corresponding to LSEctin His-CRD, were sent for sequencing to Genewiz (Takeley, United Kingdom).

8.2 Strategies for Recombinant Protein Expression

8.2.1 Expression as soluble folded CLR

This strategy was based on the use of ompA, pelB or DsbA protein signal sequences and the capability of the lectin constructs to express the corresponding proteins was tested in *E.coli*. Protein expression was investigated by preparing samples of total expression (whole cells **WC**), inclusion bodies (pellet **P** after cell disruption) and soluble fraction (supernatant **S** after cell disruption). Different bacterial strains were tested:

- Tuner cells, that are lacZY deletion mutants of BL21 and enables adjustable level of protein expression through a uniform IPTG induction,
- Shuffle, an engineered *E.coli* K12 that promotes disulphide bond formation in the cytoplasm, and
- conventional BL21(DE3) co-expressed with the helper plasmid pTUM4 created in Pr Arne Skurra (laboratory Technische Universitaet Muenchen, Freising-Weihenstephan, Germany), which contains four periplasmic chaperones and folding catalysts (the thiol-disulphide oxidoreductases DsbA and DsbC that catalyze the formation of disulphide bridges and the peptidyl-prolyl cis/trans-isomerases with chaperone activity, FkpA and SurA)[161] (Fig.70).

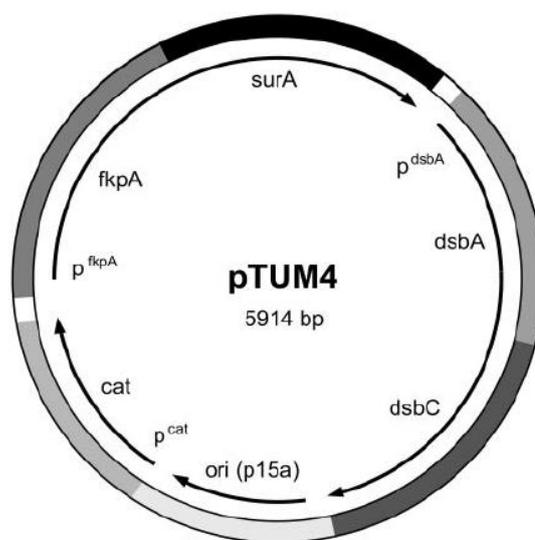


Fig.70 pTUM4 helper plasmid, from [161].

Different plasmids were used according to the strain considered:

- *pASK-IBA6* expression plasmid, that already contained the *ompA* signal sequences was used with BL21(DE3) expression systems.
- *pET30b* vector was used for Tuner cells with constructs containing a DsbA or *pelB* signal sequence.
- *pASK-IBA7+* for cytoplasmic expression in SHuffle strain with no signal peptide was exploited.

Considering the number of strains, constructs and lectins tested, it is not possible to exhaustively present all the results of soluble expression attempts. Therefore, only the example of Dectin-2 ECD will be presented for all the constructs/conditions tested. Besides, it is the only lectin for which there might be future exploitability.

- As previously mentioned, *pASK6 ompA* construct was employed for expression in BL21(DE3)-*pTUM4* strain. Different temperatures (37, 20 and 25 °C) were tested as periplasmic expression can be sensitive to this parameter, with an optimal temperature around 25 °C.

SDS-PAGE results indicate that dectin-2 *ompA* Strep-ECD strain (23.3 kDa, when expressed into periplasm: 21.3 kDa) is expressed in high quantities in the insoluble fractions (red arrow) (Fig.71a). Anti-Strep western blot was also performed and confirmed the absence of the protein in the soluble fraction (Fig.69b).

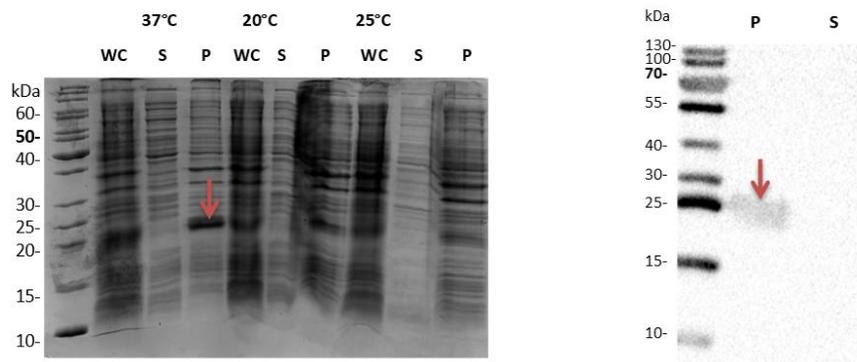


Fig.71 Dectin-2 *ompA* Strep-ECD expression. a) SDS PAGE analysis of *pASK6* dectin-2 *ompA* Strep-ECD expression in BL21(DE3)-*pTUM4* strain (23.3 kDa, when expressed into periplasm: 21.3 kDa). WC=whole cell, P=pellet, S=supernatant b) Fig. Anti-Strep western blot analysis of dectin-2 *ompA* Strep-ECD expression in BL21(DE3)-*pTUM4* strain (23.3 kDa when expressed into periplasm: 21.3 kDa) . P=pellet, S=supernatant. The red arrows indicate the band corresponding to dectin-2 *ompA* Strep-ECD.

- *pASK7+* constructs were then employed for soluble expression in SHuffle strain cytoplasm. Different temperatures (30 °C for 3h or 16 °C overnight) were tested (Fig.72).

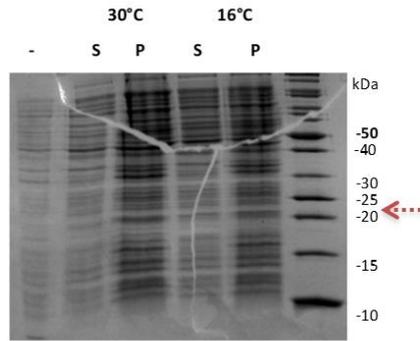


Fig.72 Dectin-2 Strep-ECD expression. SDS-PAGE analysis of pASK7+ dectin-2 Strep-ECD expression in SHuffle (21.4 kDa). - =before induction, P=pellet, S=supernatant. The red dotted arrow indicates the band where dectin-2 Strep-ECD should have been.

In this case, SDS-PAGE results indicate that dectin-2 Strep-ECD is not over-expressed at all, nor even in the insoluble fraction (Fig.72, red dotted arrow). That was confirmed by anti-Strep Western Blot (data not shown).

- Finally, pET30b vectors with constructs containing DsbA (data not shown) or pelB signal peptides were used for expression in Tuner strain. Expression of pET30b dectin-2 pelB His-ECD (23.26 kDa, when expressed into periplasm: 21.05 kDa) at different temperatures (37 and 22 °C) (Fig.73).

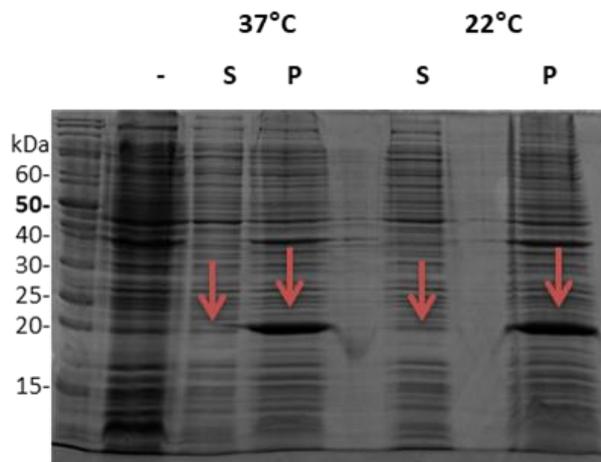


Fig.73 Dectin-2 pelB His-ECD expression. SDS-PAGE analysis of pET30b dectin-2 pelB His ECD (23.26 kDa, when expressed into periplasm: 21.05 kDa) expression in Tuner. - =before induction, P=pellet, S=supernatant. The red arrows indicate the band corresponding to dectin-2 pelB His ECD.

All the constructs appeared to express the corresponding CLRs in low quantities in the insoluble fraction (Fig.73, red arrows). Western blot was performed on pET30b dectin-2 pelB His-ECD expression in Tuner cells and revealed the presence of the protein in the soluble fraction (Fig.74, red arrows).

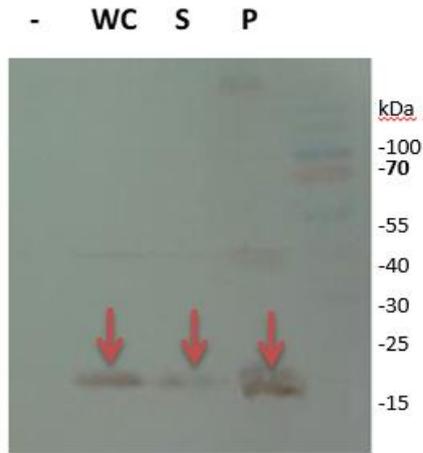


Fig.74 Dectin-2 pelB His-ECD Western Blot. Anti-His Western Blot analysis of pelB dectin-2 His-ECD expression in Tuner strain (23.26 kDa, when expressed into periplasm: 21.05 kDa). - = without induction WC= whole cells, P=pellet, S=supernatant. The red arrows indicate the band corresponding to dectin-2 pelB His ECD.

It was therefore decided to proceed to the purification of the soluble fraction and cells obtained from 1 L of culture were disrupted with a micro fluidizer (**Microfluidics Corp**). The obtained sample was purified over a HisTrap column. Elution was performed with 0.5 M imidazole buffer. 20 mM of imidazole was added to the buffer during the washing step in order to remove contaminants at this stage (Fig.75).

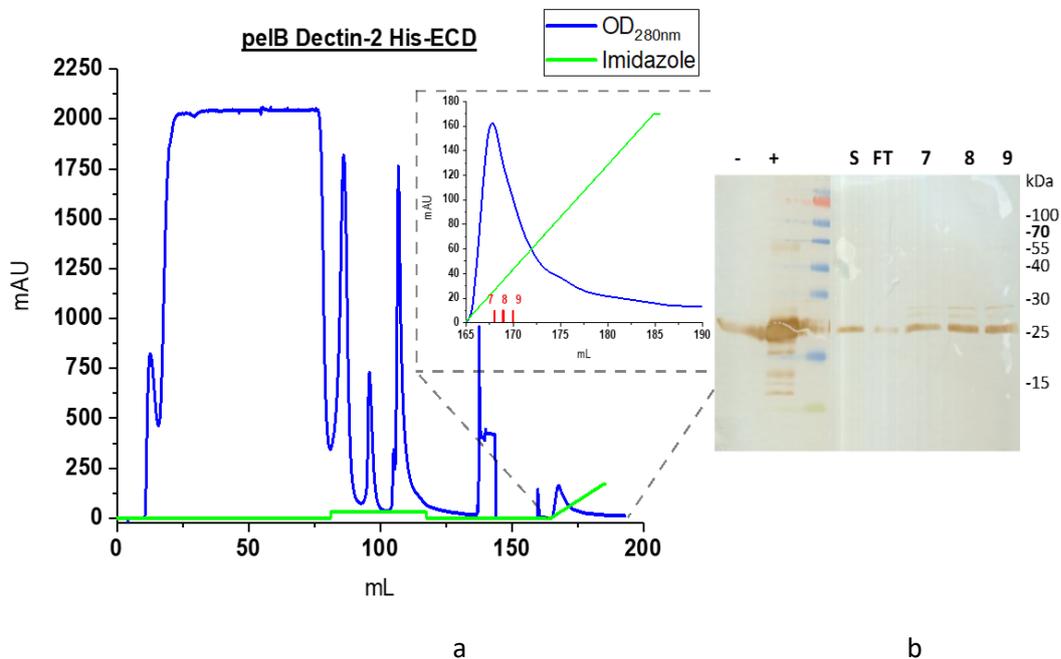


Fig.75 Dectin-2 pelB His-ECD. a) Chromatogram of dectin-2 pelB His-ECD purification on a HisTrap column and b) Western Blot analysis of pelB dectin-2 His-ECD purification (23.26 kDa, when expressed into periplasm: 21.05 kDa). Sample before (-) and after (+) induction. S=soluble fraction (before purification), FT=flow through. 7,8,9 = elution fractions.

pelB dectin-2 His-ECD is the only construct targeting periplasmic expression that was sufficiently expressed to be purified. However, the amount of protein obtained was too low (few micrograms). Moreover, functionality assay on a Mannan column was performed and did not allow the recovery of any functional protein. At that time, other functionality assays were not available in the laboratory (e.g. LECprofile assay). Being in the impossibility to assess dectin-2 ECD functionality, we did not pursue with this strategy.

Indeed, despite the large range of constructs tested, most of the approaches investigated for functional expression led to either low expression level or insoluble expression. Therefore, the alternative strategy that consists in high yield production of insoluble proteins in inclusion bodies and their subsequent refolding was considered and periplasmic expression strategy was stopped. Nevertheless, periplasmic expression optimisation will still be explored latter on in the group. Table 11

Construct	Promotor	Cell	Culture Condition	Induction	Expression	Sol/Insol Expression
pASK6 ompA Strep-ECD BDCA2/DC-SIGNR/Dectin-1/ Dectin-2/MCL/Mincle	Tet	BL21 (DE3) + pTUM4	O/N at 22°C	Anhydrotetracycline 200mg/ml	+++ (I)	insol. (II)
pASK6 ompA His-ECD BDCA2/Dectin-1/Dectin-2/MCL/Mincle	Tet	C41 (DE3)	3h at 37°C O/N at 20°C O/N at 25°C	Anhydrotetracycline 200 mg/ml	++ (I)	insol. (III)
		BL21 (DE3) + pTUM4	37 °C at 3h O/N at 20 °C O/N at 25 °C	Anhydrotetracycline 200 mg/ml	++ (I)	insol. (III)
pET30b pelb Strep-ECD BDCA2/Dectin-1/Dectin-2 Dectin-1	T7	BL21 (DE3) + pTUM4	O/N at 22°C	DE3IPTG 1 mM	-/+ (III)	Dectin 1: soluble (III) Dectin 2 : insol. (III) BDCA2:insol. (III)
		Tuner (DE3)	O/N at 22°C	IPTG 0.1 mM IPTG 0.5 mM	- (III)	
pET30b DsbA Strep-ECD Dectin-1/Dectin-2 Dectin-1	T7	BL21 (DE3) + pTUM4	O/N at 22°C	IPTG 1mM	-/+ (III)	Dectin 1: soluble (III) Dectin 2 : not tested (III)
		Tuner	O/N at 22°C	IPTG 0.1 mM IPTG 0.5 mM	- (III) - (III)	
pET30b pelb His-ECD Dectin-1/Dectin-2 Dectin-1	T7	BL21 (DE3) + pTUM4	O/N at 22°C	IPTG 1 mM	-/+ (III)	Dectin 1: soluble* (III) Dectin 2 : soluble** (III)
		Tuner (DE3)	O/N at 22°C	IPTG 0.1 mM IPTG 0.5 mM	- (III) - (III)	
pET30b DsbA His-ECD Dectin-1 Dectin-1	T7	BL21 (DE3) + pTUM4	O/N at 22°C	IPTG 1 mM	-/+ (III)	Dectin 1: soluble (III)
		Tuner (DE3)	O/N at 22°C	IPTG 0.1 mM IPTG 0.5 mM	- (III) - (III)	
pASK7+ Strep-ECD Dectin-2	T7	SHuffle	3h at 20°C	IPTG 1 mM	- (III)	
			O/N at 16°C	IPTG 1 mM	- (III)	

Table.11 Summary of all the attempts that have been made for soluble expression. Analysis of the fractions has been made as follows: I SDS-PAGE analysis performed on fractions before and after induction; II SDS-PAGE analysis performed on soluble (supernatant) or insoluble (pellet) fractions, III anti-Strep or anti-His Western Blot analysis performed on soluble (supernatant) or insoluble (pellet) fractions. * The soluble fraction was purified over a His-Trap column, an anti-His Western blot of the purified fractions revealed very dirty samples and poorly concentrated. ** The soluble fraction was purified over a His-Trap column, a anti-His Western blot of the purified fractions was really satisfactory, however the subsequent Mannan-Agarose column did not retain the protein.

8.2.2 EXPRESSION AS INCLUSION BODIES

All the below mentioned CLRs have been cloned into pEt30b vector and expressed in BL21(DE3) *E.coli* strain. Cells obtained from 1 L of culture were disrupted by sonication and inclusion bodies were collected and washed by several steps of ultracentrifugation. Finally, a 6 M Gdn-HCl buffer was used for their solubilisation and refolding was achieved through the flash dilution of the protein sample (at 2 mg/ml) into 25 mM Tris-HCl pH 8, 1.25 M NaCl, 25 mM CaCl₂ refolding buffer. Dialysis was then performed to remove the Gdn-HCl. Variations of the above-mentioned refolding conditions will be precised when required in the following sections.

The presence of the protein at each step was assessed by SDS-PAGE analysis in reductive (with β -mercaptoethanol) or in non-reductive (without β -mercaptoethanol) conditions. The absence of β -mercaptoethanol allowed the visualization of inter protein molecular disulphide bridges that could result from incorrect *in vitro* refolding.

8.2.2.1 ECD constructs

Among the many constructs analysed, three different examples of ECD purification strategies will be detailed:

-**DC-SIGNR ECD** construct was designed without any affinity tag. From the literature, mannose appeared as a suitable high affinity ligand for DC-SIGNR[63]. Therefore, Mannan affinity column, already available in the laboratory, was used for its purification.

-**Dectin-2 ECD** also binds mannose type ligands and, recently, Man α 1-2 has been identified as the minimum binding epitope [78]. However, all the attempts of purification through Mannan column were unsuccessful and we concluded that the affinity of dectin-2 for the Mannan column was probably weaker than that of DC-SIGNR. For this reason, Strep-tagged dectin-2 ECD construct was expressed and purified over a StrepTactin column. Besides, the Strep tag would enable the oriented functionalization of dectin-2 onto a SPR sensor chip for direct interaction studies.

-**Mincle ECD** interacts with cord factor and no affinity column was available for its purification. Mincle ECD was constructed with a His tag and expressed and purified over a HisTrap column.

8.2.2.1.1 DC-SIGNR ECD over-expression and purification results

DC-SIGNR ECD (37.28 kDa) shares 77% of identity with DC-SIGN, therefore refolding and purification were performed following the protocol already established for DC-SIGN ECD [57]. DC-SIGNR ECD was overexpressed in BL21(DE3) strain in good quantity (Fig.76, red arrow).

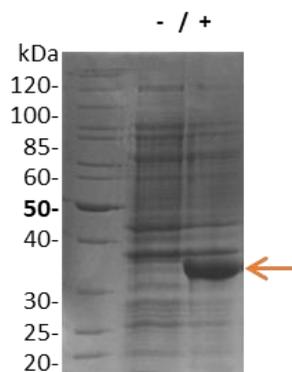


Fig.76 DC-SIGNR ECD expression. SDS-PAGE (12%) analysis of DC-SIGNR ECD (37.28 kDa, red arrow) expression in BL21(DE3) cells. Sample before (-) and after (+) induction with 1 mM IPTG.

The glycan binding properties of this lectin and its calcium requirement were used to perform Mannan-Agarose affinity chromatography with an elution using EDTA buffer that allowed the recovery of functional proteins. Several runs of purification over an exclusion chromatography (Superose-12 column) were performed with the sample eluted from the Mannan-Agarose column to restore the Ca^{2+} buffer and eliminate aggregated proteins. The two steps purification was performed in an automated mode thanks to an Akta-Xpress platform (MP3 platform of the Institute). The sample eluted from the Mannan-agarose column is stored transiently in a Superloop and re-injected by 2 mL fractions onto the Superpose 12 (column vol. : 120 mL). Thus, 4 to 5 consecutive injections are required to process all the DC-SIGNR sample eluted from the Mannan-agarose column. An example of the Mannan-Agarose and Superose 12 column coupled purifications is shown in Fig.77a. All steps of purification were monitored by SDS-PAGE (Fig.77b).

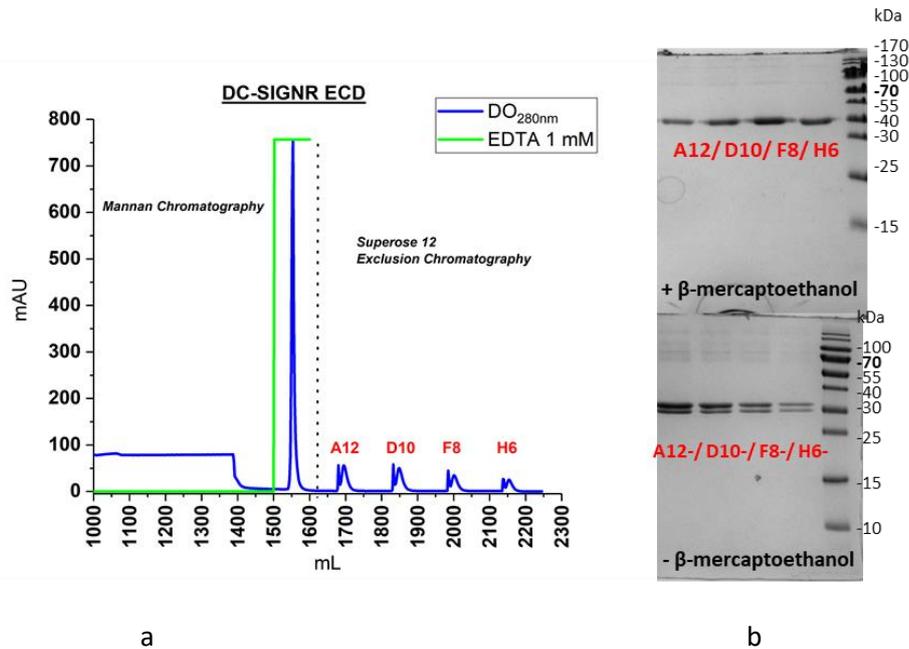


Fig.77 DC-SIGNR ECD purification (a) Chromatogram of DC-SIGNR ECD purification on a Mannan-Agarose column coupled to four consecutive Superose 12 columns ab) the SDS-PAGE analysis of DC-SIGNR ECD purification (37.28 kDa) in reductive (+ β -mercaptoethanol) /non-reductive conditions (- β -mercaptoethanol).

In the SDS-PAGE without β -mercaptoethanol, a partial reduction of disulphide bridges was observed and explains the presence of two bands of close molecular weight.

Generally, the quantity of protein purified from 1L of culture reaches the very satisfactory yield of 10 mg, although far from what was obtained for DC-SIGN ECD (50 mg).

DC-SIGNR ECD produced through this protocol has been used in interaction studies that have led to publications 1 (8.1.3), 2 (8.2.1.1) and 3 (8.2.2.1).

8.2.2.1.2 Dectin-2 Strep-ECD over-expression and purification results

Dectin-2 Strep-ECD (21.4 kDa) contains a N-terminal Strep Tag and was overexpressed in BL21(DE3) in good quantities (Fig.78, red arrow).

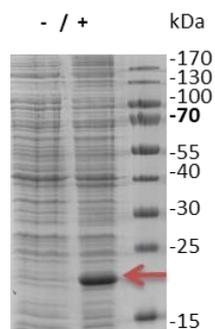


Fig.78 Dectin-2 Strep-ECD expression. SDS-PAGE analysis of dectin-2 Strep-ECD expression in BL21(DE3) (21.4

kDa, red arrow). Sample before (-) and after (+) induction.

The refolding was achieved using a starting optimised protein concentration of 1 mg/mL of protein in 6M Gdn-HCl. Refolding buffer optimization was attempted using two concentrations of Tris-HCl pH 8 buffer (25 and 200 mM) without any improvement. The refolded proteins were purified over a StrepTactin column and eluted with D-desthiobiotin (Fig.79).

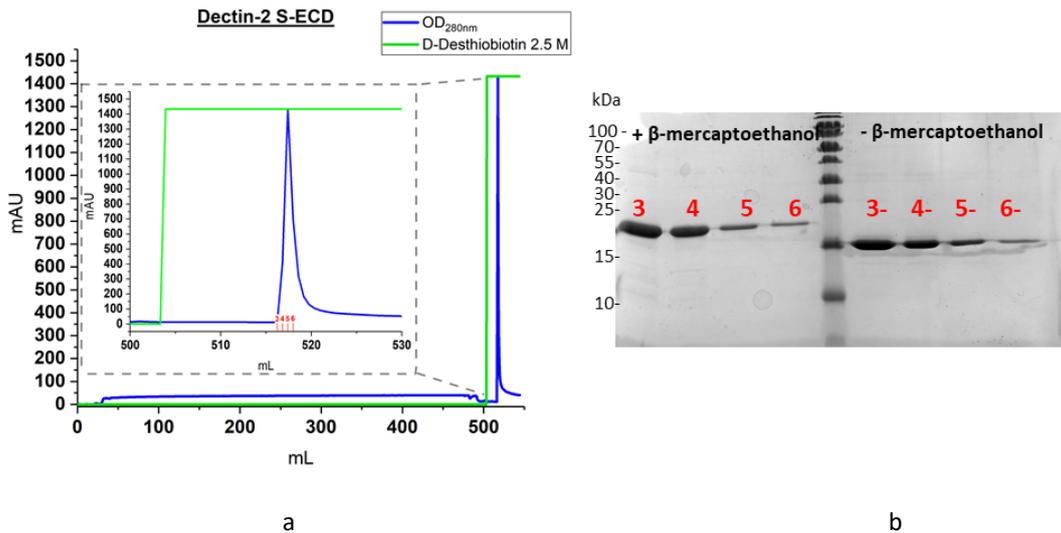


Fig.79 Dectin-2 Strep-ECD purification a) Chromatogram of dectin-2 Strep-ECD purification on a StrepTactin column b) SDS-PAGE analysis of dectin-2 Strep-ECD purification. Reduced (+ β -mercaptoethanol) and non-reduced (- β -mercaptoethanol) (21.4 KDa).

Notably, when the disulphide bridges are maintained, the protein maintains the globular shape and can run faster compared to when it is not elongated (condition with β -mercaptoethanol). Indeed, the SDS-PAGE in the absence of β -mercaptoethanol showed an apparent MW of the protein that is 5 kDa lighter than the actual MW (Fig.79b).

Dectin-2 Strep-ECD is obtained in poor quantity (0.5 mg/L culture) but with high purity. This is mainly due to a very low yield of the refolding step.

Finally, LectPROFILE platform from GLYcoDIAG Company (section 6.2.1) was used to assess the functionality of the protein. Man-BSA, Glc-BSA and Thyroglobulin were used as positive controls while Gal-BSA was the negative one, as dectin-2 is supposed to bind mannose type ligands (Fig.80).

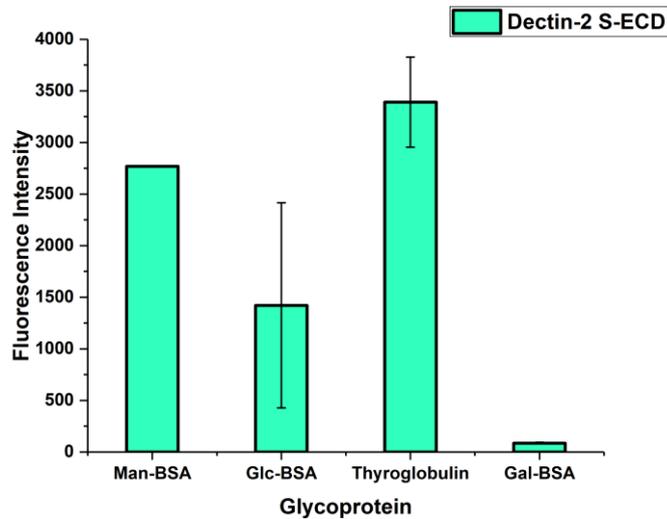


Fig.80 Dectin-2 Strep-ECD LECprofile analysis. Dectin-2 was used at a concentration of 20 µg/mL and 80 µg/mL concentration was used for the incubation with glycoproteins Man-BSA, Glc-BSA and Gal-BSA and the neoglycoprotein Thyroglobulin.

The experiment proved the functionality of dectin-2 S-ECD by showing a strong interaction with mannose-type glycoproteins and no interaction with galactose-BSA.

Dectin-2 S-ECD produced through this protocol have been used in studies that have led to publication **3** (8.2.2.1).

8.2.2.1.3 Mincle His-ECD over-expression and purification results

Mincle His-ECD (22.34 kDa) was expressed with a N-terminal His tag in BL21(DE3) in high quantity (Fig.81, red arrow).

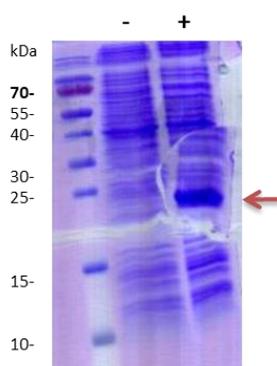


Fig.81 Expression of mincle His-ECD. SDS-PAGE analysis of Mincle His-ECD over-expression in BL21(DE3) (22.34 kDa, red arrow). Sample before (-) and after (+) induction.

The refolding was achieved using a starting protein concentration of 1 mg/mL and the following refolding buffer: 200 mM Tris-HCl pH 8, 1,25 M NaCl, 25 mM CaCl₂.

The refolded protein was purified over a HisTrap column with an elution by an imidazole containing buffer (data not shown), followed by an exclusion chromatography *via* Toyopearl HW 50S column (Fig.82).

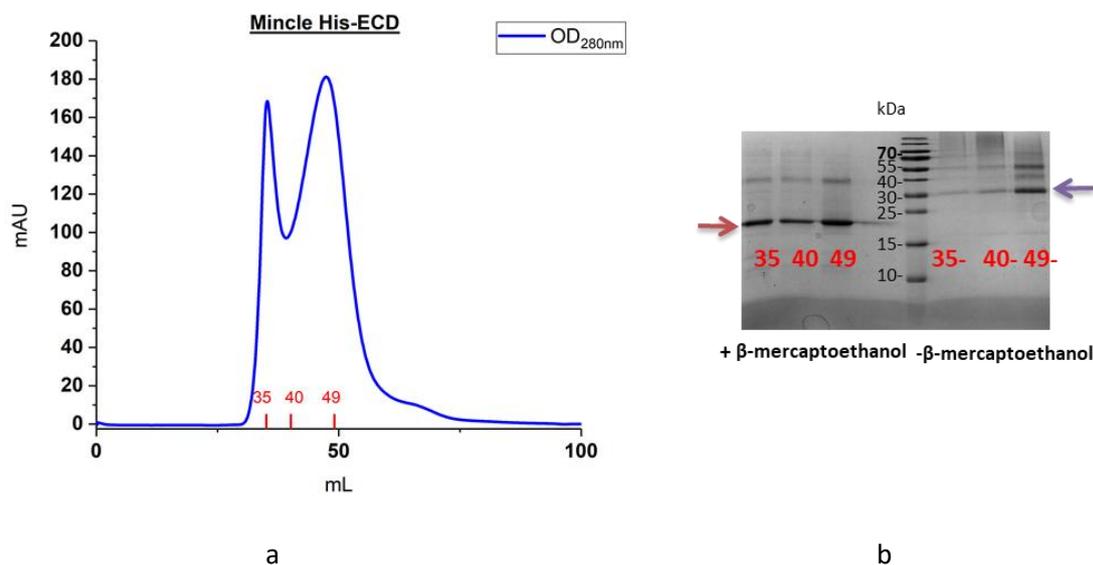


Fig.82 Mincle His-ECD purification a) Chromatogram of Mincle His-ECD exclusion chromatography purification on a Toyopearl column and b) the SDS-PAGE analysis of Mincle His-ECD purification (22.34 kDa) in reductive (+ β-mercaptoethanol) /non-reductive conditions (- β-mercaptoethanol). The red arrow indicate the monomeric mincle-His-ECD, the violet arrow the hypothetical dimer.

The non-reductive SDS-PAGE analysis indicated that the first peak (fr.35, red arrow) mainly contained aggregated protein, while the second one (fr.49, violet arrow) contained dimeric protein. As mentioned above, in the absence of β-mercaptoethanol, the protein maintains its globular shape and, therefore, on SDS-PAGE, it would run faster. In this case, the dimer would run up to a MW of 44 kDa. However, in oxidative condition, the monomer would be 5 kDa lighter, for the dimer this can correspond to 10 kDa.

Both peaks were recovered and 370 of protein μg/L of culture were obtained from the first peak, 1.27 mg/L of culture from the second one. Both sample were labelled with Cy3 fluorophore for FACS analysis performed during a secondment in Hannover in the laboratory of Prof. Bernd Lepenies. Unpublished results from Lepenies research group have previously shown an interaction between Group A Streptococcus (GAS) and murine Mincle. Since murine Mincle is considered as a model for the human Mincle [95], FACS analysis was used to assess whether recombinant human Mincle would interact with GAS. Briefly, Cy3-labelled Mincle His-ECD was incubated with GAS cells labelled with SYTO61 fluorophore and FACS experiment was performed. These cytofluorimeter analysis revealed an interaction only between GAS and proteins collected from the second peak (green histogram) while no interaction was detected after the incubation with the first peak collection (orange histogram) (Fig.83).

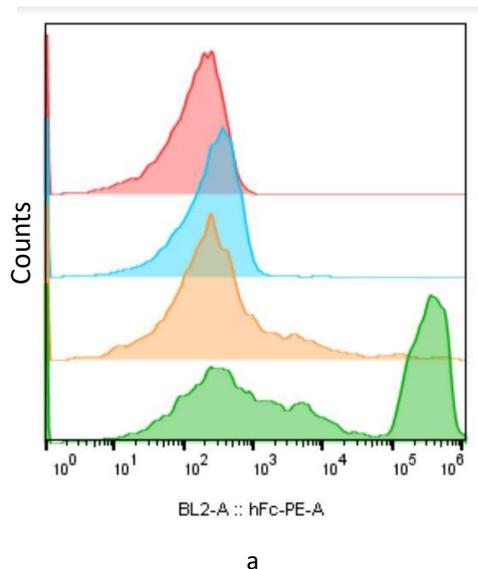


Fig.83 GAS recognition by Mincle His-ECD. Representative histograms of GAS interaction with Mincle His-ECD from two distinct elution peaks) in red SYTO61-GAS, in blue the Cy3-NeutrAvidin®, in orange Mincle His-ECD from first peak and in green Mincle His-ECD from second peak.

These results have led to the conclusion that Mincle His-ECD from the second elution peak was properly refolded and functional, while the first peak is mainly characterized by aggregated non-functional protein. Other assay will be performed to confirm the hypothesis.

8.2.2.2 CRD constructs

CRD constructs were designed after their ECD counterparts. As functional periplasmic ECD expression has been unsuccessful, CRD constructs enabling cytoplasmic expression in inclusion bodies were designed. Previous trials from our group [162] and some of our unpublished data suggest that a His Tag is more favourable than a Strep Tag for refolding attempts. Therefore, all CRD constructs bore N-terminal His Tag. The refolded proteins were purified over a HisTrap column and eluted by an imidazole buffer. Imidazole was removed by two cycles of dialysis. His Tag cleavage was enabled by the presence of the sequence specifically recognized by the protease Xa between the CRD sequence and the His-tag. Then an exclusion chromatography allowed the removal of the protease and of the N-terminal cleaved tag.

Finally, purified CRDs deleted of the His-tag were used to produce TETRALEC molecules (paragraph 3.3.2.1). First of all, they were submitted to a sortase driven biotinylation which kinetic was followed by ESI-MS. A final HisTrap column was then used to separate the His-tagged sortase from the biotinylated CRD recovered in the flow through. Purified biotinylated CRD was finally coupled to Cy3 labelled NeutrAvidin® to obtain one molecule of TETRALEC.

8.2.2.2.1 DC-SIGNR His-CRD

DC-SIGNR His-CRD (17.2 kDa) expression in BL21(DE3) cells was relatively efficient (Fig.84a) and the production yield was very good compared to DC-SIGN CRD: 20 mg of proteins/L of culture were recovered from the HisTrap purification (Fig84b and c) against 10 mg for DC-SIGN CRD.

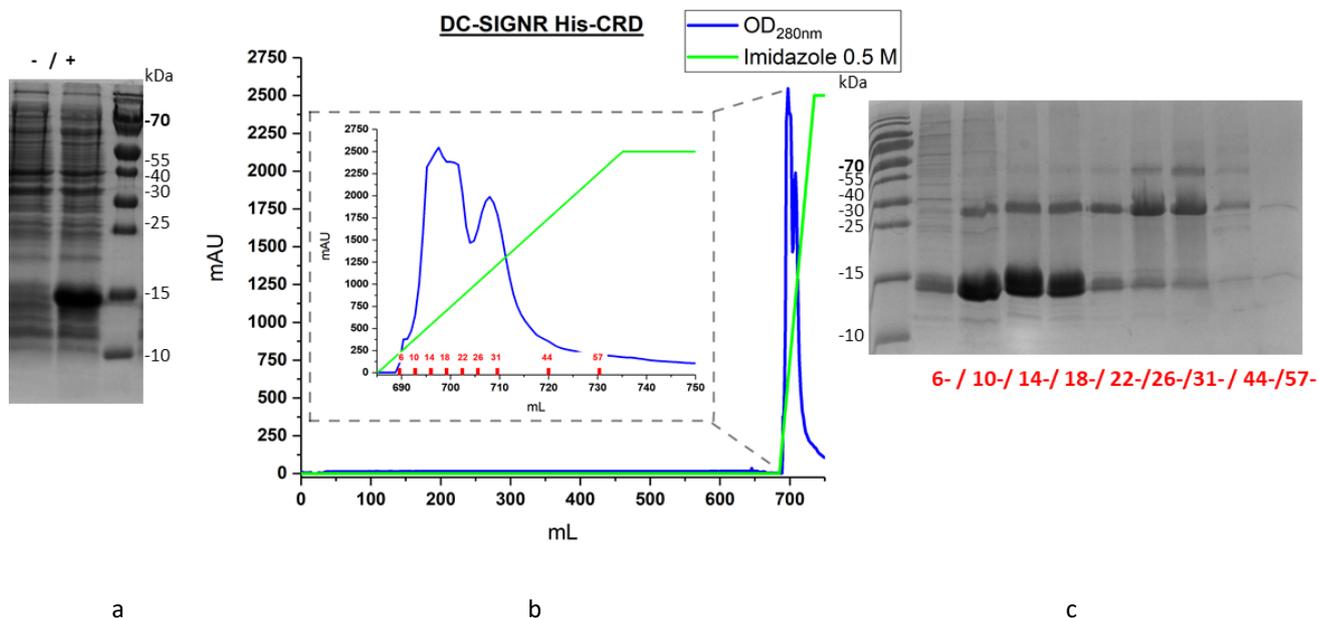


Fig.84 DC-SIGNR His-CRD expression and purification a) SDS-PAGE analysis of DC-SIGNR His-CRD expression (17.2 kDa) in BL21(DE3), sample before (-) and after (+) induction with 1 mM IPTG. b) Chromatogram of DC-SIGNR His-CRD HisTrap purification and c) SDS-PAGE analysis of DC-SIGNR His-CRD HisTrap purification in non-reductive conditions.

The imidazole gradient enabled the elution of two sequential peaks of protein. SDS PAGE analysis in non-reductive condition showed that they both contain monomeric and dimeric species. Indeed, unwanted intermolecular disulphide bridge can result in protein dimerization. Fractions 10 to 22, corresponding to the enriched monomeric form, were collected. The dimeric form is supposed to be eluted at higher imidazole concentration due to the stronger interaction of two His-tags with the His-Trap column. A first attempt to remove the His-tag was performed on 2 mg of protein in the presence of imidazole but factor Xa digestion was incomplete, proving the importance of the dialysis step (elimination of imidazole) for cleavage (Fig.85a). Factor Xa had a molecular weight of approximately 43 kDa, consisting of two disulphide-linked chains of approximately 27 kDa and 16 kDa. On SDS-PAGE, the reduced chains had apparent molecular weights of 30 kDa (Fig. 85a, blue arrow) and 20 kDa (Fig. 85a, green arrow). DC-SIGNR CRD molecular weight after His-Tag removal is 15.7 kDa (red arrow). The cleavage was followed by an exclusion chromatography performed on a Toyopearl HW 50S column (Fig.85b) and revealed the presence of aggregates in the sample (Fig.85c). 1,18 mg/L of culture were recovered at this stage. A double band is present in Fig.85c and might represent the partial uncut of the His-tag.

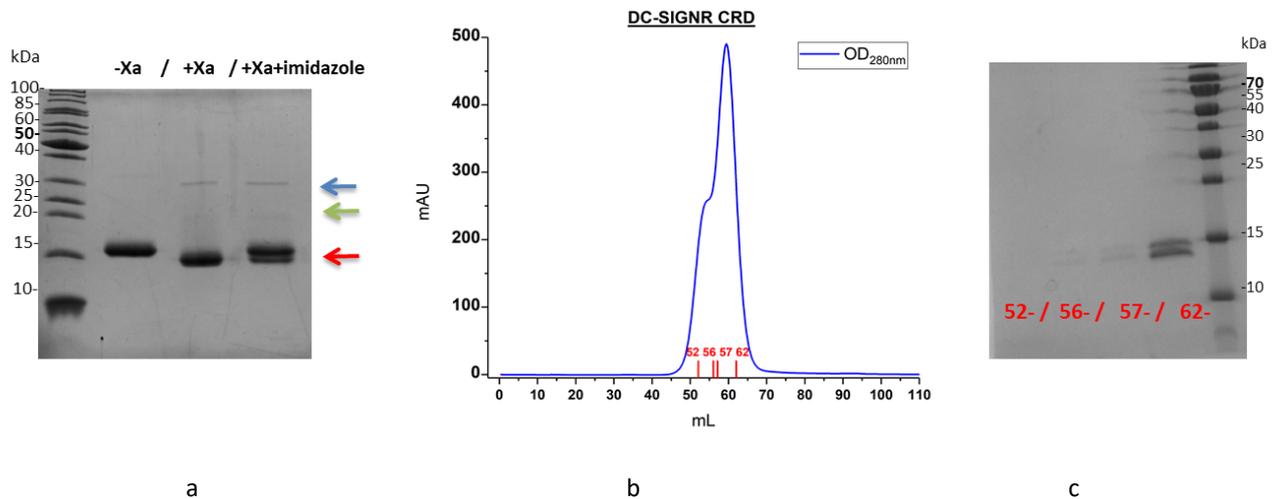


Fig.85 DC-SIGNR His-CRD His tag cleavage and purification a) SDS-PAGE analysis of the His Tag cleavage without protease Xa (-Xa), with Xa (+Xa) and with Xa in the presence of imidazole (+Xa + imidazole). The red arrow indicates DC-SIGNR CRD, the green arrow the 16 kDa chain of factor Xa, the blue arrow the 27 kDa chain of factor Xa b) Chromatogram of DC-SIGNR CRD size exclusion chromatography purification on Toyopearl HW 50S column and c) SDS-PAGE analysis of DC-SIGNR CRD size exclusion purification (15.86 KDa) in non-reductive condition.

Finally, the SrtA (28.13 kDa) catalysed reaction enabled a complete N-terminal protein labelling with the corresponding biotinylated peptide after 6h of reaction as demonstrated by ESI-MS (Fig.86a). The final purification was performed on a HisTrap column to remove the his-tagged sortase (Fig.86b). The biotinylated DC-SIGNR CRD (0.2 mg) was recovered in the flow through (red arrow) (Fig.86c). Nevertheless, the SDS-PAGE analysis of this last purification step revealed a contamination by the enzyme of DC-SIGNR CRD sample and the presence of DC-SIGNR CRD in the putative SrtA fractions (violet arrow) after imidazole elution, which indicates the presence of remaining SrtA-DC-SIGNR CRD complex.

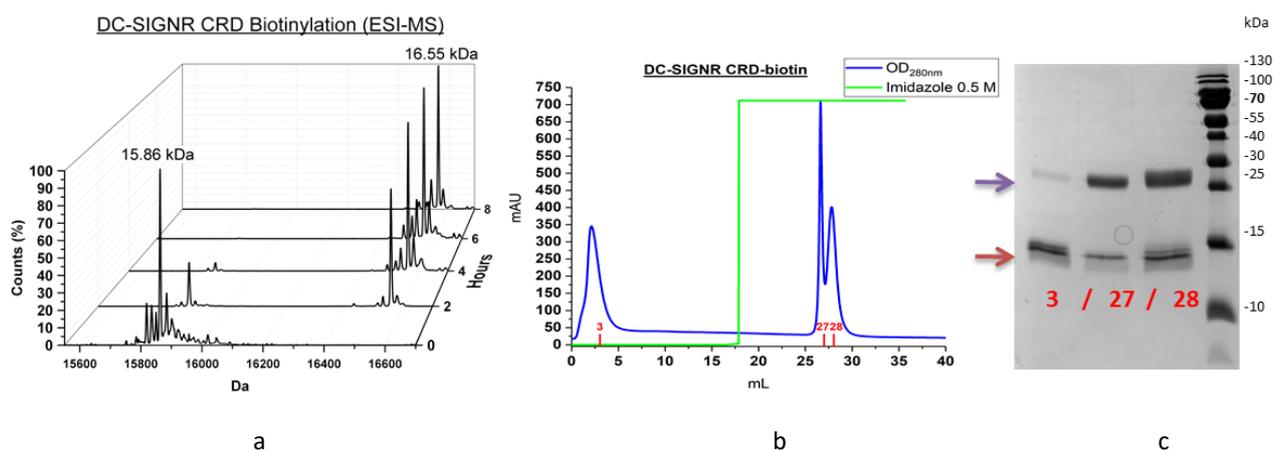


Fig.86 DC-SIGNR CRD biotinylation a) sortase reaction followed by ESI-MS. b) Chromatogram of biotin-DC-SIGNR CRD HisTrap purification and b) SDS-PAGE analysis of DC-SIGNR CRD HisTrap purification (16.55 KDa) in non-reductive condition.

The purified biotinylated DC-SIGN-CRD was then complexed via its biotin moiety to a tetramer of NeutrAvidin previously functionalized with Cy3-fluorophore (DOL 0,5), achieving the formation of TETRALEC

(0.4 mg). Some more details about the preparation and characterization of the DC-SIGNR-CRD TETRALEC complex formation will be presented in the preliminary version of publication XX in this manuscript below.

8.2.2.2.2 BDCA2 and LSEctin His-CRD

The over-expression of BDCA2 (17.11 kDa) and LSEctin His-CRD (16.42 kDa) in BL21(DE3) were very satisfactory (Fig.87a and b, red arrows).

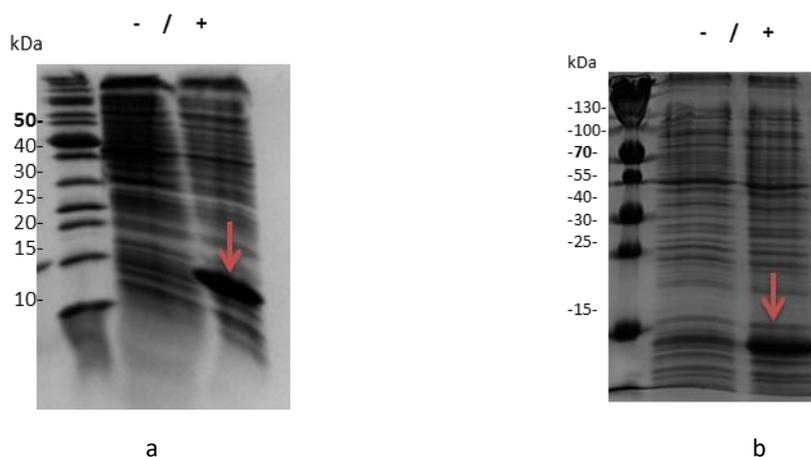


Fig.87 BDCA2 and LSEctin His-CRD expression. SDS-PAGE analysis of a) BDCA2 (17.11 kDa, red arrow) and b) LSEctin (16.42 kDa, red arrow) His-CRD over-expression in BL21(DE3), sample before (-) and after (+) induction.

BDCA2 His-CRD was purified after refolding over a HisTrap column (data not shown) with good yield of purification (4 mg/L culture). However, several assays of biotinylation performed after HisTag cleavage were unsuccessful. For some unknown reasons, the protein dramatically precipitated during the SrtA (28.13 kDa, violet arrow) reaction. The SDS-PAGE analysis of purification did not reveal any presence of biotinylated BDCA2 CRD (16.38 kDa, Fr 2/3/4) while it showed a partial enzymatic degradation (Fig.88 a and b).

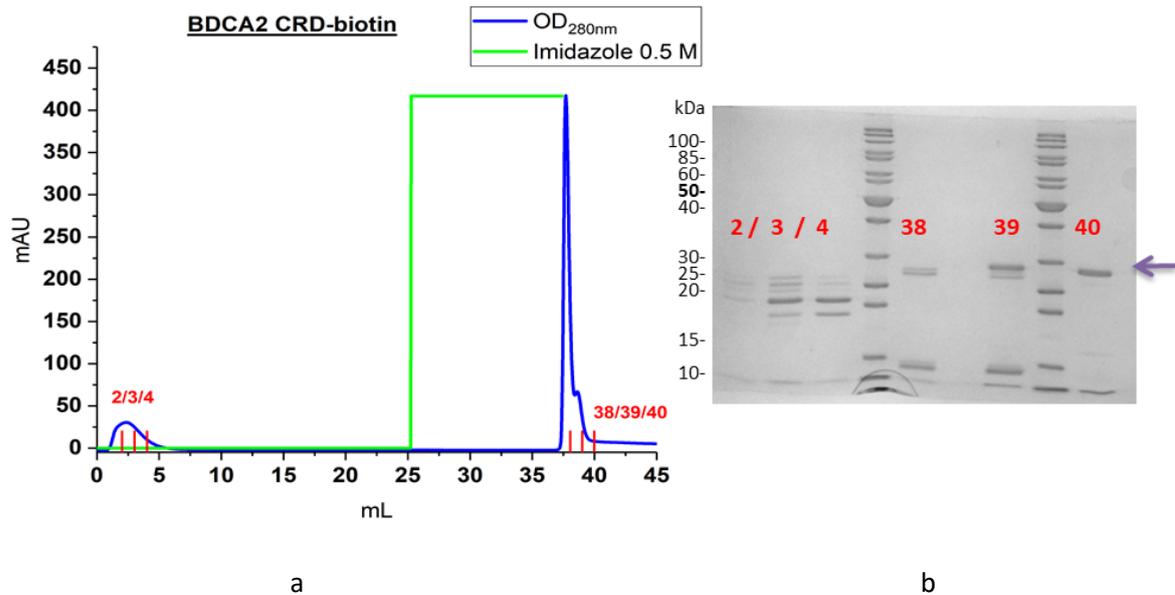


Fig.88 BDCA2 biotinylation a) Biotinylated BDCA2 CRD (16.38 kDa) from HisTrap purification and b) SDS-PAGE analysis after HisTrap purification.

Since the site-specific biotinylation, using the sortase strategy, did not function for BDCA2, an unspecific biotinylation, targeting lysine residues onto CRD surface, was performed. This strategy will produce a BDCA2 TETRALEC, with CRD randomly oriented with respect to the NeutrAvidin®. The degree of biotin labelling has been analysed and a degree of biotinylation of 2 was obtained per BDCA2 CRD. Finally, as Gal β 1-4GlcNAc β 1-2Man was shown to be a high affinity ligand binding for BDCA2 [54], a LectPROFILE assay was performed to assess BDCA2 His-CRD functionality through binding to the Gal β 1-4GlcNAc β 1-2Man of Asialofetuin (Fig.89). Man-BSA and Gal-BSA were used as negative control.

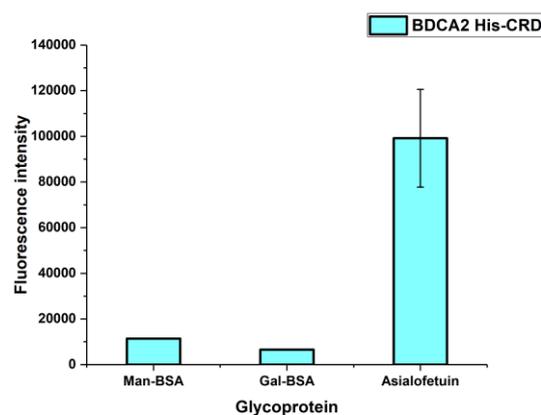


Fig.89 BDCA2 His-CRD LECTprofile analysis. BDCA2 was used at a concentration of 20 μ g/mL and 80 μ g/mL concentration was used for the incubation with glycoproteins Man-BSA, Gal-BSA and the neoglycoprotein Asialofetuin.

We started to work on **LSEctin His-CRD** by the end of my second year of PhD. Careful optimization has still to be performed on its purification. LSEctin His-CRD was refolded using 1 mg/mL starting concentration and subsequently purified on a HisTrap column (Fig.90a) with no further exclusion chromatography. However, this single step of purification was not sufficient to provide homogeneous sample, as demonstrated by the SDS-PAGE analysis of the purification (Fig.90b), which indicates the presence of both monomeric (fr.19) and dimeric species (fr.24 and 30).

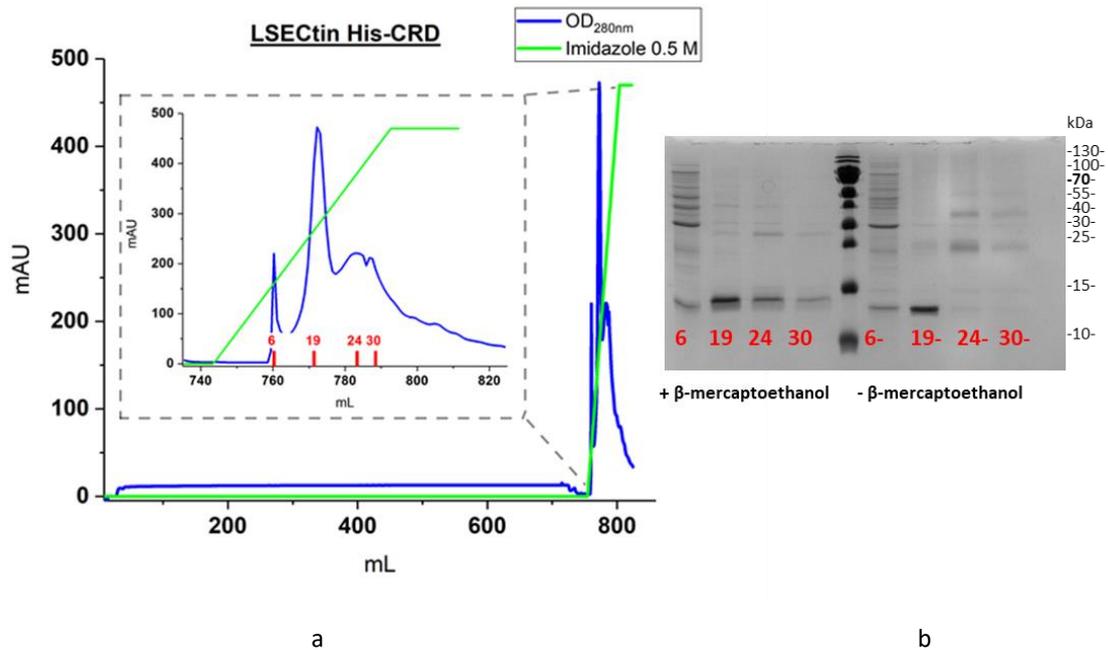


Fig.90 LSEctin His-CRD purification a) Chromatogram of LSEctin His-CRD HisTrap purification and b) the SDS-PAGE analysis of LSEctin His-CRD HisTrap purification (16.42 KDa) in reductive (+ β -mercaptoethanol)/non-reductive conditions (- β -mercaptoethanol).

The peak corresponding to the monomeric fractions (fraction 19 to 24) was collected and led to 0.8 mg of protein/L of culture. Again, LectPROFILE assay was used to assess LSEctin His-CRD functionality (Fig.91) and revealed binding mainly towards mannose-BSA but also towards GlcNac and to a lesser extend to Gal-BSA, as already shown by *Powlesland et al* study [88]. This result confirmed the correct folding and functionality of the protein.

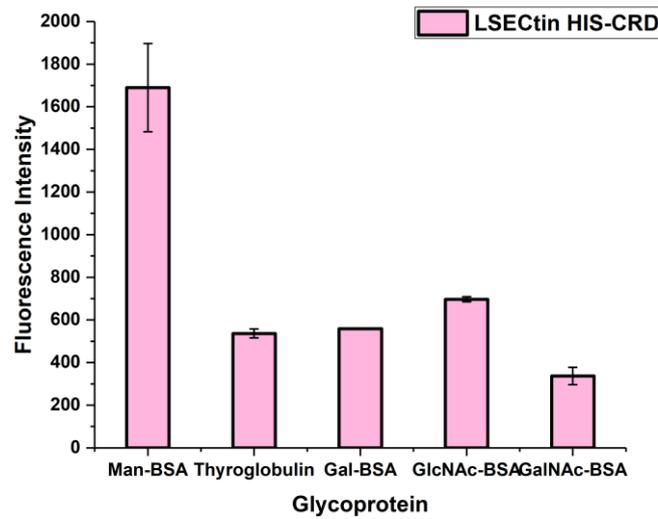


Fig.91 LSEctin His-CRD LectPROFILE analysis. LSEctin was used at a concentration of 20 $\mu\text{g}/\text{mL}$ and 80 $\mu\text{g}/\text{mL}$ concentration was used for the incubation with glycoproteins Man-BSA, Gal-BSA, GlcNAc-BSA, GalNAc-BSA and the neoglycoprotein Thyroglobulin.

No assay of site specific biotinylation was performed yet.

The LSEctin His-CRD produced here has been used in the study that led to publication 2 (8.2.1.1).

8.2.2.2.3 Dectin-1, Dectin-2, Mincle, MCL His-CRD

The over-expression of dectin-1 (18.83 kDa), dectin-2 (17 kDa), Mincle (17.91 kDa) and MCL (17.32 kDa) His-CRDs in BL21(DE3) cells were satisfactory (Fig.92), as indicated by the red arrows.

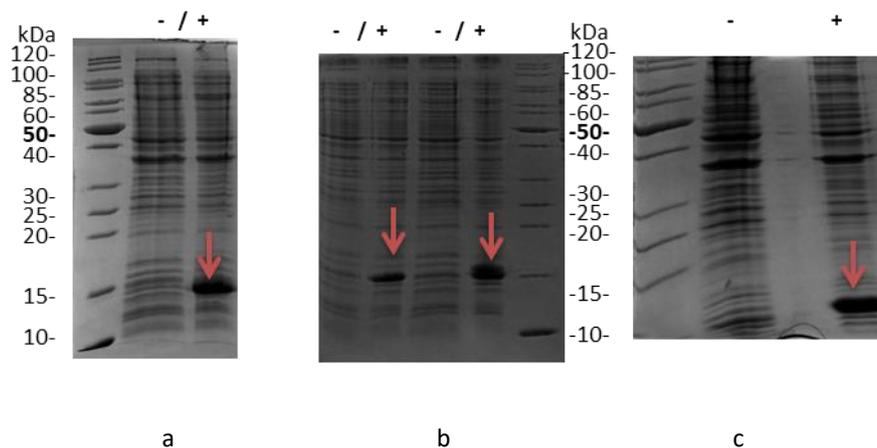


Fig.92 Dectin-1, Dectin-2, Mincle and MCL HIS-CRD expressions. SDS-PAGE analysis of dectin-1 (18.83 kDa), dectin-2 (17 kDa), Mincle (17.91) and MCL (16.42 kDa) CRD over-expressions in BL21(DE3), sample before (-) and after (+) induction. The red arrows indicate the band corresponding to the four CLR.

Dectin-1 His-CRD (18.83 kDa) was refolded using 1 mg/mL of starting protein concentration and purified on a HisTrap column (Fig.93a) with a poor yield of purification. SDS-PAGE analysis (Fig.93b) revealed the

presence of aggregated proteins and only a small fraction of contaminated monomeric dectin-1 His CRD was only present in the first elution peak (Fig93b, red arrow).

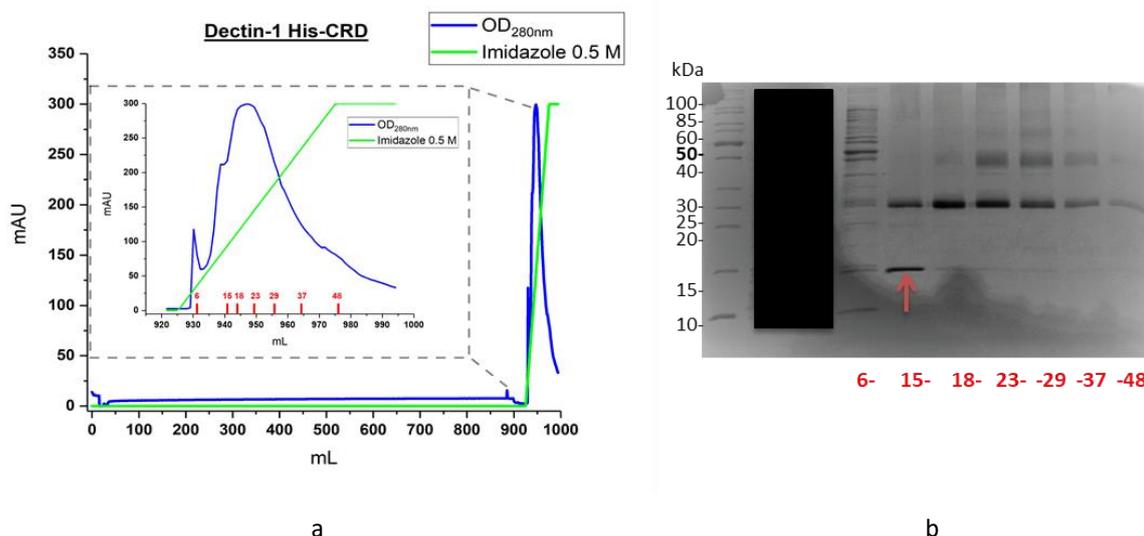


Fig.93 Dectin-1 purification a) Chromatogram of dectin-1 His-CRD HisTrap (Nickel) purification and b) SDS-PAGE analysis of dectin-1 His-CRD purification (18.83 kDa, red arrow) in non-reductive conditions.

The fractions containing mainly monomeric protein (12-25 fractions) were pooled and allowed the recovery of 1.3 mg of protein. Those fractions were labelled with Cy3 fluorophore and FACS analysis on GAS (section 8.1.1.2.1.3, Mincle His-ECD) was performed to assess the protein functionality. As for Mincle, unpublished results from Lepenies research group showed an interaction between GAS and murine dectin-1 and we wanted to confirm this interaction with the refolded human dectin-1 CRD that we have produced. Briefly, Cy3-labelled dectin-1 His-CRD was incubated with GAS cells labelled with SYTO61 and FACS analysis was performed (Fig.94). SYTO61-GAS and Cy3-NeutrAvidin© were used as negative control.

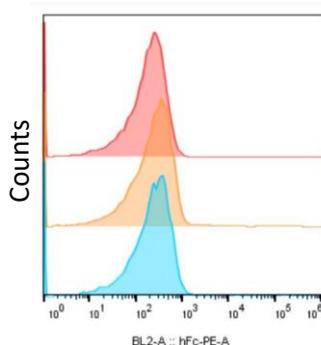


Fig.94 GAS recognition by dectin-1 His-CRD. Representative histograms of GAS interaction with dectin-1 His-CRD (in red SYTO61-GAS, in orange Cy3-NeutrAvidin, in blues dectin-1 His-CRD).

FACS analysis revealed no interaction with GAS cells. However, murine and human dectin-1 have difference on ligand recognition. Therefore, additional functionality assays need to be performed to establish the protein functionality.

Mincle His-CRD (17.91 kDa) was refolded using 1 mg/mL of starting protein concentration and purified on a HisTrap column (Fig.95a). The various elution peaks analysed by SDS-PAGE revealed the presence of monomeric protein in only two fractions (Fig.95b, red arrows). These fractions were further purified on a Toyopearl HW 50S gel filtration column (Fig.95c).

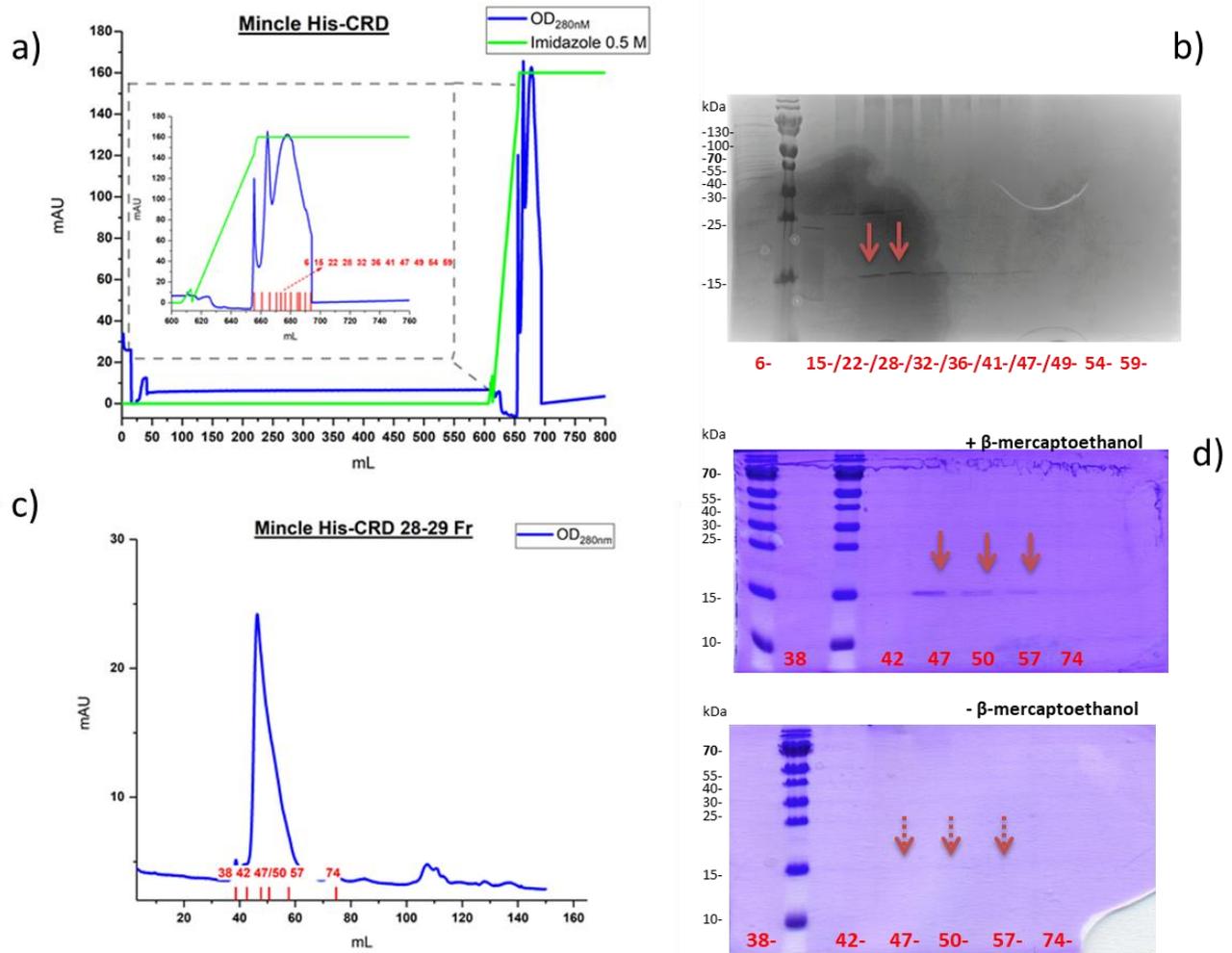


Fig.95 Mincle His-CRD purification a) Chromatogram of minclle His-CRD HisTrap purification and b) SDS-PAGE analysis of Minclle His-CRD His Trap purification (17.91 kDa, red arrows) in non-reductive conditions. c) Chromatogram of minclle His-CRD size exclusion purification on Toyopearl HW 50S column and b) SDS-PAGE analysis of minclle His-CRD size excusion purification (17.91 kDa) in reductive (+ β -mercaptoethanol, red arrows) and non-reductive conditions (- β -mercaptoethanol, red dotted arrows).

The SDS-PAGE analysis in reductive conditions showed the presence of monomeric minclle His-CRD species (Fig.95d, red arrows), which surprisingly disappeared in the absence of β -mercaptoethanol (Fig.95d, red dotted arrows). This point and the elution volume from the Toyopearl size exclusion column suggest that minclle His-CRD is present only in an aggregate form, probably implying intermolecular disulfide bridges, and is therefore not suitable for further analysis since not properly folded.

Dectin-2 His-CRD (17 kDa) seemed to be more promising than dectin-1 and mincle His-CRD. The imidazole elution from the HisTrap column revealed the presence of monomeric dectin-2 His-CRD (red arrows), with quite a good yield of purification (8 mg/L of culture) (Fig.96).

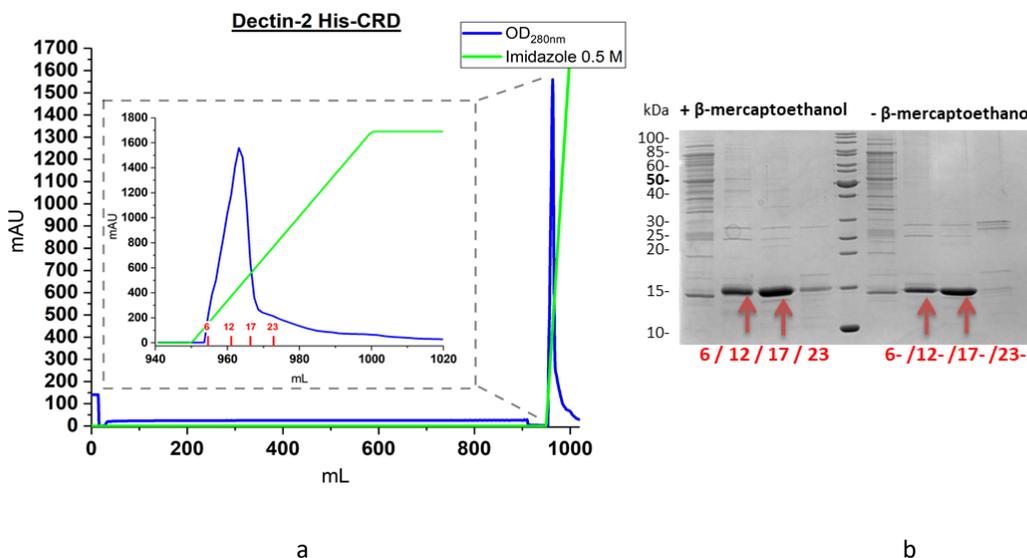


Fig.96 Dectin-2 His-CRD purification a) Chromatogram of Dectin-2 His-CRD HisTrap purification and b) SDS-PAGE analysis of Dectin-2 His-CRD HisTrap purification (17 kDa, red arrows) in reductive and non-reductive conditions.

However, the SDS-PAGE condition without β -mercaptoethanol did not reveal the expected shift of MW, leading to the hypothesis that problem have occurred at S-S level. LectPROFILE assay was then used to assess the functionality of dectin-2 His-CRD. Man-BSA, Glc-BSA and GlcNAc-BSA were awaited to be recognised while Gal-BSA was used as a negative one (Fig.97).

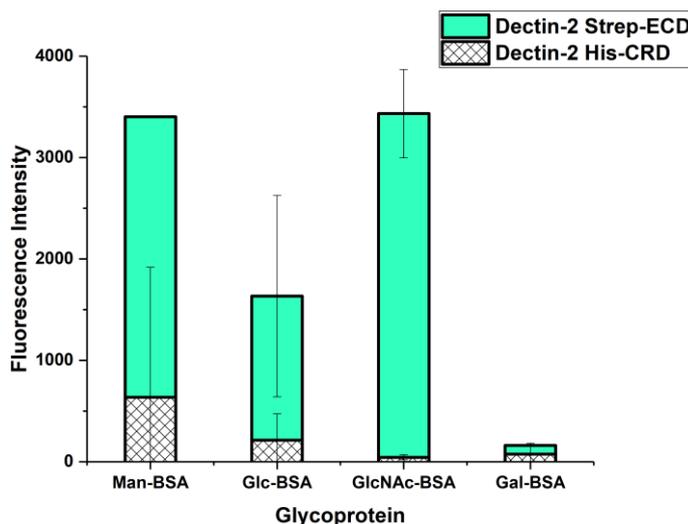


Fig.97 Comparison between Dectin-2 His-CRD and Dectin-2 Strep-ECD LectPROFILE analysis. Dectin-2 His-CRD was used at a concentration of 100 μ g/mL and 80 μ g/mL concentration was used for the incubation with glycoproteins Man-BSA, Glc-BSA, GlcNAc-BSA and Gal-BSA. For Dectin-2 Strep-ECD were used the same conditions of Figure 78.

Unexpectedly, no significant recognition could be observed. By comparing the LECprofile from dectin-2 His-CRD and dectin-2 Strep-ECD, only low fluorescence intensity signals were observed for the CRD interaction with the positive controls. Circular dichroism (CD) analysis was then performed to assess the dectin-2 CRD secondary structure integrity in comparison with the one of another CLR produced in the lab (DCIR) that was correctly folded (data not published) (Fig.98).

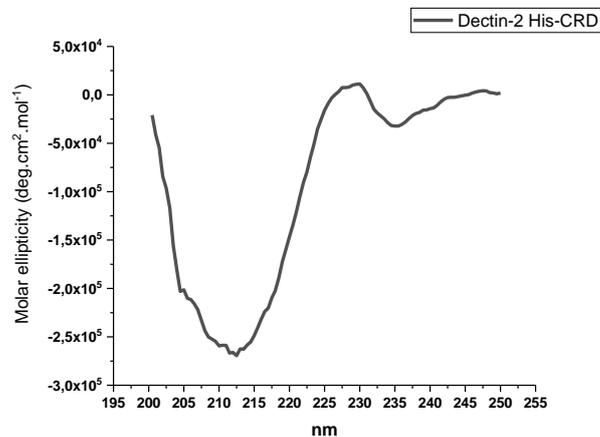


Fig.98 Dectin-2 His-CRD CD spectrum.

CD spectrum revealed a shift of the intensity towards 210 nm, which is not a typical feature of β -sheet. This might indicate the presence of unfolded protein that could explain the production of non-functional protein.

MCL His-CRD (17.32 kDa) was awaited to be difficult to handle as no ligand has been identified so far for this lectin making the assessment of the protein functionality after refolding difficult.

The purification over a HisTrap column revealed a good yield (10mg/L culture), enabling the recovery of the monomeric MCL His-CRD (Fig.99).

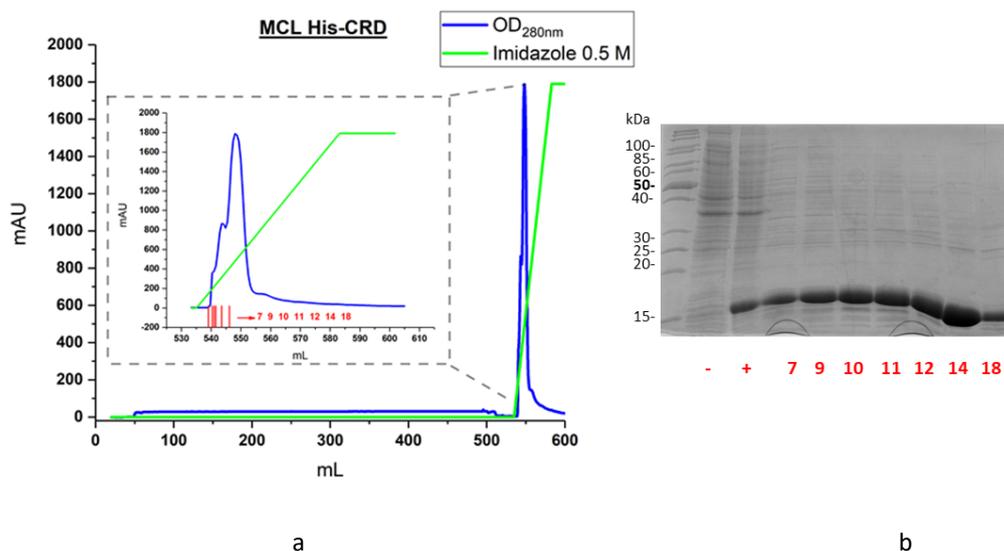


Fig.99 MCL His-CRD purification a) Chromatogram of MCL His-CRD HisTrap purification and b) SDS-PAGE analysis of MCL His-CRD HisTrap purification (16.42 kDa) in reductive conditions.

Since there is no tool to assess the protein functionality, Circular Dichroism analysis was performed to at least verify the integrity of the secondary structure (Fig.100).

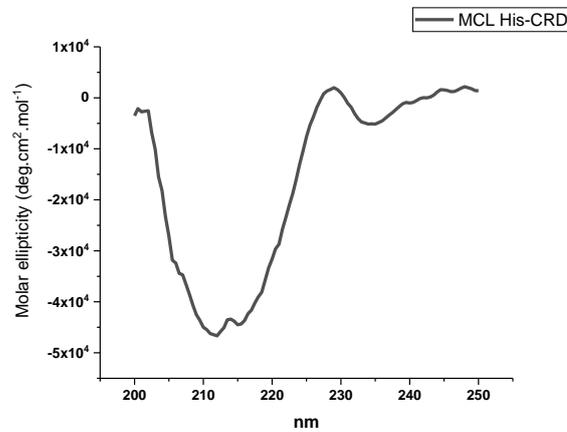


Fig.100 MCL His-CRD CD spectrum.

The CD spectra obtained confirmed the integrity of the secondary structure of MCL His-CRD, with the typical signature of β -sheets, and is fully comparable DCIR His-CRD CD spectrum (data not shown), which has the same fold. As for DC-SIGNR His-CRD, attempts were made towards the his-tag cleavage in the presence of low concentration of imidazole in order to avoid the dialysis step. However, the presence of imidazole again inhibited the cleavage, resulting in a less efficient his tag removal (Fig.101a). Purification of the cleaved protein using size exclusion column Toyopearl HW 50S revealed a very symmetrical elution peak suggesting a well-defined protein species with no aggregated form (1 mg/L of culture) (Fig.101b and c). These biochemical characterisations (CD+size exclusion) were very promising regarding the correct sample refolding and homogeneity. Indeed, despite the absence of known ligand to definitively validate its functionality, we pursued the process toward the generation of a MCL CRD-TETRALEC, a tool that could be of great help for further ligand identification.

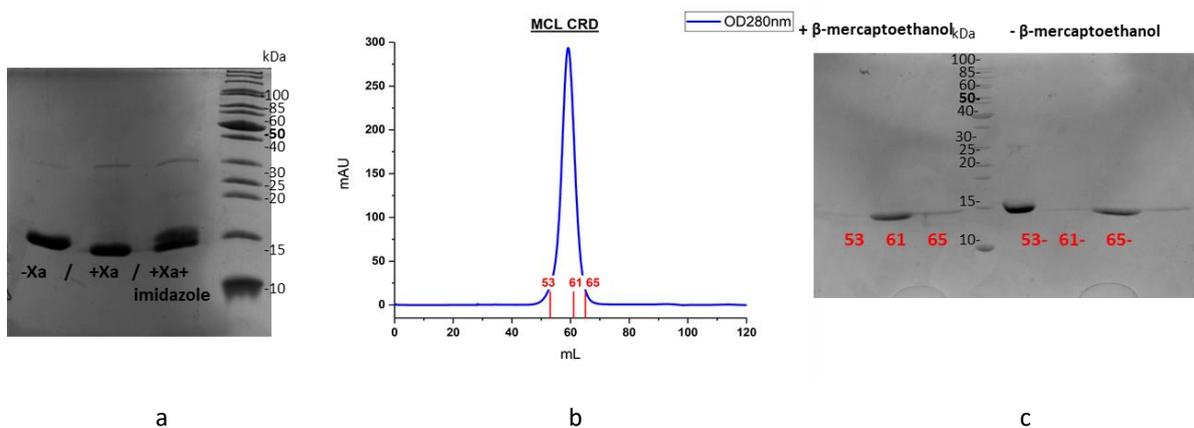


Fig.101 MCL His-CRD His tag cleavage and purification a) SDS-PAGE analysis of MCL CRD His-tag cleavage

without protease Xa (-Xa), with Xa (+Xa) and with Xa and in the presence of imidazole buffer (+Xa +imidazole)
 b) Chromatogram of MCL CRD size exclusion Toyopearl purification and c) the SDS-PAGE analysis of MCL CRD size exclusion purification (15.907 KDa) in reductive (+ β -mercaptoethanol) and non-reductive conditions (- β -mercaptoethanol).

The labelling of MCL using the orientated SrtA strategy, as used for DC-SIGNR CRD, was also successful (Fig.102a), leading to 0.55 mg of MCL CRD specifically biotinylated at its N-terminus (fraction 1, Fig.102c) after His-Trap SrtA removal (fraction 2, Fig.102b)

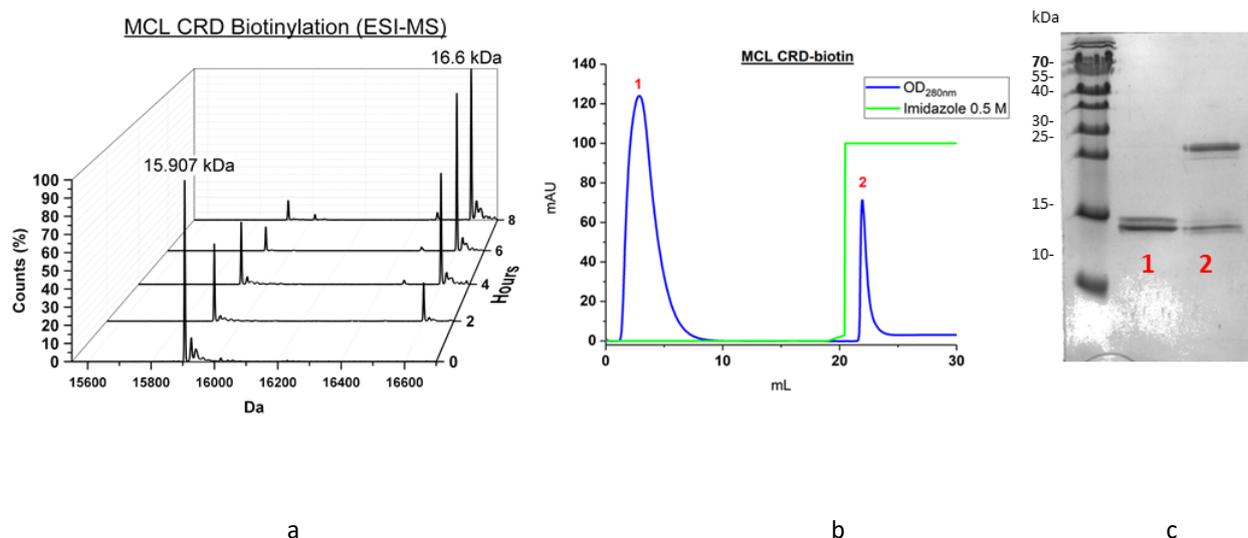


Fig.102 MCL CRD-biotinylation a) MCL-CRD biotinylation followed by ESI-MS, b) chromatogram of biotin-MCL CRD HisTrap purification and c) SDS-PAGE analysis of biotin-MCL-CRD after HisTrap purification (16.6 KDa) in non-reductive condition.

In the SDS-PAGE analysis without β -mercaptoethanol, a partial reduction of disulphide bridges was observed also in this case and explain the presence of two bands of close molecular weight. Nevertheless, other investigation should be performed.

The purified biotinylated MCL-CRD was then complexed *via* its biotin moiety to a tetramer of NeutrAvidin previously functionalized with Cy3 fluorophore, achieving the formation of MCL-TETRALEC.

Finally, Table 12 sums up of expression and yield of purification of the ECD constructs. The same table for CRD constructs is found in Table 13, with in addition the achievement of TETRALEC strategy.

ECD constructs	Expression	Purification mg/mL culture
<i>Langerin</i>	Green	7
<i>DC-SIGN</i>	Green	50
<i>DC-SIGNR</i>	Green	10
<i>Dectin-2 Strep</i>	Green	0,5
<i>Mincl His</i>	Green	1,27

Table 12. ECD expression and purification yield. Green = good, yellow = to improve

His-CRD constructs	Expression	Purification mg/mL culture	Functionality	Circular dichroism	TETRALEC	Random TETRALEC
<i>BDCA2</i>		4		Not necessary		
<i>DC-SIGNR</i>		20		Not necessary		Perspective
<i>Dectin-1</i>		1,3	?	Perspective	Perspective	Perspective
<i>Dectin-2</i>		4		✓	Perspective	Perspective
<i>LSEctin</i>		0,5		Not necessary	Perspective	Perspective
<i>MCL</i>		4	?	✓		Perspective
<i>Mincle</i>		X				

Table 13. CRD expression, purification yield, functionality assay, CD assay and TETRALEC formation. Green = good, yellow = to improve, red = not achieved.

8.3 TETRALEC, Artificial Tetrameric Lectins: a Tool to Screen Ligand and Pathogen Interactions

The following section presents a preliminary version for a future article (target Journal: Glycobiology) investigating the binding enhancement enabled by the TETRALEC multivalency. Briefly, two natural and two artificial constructs of DC-SIGNR were considered: the natural tetrameric DC-SIGNR ECD and the monomeric DC-SIGNR CRD, the artificial tetrameric DC-SIGNR-TETRALEC and a dimeric DC-SIGNR-Fc fusion protein. SEC-MALS analysis was performed in order to confirm their oligomerisation level while mannose-based glycan microarrays were used to assess the impact of multivalency on ligand recognition. Finally, FACS analysis on *Candida albicans* proved the functionality of DC-SIGNR-TETRALEC by entire cell recognition and show the discovery of a new CLR/pathogen interaction.

The main outcome of these studies:

Site specific TETRALEC strategy can be used to enhance multivalency at protein level. In addition, it was shown for the first time the interaction between DC-SIGNR and *Candida albicans*.

Contributions:

SEC-MALS analysis has been conducted with the support of the SEC-MALS analysis (PAOL platform in our laboratory), while glycan array were performed in Dr. N. Reichardt's laboratory. Finally, Joao Monteiro, a PhD student in B. Lepenies laboratory in Hannover, performed FACS analysis. They also produced DC-SIGNR-Fc fusion protein.

My contribution to this study:

I have prepared DC-SIGNR ECD, DC-SIGNR CRD and DC-SIGNR-TETRALEC.

I have performed the glycan array experiments and contributed to the result interpretation.

I have participated to the FACS experiments and result interpretation.

I have produced the figures.



TETRALEC, Artificial Tetrameric Lectins: a tool to screen ligand and pathogen interaction

Silvia Achilli¹, Joao Monteiro², Sonia Serna³, Michel Thepaut¹, Aline Leroy¹, Christine Ebel¹, Bernd Lepenies², Niels Reichardt³, Franck Fieschi¹, Corinne Vivès¹

¹Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

²University of Veterinary Medicine Hannover, Immunology Unit & Research Center for Emerging Infections and Zoonoses (RIZ),

³Glycotechnology laboratory, CIC biomaGUNE, Paseo Miramón 182, 20014, Donostia/San Sebastián, Spain

Target Journal: Glycobiology

Introduction

Glycans are essential for many biological processes ranging from critical roles in the maintenance of cell or tissue structure, molecular signal transduction, and cell recognition. The mechanisms by which they perform these diverse functions involve the interaction of the glycan with another endogenous or exogenous molecule. For instance, many cell-cell-interactions are carbohydrate driven¹. Detection of pathogens such as viruses, fungi and bacteria is mediated by recognition of glycans express on the microorganism surface. *Candida albicans*, for instance, exposes hypermannosylated N glycans its the surface². Indeed, the human immune system possesses Pattern Recognition Receptors (PRRs), expressed on dendritic cells, which are able to recognise pathogenic molecular motifs and activate the immunity³. Amongst those receptors, C-type lectin receptors (CLRs) are carbohydrate-binding proteins that are specifically involved in the recognition and the uptake of alterned-self and non-self glycans through their Ca²⁺dependent carbohydrate recognition domain (CRD). The crucial role played by CLRs in the balance of immune responses offers to CLR-glycan interaction pharmaceutical applications.⁴

Human CLRs are generally characterized by low affinity for their glycan partners. Single interaction between the protein and the isolated monosaccharide or small oligosaccharide usually involves high degree of solvent exposure, thus leading to a poor affinity, in the millimolar range ⁵. Dectin-2 for example, a CLRs involved in the response against fungi⁶, has a dissociation constant (K_d) of 2.5 mM when bound to its natural disaccharide ligand Man α 1-2Man⁷. This apparent drawback, which is a typical feature of myeloid lectins, results from the necessity to recognise a set of different ligands. The globular structure of the CRD, in fact, does not contain any cavity, therefore the recognition of carbohydrates occurs through a largely open binding site, centred on the Ca²⁺, limiting in many cases the level of selectivity achieved.

Nature counters with the « low affinity » issue by exploiting multiple binding events. The accumulation of weak affinity bindings leads to an apparent strong interaction, an effect called avidity. While affinity refers to the direct interaction of a single CRD with a monovalent ligand, avidity reflects the overall strength of multivalent interactions⁸. Multivalent binding plays a crucial role the cell-surface recognition. It could be reached by different means. On the one hand, multimerisation can be achieved on the CLR side.⁹ The result is either a clustering of single CLR in micro domains, multiple CRDs along a single polypeptide chain¹⁰ or through oligomerisation of CRD-containing receptors¹¹ (Fig.1) .In some cases an oligomerization domain, termed neck, serves as stalk to project CRDs from the cell membrane⁹.

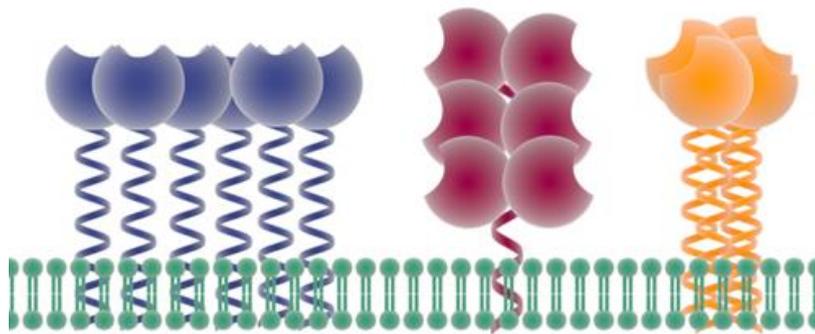


Fig. 1: Strategy allowing avidity at the cell surface. Single CLR in micro domains (blues), or multiple CRDs in a single polypeptide chain (red) or polypeptide oligomers each containing a single CRD (orange).

On the other hand, multivalent ligands also participate to high-avidity binding and contribute to the « glycan cluster effect » or « the velcro effect » that occurs when the multivalency is reached on both protein and sugar/glycan side. The interactions of DC-SIGN, a CLR expressed on DCs , implicated in viral and bacterial infections¹² with its ligands remarkably illustrates this multivalency enhancement. Firstly, DC-SIGN is expressed on cell surfaces in a tetrameric form and these tetramers are further clustered into microdomains in lipid rafts, leading to a concentrated presentation of the binding sites at the cell surface. Secondly, surface glycans of the pathogens recognised by DC-SIGN, as for instance the high mannose type glycans of HIV envelope glycoprotein gp120, are also presented as clusters with an unusual density¹³.

In vitro assays performed to detect and characterize CLR/glycan interactions have to be technically adapted to mimic this multivalency, otherwise they would fail to reveal low binding strength events. At the ligand level, one way to increase valency is to present artificial multivalent ligands. Those would include both low valency compounds, such as short polymers, glycoclusters or peptide conjugates and high valency compounds, such as dendrimers, liposomes or nanoparticles. The glycoclusters, for instance could be optimised to expose a controlled and optimal number of sugars/derivatives with an optimised presentation mode¹⁴. Glycan arrays, which consist of a series of glycan molecules attached onto a supporting material and used as a platform for biological sample screening are also a mean to present a dense surface of glycans¹⁵. On the protein side, lectin arrays¹⁵ , the counterpart of glycan arrays, with lectin coated surfaces,

and Surface Plasmon Resonance ¹⁶, with lectin functionalised sensorchips, provide solid supports to study the interaction between the ligand and a dense protein surface. However, in solution study of protein/glycan in soluble phase can provide supplementary and complementary information to the one obtained by solid phase investigation but the artifices mentioned above developed to build up multivalency could not be employed in solution. One strategy that has been considered to force CLR multimerisation resides in the development of Fc-constructs. This approach based on the fusion of a CRD domain with an immunoglobulin Fc domain enables dimerization ¹⁷. Nevertheless, this method provides limited oligomerization enhancement and higher level of multivalency may be required in some case. One way to create artificial protein tetramer is to proceed to a random labelling of the protein with a biotin tag and couple it to a tetrameric molecule of Streptavidine. This strategy presents two major drawbacks: the degree of labelling remains random and any of the accessible lysines can be targeted which may affect the lectin/sugar interactions. An alternative consists in appending a biotinylation sequence at the C-terminus end of the protein and co-expressing it with a biotin ligase. The protein, biotinylated on a single and specific lysine, can then be complexed to Streptavidin to provide tetrameric complexes ¹⁸.

Here, we propose a strategy that enables site specific biotin labelling of the N-terminus of lectin CRD by exploiting the bacterial enzyme Sortase A (SrtA) based method ¹⁹. In the Gram-positive *Staphylococcus aureus*, SrtA catalyses the anchorage of target proteins, including virulence factors, to the cell wall. The enzyme cleaves the LPXTG motif present in the proteins and links them to the amino terminal group of five glycines of the peptidoglycan ²⁰. Here, SrtA is used to enzymatically couple the N-Terminus of the protein CRD to a biotinylated peptide. The resulting biotinylated CRD is then complexed with NeutrAvidin® to obtain a final molecule exposing four glycan binding sides named hereafter TETRALEC.

In order to set up the TETRALEC strategy, a first construct was prepared with the CRD of **DC-SIGNR** ²¹. Lymph node-specific intercellular adhesion molecule-3-grabbing integrin or DC-SIGN related (L-SIGN or DC-SIGNR) natural extracellular domain (ECD) is organized in a tetrameric conformation exposing four carbohydrate domains²². Therefore, a direct comparison in ligand interaction experiments with the artificial tetrameric TETRALEC can be performed. Two other DC-SIGNR constructs will be analysed in parallel: the monomeric DC-SIGNR CRD and an artificial dimer formed by an Fc fusion with DC-SIGNR CRD (Fig.2). We have characterised the oligomeric status by SEC-MALS of all the multimeric constructs. All the constructs were then used to screen in a comparative way a library of N-glycan microarray in order to evaluate the impact of the number of CRD presented and their geometry on the avidity-based recognition process. This library comprises high mannose N glycans such as the hypermannosylated N glycans exposed by pathogenic microorganism, e.g. *Candida albicans*. *C. albicans* is a benign colonizer of skin and mucosal surfaces in oral, gastrointestinal and genital tracts²³. As an opportunistic fungal pathogen, an impaired immune system can result in infection dissemination in forms of candidiasis and systemic fungal infection (candidaemia). The highly glycosylated cell wall of *C. albicans* is constituted by α -(1,6)-linked mannosyl repeat units with side

chains of α - or β -(1,2)-linked mannosyl units connected to the backbone by a α -(1,2)-linkage as well as phosphate di-mannosyl esters and α -(1,3)-linked mannosyl units². This glycosylation pattern contributes for its immunological signature and O-linked and N-linked mannans are the main pathogen associated molecular patterns (PAMPs) that engage pattern recognition receptors (PRRs), like Toll-like receptors (TLRs) and C-type lectins (CLRs)^{24,25}. DC-SIGN interaction with N-mannosylated glycans of *C. albicans* is crucial for detection, binding and phagocytosis by human dendritic cells (DCs)². DC-SIGNR shares 77% of amino acid homology to DC-SIGN and is a calcium-dependent mannose-specific CLR. DC-SIGNR is expressed in endothelia cells of lymph nodes, liver and placenta but not on DCs and is implicated in recognition of a broad range of pathogens, like HIV, SARS-CoV, *M. tuberculosis* and *L. infantum*²². In addition, DC-SIGNR is also able to bind and internalize von Willebrand factor (VWF), a plasma glycoprotein responsible for platelet adhesion and aggregation²⁶. Despite an overlap in mannosylated glycans recognition of DC-SIGNR with DC-SIGN²⁷, no binding of DC-SIGNR and fungi has yet been described to the best of our knowledge. Therefore, we used the different constructs to investigate, by a cytofluorimetric approach, interactions between DC-SIGNR and *C. albicans*.

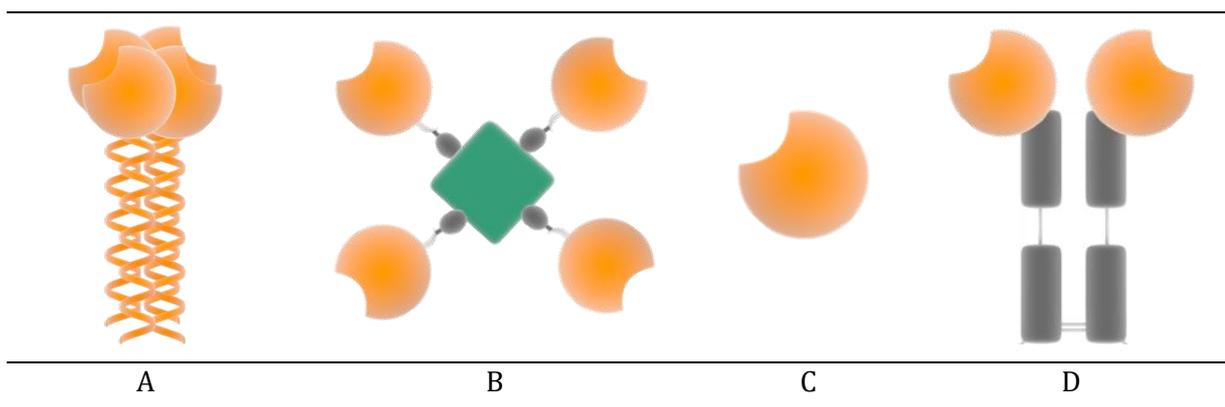


Fig.2: Four investigated constructs. A) ECD, b) TETRALEC, C) CRD and D) Fc-CRD.

Results

Design and synthesis of the TETRALEC

The N-terminus of the protein CRD was extended with a poly-Gly chain (GGG) in order to be specifically functionalized by a biotinylated peptide during the reaction catalysed by the recombinant SrtA (Figure 4a).

The number of required glycines depends on the N-terminus accessibility but Theile et al ²⁸ claim that one residue should be sufficient. The peptidic motif LPRT-OMe recognized by the Sortase A was appended to the biotin. The methylated extremity is an expedient that drastically reduces the reversibility of the reaction ¹⁹.

The kinetic of the biotinylation reaction was followed by ESI mass spectrometry (ESI-MS). An intermediate thio-ester sortase-biotin first occurs, followed by the biotinylated peptide association to DC-SIGNR GGG-CRD. As showed in figure 4b, the reaction could be considered completed after 4 to 6h of reaction.

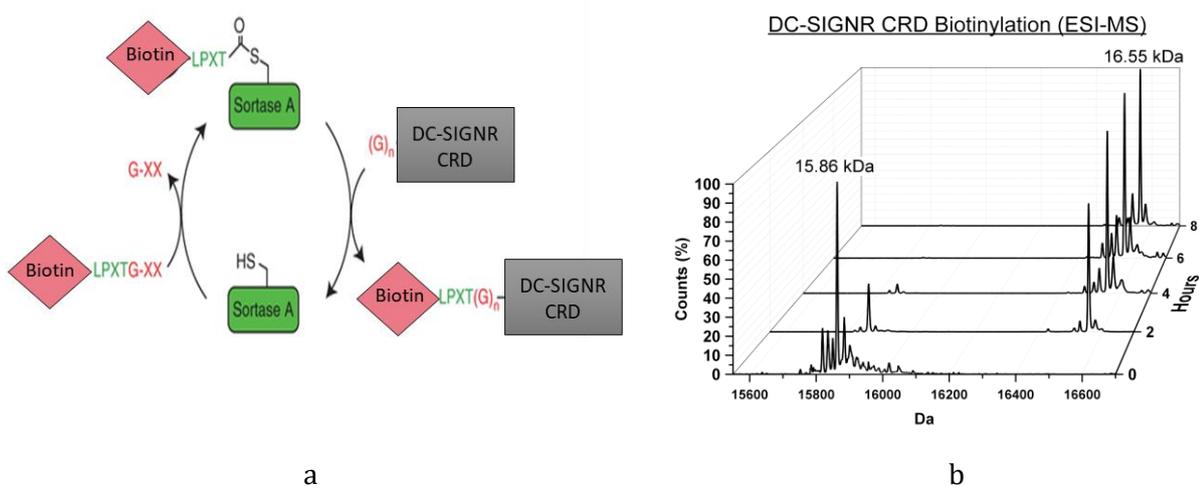


Figure 4 – **The kinetic of DC-SIGNR biotinylation followed by ESI-MS.** The peak at 15.862 kDa corresponds to the untagged CRD and the peak at 16.577 kDa corresponds to the biotinylated CRD (15.862 + 0.694 kDa). After 6h of reaction, the biotilylation is completed.

The biotinylated DC-SIGNR-CRD was then purified and complexed via its biotin moiety to a tetramer of NeutrAvidin© previously labelled with Cy3 fluorophore.

TETRALEC Structural characterization

SEC-MALS analysis was performed on DC-SIGNR-TETRALEC, to confirm the expected tetrameric presentation of DC-SIGNR-CRD by one molecule of NeutrAvidin©, and on DC-SIGNR-ECD and DC-SIGNR-Fc. For the **TETRALEC** analysis, in addition to the refractive index detection, acquisition at two wavelengths were considered: 280 nm, with the contribution of both Cy3-NeutrAvidin and DC-SIGNR CRD, and 555 nm, with the solo contribution of Cy3-NeutrAvidin. The Cy3-NeutrAvidin alone eluted as main peak at 9.2 mL with a small shoulder at 8.2 mL of larger species. The extracted molar mass along the peak decreased slightly with the elution volume suggesting some heterogeneous aggregated state. We observed a plateau after 9.4 mL. The extracted molar mass was about 120 kDa corresponding to a dimer of the natural tetramer of NeutrAvidin (MW_{theo} = 2x58kDa).

We combined the refractive index detection and the 550 nm detection to determine an experimental extinction coefficient for Cy3-NeutrAvidin: $2890 \text{ ml} \cdot (\text{g} \cdot \text{cm})^{-1}$ close to the theoretical value ($\text{exptheo} = 3100 \text{ ml} \cdot (\text{g} \cdot \text{cm})^{-1}$).

The complex eluted quite similarly to the NeutrAvidin alone with the main peak eluting at 9.1 mL. We observed also an additional shoulder at 8 mL (3%). Two additional contributions were detected at 10.5 mL (5%) and 11.5 mL (28%) corresponding to free DC-SIGNR CRD with no 550 nm absorption (Figure 5). The extracted molar masses of 38 and 17 kDa indicate a dimer and monomer of DC-SIGN CRD, respectively. Finally, the main peak at 9.1 mL shows a slightly decreasing molar mass with the elution volume, which could be due to a contamination by the shoulder. The mean molar mass is $144 \pm 1 \text{ kDa}$. Taking into consideration the theoretical extinction coefficient at 280 nm for NeutrAvidin-Cy3, the analysis gives molar masses for Cy3-NeutrAvidin of 69 kDa and for DC-SIGNR CRD of 74 kDa, thus a 4:4 complex ($M_w \text{ theo } 58 - 66 \text{ kDa}$). Unfortunately, the analysis at 550 nm gives molar masses for Cy3-NeutrAvidin of 110 to 75 kDa and, and for DC-SIGNR CRD of 40 to 50 kDa, compatible with improbable stoichiometries going from 7:2 to 5:3 of Cy3-NeutrAvidin:DC-SIGNR CRD. The discrepancy between the analysis considering 280 and 550 nm acquisition are related to bias in the determination of one or the various extinction coefficients. We consider that the presence of free Cy3-NeutrAvidin is unlikely because of the evidence of a peak of free DC-SIGNR CRD in excess. The SEC-MALS experiment presented here would not give by itself a definitive conclusion but however support the formation of the 4:4 TETRALEC stoichiometry.

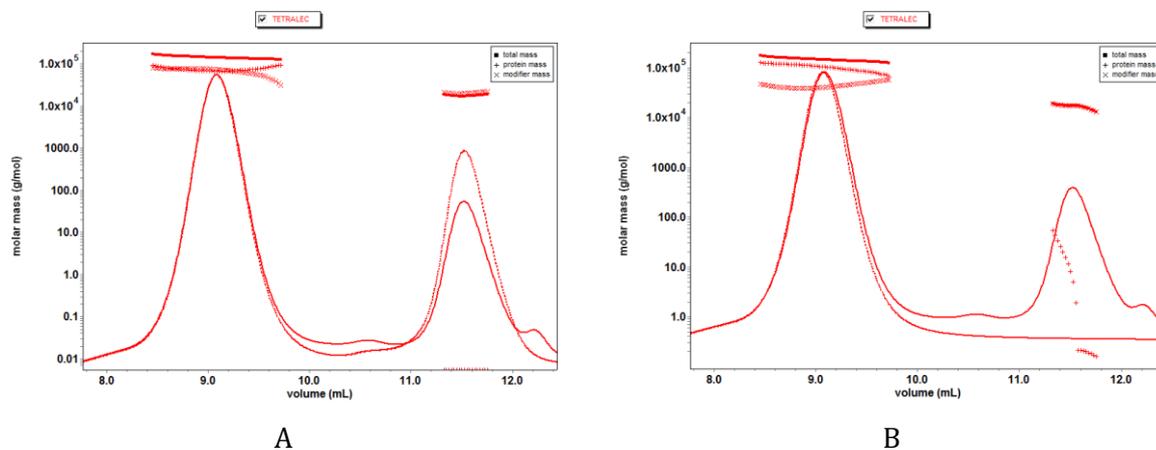


Figure 5 – PAOL analysis of the TETRALEC. a) Measurement of the elution peak at 280 nm b) Measurement of the elution peak at 555 nm. Refractive Index is represented as line, Absorbance as dot. + = NeutrAvidin and x = DC-SIGNR CRD

The artificial construct **DC-SIGNR Fc-CRD** gave a main contribution at 8.9 mL, with a M_w = of 86.5 kDa and $R_H = 9 \text{ nm}$. This contribution corresponds to a dimer of DC-SIGNR-Fc ($M_w \text{ theo } = 79.4 \text{ kDa}$) (Figure 6). Finally, the analysis on the natural tetrameric **DC-SIGNR-ECD** gave a main contribution at 7.4 mL, with a M_w = of 155.6 kDa and $R_H = 7.7 \text{ nm}$. This contribution corresponds to a tetramer of DC-SIGNR-ECD ($M_w \text{ theo } = 148.8 \text{ kDa}$) (Figure 6).

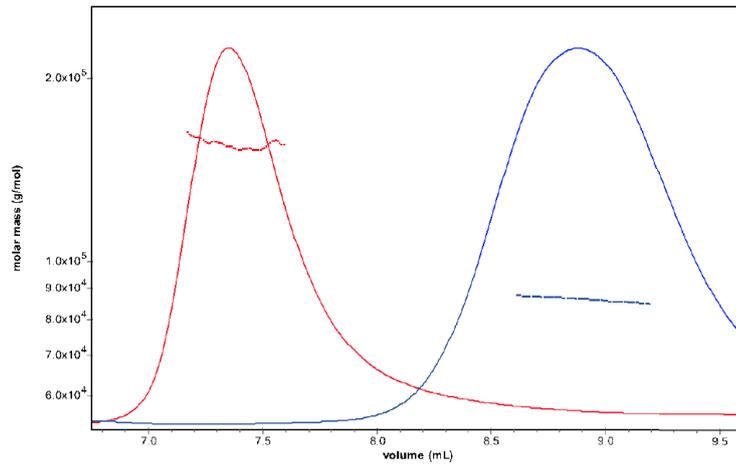


Figure 6 – SEC-MALS analysis of ECD construct (red) and Fc-CRD construct (in blue). The considered wavelength is 280 nm.

The results described above confirmed the tetrameric nature of DC-SIGNR-ECD and the dimeric nature of Fc-CRD.

Validation on glycan array

Once the oligomeric status was assessed, N-glycan microarrays were exploited to evaluate the impact of the protein multivalency on ligand binding. DC-SIGNR natural ligands are N-linked high-mannose oligosaccharides, presented on several pathogens, while blood group antigens are not recognized. Thus, a 135 N-glycan microarray (SI) was exploited to compare the three Cy3 labelled constructs DC-SIGN CRD, ECD and TETRALEC. Since DC-SIGN Fc-CRD binding was revealed by a different system using labelled anti-Fc antibodies, a direct comparison of the Fc-CDR data with the other constructs was not possible. To accurately appreciate the effect of multivalency, the concentration of active site had to be identical in each experiment. For this reason, while a concentration of 1 μ M of the tetrameric DC-SIGNR-ECD and TETRALEC was fixed for the incubation, 4 μ M and 2 μ M were used for the monomeric DC-SIGNR-CRD and the dimeric DC-SIGNR Fc-CRD, respectively.

These three constructs recognised the same ligands (Figure 7) but with different binding intensity. Monomeric CRD (0.2 DOL) was characterised by a binding weaker than the natural tetrameric ECD, and close to the background, but exhibited the same tendency of glycan recognition (Figure 7A). **GL30-42-45** and **GL65** were indeed the only binders emerging from the background for the CRD and showed to be intensively recognised also by the ECD (0.4 DOL) and TETRALEC (0.5 DOL). Incubation with DC-SIGNR ECD with a DOL 0.95 was also performed a comparable binding intensity to the 0.4 one. The CRD artificial tetrametisation in the TETRALEC led to an increase of the binding, with fluorescent intensity three-fold higher than the CRD, comparable to the ECD data (Figure 7B). Another noticeable feature of DC-SIGN TETRALEC binding profile is an increased specificity in ligand recognition compared to the two other

constructs. This presentation mode of CRD compared to the ECD seems to exacerbate differences between ligands with lower signal for the weaker ones. While, CRD and ECD constructs bind to a broad range of ligands, solely 10 glycans were strongly recognised by DC-SIGNR TETRALEC (Figure 8B). **GL45** was shown to be among the highest binders and it is an atypical non-natural branching pattern presenting the Man α 1-3(Man α 1-6Man) trisaccharide. The other glycans that were poorly recognised by DC-SIGN CRD and ECD constructs showed no binding to DC-SIGNR TETRALEC. As additional surprisingly remark, **GL43**, **GL125** and **GL134**, which showed good interaction with the ECD construct, did not show any binding towards DC-SIGNR TETRALEC. **GL43** and **GL134** possess the same branching on the 6-arm (Figure 8A). This behaviour could be explained by a more constrained spacing between the CRDs in the TETRALEC construct that could exclude some avidity binding effect with very weak ligands. Further analysis and experiments have to be performed to better understand and define this hypothesis. The Fc-CRD construct (Figure 8B) followed the same tendency of recognition observed for the other three constructs and share with the TETRALEC the increased specific recognition for some ligands. The full glycan array is included in [SI](#).

B)

A)

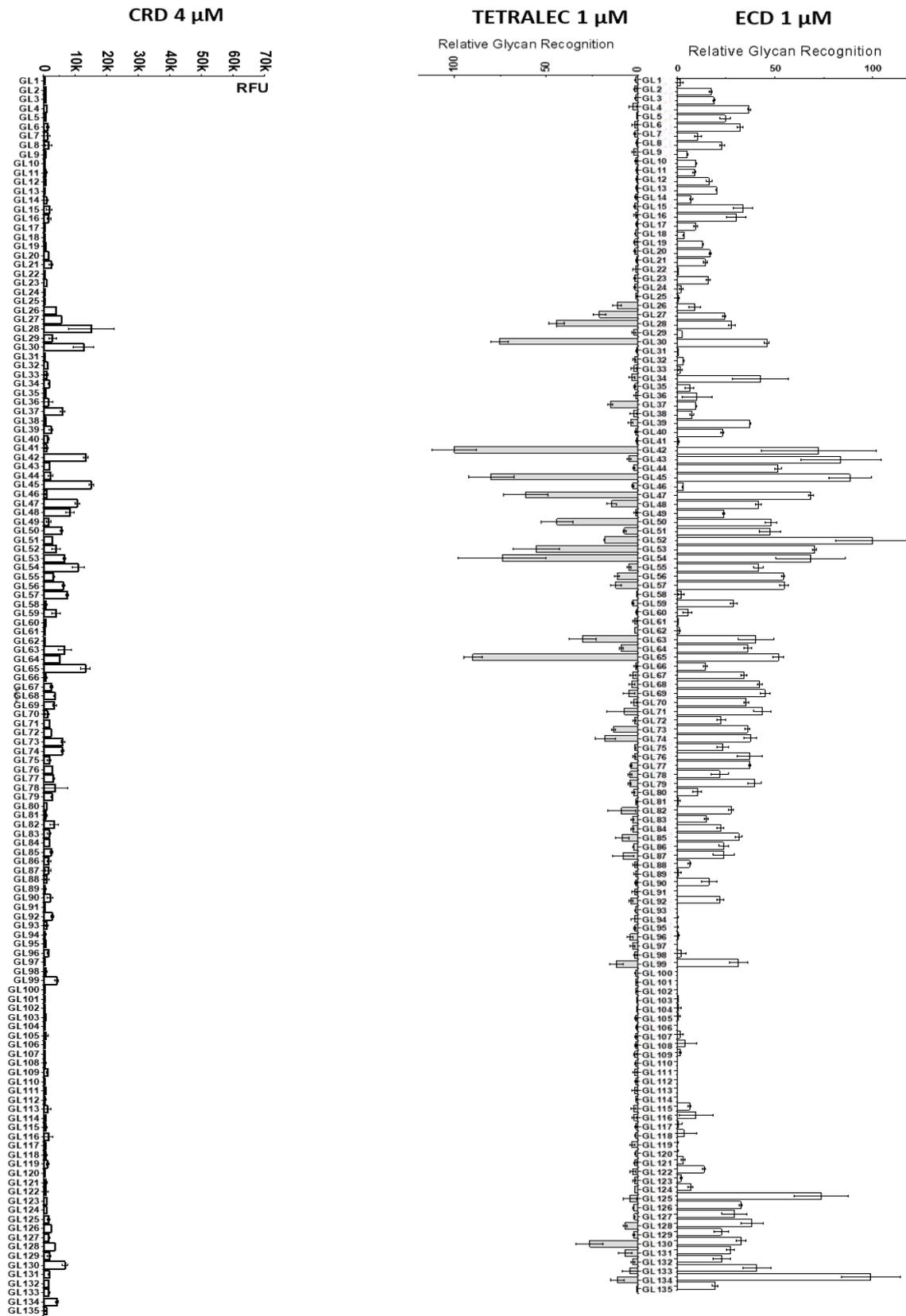
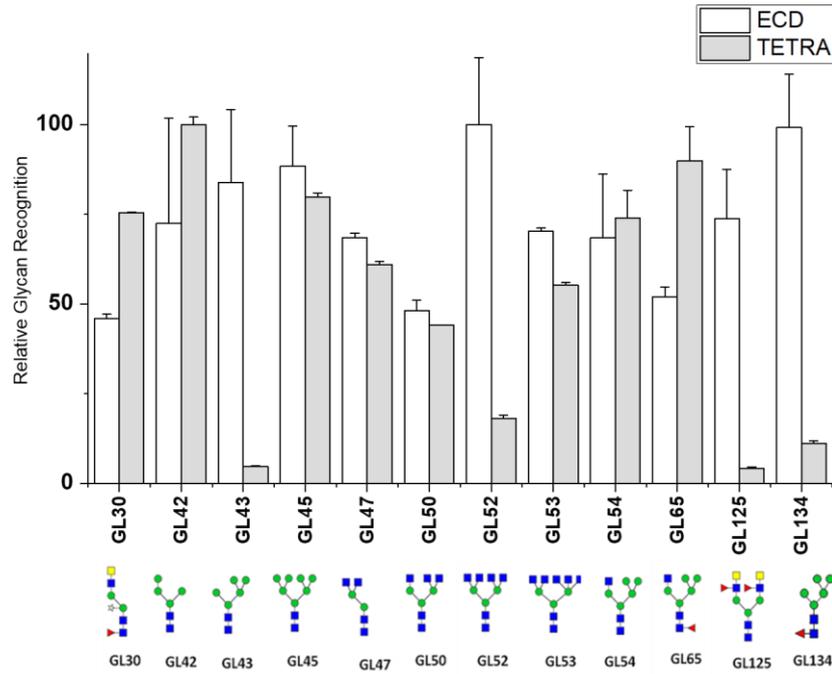
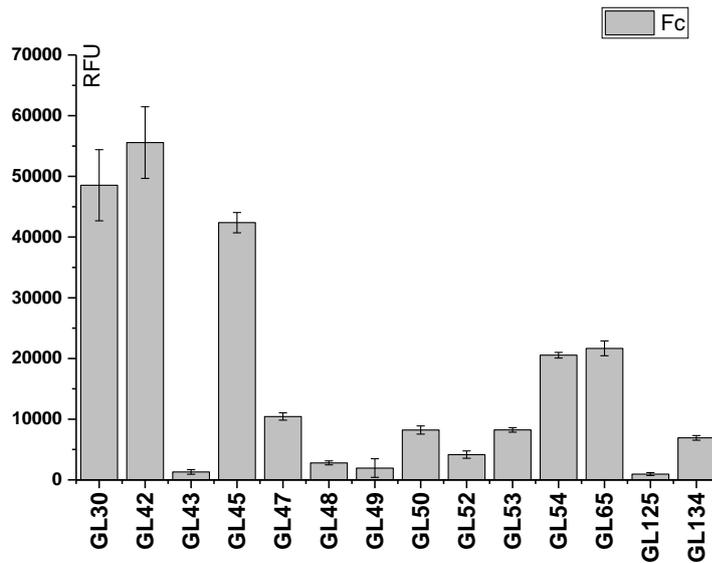


Figure 7 – A) Glycan array incubation with DC-SIGNR CRD at a concentration of 4 μM. B) Relative glycan recognition of DC-SIGNR ECD and TETRALEC incubated a concentration of 1 μM. The normalisation was performed on the glycan with the highest fluorescence.



A



B

Figure 8 – A) Zoom on the signals of the twelve highest binders for DC-SIGNR ECD and TETRALEC. B) Zoom on the signals of the same binder for DC-SIGNR CRD-Fc.

Binding to *Candida albicans*: cytofluorimetry assay

Interaction of DC-SIGNR with the pathogenic fungus *C.albicans* was investigated by cytofluorimetry. DC-SIGN was used as a positive control, taking into account the well-described role of this CLR in the recognition of *C. albicans* and DC-SIGN-mediated anti-fungal immune responses^{2,29,30}. Binding to heat-killed

C. albicans (HKCA) was observed for all DC-SIGNR-Cy3 constructs (Figure 9A), albeit to a minor extent when compared to the positive control, DC-SIGN ECD.

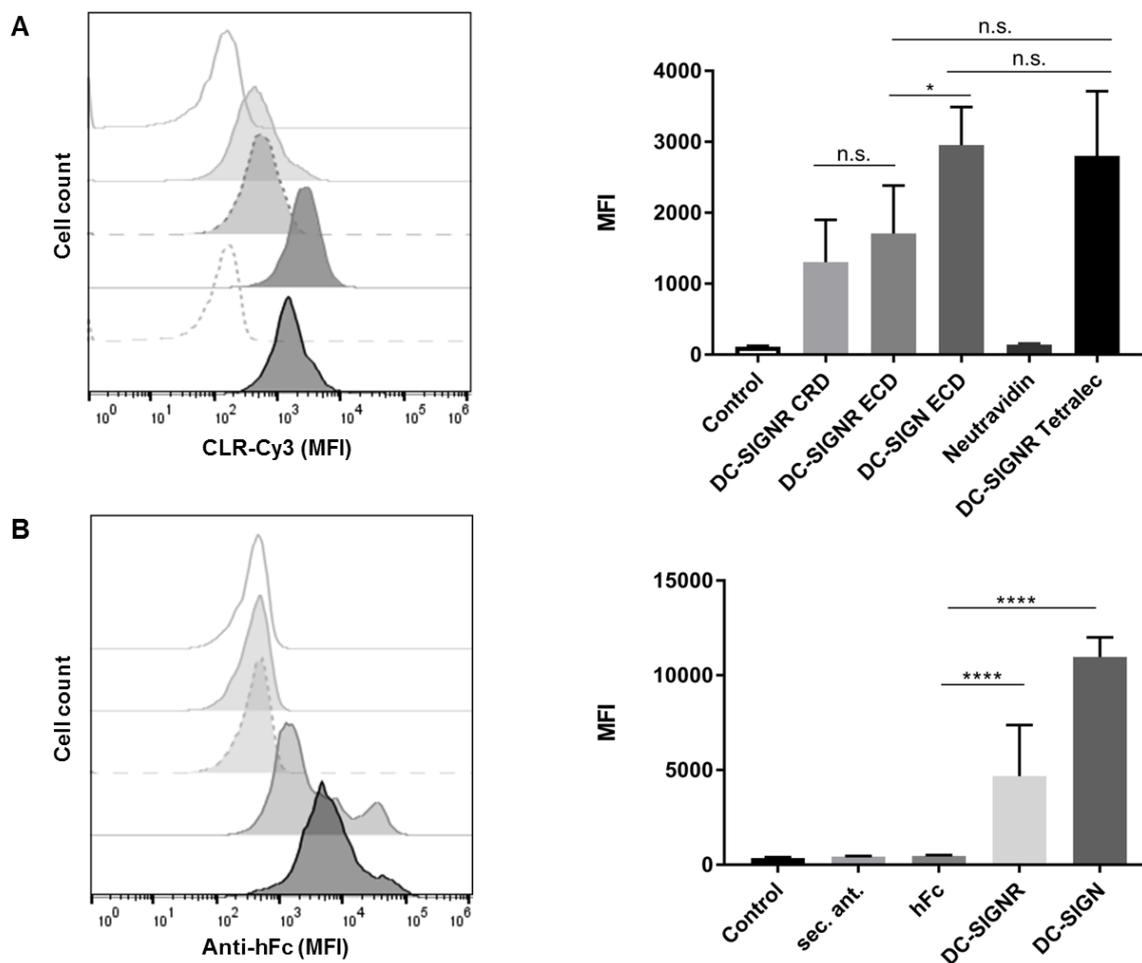


Figure 9 – Heat-killed *Candida albicans* recognition by DC-SIGNR. (A, left) Representative experiment of the histograms obtained for HKCA binding by the different DC-SIGNR-Cy3 constructs. DC-SIGN ECD was used as a positive control. (...) (A, right) Mean fluorescent intensity (MFI) of DC-SIGNR-Cy3 constructs binding to HKCA. Data depicted are the average of at least three independent experiments. (B, left) Representative experiment of CLR-Fc fusion proteins recognition of HKCA. Fc and the secondary antibody were used as negative controls, while DC-SIGN-Fc is the positive control. (B, right) Average of the MFI values obtained. Data showed are the average of at least five independent experiments. Statistical analysis of the MFI results was performed using the unpaired Student's t test, where p-values of <0.05 were considered to be significant (ns = not significant, * p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001, **** p ≤ 0.0001).

In addition, it is noticeable that DC-SIGNR TETRALEC presented the same binding profile than DC-SIGN, indicating that this construct possibly presents a higher avidity towards the mannosylated structures of *C. albicans*, which goes in accordance with the binding data obtained in the glycan array. Neutravidin did not impact binding to *C. albicans*, since no binding was observed when using the negative control neutravidin-Cy3. Recognition of HKCA occurred whatever DCSIGNR oligomeric status, although a tendency of reduced binding is observed for DC-SIGNR CRD, which did not reach statistical significance. The binding profile to HKCA was also evaluated using CLR-Fc fusion proteins (Figure 9B). The Fc fragment works as a primary

antibody, hence enabling detection of binding events with an anti-Fc antibody. Oppositely to the DC-SIGNR-Cy3 constructs that present direct labeling of the CLRs, the CLR-Fc fusion proteins employ an indirect labeling strategy. Thus, the labelling intensity cannot be directly compared to evaluate the impact of the dimeric presentation against the monomeric or tetrameric presentations for HKCA binding. DC-SIGNR-Fc also showed binding to HKCA, although to a lower degree than DC-SIGN-Fc (Figure 9B). The negatives controls (Fc and secondary antibody alone) did not evidenced binding to HKCA, when compared with the CLR-Fc fusion proteins.

Discussion

The crucial roles played by CLRs in many biological processes including pathogen recognition mechanism and/or modulation of immune response place them as strategic targets for pharmacological approaches. Besides several CLRs remain “orphan” with no ligand identified. Therefore, the screening and optimisation of CLR ligands has interested many research groups. However, while multimerisation of either the glycan or the CLR can be achieved onto a solid surface to embark upon such *in vitro* investigations, increasing the valency of the protein to perform in solution studies is less obvious. The most commonly chosen option is the creation of Fc fusion protein and it has been successfully used to screen for new CLR pathogen ligands. Such constructs allowed the identification of various interacting pairs: Mincle /*P. carinii*³¹, Mincle/*Streptococcus pneumoniae*³², SIGNR3/*Lactobacillus acidophilus*³³. Or MDL-1 (CLEC5A)/*Listeria monocytogenes*³⁴. An alternative strategy has been developed by Drickamer group. They formed a complex *in vitro* between biotinylated lectin CRDs produced in *E. coli* and StreptAvidin. The biotinylation was enabled by the addition of a 15 AA sequence at the C-terminal end of the CRD³⁵. This sequence contains a single Lysine in an appropriate context, which will be specifically biotinylated thanks to the co-expression of the bacterial biotin ligase birA. Complexation with streptavidin enabled the creation of a molecule presenting 4 CRDs. This strategy was successfully utilized to facilitate purification of human DC-SIGN homologues on mannose-Sepharose and perform glycan array screening³⁶. It was also used to study the interaction between artificial peptide-MHC oligomers with cell surface TCRs³⁷.

In this work, we present an alternative way to enable biotin-mediated tetramerisation of a DC-SIGNR CRD using peptide ligation with SrtA. This strategy only requires the introduction of a few (in theory even a single) glycine residues at the N-terminus of the protein of interest and could transfer under mild conditions any substrate carrying a LPXTG motif²⁸. In our case a biotin tag was appended and the reaction was made irreversible using a methyl ester peptide, but such an approach is potentially broadly applicable to incorporate various functions to the N-terminus of protein and not only a biotin. The enzymatic activity of SrtA has been widely used to link fluorescent tags³⁸, glycosylphosphatidylinositol (GPI) mimics³⁹ or even

PEG chains⁴⁰. The functionalization being post purification, and not co-translationally like the BirA approach, it is not limited to protein recombinantly produced in *E.coli*.

We also improve the quality of the final TETRALEC complex by changing the biotin partner for NeutrAvidin. NeutrAvidin is a deglycosylated tetrameric protein derived from avidin, its biochemical characteristics reduce non-specific binding compared to StreptAvidin⁴¹. The biotin and avidin derivatives interaction is one of the strongest non-covalent interactions ($K_D=10^{-15}M$)⁴².

Although similar artificial oligomeric constructs have been used to study Glycan/CLR interactions, to our knowledge they have never been through a complete characterisation prior starting binding studies. The objective of our approach was, besides a novel conception of artificial tetrameric lectins, to ensure that these synthetic constructs retain unaltered characteristics. For that purpose, we compared the TETRALEC behaviour in term of glycan recognition via glycan array with DC-SIGNR natural tetramer and evaluated the gain in term of binding triggered by the increase of valency by a comparison with dimeric and monomeric DC-SIGN CRD. A SEC-MALS analysis of all oligomeric lectins was performed, it awaited oligomeric status of the natural DC-SIGNR ECD and the DC-SIGN Fc-CRD construct and allowed us to assume that in the TETRALEC construct four biotinylated DC-SIGNR-CRD were successfully complexed on a molecule of NeutrAvidin®.

Very interestingly, the overall recognition glycan pattern was comparable for all construct with a clear binding enhancement triggered by the increase of valency. Besides for some reasons still to elucidate, it seems that the TETRALEC construct had an increased ligand selectivity. Those data clearly legitimate the use of our artificial lectin to perform ligand screening studies.

Like many pathogens, *C. albicans* first encounter with host defense involves its detection and clearance by the innate immune system, where CLRs expressed in the surface of epithelia, endothelial and myeloid cells play a pivotal role. Since *C. albicans* has a unique and highly mannosylated cell wall, where N-glycans account for more than 90% of the glycans present at the surface, we investigated the possibility of DC-SIGNR to interact with this fungal pathogen. Using the three different Cy3 labeled constructs of DC-SIGNR (DC-SIGNR CRD, DC-SIGNR ECD, DC-SIGNR TETRALEC) we evaluated two main points: first, if DC-SIGNR is able to bind to HKCA and, in second place, how the oligomerization state of the impact pathogen recognition. All three constructs recognised HKCA with an intensity for the TETRALEC construct almost similar to that observed for DC-SIGN. On the contrary, NeutrAvidin-Cy3, the negative control, showed no interaction, indicating that this conjugation strategy represents a useful tool to identify novel CLR-pathogen interactions, with minimal unspecific binding. The different oligomerizations states of DC-SIGNR did not

significantly impact recognition of HKCA. CLR-Fc fusion have been used extensively to identify novel CLR-pathogen interactions, namely fungi⁴³. In our case, the dimeric presentation of the Fc fusions denotes a markedly strong interaction between HKCA and DC-SIGN and the identification of a novel interaction to HKCA by DC-SIGNR. The preferential binding differences between DC-SIGN and DC-SIGNR to pathogens that possess high mannose oligosaccharides may be associated with the different capacities of the CLRs to accommodate the differential spatial conformations of sugar epitopes in their binding pockets due to their distinct properties of the neck domains⁴⁴.

The identification of an interaction between *C. albicans* and DC-SIGNR may open new fields of investigations and could be considered in pathologies due to *C. albicans* infections involving cells expressing DCSIGNR. In immunocompromised mice, gastrointestinal (GI) candidiasis can occur and destroy the gut mucosa⁴⁵ where DCSIGNR is expressed in the sinusoidal endothelial cells. In addition, DC-SIGNR is reported to be expressed in the placental villi and *C. albicans* is shown to cause chorioamnionitis, an inflammation of the fetal membranes due to a bacterial infection, after placenta invasion^{46,47}.

Materials and Methods

Cloning

Standard pUC57 plasmids containing optimized synthetic human genes encoding human DC-SIGNR ECD (amino acids 78-399) and CRD (amino acids 264-399) designed for the efficient production in *E. coli* were manufactured by GeneCust Europe (Luxembourg). PCR amplification using suitable primers and restriction enzyme digestion were used to sub-clone into the pET30-b (Novagen) DC-SIGNR ECD between the NdeI and HindIII restriction sites and DC-SIGNR CRD between the XbaI and HindIII sites. The sequencing of each construction was done by Genewiz (Takeley, Royaume Uni).

Protein expression and purification

DC-SIGNR ECD was expressed in *E. coli* BL21(DE3) in 1 liter of LB medium supplemented with 50 µg/ml kanamycin at 37 °C. Expression was induced by addition of 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG) when the culture had reached an A_{600 nm} of 0.8 and maintained for 3h. The protein was expressed in the bacterial cytoplasm as inclusion bodies. Cells were harvested by a 20-min centrifugation at 5000 g at 4 °C. The pellet was resuspended in 30 mL of a solution containing 150 mM NaCl, 25 mM Tris-HCl, pH 8 and one anti-protease mixture tablet (Complete EDTA free, Roche). Cells were disrupted by sonication and cell debris eliminated by centrifugation at 100,000 g for 45 min at 4 °C in a Beckman 45Ti rotor. The pellet was solubilized in 30 mL of 6 M guanidine-HCl containing 25 mM Tris-HCl pH 8, 150 mM NaCl and 0,01% (v/v) beta-mercaptoethanol. The mixture was centrifuged at 100,000g for 45 min at 4°C and the supernatant was diluted 5-fold with 25 mM Tris-HCl pH 8, 1,25 M NaCl and 25 mM CaCl₂ by slow addition with stirring. The

diluted mixture was dialyzed against 10 volumes of 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mM CaCl₂ (buffer A) with 3 buffer changes. After dialysis, insoluble precipitate was removed by centrifugation at 100,000g for 1h at 4°C. The supernatant containing DC-SIGNR ECD was loaded on Mannan agarose column (Sigma) for purification by affinity chromatography equilibrated with buffer A. After loading, DC-SIGNR ECD was tightly bound to the column and eluted in the same buffer without CaCl₂ but supplemented with 1 mM EDTA (buffer B). This step was followed by SEC (Size Exclusion Chromatography) using a Superose 6 column (GE Healthcare) equilibrated with buffer A. Fractions were analysed by SDS-PAGE (12%) and DC-SIGNR ECD containing fractions were pooled and concentrated by ultrafiltration (YM10 membrane from Amicon).

DC-SIGNR CRD was expressed in *E. coli* BL21(DE3) in 1 liter of LB medium supplemented with 50 µg/ml kanamycin at 37 °C. Expression was induced by addition of 1 mM IPTG when the culture had reached an A_{600 nm} of 0.8 and maintained for 3h. The protein was expressed in the cytoplasm as inclusion bodies. Cells were harvested by a 20 min centrifugation at 5000 g at 4 °C. The pellet was resuspended in 30 mL of a solution containing 150 mM NaCl, 25 mM Tris-HCl, pH 8 and one anti-protease mixture tablet (Complete EDTA free, Roche). Cells were disrupted by sonication and cell debris eliminated by centrifugation at 100,000 g for 45 min at 4 °C in a Beckman 45Ti rotor. The pellet was solubilized in 30 mL of 6 M guanidine-HCl containing 25 mM Tris-HCl pH 8, 150 mM NaCl and 0,01% beta-mercaptoethanol. The mixture was centrifuged at 100,000 g for 45 min at 4°C and the supernatant was diluted 5-fold with 200 mM Tris-HCl pH 8, 1,25 M NaCl and 25 mM CaCl₂ by slow addition under stirring. The diluted mixture was dialyzed against 10 volumes of Buffer A with 3 buffer changes. After dialysis, insoluble precipitate was removed by centrifugation at 100,000 g for 1h at 4°C. The supernatant containing the His tagged DC-SIGNR CRD was loaded onto a HisTrap (GE Healthcare) at 4°C. Unbound proteins were washed away with buffer A before DC-SIGNR CRD was eluted with buffer C (150 mM NaCl, 25 mM Tris-HCl, pH 8, 4 mM CaCl₂, 0.5 M imidazole). Eluted fractions were analyzed by SDS-PAGE (15%) and the DC-SIGNR CRD containing fractions were pooled and concentrated by ultrafiltration (YM10 membrane from Amicon).

Each construct was checked by N-terminal amino acid sequencing and mass spectrometry.

Labelling

A total of 250 µL of 3.77 mg/mL solutions of DC-SIGNR ECD, 212,5 µL of 4.75 mg/mL solutions of DC-SIGNR CRD in 25 mM HEPES pH 7,25, 4 mM CaCl₂ and 100 µL of 3.47 mg/mL solutions of NeutrAvidin in PBS pH 7.4 were prepared. 1 µL, 4 µL and 2 µL, respectively of 10 mg/mL Cy3-NHS ester (Gene Copoeia) were added to the solutions and the reactions were gently shaken at RT for 2 h and then at 4°C for 4h. Excess dye was removed by two dialysis (3.5k Z-lyser from Thermo Scientific) of 3h against 25 mM Tris pH8, 150 mM NaCl, 4 mM CaCl₂. The amount of attached Cy3 was estimated spectrophotometrically based on the dye molar

extinction coefficient (ϵ 150 000 $\text{cm}^{-1}\text{M}^{-1}$) and proteins extinction coefficient ϵ . The obtained degree of labelling (DOL) was 0.4, 0.2 and 0.5.

TETRALEC formation

His tag cleavage. His-GGG-CRD was cleaved using factor Xa (**Thermo Fischer**) following the ratio recommended by the company: 1 μg of factor Xa per 50 μg of His-GGG-CRD protein at 1 mg/mL . The reaction was performed overnight at RT under agitation and then injected into Toyopearl® exclusion chromatography column previously equilibrated in 25 mM Tris pH 8, 150 mM NaCl, 4 mM CaCl_2 buffer. 1 mL/min flow rate of was maintained during the purification. Eluted fractions were pooled and concentrated up to 1 mg.

Sortase-directed biotinylation and TETRALEC formation. The protocol from¹⁹ was used for the biotinylation of GGG-CRD. The protein exposing three glycines at the N-terminus (1 equivalent) was mixed with the peptide biotin-LPRT-OMe (MW= 725.9 Da, **Covalab**) (5 eq.) and His-tag Sortase A (SrtA) (0.3 eq) from *S.aureus*, recombinantly produced in the lab, in 25 mM Tris pH 8, 150 mM NaCl, 4 mM CaCl_2 buffer. The reaction was incubated at 37°C for 6h under agitation. The kinetic of reaction was followed by ESI-MS (electrospray ionization mass spectrometry): 10 μL of reaction was analysed at 0h, 2h, 4h, 6h,8h and overnight. When the reaction was completed, the solution was loaded onto a 1 mL HisTrap column previously equilibrated in 25 mM Tris pH, 150 mM NaCl, 4 mM CaCl_2 buffer. After column washing, the elution step was performed using 25 mM Tris pH, 150 mM NaCl, 4 mM CaCl_2 , 0.5 M Imidazole buffer. 1 mL/min flow rate of buffer was maintained during the purification. The His tagged sortase was retained by the HisTrap column while the untagged biotin-CRD was eluted during the washing step and was pooled and dialysed against 25 mM Tris pH, 150 mM NaCl, 4 mM CaCl_2 to eliminate un-reacted biotin.

Finally, NeutrAvidin© (MW = 14.5 kDa, **Thermo Fisher**) sample previously labelled with Cy3-fluorophore (2.9 mg/mL , DOL=0.5) was mixed to Biotin-CRD with a molar ratio of 1:1 and the reaction was incubated overnight at 4°C under agitation. The obtained CRD-TETRALEC was frozen in liquid nitrogen for storage at -80°C.

SEC-MALS analysis

The samples were centrifuged for 15 min at 20800g just before the experiment. The elution buffer, 25mM Tris pH 7.5, 150mM NaCl, 4mM CaCl_2 , was filtered at 0.1 μm SEC-MALS experiments were conducted on a HPLC (Schimadzu, Kyoto, Japan) consisting of a degasser DGU-20AD, a LC-20AD pump, a autosampler SIL20-ACHT, a communication interface CBM-20A, a UV-Vis detector SPD-M20A and a fraction collector FRC-10A, a column oven XL-Therm (WynSep, Sainte Foy d'Aigrefeuille, France) and a static light scattering detector miniDawn Treos, a dynamic light scattering detector DynaPro NANOSTAR, a refractive index detector

Optilab rEX (Wyatt, Santa-Barbara, USA). The samples were stored at 4°C, and a volume of 50 µl was injected, on a KW 802.5 column, equilibrated at 4°C, with the elution buffer, at a flow rate of 0.5 ml/min.

Glycan array analysis

Cy3 labelled C-type lectins were diluted in incubation buffer (25 mM Tris-HCl, 150 mM NaCl, 4 mM CaCl₂, pH7.5 containing 0.5% (w/v) bovine serum albumin (BSA) and 0.005% (v/v) Tween®-20. Lectin solutions (200 µL per array) were used to incubate individual wells on a glycan array slide at 4 °C for 18 hours. Arrays were washed with incubation buffer without BSA, H₂O and dried in a slide spinner.

Fluorescence measurements were performed on a microarray scanner (Agilent G2565BA, Agilent Technologies) at 10 µm resolution. Quantification of fluorescence was performed by ProScanArray® Express software (Perkin Elmer) employing an adaptive circle quantification method from 50 µm (minimum spot diameter) to 300 µm (maximum spot diameter). Average RFU values with local background subtraction of four spots and standard deviation of the mean were reported using Microsoft Excel and GraphPad Prism® software.

Generation of human DC-SIGNR-Fc fusion protein

The production of human DC-SIGNR-Fc and DC-SIGN fusion proteins were performed as previously described ^{48, 49}. Briefly, a human cDNA library was used as template (GE Dharmacon, Lafayette, CO, USA) and specific primers to amplify the CRD of DC-SIGNR were generated (Eurofins Genomics, Ebersberg, Germany). The DC-SIGNR and DC-SIGN primers were as follows: DC-SIGNR forward, gaattcctatcaagaactgaccgattg; DC-SIGN forward, gaattcgtccaaggtccccagctccat; DC-SIGNR reverse ccatggattcgtctctgaagcaggc; and DC-SIGN reverse, ccatggacgcaggaggggggtttgggggt. PCR was used to amplify the cDNA, followed by ligation into a pFuse-hlgG1-Fc expression vector (InvivoGen, San Diego, CA, USA). The DC-SIGNR-Fc expression vector was used to transiently transfect CHO-S cells with MAX reagent (Invivogen). After 4 days of transfection, the supernatant was collected and the fusion proteins were purified with a HisTrap protein G HP column (GE Healthcare, Piscataway, NJ, USA). Protein purity was confirmed by SDS-PAGE with subsequent Coomassie staining. Western blot using an anti-human IgG-horseradish peroxidase (HRP) antibody (Dianova, Hamburg, Germany) was also performed to detect the presence of the fusion protein.

Flow cytometry-based binding to *Candida albicans*

Heat-killed *Candida albicans* (InvivoGen, San Diego, CA, USA) was stained for 15 min with 1µM of DNA-staining dye Syto61 (ThermoFisher Scientific) at 4°C. The samples were subsequently washed two times with 1x PBS. Then, samples were incubated for 1h either with 250 ng of the respective CLR-Fc fusion proteins in lectin-binding buffer (50 mM HEPES, 5 mM MgCl₂, 5 mM CaCl₂, pH 7.4) or with 1 µM of the DC-

SIGNR-HisTag and DC-SIGNR-Neutravidin constructs in its respective lectin-binding buffer (25 mM Tris, 150 mM NaCl, 4 mM CaCl₂, pH 8.0). After washing once with the lectin-binding buffer, the pellet was suspended in a 1:200 PE-conjugated goat anti-human Fc antibody (Dianova) and incubated for 20 min at 4°C, for detection of the bound CLR-Fc fusion proteins. Finally, cells were washed two times and flow-cytometric analysis was performed using an Attune NxT Flow Cytometer (ThermoFisher Scientific). The gating strategy applied was a first gate in the *Candida albicans* population, followed by a single cell population gating for doublet exclusion. In the single cell population gate, Syto61 positive cells were selected and further analyzed for CLR binding. The same gating strategy was performed for all experimental conditions within one experiment. Flow cytometry data were analyzed using the FlowJo version 10 software (FlowJo, Ashland, OR, USA).

References

1. Schnaar RL. Glycobiology simplified: diverse roles of glycan recognition in inflammation. *J Leukoc Biol.* 2016;99(6):825-838. doi:10.1189/jlb.3RI0116-021R
2. Cambi A, Netea MG, Mora-Montes HM, et al. Dendritic Cell Interaction with *Candida albicans* Critically Depends on N-Linked Mannan. *J Biol Chem.* 2008;283(29):20590-20599. doi:10.1074/jbc.M709334200
3. Levitz SM, Golenbock DT. Beyond empiricism: Informing vaccine development through innate immunity research. *Cell.* 2012;148(6):1284-1292. doi:10.1016/j.cell.2012.02.012
4. J VS, J GJ, Yvette K. Dendritic cells and C-type lectin receptors: coupling innate to adaptive immune responses. *Immunol Cell Biol.* 2008;86(7):580-587. doi:10.1038/icb.2008.55
5. Sager CP, Eriş D, Smieško M, Hevey R, Ernst B. What contributes to an effective mannose recognition domain? Hoffmann-Röder A, ed. *Beilstein J Org Chem.* 2017;13:2584-2595. doi:10.3762/bjoc.13.255
6. SX T, DL M, JP R, Blagojevic M, JR N. Epithelial discrimination of commensal and pathogenic *Candida albicans*. *Oral Dis.* 2016;22(S1):114-119. doi:10.1111/odi.12395
7. Feinberg H, Jégouzo SAF, Rex MJ, Drickamer K, Weis WI, Taylor ME. Mechanism of pathogen recognition by human dectin-2. *J Biol Chem.* 2017;292(32):13402-13414. doi:10.1074/jbc.M117.799080
8. Varki A GP. Biological Functions of Glycans. In: *Essentials of Glycobiology [Internet]. 3rd Edition.* ; 2017.
9. Taylor ME, Drickamer K SR. Discovery and Classification of Glycan-Binding Proteins. In:

Essentials of Glycobiology [Internet]. 3rd Edition. ; 2017.

10. East L, Isacke CM. The mannose receptor family. *Biochim Biophys Acta - Gen Subj.* 2002;1572(2-3):364-386. doi:10.1016/S0304-4165(02)00319-7
11. Tabarani G, Thépaut M, Stroebel D, et al. DC-SIGN Neck Domain Is a pH-sensor Controlling Oligomerization: SAXS AND HYDRODYNAMIC STUDIES OF EXTRACELLULAR DOMAIN. *J Biol Chem.* 2009;284(32):21229-21240. doi:10.1074/jbc.M109.021204
12. Sprokholt JK, Overmars RJ, Geijtenbeek TBH. DC-SIGN in Infection and Immunity BT - C-Type Lectin Receptors in Immunity. In: Yamasaki S, ed. Tokyo: Springer Japan; 2016:129-150. doi:10.1007/978-4-431-56015-9_9
13. Scanlan CN, Offer J, Zitzmann N, Dwek RA. Exploiting the defensive sugars of HIV-1 for drug and vaccine design. *Nature.* 2007;446:1038. <http://dx.doi.org/10.1038/nature05818>.
14. Berzi A, Ordanini S, Joosten B, et al. Pseudo-Mannosylated DC-SIGN Ligands as Immunomodulators. *Sci Rep.* 2016;6:35373. <http://dx.doi.org/10.1038/srep35373>.
15. Gemeiner P, Mislovičová D, Tkáč J, et al. Lectinomics: II. A highway to biomedical/clinical diagnostics. *Biotechnol Adv.* 2009;27(1):1-15. doi:10.1016/J.BIOTECHADV.2008.07.003
16. Tamburrini A, Achilli S, Vasile F, et al. Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist. *Bioorg Med Chem.* 2017;25(19):5142-5147. doi:10.1016/J.BMC.2017.03.046
17. Czajkowsky DM, Hu J, Shao Z, Pleass RJ. Fc-fusion proteins: new developments and future perspectives. *EMBO Mol Med.* 2012;4(10):1015-1028. doi:10.1002/emmm.201201379
18. Chapman-Smith A, Cronan Jr JE. The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci.* 1999;24(9):359-363. doi:10.1016/S0968-0004(99)01438-3
19. Antos JM, Chew G-L, Guimaraes CP, et al. Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity. *J Am Chem Soc.* 2009;131(31):10800-10801. doi:10.1021/ja902681k
20. J. BW, H. DA, J. CC, K. RA, C. SC, Ravi AK. Molecular features of the sortase enzyme family. *FEBS J.* 2015;282(11):2097-2114. doi:10.1111/febs.13288
21. Bashirova AA, Geijtenbeek TBH, van Duijnhoven GCF, et al. A Dendritic Cell-Specific Intercellular Adhesion Molecule 3-Grabbing Nonintegrin (Dc-Sign)-Related Protein Is Highly Expressed on Human Liver Sinusoidal Endothelial Cells and Promotes HIV-1 Infection. *J Exp Med.* 2001;193(6):671-678.

22. Khoo U-S, Chan KYK, Chan VSF, Lin CLS. DC-SIGN and L-SIGN: the SIGNs for infection. *J Mol Med*. 2008;86(8):861-874. doi:10.1007/s00109-008-0350-2
23. Kim J, Sudbery P. *Candida albicans*, a major human fungal pathogen. *J Microbiol*. 2011;49(2):171. doi:10.1007/s12275-011-1064-7
24. Barreto-Bergter E, Figueiredo RT. Fungal glycans and the innate immune recognition. *Front Cell Infect Microbiol*. 2014;4:145. doi:10.3389/fcimb.2014.00145
25. Tang J, Lin G, Langdon WY, Tao L, Zhang J. Regulation of C-Type Lectin Receptor-Mediated Antifungal Immunity. *Front Immunol* . 2018;9:123.
26. Rydz N, Swystun LL, Notley C, et al. The C-type lectin receptor CLEC4M binds, internalizes, and clears von Willebrand factor and contributes to the variation in plasma von Willebrand factor levels. *Blood*. 2013;121(26):5228-5237. doi:10.1182/blood-2012-10-457507
27. Guo Y, Feinberg H, Conroy E, et al. Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct & Mol Biol*. 2004;11:591. <http://dx.doi.org/10.1038/nsmb784>.
28. Theile CS, Witte MD, Blom AEM, Kundrat L, Ploegh HL, Guimaraes CP. Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nat Protoc*. 2013;8(9):1800-1807. doi:10.1038/nprot.2013.102
29. Takahara K, Arita T, Tokieda S, et al. Difference in Fine Specificity to Polysaccharides of *Candida albicans* Mannoprotein between Mouse SIGNR1 and Human DC-SIGN. Deepe GS, ed. *Infect Immun*. 2012;80(5):1699-1706. doi:10.1128/IAI.06308-11
30. Alessandra C, Karlijn G, M. de VIJ, et al. The C-type lectin DC-SIGN (CD209) is an antigen-uptake receptor for *Candida albicans* on dendritic cells. *Eur J Immunol*. 2003;33(2):532-538. doi:10.1002/immu.200310029
31. Hoving JC. Pneumocystis and interactions with host immune receptors. Sheppard DC, ed. *PLoS Pathog*. 2018;14(2):e1006807. doi:10.1371/journal.ppat.1006807
32. Rabes A, Zimmermann S, Reppe K, et al. The C-Type Lectin Receptor Mincle Binds to *Streptococcus pneumoniae* but Plays a Limited Role in the Anti-Pneumococcal Innate Immune Response. Miyaji EN, ed. *PLoS One*. 2015;10(2):e0117022. doi:10.1371/journal.pone.0117022
33. Lightfoot YL, Selle K, Yang T, et al. SIGNR3-dependent immune regulation by *Lactobacillus acidophilus* surface layer protein A in colitis. *EMBO J*. 2015;34(7):881-895. doi:10.15252/embj.201490296

34. Chen S-T, Li F-J, Hsu T, et al. CLEC5A is a critical receptor in innate immunity against *Listeria* infection. *Nat Commun.* 2017;8(1):299. doi:10.1038/s41467-017-00356-3
35. Schatz PJ. Use of Peptide Libraries to Map the Substrate Specificity of a Peptide-Modifying Enzyme: A 13 Residue Consensus Peptide Specifies Biotinylation in *Escherichia coli*. *Bio/Technology.* 1993;11:1138. <http://dx.doi.org/10.1038/nbt1093-1138>.
36. Powlesland AS, Ward EM, Sadhu SK, Guo Y, Taylor ME, Drickamer K. Widely divergent biochemical properties of the complete set of mouse DC-SIGN-related proteins. *J Biol Chem.* 2006;281(29):20440-20449. doi:10.1074/jbc.M601925200
37. Stone JD, Artyomov MN, Chervin AS, Chakraborty AK, Eisen HN, Kranz DM. Interaction of Streptavidin-Based Peptide-MHC Oligomers (Tetramers) with Cell-Surface T Cell Receptors(). *J Immunol.* 2011;187(12):6281-6290. doi:10.4049/jimmunol.1101734
38. Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Sortagging: a versatile method for protein labeling. *Nat Chem Biol.* 2007;3:707. <http://dx.doi.org/10.1038/nchembio.2007.31>.
39. Guo X, Wang Q, Swartz BM, Guo Z. Sortase-Catalyzed Peptide-Glycosylphosphatidylinositol Analogue Ligation. *J Am Chem Soc.* 2009;131(29):9878-9879. doi:10.1021/ja903231v
40. Parthasarathy R, Subramanian S, Boder ET. Sortase A as a Novel Molecular "Stapler" for Sequence-Specific Protein Conjugation. *Bioconjug Chem.* 2007;18(2):469-476. doi:10.1021/bc060339w
41. Nguyen TT, Sly KL, Conboy JC. Comparison of the Energetics of Avidin, Streptavidin, NeutrAvidin, and Anti-Biotin Antibody Binding to Biotinylated Lipid Bilayer Examined by Second-Harmonic Generation. *Anal Chem.* 2012;84(1):201-208. doi:10.1021/ac202375n
42. Jain A, Cheng K. The principles and applications of avidin-based nanoparticles in drug delivery and diagnosis. *J Control Release.* 2017;245:27-40. doi:10.1016/J.JCONREL.2016.11.016
43. Stappers MHT, Clark AE, Amanianda V, et al. Recognition of DHN-melanin by a C-type lectin receptor is required for immunity to *Aspergillus*. *Nature.* 2018;555:382. <http://dx.doi.org/10.1038/nature25974>.
44. dos Santos Á, Hadjivasiliou A, Ossa F, et al. Oligomerization domains in the glycan-binding receptors DC-SIGN and DC-SIGNR: Sequence variation and stability differences. *Protein Sci.* 2017;26(2):306-316. doi:10.1002/pro.3083
45. Kobayashi-Sakamoto M, Tamai R, Isogai E, Kiyoura Y. Gastrointestinal colonisation and

systemic spread of *Candida albicans* in mice treated with antibiotics and prednisolone.

Microb Pathog. 2018;117:191-199. doi:10.1016/J.MICPATH.2018.02.043

46. Fumitake I, Tomoharu O, Tadahiro Y, et al. Premature delivery due to intrauterine *Candida* infection that caused neonatal congenital cutaneous candidiasis: A case report. *J Obstet Gynaecol Res.* 2012;39(1):341-343. doi:10.1111/j.1447-0756.2012.01938.x
47. Delprado WJ, Baird PJ, Russell P. Placental candidiasis: report of three cases with a review of the literature. *Pathology.* 1982;14(2):191-195. doi:10.3109/00313028209061293
48. Maglinao M, Eriksson M, Schlegel MK, et al. A platform to screen for C-type lectin receptor-binding carbohydrates and their potential for cell-specific targeting and immune modulation. *J Control Release.* 2014;175:36-42. doi:10.1016/J.JCONREL.2013.12.011
49. Mayer S, Moeller R, Monteiro JT, Ellrott K, Josenhans C, Lepenies B. C-Type Lectin Receptor (CLR)–Fc Fusion Proteins As Tools to Screen for Novel CLR/Bacteria Interactions: An Exemplary Study on Preselected *Campylobacter jejuni* Isolates. *Front Immunol.* 2018;9:213. doi:10.3389/fimmu.2018.00213

9. Screening: identification of selective ligands towards human CLRs

The ligand microarray platform from Niels Reichardt's laboratory was exploited for high throughput screening of a panel of glycans (synthesized by the laboratory of Niels Reichardt) and glycomimetics (synthesized by the laboratory of Anna Bernardi) with the fluorescently labelled recombinant CLRs that we produced. After an overnight incubation of the Cy3-labelled CLRs with the glycans/glycomimetics spotted in quadruplicate on a slide, the fluorescence was measured using a green channel laser. The interaction was revealed using Agila Scanner and the green spot obtained (Fig.103) possessed a colour intensity proportional to the strength of the interaction. After this first visualisation, ProScanArray® Express software allowed a proper evaluation of the shape of each single spot and a relative quantification of the interaction. The final considered output was the median relative fluorescence intensity per glycan. All the histograms presented in this chapter correspond to these values.

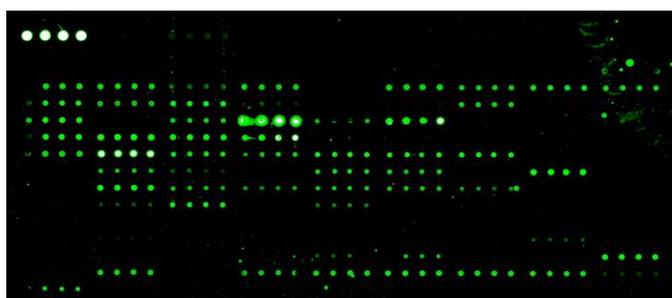


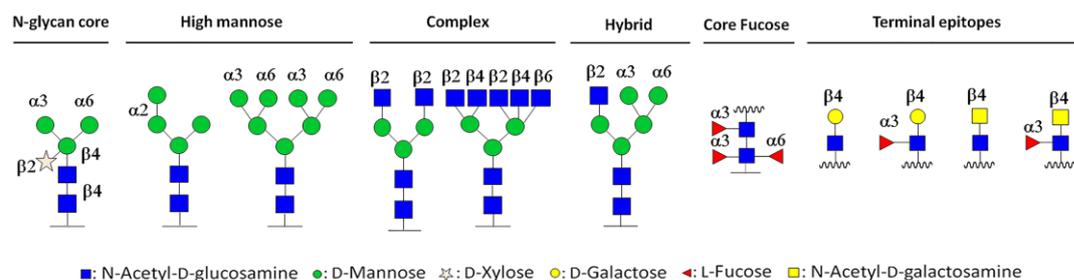
Fig.103 Example of interaction detection by Agila Scanner.

This qualitative characterization was explored for two purposes:

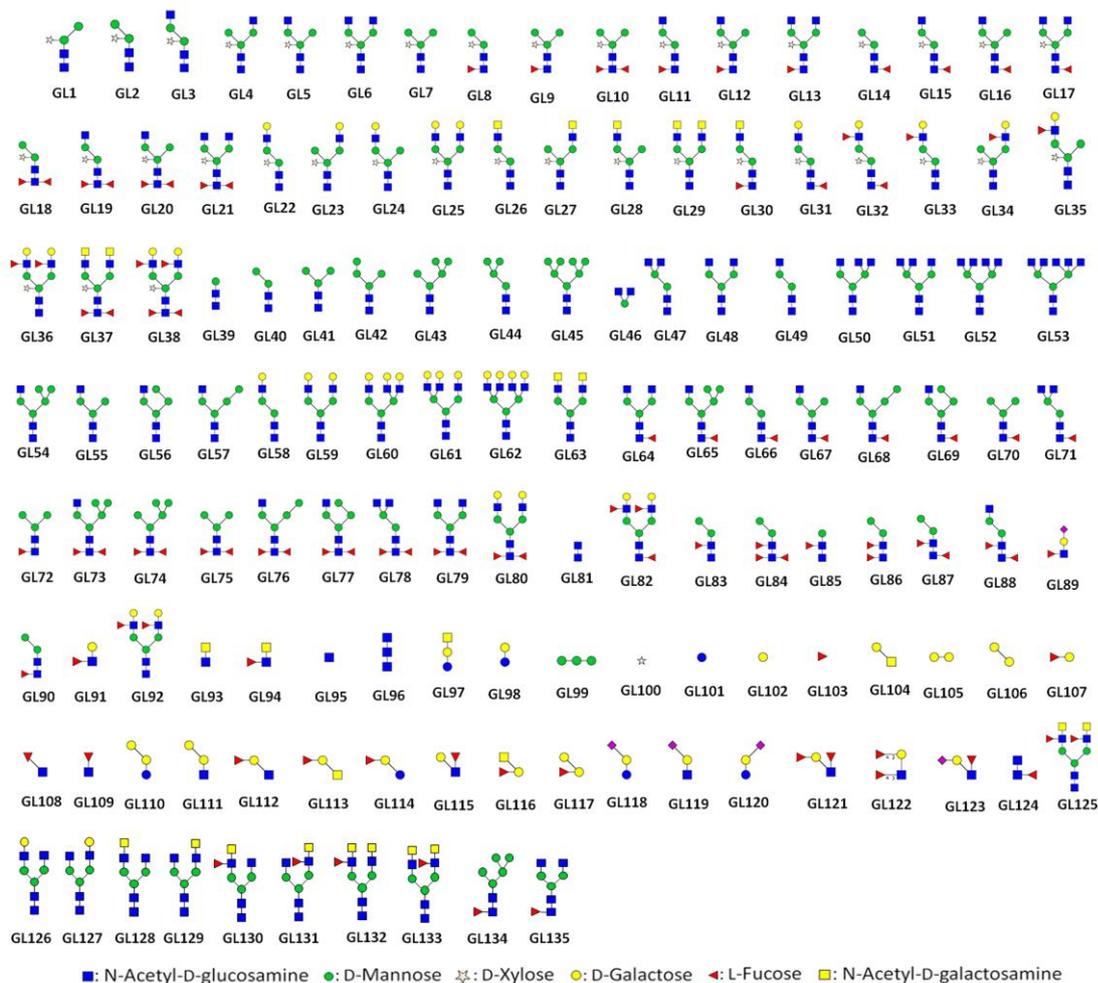
- 1) *in vitro* refolding can result in non-functional CLRs, as explained in chapter 8.1, and, when possible, lectin array or FACS analysis were performed to assess the protein functionality. Nevertheless, glycan microarray was used as an additional technique to confirm the effective ligand recognition by the refolded CLRs.
- 2) These glycans and glycomimetics microarrays were for the first time tested with the panel of human CLRs that we have produced. While some of the identified interactions were expected, others constitute new hits that need to be investigated in more details.

9.1 Glycan array

The glycan array was composed of 135 N-, complex and hybrid glycans, with structures characteristic of parasites and fungi. All the glycans were printed by Sonia Serna, a researcher of Niels Reichardt's laboratory. Figure 104a indicates the glycan stereochemistry and Figure 104b gathers the panel of glycans that were screened.



a



b

Fig.104 Glycan array a) glycan stereochemistry and b) panel of the 135 glycans printed on the array.

The following CLRs were tested: BDCA2, LSEctin and MCL CRDs, DC-SIGN, DC-SIGNR, dectin-2, langerin ECDs. DC-SIGNR and MCL TETRALEC molecules were also tested, together with the random TETRALECs of BDCA2 and LSEctin.

Particular focus was given to the interaction with six new glycans, namely **GL126,127,128,129,130,131**. These glycans are positional isomers and very interestingly three CLRs show a clear recognition selectivity for one elongated branch over the other. These results are described in the following revised paper.

9.1.1 Chemoenzymatic Synthesis of N-glycan Positional Isomers and Evidence for Branch Selective Binding by Monoclonal Antibodies and Human C-type Lectins Receptors

The incomplete enzymatic elongation during the N-glycosylation performed by our collaborators at the CICbiomaGUNE for the glycan array synthesis gave rise to isomeric structure that only differs in one position. We explored the potentially selective interaction of these positional antennae isomers with three CLRs: DC-SIGN ECD, DC-SIGNR ECD and LSECtin CRD. DC-SIGN showed significant binding to the mono-galactosylated biantennary glycan but not to the positional isomer galactosylated in the 3-arm only. Incubation with DC-SIGNR gave a different picture. It bound to the opposite mono galactosylated positional isomers (preference for a 6-arm extension). LSECtin behaved similarly to DC-SIGN with a preferential interaction for galactose or a N-acetyl galactose capping the 6-arm. Therefore, we identified pairs of positional N-glycan isomers that selectively bind to DC-SIGN, DC-SIGNR and LSECtin.

The main outcome of these studies:

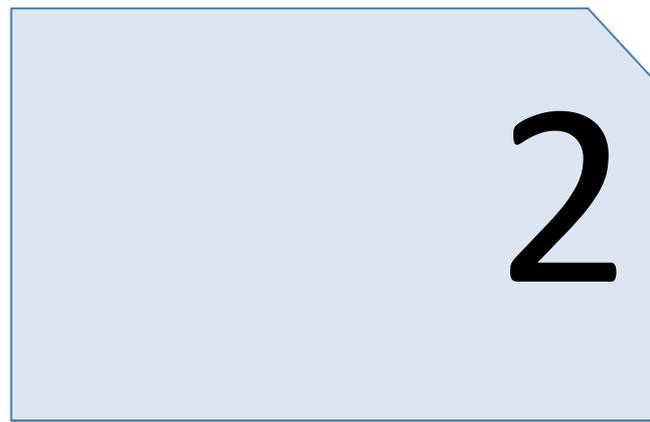
The exploration of potentially selective interaction of positional antennae opens the door for further improvement on selectivity and affinity of the identified epitopes.

Contributions:

All the glycans were designed and printed by the group of Niels Reichardt, while the antibodies were provided by the group of Hokke (NL).

My contribution to this study:

I have prepared DC-SIGN ECD, DC-SIGNR ECD and LSECtin CRD required for glycan microarray experiments and performed all the described assays involving CLRs, analysed data and participated to the preparation of the paper manuscript.





Chemoenzymatic Synthesis of N-glycan Positional Isomers and Evidence for Branch Selective Binding by Monoclonal Antibodies and Human C-type Lectin Receptors

Begoña Echeverria,^{†,‡} Sonia Serna,^{†,‡} Silvia Achilli,^{‡,‡} Corinne Vivès,[‡] Julie Pham,[†] Michel Thépaut,[‡] Cornelis H. Hokke,[§] Franck Fieschi,[‡] and Niels-Christian Reichardt^{*,†,||}

[†]CIC biomaGUNE, Glycotechnology Laboratory, Paseo Miramón 182, 20014 San Sebastian, Spain

[‡]Université Grenoble Alpes, CEA, CNRS, IBS, F-38000 Grenoble, France

[§]Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands

^{||}CIBER-BBN, Paseo Miramón 182, 20014 San Sebastian, Spain

Supporting Information

ABSTRACT: Here, we describe a strategy for the rapid preparation of pure positional isomers of complex N-glycans to complement an existing array comprising a larger number of N-glycans and smaller glycan structures. The expanded array was then employed to study context-dependent binding of structural glycan fragments by monoclonal antibodies and C-type lectins. A partial enzymatic elongation of semiprotected core structures was combined with the protecting-group-aided separation of positional isomers by preparative HPLC. This methodology, which avoids the laborious chemical differentiation of antennae, was employed for the preparation of eight biantennary N-glycans with Gal β 1,4GlcNAc (LN), GalNAc β 1,4GlcNAc (LDN), and GalNAc β 1,4-[Fuc α 1,3]GlcNAc (LDNF) motifs presented on either one or both antennae. Screening of the binding specificities of three anti-Le^x monoclonal IgM antibodies raised against *S. mansoni* glycans and three C-type lectin receptors of the innate immune system, namely DC-SIGN, DC-SIGNR, and LSECtin, revealed a surprising context-dependent fine specificity for the recognition of the glycan motifs. Moreover, we observed a striking selection of one individual positional isomer over the other by the C-type lectins tested, underscoring the biological relevance of the structural context of glycan elements in molecular recognition.

N-glycosylation is a very common modification of the majority of eukaryotic proteins. N-glycans have important intracellular functions in protein folding, targeting, and degradation and play extracellular roles in protein stability, antigenicity, and molecular binding and signaling in cellular communication.^{1–3} The heterogeneity generally observed in protein glycosylation is a consequence of the nontemplate nature of glycan biosynthesis that is controlled by expression levels of glycosyltransferases and transporter proteins, nucleotide sugar donor concentrations, and protein–protein and protein–carbohydrate interactions around the glycosylation sites. Hence, glycan profiles are not only species-, tissue-, cell-, or protein-specific but often altered under pathological or changing physiological conditions, underscoring their potential as disease markers.⁴

The structural heterogeneity of complex N-glycans is largely due to the variable enzymatic elongation of terminal N-acetylglucosamine (GlcNAc) moieties with galactose, sialic acid, and fucose residues on bi-, tri-, and tetra-antennary glycans. The majority of N-glycan structures found in nature have nonidentical antennae, giving rise to isomeric and structural

variations in the terminal motifs that form specific ligands for glycan-binding proteins.^{5–7}

Molecular recognition processes based on carbohydrate–protein interactions have been found to mediate a plethora of biological processes including cell adhesion, cell development, protein trafficking, or the differentiation of self- and nonself by the immune system.^{4,8} The usually weak monovalent interactions between carbohydrates and proteins are reinforced by a multivalent presentation of glycans on cell surfaces and through multimerization of receptors and glycan binding domains. Therewith, affinities can be increased up to 6 orders of magnitude, leading to binding strengths comparable to those of protein–protein interactions.^{9,10} On the molecular level, carbohydrate–protein interactions are stabilized by hydrogen and water mediated hydrogen bonding, calcium coordination, hydrophobic interactions between aliphatic amino acid side chains with CH groups, and CH– π interactions between

Received: May 9, 2018

Accepted: June 12, 2018

Published: June 12, 2018

aromatic amino acid side chains and sugar residues.^{11–13} Ultimately, glycan motif presentation, *e.g.*, by glycan clustering, as repeating units in polysaccharides or on the antennae glycans of secreted glycoproteins, can influence the strength and specificity of glycan protein interactions.^{8,10}

In relation to pathogen recognition and the discrimination between self and nonself on the basis of glycan determinants, two important classes of carbohydrate binding proteins are antiglycan antibodies and lectins of the innate immune system. Immune responses to viral, bacterial, and parasitic infections are often orchestrated by glycan-induced molecular mechanisms.

In various parasitic worm infections, including the major human disease schistosomiasis, the antibody response is dominated by antibodies directed against antigenic glycans expressed on the parasite surface or on secreted glycoprotein antigens.¹⁴ Knowledge of the exact binding specificities of these (monoclonal) antibodies is highly relevant for the development of diagnostic assays, in imaging, and in the development of glycan-targeting therapeutics.

One important class of mammalian glycan-binding receptors of the innate immune system is the C-type lectin receptors (CLRs). CLRs are pattern recognition receptors (PRRs) located on immune cells and involved in the recognition and uptake of both self and nonself glycoconjugates. Beyond their important functions in pathogen and/or antigen recognition and uptake, some CLRs are multifunctional receptors mediating cell adhesion, glycoprotein clearance, or further signaling events.^{10,15} With few exceptions, CLRs with sugar binding affinity share a common fold for their carbohydrate binding domain and a well-defined primary sugar-binding site that involves the coordination of calcium ions.¹⁵ On the basis of the specificity of their carbohydrate binding domain (CRD), CLRs are roughly grouped into mannose/fucose and galactose/GlcNAc binding receptors, but the fine specificity and additional affinity is gained through binding of oligosaccharides to an extended binding domain or *via* secondary binding sites.^{15,16}

In this study, we have explored the influence of the structural context of N-glycan antenna ligands for binding of three anti-Lewis X antibodies directed against *S. mansoni* glycans and three C-type lectins, DC-SIGN, DC-SIGNR, and LSECtin.^{17,18} DC-SIGN and DC-SIGNR are closely related CLRs that bind Ebola, hepatitis C, and human immune deficiency viruses *via* high mannose and complex type N-glycans of the viral surface proteins. While DC-SIGN induces specific immune responses upon interaction with numerous pathogens, DC-SIGNR in contrast seems to be only an adhesion receptor.^{17,19} Both receptors have also been reported to enhance trans-infection of T-cells even at low viral loading by capture and efficient surface presentation of HIV to CD4 presenting cells.^{20,21} L-SECtin recognizes Ebola virus *via* complex-type N-glycans²² of the viral envelope protein, as well as West Nile filovirus and SARS coronavirus,²³ but not HIV or hepatitis C virus.²⁴ It also interacts with activated T-cells.²⁵

Targeting antigen to densely expressed dendritic cell receptors such as DC-SIGN is an attractive strategy for vaccine efficacy improvement and immunotherapy²⁶ and has been achieved both with antilectin antibodies and with glycans recognized by the targeted lectin receptor.^{27–30} A lack of specificity due to an overlap in glycan binding specificities of lectin receptors can be problematic for dendritic cell targeting approaches, especially if various receptors with opposing functions are responsive. The identification of high affinity and selective ligands that specifically target a single receptor on

dendritic cells is an active area of research not only for immunotherapy and vaccine development but also for understanding lectin receptor function in antigen recognition, uptake, and processing. Previous glycan array profiling experiments have shown that DC-SIGN and DC-SIGNR both recognize Man α 1,2Man residues, preferentially in high mannose type glycans, but that only DC-SIGN shows affinity to blood group antigens by binding to Fuc α 1,3GlcNAc and Fuc α 1,4GlcNAc residues.^{17,31,31} Recently, Guo *et al.* have shown differentiated binding patterns of tetravalent DC-SIGN R and DC-SIGN as a function of the mannose residue density on glycan functionalized quantum dots. While DC-SIGN bound efficiently to quantum dots displaying mannose residues at high density, the spatially differently oriented DC-SIGNR CRDs could not engage in a sufficient number of contacts for strong binding.³² Other screening efforts have also shown strong binding of DC-SIGN to biantennary GlcNAc terminating glycans, which is however completely lost in the presence of a core α 1,2 xylose residue.³³

All binding studies in this work were performed on an expanded version of our existing synthetic glycan array. It includes a large number of N-glycans reflecting a significant part of the structural heterogeneity found for antennae type and number, core modifications, and terminal sugar residues and an additional collection of smaller glycan fragments and blood group antigens.^{33,34} For this study, we broadened our glycan array with a series of complex biantennary N-glycans presenting Gal β 1,4GlcNAc (LacNAc), GalNAc β 1,4GlcNAc (LDN), and GalNAc β 1,4[Fuc α 1,3]GlcNAc (LDNF) motifs on either one or both antennae (Figures 2 and 3). In 2012, Benevides *et al.* observed the unusual specificity of a lectin from the legume *Platypodium elegans* for asymmetrically branched glycans.³⁵ A year later, the Boons group presented a strategy for the synthesis of N-glycans with differentiated antennae and found significant differences in the binding affinities of some antennae isomers to certain plant lectins and to hemagglutinins involved in influenza virus binding. This work suggested a role for positional isomers in molecular recognition processes.³⁶ Since then, other reports of differential binding patterns of asymmetrically branched glycans and strategies for their synthesis have appeared.^{37–39} The chemo-enzymatic synthesis of asymmetrically extended complex multiantennary glycan isomers was first reported by the Boons group, based on an orthogonally protected central pentasaccharide scaffold and on a set of partially protected advanced glycosyl donors.³⁶ Subsequent enzymatic glycosylations elongated only fully deprotected antennae residues while acetate protected terminal sugars remained unchanged. However, this elegant approach suffered from incomplete enzymatic reactions, leading to compound mixtures that were arduous to separate.⁴⁰ Moreover, an elaborate chemical synthesis has to be accomplished for every underlying scaffold structure. This prompted us to seek an alternative approach for the preparation of N-glycan positional isomers.

Recently, we have developed PADS, the Protecting-group-Aided Detection and Separation of glycans, as a very robust methodology to quickly prepare libraries of all enzymatically accessible positional isomers from just a handful of multi-antennary basic glycan scaffolds.³⁸ LDN and LDNF constitute CLR ligands that are very selectively expressed on mammalian N-glycoproteins⁴¹ but that are highly abundant for instance in helminth parasites, including *S. mansoni*, one of the most significant worm parasites of humans. We studied the interaction of positional N-glycan isomers with several antiglycan

Scheme 1. Chemical Synthesis of Biantennary N-Glycan 1: (a) 10% TMSOTf, CH₂Cl₂, m.s., -20 °C, 76%; (b) BF₃·OEt₂, EtSH, CH₂Cl₂, 79%; (c) 4, 10% TMSOTf, CH₂Cl₂, m.s., -45 °C, 52%; (d, i) NH₂(CH₂)₂NH₂, *n*BuOH; (ii) Ac₂O, pyridine; (iii) NaOMe, MeOH; Below: Pictogram Presentation of Semiprotected Glycans 1 and 2 with Benzyl Protected Monosaccharide Residues Boxed in Light Blue

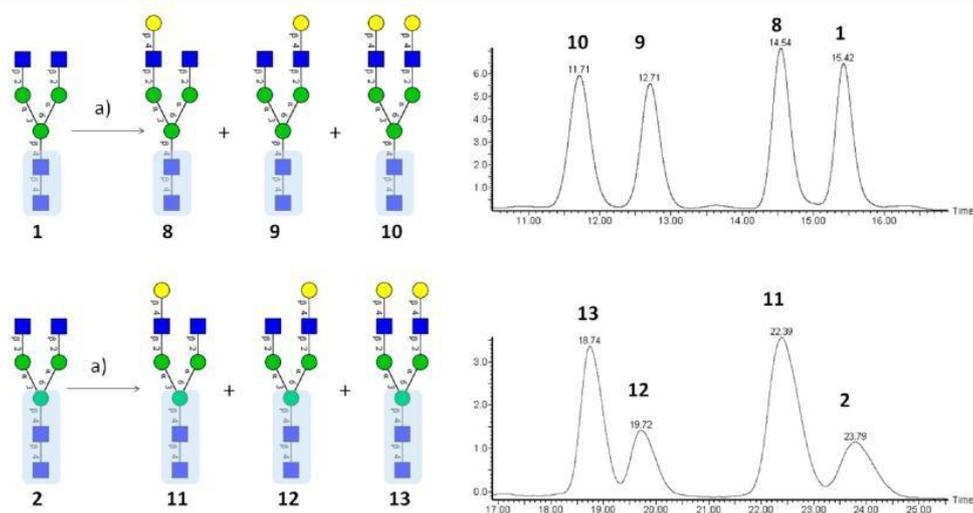
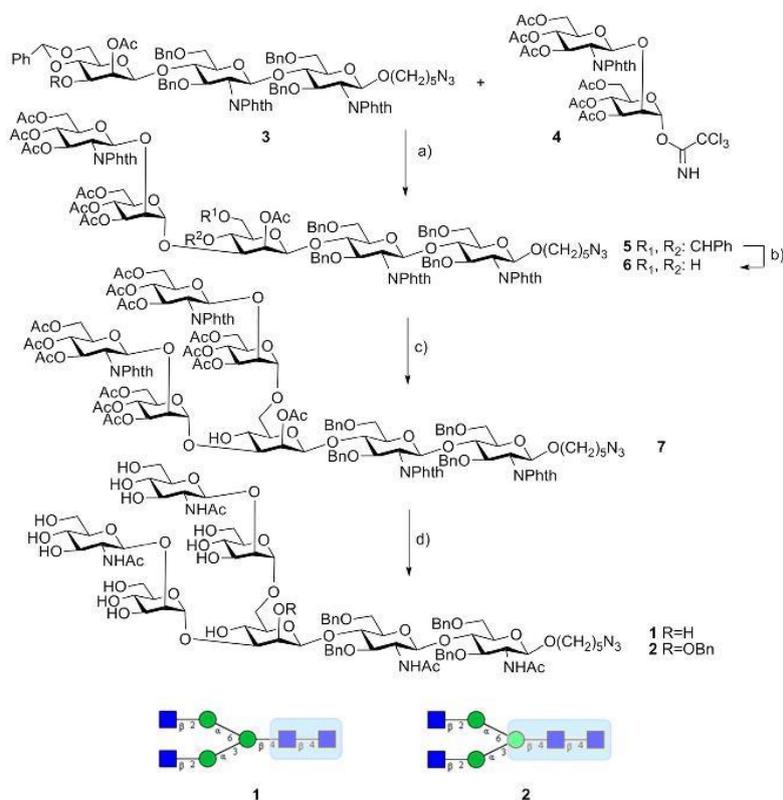


Figure 1. Galactosylation reaction of 1 and 2 and the corresponding purification chromatograms. (a) GalT, UDP-Gal, HEPES 50 mM at pH 8.5, MnCl₂ 10 mM, 37 °C. Galactosylation reaction of 1, product distribution: 8, 26.5%; 9, 22.3%; 10, 25.4%; 1, 25.8%. Galactosylation reaction of 2, product distribution: 11, 44.1%; 12, 12.4%; 13, 29.2%; 2, 14.3%.

monoclonal antibodies and human C-type lectins involved in pathogen and self recognition. These new binding data support the idea of isomer-specific roles in carbohydrate mediated molecular recognition.

RESULTS AND DISCUSSION

Synthesis. Enzymatic reactions toward the preparation of biantennary N-glycans presenting LDN and LDNF motifs were performed on the two partially benzylated glycan scaffolds 1 and

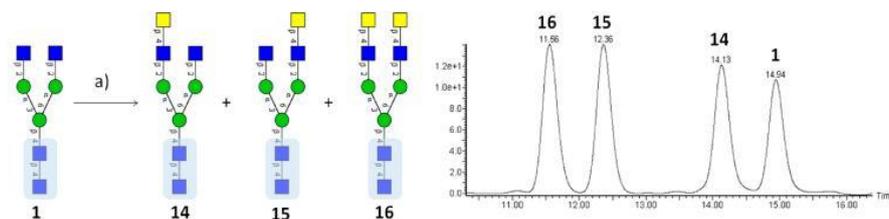


Figure 2. Enzymatic N-acetyl-galactosamination of **1**. (a) GalT mutant, UDP-GalNAc, HEPES 50 mM pH 8.5, MnCl_2 10 mM, 37 °C. Product distribution: **14**, 24.4%; **15**, 27.5%, **16**, 27.6%, **1**, 20.5%.

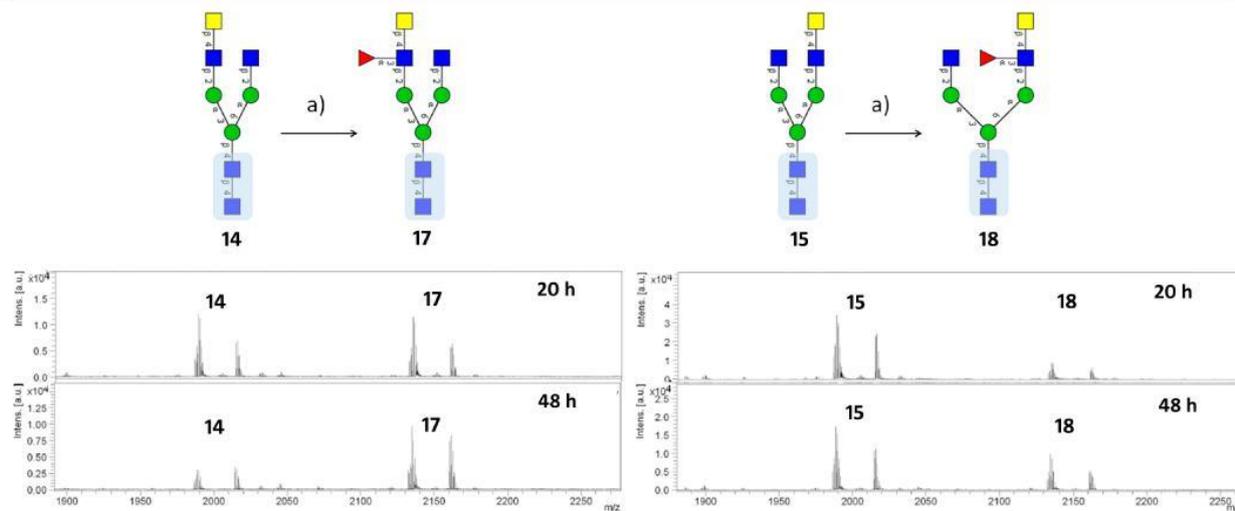


Figure 3. MALDI-TOF MS spectra of the fucosylation of **14** (left) and **15** (right) after 20 h and 48 h, respectively. (a) GDP-Fuc, CeFUT6, MES 40 mM, MnCl_2 10 mM.

2, with compound **2** presenting an additional benzyl group at C2 of the central mannose residue. The synthesis of both scaffolds **1** and **2**⁴² follows a [3 + 2 + 2] building block approach and is described in detail in the [Methods](#) section (see also [Scheme 1](#)).

Enzymatic Synthesis. The enzymatic extension of **1** and **2** synthesis to glycans presenting terminal LN (*N*-acetyl lactosamine), LDN, and LDNF moieties was carried out with a recombinant bovine milk galactosyltransferase (GalT), a mutant of bovine milk galactosyltransferase with GalNAc-transferase (GalNAcT) activity,⁴³ and a Lewis X type fucosyltransferase (FucT) from *C. elegans*. The presence of benzyl groups distant to the acceptor site had little or no influence on the outcome of the enzymatic reactions. To ensure a largely equimolar product distribution, we monitored the enzymatic reaction by UPLC-MS and carefully adjusted the substrate concentration, the incubation time, and the amounts of both added enzyme and nucleotide donor UDP-Galactose. Tuning of reaction conditions including time, enzyme, and donor concentrations afforded both regioisomers **8** and **9**, alongside the bis-elongated product **10** and the starting glycan **1** in nearly equimolar amounts, in line with the known indiscriminate action of the enzyme on the 3- and 6-arms of biantennary N-glycans⁴⁴ ([Figure 1](#)).

For the galactosylation of **2**, however, with an additional 2-*O*-benzyl group in the central mannose residue ([Scheme 1](#)), we observed the preferred formation of one regio-isomer over the other ([Figure 1](#)). Separation of the reaction components by preparative HPLC and structural assignment of both isomers by

NMR revealed the 3-arm as the preferred branch for galactosylation (**11**, 44.1%). The added steric bulk conveyed by the benzyl group apparently gave rise to an unfavorable interaction with the enzyme for the elongation of the GlcNAc residue on the 6-arm. Consequently, only 12.4% of the 6-isomer **12** was formed. This protecting group induced selectivity was surprising, and a more systematic investigation of the steric effects of residual protecting groups as a means of modulating the selectivity of glycosyltransferases and glycosidases is currently underway in our lab.

For the rapid preparation of near equimolar amounts of all regio-isomers, however, the 2-*O*-benzyl-protected glycan **2** seemed less suited and was therefore abandoned as a scaffold for further enzymatic reactions. For the preparation of terminal GalNAc glycans, we incubated glycan **1** with UDP-GalNAc and GalNAcT and carefully adjusted the amount of enzyme, of nucleotide donor and the reaction time to produce similar amounts of both regioisomers **14** and **15**, starting material **1**, and bis-glycosylated product **16**, which were easily separated by preparative HPLC ([Figure 2](#)). Benzyl protected aliquots of **14**, **15**, and **16** were subjected to further enzymatic derivatization with a Lewis type fucosyltransferase to produce the LDNF immunogenic element on one or both arms of the biantennary N-glycan. Compounds **14** and **15** were incubated with CeFUT6 and GDP-fucose and the enzymatic conversion monitored for each reaction by MALDI-TOF mass spectrometry ([Figure 3](#)). After 20 h of incubation, around 50% of isomer **14** had reacted, and after 48 h, the fucosylation had reached completion. The

isomer **15** with the GalNAc residue on the 6-arm reacted substantially slower, showing only 10% conversion to the fucosylated product **18** after 20 h (Figure 3). The reaction was left for 48 h until no further conversion to **18** was observed by MALDI-TOF MS, and the unreacted starting material was removed by preparative HPLC.

Treatment of the glycan **16** with CeFUT6, an α -1,3 fucosyltransferase that catalyzes the addition of a fucose residue at the distal GlcNAc of the chitobiose core in *C. elegans* N-glycans⁴⁵ and which also shows activity toward terminal LN and LDN residues,⁴⁶ provided a mixture of regioisomers **19** and **20** and bis-glycosylated glycan **21**. Although the enzyme adds fucose on GalNAc residues of both antennae, the chromatographic profile (Figure 4) suggests preferential glycosylation of

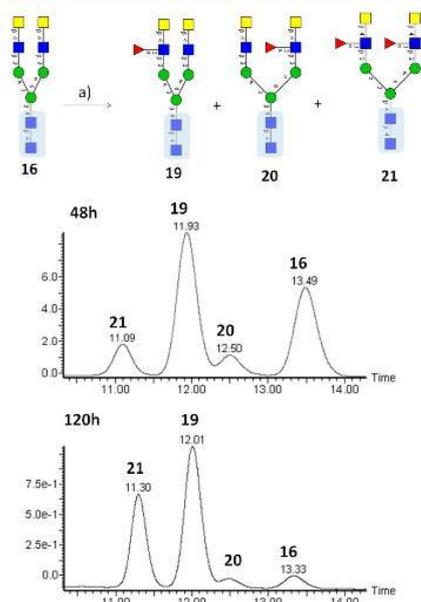


Figure 4. α -1,3-Fucosylation of compound **16**. (a) GDP-Fuc, CeFUT6, MES 40 mM, MnCl_2 at 10 mM. Chromatogram of the reaction after 48 and 120 h, respectively. Product distribution after 48 h: **19**, 55.6%; **20**, 4.6%; **21**, 11.8%; **16**, 28.0%. After 120 h: **19**, 55.8%; **20**, 1.7%; **21**, 35.7%; **16**, 6.8%.

the 3-arm prior to addition on the 6-arm, a selectivity which we believe has not been described before. Even after 5 days of incubation with CeFUT6 and the GDP-fucose donor and after most starting glycan had been consumed, only trace amounts of

the 6-isomer **20** (1.7%) were formed. To gain access to useful amounts of region-isomer **20**, we treated the monofucosylated glycan **18** with GalNAcT and UDP-GalNAc.

After purification by semipreparative HPLC, compounds **8–10**, **14–15**, and **17–21** groups were deprotected by Pd/C catalyzed continuous flow hydrogenation (for details, refer to the SI) providing eight new positional isomers **22–29** with LN, LDN, and LDNF structural elements (Figure 5), which were included in our N-glycan microarray alongside our existing collection of 126 synthetic glycans onto NHS-activated glass microarray slides (see SI for all structures included on the glycan array).

Carbohydrate Fine Specificity of Anti-Le^x Antibodies Derived from *S. mansoni* Infected Mice. Printed arrays were first used to study the glycan binding specificities of monoclonal antibodies produced by hybridomas generated from the spleen cells of *S. mansoni* infected mice.^{47,48} Schistosomiasis is an infectious disease caused by blood flukes of the *Schistosoma* genus that affects more than 200 million people worldwide.⁴⁹ *S. mansoni* parasites are highly glycosylated during all developmental stages and present a structurally rich variety of glycans including fucosylated LN and LDN motifs on their cell surface and secreted glycoproteins. Both innate and adaptive immune responses toward schistosome glycans are generated upon infection in humans, including the formation of antiglycan antibodies.^{48,50} Several years ago, Remoortere *et al.* studied the glycan binding specificities of a library of 188 mAbs derived from *Schistosoma*-infected or immunized mice employing a number of small synthetic glycan conjugates.^{49,51,52} Antibodies were grouped by binding specificity for Le^x, LDN, and LDNF epitopes, found in different developmental stages of *S. mansoni* and in addition classified for distinguishing monomeric from polymeric epitope presentation. The antibodies were then employed to visualize the spatial distribution of glycan epitopes in various stages of the parasite by immunofluorescence assays.^{49,53} Here, we have screened a selection of mAbs known to bind Le^x epitopes to assess the effect of epitope presentation within the context of a larger N-glycan by glycan array methodology. Printed glycan arrays were incubated with diluted antibody solutions, and binding was detected by incubation with a fluorescently labeled secondary antimouse immunoglobulin antibody. Figure 6 shows the specificities of three Le^x binding monoclonal IgM antibodies for glycans presenting the Le^x epitope. The mAb 128–4F9 recognizes Le^x only as the free nonextended trisaccharide epitope, but not within the context of a larger N-glycan as neither GL32–GL36, GL38, GL82, nor GL92 presenting the same epitope on the 3- and/or the 6-arm are bound. In contrast, the mAb 99–1G3 recognizes Le^x

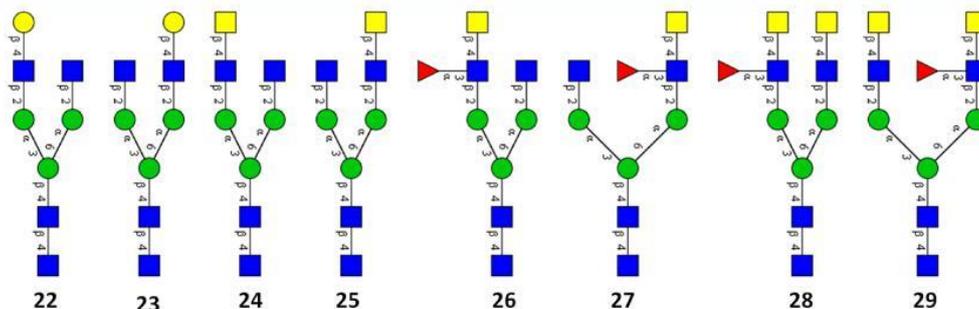


Figure 5. Pictogram representation of the eight newly synthesized positional LN, LDN, and LDNF isomers.

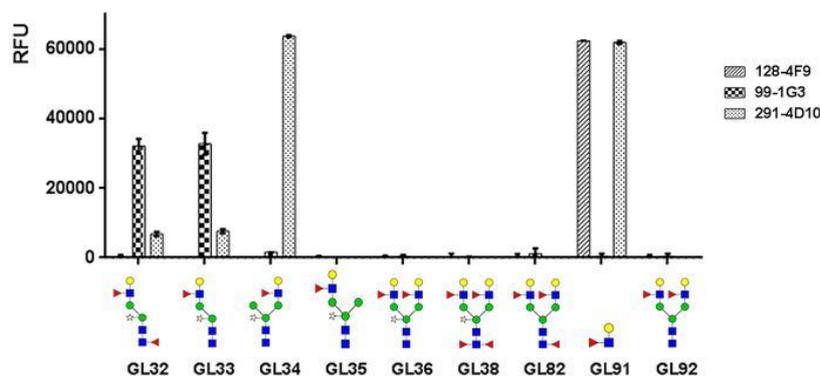


Figure 6. Glycan binding profile of *S. mansoni*-related monoclonal IgM antibodies 128–4F9, 99–1G3, and 291–4D10 toward Le^x containing structures. RFU values were analyzed after incubation with Alexa Fluor-555 antimouse IgM. Each histogram represents the mean RFU values for four spots with the SD of the mean.

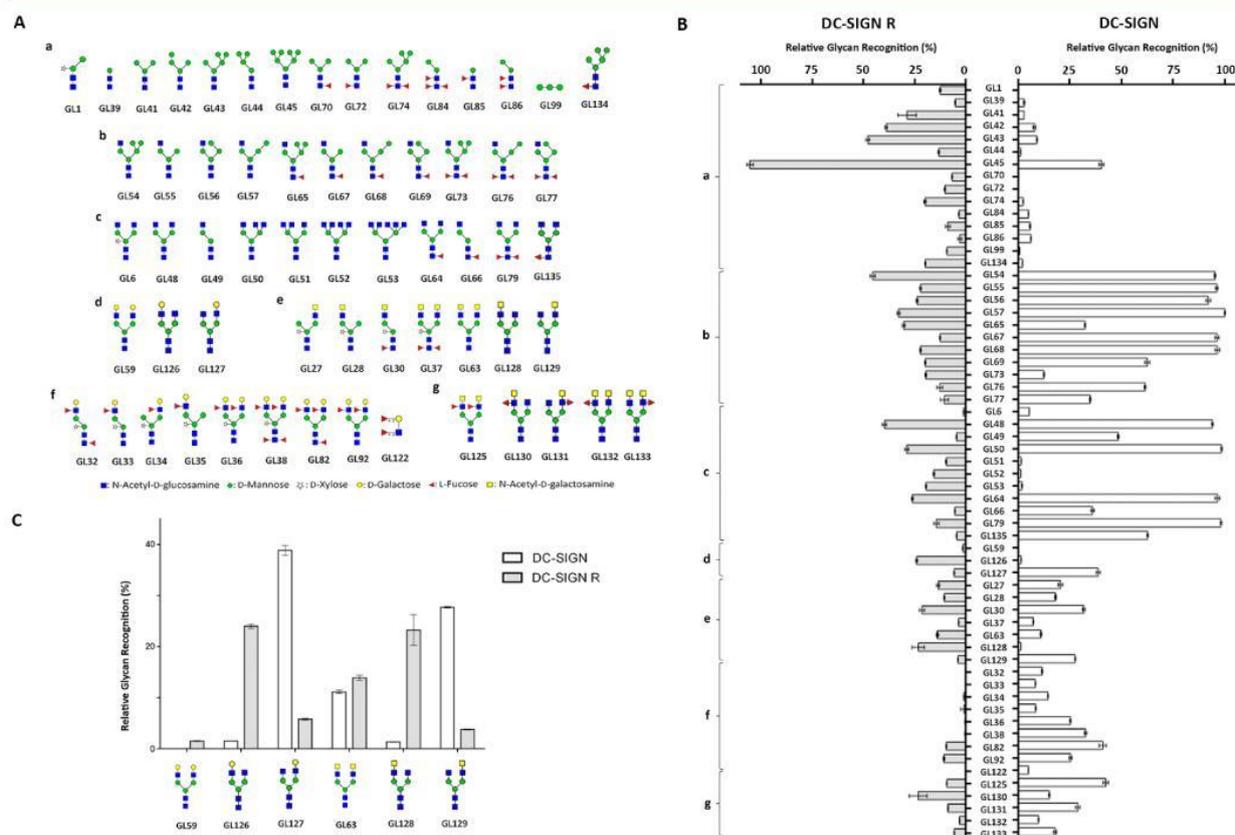


Figure 7. Glycan binding profile of DC-SIGN ECD and DC-SIGNR ECD. (A) Schematic representation of glycan structures binding to DC-SIGN and/or DC-SIGNR grouped by structure: (a) high mannose related structures, (b) hybrid N-glycans, (c) complex GlcNAc terminating N-glycans, (d) galactose containing N-glycans. (e) LDN containing N-glycans, (f) Le^x containing N-glycans and Le^y, (g) LDNF containing N-glycans. (B) Histograms representing the % RFU values for DC-SIGN and DC-SIGNR binding. (C) Comparison of %RFU values of monogalactosylated and N-glycans with terminal single GalNAc residues in DC-SIGN and DC-SIGNR binding.

exclusively on the 3-arm of the two core xylosylated glycans (GL32, GL33), while other compounds presenting the epitope in a spatially different manner, including the free epitope, are not bound. The third Le^x binding mAb 291–4D10 displays yet another specific binding profile as it recognizes both the free epitope and a structure presenting the epitope on the 6-arm of a core-xylosylated N-glycan (GL34). Compounds presenting the

epitope on the 3-arm (GL35) or bis-functionalized compounds, GL36 and GL38, were however not bound at all by 291–4D10. A rationale for this highly specific context-dependent binding is difficult to provide without additional structural information regarding the carbohydrate antibody complex. However, a possible explanation for the observed fine-specificity of these mAbs would be that their binding site extends beyond the Le^x

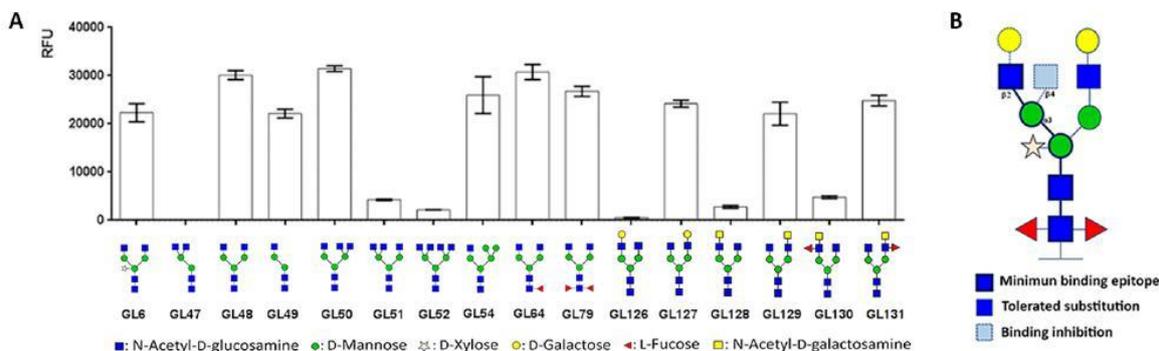


Figure 8. (A) Cy3-labeled L-SECTin CRD binding profile toward selected N-glycan structures. Each histogram represents the mean RFU values for four spots with the SD of the mean. (B) Preferred epitope found in complex and hybrid N-glycans for binding toward L-SECTin CRD.

trisaccharide itself. The mAbs 99–1G3 and 291–4D10 appear to be selective for the α 3- and α 6-arms of core xylosylated glycans, respectively, indicating that the mannose residue and the context of the trimannosyl core may take part in the mAb–glycan interactions. The observed specificity might also be the result of compound exclusion due to sterical clashes of antennae residues with the lectin rather than binding of a larger epitope *via* an extended binding groove. In contrast, 128–4F9 may prefer a more freely accessible Le^x motif such as in the context of the synthetic trisaccharide-linker conjugate. Crystallographic studies⁵⁴ indicated that glycan-binding pockets of Fab fragments can indeed accommodate oligosaccharides larger than a trisaccharide, including a linear trimer of Le^x. Interestingly, an at that time unspecified difference in the binding of Le^x-containing N-glycans between 99 and 1G3 and 291–4D10 was observed earlier using a glycan microarray of naturally occurring glycans isolated from *S. mansoni*.⁵⁵ N-glycans containing a single Le^x antenna were bound by 291–4D10, but not by 99–1G3, suggesting that the so far unspecified location of the Le^x motif in these cercarial glycans was on the α 6-arm rather than the α 3-arm. Therefore, unraveling the fine specificity of antiglycan mAbs by defined synthetic glycan arrays is helpful in defining isomeric and asymmetric glycan structures from natural sources.

The exquisite antiglycan specificity of some monoclonal antibodies found in this study highlights the importance of epitope selection and presentation and could have important consequences, *e.g.*, in immunoassays employing natural or synthetic glycan antigens such as those that detect allergen-directed IgEs.⁵⁶

Screening of the Carbohydrate Binding Specificity of C-Type Lectin Receptors. Here, printed arrays including the novel structures 22–29 were used to screen the C-type lectins DC-SIGN, DC-SIGNR, and LSECTin. Both DC-SIGN and DC-SIGNR extracellular domains (ECDs) were incubated on our array seeking ligands with selectivity for one of these two closely related lectins. Figure 7 shows strong binding of DC-SIGN to hybrid (GL54–GL57) and complex type glycans (GL48–GL50) and N-glycans presenting Le^x epitopes (GL82 and GL92). One of the most strongly bound glycans in our collection was, as previously shown, the complex type biantennary glycan GL48.³³ We also observed a significant binding to the biantennary glycan monogalactosylated on the 6-arm (GL127) but not to the positional isomer galactosylated on the 3-arm (GL126). A similar behavior was observed for N-glycans with a single terminal N-acetylgalactosamine (GalNAc) residue (GL128 and GL129). Incubation of DC-SIGNR on the

same arrays gave a slightly different picture than DC-SIGN (Figure 7). The strongest binding glycan was the high mannose glycan GL45 with an atypical non-natural branching pattern presenting the Man α 1–3(Man α 1–6Man) trisaccharide on both 3 and 6 arms. GL45 was less strongly bound by DC-SIGN. Even more interestingly, we observed opposite binding behavior of DC-SIGNR and DC-SIGN toward the monogalactose and mono-GalNAc positional isomers. Both galactosylated and GalNAc-terminating glycans GL126 and GL128 were specifically recognized by DC-SIGNR, while the corresponding positional isomers GL127 and GL129 were bound with far lower affinity.

Finally, we looked at the carbohydrate binding profile of the C-type lectin L-SECTin CRD to our glycan array. L-SECTin is widely expressed in lymph node and liver endothelial tissue, peripheral blood, and on macrophages.¹⁸ Its carbohydrate binding specificities have not been studied in much detail yet. A report from Dominguez-Soto *et al.* mentions binding to GlcNAc functionalized sepharose but not to mannan or GalNAc analogues.¹⁸ A later published binding profile to an array of 300 glycans showed high affinity of the CRD to GlcNAc β -1,2-Man containing glycan fragments, while fucosylated and mannosylated glycans were only identified as low affinity binders at very high receptor concentrations.²² Our results indicate an even higher selectivity of L-SECTin CRD for GlcNAc1,2-Man residues on the 3-arm of the complex and hybrid N-glycans. An additional GlcNAc residue in position C4 inhibited binding (structures GL47, Figure 8). The same GlcNAc1,2-Man disaccharide presented on the 6-arm of di-, tri-, or tetra-antennary complex N-glycans was bound to a far lesser extent (structures GL48, GL51, GL52). An additional capping with galactose on the 3-arm (GL126) completely abolishes the low binding of GlcNAc1,2-Man residues on the 6-arm. An N-acetylgalactosamine cap on the same arm is not as efficient in inhibiting this interaction (structures GL128 and GL129). The striking observation comes again from the differential interaction of LSECTin toward two positional isomers GL126 and GL127 where a galactose is capping the 3- and 6-arm, respectively. Similar behavior, albeit to a lesser extent, can be observed with two other couples of positional isomers with an LDN (GL128, GL129) motif or LDNF motif (GL130, GL131) present on the 3- and 6-arms, respectively. It is yet another example for the context specific binding of a well-defined structural element. This highlights the biochemical relevance of truncated or incompletely extended structures here for the

molecular recognition of glycans and glycoproteins by C-type lectin immune receptors.

The binding data of L-SECTin can be resumed in the preferred ligand GlcNAc β 1,2Man α 1,3Man. Substitution of this structural element with core xylose is tolerated, while capping with galactose or additional branching with β 1-4GlcNAc completely inhibits binding. Capping with GalNAc on the 3-arm or branching with β 1-6GlcNAc on the 6-arm does not affect the moderate binding observed for structures presenting a GlcNAc β 1,2Man residue on the 6-arm (structures GL48, GL51, GL52).

CONCLUSIONS

We and others have reported a growing number of examples where previously defined minimal ligand structures are differentially recognized by antibodies, lectins, and glycosyltransferases when presented in different structural contexts. This common feature in glycan recognition can be explained by the existence of more extended binding domains that can accommodate larger structural elements than previously thought. Another possible explanation could be the existence of different binding modes between an oligosaccharide and protein binding domain, with at times exquisite selectivity as suggested by some of our positional isomer selective binding data. The ever more extended structural annotation of glycan structures with simple pictograms that neglect glycosidic binding information might nurture the misconception of symmetrically equivalent antennae branches in N-glycans similar to a dendrimer. Detailed structural analysis, e.g., by NMR has shown large differences in the accessibility and the three-dimensional orientation for the different arms of the same N-glycan backbone in line with the different connectivity of monosaccharide residues. The structural asymmetry between antennae can now be visualized by NMR by attaching a paramagnetic lanthanide tag to the anomeric position of a multiantennary glycan. The induced pseudo-contact shifts decrease with the distance to the paramagnetic metal and provide spectral separation of peaks for the same residues present on different antennae.⁵⁷ The identification of a pair of positional N-glycan isomers that selectively bind to DC-SIGN and DC-SIGNR was an unexpected but welcome finding, which opens the door to further improvement on the selectivity and affinity of the identified ligands by medicinal chemistry approaches. Likewise, the discovery of the monogalactosylated N-glycan GL127 as a very strong binder of L-SECTin while the isomeric glycan does not show any binding requires further structural analysis, e.g., by NMR. Additional screening of other human lectins will show if this lead structure could eventually become a selective probe for targeting this important lectin involved in pathogen recognition and T-cell activation. Finally, the differences we observed for mAbs binding the Le^x epitope indicates that not only lectins but also the adaptive immune system (antibodies) can differentiate glycan motifs presented in different contexts.

METHODS

Chemical Synthesis of Semiprotected N-Glycan Precursors 1 and 2. Briefly, **1** was prepared by the sequential glycosylation of trisaccharide **3** with the disaccharide donor **4** on the central mannose residue. Glycosylation at position C3 under TMSOTf promotion provided pentasaccharide **5** in 76% yield. The Bn-acetal was subsequently removed by transacetalization with thioethanol and BF₃·OEt₂ to produce **6** in 79% yield. Glycosylation on OH-6 was

achieved with complete regioselectivity at low temperature and dilute reactant concentration again employing imidate **4** as a glycosyl donor and TMSOTf promotion to furnish the protected heptasaccharide **7** in 52% yield (Scheme 1). Phthalimide groups were removed by microwave assisted aminolysis with diaminoethane and the intermediate amine treated subsequently with acetic anhydride and sodium methanolate to arrive at glycan **1** with four remaining benzyl groups in the core chitobiose moiety as our key scaffold for enzymatic diversification. The synthesis of compound **2** followed an analogous route which has been described earlier.⁴²

C-Type Lectin Expression Cloning. Standard pUC57 plasmids containing optimized synthetic human genes encoding human DC-SIGN R ECD (amino acids 80–399) and LSECTin CRD (amino acids 162–292) designed for efficient production in *E. coli* were manufactured by GeneCust Europe. PCR amplification using suitable primers and restriction enzyme digestion were used to subclone DC-SIGN R ECD and N-terminal 6-His tagged LSECTin CRD into the pET30-b (Novagen) between the NdeI and HindIII restriction sites. The sequencing of each construction was done by Genewiz.

Protein Expression and Purification. The DC-SIGN extracellular domain (DC-SIGN ECD) was produced and purified as previously described.⁵⁸ DC-SIGNR ECD and L-SECTin CRD were expressed in *E. coli* BL21(DE3) in 1 L of LB culture supplemented with 50 μ g/mL kanamycin at 37 °C. Expression was induced by the addition of 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG) when the culture had reached an A_{600nm} of 0.8 and maintained for 3 h. The protein was expressed in the cytoplasm as inclusion bodies. Cells were harvested by a 20 min centrifugation cycle at 5000g at 4 °C. The pellet was resuspended in 30 mL of a solution containing 150 mM NaCl, 25 mM Tris-HCl at pH 8, and one antiprotease mixture tablet (Complete EDTA free, Roche). Cells were disrupted by sonication and cell debris eliminated by ultracentrifugation at 100 000g for 45 min at 4 °C with a Beckman 45Ti rotor. The pellet was solubilized in 30 mL of 6 M guanidine-HCl containing 25 mM Tris-HCl at pH 8, 150 mM NaCl, and 0.01% (v/v) β -mercaptoethanol. The mixture was centrifuged at 100 000g for 45 min at 4 °C, and the supernatant was diluted 5-fold with 1.25 M NaCl, 25 mM CaCl₂, and 25 mM Tris-HCl or 200 mM Tris-HCl at pH 8 (for DC-SIGNR or L-SECTin respectively) by slow addition with stirring. The diluted mixture was dialyzed against 10 volumes of 25 mM Tris-HCl at pH 8, 150 mM NaCl, and 4 mM CaCl₂ (buffer A) with three buffer changes. After dialysis, insoluble precipitate was removed by centrifugation at 100 000g for 1 h at 4 °C.

The supernatant containing DC-SIGNR ECD was loaded onto a mannan agarose column (Sigma) equilibrated with buffer A for purification by affinity chromatography. After loading, DC-SIGNR ECD was tightly bound to the column and eluted in the same buffer without CaCl₂ but supplemented with 1 mM EDTA (buffer B). This step was followed by Size Exclusion Chromatography (SEC) using a Superose 6 column equilibrated with buffer A. Fractions were analyzed by SDS-PAGE (12%), and DC-SIGN R ECD containing fractions were pooled and concentrated by ultrafiltration (YM10 Amicon membrane from Merck-Millipore).

The supernatant containing the His tagged LSECTin CRD was loaded onto a HisTrap HP column (GE Healthcare) at 4 °C. Unbound proteins were washed away with buffer A before LSECTin CRD was eluted with buffer C (150 mM NaCl, 25 mM Tris-HCl at pH 8, 4 mM CaCl₂, 0.5 M imidazole). Eluted fractions were analyzed by SDS-PAGE (15%), and the LSECTin CRD containing fractions were pooled and concentrated by ultrafiltration (YM10 Amicom membrane from Merck-Millipore). Each construct was checked by N-terminal amino acid sequencing and mass spectrometry.

Cy3 Labeling of DC-SIGN ECD, DC-SIGNR ECD, and LSECTin CRD. A total of 150 μ L of 10 mg mL⁻¹ solutions of DC-SIGN ECD, 250 μ L of 3.77 mg mL⁻¹ solutions of DC-SIGNR ECD, and 1 mL of 0.5 mg mL⁻¹ solution of LSECTin CRD in 25 mM HEPES at pH 7.25 and 4 mM CaCl₂ were prepared. Corresponding volumes (4.5, 3, and 2 μ L, respectively) of 10 mg mL⁻¹ Cy3-NHS ester (GeneCopoeia) were added to the DC-SIGN ECD, DC-SIGNR ECD, and LSECTin CRD solutions individually, and the reactions were gently shaken at RT for 2 h and then at 4 °C for 4 h. Excess dye was removed by two dialyses (3.5k

Z-lyser from Thermo Scientific) of 3 h against 25 mM Tris at pH 8, 150 mM NaCl, and 4 mM CaCl₂. The amount of attached Cy3 was estimated spectrophotometrically based on the dye (ϵ_{550} 150 000 cm⁻¹ M⁻¹) and protein molar absorption coefficients (ϵ_{280} DC-SIGN ECD 70 400 cm⁻¹ M⁻¹, ϵ_{280} DC-SIGNR ECD 60 890 cm⁻¹ M⁻¹, ϵ_{280} LSEctin CRD 48 845 cm⁻¹ M⁻¹). The obtained degree of labeling (DOL) was 0.95 for both DC-SIGN ECD and DC-SIGNR ECD and 0.2 for LSEctin CRD.

Microarray Preparation. Glycan microarrays were prepared as previously described.³³ Briefly, ligand solutions (50 μ M, 1.25 nL, five drops, drop volume: 250 pL) in sodium phosphate buffer (300 mM, pH 8.4, 0.005% Tween-20) were spatially arrayed employing a robotic noncontact piezoelectric spotter (SciFLEXARRAYER S11, Scienion) onto *N*-hydroxysuccinimide (NHS) activated glass slides (Nexterion H, Schott AG). After printing, the slides were placed in a 75% humidity chamber at 25 °C for 18 h. The remaining NHS groups were quenched with 50 mM solution of ethanolamine in 50 mM sodium borate buffer, at pH 9.0, for 1 h. The slides were washed with PBST (PBS solution containing 0.05% Tween 20), PBS, and water. The slides were dried in a slide spinner and stored at -20 °C until use.

Incubation with Monoclonal Antibodies. Monoclonal antibodies (mAbs) were produced as previously described.⁴⁹ Culture supernatants containing IgM mAbs against Le^x epitopes (128-4F9, 99-1G3, and 291-4D10) were diluted 1:200 in PBS containing 1% (w/v) bovine serum albumin (BSA) and 0.01% (v/v) Tween-20. MAb solutions (200 μ L per array) were used to incubate individual wells on a glycan array slide at RT for 1 h. The slides were washed with PBST (PBS containing 0.05% Tween-20) followed by PBS. Slides were then incubated with antimouse IgM-555 (1:1000) in PBS containing 1% BSA and 0.01% Tween-20 for 1 h in the dark. Arrays were washed from unbound secondary antibodies with PBST, PBS, and water. Microarrays were dried in a slide spinner, and fluorescence measurements were performed on a microarray scanner (Agilent G2565BA, Agilent Technologies) at 10 μ m resolution. Quantification of fluorescence was performed by ProScanArray Express software (PerkinElmer) employing an adaptive circle quantification method from 50 μ m (minimum spot diameter) to 300 μ m (maximum spot diameter). Average RFU values with local background subtraction of four spots and standard deviation of the mean were reported using Microsoft Excel and GraphPad Prism software.

Incubation with C-type Lectins. Cy3 labeled C-type lectins were diluted in incubation buffer (25 mM Tris-HCl, 150 mM NaCl, 4 mM CaCl₂, pH 7.5 containing 0.5% bovine serum albumin (BSA), and 0.005% Tween-20). C-type lectin solutions (200 μ L per array) were used to incubate individual wells on a glycan array slide at 4 °C for 18 h. Arrays were washed with incubation buffer without BSA and with water and dried in a slide spinner. Fluorescence measurements and analysis were performed as described above for monoclonal antibodies.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.8b00431.

Protocols for the chemo-enzymatic synthesis of asymmetric N-glycan structures, list of glycans printed on microarrays, and full histograms and images of microarray experiments (DOCX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: nreichardt@cicbiomagune.es.

ORCID

Niels-Christian Reichardt: 0000-0002-9092-7023

Author Contributions

[†]These authors contributed equally to the work.

Funding

We acknowledge funding from the Spanish Ministry of Economy and Competitiveness (MINECO, grants CTQ2014-58779-R and CTQ2017-90039-R) to N.-C.R. This project has also received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 642870 (ETN-ImmunoShape). For human CLRs production, this work used the Multistep Protein Purification Platform of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology.

Notes

The authors declare no competing financial interest.

■ REFERENCES

- (1) Varki, A. (2017) Biological roles of glycans. *Glycobiology* 27, 3–49.
- (2) Van Kooyk, Y., and Rabinovich, G. A. (2008) Protein-glycan interactions in the control of innate and adaptive immune responses. *Nat. Immunol.* 9, 593–601.
- (3) Ohtsubo, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* 126, 855–867.
- (4) Hart, G. W., and Copeland, R. J. (2010) Glycomics hits the big time. *Cell* 143, 672–6.
- (5) Gagneux, P., and Varki, A. (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9, 747–55.
- (6) Cummings, R. D. (2009) The repertoire of glycan determinants in the human glycome. *Mol. Biosyst.* 5, 1087–1104.
- (7) Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* 13, 448–62.
- (8) Li, R. E., van Vliet, S. J., and van Kooyk, Y. (2018) Using the glycan toolbox for pathogenic interventions and glycan immunotherapy. *Curr. Opin. Biotechnol.* 51, 24–31.
- (9) Liang, P.-H., Wang, S.-K., and Wong, C.-H. (2007) Quantitative analysis of carbohydrate-protein interactions using glycan microarrays: determination of surface and solution dissociation constants. *J. Am. Chem. Soc.* 129, 11177–84.
- (10) Dam, T. K., and Brewer, C. F. (2010) Lectins as pattern recognition molecules: the effects of epitope density in innate immunity. *Glycobiology* 20, 270–9.
- (11) Lee, Y. C., and Lee, R. T. (1995) Carbohydrate-protein interactions: basis of glycobiology. *Acc. Chem. Res.* 28, 321–327.
- (12) Weis, W. I., and Drickamer, K. (1996) Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* 65, 441–73.
- (13) Wu, A. M. (2012) *The Molecular Immunology of Complex Carbohydrates—2*, Springer Science & Business Media.
- (14) Hokke, C. H., and van Diepen, A. (2017) Helminth glycomics - glycan repertoires and host-parasite interactions. *Mol. Biochem. Parasitol.* 215, 47–57.
- (15) Drickamer, K., and Taylor, M. E. (2015) Recent insights into structures and functions of C-type lectins in the immune system. *Curr. Opin. Struct. Biol.* 34, 26–34.
- (16) Taylor, M. E., and Drickamer, K. (2009) Structural insights into what glycan arrays tell us about how glycan-binding proteins interact with their ligands. *Glycobiology* 19, 1155–62.
- (17) Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. A., Alvarez, R., Blixt, O., Taylor, M. E., Weis, W. I., and Drickamer, K. (2004) Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat. Struct. Mol. Biol.* 11, 591–598.
- (18) Dominguez-Soto, A., Aragoneses-Fenoll, L., Martin-Gayo, E., Martinez-Prats, L., Colmenares, M., Naranjo-Gomez, M., Borrás, F. E., Muñoz, P., Zubiaur, M., Toribio, M. L., Delgado, R., and Corbi, A. L. (2007) The DC-SIGN-related lectin LSEctin mediates antigen capture and pathogen binding by human myeloid cells. *Blood* 109, 5337–5345.

- (19) Den Dunnen, J., Gringhuis, S. I., and Geijtenbeek, T. B. H. (2009) Innate signaling by the C-type lectin DC-SIGN dictates immune responses. *Cancer Immunol. Immunother.* 58, 1149–57.
- (20) Geijtenbeek, T. B., Kwon, D. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C., Middel, J., Cornelissen, I. L., Nottet, H. S., KewalRamani, V. N., Littman, D. R., Figdor, C. G., and van Kooyk, Y. (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 100, 587–597.
- (21) Pöhlmann, S., Soilleux, E. J., Baribaud, F., Leslie, G. J., Morris, L. S., Trowsdale, J., Lee, B., Coleman, N., and Doms, R. W. (2001) DC-SIGNR, a DC-SIGN homologue expressed in endothelial cells, binds to human and simian immunodeficiency viruses and activates infection in trans. *Proc. Natl. Acad. Sci. U. S. A.* 98, 2670–5.
- (22) Powlesland, A. S., Fisch, T., Taylor, M. E., Smith, D. F., Tissot, B., Dell, A., Pöhlmann, S., and Drickamer, K. (2008) A novel mechanism for LSECtin binding to Ebola virus surface glycoprotein through truncated glycans. *J. Biol. Chem.* 283, 593–602.
- (23) Gramberg, T., Hofmann, H., Möller, P., Lalor, P. F., Marzi, A., Geier, M., Krumbiegel, M., Winkler, T., Kirchhoff, F., Adams, D. H., Becker, S., Münch, J., and Pöhlmann, S. (2005) LSECtin interacts with filovirus glycoproteins and the spike protein of SARS coronavirus. *Virology* 340, 224–236.
- (24) Gramberg, T., Soilleux, E., Fisch, T., Lalor, P. F., Hofmann, H., Wheeldon, S., Cotterill, A., Wegele, A., Winkler, T., Adams, D. H., and Pöhlmann, S. (2008) Interactions of LSECtin and DC-SIGN/DC-SIGNR with viral ligands: Differential pH dependence, internalization and virion binding. *Virology* 373, 189–201.
- (25) Tang, L., Yang, J., Tang, X., Ying, W., Qian, X., and He, F. (2010) The DC-SIGN family member LSECtin is a novel ligand of CD44 on activated T cells. *Eur. J. Immunol.* 40, 1185–91.
- (26) Figdor, C. G., de Vries, I. J. M., Lesterhuis, W. J., and Melief, C. J. (2004) Dendritic cell immunotherapy: mapping the way. *Nat. Med.* 10, 475–480.
- (27) Shortman, K., Lahoud, M. H., and Caminschi, I. (2009) Improving vaccines by targeting antigens to dendritic cells. *Exp. Mol. Med.* 41, 61–66.
- (28) Kreutz, M., Tacke, P. J., and Figdor, C. G. (2013) Targeting dendritic cells—why bother? *Blood* 121, 2836–2844.
- (29) Caminschi, I., Lahoud, M. H., and Shortman, K. (2009) Enhancing immune responses by targeting antigen to DC. *Eur. J. Immunol.* 39, 931–938.
- (30) Brzezicka, K., Vogel, U., Serna, S., Johannsen, T., Lepenies, B., and Reichardt, N.-C. (2016) Influence of Core β -1,2-Xylosylation on Glycoprotein Recognition by Murine C-type Lectin Receptors and Its Impact on Dendritic Cell Targeting. *ACS Chem. Biol.* 11, 2347–56.
- (31) Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., Bryan, M. C., Fazio, F., Calarese, D., Stevens, J., Skehel, J. J., van Die, I., Burton, D. R., Wilson, I. A., Cummings, R., Bovin, N., Wong, C. H., Paulson, J. C., Razi, N., and Stevens, D. J. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* 101, 17033.
- (32) Guo, Y., Sakonsinsiri, C., Nehlmeier, I., Fascione, M. A., Zhang, H., Wang, W., Pöhlmann, S., Turnbull, W. B., and Zhou, D. (2016) Compact, polyvalent mannose quantum dots as sensitive, ratiometric FRET probes for multivalent protein-ligand interactions. *Angew. Chem., Int. Ed.* 55, 4738–4742.
- (33) Brzezicka, K., Echeverria, B., Serna, S., van Diepen, A., Hokke, C. H., and Reichardt, N.-C. (2015) Synthesis and microarray-assisted binding studies of core xylose and fucose containing N-glycans. *ACS Chem. Biol.* 10, 1290–302.
- (34) Yang, Y. M., Li, X. H., Brzezicka, K., Reichardt, N.-C., Wilson, R. A., van Diepen, A., and Hokke, C. H. (2017) Specific anti-glycan antibodies are sustained during and after parasite clearance in *Schistosoma japonicum*-infected rhesus macaques. *PLoS Neglected Trop. Dis.* 11, e0005339.
- (35) Benevides, R. G., Ganne, G., da Conceição Simões, R., Schubert, V., Niemietz, M., Unverzagt, C., Chazalet, V., Breton, C., Varrot, A., Cavada, B. S., and Imberty, A. (2012) A lectin from *Platygodium elegans* with unusual specificity and affinity for asymmetric complex N-glycans. *J. Biol. Chem.* 287, 26352–26364.
- (36) Wang, Z., Chinoy, Z. S., Ambre, S. G., Peng, W., McBride, R., de Vries, R. P., Glushka, J., Paulson, J. C., and Boons, G.-J. (2013) A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans. *Science* 341, 379–83.
- (37) Shivatare, S. S., Chang, S.-H., Tsai, T.-I., Tseng, S. Y., Shivatare, V. S., Lin, Y.-S., Cheng, Y.-Y., Ren, C.-T., Lee, C.-C. D., Pawar, S., Tsai, C.-S., Shih, H.-W., Zeng, Y.-F., Liang, C.-H., Kwong, P. D., Burton, D. R., Wu, C.-Y., and Wong, C.-H. (2016) Modular synthesis of N-glycans and arrays for the hetero-ligand binding analysis of HIV antibodies. *Nat. Chem.* 8, 338–346.
- (38) Echeverria, B., Etxebarria, J., Ruiz, N., Hernandez, A., Calvo, J., Habeger, M., Reusch, D., and Reichardt, N.-C. (2015) Chemo-Enzymatic Synthesis of (13)C Labeled Complex N-Glycans As Internal Standards for the Absolute Glycan Quantification by Mass Spectrometry. *Anal. Chem.* 87, 11460–7.
- (39) Li, L., Liu, Y., Ma, C., Qu, J., Calderon, A. D., Wu, B., Wei, N., Wang, X., Guo, Y., Xiao, Z., Song, J., Sugiarto, G., Li, Y., Yu, H., Chen, X., and Wang, P. G. (2015) Efficient chemoenzymatic synthesis of an N-glycan isomer library. *Chem. Sci.* 6, 5652–5661.
- (40) Prudden, A. R., Chinoy, Z. S., Wolfert, M. A., and Boons, G.-J. (2014) A multifunctional anomeric linker for the chemoenzymatic synthesis of complex oligosaccharides. *Chem. Commun.* 50, 7132–7135.
- (41) Kawar, Z. S., Haslam, S. M., Morris, H. R., Dell, A., and Cummings, R. D. (2005) Novel poly-GalNAc β 1–4GlcNAc (LacdiNAc) and fucosylated poly-LacdiNAc N-glycans from mammalian cells expressing β 1,4-N-acetylgalactosaminyltransferase and α 1,3-fucosyltransferase. *J. Biol. Chem.* 280, 12810–9.
- (42) Serna, S., Etxebarria, J., Ruiz, N., Martin-Lomas, M., and Reichardt, N.-C. (2010) Construction of N-glycan microarrays by using modular synthesis and on-chip nanoscale enzymatic glycosylation. *Chem. - Eur. J.* 16, 13163–75.
- (43) Ramakrishnan, B., and Qasba, P. K. (2002) Structure-based design of β 1,4-galactosyltransferase I (β 4Gal-T1) with equally efficient N-acetylgalactosaminyltransferase activity: point mutation broadens β 4Gal-T1 donor specificity. *J. Biol. Chem.* 277, 20833–9.
- (44) Narasimhan, S., Freed, J. C., and Schachter, H. (1985) Control of glycoprotein synthesis. Bovine milk UDPgalactose: N-acetylglucosamine 4- β -galactosyltransferase catalyzes the preferential transfer of galactose to the GlcNAc β 2, 2Man α 1, 3-branch of both bisected and nonbisected complex biantennary asparagine-linked oligosaccharides. *Biochemistry* 24, 1694–1700.
- (45) Yan, S., Serna, S., Reichardt, N.-C., Paschinger, K., and Wilson, I. B. (2013) Array-assisted characterization of a fucosyltransferase required for the biosynthesis of complex core modifications of nematode N-glycans. *J. Biol. Chem.* 288, 21015–21028.
- (46) Nguyen, K., van Die, I., Grundahl, K. M., Kawar, Z. S., and Cummings, R. D. (2007) Molecular cloning and characterization of the *Caenorhabditis elegans* α 1, 3-fucosyltransferase family. *Glycobiology* 17, 586–599.
- (47) Prasanphanich, N. S., Mickum, M. L., Heimburg-Molinario, J., and Cummings, R. D. (2013) Glycoconjugates in host-helminth interactions. *Front. Immunol.* 4, 240.
- (48) Hokke, C., and Yazdanbakhsh, M. (2005) Schistosome glycans and innate immunity. *Parasite Immunol.* 27, 257–264.
- (49) Van Remoortere, A., Hokke, C. H., van Dam, G. J., van Die, I., Deelder, A. M., and van den Eijnden, D. H. (2000) Various stages of *Schistosoma* express Lewisx, LacdiNAc, GalNAc β 1–4 (Fuc α 1–3) GlcNAc and GalNAc β 1–4 (Fuc α 1–2Fuc α 1–3) GlcNAc carbohydrate epitopes: detection with monoclonal antibodies that are characterized by enzymatically synthesized neoglycoproteins. *Glycobiology* 10, 601–609.
- (50) Van Die, I., and Cummings, R. D. (2010) Glycan gimmickry by parasitic helminths: a strategy for modulating the host immune response? *Glycobiology* 20, 2–12.
- (51) van Remoortere, A., Bank, C. M., Nyame, A. K., Cummings, R. D., Deelder, A. M., and van Die, I. (2003) *Schistosoma mansoni*-infected mice produce antibodies that cross-react with plant, insect, and

mammalian glycoproteins and recognize the truncated biantennary N-glycan Man₃GlcNAc₂-R. *Glycobiology* 13, 217–225.

(52) Carvalho de Souza, A., van Remoortere, A., Hokke, C. H., Deelder, A. M., Vliegthart, J. F., and Kamerling, J. P. (2005) Determination of the specificity of monoclonal antibodies against *Schistosoma mansoni* CAA glycoprotein antigen using neoglycoconjugate variants. *Biol. Chem.* 386, 901–908.

(53) Peterson, N. A., Hokke, C. H., Deelder, A. M., and Yoshino, T. P. (2009) Glycotope analysis in miracidia and primary sporocysts of *Schistosoma mansoni*: differential expression during the miracidium-to-sporocyst transformation. *Int. J. Parasitol.* 39, 1331–1344.

(54) De Geus, D. C., van Roon, A.-M. M., Thomassen, E. A., Hokke, C. H., Deelder, A. M., and Abrahams, J. P. (2009) Characterization of a diagnostic Fab fragment binding trimeric Lewis X. *Proteins: Struct., Funct., Genet.* 76, 439–447.

(55) Smit, C. H., Homann, A., van Hensbergen, V. P., Schramm, G., Haas, H., van Diepen, A., and Hokke, C. H. (2015) Surface expression patterns of defined glycan antigens change during *Schistosoma mansoni* cercarial transformation and development of schistosomula. *Glycobiology* 25, 1465–79.

(56) Altmann, F. (2007) The role of protein glycosylation in allergy. *Int. Arch. Allergy Immunol.* 142, 99–115.

(57) Canales, A., Mallagaray, A., Pérez-Castells, J., Boos, I., Unverzagt, C., André, S., Gabius, H.-J., Cañada, F. J., and Jiménez-Barbero, J. (2013) Breaking Pseudo-Symmetry in Multiantennary Complex N-Glycans Using Lanthanide-Binding Tags and NMR Pseudo-Contact Shifts. *Angew. Chem.* 125, 14034–14038.

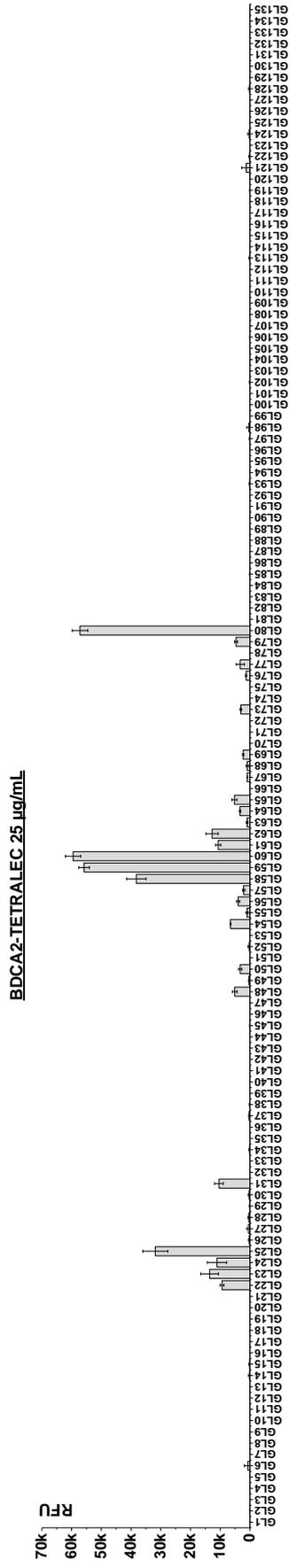
(58) Tabarani, G., Thépaut, M., Stroebel, D., Ebel, C., Vivès, C., Vachette, P., Durand, D., and Fieschi, F. (2009) DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain. *J. Biol. Chem.* 284, 21229–40.

9.1.2 Other glycan array screening

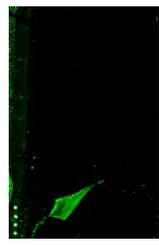
BDCA2 random TETRALEC was also incubated on the glycan array (Fig.105a). BDCA2 ligand recognition has been studied extensively by [54] and Gal β 1-3/4GlcNAc β 1-2Man was found to be the preferred binding epitope. The analysis performed with the random TETRALEC revealed a selective recognition of this specific epitope (**GL22-25**, **GL31**, **GL58-62**, **GL80**). In addition, when mannose is functionalised by fucose, the interaction is completely abolished (**GL31-38**, **GL81/92**). Nevertheless, positional isomers **GL126** and **GL127**, carrying the recognised motif, did not show any binding. A possible explicative hypothesis could be done taking into account the presence of an additional GlucNAc in the other arm, making the TETRALEC not suitable for proper interaction. Finally, no binding occurred on mannose terminal glycans.

MCL TETRALEC was incubated on the glycan array at a concentration of 50 μ g/mL. As expected, no specific recognition was observed for the glycans and all the values were considered as background binding (Fig.105b). An incubation at 100 μ g/mL was performed as well (data not shown) but did not reveal any binding enhancement.

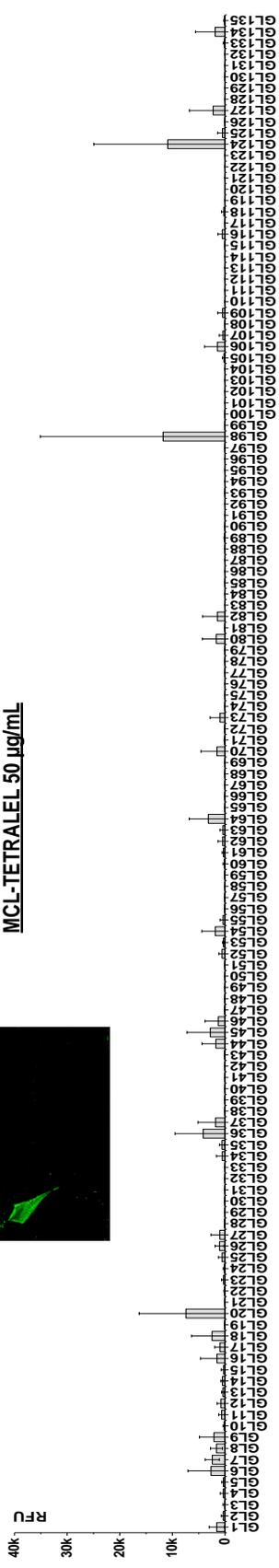
Indeed, no glycolipids were included in the glycan microarray used. In the future, collaboration will be established to test MCL.



a)



MCL-TETRALEC 50 µg/mL



b)

Fig.105 Glycan microarray incubation with a) BDCA2 random TETRALEC and b) MCL-TETRALEC

9.2 Glycomimetic array

The glycomimetic array contained 40 glycomimetics derived from Man $\alpha 1 \rightarrow 2$ Man and 11 fucose-derived compounds (Fig.106). All the glycomimetics were designed for discovery purpose and not to specifically target one lectin.

Four compounds were used as control : the linker alone (19.1), mannose glycosyl compound (19.2) mannose psDi (20) and Man030 (21) (chapter 3.2.1 and 8.3). All the glycomimetics were printed on the slides by Laura Medve, a PhD student from Anna Bernardi's group.

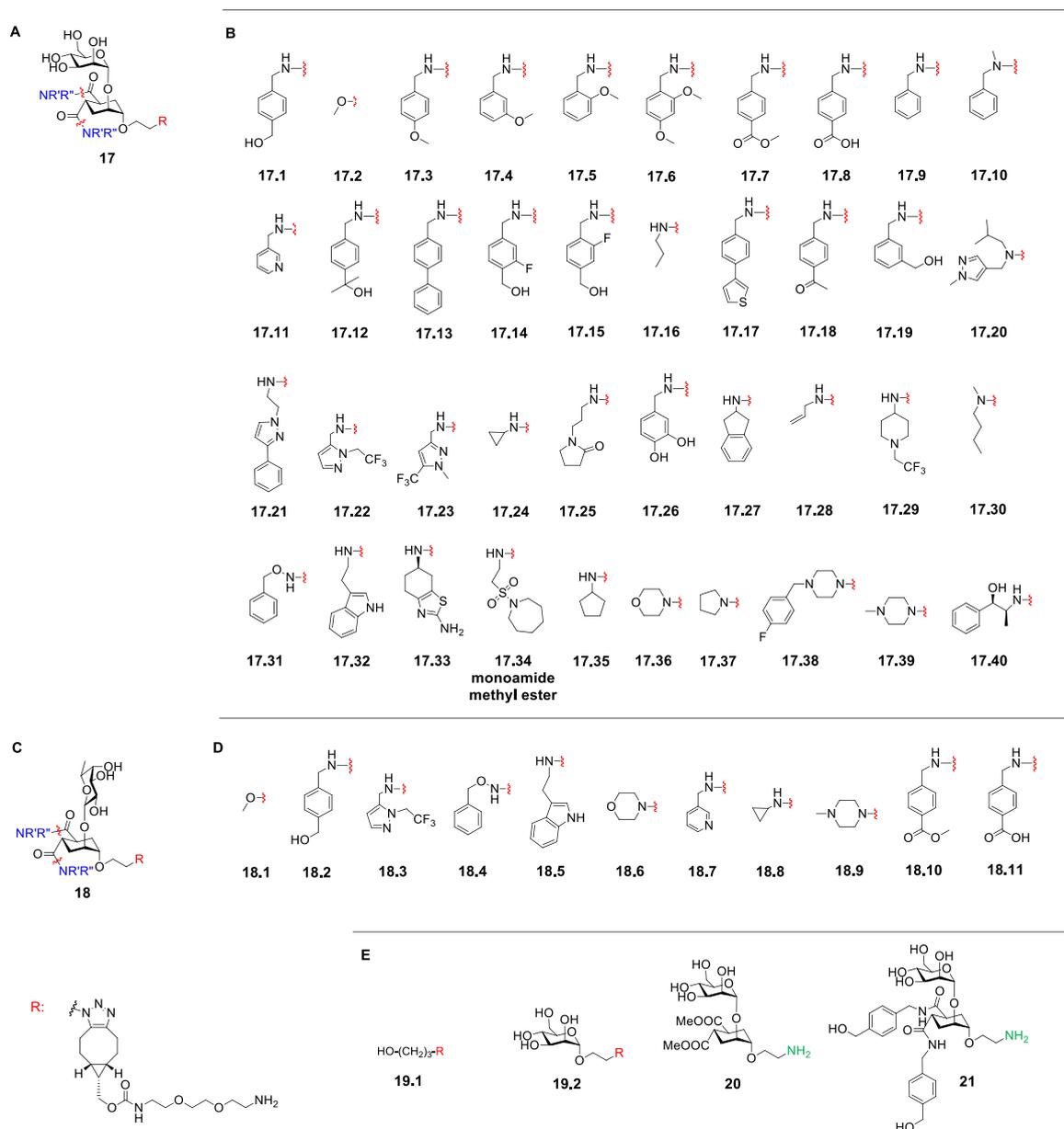


Fig.106 Glycomimetic array A) General structure of the mannose-based glycomimetics and B) substituents of the mannose-based glycomimetics. C) General structure of the fucose-based glycomimetics and D) substituents of the fucose-based glycomimetics. E) Control compounds.

Since all the glycomimetics were derived from mannose and fucose, no interaction was expected with

BDCA2 nor LSECtin, while langerin, DC-SIGNR and dectin-2 were potential candidates in addition to DC-SIGN. The following paper is based on the screening of this glycomimetic array against these four CLRs and it also includes the biophysical characterization of a few specific interactions.

9.2.1 On-chip screening of a glycomimetic library with C-type lectins reveals structural features responsible for preferential binding of dectin-2 over DC-SIGN/R and langerin

The main outcome of these studies:

This new glycomimetic array provides a fast tool for screening of the interaction between non glyco compounds and four human myeloid CLRs, namely DC-SIGN, DC-SIGNR, langerin and dectin-2 ECDs. The qualitative interactions for DC-SIGN and DC-SIGNR with four glycomimetics that were identified were confirmed by quantitative biophysical characterization by SPR. The most interesting outcome of this study was the identification of glycomimetics potentially selective towards dectin-2 over DC-SIGN.

Contributions:

All the mannose and fucose-based glycomimetics were designed by the group of Anna Bernardi and printed and validated by the group of Niels Reichardt and Anna Bernardi.

My contribution to this study:

I have prepared and labelled DC-SIGN, DC-SIGNR, Dectin-2 and langerin required for glycomimetic microarray experiments and performed all the assays involving CLRs. Moreover, I have prepared DC-SIGN and DC-SIGNR for the SPR assay, performed the SPR experiments, analysed data, participated to the preparation of the manuscript and prepared some of the figures.



Carbohydrates

On-Chip Screening of a Glycomimetic Library with C-Type Lectins Reveals Structural Features Responsible for Preferential Binding of Dectin-2 over DC-SIGN/R and Langerin

Laura Medve,^[a] Silvia Achilli,^[b] Sonia Serna,^[c] Fabio Zuccotto,^[d] Norbert Varga,^[a] Michel Thépaut,^[b] Monica Civera,^[a] Corinne Vivès,^[b] Franck Fieschi,^[b] Niels Reichardt,^[c, e] and Anna Bernardi^{✉[a]}

Abstract: A library of mannose- and fucose-based glycomimetics was synthesized and screened in a microarray format against a set of C-type lectin receptors (CLRs) that included DC-SIGN, DC-SIGNR, langerin, and dectin-2. Glycomimetic li-

gands able to interact with dectin-2 were identified for the first time. Comparative analysis of binding profiles allowed their selectivity against other CLRs to be probed.

Introduction

Lectins are sugar-binding proteins that engage in interactions with endogenous and exogenous glycans. The interactions between lectins and carbohydrates are involved in many fundamental biological events, from cell adhesion to antigen recognition and internalization, inflammation, or quality control in protein folding. The most abundant class of animal lectins are the C-type lectin receptors (CLRs) serving a broad range of functions. They are involved in pathogen recognition and in prevention of autoimmunity by contributing to the immune system's ability to identify carbohydrate-based pathogen-associated molecular patterns (PAMP) and damaged-self-associated molecular patterns (DAMP). They take part in signal transduction, cell trafficking, and in the induction of T-cell differentiation. Their name, C-type lectins, indicates the presence of a Ca^{2+} ion in their carbohydrate recognition domain (CRD). This

ion is the primary site of carbohydrate interaction and typically coordinates two vicinal hydroxyl groups on a sugar ring. Many C-type lectin receptors are yet to be explored and described in detail with regard to their CRD structure, carbohydrate binding specificities, and the molecular factors governing the interaction with glycans. However, it is known that the CRDs of CLRs feature evolutionarily conserved groups of residues that coordinate the Ca^{2+} ion and determine the monosaccharide binding specificity of the CLR. So, a Glu-Pro-Asn (EPN) motif results in a preference for mannose (Man), *N*-acetylglucosamine (GlcNAc), fucose (Fuc), or glucose (Glc) residues, whereas a Gln-Pro-Asp (QPD) sequence leads to recognition of galactose (Gal) and *N*-acetylgalactosamine (GalNAc).^[1] Available data also suggests that the extended binding sites of CLRs often display a higher affinity towards larger glycan structures, thanks to additional interactions occurring in the vicinity of the primary Ca^{2+} site.^[2] Thus, complex glycans exposing the same monosaccharide can bind to CLRs with very different affinities, depending on the accessibility of the recognition element and on additional features of the lectin binding sites. Structural studies also demonstrate that secondary binding sites can alter the affinity and specificity of the CLR towards ligands. As an example, langerin,^[3] a transmembrane CLR expressed on Langerhans-cells, and the dendritic-cell specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN), a CLR expressed by dendritic cells (DCs)^[4] share similar primary binding sites with similar specificity for oligomannosides, but langerin has an additional calcium-independent sugar-binding site and binds to large sulfated glycosaminoglycans, whereas DC-SIGN does not.^[5]

Given the key role played by lectins in biological systems, many research groups have turned their attention towards the development of glycomimetic molecules to be used as selective probes for the study of sugar-protein interactions and/or for medicinal chemistry purposes.^[6] Glycomimetics present sev-

[a] L. Medve, N. Varga, Dr. M. Civera, Prof. A. Bernardi
Dipartimento di Chimica, Università degli Studi di Milano
Via Golgi 19, 20133 Milano (Italy)
E-mail: anna.bernardi@unimi.it

[b] S. Achilli, Dr. M. Thépaut, Dr. C. Vivès, Prof. F. Fieschi
Univ. Grenoble Alpes, CEA, CNRS
Institut de Biologie Structurale
38000 Grenoble (France)

[c] Dr. S. Serna, Dr. N. Reichardt
Glycotechnology laboratory, CIC biomaGUNE
Paseo Miramón 182, 20014
Donostia-San Sebastián (Spain)

[d] Dr. F. Zuccotto
University of Dundee, Dundee (UK)

[e] Dr. N. Reichardt
CIBER-BBN
20014 Donostia-San Sebastián (Spain)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/chem.201802577>

eral advantages as drug candidates over natural glycans, since they can be made metabolically more stable, more bioavailable, and possibly more active and selective than natural oligosaccharides. Our previous studies have focused on inhibiting the dendritic cell receptor DC-SIGN, a CLR implicated in viral and bacterial infections.^[7] The primary binding site of DC-SIGN CRD recognizes mannose oligosaccharides, L-fucose residues in Lewis-type blood antigens,^[4] and biantennary *N*-glycans with terminal GlcNAc moieties.^[8] Additional druggable sites on the lectin have been described recently.^[9] DC-SIGN-mediated adhesion to dendritic cells is the first step of several viral infections, notably by HIV and Ebola viruses.^[10] Glycomimetic antagonists of DC-SIGN have mainly been designed starting from high mannose glycans, like Man₉ (Man)₉(GlcNAc)₂ (**1**, Figure 1), or from Le^x (Fucα1,3-(Galβ1,4)-GlcNAc) (**2**)-type structures. In particular, pseudo-di and tri-mannoside fragments (**3–6**) were synthesized^[11] as mimics of the D2 and D3 arms of Man₉ (Figure 1). When used in multivalent constructs, they were found to block DC-SIGN-mediated infection with activities up to the nanomolar level both in HIV and Ebola infection models.^[12] Notably, the bisbenzylamide derivatives **6a**^[11c] and **6b**^[11d] also exhibited strong selectivity towards DC-SIGN and did not bind to langerin, a CLR that shares with DC-SIGN a similar set of ligands but, rather than spreading the infection, facilitates HIV eradication.^[13] These results suggested that, with appropriate modifications, the structure of **6** could represent a general template to generate a diverse library of glycomimetics containing one natural monosaccharide as the lectin-target-

ing element and a tuning unit, which could provide additional functional elements for interaction with the lectin in the proximity of the primary binding site. As mentioned above, the structure of **6** derives from mimicry of the Manα1-2Man disaccharide, the terminal unit of the D1–D3 arms of Man₉, and a common disaccharide ligand for DC-SIGN (PDB: 2IT6) and for other CLRs of the immune system with similar specificity, such as DC-SIGNR,^[14] langerin (PDB: 3P5F),^[15] and dectin-2 (PDB: 5VYB).^[16] Screening such a library against these CLRs in a microarray format may become a potent tool for glycomimetic drug discovery.^[17] Glycan microarrays, as introduced and developed over the past 15 years, have been an essential tool in the characterization of lectin specificity,^[18] and have been used to pave the way for the therapeutic exploitation of vital lectin-sugar interactions.

Herein, we describe the synthesis of a mannose- and fucose-based glycomimetic library and its on-chip screening against a set of human CLRs that led to the discovery of hit ligands able to interact with dectin-2.^[16]

Results and Discussion

Design and synthesis of the library

The set of bis(benzylamide) compounds previously designed as DC-SIGN antagonists was directly expanded into a mannose-based library by adopting the described route^[11c] that starts from diacid **7**, with small modifications (Scheme 1).

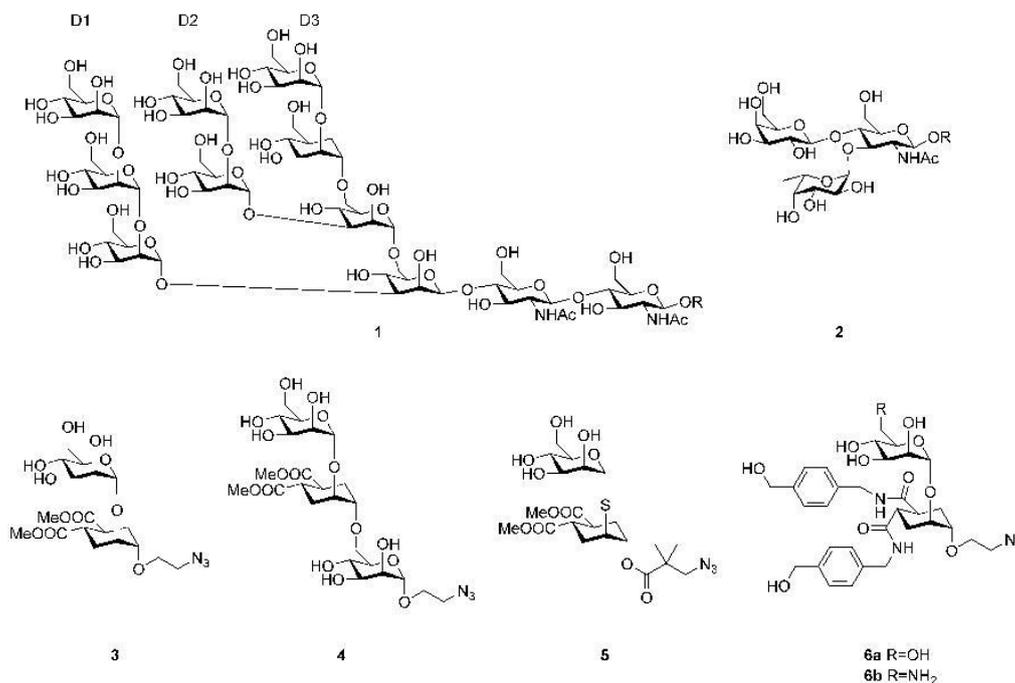
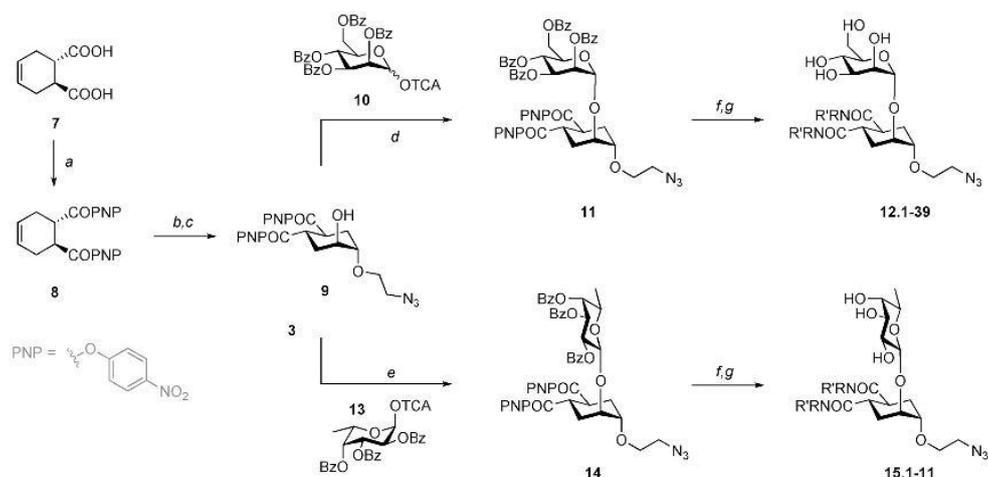


Figure 1. Natural DC-SIGN ligands: Man₉ (**1**)—the D1–D3 arms are the template for the design of mimics **3–6**; L-Fuc containing Le^x blood group antigen (**2**). Linear fragment mimics of Man₉ arms: pseudo-mannobioside (**3**) and pseudo-mannotriose (**4**); the pseudo-thiomannobioside (**5**) and two derivatives of pseudo-mannobioside (**6**).



Scheme 1. Synthesis and general structure of mannose-(12.1–39) and fucose-derived (15.1–11) members of the glycomimetic library. a) *p*-Nitrophenyl trifluoroacetate, pyridine, DMF, 50 °C, 18 h, 73%; b) *m*CPBA, CH₂Cl₂, 16 h, 94%; c) 2-azidoethanol, Cu(OTf)₂, CH₂Cl₂, 70%; d) TMSOTf, –30 °C, 20 min, 74%; e) TMSOTf, –30 °C, 3.5 h, 74%; f) R'R'NH, THF/DMF, (Et₃N), 1 h–2 d; g) NaOMe, MeOH, 1.5 h. *m*CPBA = *meta*-chloroperbenzoic acid.

Scale-up to a multigram scale of the protected common scaffold **11**, equipped with a versatile azido-functionality, allowed further derivatization to a collection of bisamides **12**. The glycomimetic library was not conceived to specifically target one lectin, but rather to broadly interact with CLRs that contain the EPN motif in their CRD. Therefore, the required set of amines was selected for diversity, with the help of chemoinformatic tools and based on commercial availability. Lead-like physicochemical filters were applied to a large number of amines from available commercial collections (see the Experimental Section) and the selection was sifted to exclude any structural incompatibility. The remaining structures were then clustered by chemoinformatic descriptors and representatives of the various subgroups were selected based on their availability. Overall, a collection of 38 diverse amines were selected (Figure 2), which allowed us to prepare the Man-based derivatives **12.1–39** (Scheme 1, Figure 2).

As discussed previously, Man-binding CLRs are expected to recognize L-Fuc residues as well, due to the overlapping carbohydrate-specificity imparted by the EPN motif. Therefore, in analogy to the mannobioside mimics, β -fucosylated ligands **15** were synthesized by linking the cyclohexane acceptor (**9**) with an appropriately protected fucose-trichloroacetimidate donor (**13**) to yield the bis-*p*-nitrophenylester **14**, which was reacted with a set of primary and secondary amines. This approach afforded the 11 Fuc-based ligands **15.1–11** (Scheme 1, Figure 2).

Ligand immobilization on the microarray surface

Although the azidoethyl functionalized glycomimetics could, in principle, be immobilized directly by surface-based cycloaddition,^[19] we chose to extend the short linker with an additional hetero-bi-functional spacer prior to printing the library, to improve ligand accessibility. This set-up presents the additional advantage of allowing the glycomimetics to be printed alongside existing glycan libraries, which are typically functionalized

with aminoterminated linkers.^[20] To this end, commercially available *N*-((1*R*,8*S*,9*S*)-bicyclo[6.1.0]non-4-yn-9-ylmethoxyxycarbonyl)-1,8-diamino-3,6-dioxaoctane (**16**, Scheme 2) was submitted to selective, rapid, and bioorthogonal strain-promoted azide–alkyne cycloaddition (SPAAC), as described by Agard et al.^[21] yielding conjugates **17** and **18** (Figure 2). The azido-terminated ligands reacted quantitatively overnight with **16** and the reaction products were analyzed by MALDI-TOF-MS to assess the identity of the resulting compounds, which were later robotically printed through the terminal amino functionality onto *N*-hydroxysuccinimidyl ester (NHS)-functionalized glass slides. To detect any interference of the linker on lectin binding, compound **19.1** (Figure 2E) was included as a nonglycosylated “clicked-linker”. Conjugation of 2-azidoethyl- α -mannoside to **16** provided the control mannose glycoside **19.2**. In addition, amines **20** and **21** (Figure 2) were prepared by azide reduction of two ligands, **12.1** and **12.2**, respectively. The amines were directly printed on the slides to compare their binding with the same ligands immobilized through the bifunctional linker (**17.1** and **17.2**, respectively).

The final ligand network, conjugated with the additional linker and printed onto the microarray surface, consisted of 39 α -mannosylated (**17.1–17.39**) and 11 β -fucosylated (**18.1–18.11**) disaccharide mimics (Figure 2) and included four control compounds (**19.1**, **19.2**, **20**, and **21**).

After microarray design and optimization of printing parameters, the slides were probed with a variety of recombinant and fluorescently tagged human CLRs, and the binding profile recorded with a fluorescence scanner. The fluorescence intensity of individual spots relates to the amount of bound lectin and is an estimation of the relative strength of interaction. For a selected set of ligands, the intensity values from the microarray analysis were compared to IC₅₀ values determined in a SPR competition experiment against DC-SIGN and DC-SIGNR. This comparison allowed us to critically evaluate the ranking suggested by the array interaction studies and provided a useful

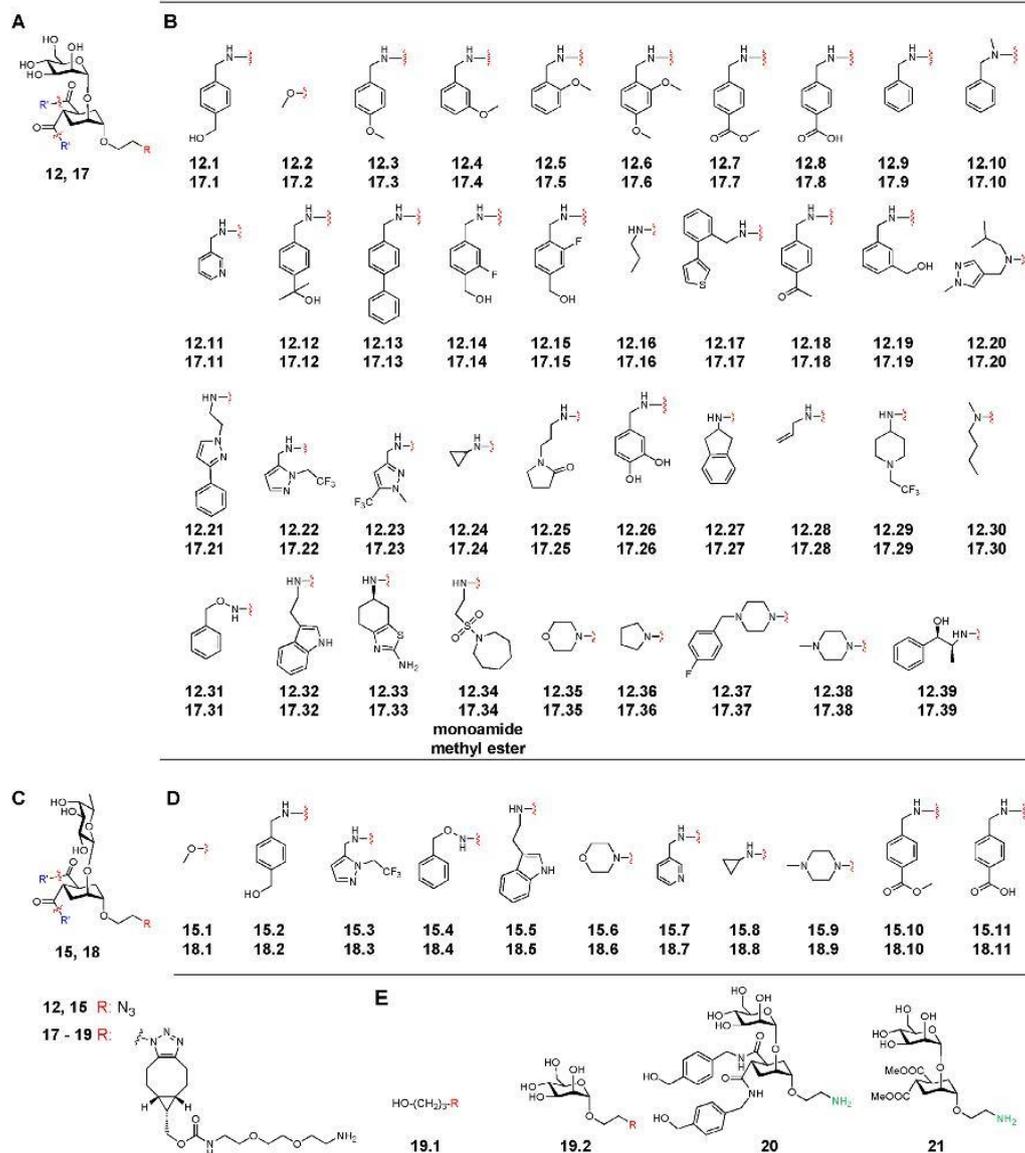
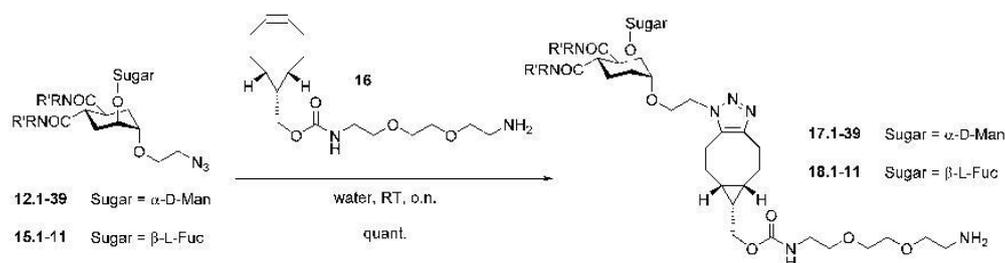


Figure 2. Glycomimetic library of 39 mannosylated, 11 fucosylated, and 4 control compounds. A) General structure of mannose glycomimetics. B) Different R' substituents present in mannose glycomimetics. C) General structure of fucose glycomimetics. D) Different R' substituents presented in fucose glycomimetics. E) Control compounds included in the library.



Scheme 2. Conjugation of azide-containing glycomimetics with cyclooctyne **16** by strain-promoted alkyne-azide cycloaddition (SPAAC).

perspective in the screenings performed with other human CLRs.

Experimental design and data analysis

Glycan microarrays were printed as previously described^[8] on commercially available NHS-ester activated glass slides. The immobilization of the ligands was confirmed by incubation with fluorescently labeled lectins of known glycan specificity: plant lectin *Concanavalin A* (ConA, D-Man)^[22] and fungal *Aleuria aurantia* lectin (AAL, L-Fuc).^[23] These experiments allowed the optimization of the spotting conditions and the validation of the array. In particular, it was observed that the addition of 10% DMSO to the printing buffer afforded the most homogeneous fluorescent images.

On the array, ConA was found to recognize most of the printed Man-based glycomimetics **17.n** as well as the controls **19.2**, **20**, and **21**, whereas no binding to the non-mannosylated linker **19.1** and to the fucose-based glycomimetics **18.n** was observed (see the Supporting Information, Figure SI-1). Interestingly, many of the Man-based mimics appeared to interact with ConA more efficiently than mannose itself (**19.2**), supporting the hypothesis that secondary interactions can contribute to the overall affinity of lectin ligands. The intensity of the signals obtained for **17.1** and **17.2** was somewhat higher than that of the short-linker controls **20** and **21**, which suggests a better accessibility of the former compounds on the array.

As expected, screening the array with fucose-specific AAL resulted in a totally different interaction profile, involving exclusively the Fuc-based ligands **18.n** (see the Supporting Information, Figure SI-2). Substitutions to the fucose core apparently did not affect recognition, as the binding intensity seems largely the same for all compounds. No binding was observed to the linker control **19.1** for this or any other of the tested lectins.

Interaction with human CLRs

DC-SIGN and langerin extracellular domains (ECD) were produced as previously described.^[3a,24] DC-SIGNR ECD and dectin-2 ECD were expressed in *E.coli* and purified as detailed in the Supporting Information. All CLRs were labeled with Cy3 and the degree of labeling was estimated as described in the Supporting Information.

The analysis was performed under the optimized conditions described above. Results are reported in Figures 3–5 (see below), in which ligands are grouped by chemical features (type of monosaccharide, degree of amide substitution) rather than by numbering. A comparative heat map for the four proteins is reported in the Supporting Information, Figure SI-3.

DC-SIGN

DC-SIGN is a transmembrane protein expressed primarily on the surface of dendritic cells in dermal mucosa and on various other antigen presenting cells of the myeloid lineage. DC-SIGN has a dual role as a cell surface pattern-recognizing receptor

and as a mediator for T-cell activation. Additionally, a number of pathogens, most notably the HIV virus, are known to exploit DC-SIGN in the initial steps of host invasion.^[10a,25] For these reasons, DC-SIGN has been actively investigated, both as a target for discovery of anti-adhesive antiviral therapies and for its potential role in immunoregulation.^[26]

The lectin recognizes highly mannosylated oligosaccharides often found on viral and bacterial cell surfaces; the four Lewis-type blood group antigens (Le^x, Le^y, Le^a, and Le^b) and mannan-capped lipoarabinomannan and phosphatidylinositol-mannosides expressed on mycobacterial surfaces.^[27] On the array (Figure 3A), many of the mannosylated structures, but essentially none of the β -fucosylated ligands are recognized by the tetrameric DC-SIGN extracellular domain (ECD). Although DC-SIGN is known to bind fucosylated oligosaccharides, they all consist of α -fucosides, whereas, to the best of our knowledge, β -fucosides have not been explored before. The mannose-based ligands identified by red bars in Figure 4 had originally been designed as DC-SIGN antagonists and tested by SPR as inhibitors of DC-SIGN binding to mannosylated BSA.^[11c] Both the SPR data and the microarray results indicate that all these compounds have a similar affinity for the lectin, with the methyl ester **17.2** being the least effective of the series. None of the additional modifications of the amide functionality attempted in this library was found to improve on the previous design. On the other hand, very low fluorescence intensity was detected for all tertiary amide derivatives (**17.10**, **17.20**, **17.30**, **17.36**, **17.37**, **17.38**, and **17.39**) on the chip. This observation confirmed early data from previous SPR screenings in our groups, which had indicated that tertiary amides on the pseudo-dimannoside scaffold were significantly reducing the binding affinity towards DC-SIGN.^[28] The current set of data strongly suggests that this feature can be reliably used to generate selectivity against DC-SIGN. Overall, these data allowed the validation of microarray results and showed that the screening technique implemented is robust and adequate for fast analysis of binding activity.

DC-SIGNR

The DC-SIGN-related homologue, DC-SIGNR (or L-SIGN) is expressed primarily on sinusoidal endothelial cells in lymph nodes, the liver, the lungs, the gastrointestinal tract, and capillary endothelial cells in the placenta.^[29] The two homologues exhibit 73% identity at the nucleic acid level and 77% identity in the amino acid sequences,^[30] but display differences in the coiled-coil neck domains that affect the spatial arrangement of the four CRDs. As a result, the two CLRs can show different avidity towards the same multivalent ligand.^[24,31] Similarly to DC-SIGN, DC-SIGNR is a pathogen receptor for HIV-1, HCV, SARS-coronavirus and *Mycobacterium tuberculosis*, recognizes influenza A and interacts with lymphocytes.^[32]

The microarray screening results are shown in Figure 3B. Many of the mannose-based ligands interact with the protein more effectively than mannose itself (**19.2**) and the binding profile is rather similar to that observed for DC-SIGN. Four ligands (**17.11**, **17.15**, **17.19**, and **17.27**, all shown as blue bars

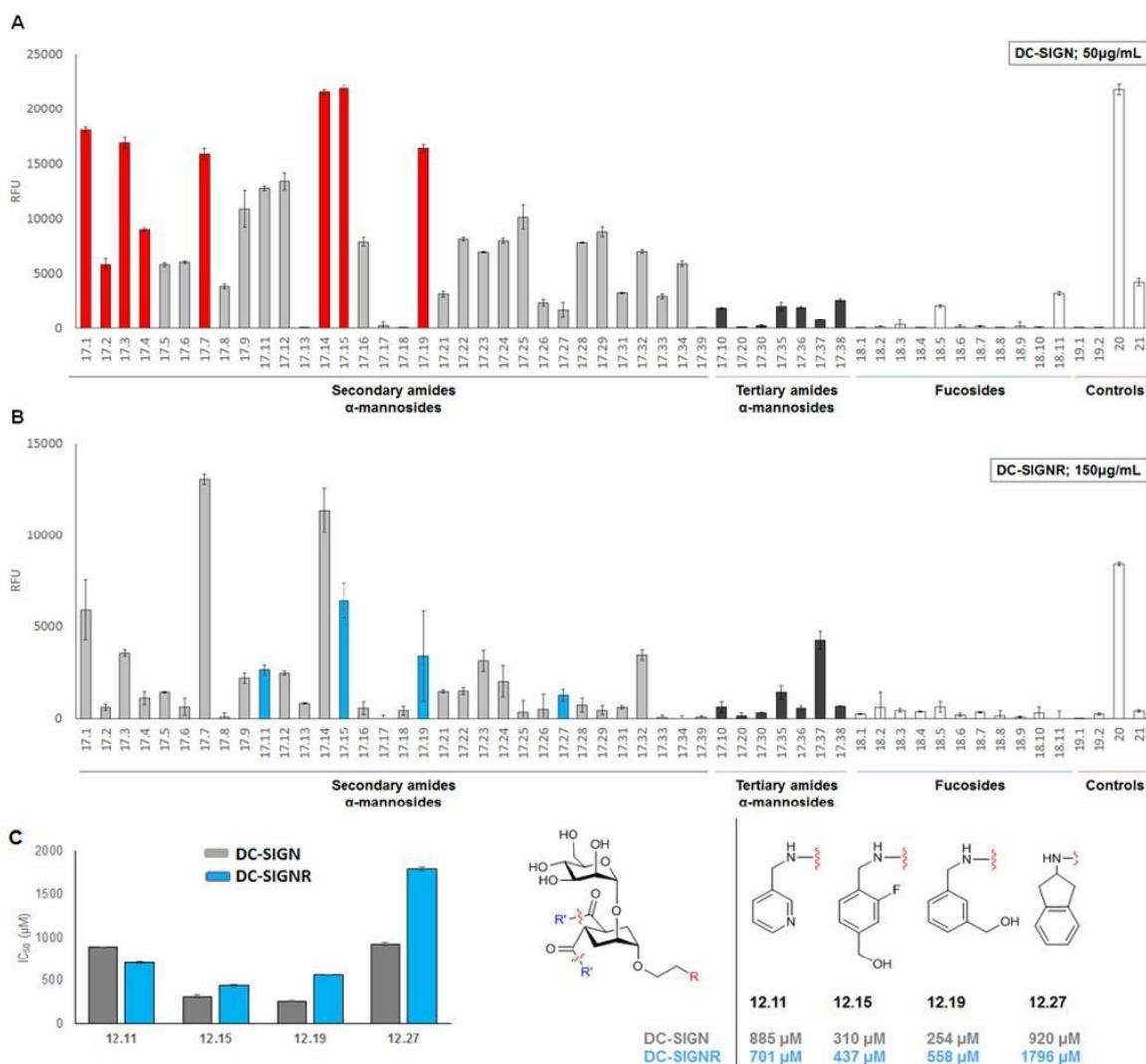


Figure 3. A) DC-SIGN glycomimetic ligand binding profile at $50 \mu\text{g mL}^{-1}$ DC-SIGN ECD (extracellular domain, tetramer). The red bars indicate previously known DC-SIGN ligands (ref. [11c]). The IC_{50} values measured for these ligands in ref. [11c] are collected in Figure S1-4, Supporting Information. B) DC-SIGNR glycomimetic ligand binding profile at $150 \mu\text{g mL}^{-1}$ DC-SIGNR ECD. Ligands selected for SPR studies are highlighted in blue. C) Left: ligand inhibitory potency towards DC-SIGN and DC-SIGNR. IC_{50} values measured in SPR inhibition assays (Man-BSA immobilized on the surface). Right: the structure of the ligands and the IC_{50} values for DC-SIGN (grey) and DC-SIGNR (blue).

in Figure 3B) were selected for further analysis and the affinity of the corresponding recognition elements **12.11**, **12.15**, **12.19**, and **12.27** for DC-SIGN and DC-SIGNR was evaluated using an SPR inhibition assay that measures their ability to inhibit protein binding to mannosylated bovine serum albumin (Man-BSA) (Figure 3C). The results show some obvious differences in the way the ligands are classified by the two assays. In particular, the IC_{50} value measured by SPR for **12.11** and DC-SIGN is higher than expected based on the array profile. This may simply reflect the intrinsic differences of the physical processes interrogated in the two assays: in the microarray setup, direct binding of the tetravalent lectins to a multivalent-functionalized surface is observed; in the SPR inhibition assay, the

ligands are scored based on the strength of their monovalent interaction with the protein. A strong dependence of the affinity values measured for carbohydrate-protein complexes and even an influence on the observed binding selectivity of lectins on the physical format of the assay was noted early on and repeatedly confirmed for various systems.^[33] Once again, our data suggest that multiple analytical methods need to be applied to fully characterize the interaction of lectins with synthetic or natural ligands. This represents an additional element of complexity for the discovery of selective lectin antagonists. However it is still possible, using data from various techniques, to identify ligands of potential interest in drug discovery programs and to ascertain the structural features that concur to

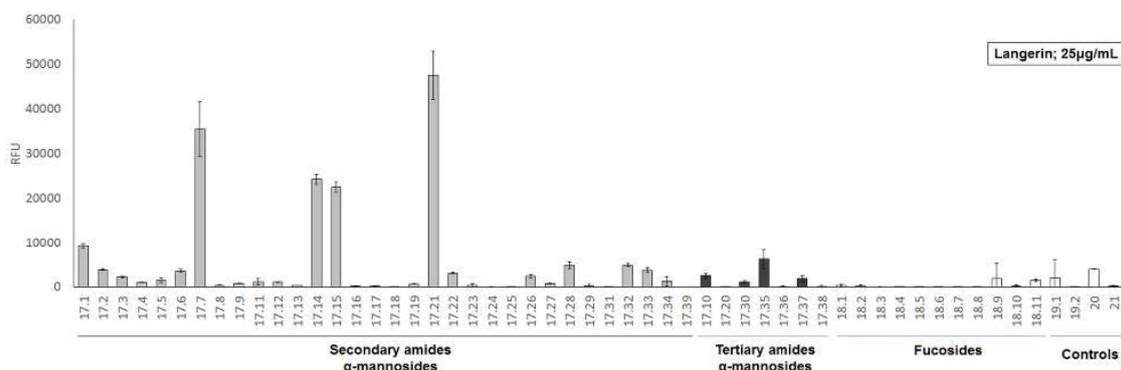


Figure 4. Langerin glycomimetic ligand binding profile at $25 \mu\text{g mL}^{-1}$ lectin concentration.

determine their activity and selectivity. We have obtained here, for the first time, microarray and SPR data on the recognition of glycomimetic ligands by DC-SIGNR, which will be the base for further elaboration of these structures.

Langerin

Langerin is a trimeric CLR abundantly expressed on Langerhans cells in the epidermis, and at lower levels on CD1c^+ myeloid dendritic cells and lamina propria of the human colon.^[34] Apart from the protective role against HIV mentioned above, the lectin binds to other pathogens, such as *Candida*, *Saccharomyces*, *Malassezia furfur*, and *Mycobacterium leprae*.^[35] Although langerin and DC-SIGN share many of their natural ligands, differences can be found in their specificity towards fucosylated glycans. DC-SIGN exhibits a good recognition of many fucose-based Lewis-type ligands (Le^x , Le^a , Le^b , and Le^y), as well as of the A, B, and H blood group antigens. Langerin binds with good affinity only the blood group antigens B and A, whereas Le^a , Le^b , Le^y , and Le^x are poorly recognized.^[15,36] Moreover, opposite to DC-SIGN, langerin selectively recognizes sulfated Gal, GalNAc, and glycosaminoglycans.^[5,11d,15] Additionally, a divergent structural organization and their distinct expression locations suggest fundamentally different biological roles for these two CLRs.

In the microarray assay, the signal for trimeric langerin ECD raised above noise level only for four ligands, three of which (17.7, 17.14, and 17.15) are known ligands of DC-SIGN (Figure 4). The corresponding recognition elements 12.7, 12.14, and 12.15 had been previously assessed also against langerin, using the SPR inhibition experiment described above, and found to be poor competitors of immobilized Man-BSA.^[11c] SPR inhibition studies were repeated for 12.15 and performed for the first time with 12.21. Neither of them could inhibit langerin binding to immobilized Man-BSA, up to millimolar concentrations (Supporting Information, Figure SI-5). Some inhibitory activity of 12.15 could be observed in SPR competition assays, but only when challenging a weaker interaction using a surface functionalized with Le^a -BSA, which is a poor langerin ligand (Supporting Information, Figure SI-5). This experiment allowed the evaluation of an IC_{50} value of 1.8 mM. For 12.21, a

sharp drop in langerin activity could be observed above 1 mM concentration of the ligand, but the data could not be fitted to a binding isotherm. The fact that ligands displaying little inhibitory activity in the SPR experiment can still light up on the microarray may depend on the avidity of the polyvalent presentation generated upon printing them on the chip. However, we cannot rule out at this stage that these molecules, through their amide substituents, may be interacting in a noncompetitive fashion, that is, with a different site than the carbohydrate binding site on the ECD.

Dectin-2

Dendritic cell-associated C-type lectin 2, dectin-2, is a predominantly macrophage and monocyte associated CLR,^[37] with a known specificity for mannose and a preference for $\text{Man}\alpha 1\text{-}2\text{Man}$ recognition.^[16] Dectin-2 binds to bacteria such as *Klebsiella pneumoniae* and *Mycobacterium tuberculosis* and fungi such as *Candida albicans*.^[38] Its antifungal activity has been demonstrated in animal-models.^[39] Upon ligand binding, dectin-2 is able to promote signaling, cytokine secretion, and, finally, the initiation of a Th17 immune response.^[40]

The binding profile of dectin-2 ECD towards the glycomimetic library is shown in Figure 5A. It can be observed that some of the mannosides appear to interact more strongly than mannose 19.2, a weak binder of dectin-2. Remarkably, dectin-2 exhibits an affinity towards some tertiary amide structures (17.10, 17.20, 17.30, 17.36, 17.37, 17.38, and 17.39, showcased in Figure 5B) and β -fucosylated ligands, which are not or barely recognized by DC-SIGN (Figure 3A). Although the two lectins display a similar profile for mannosides bearing secondary amide structures, they clearly differ in the fucoside section and more strikingly so in the mannoside-bearing tertiary amide groups section. This suggests the possibility of an unprecedented selectivity between the two CLRs towards glycomimetic compounds, which may be related to the different nature of the two binding sites.^[16]

Indeed, the X-ray structure of the dectin-2 in complex with $\text{Man}_6\text{GlcNAc}_2$ has been recently solved.^[16] Figure 6 shows dectin-2 CRD superimposed to the X-ray structure of the DC-SIGN complex with the pseudo-dimannoside 3.^[41] The overlay

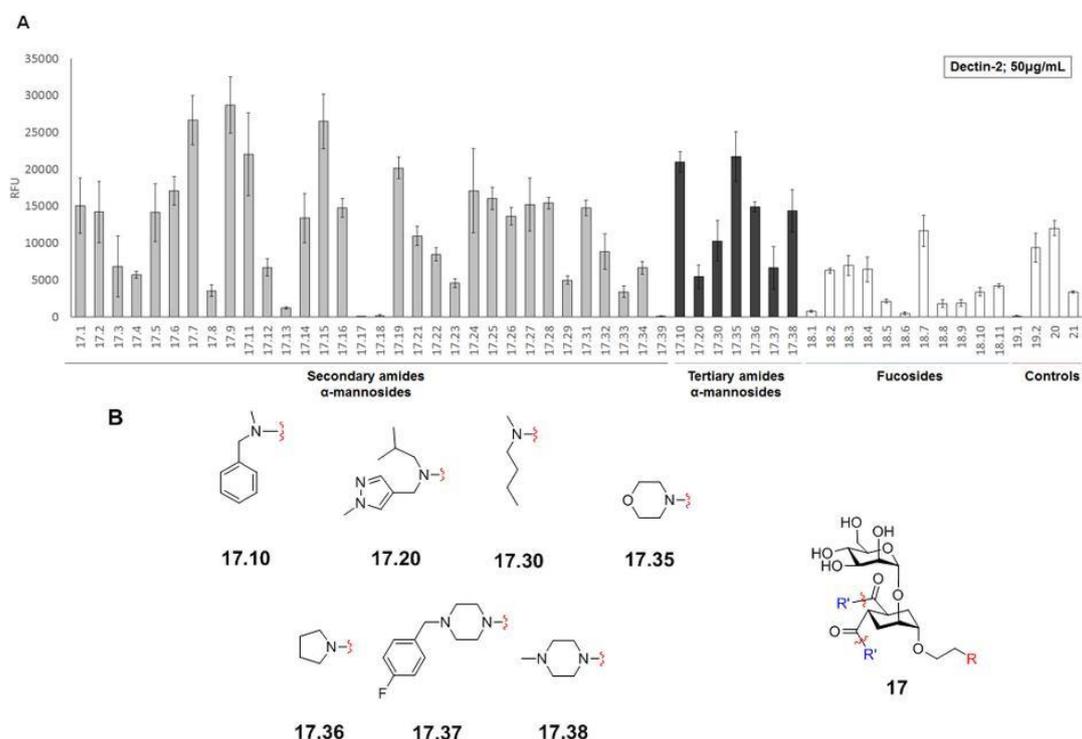


Figure 5. Dectin-2 binding profile. Mannosides bearing secondary amide groups are shown as grey bars, mannosides bearing tertiary amide groups are shown as black bars, fucosides, and controls are shown as white bars. B) Structure of mannose ligands of dectin-2 that do not interact with DC-SIGN.

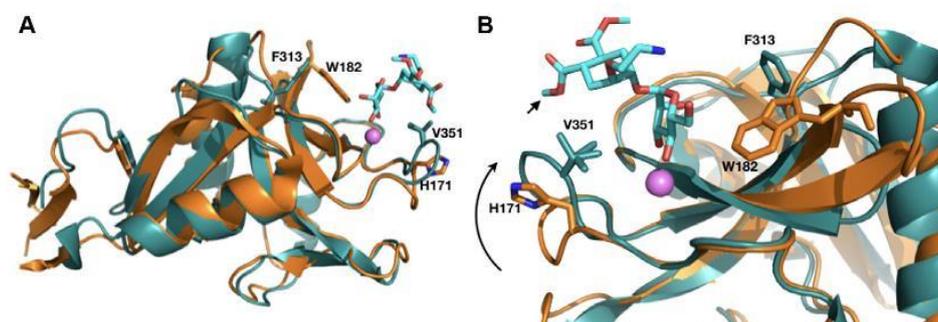


Figure 6. A) Alignment of DC-SIGN complex with **3** (blue protein, cyan ligand; PDB: 2XR5) and dectin-2 complex of $\text{Man}_3\text{GlcNAc}_2$ (orange protein, ligand not shown PDB: 5VYB). The Ca^{2+} ion is shown as a pink sphere. B) Zoom on the Ca^{2+} binding site viewed from the opposite direction. The curved arrow highlights the different orientation of the two loops in the two proteins. The small arrow points to the cyclohexene ring of **3** and to the position of the amide group in the glycomimetic ligands **12.1–n**.

shows a highly conserved tertiary structure with a difference in the loops in close proximity of the Ca^{2+} binding site that contain V351 for DC-SIGN and H171 for dectin-2 (to the right of the ligand in Figure 6A). At the other side of the Ca^{2+} ion, the X-ray structure of dectin-2 shows a very shallow surface, lined by a Trp side chain (W182, Figure 6A). Alignment of the dectin-2 and DC-SIGN sites shows that the DC-SIGN binding region (blue) is more confined, by Phe313 on one side and Val351 on the opposite side of the Ca^{2+} ion (Figure 6B). The

Val351 side chain is in close contact with the cyclohexene ring of **3** and near the amide side chains of the amide derivatives (Figure 6B). In dectin-2, the corresponding loop is more open and the valine residue is replaced by a histidine in this position (His171). The different loop orientation, as highlighted in Figure 6, allows more available space between the protein and the position of the amide side chain of the mimics (as indicated by the curved arrow in Figure 6B). As a result, this lectin may be able to accommodate mannobioside mimics with

larger groups, such as tertiary amides, for the interaction with its CRD. Further investigation on these ligands is underway.

Conclusions

We have set up and optimized a glycomimetic microarray to use as a primary screening tool for mannose/fucose selective C-type lectins. A doubly-functionalized cyclooctyne linker was used for the fast immobilization of glycomimetic structures carrying an azide-terminated side chain. The array was validated with plant or fungal lectins of known specificity and then interrogated with a set of four human C-type lectins: DC-SIGN, DC-SIGNR, langerin, and dectin-2. Appropriate controls showed that, for all the lectins examined so far, the linker does not interfere with the binding process. The glycomimetics used are based on a central cyclohexane scaffold, carrying either an α -mannose or a β -fucose residue and further diversified by the presence of different amide appendages. The mannose based glycomimetics are structurally derived from the Man α -1-2Man (mannobioside) natural disaccharide motif. Interestingly, this disaccharide is a common natural ligand of all the C-type lectins tested in this study, which on the contrary differ strongly in their ability to interact with fucose-containing oligosaccharides. Thus, screening of mannose and fucose based glycomimetics is potentially of high interest in the search for selective ligands. In fact, only dectin-2 appeared to interact with the β -fucosides on the array, although less effectively than with most mannose-based derivatives.

The screening also revealed that the CLRs studied differentially respond to the amide substituents of the mimics, generating different binding profiles. Whereas langerin was found to bind weakly to most of the structures examined, DC-SIGN and DC-SIGNR displayed a rather good tolerance to secondary amide substituents on the pseudo-mannobioside structure and some similarity in the recognition profile. Most interestingly, a set of mannosides carrying tertiary amide substituents were found to selectively recognize dectin-2 over DC-SIGN, which may be explained by the known structure of the two lectins' binding site. Some of the fucose-derived glycomimetics loaded on the chip also displayed a selectivity for dectin-2 over DC-SIGN and DC-SIGNR. Thus these screening campaigns simultaneously provided the first discovery of glycomimetic ligands for dectin-2 and gave important indications for the design and optimization of dectin-2 selective antagonists.

The affinity of selected compounds for DC-SIGN, DC-SIGNR, and langerin was also measured by SPR inhibition experiments. The ligand ranking obtained in these solution assays differed quantitatively from that inferred from array binding profiles. As often observed in the study of sugar-protein interactions, the affinity and binding selectivity measured for lectins can strongly depend on the format of the assay, in ways that complicate the discovery process. In the case at hand, the clustered ligand presentation on the array can strongly influence the avidity of the system in ways that cannot be reproduced by binding inhibition experiments, where the monovalent ligand in solution is competing against an immobilized glycoprotein. Thus, the SPR and array binding data should be regarded as comple-

mentary information, describing different features of the ligand-lectin interaction. The microarray format of the test we propose here allows the binding profiles of lectins to be analyzed even if they are available only in minute quantities, as is the case for dectin-2 in this study, and may provide structural information useful in the design of multivalent inhibitors that mimic the dense ligand presentation of the array surface. Further characterization of the binding properties of dectin-2 binders, as well as their structural optimization will be the object of active investigation in our laboratories.

Experimental Section

General

Chemicals were purchased from Sigma-Aldrich or Acros Organics, and specific amines from Key Organics, Crea-Chim, Vitas M Labs, Life Chemicals, Alinda Chemicals, Chem Bridge, or Enamine BB (suppliers indicated in the Supporting Information, Characterization of the ligands) and were used without further purification. All reaction solvents were dried over activated 4 or 3 Å molecular sieves. TLC was carried out using 60 F₂₅₄ TLC plates and visualized by UV irradiation (254 nm) or by staining with cerium molybdate, potassium permanganate, or ninhydrin solution. *Canavalia ensiformis* lectin (ConA) and *Aleuria aurantia* lectin (AAL) were purchased from VectorLabs and labeled with Alexa Fluor® 555 NHS Ester (Succinimidyl Ester) (Thermo Fisher Scientific). Human CLRs were prepared and labeled as described below.

Library design

Markush structures were used to search the compound collections from eMolecules (6 585 694 compounds, August 2013) and Molport (10010542, December 2013) for primary and secondary amines. In particular, only amines with $150 \leq MW \leq 275$, number of heavy atoms 9–20, number of rotatable bonds < 6 , number of rings > 0 and no undefined stereocenters were considered. Amines containing potentially reactive species^[42] and PAINS^[43] were also removed. The resulting compounds were clustered and the cluster center selected (ECFP4 fingerprints as molecular descriptors, maximum distance between cluster members Tanimoto=0.6). To reduce the compounds to a number amenable to visual inspection 20% of the cluster centers was selected maximizing molecular diversity (based on FCFP4 fingerprints). This resulted in a set of 1116 primary amines (630 aliphatic and 486 aromatic) and 796 secondary amines (472 aliphatic and 324 aromatic). After visual inspection and confirmed commercial availability a set of 38 compounds was selected for acquisition.

Synthesis of the ligands

Mannosylated scaffold **11** was prepared according to ref [11c] from acceptor **9** and the known donor **10**. Compound **9**, in turn, was prepared from **8** in two steps, according to ref [11c].

Bis(4-nitrophenyl)-(1S,2S)cyclohex-4-ene-1,2-dicarboxylate (**8**)

Diacid **7** (4.23 g, 24.86 mmol, 1 mol equiv) was dissolved in dry DMF under N₂ and pyridine (5.23 mL, 64.63 mmol, 2.6 mol equiv) was added dropwise to the solution. 4-nitrophenyl trifluoroacetate (14.03 g, 59.66 mmol, 2.4 mol equiv) was added to the mixture and the reaction was stirred overnight at 50 °C. After completion (R_f

(product) = 0.58 in toluene : EtOAc = 8:2 + 0.1% acetic acid), the reaction was diluted with dichloromethane (200 mL) and washed twice with 0.5 M HCl (100 mL), twice with cold, saturated NaHCO₃ (50 mL) and twice with water (50 mL). The organic phase was dried over Na₂SO₄, filtered, and concentrated in vacuo. The obtained crystals were washed with cooled diethyl ether and filtered to yield the pure product **8** as a white powder. Yield: 73%; [α]_D: +130 (c = 1 in CHCl₃ at 20 °C); ¹H NMR (400 MHz, CDCl₃): δ = 8.28–8.22 (m, 4H, H₁₀), 7.27–7.22 (m, 4H, H₉), 5.83 (app. d, *J* = 2.8 Hz, 2H, H₄, H₅), 3.27–3.19 (m, 2H, H₁, H₂), 2.78–2.68 (m, 2H, H_{3ps-eq}, H_{6ps-eq}), 2.48–2.37 ppm (m, 2H, H_{3ps-ax}, H_{6ps-ax}); ¹³C NMR (100 MHz, CDCl₃): δ = 172.8 (C₇); 155.4 (C₁₁); 145.7 (C₈); 125.5 (C₁₀); 124.9 (C₅, C₄); 122.5 (C₉); 41.5 (C₁, C₂); 28.0 ppm (C₃, C₆); MS (ESI): *m/z*: calcd for [C₂₀H₁₆N₂O₈Na]⁺: 435.08 [M+Na]⁺; found: 435.33; m.p. 171 °C.

2,3,4-Tri-*O*-benzoyl- α -L-fucopyranosyl-1-trichloroacetimidate **13**

L-(–)-Fucose (200 mg, 1.22 mmol, 1 mol equiv) was dissolved at –40 °C in pyridine (1 mL) under a N₂ atmosphere and benzoyl chloride (636 μ L, 5.48 mmol, 4.5 mol equiv) was added dropwise to the solution. After 2 h, the starting unprotected sugar was not detected by TLC anymore (*R*_f = 0.1 in toluene : EtOAc = 9:1), so the reaction was left to warm to room temperature and stirred until there was only one major spot visible on the TLC plate (*R*_f = 0.24 in toluene/EtOAc = 97:3). Upon completion, the reaction mixture was diluted with water and extracted with CH₂Cl₂. The joint organic phases were dried over Na₂SO₄, filtered, and concentrated in vacuo, to yield the product 1,2,3,4-tetra-*O*-benzoyl-L-fucopyranose as a white foam. Yield: quant. (α/β = 9:1). Under these conditions, only a small amount of fuco-furanose form is obtained, typically around 1% by ¹H NMR, so that no additional purification is needed before the next step. ¹H NMR (400 MHz, CDCl₃): α -anomer: δ = 8.17–7.21 (m, 20H, OBz), 6.87 (d, 1H, H₁, *J*_{1,2} = 3.7 Hz), 6.08 (dd, 1H, H₃, *J*_{3,4} = 3.2, *J*_{3,2} = 10.8 Hz), 5.99 (dd, 1H, H₂, *J*_{2,3} = 10.8, *J*_{2,1} = 3.7 Hz), 5.90 (dd, 1H, H₄, *J*_{4,5} = 1.1, *J*_{4,3} = 3.2 Hz), 4.64 (dd, 1H, H₅, *J*_{5,6} = 6.6, *J* = 12.9 Hz), 1.32 (d, 3H, H₆, *J* = 6.4 Hz); β -anomer: δ = 8.17–7.21 (m, 20H, OBz), 6.21 (d, 1H, H₁, *J*_{1,2} = 8.3 Hz), 6.06 (m, 1H, H₂), 5.81 (dd, 1H, H₄, *J*_{4,5} = 1.0, *J*_{4,3} = 3.4 Hz), 5.72 (m, 1H, H₃), 4.34 (dd, 1H, H₅, *J*_{5,6} = 5.7, *J* = 12.5 Hz), 1.39 (d, 3H, H₆, *J* = 6.4 Hz); MS (ESI): *m/z*: calcd for [C₃₄H₂₈O₉Na]⁺ [M+Na]⁺: 603.16; found: 603.35.

1,2,3,4-Tetra-*O*-benzoyl-L-fucopyranose (707.3 mg, 1.218 mmol, 1 mol equiv) was dissolved in dry THF and cooled to 0 °C under a N₂ atmosphere. 2 M MeNH₂ in THF (731 μ L, 1.462 mmol, 1.2 mol equiv) was added dropwise to the solution. After 1 h, the major spot on the TLC had a slightly lower *R*_f than the starting material and a new spot started to appear further below. The solution was concentrated in vacuo, and the product was purified by flash chromatography (*R*_f = 0.18 and 0.15 for α and β anomers in toluene/EtOAc = 9:1) to yield pure 2,3,4-tri-*O*-benzoyl-L-fucopyranose as a colourless oil. Yield: 74% (α/β = 2:1); ¹H NMR (400 MHz, CDCl₃): α -anomer: δ = 8.18–7.23 (m, 15H, OBz), 6.04 (dd, 1H, H₃, *J*_{3,2} = 10.7, *J*_{3,4} = 3.2 Hz), 5.81–5.77 (m, 2H, H₁, H₂), 5.68 (m, 1H, H₂), 4.69 (dd, 1H; H₅, *J*_{5,6} = 6.5, *J*_{5,4} = 13.1 Hz), 1.30 (d, 3H, H₆, *J*_{6,5} = 6.5 Hz); β -anomer: δ = 8.18–7.23 (m, 15H, OBz), 5.73 (dd, 1H, H₄, *J*_{4,5} = 1.0, *J*_{3,4} = 3.5 Hz), 5.71–5.66 (m, 1H, H₃), 5.58 (dd, 1H, H₂, *J*_{2,1} = 7.9, *J*_{2,3} = 10.4 Hz), 4.99 (d, 1H, H₁, *J*_{1,2} = 7.9 Hz), 4.13 (dd, 1H, H₅, *J*_{5,6} = 6.1, *J*_{5,4} = 12.8 Hz), 1.39 ppm (d, 3H, H₆, *J*_{6,5} = 6.5 Hz); MS (ESI): *m/z*: calculated for [C₂₇H₂₄O₈Na]⁺ [M+Na]⁺: 499.14; found: 499.84

2,3,4-Tri-*O*-benzoyl-L-fucopyranose (460 mg, 0.969 mmol, 1 mol equiv) was dissolved in dry dichloromethane (5 mL) and cooled to 0 °C. DBU (58 μ L, 0.388 mmol, 0.4 mol equiv) and trichloroacetonitrile (972 μ L, 9.694 mmol, 10 mol equiv) were added to the solution

which was left to warm to room temperature and stirred until the starting material had disappeared. The solution was concentrated in vacuo, and the product was purified by flash chromatography (*R*_f = 0.34 in hexane/EtOAc = 9:1) to yield the pure product as a colourless oil (only α -anomer). Yield: 93%; ¹H NMR (400 MHz, CDCl₃): δ = 8.60 (s, 1H, NH), 8.18–7.25 (m, 15H, OBz), 6.83 (d, 1H, H₁, *J*_{1,2} = 3.5 Hz), 6.04 (dd, 1H, H₃, *J*_{3,2} = 10.9, *J*_{3,4} = 3.5 Hz), 5.91 (dd, 1H, H₂, *J*_{2,3} = 10.9, *J*_{2,1} = 3.5 Hz), 5.89–5.86 (m, 1H, H₄), 4.65 (dd, 1H, H₅, *J*_{5,6} = 6.5, *J*_{5,4} = 6.3 Hz), 1.30 ppm (d, 3H, H₆, *J*_{6,5} = 6.5 Hz); MS (ESI): *m/z*: calcd for [C₂₉H₂₄O₈ClNa]⁺ [M+Na]⁺: 642.05; found: 642.37.

1,2-Cyclohexanedicarboxylic acid, (1*S*,2*S*,4*S*,5*S*)-4-(2-azidoethoxy)-5-[(2,3,4-tri-*O*-benzoyl- β -D-fucopyranosyl)oxy]-1,2-bis(*p*-nitrophenylester) (**14**)

A mixture of acceptor **9** (250 mg, 0.4013 mmol, 1 mol equiv) and donor **13** (207 mg, 0.4013 mmol, 1 mol equiv) was co-evaporated from toluene three times. Powdered and activated 4 Å molecular sieves (acid washed) were added and the mixture was kept under vacuum for a few hours and then dissolved in dry CH₂Cl₂ (4 mL). The solution was cooled to –30 °C and TMSOTf (9 μ L, 0.0401 mmol, 0.2 mol equiv) was added to the mixture. The reaction was stirred at –30 °C for 2 h and at RT for an additional 1.5 h and was then quenched with Et₃N. The mixture was filtered over a Celite pad. The solution was concentrated in vacuo, and the product was purified by flash chromatography (*R*_f = 0.22 in hexane/EtOAc = 6:4) to yield the pure product **14** as a colorless foam. Yield: 72%; ¹H NMR (400 MHz, CDCl₃): δ = 8.23–6.83 (m, 23H, H_{Ar}), 5.75–5.63 (m, 2H, H₂, H₄), 5.57 (dd, 1H, H₃, *J*_{3,2} = 10.5, *J*_{3,4} = 3.4 Hz), 4.86 (d, 1H, H₁, *J*_{1,2} = 8.0 Hz), 4.14 (m, 1H, C₂), 4.06 (q, 1H, H₅, *J*_{5,4} = 12.9, *J*_{5,6} = 6.3 Hz), 3.87 (m, 1H, C₁), 3.84–3.77 (m, 1H, H_{7a}), 3.69–3.62 (m, 1H, H_{7b}), 3.41–3.31 (m, 2H, H_{8a,b}), 3.22–3.13 (m, 1H, C₄), 3.02–2.92 (m, 1H, H₅), 2.36–2.27 (m, 1H, C₃ or 6_{eq}), 2.21–2.12 (m, 1H, H₃ or 6_{eq}), 2.08–1.90 ppm (m, 2H, H_{3ax,6ax}); ¹³C NMR (100 MHz, CDCl₃): δ = 172.5, 172.3 (C₉); 166.3, 165.9, 165.7 (CO_{Bz}); 155.5, 155.3 (C₁₀); 145.8, 145.7 (C₁₃); 133.9, 133.7, 133.5 (CH_{Bz}); 130.3, 130.0, 130.0 (CH_{Bz}); 129.5, 129.1 (C_{quatBz}); 128.9, 128.7, 128.6 (CH_{Bz}); 125.5, 125.4 (C₁₂); 122.7, 122.6 (C₁₁); 100.2 (C₁); 75.0 (C_{C1}); 72.7 (C_{C2}); 72.0 (C₃); 71.3 (C₂); 70.4 (C₅); 70.1 (C₄); 69.6, 68.7 (C₄); 51.2 (C₈); 39.0, 39.0 (C_{C4,C5}); 27.5, 27.2 (C_{C3}, C_{C6}); 16.9, 16.7 ppm (C₆); MS (ESI): *m/z*: calcd for [C₄₉H₄₃N₅O₁₇]⁺: 996.26; found: 996.86.

General procedure for the synthesis of bisamides **12** and **15**

Amine coupling: Scaffold **11** or **14** (1 mol equiv) was dissolved in dry THF or DMF (see the Supporting Information) under N₂ and the amine (3 mol equiv) was added to the solution. For amines sold as ammonium salts and amines with low reactivity (see the Supporting Information) 3 mol equiv of Et₃N were also added. The mixture was stirred at RT from 1 h to 2 days, and monitored by TLC or NMR. Upon completion, the solution was washed with 1 M HCl, 1 M NaOH and water on supported liquid extraction cartridges (Biotage ISOLUTE® HM-N). The crude was purified by flash chromatography (CH₂Cl₂ with gradient of methanol from 0 to 20%) or used without purification in the following Zemplén-deprotection if the purity was satisfying.

Zemplén debenzoylation: The benzoyl-protected bisamide (1 mol equiv) was dissolved in distilled MeOH and 1 M freshly prepared NaOMe in MeOH was added to the solution (1.5 mol equiv NaOMe) to a 0.1 M final concentration of the substrate. After completion, the reaction was neutralized with Amberlite® IR120 hydrogen form ion-exchange resin, filtered, and concentrated in vacuo.

The crude was purified by direct or reverse-phase flash chromatography, yielding the pure product **12** or **15**.

Library characterization: Ligands **12.1**, **12.3**, **12.4**, **12.7**, **12.12**, **12.14**, **12.15**, **12.18**, and **12.19** were previously described by Varga et al.,^[11d] ligand **12.2** was described by Reina et al.,^[44] whereas the characterization of the other ligands is detailed in the Supporting Information.

Conjugation of the ligands with the bifunctional linker **16**

A 10 mM solution of the ligands was prepared in water and when necessary for complete solubilization, 5% DMSO (Thermo Scientific Molecular Probes™) was added. The compounds were stirred overnight at room temperature with equimolar amounts of **16** (*N*-[(1*R*,8*S*,9*S*)-bicyclo[6.1.0]non-4-yn-9-ylmethoxyxycarbonyl]-1,8-diamino-3,6-dioxaoctane, BCN-amine, Sigma Aldrich) in water and the conversion of the SPAAC-reactions was monitored by MALDI-TOF mass spectrometry by using 2,5-dihydroxybenzoic acid (DHB) as a matrix (5 mg mL⁻¹ in CH₃CN/0.1% aqueous TFA, 3:7 containing 0.005% NaCl). The MALDI data are reported in the Supporting Information.

Ligand printing and screening

Stock solutions of the conjugates 1 mM in water were diluted with sodium phosphate buffer (300 mM, pH 8.5, 0.005% Tween® 20, 10% DMSO) to a final concentration of 50 μM. 40 μL of each solution was placed into a 384 well source plate (Scienion, Berlin, Germany), which was stored at -20 °C and reused if necessary. These solutions (750 pL, 3 drops of 250 pL) were spotted onto NHS-functionalized glass slides (Nexterion® Slide H-Schott AG, Mainz, Germany). Ligands were spotted in 4 replicates (9 different ligands per row), establishing the complete microarray that was printed in 7 copies onto each slide. After printing, the slides were placed in a 75% humidity chamber (saturated NaCl solution) at room temperature overnight. The unreacted NHS groups were quenched by placing the slides in a 50 mM solution of ethanolamine in sodium borate buffer 50 mM, pH 9.0, for 1 h.

The immobilized ligands were probed with solutions of fluorescently labeled (Alexafluor555) plant and fungal lectins. Solutions of *Concanavalin A* (ConA-555, 1 μg mL⁻¹) and *Aleuria aurantia* lectin (AAL-555, 15 μg mL⁻¹) were prepared in PBS containing 2 mM CaCl₂, 2 mM MgCl₂, and 0.005% Tween-20. For incubations, 200 μL of the lectin solution was applied to each microarray by using 8 Well ProPlate™ Slide Module incubation chambers for 1 h in the dark at room temperature. The slides were washed under standard conditions (PBS and water), dried with argon, and the introduced fluorescence was analyzed with a microarray scanner.

The immobilized ligands were probed with solutions of fluorescently labeled C-type lectins. Solutions of Cy3-labelled DC-SIGN ECD-Cy3 (50 μg mL⁻¹, DOL: 0.3) and DC-SIGNR ECD (150 μg mL⁻¹, DOL: 0.95), langerin-Cy3 (25 μg mL⁻¹, DOL: 0.7) and dectin-2-Cy3 (50 μg mL⁻¹, DOL: 0.4) were diluted in TBS (50 mM Tris-HCl, 150 mM NaCl, pH 8.0) containing 4 mM CaCl₂, 0.5% BSA and 0.005% Tween® 20. For incubations, 200 μL of each lectin solution was applied to each subarray by using 8 Well ProPlate Module incubation chambers. The microarray was incubated by gentle shaking overnight in the dark at 4 °C. The slides were washed using TBS containing 4 mM CaCl₂ and water, dried with argon, and the fluorescence was analyzed with a microarray scanner.

Expression and purification of CLR

DC-SIGN extracellular domain (DC-SIGN ECD) and langerin extracellular domain (langerin ECD) constructs were produced and purified as previously described.^[3a,24]

DC-SIGNR ECD and dectin-2 ECD were expressed in *E. coli* BL21(DE3) in 1 L of LB medium supplemented with 50 μg mL⁻¹ kanamycin at 37 °C. Expression was induced by addition of 1 mM isopropyl 1-thio- β -galactopyranoside (IPTG) when the culture had reached an A_{600 nm} of 0.8 and was maintained for 3 h. The protein was expressed in the cytoplasm as inclusion bodies. Cells were harvested by a 20 min centrifugation at 5000 g at 4 °C. The pellet was re-suspended in 30 mL of a solution containing 150 mM NaCl, 25 mM Tris-HCl, pH 8 and one antiprotease mixture tablet (Complete EDTA free, Roche). Cells were disrupted by sonication and cell debris eliminated by centrifugation at 100000 g for 45 min at 4 °C in a Beckman 45Ti rotor. The pellet was solubilized in 30 mL of 6 M guanidine-HCl containing 25 mM Tris-HCl pH 8, 150 mM NaCl and 0.01% β -mercaptoethanol. The mixture was centrifuged at 100000 g for 45 min at 4 °C and the supernatant was diluted 5-fold, by slow addition with stirring, with 1.25 M NaCl, 25 mM CaCl₂, and 25 or 200 mM Tris-HCl pH 8 for DC-SIGNR and dectin-2 ECD, respectively. The diluted mixture was dialyzed against ten volumes of 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mM CaCl₂ (buffer A) with three buffer changes. After dialysis, insoluble precipitate was removed by centrifugation at 100000 g for 1 h at 4 °C. The supernatant containing DC-SIGNR ECD was loaded on Mannan agarose column (Sigma) for purification by affinity chromatography equilibrated with buffer A. After loading, DC-SIGNR ECD was tightly bound to the column and eluted in the same buffer without CaCl₂ but supplemented with 1 mM EDTA (buffer B). This step was followed by SEC (size exclusion chromatography) by using a Superose 6 column (GE Healthcare) equilibrated with buffer A. Fractions were analyzed by SDS-PAGE (12%) and DC-SIGNR ECD containing fractions were pooled and concentrated by ultrafiltration (YM10 membrane from Amicon). The supernatant containing the Strep tagged dectin-2 ECD was loaded onto a StrepTrap HP column (GE Healthcare) at 4 °C. Unbound proteins were washed away with buffer A before dectin-2 ECD was eluted with buffer C (150 mM NaCl, 25 mM Tris-HCl, pH 8, 4 mM CaCl₂, 2.5 mM D-desthiobiotin). Eluted fractions were analyzed by SDS-PAGE (15%) and dectin-2 ECD-containing fractions were pooled and concentrated by ultrafiltration (YM10 membrane from Amicon).

Each protein construct was checked by N-terminal amino acid sequencing and mass spectrometry.

The labeling procedure is described in the Supporting Information.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 642870 (Immunoshape). Support by the Italian Ministry of Research through a PRIN grant (prot. 2015RNWJAM 002) and the Spanish Ministry of Economy, Industry and Competitiveness (grant CTQ2017-90039-R to N.R.) is acknowledged. HMRS analyses were obtained from the UNITECH "COSPECT" platform at the University of Milan. For human CLR ECD production, this work used the Multistep Protein Purification Platform (MP3) and the SPR platform for the competition test of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from

FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology. F.F. also acknowledges the support of the French ANR for Glyco@Alps (ANR-15-IDEX-02).

Conflict of interest

The authors declare no conflict of interest.

Keywords: carbohydrates · C-type lectins · drug discovery · glycomimetics · microarrays

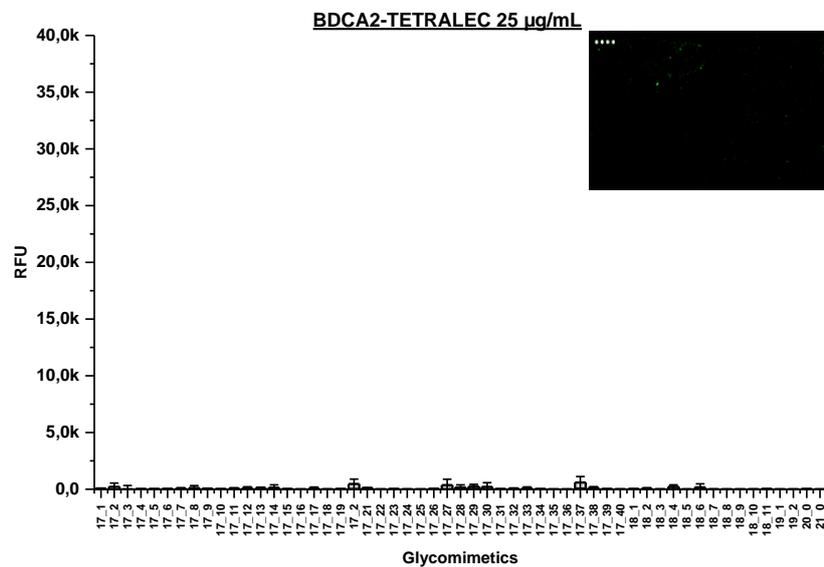
- [1] K. Drickamer, *Nature* **1992**, *360*, 183.
- [2] G. Navarra, P. Zihlmann, R. P. Jakob, K. Stangier, R. C. Preston, S. Rabbani, M. Smiesko, B. Wagner, T. Maier, B. Ernst, *ChemBioChem* **2017**, *18*, 539–544.
- [3] a) M. Thépaut, J. Valladeau, A. Nurisso, R. Kahn, B. Arnou, C. Vivès, S. Saeland, C. Ebel, C. Monnier, C. Dezutter-Dambuyant, A. Imberty, F. Fieschi, *Biochemistry* **2009**, *48*, 2684–2698; b) L. Chatwell, A. Holla, B. B. Kaufner, A. Skerra, *Mol. Immunol.* **2008**, *45*, 1981–1994.
- [4] 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.
- [5] E. Chabrol, A. Nurisso, A. Daina, E. Vassal-Stermann, M. Thepaut, E. Girard, R. R. Vives, F. Fieschi, *PLoS One* **2012**, *7*, e50722.
- [6] a) T. Klein, D. Abgottsson, M. Wittwer, S. Rabbani, J. Herold, X. Jiang, S. Kleeb, C. Luthi, M. Scharenberg, J. Bezencon, E. Gubler, L. Pang, M. Smiesko, B. Cutting, O. Schwardt, B. Ernst, *J. Med. Chem.* **2010**, *53*, 8627–8641; b) S. Cecioni, A. Imberty, S. Vidal, *Chem. Rev.* **2015**, *115*, 525–561; c) S. Sattin, A. Bernardi, *Carbohydr. Chem.* **2015**, *41*, 1–25; d) C. Colombo, A. Bernardi, in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier **2017**.
- [7] J. K. Sprockholt, R. J. Overmars, T. B. Geijtenbeek, in *C-Type Lectin Receptors in Immunity*, Springer, Berlin, **2016**, pp. 129–150.
- [8] K. Brzezicka, B. Echeverria, S. Serna, A. van Diepen, C. H. Hokke, N. C. Reichardt, *ACS Chem. Biol.* **2015**, *10*, 1290–1302.
- [9] J. Aretz, H. Baukmann, E. Shanina, J. Hanske, R. Wawrzinek, V. A. Zapol'skii, P. H. Seeberger, D. E. Kaufmann, C. Rademacher, *Angew. Chem. Int. Ed.* **2017**, *56*, 7292–7296; *Angew. Chem.* **2017**, *129*, 7398–7402.
- [10] a) 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; b) C. P. Alvarez, F. Lasala, J. Carrillo, O. Muñoz, A. L. Corbí, R. Delgado, *J. Virol.* **2002**, *76*, 6841–6844.
- [11] a) S. Sattin, F. Fieschi, A. Bernardi, in *Carbohydrate Chemistry: State of the Art and Challenges for Drug Development*, Imperial College Press, **2015**, pp. 379–394; b) N. Obermajer, S. Sattin, C. Colombo, M. Bruno, U. Svajger, M. Anderluh, A. Bernardi, *Mol. Diversity* **2011**, *15*, 347–360; c) 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; d) V. Porkolab, E. Chabrol, N. Varga, S. Ordanini, I. Sutkeviciute, M. Thepaut, M. J. Garcia-Jimenez, E. Girard, P. M. Nieto, A. Bernardi, F. Fieschi, *ACS Chem. Biol.* **2018**, *13*, 600–608; e) A. Tamburrini, S. Achilli, F. Vasile, S. Sattin, C. Vives, C. Colombo, F. Fieschi, A. Bernardi, *Bioorg. Med. Chem.* **2017**, *25*, 5142–5147.
- [12] a) 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; b) J. Luczkowiak, S. Sattin, I. Sutkeviciute, J. J. Reina, M. Sanchez-Navarro, M. Thepaut, L. Martinez-Prats, A. Daggetti, F. Fieschi, R. Delgado, A. Bernardi, J. Rojo, *Bioconjugate Chem.* **2011**, *22*, 1354–1365; c) A. Berzi, J. J. Reina, R. Ottria, I. Sutkeviciute, P. Antonazzo, M. Sanchez-Navarro, E. Chabrol, M. Biasin, D. Trabattoni, I. Cetin, J. Rojo, F. Fieschi, A. Bernardi, M. Clerici, *AIDS* **2012**, *26*, 127–137; d) S. Ordanini, N. Varga, V. Porkolab, M. Thepaut, L. Belvisi, A. Bertaglia, A. Palmioli, A. Berzi, D. Trabattoni, M. Clerici, F. Fieschi, A. Bernardi, *Chem. Commun.* **2015**, *51*, 3816–3819.
- [13] L. de Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. de Jong, T. de Gruijl, V. Piguet, Y. van Kooyk, T. B. Geijtenbeek, *Nat. Med.* **2007**, *13*, 367–371.
- [14] H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger, W. I. Weis, *J. Biol. Chem.* **2006**, *282*, 4202–4209.
- [15] H. Feinberg, M. E. Taylor, N. Razi, R. McBride, Y. A. Knirel, S. A. Graham, K. Drickamer, W. I. Weis, *J. Mol. Biol.* **2011**, *405*, 1027–1039.
- [16] H. Feinberg, S. A. F. Jegouzo, M. J. Rex, K. Drickamer, W. I. Weis, M. E. Taylor, *J. Biol. Chem.* **2017**, *292*, 13402–13414.
- [17] a) O. Blixt, S. Han, L. Liao, Y. Zeng, J. Hoffmann, S. Futakawa, J. C. Paulson, *J. Am. Chem. Soc.* **2008**, *130*, 6680–6681; b) C. D. Rillahan, E. Schwartz, R. McBride, V. V. Fokin, J. C. Paulson, *Angew. Chem. Int. Ed.* **2012**, *51*, 11014–11018; *Angew. Chem.* **2012**, *124*, 11176–11180; c) C. D. Rillahan, E. Schwartz, C. Rademacher, R. McBride, J. Rangarajan, V. V. Fokin, J. C. Paulson, *ACS Chem. Biol.* **2013**, *8*, 1417–1422; d) J. Y. Hyun, C. W. Park, Y. Liu, D. Kwon, S. H. Park, S. Park, J. Pai, I. Shin, *ChemBioChem* **2017**, *18*, 1077–1082.
- [18] a) S. Fukui, T. Feizi, C. Galustian, A. M. Lawson, W. Chai, *Nat. Biotechnol.* **2002**, *20*, 1011–1017; b) W. G. T. Willats, S. E. Rasmussen, T. Kristensen, J. D. Mikkelsen, J. P. Knox, *Proteomics* **2002**, *2*, 1666–1671; c) S. Park, I. Shin, *Angew. Chem. Int. Ed.* **2002**, *41*, 3180–3182; *Angew. Chem.* **2002**, *114*, 3312–3314; d) D. Wang, S. Liu, B. J. Trummer, C. Deng, A. Wang, *Nat. Biotechnol.* **2002**, *20*, 275.
- [19] C. Zilio, A. Bernardi, A. Palmioli, M. Salina, G. Tagliabue, M. Buscaglia, R. Consonni, M. Chiari, *Sens. Actuators B* **2015**, *215*, 412–420.
- [20] S. Serna, J. Etxebarria, N. Ruiz, M. Martin-Lomas, N. C. Reichardt, *Chem. Eur. J.* **2010**, *16*, 13163–13175.
- [21] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
- [22] I. Goldstein, C. Hollerman, E. Smith, *Biochemistry* **1965**, *4*, 876–883.
- [23] J. Olausson, L. Tibell, B.-H. Jonsson, P. Pahlsson, *Glycoconjugate J.* **2008**, *25*, 753.
- [24] G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand, F. Fieschi, *J. Biol. Chem.* **2009**, *284*, 21229–21240.
- [25] S. I. Gringhuis, M. van der Vlist, L. M. van den Berg, J. den Dunnen, M. Litjens, T. B. Geijtenbeek, *Nat. Immunol.* **2010**, *11*, 419–426.
- [26] a) Y. van Kooyk, W. W. Unger, C. M. Fehres, H. Kalay, J. J. Garcia-Vallejo, *Mol. Immunol.* **2013**, *55*, 143–145; b) C. A. Aarnoudse, M. Bax, M. Sanchez-Hernandez, J. J. Garcia-Vallejo, Y. van Kooyk, *Int. J. Cancer* **2008**, *122*, 839–846.
- [27] B. J. Appelmelk, I. van Die, S. J. van Vliet, C. M. J. E. Vandenbroucke-Grauls, T. B. H. Geijtenbeek, Y. van Kooyk, *J. Immunol.* **2003**, *170*, 1635–1639.
- [28] S. Sattin, PhD thesis, Università degli Studi di Milano **2009**.
- [29] S. Pohlmann, E. J. Soilleux, F. Baribaud, G. J. Leslie, L. S. Morris, J. Trowsdale, B. Lee, N. Coleman, R. W. Doms, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2670.
- [30] E. J. Soilleux, R. Barten, J. Trowsdale, *J. Immunol.* **2000**, *165*, 2937–2942.
- [31] Y. Guo, I. Nehlmeier, E. Poole, C. Sakonsiniri, N. Hondow, A. Brown, Q. Li, S. Li, J. Whitworth, Z. Li, A. Yu, R. Brydson, W. B. Turnbull, S. Pohlmann, D. Zhou, *J. Am. Chem. Soc.* **2017**, *139*, 11833–11844.
- [32] F. Zhang, S. Ren, Y. Zuo, *Int. Rev. Immunol.* **2014**, *33*, 54–66.
- [33] a) N. Horan, L. Yan, H. Isobe, G. M. Whitesides, D. Kahne, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11782; b) R. Raman, K. Tharakaraman, V. Sasisekharan, R. Sasisekharan, *Curr. Opin. Struct. Biol.* **2016**, *40*, 153–162; c) R. J. E. Li, S. J. van Vliet, Y. van Kooyk, *Curr. Opin. Biotechnol.* **2018**, *51*, 24–31; d) P.-H. Liang, S.-K. Wang, C.-H. Wong, *J. Am. Chem. Soc.* **2007**, *129*, 11177–11184.
- [34] M. Merad, F. Ginhoux, M. Collin, *Nat. Rev. Immunol.* **2008**, *8*, 935–947.
- [35] a) M. A. de Jong, L. E. Vriend, B. Theelen, M. E. Taylor, D. Fluitsma, T. Boekhout, T. B. Geijtenbeek, *Mol. Immunol.* **2010**, *47*, 1216–1225; b) M. van der Vlist, T. B. Geijtenbeek, *Immunol. Cell Biol.* **2010**, *88*, 410–415.
- [36] A. Holla, A. Skerra, *Protein Eng. Des. Sel.* **2011**, *24*, 659–669.
- [37] M. J. Fernandes, A. A. Finnegan, L. D. Siracusa, C. Brenner, N. N. Iscove, B. Calabretta, *Cancer Res.* **1999**, *59*, 2709–2717.
- [38] a) K. Sato, X. L. Yang, T. Yudate, J. S. Chung, J. Wu, K. Luby-Phelps, R. P. Kimberly, D. Underhill, P. D. Cruz, Jr., K. Arizumi, *J. Biol. Chem.* **2006**, *281*, 38854–38866; b) E. P. McGreal, M. Rosas, G. D. Brown, S. Zamze, S. Y. Wong, S. Gordon, L. Martinez-Pomares, P. R. Taylor, *Glycobiology* **2006**, *16*, 422–430.
- [39] a) O. Gross, H. Poeck, M. Bscheider, C. Dostert, N. Haneschlager, S. Andres, G. Hartmann, A. Tardivel, E. Schweighoffer, V. Tybulewicz, A. Mocsai, J. Tschopp, J. Ruland, *Nature* **2009**, *459*, 433–436; b) B. Kerscher, J. A. Willment, G. D. Brown, *Int. Immunol.* **2013**, *25*, 271–277.

- [40] M. J. Robinson, F. Osorio, M. Rosas, R. P. Freitas, E. Schweighoffer, O. Gross, J. S. Verbeek, J. Ruland, V. Tybulewicz, G. D. Brown, L. F. Moita, P. R. Taylor, C. Reis e Sousa, *J. Exp. Med.* **2009**, *206*, 2037–2051.
- [41] M. Thépaut, C. Guzzi, I. Sutkevičute, S. Sattin, R. Ribeiro-Viana, N. Varga, E. Chabrol, J. Rojo, A. Bernardi, J. Angulo, P. M. Nieto, F. Fieschi, *J. Am. Chem. Soc.* **2013**, *135*, 2518–2529.
- [42] a) R. F. Bruns, I. A. Watson, *J. Med. Chem.* **2012**, *55*, 9763–9772; b) R. Brenk, A. Schipani, D. James, A. Krasowski, I. H. Gilbert, J. Frearson, P. G. Wyatt, *ChemMedChem* **2008**, *3*, 435–444.
- [43] J. B. Baell, G. A. Holloway, *J. Med. Chem.* **2010**, *53*, 2719–2740.
- [44] J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martinez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo, A. Bernardi, *ChemMedChem* **2007**, *2*, 1030–1036.

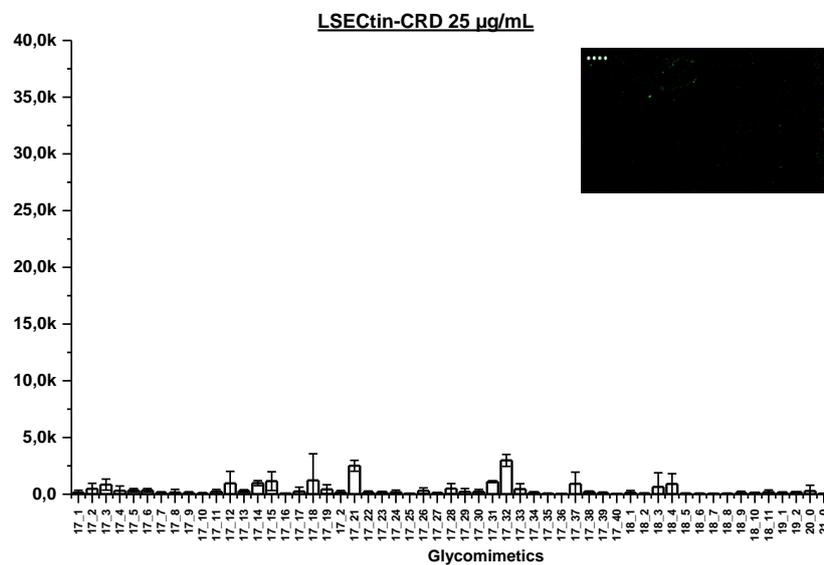
Manuscript received: May 22, 2018
Revised manuscript received: July 3, 2018
Accepted manuscript online: July 5, 2018
Version of record online: ■■■■■, 0000

9.2.2 Other glycomimetic array screening

BDCA2 and LSEctin were not expected to interact with the printed glycomimetics. Nevertheless, incubation with BDCA2 random TETRALEC and LSEctin CRD (Fig.107) was performed but all the values were considered as background binding.



A



b

Fig.107 Other glycomimetic arrays a) BDCA2 random TETRALEC and b) LSEctin CRD incubation on the glycomimetic array.

As expected, no binding was observed with any of the glycomimetic.

10. Characterization of new glycomimetics specific to DC-SIGN

This chapter focuses on the optimization of glycomimetics in term of affinity and specificity towards DC-SIGN. A tight collaboration with the chemistry laboratory of Prof. Anna Bernardi was established years ago to develop inhibitors selective for DC-SIGN in the context of HIV study. As already mentioned in chapter **2.3.5**, DC-SIGN and langerin both interacts with the mannose of the HIV gp120 but lead to different responses: DC-SIGN is involved in the spreading of the virus, while langerin blocks the infection. For this reason, glycomimetics that selectively inhibit DC-SIGN interaction with gp120 without affecting langerin interaction have to be developed. Since the two CLRs share the same natural disaccharide ligand Man α 1-2Man (Fig.108a), our team has been involved in the characterization of compounds that mimic Man α 1-2Man modified to be selective towards DC-SIGN. The **first generation** of glycomimetics led to the development of psDi (section **3.2.1**) [112] (Fig.108b) as lead compound for the rational synthesis of a **second-generation** compound: Man030 (section **3.2.1**)[163] (Fig.108b) .

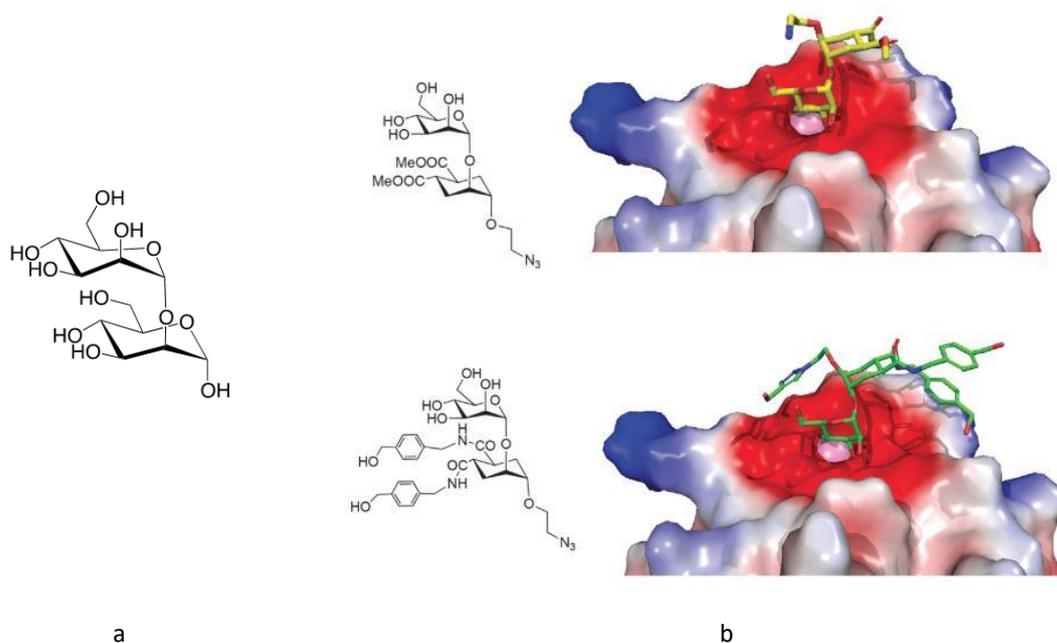


Fig.108 First and second generation glycomimetics against DC-SIGN a) Man α 1-2Man structure and b) psDi/DC-SIGN CRD and Man030/DC-SIGN CRD structures. Adapted from [111].

Man030 was characterized by a good affinity towards DC-SIGN and implementation towards its selectivity was achieved by studying the structure of DC-SIGN and langerin CRDs. Langerin possesses two lysines in proximity of the binding site creating a positively charged environment. Man030 was, therefore, implemented into 6NH₂-Man030 (section **3.2.1**) (Fig.109), with a positive charge in C6 position to establish electrostatic repulsion with the positive charges of langerin lysines. This new glycomimetic combines

affinity enhancement and strong selectivity towards DC-SIGN, with a complete exclusion of binding towards langerin ($IC_{50} > 4400 \mu M$).

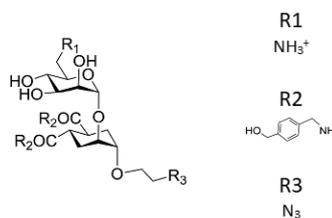


Fig.109 6NH₂-Man030 structure, adapted from[111].

Following up the project, a **third generation** of glycomimetics was developed, exploring two different but not mutually exclusive approaches:

- 1) To enhance the robustness of the glycomimetic and its stability against glycosidases, thio glycomimetics were synthesized by the group of Anna Bernardi.

In addition, C-glycosides have been synthesized by the group of Prof. J. Moravcova (Institute of Chemical Technology, Praha, Czech Republic)

- 2) To enhance affinity towards DC-SIGN, computational screening was performed by the group of Sonsoles Martin-Santamaria (Centro de Investigaciones Biológicas, CSIC-CIB, Madrid, Spain) to identify fragments that could interact with the protein in the vicinity of the Ca²⁺ binding site. The X-ray structure of the psDi/DC-SIGN complex (2XR5) indicated that the mannose 2-OH position (Fig.110, black arrow) was oriented towards the protein surface and not towards the solvent. Thus, this position of the ligand, that haven't been exploited up to now by us, could be a site of optimization if fragments are observed in this area by the above mentioned computational screening.

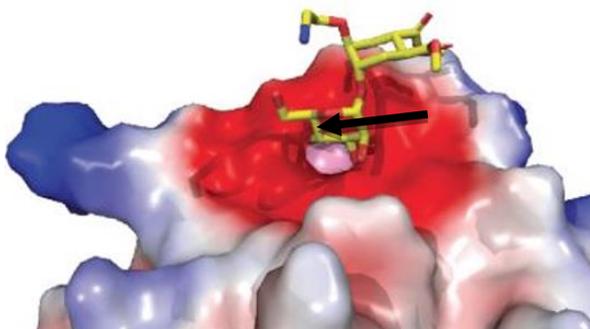


Fig.110 psDi/DC-SIGN CRD X-ray structure. The arrow indicate the mannose 2-OH position. Adapted from [111].

Indeed, appropriate candidates obtained from the docking were characterized by a possible site for an ammonium ion and an aromatic ring therefore, one of our approach consisted in modification of the ligand in order to reach this “ammonium pocket”.

10.1 Robustness enhancement

10.1.1 Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist.

Glycosidases hydrolysis constitutes a major problematic for compounds design for medicinal applications. Changing the glycosidic linkage of the glycomimetic may enhance its resistance towards enzymatic degradation. The following article by *Tamburrini et al* addresses this question by substituting the O-glycosidic linkage by a S-bound in the PsDi glycomimetic, creating the thio-psDi (Fig.111).

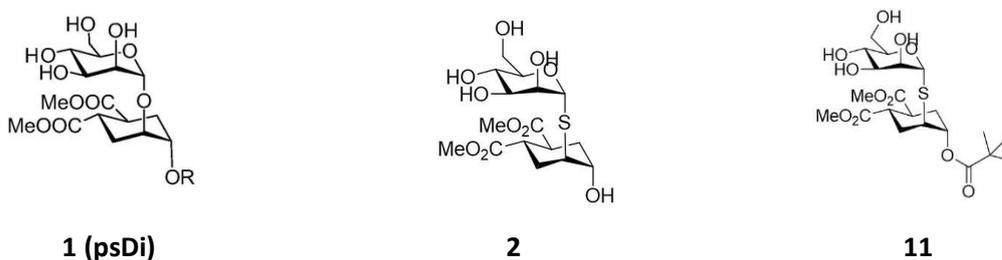


Fig.111 psDi and thio-psDi structures.

SPR competition assays were performed to investigate whether the S-linkage impacts the inhibition power of compounds 2 and 11 towards DC-SIGN interaction with the mannosylated surface.

The IC_{50} values obtained ($2 = 801,1 \pm 13,79 \mu M$; $11 = 782.13 \pm 8,82 \mu M$) showed that thio-psDi acts as a DC-SIGN antagonist, with an affinity similar to PsDi. ($714.7 \pm 9,2 \mu M$)

The main outcome of these studies:

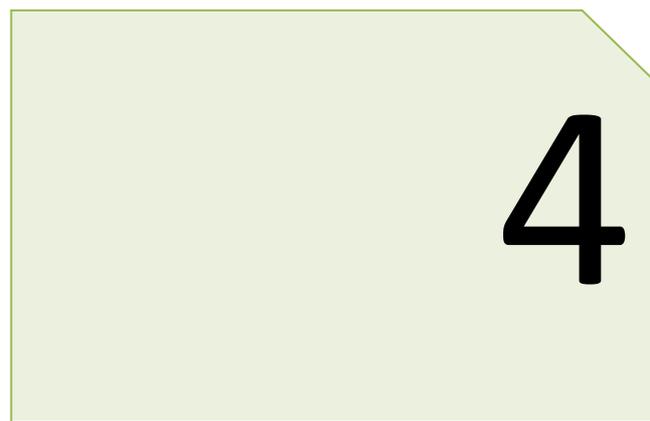
Changing the glycosidic nature of the glycomimetic bound towards S-linkage does not affect the capacity of the compound to inhibit DC-SIGN interaction with the mannosylated surface.

Contributions:

All the compounds were designed and synthesized by the group of Pr. Anna Bernardi.

My contribution to this study:

I have prepared DC-SIGN required for the SPR experiments and performed all the described SPR assays, analysed data and participated to the preparation of the manuscript.



4



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist

Alice Tamburrini^a, Silvia Achilli^b, Francesca Vasile^a, Sara Sattin^a, Corinne Vivès^b, Cinzia Colombo^a, Franck Fieschi^b, Anna Bernardi^{a,*}

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

^b Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

ARTICLE INFO

Article history:

Received 31 January 2017

Revised 20 March 2017

Accepted 21 March 2017

Available online 22 March 2017

Keywords:

Glycomimetics

DC-SIGN

Thioglycosides

Mannobioside

NOESY

ABSTRACT

The synthesis and conformational analysis of pseudo-thio-1,2-dimannoside are described. This molecule mimics mannobioside (Man α (1,2)Man) and is an analog of pseudo-1,2-dimannoside, with expected increased stability to enzymatic hydrolysis. A short and efficient synthesis was developed based on an epoxide ring-opening reaction by a mannosyl thiolate, generated *in situ* from the corresponding thioacetate. NMR-NOESY studies supported by MM3^{*} calculations showed that the pseudo-thio-1,2-dimannoside shares the conformational behavior of the pseudo-1,2-dimannoside and is a structural mimic of the natural disaccharide. Its affinity for DC-SIGN was measured by SPR and found to be comparable to the corresponding *O*-linked analog, offering good opportunities for further developments.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past decades, an increased understanding of glycans ability to encode biochemical information relevant to both physiological and pathological conditions has stimulated a fast growth of glycochemical research.^{1–4} One of the important objectives in this field is the development of glycomimetic structures, targeted against sugar binding proteins (called lectins) and able to alter the read-out of sugar-encoded information. In this context, glycomimetics have been proposed both as probes of biological processes and for medicinal applications.^{5–9}

Many modifications have been introduced in the structure of carbohydrates to generate glycomimetics with improved drug-like characteristics and stability to enzymatic degradation.¹⁰ In particular, thioglycosides have raised a great deal of interest, because the *S*-glycosidic linkage is typically more resistant toward both acid-catalyzed and enzymatic hydrolysis than *O*-glycosides,^{11,12} and yet oxygen and sulfur share similar bonding and geometries, allowing structural similarity. From a structural point of view, the C–S bond is longer than the corresponding C–O bond by about 0.4 Å and the *exo*-anomeric effect in thiodisaccharides is less pronounced than in regular sugars.¹³ As a result, an increased conformational flexibility of the glycosidic torsions is often observed.¹⁴

We previously reported that the pseudo-1,2-dimannoside **1** is a structural mimic of 1,2-mannobioside (Man α (1,2)Man) and a slow-reacting substrate of jack-bean mannosidase.¹⁵ Like the natural disaccharide, this molecule is an antagonist¹⁶ of the dendritic cell receptor DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3 (ICAM-3)-Grabbing Non-integrin), a tetrameric C-type lectin involved in the recognition of viruses and pathogens at the mucosal level.¹⁷ Based on this framework, we have developed a family of mannose-based DC-SIGN monovalent and polyvalent ligands that were shown to work as potent inhibitors of DC-SIGN mediated viral infections.^{18–21} In an effort to improve the hydrolytic stability of **1** and to possibly simplify its synthesis, we have examined the corresponding pseudo-thio-1,2-dimannoside **2** (Fig. 1).

We herein report the synthesis of **2**, an analysis of its conformation based on NMR-NOESY studies (in D₂O) supported

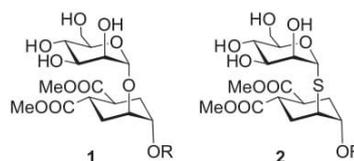


Fig. 1. Pseudo-1,2-dimannoside **1** and its thio analog **2**.

* Corresponding author.

E-mail address: anna.bernardi@unimi.it (A. Bernardi).

by computational modeling, and the evaluation of its interaction with DC-SIGN by Surface Plasmon Resonance (SPR).

2. Results

2.1. Synthesis of the pseudo-thio-1,2-dimannoside

The synthesis of S-linked glycosides can be readily achieved by reaction of thiolates with glycosyl donors.²² For the case at hand, however, a more expeditious approach was employed, exploiting nucleophilic opening of epoxide **3**¹⁶ (Scheme 1) by a glycosyl thiol (1-thio sugar). Epoxide **3** is indeed a known intermediate in the synthesis of **1**, while the required glycosyl thiol can be generated *in situ* from 2,3,4,6-tetra-O-acetyl-1-S-acetyl- α -D-mannopyranose **4** by selective S-deacetylation (Scheme 1).^{23,24} This allows, in principle, a one-pot synthesis of the pseudo-thio-1,2-dimannoside framework **5**.

2,3,4,6-Tetra-O-acetyl-1-S-acetyl- α -D-mannopyranose **4** is a known compound,²³ which is most conveniently prepared from penta-O-acetylmannopyranose **6** by acid catalyzed reaction with thioacetic acid **7** (Scheme 2). The reaction affords **4** as a 9:1 α : β mixture (61% yield), from which pure α -**4** can be chromatographically isolated.

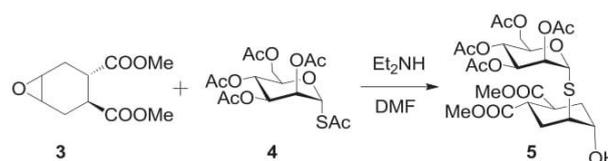
Selective thio-deacetylation of purified α -**4** with excess Et₂NH affords the anomeric thiol **8**, in equilibrium with the more reactive thiolate **9** (Scheme 3). Alkylation of **9** with primary iodides has recently been reported for the synthesis of S-linked 1,6-oligomannosides.²⁴ Alkylation with **3** (DMF, room temperature) gave almost full conversion of the epoxide in 4 h (as evaluated by TLC and ¹H NMR monitoring), affording product **5** in 68% yield after chromatography. Addition of MeOH (5 mol equiv) as a supplementary proton source did not modify the outcome of the reaction.

Compound **5** was obtained as a single isomer from a completely selective opening of **3**, affording *trans*-diaxial cyclohexane disubstitution[#] (Scheme 1). Analysis of the coupling constants in the ¹H NMR spectrum (CD₃OD, Fig. 2) clearly shows that the conformation of the cyclohexane ring depicted in Scheme 1 and Fig. 2, carrying both carbomethoxy groups in the equatorial position is populated by $\geq 90\%$. This feature is characteristic of vicinal *trans*-cyclohexanedicarboxylic acid derivatives,²⁵ including the pseudo-1,2-dimannoside **1**,¹⁵ and it allows the distal substituents in **2** to occupy the axial position, thus mimicking the 3D features of an α -mannopyranose residue.

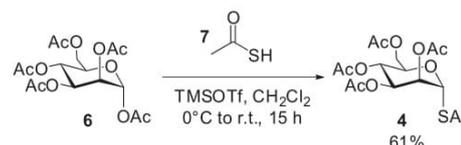
The tetra-O-acetyl-pseudo-thio-1,2-dimannoside **5** was deprotected (MeONa in MeOH), giving **2** (R = H) in quantitative yield (Scheme 4). To allow the synthesis of multivalent derivatives of **2**, we also pursued selective functionalization of the free alcohol group in the 1' position of **5**. As a model reaction, we initially showed that **5** could be transformed in the pivaloate **10** (pivaloyl chloride, CH₂Cl₂, pyridine) and selectively deacetylated under Zemplén conditions (0.02 M MeONa in MeOH, room temperature, 5 min) to afford **11** (Scheme 4) in quantitative yield.

This set the stage for the synthesis of **14** (Scheme 5), a pseudo-thio-1,2-dimannoside functionalized with an azide tether, which allows for further manipulation and transformation into multivalent constructs. Compound **14** was obtained by reaction of **5** with 2,2-dimethyl-3-azido-propanoyl chloride **12**, synthesized according to a method described by Sewald et al.²⁶ (with modifications), to afford the ester **13** in 79% yield. Carefully controlled Zemplén deacetylation of **13**, as shown above, afforded **14** in quantitative yield.

[#] A second stereoisomer ($\leq 3\%$) was observed in the ¹H NMR spectrum (CDCl₃) of the crude reaction product, deriving from opening of the minor enantiomer of **3** that has 94% e.e. This stereoisomer could be chromatographically removed from **5** (4:6 Hex:EtOAc).



Scheme 1. One-pot synthesis of **5**.



Scheme 2. Synthesis of **4** from penta-O-acetylmannopyranose **6**.

2.2. Conformational analysis of the pseudo-thio-1,2-dimannoside

The conformation of **2** (R = H) was investigated by NOESY experiments; the data collected were correlated to computational models and compared to similar information on the pseudo-1,2-dimannoside **1**.

A previous NMR study on **1**¹⁵ had revealed that this mimic has a conformational distribution similar to the natural disaccharide Man α (1,2)Man. Two conformations, dubbed *E* (for Extended) and *S* (for Stacked), are in fast equilibrium around the (pseudo)glycosidic bonds and are characterized by two mutually exclusive NOE contacts: the H1-H1' contact for the *S* conformer, and the H1-H3'_{eq} contact for the *E* conformer (Fig. 3a, the unconventional numbering of the cyclohexane ring was adopted for easier comparison with natural sugars). A qualitative analysis of the NOE data allowed deducing that the *E* conformer of **1** is the preferred conformation in solution (D₂O). Interestingly, the *E* conformation was also the one observed in the X-ray of the DC-SIGN:**1** complex.²⁷

A similar analysis performed on **2** (R = H, D₂O, 500 MHz) showed an intense H1-H3'_{eq} cross-peak (Fig. 4), while the H1-H1' contact was revealed as a weak cross-peak (Fig. 4, inset) using 800 ms mixing time. A conformational search of **2** performed with the MM3* force field supported the existence of both *S* and *E* conformers of similar strain energy (Fig. 3b). Both R = Me and R = H were calculated, obtaining the same trends. Data are shown for R = Me). However, due to the increased length of the C–S bond relative to the C–O bond, the calculated H1-H1' distance in the *S* conformation of **2** (**S** in Fig. 3b) is 3.1 Å, as opposed to 2.6 Å in **1** (Fig. 3a), in qualitative agreement with a reduced intensity of the H1-H1' cross-peak. Overall, the data are consistent with both the pseudo-1,2-dimannoside **1** and the pseudo-thio-1,2-dimannoside **2** populating the same conformational space and mimicking the conformational features of the native Man α (1,2)Man sugar.

2.3. Surface plasmon resonance inhibition studies with DC-SIGN

In order to investigate the biological activity of the pseudo-thio-1,2-dimannoside and compare it to the O-analog, compounds **2** (R = H) and **11** were tested by a well-established Surface Plasmon Resonance (SPR) inhibition assays.²⁷ In this test, increasing concentrations of the glycomimetics are used to inhibit the binding of DC-SIGN ECD (extra cellular domain) to mannosylated BSA immobilized on an SPR sensor chip. The calculated IC₅₀ values for **1** (R = CH₂CH₂N₃), **2** and **11** in the same campaign (720 μ M, 776 μ M and 776 μ M, respectively) are remarkably similar to one another, showing that the replacement of the glycosidic oxygen atom with a sulfur atom is not detrimental to the binding affinity

in ppm. The ^1H and ^{13}C 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) or ThermoFischer LCQ apparatus (ESI ionization). Specific optical rotation values were measured using a Perkin-Elmer 241, at 589 nm in a 1 dm cell.

4.2. NMR NOESY experiments

NMR spectra of **2** ($R = \text{H}$) were obtained using 5 mg of **2** in 500 μL D_2O (24 mM). NOESY spectra were acquired using a Bruker 500 MHz instrument at 300 K. Spectra (8 scans and 256 increments) were collected using three mixing times (300, 600 and 800 ms). Water suppression was achieved by excitation-sculpting pulse sequence.

4.3. Synthesis

Compound **3** was prepared as previously described.¹⁶ Compound **4** was prepared from commercially available penta-*O*-acetylmannopyranose **6** following a reported procedure²³ and 2,2-dimethyl-3-azido-propanoylchloride **12** was synthesized from commercially available 2,2-dimethyl-3-chloro-propanoic acid according to Sewald,²⁶ with modifications.

4.3.1. 2,3,4,6-Tetra-*O*-acetyl-1-*S*-acetyl- α -*D*-mannopyranose **4**

To a solution of penta-*O*-acetylmannopyranose **6** (1.021 g, 2.62 mmol) in dry CH_2Cl_2 (1.75 mL), thioacetic acid **7** (558 μL , 7.85 mmol) was added under nitrogen atmosphere. TMSOTf (473 μL , 2.62 mmol) was added dropwise at 0 °C and the resulting mixture was stirred at room temperature overnight. After 15 h, TLC showed complete conversion of the starting material. The reaction mixture was diluted with CH_2Cl_2 and washed with satd aq NaHCO_3 and H_2O . The organic phase was dried over anhydrous Na_2SO_4 , filtered and concentrated in vacuo. The crude material was purified by flash chromatography (2:8 CH_2Cl_2 : i Pr $_2\text{O}$) to afford **4** (654 mg, 1.61 mmol) in 61% yield. TLC: $R_f = 0.35$ (2:8 CH_2Cl_2 : i Pr $_2\text{O}$); ^1H NMR (400 MHz, CDCl_3): $\delta = 5.93$ (d, $J = 1.6$ Hz, 1H), 5.33 (dd, $J = 10$ Hz, 1H), 5.31 (dd, $J = 3.2$ Hz, $J = 1.6$ Hz, 1H), 5.08 (dd, $J = 10$ Hz, $J = 3.2$ Hz, 1H), 4.26 (dd, $J = 12.4$ Hz, 4.8 Hz, 1H), 4.05 (dd, $J = 12.4$ Hz, 2.4 Hz, 1H), 3.91 (ddd, $J = 10$ Hz, 4.8 Hz, 2.4 Hz, 1H), 2.41 (s, 3H), 2.16, 2.06, 2.02, 1.97 (s, 3H each); ^{13}C NMR (100 MHz, CDCl_3): $\delta = 190.3$, 170.5, 169.8, 169.7, 169.4, 80.1, 72.4, 70.9, 69.8, 65.6, 62.1, 31.2, 20.8, 20.6, 20.6, 20.5.

4.3.2. Compound **5**

Diethylamine (75 μL , 0.72 mmol) was added to a solution of 2,3,4,6-tetra-*O*-acetyl-1-*S*-acetyl- α -*D*-mannopyranose **4** (200 mg, 0.49 mmol) and epoxide **3** (81 mg, 0.38 mmol) in dry DMF (630 μL). The reaction mixture was stirred for 4 h, then diluted with EtOAc, washed with 1 M HCl and H_2O . The organic phase was dried over anhydrous Na_2SO_4 , filtered and concentrated in vacuo. The crude material was purified by flash chromatography (4:6 Hex:EtOAc) to afford **5** (149 mg, 0.26 mmol) in 68% yield. TLC: $R_f = 0.3$ (4:6 Hex:EtOAc); $[\alpha]_D^{25}$ (CHCl_3 , c 1.03): +68; ^1H NMR (400 MHz, CDCl_3): $\delta = 5.37$ (dd, $J = 3.2$ Hz, 1.5 Hz, 1H), 5.33 (d, $J = 1.3$ Hz, 1H), 5.26 (t, $J = 9.9$ Hz, 1H), 5.19 (dd, $J = 9.9$ Hz, 3.2 Hz, 1H), 4.40–4.34 (m, 1H), 4.24 (dd, $J = 12$ Hz, 6.4 Hz, 1H), 4.11 (dd, $J = 12.3$ Hz, 2.6 Hz, 1H), 3.86–3.78 (m, 1H), 3.71 (s, 3H), 3.69 (s, 3H), 3.22–3.14 (m, 1H), 3.10–2.98 (m, 2H), 2.2–2.11 (m, 4H), 2.08 (s, 3H), 2.05 (s, 3H), 1.99 (s, 3H), 1.93–1.76 (m, 2H); ^1H NMR (400 MHz, CD_3OD): $\delta = 5.45$ (bs, 1H); 5.37 (dd, $J = 3.2$ Hz, 1.5 Hz, 1H), 5.24 (t, $J = 9.9$ Hz, 1H), 5.19 (dd, $J = 9.9$ Hz, 3.2 Hz, 1H), 4.38–4.32 (m, 1H), 4.25 (dd, $J = 12$ Hz, 6.4 Hz, 1H), 4.12 (dd, $J = 12$ Hz,

2.6 Hz, 1H), 4.0 (q, $J = 3.3$ Hz, 1H), 3.68 (s, 6H), 3.20 (q, $J = 3.6$ Hz, 1H), 3.05 (dt, $J = 11$ Hz, 5.5 Hz, 1H), 2.92 (dt, $J = 11$ Hz, 3.8 Hz, 1H), 2.22–2.32 (m, 1H), 2.15 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H), 2.06–2.00 (m, 1H) 1.97 (s, 3H), 1.97–1.88 (m, 2H); ^{13}C NMR (100 MHz, CDCl_3): $\delta = 174.1$, 173.8, 170.6, 169.9, 169.7, 169.6, 82.6, 71.1, 69.5, 69.3, 69.1, 66.3, 62.5, 52.2, 52.1, 48.7, 40.3, 39.6, 31.3, 28.6, 20.8, 20.6, 20.6, 20.5; MS (ESI) calcd for $\text{C}_{24}\text{H}_{34}\text{O}_{14}\text{S}$ [$\text{M} + \text{Na}$] $^+$ m/z : 601.16, found 601.42; HR-MS (ESI) calcd for $\text{C}_{24}\text{H}_{34}\text{O}_{14}\text{S}$ [$\text{M} + \text{Na}$] $^+$ m/z : 601.15615, found 601.15571.

4.3.3. Compound **10**

To a solution of compound **5** (40 mg, 0.069 mmol) in dry pyridine (150 μL), DMAP (8 mg, 0.069 mmol) and pivaloyl chloride (43 μL , 0.35 mmol) were added. The mixture was stirred at room temperature under nitrogen atmosphere. After 5 h, a second aliquot of pivaloyl chloride (25 μL , 0.21 mmol) was added. After reaction completion (24 h, monitored by TLC, 6:4 Hex:EtOAc) the mixture was diluted with EtOAc, washed with 1 M HCl, H_2O and satd aq NaHCO_3 . The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. Flash chromatography (6:4 Hex:EtOAc) yielded product **10** (27 mg) in 59% yield. TLC: R_f 0.3 (6:4 Hex:EtOAc); $[\alpha]_D^{25}$ (CHCl_3 , c 1.61): +75; ^1H NMR (400 MHz, CDCl_3): $\delta = 5.38$ (d, $J = 1.5$ Hz, 1H), 5.35 (dd, $J = 3.2$ Hz, 1.6 Hz, 1H), 5.29 (t, $J = 9.7$ Hz, 1H), 5.18 (dd, $J = 10.0$ Hz, 3.3 Hz, 1H), 5.08–5.04 (m, 1H), 4.36–4.27 (m, 2H), 4.16–4.07 (m, 1H), 3.7 (s, 3H), 3.69 (s, 3H), 3.34–3.29 (m, 1H), 3.01–2.87 (m, 2H), 2.15 (s, 3H), 2.13–2.00 (m, 10H), 1.99 (s, 3H) 1.2 (s, 9H); ^{13}C NMR (100 MHz, CDCl_3): $\delta = 177.0$, 174.2, 174.1, 170.5, 169.8, 169.7, 169.5, 82.9, 70.8, 69.7, 69.6, 69.2, 65.9, 62.3, 52.0, 52.0, 43.6, 39.6, 39.5, 38.8, 28.7, 27.9, 27.0, 20.8, 20.6, 20.5, 20.5; MS (ESI) calcd for $\text{C}_{29}\text{H}_{42}\text{O}_{15}\text{S}$ [$\text{M} + \text{Na}$] $^+$ m/z : 685.21, found 685.29.

4.3.4. 3-Azido-2,2-dimethylpropanoyl chloride **12**

2,2-Dimethyl-3-chloro-propanoic acid (400 mg, 2.97 mmol) was dissolved in dry DMF (14.6 mL) and NaN_3 (990 mg, 15.2 mmol) was added. The mixture was stirred at 50 °C under nitrogen atmosphere. After 3 days, the reaction was allowed to reach room temperature. The mixture was diluted with CH_2Cl_2 and filtered over a celite pad to remove unreacted NaN_3 . The solvent was evaporated at reduced pressure giving 3-azido-2,2-dimethylpropanoic acid in quantitative yield (425 mg). The analytical data were fully consistent with those reported in the literature³² (^1H NMR (400 MHz, CDCl_3): $\delta = 3.42$ (s, 2H), 1.25 (s, 6H)) and the compound was used without further purification.

The crude azide (143 mg, 1.00 mmol) was dissolved in dry CH_2Cl_2 (3.3 mL) and oxalyl chloride (254 μL , 2.00 mmol) was added drop wise to the solution. After the addition of one drop of dry DMF, the mixture was refluxed for 3 h. The solution was cooled to room temperature and the solvent was carefully removed in vacuo to yield 3-azido-2,2-dimethylpropanoyl chloride **12**. The complete conversion was confirmed by NMR analysis and the product was used in the next step without further purification. ^1H NMR (400 MHz, CDCl_3): $\delta = 3.53$ (s, 2H), 1.35 (s, 6H).

4.3.5. Compound **13**

To a solution of compound **5** (30 mg, 0.052 mmol) in dry CH_2Cl_2 (400 μL) and Et $_3\text{N}$ (16 μL , 0.117 mmol), DMAP (2 mg, 0.015 mmol) was added. The solution was cooled to 0 °C and compound **12** (19 mg, 0.116 mmol) was added. The mixture was stirred at room temperature under nitrogen atmosphere. After reaction completion (3 h, monitored by TLC, 1:1 Hex:EtOAc) the mixture was diluted with CH_2Cl_2 , washed with satd aq NaHCO_3 and 1 M HCl. The organic layer was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. Flash chromatography (6:4 Hex:EtOAc) yielded product **13** (29 mg) in 79% yield. TLC: R_f 0.5 (1:1 Hex:EtOAc); $[\alpha]_D^{25}$ (CHCl_3 , c 1.08): +67; ^1H NMR (400 MHz, CDCl_3):

$\delta = 5.39$ (m, 1H), 5.37–5.34 (m, 1H), 5.30 (t, $J = 9.5$ Hz, 1H), 5.19 (dd, $J = 9.7$ Hz, 3.6 Hz, 1H), 5.14–5.11 (m, 1H), 4.37–4.30 (m, 2H), 4.15–4.07 (m, 1H), 3.71 (s, 3H), 3.71 (s, 3H), 3.42 (s, 2H), 3.36–3.33 (m, 1H), 3.03–2.89 (m, 2H), 2.17 (s, 3H), 2.14–2.06 (m, 4H), 2.18 (s, 3H), 2.05 (s, 3H), 2.00 (s, 3H), 1.23 (s, 3H), 1.23 (s, 3H); ^{13}C NMR (100 MHz, CDCl_3): $\delta = 174.6, 174.5, 174.3, 170.8, 170.2, 170.0, 169.9, 83.3, 71.1, 70.9, 70.0, 69.6, 66.3, 62.6, 59.9, 52.4, 44.2, 43.8, 40.0, 39.9, 29.1, 28.2, 23.3, 23.3, 21.1, 20.9, 20.9, 20.9$; MS (ESI) calcd for $\text{C}_{29}\text{H}_{41}\text{N}_3\text{O}_{15}\text{S} [\text{M}+\text{Na}]^+ m/z$: 726.2; found 726.8.

4.3.6. General procedure for the Zemplén deacetylation

A MeONa/MeOH solution was added to the substrate dissolved in dry MeOH under nitrogen atmosphere (substrate and MeONa final concentration 0.02 M) and the reaction mixture was stirred at room temperature. After completion, (approx. 10 min, monitored by TLC) the reaction was quenched with AMBERLITE IR-120 H^+ till neutral pH, the resin was filtered off and the solvent was evaporated at reduced pressure, obtaining the pure product.

4.3.6.1. *Pseudo-thio-1,2-dimannoside 2 (R = H)*. TLC: R_f 0.4 (8:2 CH_2Cl_2 :MeOH); $[\alpha]_D^{19}$ (MeOH, c 0.51): +155; ^1H NMR (400 MHz, D_2O): 5.38 (d, $J = 1.5$ Hz, 1H), 4.12–4.07 (m, 1H), 4.06 (dd, $J = 1.5$ Hz, 3.2 Hz, 1H), 3.95 (ddd, $J = 1.9$ Hz, 6.2 Hz, 9.0 Hz, 1H), 3.85 (dd, $J = 2.2$ Hz, 12.4 Hz, 1H), 3.77–3.71 (m, 2H), 3.70 (s, 3H), 3.69 (s, 3H), 3.64 (dd, $J = 9.8$ Hz, 9.5 Hz, 1H), 3.25–3.19 (m, 1H), 3.06–2.95 (m, 2H), 2.27–2.16 (m, 1H), 2.11–2.03 (m, 1H), 2.00–1.92 (m, 2H); ^{13}C NMR (100 MHz, D_2O): $\delta = 177.3, 177.0, 84.9, 73.4, 71.7, 71.0, 67.7, 67.0, 60.8, 52.6, 45.4, 40.0, 39.0, 29.9, 27.5$; HR-MS (ESI) calcd for $\text{C}_{16}\text{H}_{26}\text{O}_{10}\text{S} [\text{M}+\text{Na}]^+ m/z$: 433.11389; found 433.11346.

4.3.6.2. *Compound 11*. TLC: R_f 0.24 (9:1 CH_2Cl_2 :MeOH); $[\alpha]_D^{20}$ (MeOH, c 1.01): +128; ^1H NMR (400 MHz, D_2O): 5.41 (d, $J = 1.1$ Hz, 1H), 5.11–5.08 (m, 1H), 4.08 (dd, $J = 1.5$ Hz, 3.3 Hz, 1H), 3.95 (ddd, $J = 2.4$ Hz, 5.8 Hz, 9.4 Hz, 1H), 3.85 (dd, $J = 2.4$ Hz, 12.5 Hz, 1H), 3.79–3.72 (m, 2H), 3.71 (s, 3H), 3.70 (s, 3H), 3.69–3.63 (m, 1H), 3.42–3.38 (m, 1H), 3.05–2.92 (m, 2H), 2.23–2.01 (m, 4H), 1.21 (s, 9H); ^{13}C NMR (100 MHz, D_2O): $\delta = 180.3, 176.8, 176.8, 85.3, 73.6, 71.7, 71.0, 70.6, 66.8, 60.6, 52.7, 52.7, 42.8, 39.9, 39.6, 38.7, 28.4, 27.3, 26.2$; HR-MS (ESI) calcd for $\text{C}_{21}\text{H}_{34}\text{O}_{11}\text{S} [\text{M}+\text{Na}]^+ m/z$: 517.17140; found 517.17166.

4.3.6.3. *Compound 14*. TLC: R_f 0.37 (9:1 CH_2Cl_2 :MeOH); $[\alpha]_D^{20}$ (MeOH, c 0.81): +112; ^1H NMR (400 MHz, MeOD): $\delta = 5.40$ –5.36 (m, 1H), 5.17–5.12 (m, 1H), 3.95 (dd, $J = 1.6$ Hz, 3.2 Hz, 1H), 3.88–3.81 (m, 2H), 3.75 (dd, $J = 6.0$ Hz, 12.2 Hz, 1H), 3.72–3.68 (m, 1H), 3.69 (s, 3H), 3.67 (s, 3H), 3.62 (dd, $J = 3.2$ Hz, 9.2 Hz, 1H), 3.50 (s, 2H), 3.43–3.38 (m, 1H), 3.00–2.89 (m, 2H), 2.19–2.10 (m, 2H), 2.08–2.02 (m, 2H), 1.24 (s, 6H). ^{13}C NMR (100 MHz, MeOD): $\delta = 176.1, 176.0, 176.0, 87.1, 75.7, 73.6, 73.1, 72.4, 68.5, 62.5, 60.8, 52.6, 52.6, 45.0, 44.1, 41.1, 40.9, 30.0, 29.0, 23.3, 23.3$; HR-MS (ESI) calcd for $\text{C}_{21}\text{H}_{33}\text{N}_3\text{O}_{11}\text{S} [\text{M}+\text{Na}]^+ m/z$: 558.17280; found 558.17371.

4.4. Surface plasmon resonance (SPR) analysis

The extracellular domain (ECD) of DC-SIGN (residues 66–404) was overexpressed and purified as previously described.³³ The SPR experiments were performed on a BIACore T200 using a CM3 sensor chip. Flow cells were activated as described.³⁴ Flow cell one was functionalised with BSA and blocked with ethanolamine and subsequently used as a control surface. Flow cells 2 and 3 were

treated with BSA-Man α 1-3[Man α 1-6]Man (Dextra) (60 $\mu\text{g}/\text{ml}$) in 10 mM NaOAc pH 4 to reach different binding densities and blocked with ethanolamine. The final densities on flow cells 2 and 3 were 2579 and 2923 RU, respectively. The affinity of the various compounds for DC-SIGN ECD were evaluated via an established inhibition assay³⁵ in which DC-SIGN ECD was injected at 20 μM alone or in the presence of increasing concentration of inhibitors (ranging from 0 to 5 mM). Injections were performed at 5 $\mu\text{L}/\text{min}$ using 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl_2 , 0.05% P20 surfactant as running buffer. The surface was regenerated by the injection of 50 mM EDTA.

Acknowledgments

For DC-SIGN ECD production, this work used the Multistep Protein Purification Platform (MP3) and the SPR platform for the competition test of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology. We are grateful to EU ITN Marie-Curie program (IMMUNOSHAPÉ – Grant no. 642870) for funding Silvia Achilli. Support by the Italian Ministry of Research through a PRIN grant (prot. 2015RNWJAM_002) is acknowledged.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmc.2017.03.046>.

References

- Varki A. *Glycobiology*. 2017;27:3.
- Bertozzi CR, Kiessling LL. *Science*. 2001;291:2357.
- Hudak JE, Bertozzi CR. *Chem Biol*. 2014;21:16.
- Wang LX, Davis BG. *Chem Sci*. 2013;4:3381.
- von Itzstein M. *Nat Rev Drug Discovery*. 2007;6:967.
- Ernst B, Magnani JL. *Nat Rev Drug Discovery*. 2009;8:661.
- Bernardi A, Cheshev P. *Chem Eur J*. 2008;14:7434.
- Cecioni S, Imberty A, Vidal S. *Chem Rev*. 2015;115:525.
- Sattin S, Bernardi A. *Trends Biotechnol*. 2016;34:483.
- Werz D, Koester D, Holkenbrink A. *Synthesis*. 2010;32:17.
- Driguez H. In: Driguez H, Thiem H, eds. *Glycoscience Synthesis of Substrate Analogs and Mimetics*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1998:85.
- Driguez H. *ChemBioChem*. 2001;2:311.
- Strino F, Lii JH, Gabius HJ, Nyholm PG. *J Comput Aided Mol Des*. 2009;23:845.
- Aguilera B, Jiménez-Barbero J, Fernández-Mayoralas A. *Carbohydr Res*. 1998;308:19.
- Mari S, Posteri H, Marcou G, et al. *Eur J Org Chem*. 2004;5119.
- Reina JJ, Sattin S, Invernizzi D, et al. *ChemMedChem*. 2007;2:1030.
- van Kooyk Y, Geijtenbeek TB. *Nat Rev Immunol*. 2003;3:697.
- Varga N, Sutkevičiute I, Guzzi C, et al. *Chem Eur J*. 2013;19:4786.
- Varga N, Sutkevičiute I, Ribeiro-Viana R, et al. *Biomaterials*. 2014;35:4175.
- Berzi A, Varga N, Sattin S, et al. *Viruses*. 2014;6:391.
- Ordanini S, Varga N, Porkolab V, et al. *Chem Commun*. 2015;51:3816.
- Pachamuthu K, Schmidt RR. *Chem Rev*. 2006;106:160.
- Shu P, Zeng J, Tao J, Zhao Y, Yao G, Wan Q. *Green Chem*. 2015;17:2545.
- Belz T, Williams SJ. *Carbohydr Res*. 2016;429:38.
- Brazdova B, Tan NS, Samoshina NM, Samoshin VV. *Carbohydr Res*. 2009;344:311.
- Sammert B, Bogner T, Nahrwold M, Weiss C, Sewald N. *J Org Chem*. 2010;75:6953.
- Thépaut M, Guzzi C, Sutkevičiute I, et al. *J Am Chem Soc*. 2013;135:2518.
- Bordoni V, Porkolab V, Sattin S, et al. *RSC Adv*. 2016;6:89578.
- Lázár L, Csávás M, Herczeg M, Herczegh P, Borbás A. *Org Lett*. 2012;14:4650.
- Csávás M, Demeter T, Herczeg M, et al. *Tetrahedron Lett*. 2014;55:6983.
- Manzano VE, Uhrig ML, Varela O. *J Org Chem*. 2008;73:7224.
- Nahrwold M, Bogner T, Eissler S, Verma S, Sewald N. *Org Lett*. 2010;12:1064.
- Tabarani G, Thépaut M, Stroebel D, et al. *J Biol Chem*. 2009;284:21229.
- Halary F, Amara A, Lortat-Jacob H, et al. *Immunity*. 2002;17:653.
- Andreini M, Doknic D, Sutkevičiute I, et al. *Org Biomol Chem*. 2011;9:5778.

10.1.2 Additional data non-presented in the article: enzymatic assay

The stability towards the hydrolysis by the α -Mannosidase enzyme from *Canavalia ensiformis* (jack-bean) was evaluated by $^1\text{H-NMR}$ spectroscopy for psDi (Fig.112a) and compound 2 (Fig.112b). The assays were performed by Alice Tamburrini (laboratory of Anna Bennardi).

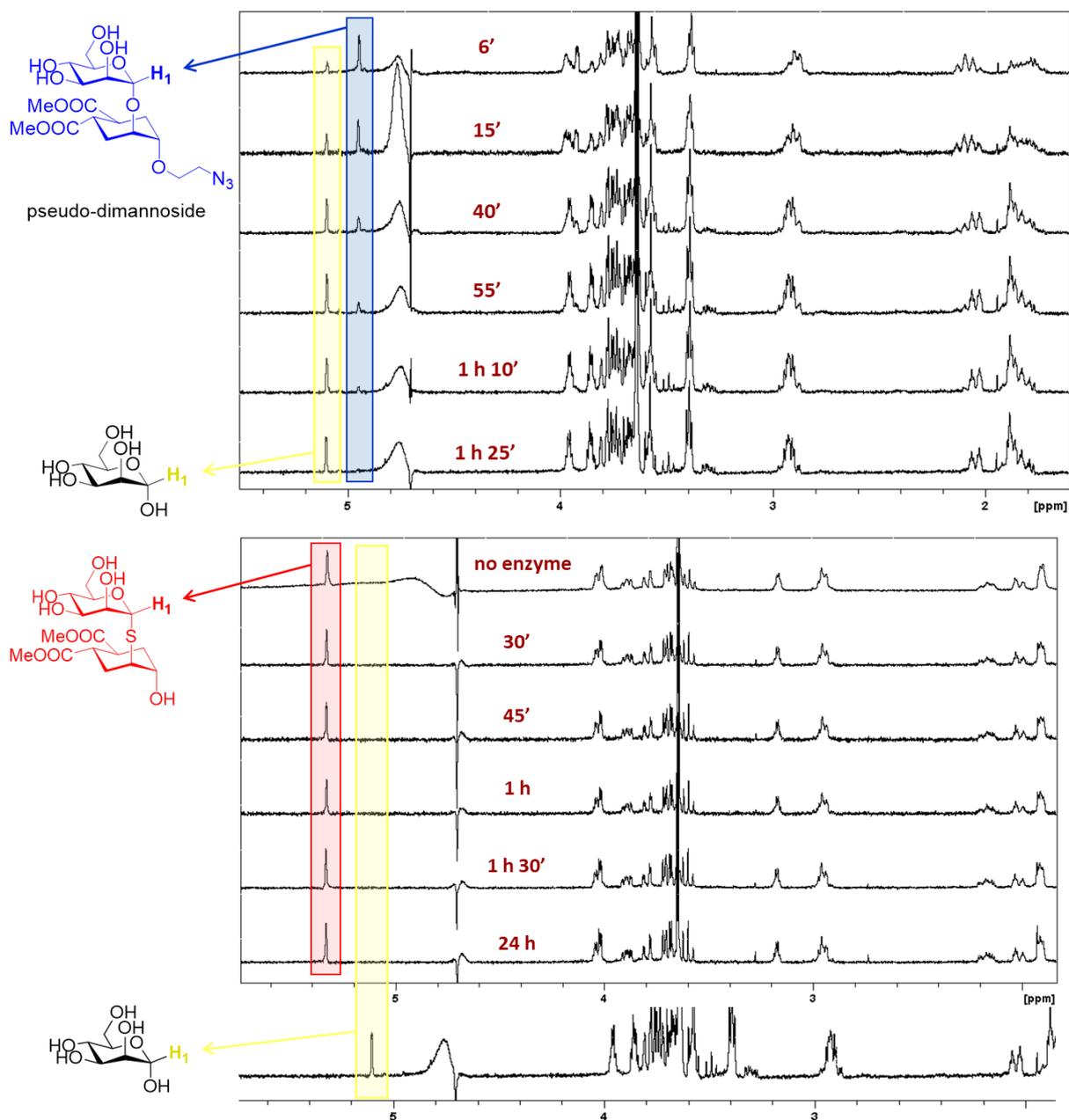


Fig.112 thio-psDi enzymatic reaction. α -Mannosidase reaction on psDi and thio-psdi

The psDi was hydrolysed after 1h30 of incubation with the α -Mannosidase, while for the thio-derivative no hydrolysis was observed even after one day of incubation. These preliminary assays proved the *in vitro* resistance towards glycosidases hydrolysis.

10.1.3 C and Si-Glycosides

With a purpose similar to the synthesis of thio-glycomimetics, C-glycosides (sugars with the anomeric oxygen replaced by a carbon) were constructed to enhance the glycomimetic resistance towards glycosidase. Those compounds were assayed by SPR for their capability to inhibit DC-SIGN ECD binding to ManBSA. The design and the synthesis of these C-glycosides was conducted in the group of Prof. J. Moravcova. In addition, a Si-linkage was also designed and tested by SPR (Fig.113).

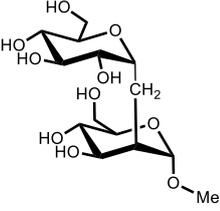
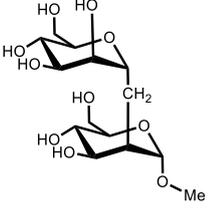
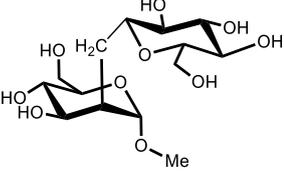
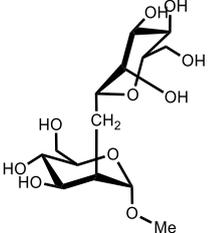
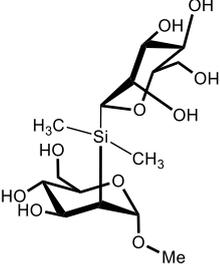
<p>ID-246-4 α-D-Glc-C-(1\rightarrow2)-D-Man</p>	
<p>ID-246-5 α-D-Man-C-(1\rightarrow2)-D-Man</p>	
<p>ID-246-11 β-D-Glc-C-(1\rightarrow2)-D-Man</p>	
<p>ID-246-12 β-D-Man-C-(1\rightarrow2)-D-Man</p>	
<p>JCH-423 β-D-Man-Si-(1\rightarrow1)-β-D-Man</p>	

Fig.113 List of C-glycosides and Si-glycoside structures.

The SPR competition assays revealed that the C-glycosides showed an inhibitory power comparable to the ones of the monosaccharide (D-mannose \approx 2.5 mM) and of Man α 1-2Man (\approx 1 mM), with a single positive exception for **ID-246-11** that presented an IC₅₀ of 885.2 \pm 1.1 μ M (Fig.114).

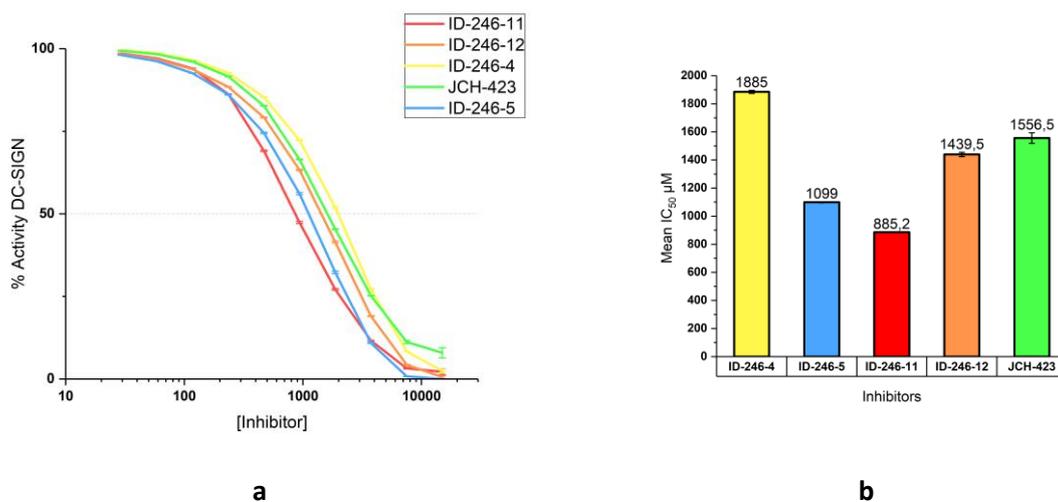


Fig.114 ID-246-4 (yellow), ID-246-5 (blues) ID-246-11 (red), ID-246-12 (orange) and JCH-423 (green) competition assay against DC-SIGN a) inhibition curves and b) IC₅₀ obtained for DC-SIGN inhibition assays

The Si- glycoside solubilisation required the addition of 4% DMSO. The inhibitory power of this compound was comparable to the one of the natural disaccharide and monosaccharide.

Globally, depending on the disaccharide used, we obtained different level of inhibition. **ID246-5** which is a C- Man α 1-2Man, maintains the same affinity observed for the O-linked one.

10.2 Development of glycomimetics selective towards DC-SIGN: the “ammonium binding pocket” strategy

The group of Sonsoles Martin-Santamaria docked (Glide) thousands of drug-like fragments (Maybridge library) on the DC-SIGN-**psDi** X-ray structure, maintaining **psDi** in its crystallographic pose. Virtual fragment screening identified several interesting moieties that could be used to modify the structure of **psDi** and would create favourable interactions with DC-SIGN binding region. Among them, they focused on the structure of fragments that interact favourably with the Phe313 region of the protein. This region, which is involved in the binding of several natural oligomannosides [63], [164] is close to the mannose 2-hydroxyl group in the DC-SIGN / **psDi** complex and could, in principle, be reached by **psDi** structures modified on that position, which is easily chemically accessible. Fig.115 shows the best candidates identified.

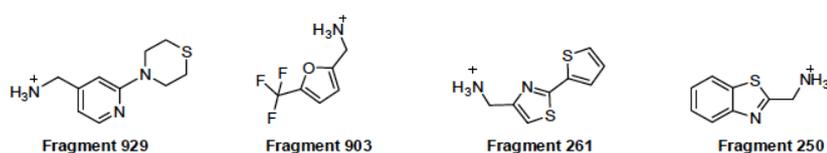


Fig.115 Best candidates from computational screening.

These candidates were characterized by the presence of an ammonium ion and an aromatic ring and they all interacted with an “ammonium binding pocket” delimited by Ser360, Phe313 and Glu358 in close proximity to the Calcium binding site. They enable cation π interactions with Phe313, hydrogen bond with Ser360 and ionic interaction with Glu358 (Fig.116). In addition to that, the aromatic ring could form π - π or cation π interactions with other residues in that area.

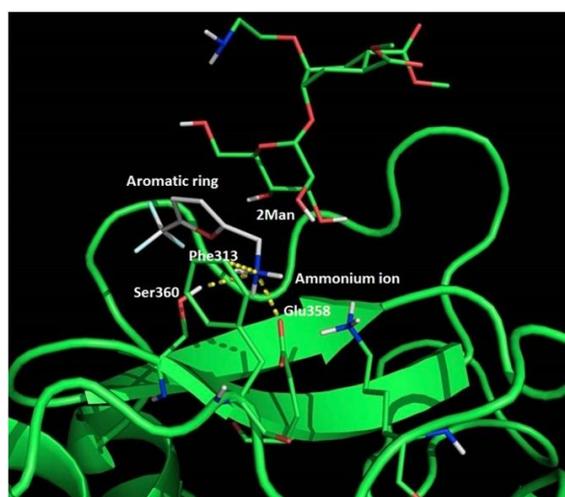


Fig.116 The “ammonium binding pocket”. Ser360, Phe313 and Glu358 interactions with ammonium ion.

On a synthetic point of view two combined strategies were investigated by the group of Anna Bernardi to reach the ammonium pocket:

- an amide bond formation with α -aminoacid derivatives (**1**)
- a Copper(I)-catalyzed Alkyne Azide Cycloaddition (CuAAC) reaction with α -aminoalkyne derivatives (**2**)

In both strategies, the 2-hydroxy group of mannose is replaced with a nitrogen atom, either an amine (**1**) or an azido group (**2**). For that purpose, **psDi** (target) was first modified into an intermediate (fig.117).

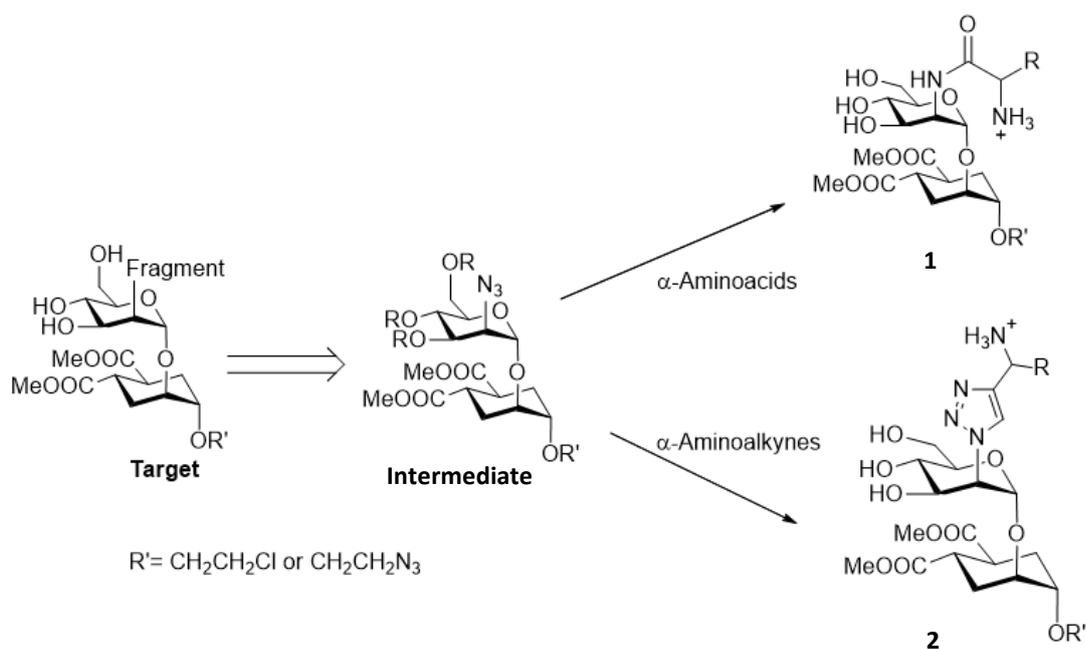


Fig.117 Synthetic strategies. α -aminoacid derivative and α -aminoacid derivative strategies.

10.2.1 Ammonium binding pocket: amino derivatives

The following figure 118 indicates the general structure of the amino derivative glycomimetics derived from psDi.

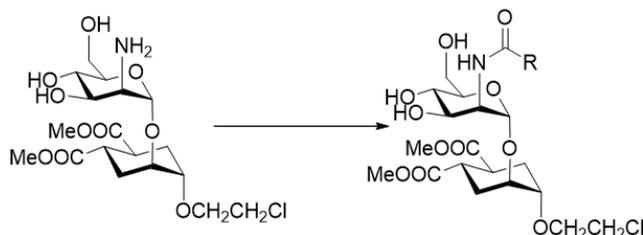


Fig.118 Synthetic strategy and overall structure of amino derivatives

Besides, computational studies have revealed that coupling an ammonium group to an aromatic ring should enhance targeting of DC-SIGN secondary binding sites. Different glycomimetics were then synthesized combining an NH_3^+ group with different terminal aromatic rings (Fig.119).

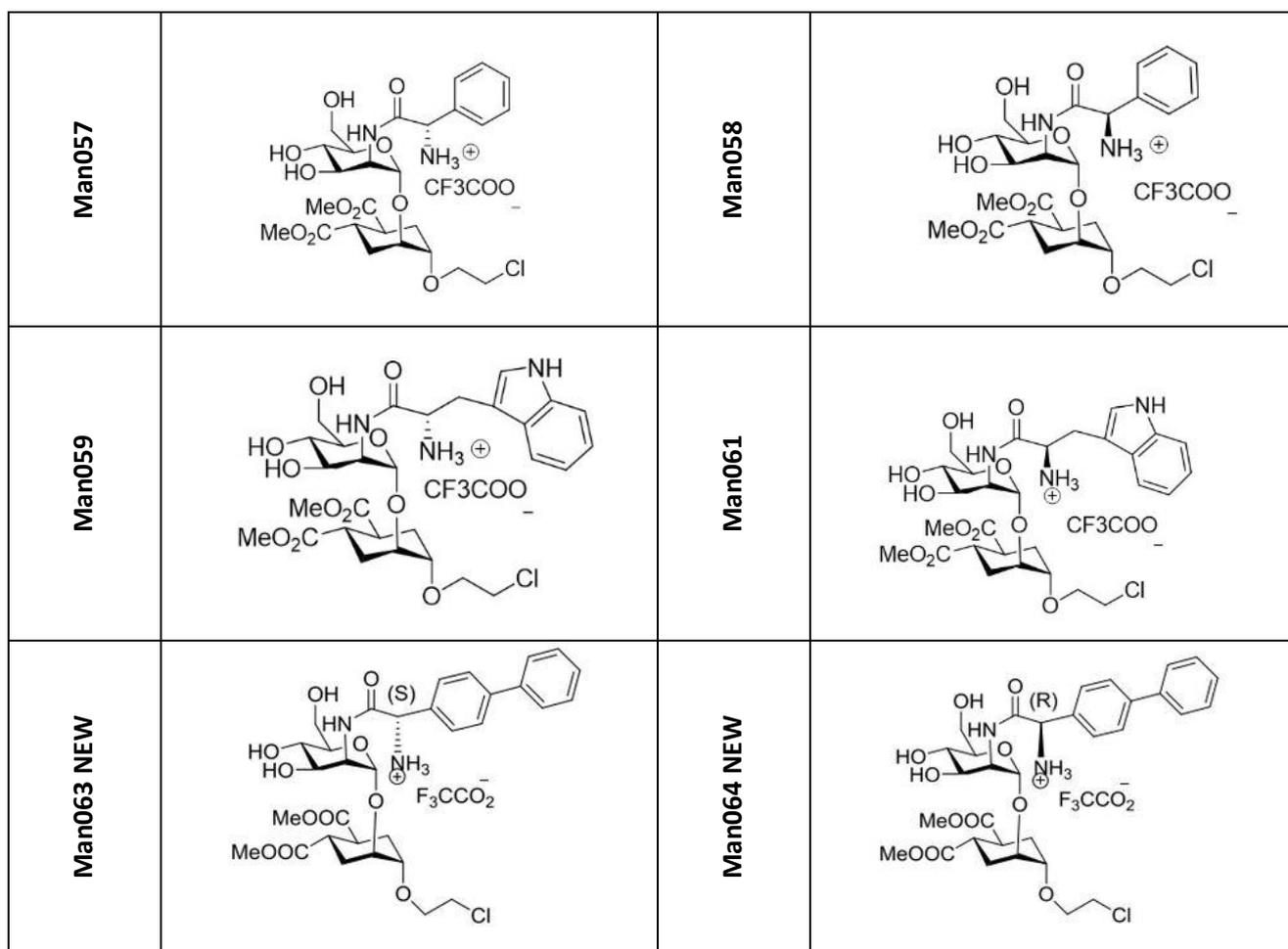


Fig.119 Amino derivatives structures.

The obtained couples of diastereoisomers were tested as potential inhibitors of DC-SIGN interaction with a mannosylated surface. The IC_{50} values obtained from the SPR inhibition assays are gathered in figure 120.

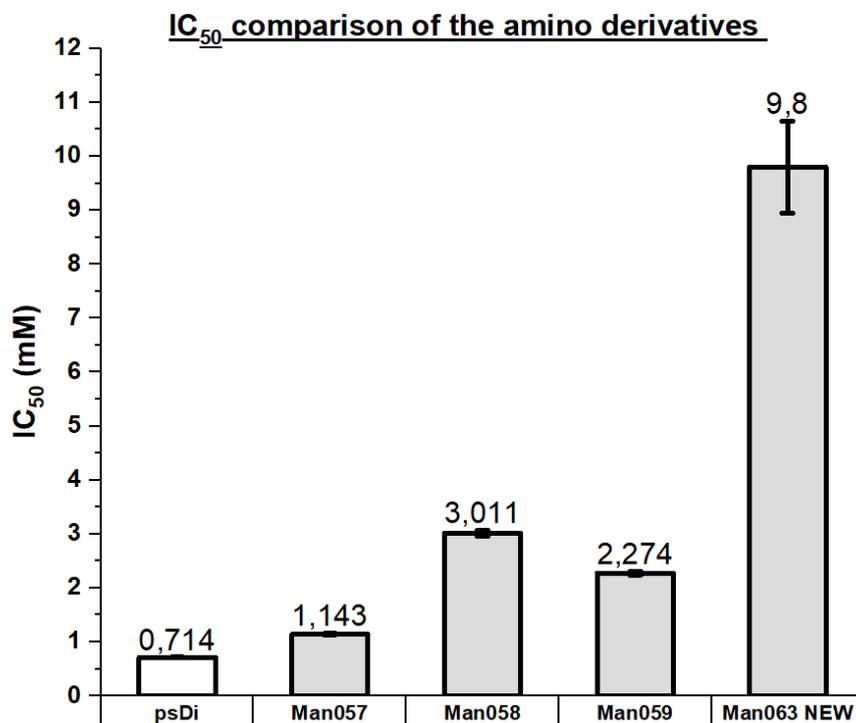


Fig.120 IC_{50} comparison of the amino derivatives (grey colour. The control psDi is in white).

The addition of aromatic rings was not beneficial. Indeed, all the compounds tested bearing an aromatic or heteroaromatic fragment had an inhibition power towards DC-SIGN interaction with the mannosylated surface lower than that of psDi. **Man061** and **Man063NEW** showed an extremely weak interaction with DC-SIGN that prevented any possible IC_{50} calculation.

These results suggest that the distance between the **psDi** scaffold and the ammonium ion pocket is probably not optimally covered by the α -aminoacid framework.

10.2.2 Ammonium binding pocket: triazole derivatives

Better results were observed with the triazole derivatives, functionalized on the 2 hydroxyl of the mannose ring by CuAAC with propargylamine (Fig.121a). The first compound analysed was **Man062** (Fig.121b).

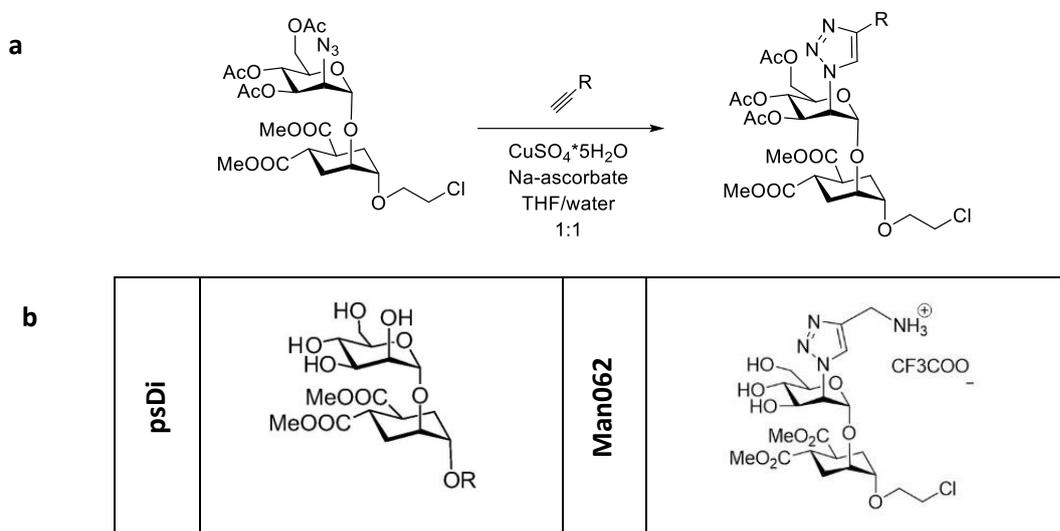


Fig.121 Triazole derivative strategy a) synthetic strategy and overall structure of triazole derivatives. b) psDi and Man062 structures .

Man062 was tested in competition assay against DC-SIGN mannosylated surface interaction and revealed an interesting IC_{50} of $113.09 \pm 11.99 \mu M$, with one order of magnitude improvement over **psDi**. In figure 122 are shown the inhibition curves and the IC_{50} comparison between the two glycomimetics.

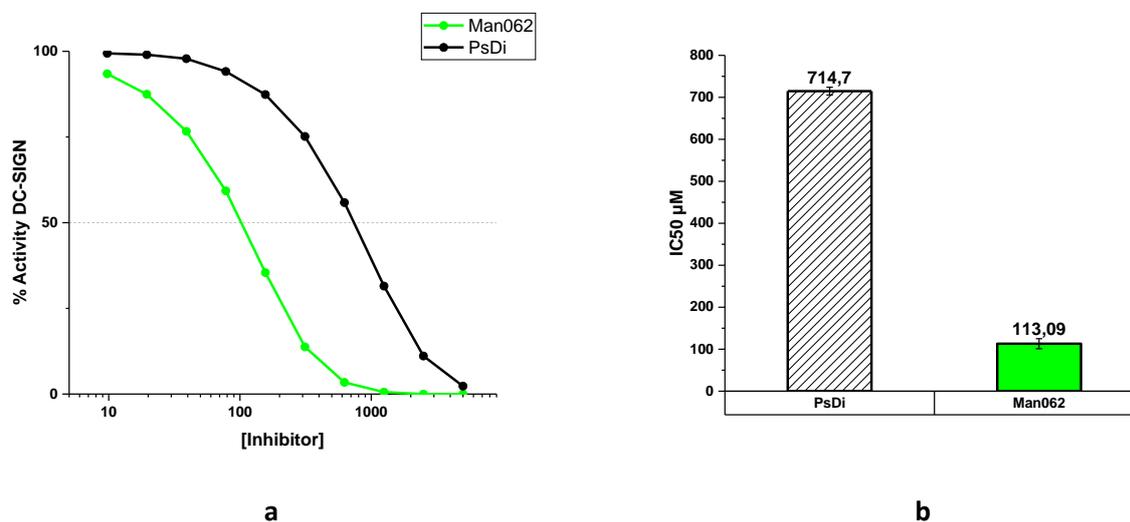


Fig.122 Man062 competition assay a) Man062 (green) and psDi (black) inhibition curves against DC-SIGN interaction with a mannosylated surface b) Man062 and psDi IC_{50} obtained for DC-SIGN inhibition assay.

Two **Man062** derivatives, **Man065**, carrying a pyridine and **Man064**, Man062 corresponding alcohol, were synthesized and tested in competition assay against DC-SIGN report to Fig.123.

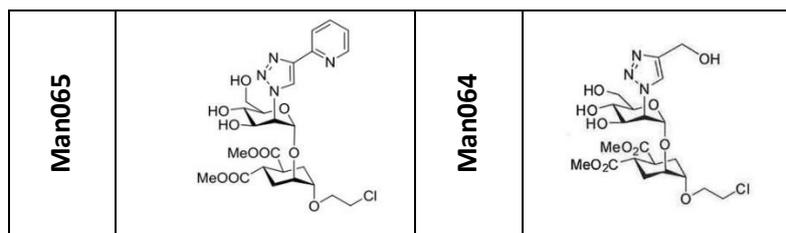


Fig.123 Man065 and Man064 structures.

Man064 provided a 3-fold affinity improvement over **psDi** (IC_{50} of $339 \pm 6.3 \mu M$) and showed that H bonds can be made also with alcohol. The triazole **Man065**, enabled an IC_{50} of $145.4 \pm 1.9 \mu M$, corresponding to a major improvement of **psDi** (Fig.124). This is a surprised finding since the pK_a of the conjugate acid (the pyridinium cation) is 5.25 and at pH used for the assay (pH 8) the pyridine does not have H bond capability. One hypothesis is that the pyridine can interact with the Phe313 but not through π - π interactions.

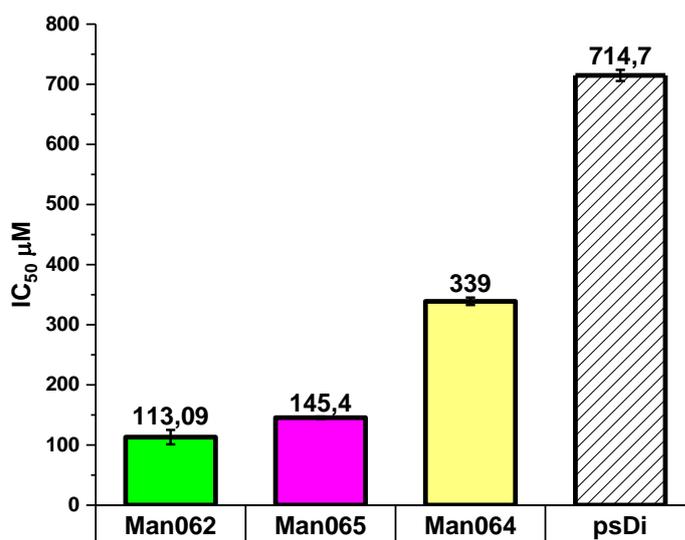


Fig.124 Man062, Man065, Man064 and psDi IC_{50} comparison.

The IC_{50} values obtained from the compounds analysed so far confirm the interactions predicted by the docking studies.

Methyl groups, meant to stabilize the positive charge were added sequentially to the nitrogen of **Man062**. Three new methylated compounds were tested: **Man066**, **Man067** and **Man068** (Fig.125).

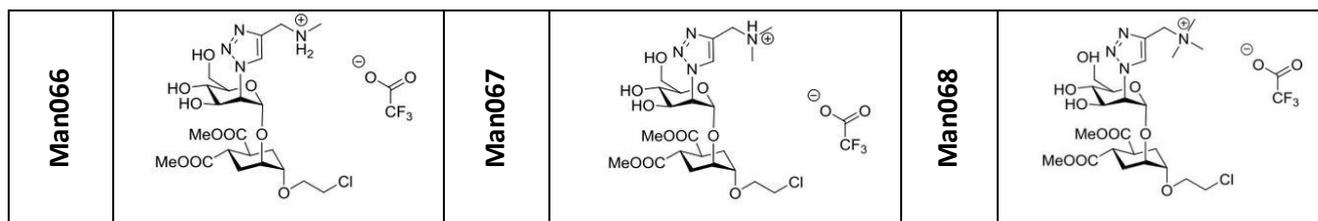


Fig.125 Man066-67-68 structures.

Figure 126 presents the inhibition curves and the IC₅₀ comparison between psDi, **Man062** and the three corresponding methylated glycomimetics. The sequential addition of methyl groups led to a significant decrease of the affinity suggesting that the ammonium ion may act as a H-bonding donor towards Ser360 and Glu358 residue. Additional experiments have to be performed to validate the hypothesis.

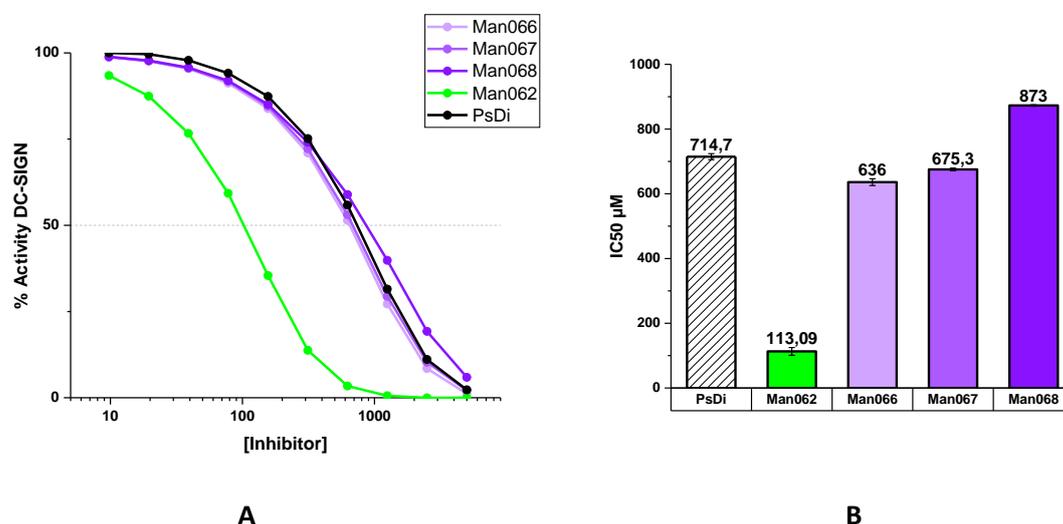


Fig.126 Man062 corresponding methylated glycomimetics inhibition assay. Man062 (green), psDi (black) and Man066 (light violet)-67 (violet)-68 (dark violet) a) inhibition curves against DC-SIGN and b) IC₅₀ histograms.

To sum up, the “ammonium pocket” strategy proved to be a success. Functionalization in C-2 can improved affinity towards DC-SIGN and the best ligands were based on an ammonium group linked through a triazole. So far, the beneficial effects of an aromatic substituent have not been confirmed.

The following figure 127 shows the overall IC₅₀ of **Man062** and its derivatives in competition assays against DC-SIGN interaction with a mannosylated surface, from the more effective to the less. The compound structures are represented below the graph.

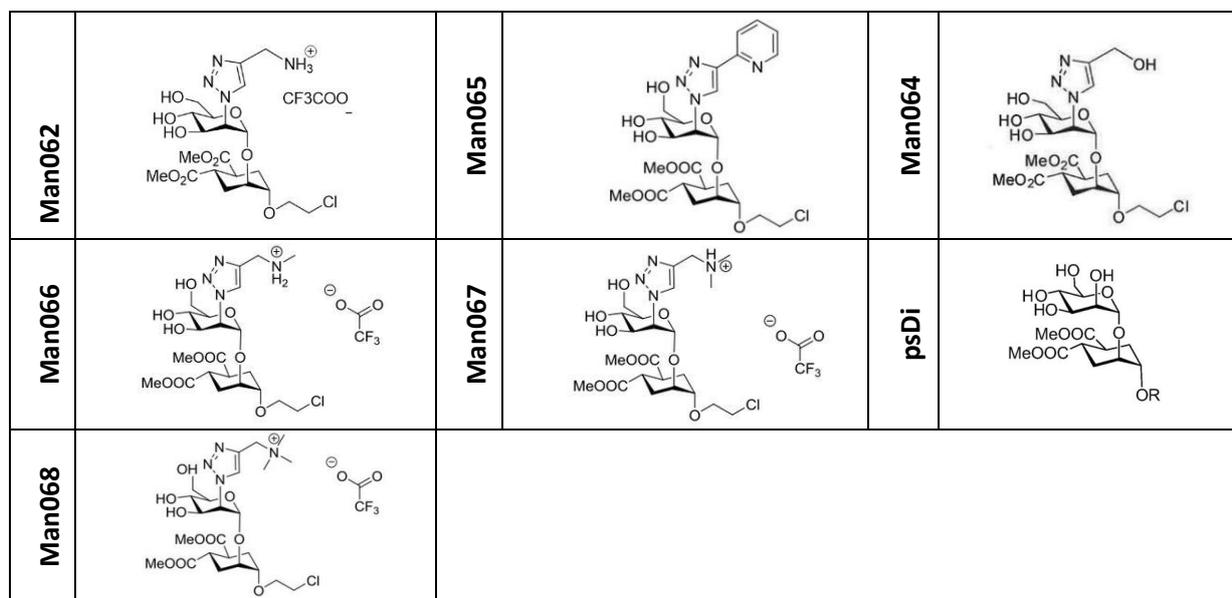
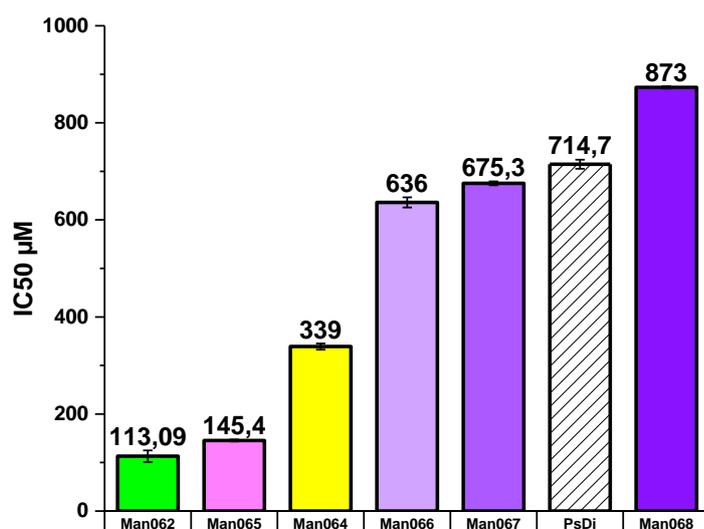


Fig.127 List of triazole derivatives IC₅₀ values and corresponding structures.

Man062 (IC₅₀ of 113.09 ± 11.99 µM) and **Man065** (145.4 ± 1.9 µM) have the highest inhibitory power among all the tested compounds. **Man062** was then chosen as lead compound to be implemented into **Man069**, which combines **Man030** backbone structure (*p*-hydroxymethylenebenzylamide residues) and Man062 functionalization in position C-2. **Man030** [163] presents an enhanced affinity towards DC-SIGN compared to psDi (psDi IC₅₀ = 0.7 mM, Man030 IC₅₀ = 0.3 mM). For this reason, **Man069** was expected to be more effective than **Man062**, which is derived from **psDi**, in the inhibition of DC-SIGN interaction with the mannosylated surface. **Man062-069** and **065** were also tested in competition assay against langerin to define whether they were selective towards DC-SIGN. Fig.128 indicates the structure of **Man069**.

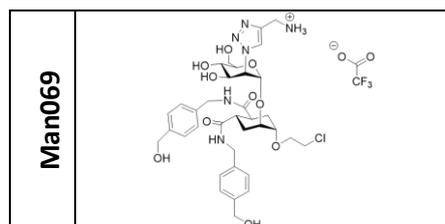


Fig.128 Man069 structure.

As expected, **Man069** showed an increased inhibitory power (IC_{50} $76.25 \pm 3.25 \mu\text{M}$) compared to **Man062** and **Man065** (Fig.129a). Langerin competition assays revealed that all three inhibitors are selective towards DC-SIGN. Nevertheless, **Man062** showed efficacy in inhibiting langerin interaction with the surface. Figure 129b presents a comparison of inhibition curves of the three inhibitors against DC-SIGN (square) and langerin interaction with a mannosylated surface (round).

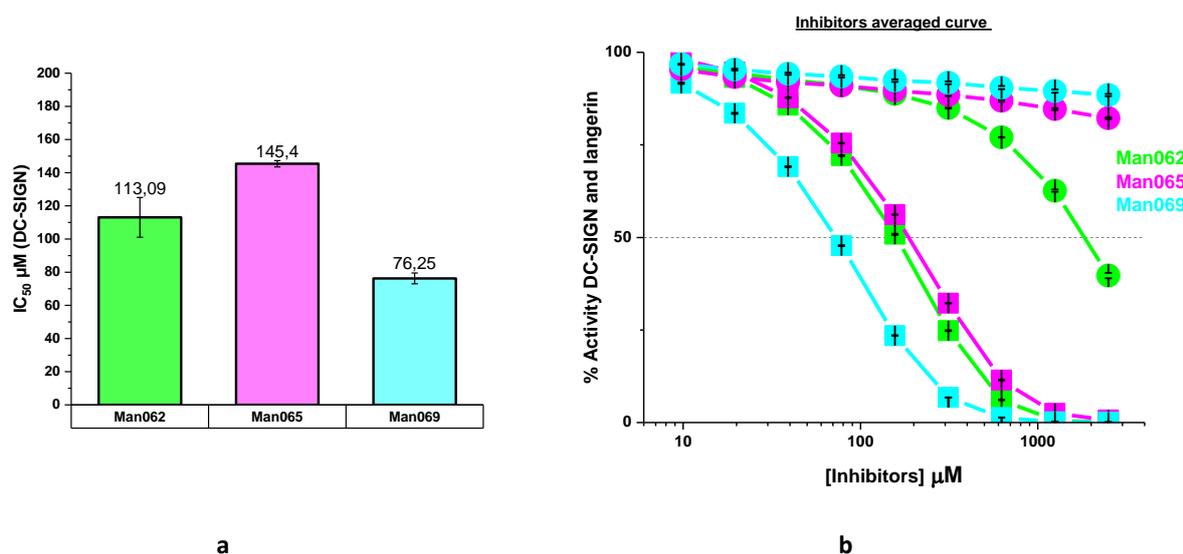


Fig.129 Man062 (green), Man065 (magenta) and Man089 (cyan) competition assay a) IC_{50} obtained for DC-SIGN inhibition assays. b)) inhibition curves of DC-SIGN (square) and langerin interacting with mannosylated surface (round).

In figure 130, examples of sensorgrams obtained for **Man062-65** and **69** competition assays against DC-SIGN and langerin interaction with a mannosylated surface are shown.

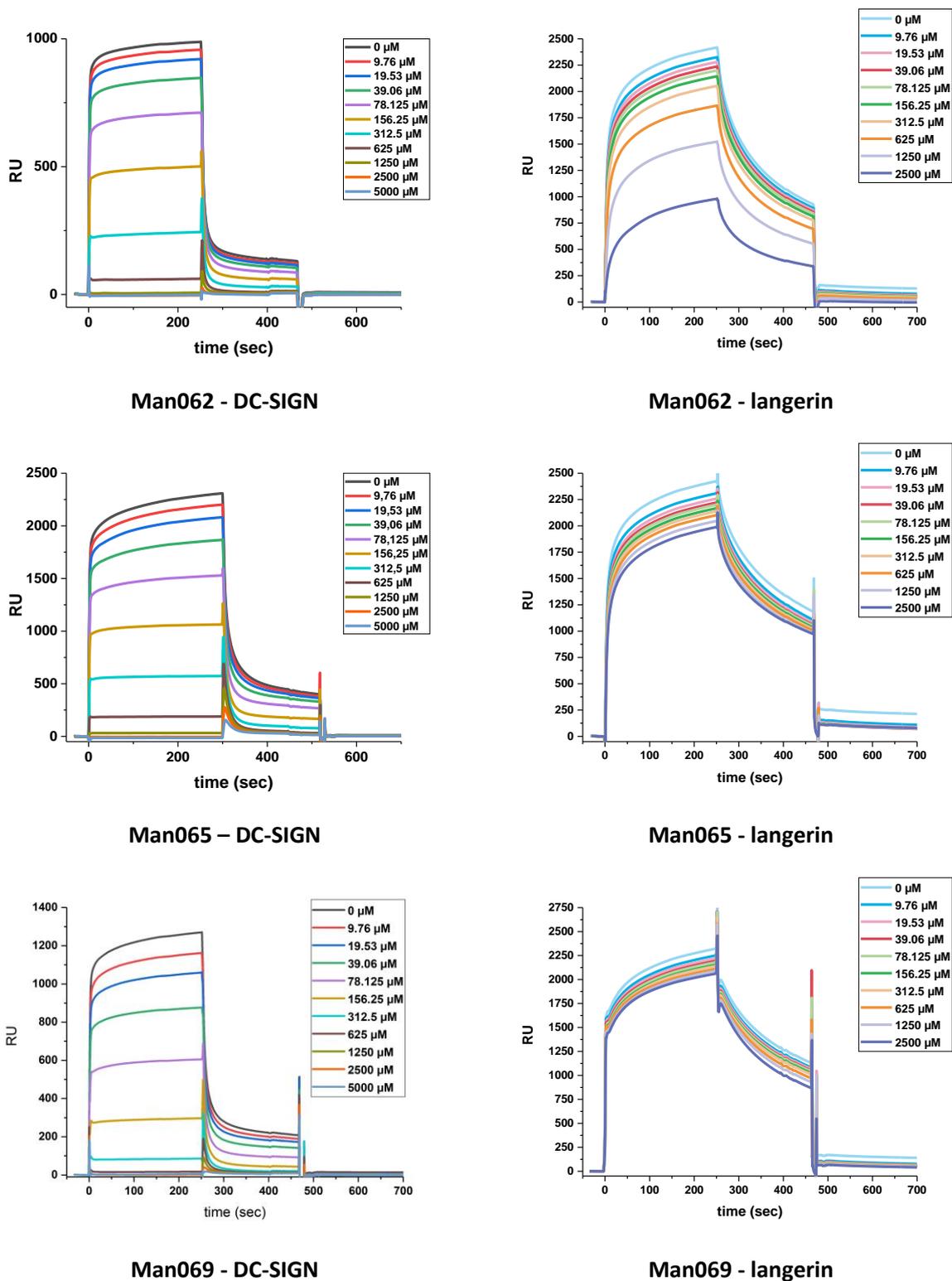


Fig.130 Examples of sensorgrams obtained for Man062-65 and 69 competition assays against DC-SIGN and langerin

We further analysed the interaction properties of **Man069** with DC-SIGN ECD by ITC. **Man069** (2.5 mM) was titrated into a lectin solution (100 μM) (Figure 131). A one binding site model fitting of the data with an

assumed stoichiometry value fixed to 1 yielded a K_d of $52.08 \pm 1.32 \mu\text{M}$, which was in agreement with the apparent affinity determined by SPR competition assay ($\text{IC}_{50} 76.25 \pm 3.25 \mu\text{M}$). A titration with Man069 alone was performed to remove the contribution of the dilution of the glycomimetic in the K_d determination. The experiment was repeated twice.

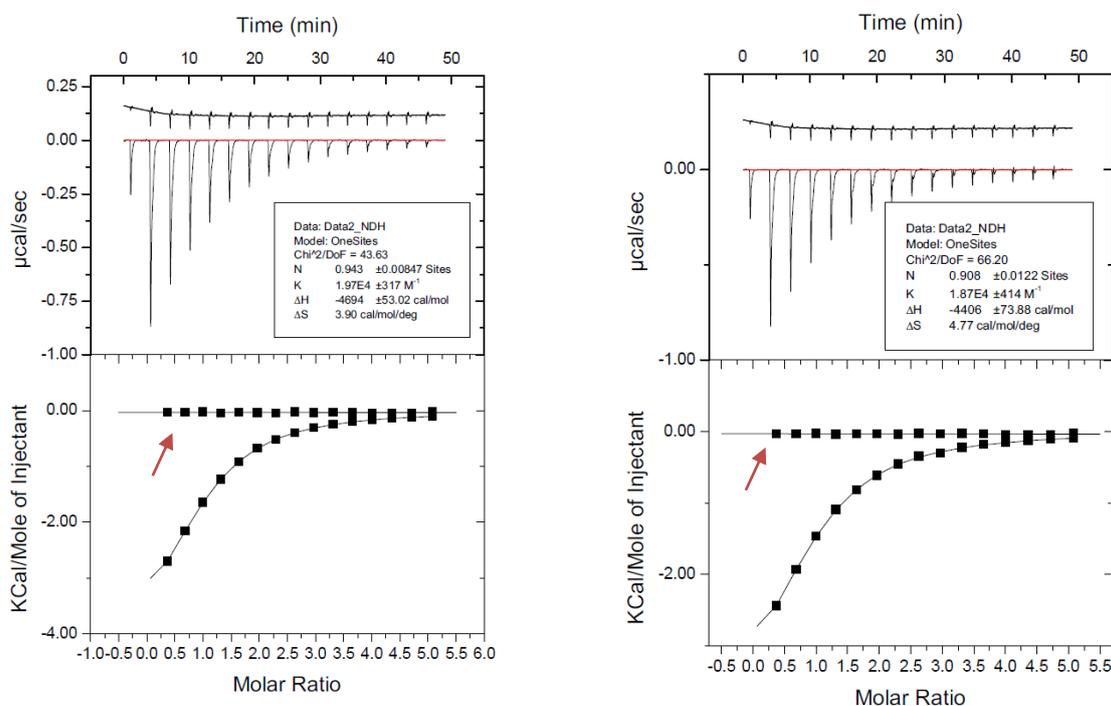


Fig.131 Titrations of glycomimetic Man069 with 2.5 mM to DC-SIGN (100 μM). Upper panel shows the titration thermogram and lower panel, the data integration with fitted curves (1:1 binding model). Both panels contain also the result of Man069 titration alone (red arrow).

Until now, the combination of the moderate affinity and the low molecular weight of these glycomimetics have precluded the evaluation of a K_{Dapp} by SPR in a direct interaction mode. However, both competition assay and ITC analysis have revealed an interaction in the medium-low μM range, there we decided to perform a Man069 titration assay on a DC-SIGN functionalized surface. The resulting K_{Dapp} from steady state fitting was found to be $52.7 \pm 2.7 \mu\text{M}$ (Figure 132), confirming the results previously obtained by SPR competition assay and ITC.

To summarize: three different assays were performed to study the interaction of Man069 new glycomimetic with DC-SIGN. The competition assay revealed an interesting $\text{IC}_{50} 76.25 \pm 3.25 \mu\text{M}$ and the assay performed with langerin showed a medium-high selectivity towards DC-SIGN. ITC analysis showed a 1:1 interaction between the protein and the glycomimetic and a K_{dapp} of $52.08 \pm 1.32 \mu\text{M}$. Finally, SPR direct

interaction assay enable to study the direct interaction between Man069 and a DC-SIGN oriented surface. This final test gave a K_{dapp} of $52.7 \pm 2.7 \mu\text{M}$.

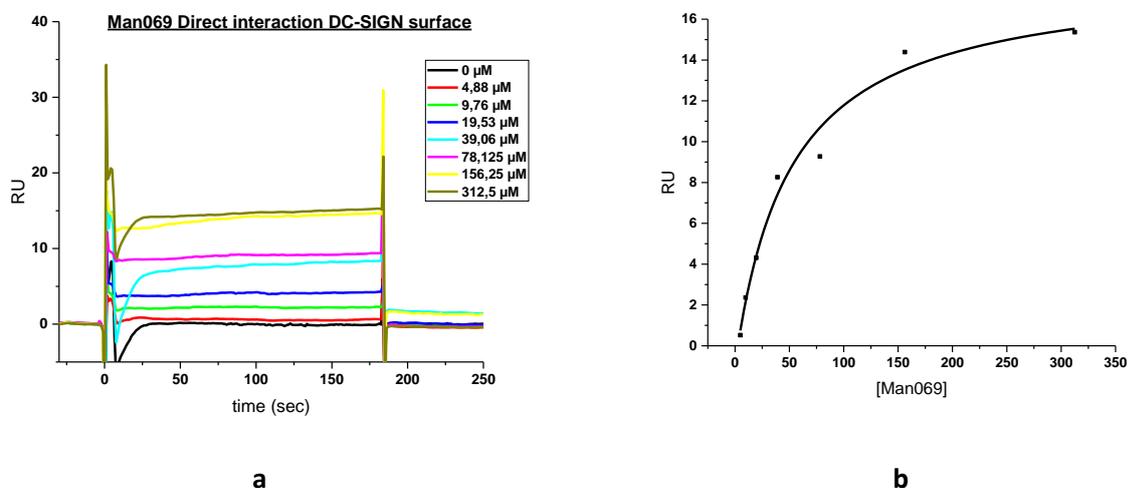
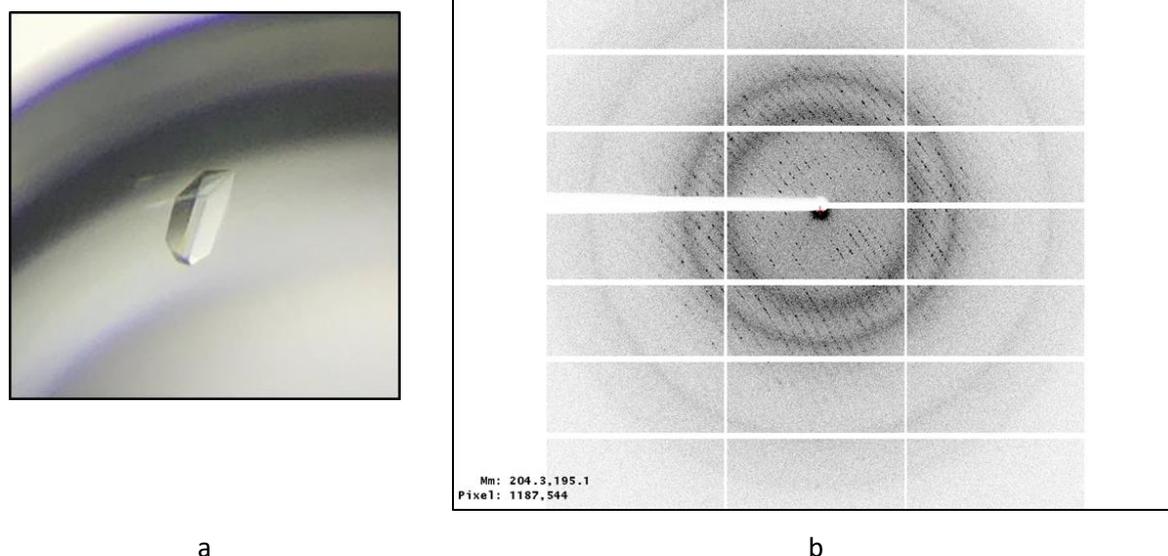


Fig.132 Man069 direct interaction assay a) sensorgram of Man069 titration of DC-SIGN surface and b) steady state fitting.

Co-crystallisation trials of DC-SIGN CRD and Man069 were performed by Dr. Michel Thépaut to identify the precise interactions mode predicted by the computational studies. HTX crystallization platform (EMBL) was used to screen conditions of co-crystallisation with the hanging-drop vapor-diffusion method at 293 K. The drop was composed of a protein/reservoir ratio of 1:1 with protein concentrated at 5.54 mg/mL in 150mM NaCl, 4mM CaCl₂, 25mM Tris pH 8, 2% (V:V) DMSO buffer and 3.25 mM Man069. Among the crystallisation hits obtained, the condition F04 (200 mM Mg(NO₃)₂, 20 % PEG 3350) from the kit PEGs-Suite Qiagen was chosen for manual optimisation screening with four different buffers (MES pH 6, HEPES pH 7, TRIS pH 8 and Bicine pH 9), concentration of PEG 3350 (15 %-25%) and Mg(NO₃)₂ concentration (from 0.15 M to 0.2 M). Finally, the best crystals were obtained in the following condition: 20% PEG 3350, 200mM Mg(NO₃)₂, 100mM MES pH 6.

Interestingly, it has been observed that in control crystallisation experiments without ligand only a few hits were obtained, supporting the idea that the ligand itself plays a role in the crystallisation process.

Crystals were directly flash frozen in liquid nitrogen using Paratone-N as cryoprotectant (Fig.133a). Data collection was performed at id30A-1 beamline (MASSIF-1), ESRF Grenoble, 3200 images (Fig.133b) were collected at 100°K, with an oscillation range of 0.05°, an exposure time of 0.039s per image, and a wavelength of 0.966Å.

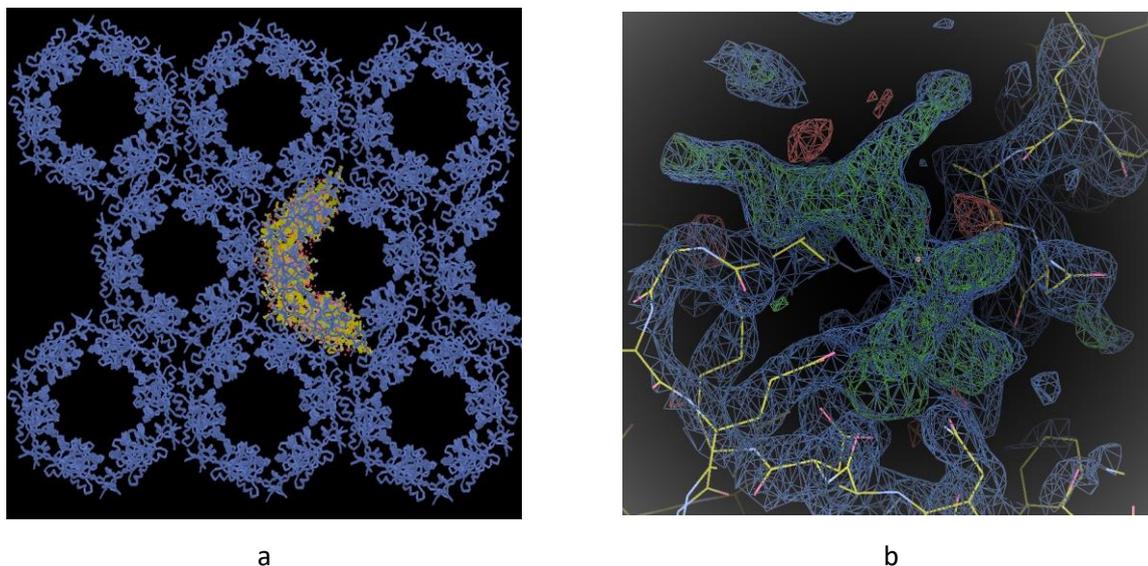


a

b

Fig.133 a) One example of obtained crystal and b) diffraction pattern.

XDS/XSCALE (version 20180126) programs were used to process data resulting in a $P2_1$ space group, with unit cell parameters (see table 13) and a data set of 65455 reflections between 40 and 2.1\AA resolution with 98.4% of completeness. Structure was solved by molecular replacement, using 1K9I DC-SIGN CRD structure and MOLREP (version 11.6.2) (Fig.134).



a

b

Fig.134 a) Crystal packing and b) evidence for ligand presence.

Model building was made alternating refinement with REFMAC (version 5.8.0218) and manual construction with COOT (version 0.8.9 EL) resulting in a final structure with the following parameters. (Table 14)

Table x: DC-SIGN CRD/Man069 complex data collection and structure refinement statistics

Data collection statistics	
Wavelength (Å)	0.966
Space group	P2 ₁
Unit cell parameters (Å)	a=105.612, b=57.507, c=107.247 α=90, β=118.666, γ=90
Resolution (Å)	94.1-2.1 (2.2-2.1)
Unique reflections	65455 (8427)
Completeness (%)	98.4 (97.9)
I/σ (I)	8.47 (2.01)
R _{merge} ^b (%)	9.9 (54.5)
Structure refinement statistics	
Resolution (Å)	40-2.1
Refinement factors	
Used reflections/free (%)	62185/3261
R _{cryst} ^c	0.1759
R _{free} ^c	0.2336
rmsd from ideality	
Bond angles (deg)	1.6610
Bond lengths (Å)	0.0166
Ramachandran plot (%)	
Most favoured regions	97.8
Additional allowed regions	1.4
Disallowed regions	0.8
Average B-factors (Å ²)	33.681

Table 14 Parameters of the crystal structure.

The solved structure revealed that **Man069** ligand, in addition to its binding to the canonical binding site, is also involved in bridging interactions with another CRD within the crystal. This explains the fact that crystallisation seems dependent of the ligand presence. As observable in Figure 135, the crystal packing is enabled by two Man069 ligands bridging each CRD to the next one in a symmetrical way. Each **Man069** molecule interacts with one of the two CRDs *via* the Ca₂₊ binding site, and with the other CRD through an

insertion of one p-hydroxymethylbenzylamide arm within the structure between a loop and helix-2. The two different binding modes have been analysed separately.

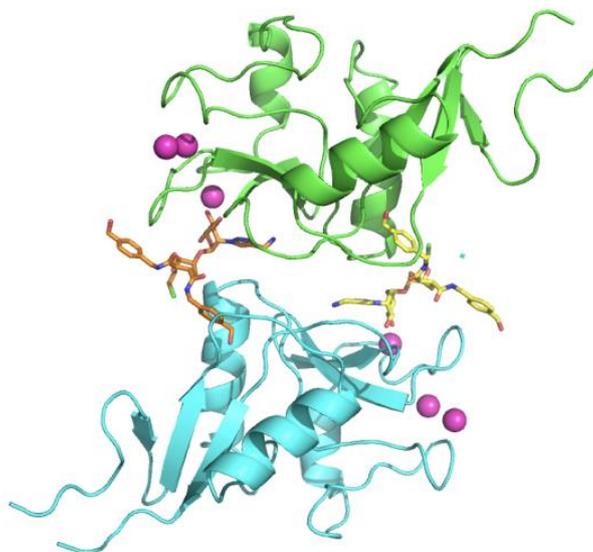


Fig.135 Man069/DC-SIGN solved X-ray structure. Two Man069 ligands bridging each CRD to the next one in a symmetrical way.

Binding mode in the canonical binding site:

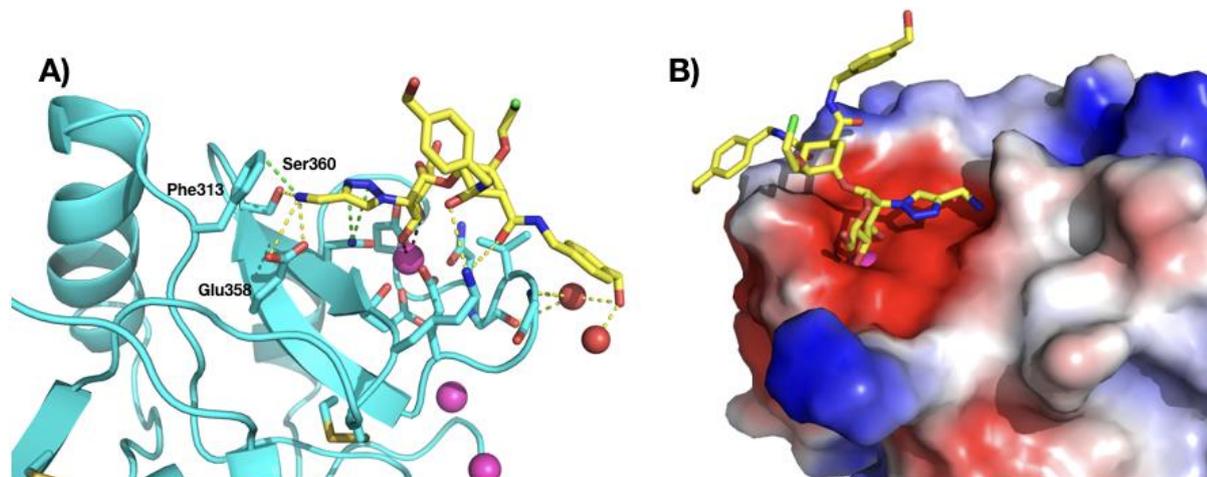


Fig.136 Man069 canonical binding site A) Man069 interacting with DC-SIGNR CRD with the canonical binding site. H bond are represented in yellow, interaction with Phe313 in green B) electrostatic surface

We can observe that **Man069** ligand binds similarly to the previously characterised psDi and Man030 glycomimetics. Remarkably, the additional ammonium group, added on position C2 of mannose with an azide linker, reaches Ser360 and Glu358 by interactions initially expected from the computational docking. Van der Waals interactions can also occur with Phe313 (Fig136A). In Figure 136B, the CRD represented as electrostatic surface highlights the negative potential favouring the ammonium binding.

Non-canonical binding site:

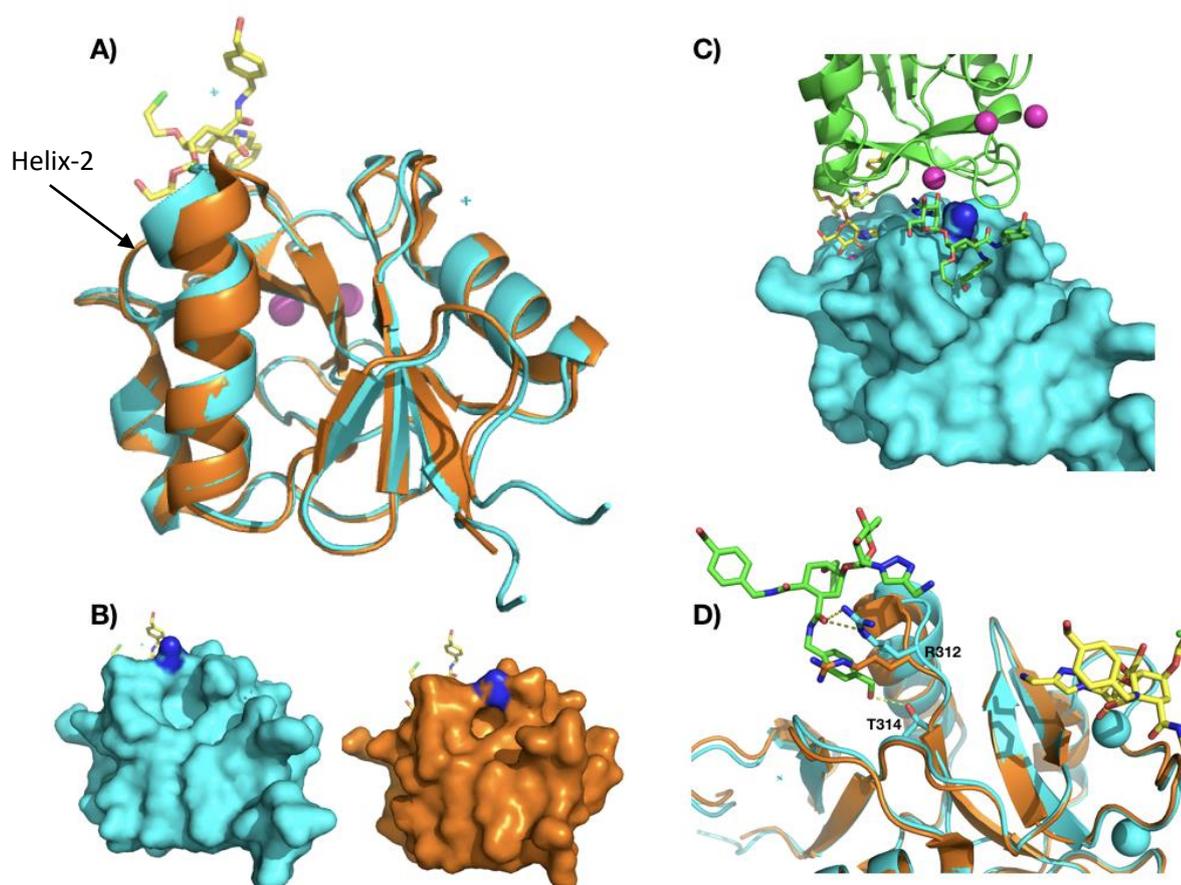


Fig.137 Man069 non-canonical binding site A) Overlay of 1K91 DC-SIGN CRD structure (orange) and DC-SIGN CRD co-crystallised with Man069 (cyan) highlighting the kink of helix-2 due to the p-hydroxymethylenebenzylamide insertion B) The comparison between the two CRDs shows the cavity generated by Arg312 movement C) zoom on the cavity D) the p-hydroxymethylenebenzylamide interaction with Thr314.

In Fig137A, the overlay of the backbone structures of DC-SIGN CRD (PDB 1K9I) with the CRD in our structure complexed to Man069 shows a good conservation of structural elements, with the exception of a kink of helix-2. This helix movement is due to the insertion into the structure of a p-hydroxymethylenebenzylamide arm between this helix and a facing loop (Fig137C), with a consequent movement of Arg312. In figure 137B is shown the cavity generated by the movement of Arg312 creating space for the p-hydroxymethylenebenzylamide. In addition, Arg312 stabilises the ligand through H bond with one carbonyl group of the ligand and finally, Thr314 is also involved in the binding through an addition H bond (Fig.137D).

11. Design of multivalent mannosylated ligands for CLR targeting

This chapter is the result of a collaboration among **three IMMUNOSHAPE groups**: our group, the group of Prof. Yvette van Kooyk (VU University Medical Center, Amsterdam, The Netherlands) and the group of Prof. Jeroen Codee (Leiden Institute of Chemistry, Leiden, The Netherlands). The three groups possess expertise in biochemistry/biophysics, cell biology/immunology and carbohydrate chemistry, respectively.

This project was the first step towards the design of a highly defined molecule for cancer vaccination, which is the final long-term objective of the IMMUNOSHAPE network. This molecule will be composed by CLR ligand for CLR targeting and immune modulation, an antigenic peptide for specific T cell activation and a TLR-ligand for DC activation.

Since the glycomimetics selective towards DC-SIGN developed by the group of Anna Bernardi have been optimised through several cycles and the ultimate ligand was finalised and characterised only at the late stage of my PhD, only classical glycans have been exploited in this study. Moreover, the context of this research project did not seek a differential response between DC-SIGN and langerin, as it was mentioned in chapter **8.3** for the design of HIV-1 inhibitors. Nevertheless, in the future, glycomimetics selective to DC-SIGN could be exploited as well and compared to the data obtained here.

For the sake of simplicity, the chapter will be divided into four different experimental sections, with a mention of the laboratory in charge of the different experiments.

Context:

CLRs expressed by APCs are attractive targets for their endocytic- and immune modulatory properties, as already mentioned in chapter **4.3**, and receptors like DC-SIGN and langerin have been targeted using a vast range of mannosylated products (chapter **3.4**). Multivalent presentation of the ligand is generally foreseen to increase the affinity towards the lectin. However, the overall affinity of a specific CLR will depend on the combination of several parameters: mannose saccharide composition, valency and presentation mode.

Here, we explored different combinations of valency and saccharide epitopes on multivalent glycoclusters, to identify the most effective approaches to target DC-SIGN and langerin with the final goal to interfere with immunological processes in cancer and infectious diseases.

11.1 Carbohydrate synthesis (Leiden)

We focused on DC-SIGN and langerin and their common high affinity ligand, high mannose (Man9). Man9 has been “deconstructed” by **Tim Hogervorst**, a PhD student of Jeroen Codee group, into different mannose cluster groups from the monovalent mannose (A), three different dimannosides (B-C-D) and a branched trimannoside (E). Four different structurally well-defined peptide scaffolds with increasing valency (one, two, three and six) were synthesized and combined with the 5 different *O*-mannosides mentioned above yielding 20 compounds. In addition, one negative control loaded with galactose was synthesized to investigate unspecific binding of the backbone. The scaffolds containing alternating azido lysines were constructed using solid phase peptide synthesis using Fmoc protection. All five mannosides were equipped with a propargyl alcohol/ether spacer. Moreover, all constructs contained a primary amine that was functionalized with biotin for Enzyme-linked immunosorbent assay (ELISA) and FACS. (Fig.138). Figure 139 presents an overview of the 20 compounds.

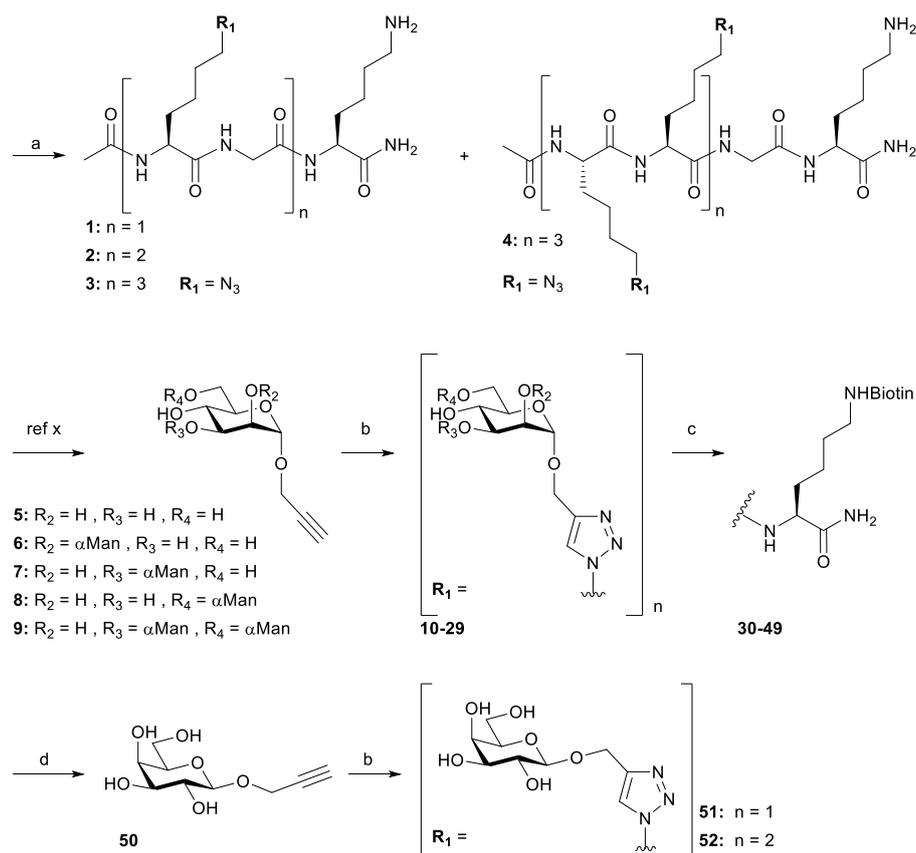


Fig.138 Cluster synthesis a) SPPS Fmoc automated synthesis, b) CuSO_4 , THPTA, NaAsc, DIPEA, c) i) Ac_2O , pyridin.

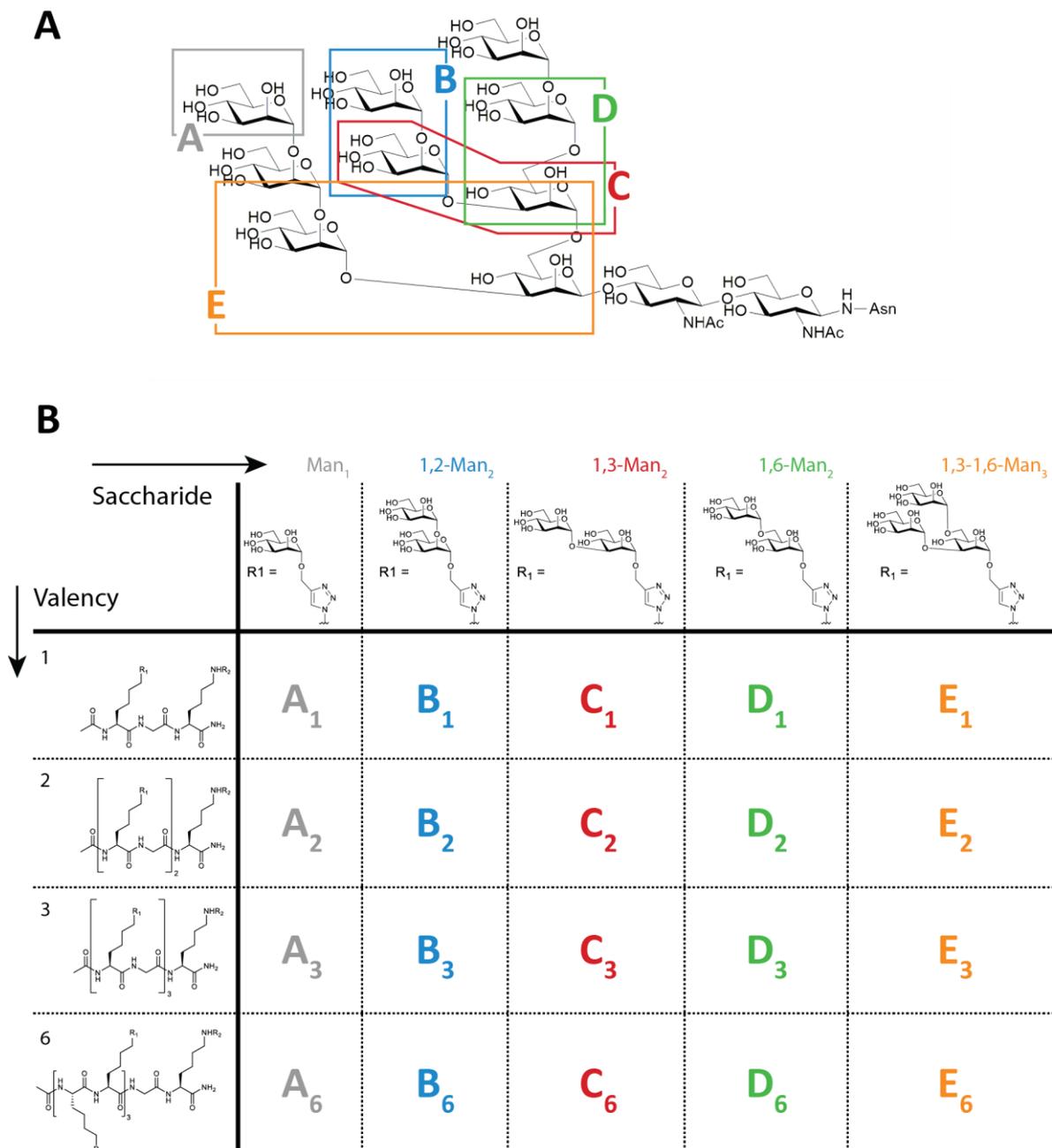


Fig.139 20 mannosylated ligands A) Structure of Man9, a high affinity ligand for DC-SIGN and Langerin. B) Schematic overview of the 20 mannosylated compounds of the library. Monomannose, dimannose, or trimannose were conjugated onto a peptide backbone in clusters of one, two, three, or six, and are coded accordingly. R₂ can be functionalized with biotin.

11.2 ELISA assays (Amsterdam)

ELISA assay was first performed in the laboratory of Prof. Yvette van Kooyk by the PhD student **Eveline Li**. Antibodies anti-Fc were used to coat the wells in order to achieve an oriented immobilization of DC-SIGN/langerin Fc-CRDs [165] available in the laboratory of Prof. Yvette van Kooyk. Biotinylated glycoclusters were incubated at increasing equimolar (the total sum of mannose units is similar per condition)

concentration and the revelation of the interaction was performed by using Strep-HRP. Fig.140a shows the results obtained for DC-SIGN, Fig140b for langerin.

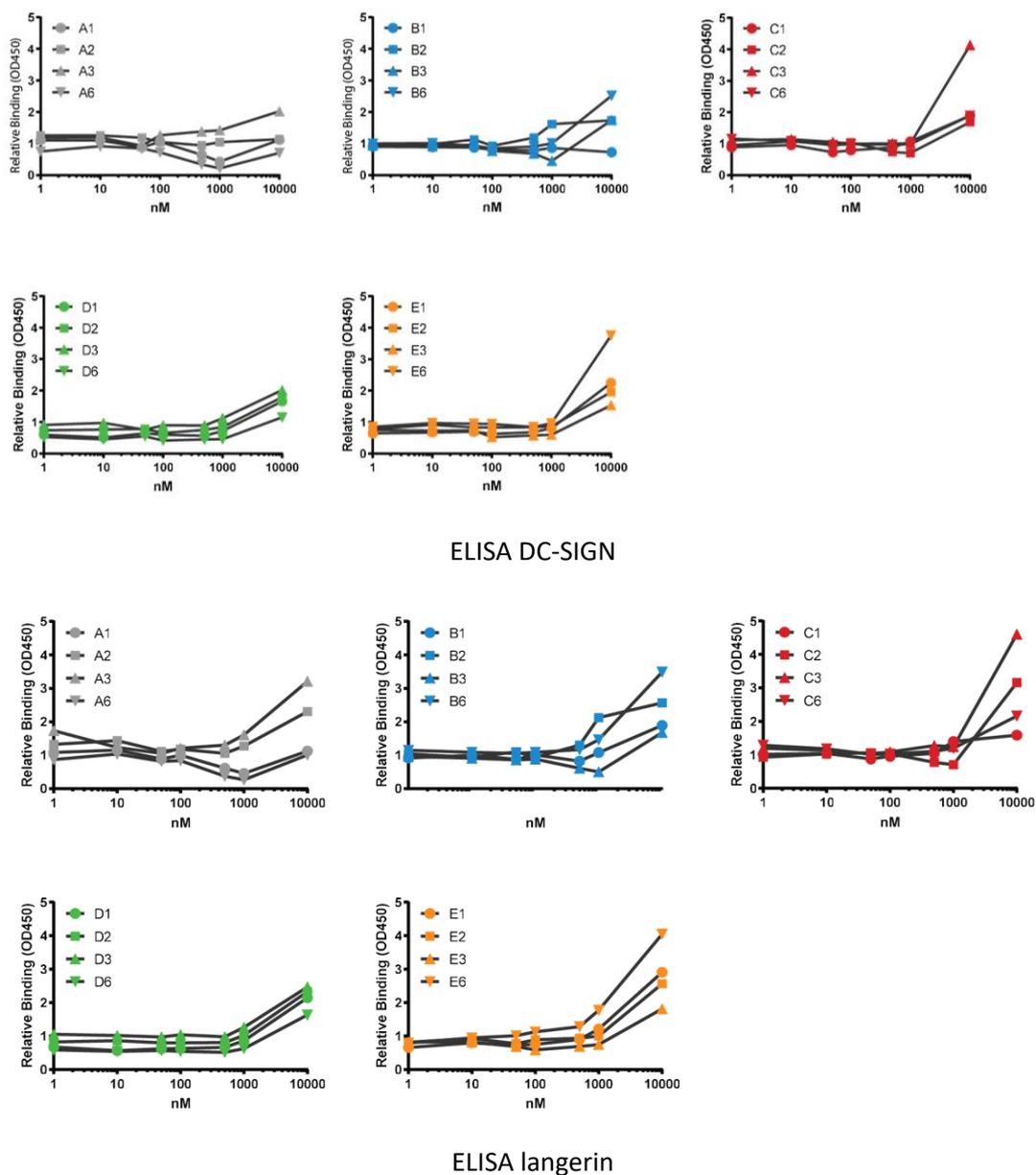


Fig.140 ELISA results performed of DC-SIGN and langerin surface. Increasing concentration of glycoclusters was used.

From this first experiment, the effect of the multivalency was detectable from 10 μ M concentration of the glycoclusters and the assay was not considered sufficient to speculate about multivalency. Indeed, a better characterization of interaction with DC-SIGN and langerin was needed and SPR experiments were performed in our laboratory.

11.3 SPR Experiments (Grenoble)

Competition assays (6.2.3) were first performed on the negative control and all the ligands from cluster 1 (monovalent clusters) using DC-SIGN and langerin ECDs.

The negative control (backbone of cluster 2 conjugated with galactose) was tested as potential competitor to evaluate whether the peptidic backbone would inhibit the interaction between DC-SIGN or langerin with the mannosylated surface. The inhibition was barely detectable, as shown by the inhibition curves in Fig141a, and indicate that the backbone had no, or very low, interactions with the two CLRs. All the monovalent clusters were then tested in a competition SPR assay against DC-SIGN and langerin (Fig141b). The amount of compounds sent by our collaborators solely allowed a starting concentration of 400 μM , which is far below the ideal one for SPR competition assays with monovalent ligands. As expected, the results obtained for mostly all compounds of cluster 1 in competition with DC-SIGN were comparable to the Gal cluster data. For B1 (1-2Man₂) it was possible to obtain an extrapolated IC₅₀ of 2 mM, confirmed by data obtained by our group from previously experiments performed with the disaccharide alone (IC₅₀ of 1-1.5 mM). In addition, an extrapolated IC₅₀ of 1.5 mM was also found for E1.

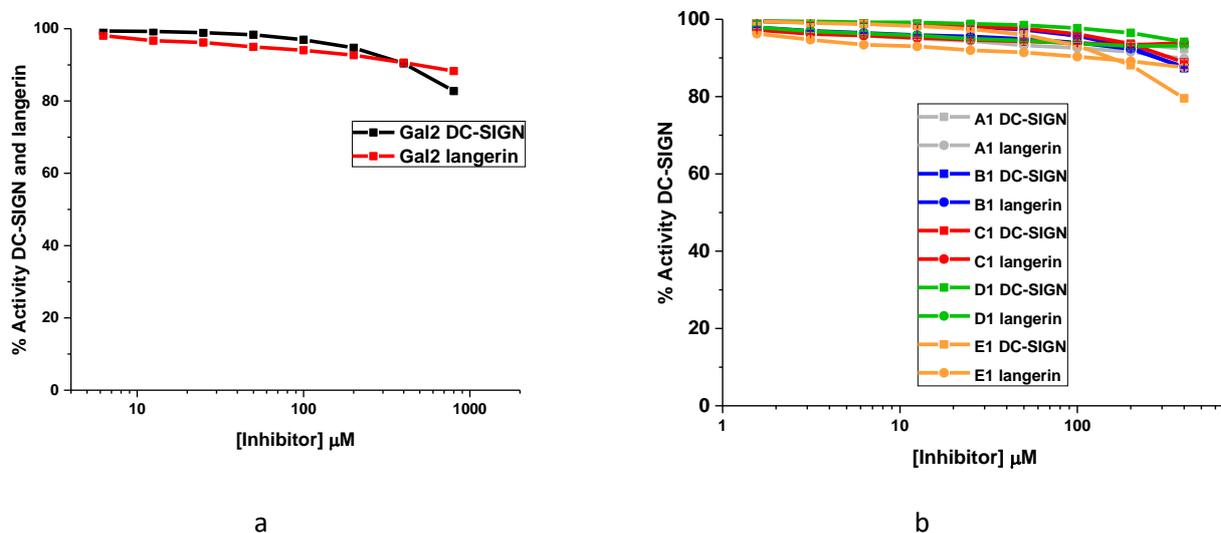


Fig.141 Inhibition curves of a) competition assay using Gal cluster 2 and b) competition assay using the five glyclusters with a valency of 1.

Depending on the quantities sent by our collaborators, we managed to test some of the clusters 2 and 3 against DC-SIGN and/or langerin. A summary of the experience is shown in Table 15

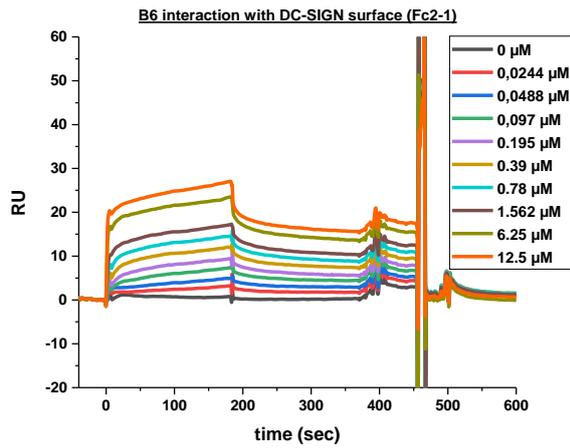
	A		B		C		E	
	<i>DC</i>	<i>lang</i>	<i>DC</i>	<i>lang</i>	<i>DC</i>	<i>lang</i>	<i>DC</i>	<i>lang</i>
2	NE	2400	ND	184	ND	1400	145	ND
3	ND	ND	102	ND	ND	280,5	ND	ND

Table 15 Summary of the competition experiments with clusters 2 and 3 against DC-SIGN and langerin. ND = not determined, NE = not exploitable. The IC₅₀ values are in μM .

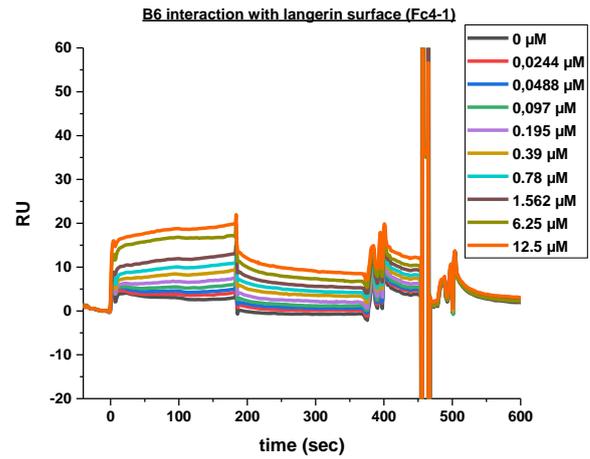
A lot of work still needs to be done on clusters 2 and 3. However, the low quantity of glycoclusters sent and the massive amount of DC-SIGN and langerin required to test all the conditions obliged us to concentrate on the hexavalent glycoclusters. Indeed, cluster 6, was expected to reach affinity in the range of low μM for which the previously described SPR inhibition assays are not appropriate. In fact, this approach can only correctly estimate the inhibitory power of compounds for which the interaction with the CLR is weaker than the CLR-surface interaction. DC-SIGN has an affinity (K_{dapp}) for the mannosylated surface of 5 μM and, therefore, inhibitions in the range of lower μM or nM would be underestimated.

For this reason, interactions between these multivalent compounds and the CLRs were monitored in real-time using direct interaction assays using a surface of oriented DC-SIGN or langerin. That was achieved by the functionalization of the surface with StrepTactin. N-terminally StrepTagII lectin ECD could then get immobilized onto the surface in an orientated way. This strategy guarantees that all CRDs were accessible to multivalent glycoclusters. Using this approach, we have been able to estimate the affinity and avidity of the different glycoclusters towards DC-SIGN and langerin. The curves of binding responses versus concentration were fitted using a steady state affinity model to obtain a K_{dapp} .

A range of increasing concentrations of the various clusters 6 was applied to DC-SIGN and langerin surfaces and the highest concentration considered was 12.5 μM . Fig.142 shows an example of sensorgrams obtained from B6 interaction with DC-SIGN and langerin surfaces. For all the hexavalent glycoclusters (data not shown, see Annexes), we reasonably considered that the equilibrium of interactions with the CLR surfaces was reached and, therefore, that it was possible to use the steady state affinity equation for the calculation of K_{dapp} (Fig.143).

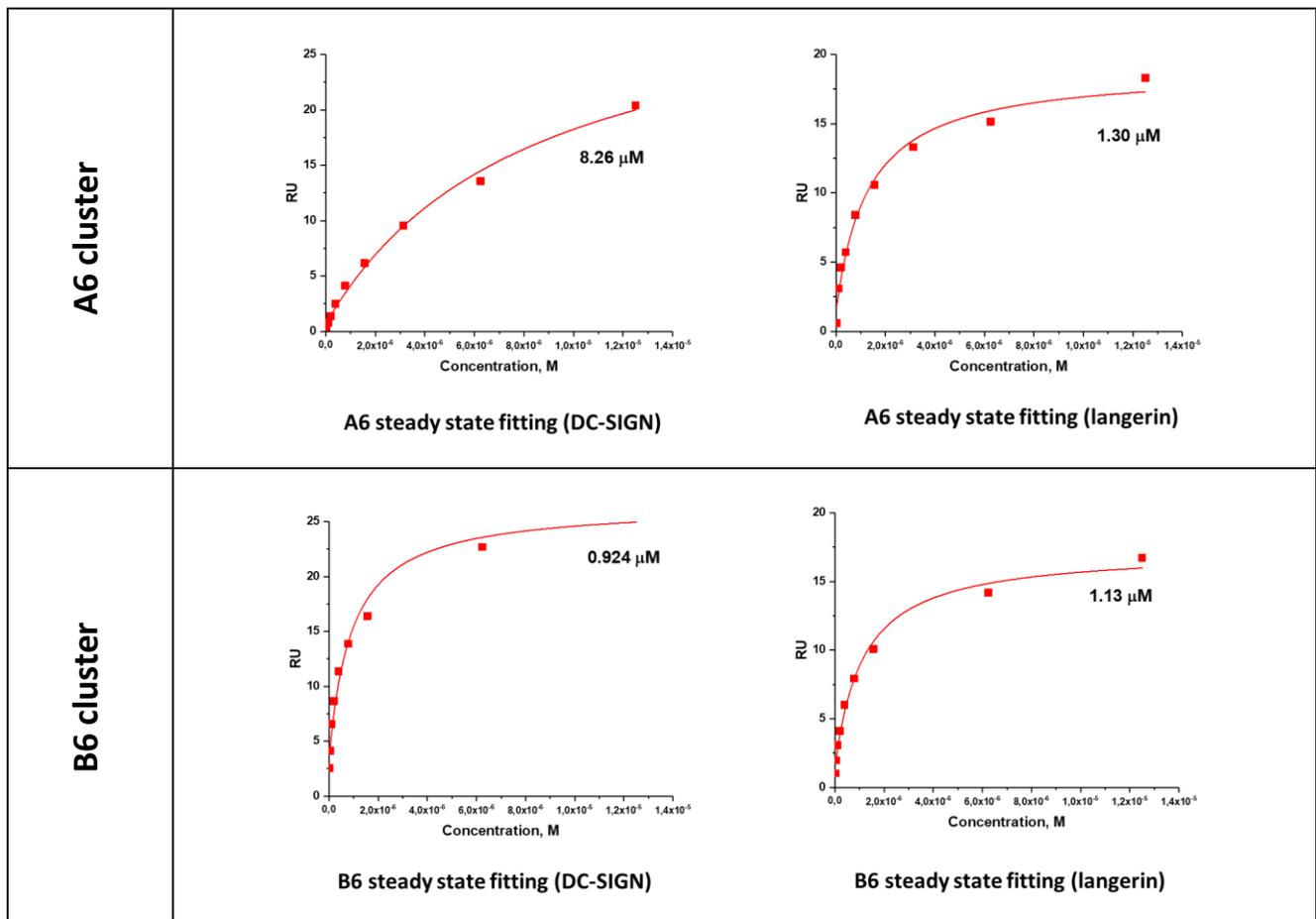


a



b

Fig.142 Sensorgrams obtained from B6 interaction with a) DC-SIGN and b) langerin. Concentration 3.125 μM was removed because of the presence of air bubbles in the sample.



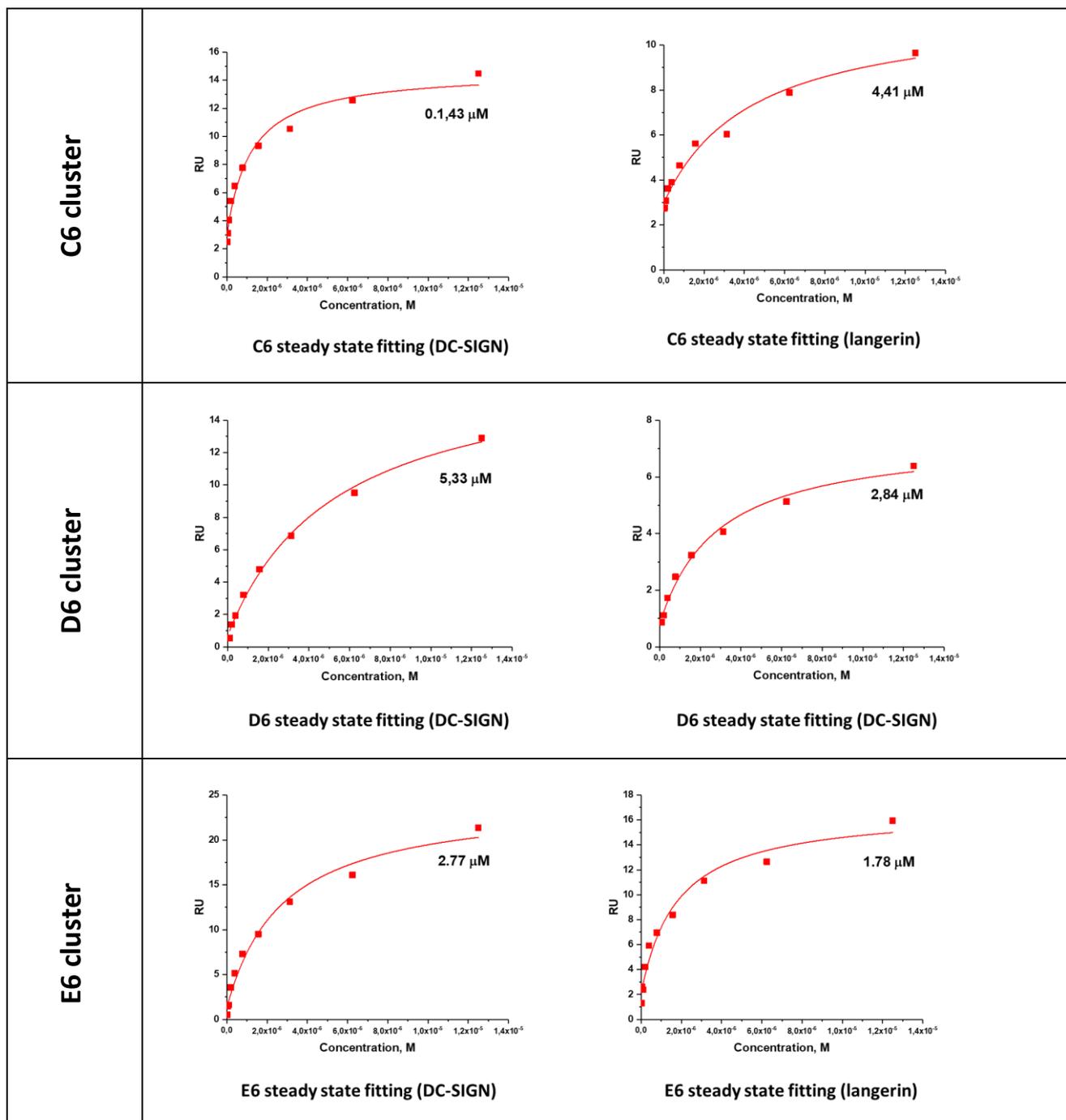


Fig.143 K_{dapp} values obtained for cluster 6 interaction with DC-surface and langerin surface by using steady state fitting. Careful of kd for C6 for DC SIGN

All clusters 6 interaction with DC-SIGN and langerin surfaces revealed K_{dapp} in the low μM range. **A6** (mannose) and **D6** (1-6Man₂) showed a weaker interaction with DC-SIGN compared to the other clusters. The most effective was cluster **B6**, proving once more that 1-2Man₂ is the favorite ligand for DC-SIGN among the dimannobiosides. **C6** proved also to be a good cluster with an average K_{dapp} of $1.177 \pm 0.253 \mu\text{M}$. Finally, **E6** which possesses features common to **C6** and **D6** (1-3,1-6Man₃) showed a K_{dapp} of 2.785 ± 0.015

μM , intermediate between the **D6** and **C6** values. For langerin, all the clusters showed a general good affinity and **B6** was confirmed as best glycocluster.

In conclusion, clusters 6 with hexavalent valency were considered promising glycoclusters for DC-SIGN and langerin targeting (Fig.144). They were included in biological assays at cellular level.

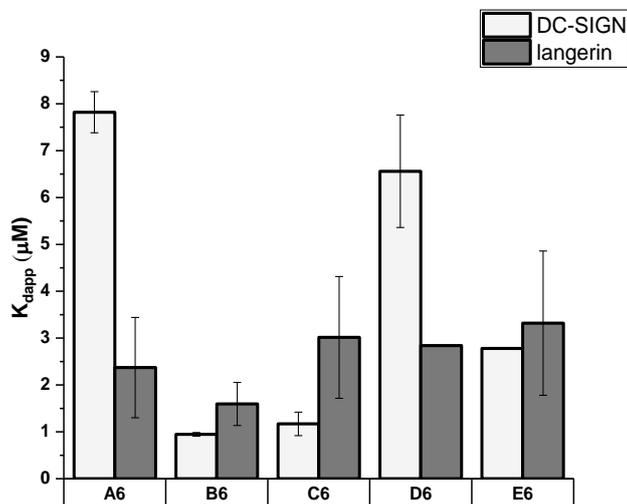


Fig.144 Comparison of the K_{dapp} values obtained for cluster 6 interaction with DC-SIGN and langerin surfaces.

11.4 FACS Assays (Amsterdam)

To measure the binding to cellular DC-SIGN and langerin and obtain preliminary biological results, **Eveline Li** stimulated with the biotinylated glycoclusters (at a concentration of 10 μM) a B-cell line, EBV OUW, transfected with langerin (Fig.145). Functionalization of the glycoclusters with biotin enabled the revelation by fluorescent streptavidin. The binding was measured by flow cytometry.

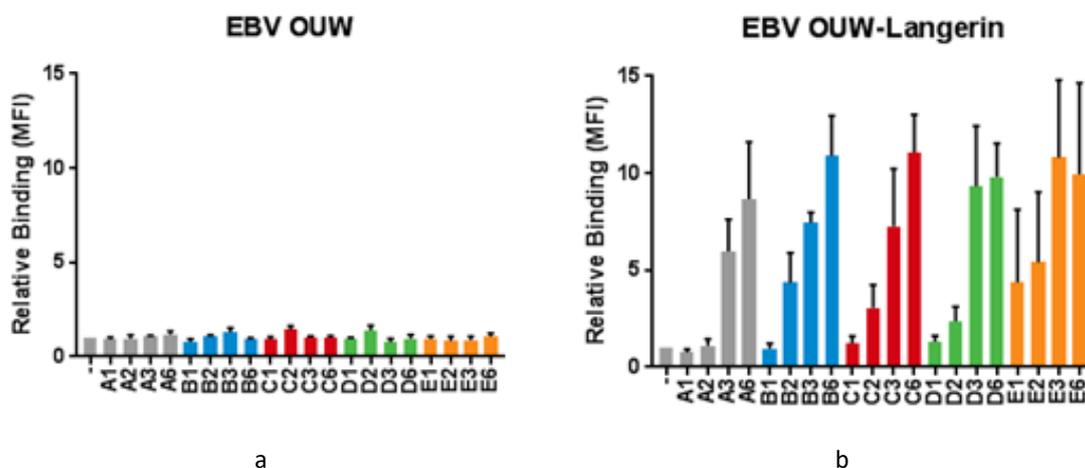


Fig.145 Interaction between biotinylated glycoclusters at a concentration of 10 μM and EBV OUW cell lined a) non-transfected and b) transfected with langerin.

On Fig.145a, no fluorescence of non-transfected EBV OUW (negative control) was measured, as expected, while langerin transfected EBV OUW (Fig.145b) showed interaction with several glycoclusters. The assay revealed that the effect of multivalency is maintained at the cellular level. However, the assay will be repeated in order to obtain better statistics and similar experiments with EBV OUW transfected with DC-SIGN are currently ongoing.

Nevertheless, we need to keep in mind that these are very preliminary biological tests that have to be confirmed. Moreover, two other biological tests are presently running at the cellular level: 1) monocyte derived DCs / glycocluster interaction tests and 2) monocyte derived DCs incubation with glycoclusters in the presence of LPS to observe cytokine production and investigate the effects on immune reaction modulation.

Part IV.

Conclusions and Perspectives

12.1 Remarks on the production of non-commercially available lectins

Three ECD constructs have been produced: DC-SIGNR ECD, Dectin-2 Strep-ECD and Mincle His-ECD. The production of DC-SIGNR ECD was efficient without any major bottleneck and reasonable yields of purification were obtained (typically 10-15 mg from 1 L of cell culture). Dectin-2 Strep-ECD and Mincle His-ECD yields of production were considerably lower. Nevertheless, the functionality of both proteins was confirmed. Dectin-2 Strep-ECD ligand specificity was further investigated and glycomimetics selective for this CLR were identified (**8.2.2.1**).

Regarding the CRD constructs, a more extended panel of CLRs have been developed. DC-SIGNR His-CRD, BDCA2 His-CRD and LSEctin His-CRD were produced as functional proteins. MCL His-CRD was successfully purified, its secondary structure integrity was confirmed but the absence of any known ligand impaired the assessment of the functionality. Dectin-1 His-CRD and Dectin-2 His-CRD were non-functional, while Mincle His-CRD production led to aggregation.

Finally, site-specific TETRALEC strategy was successfully exploited for DC-SIGNR His-CRD and MCL His-CRD, while random TETRALEC complexes were obtained for LSEctin His-CRD and BDCA2 His-CRD.

DC-SIGNR ECD, DC-SIGNR His-CRD, DC-SIGNR-TETRALEC, Dectin-2 Strep-ECD and LSEctin His-CRD ligand investigation on glycan arrays and SPR experiment will lead to three different publications.

Besides, many glycomimetics and glycoclusters synthesized by our chemist collaborators inside and outside the IMMUNOSHAPE network have been designed to target DC-SIGN and langerin. Therefore, I also intensively participated in the production of DC-SIGN and langerin ECDs in order to perform interaction studies.

Regarding BDCA2, dectin-1, dectin-2 and MCL His-ECDs, for which production was not achieved during my PhD for a matter of time, they were however successfully expressed in BL21(DE3) strain, as shown by the SDS-PAGE in Fig.146. Further efforts will be made for their production in our team.

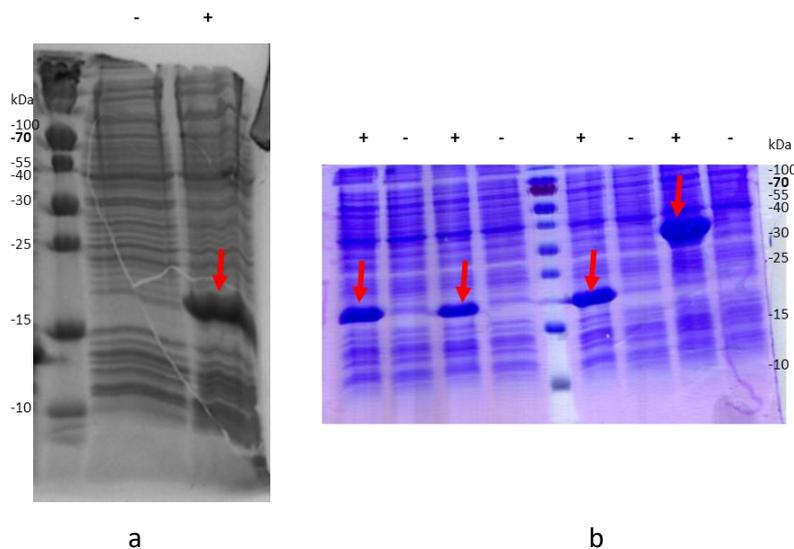


Fig.146 BDCA2, dectin-1, MCL and dectin-2 His-ECD expressions. SDS-PAGE (12%) analysis of a) BDCA2 His-ECD (22.04 kDa) b)MCL (22.29 kDa), dectin-2 (21.02 kDa), dectin-1 (21.9 kDa) and DC-SIGNR His-ECDs (38.95 kDa) expression in BL21(DE3) cells. Sample before (-) and after (+) induction with 1 mM IPTG. The red arrows indicate the ECDs.

Dectin-2 His-CRD construct was produced as non-functional. To investigate whether the initial choice we made when dectin-2 CRD construct was designed, we compared, using PyMOL software, our construct to the one that allowed dectin-2 CRD crystallisation (PDB entry 5VYB) [78].

Our design was based on sequence alignments with other CLR CRDs previously successfully expressed or crystallised. At that time, dectin-2 CRD had not been crystallised. Our comparison revealed that the sequence we selected (Fig.147 in red) is missing a N-terminal β -sheet (Fig.147 in green) present and structured in the crystallised dectin-2 CRD. This additional structural element allowed to form β -sheet including a disulphide bridge, that may be crucial for proper folding. Dectin-2 secondary structure was analysed by CD and indeed revealed an improper folding.

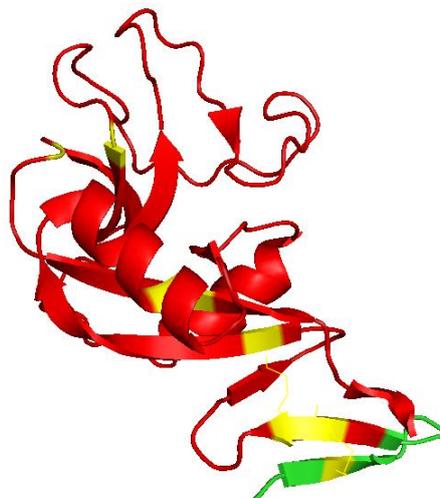


Fig.147 X-ray structure of dectin-2 (5VYB). In red, the amino acid sequence corresponding dectin-2 CRD construct that we designed, in green the additional N-terminal loop present in the crystallised CRD. Disulphide bridges are shown in yellow.

For this reason, dectin-2 CRD N-terminal sequence will be accordingly modified and the new construct will be expressed and purified again. Investigations will be also performed on Mincle CRD and a collaboration with Pr. Antonio Molinaro will be settled to study mincle and MCL proteins towards bacterial glycolipids [166].

Dectin-1 CRD functionality was refuted for its failure to bind to GAS, an interaction only established between the pathogen and murine Dectin-1 (data not published). Since human dectin-1 interacts with *Candida albicans*, new interaction assays will be performed in the laboratory of Pr. Bernd Lepenies on this pathogen [74], in line with the previous of mincle and MCL interaction.

Finally, a screening method for a rapid optimization of the refolding conditions will have to be set up. Indeed, once overexpression is achieved, the refolding step is the tighter bottleneck of the production procedure, as shown in chapter 8.1, and for most of our proteins the refolding buffer could probably be optimised. Different buffers will be screened on the nanoDSF Nanotemper platform in our group, which is an advanced Differential Scanning Fluorimetry technology. It can detect changes in the fluorescence of tryptophans which is strongly dependent to the surroundings. These changings could reflect the chemical and thermal stability of the protein. The nanoDSF monitors the temperature at which a folded protein unfolds and calculate a melting temperature (T_m). Refolding will be attempted by rapid dilution of protein solubilised in GDN-HCl into different refolding buffer solutions, containing various salts and additives. Therefore, we may assume that the conditions giving the higher T_m would be the more favourable for refolding. [167]

Lastly, we presented TETRALEC as a tool to enhance multimerisation at the protein level. Its Structural characterisation has been performed using SEC-MALS analysis to define the stoichiometry of the complex. The analysis seemed to support the formation of the 4:4 TETRALEC stoichiometry (4 NeutrAvidins : 4 DC-SIGNR CRDs) but it did not give a definitive conclusion. The additional technique that could be exploited is native Mass Spectrometry (MS). MS performed under so-called "native conditions" can be used to determine the mass of biomolecules that associate non-covalently [168] and figure 145 presents the informations that native MS can provide. For this reason, a collaboration with the MS platform at the institute([http://www.isbg.fr/preparation-d-echantillons-controles-qualite/spectrometrie-de-masse/?lang=fr](http://www.isbg.fr/preparation-d-echantillons-contrroles-qualite/spectrometrie-de-masse/?lang=fr)) may be established to definitively assess the stoichiometry of the TETRALEC complex.

In addition, in order to provide more information regarding *Candida albicans* recognition by DC-SIGNR, confocal microscopy will be used in collaboration with the group of Bernd Lepenies. Preliminary results obtained by Joao Monteiro showed that confocal experiments performed with DC-SIGNR CRD-Fc fusion protein reflects the results obtained by FACS analysis. The same approach will therefore be used with DC-SIGNR CRD, ECD and TETRALEC.

12.2 Exploiting glycan/glycomimetic array screenings

In the glycan microarray technique used throughout this work, glycans are attached onto a N-hydroxysuccinimide NHS ester activated glass surface by the reactive primary amines at their reducing termini. Fluorescently labelled CLRs are incubated with the array and, after an accurate analysis, histograms are used to represent the median relative fluorescence intensity per glycan.

134 synthetic N-glycans (mostly parasite and plant structures) available in Niels Reichardt laboratory were screened for the first time using a panel of human CLRs that we have produced. Some of the identified interactions were expected. For example, the established BDCA2 interaction with Gal β 1-3/4GlcNAc β 1-2Man [54] was confirmed by the glycan array analysis. Moreover, DC-SIGN and DC-SIGNR interactions with highly mannosylated glycans were as well awaited.

However, three CLRs (DC-SIGN, DC-SIGNR and LSECtin) showed a context-depend selectivity towards positional isomer glycans (Fig.148). Despite having the same glycan composition of the antennae, CLR show a selectivity dependency to the relative position of the antennae, highlighting that CLR recognition is sensitive to the context of glycan presentation. These findings can open the door to further improvement on selectivity of the identified epitopes. NMR studies will be conducted to investigate this selectivity rationale from a structural point of view.

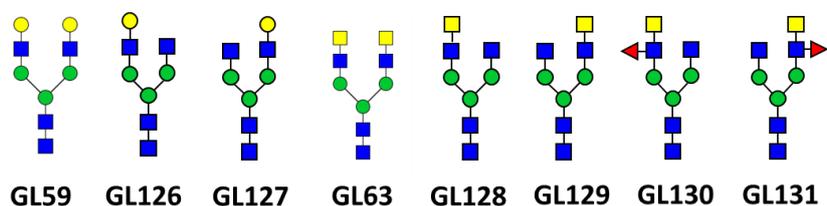


Fig.148 The six asymmetric glycans (GL-126-127-128-129-130-131) and the two non-branched corresponding glycans (GL59-63) considered in publication 8.2.1.1

In addition to the 134 N-glycans, 40 glycomimetics synthesized by the group of Anna Bernardi were also printed on the above-mentioned slides and screened for their interaction with our human CLRs. Two main outcomes resulted from this study.

- 1) DC-SIGN and DC-SIGNR screening of glycan array was followed by a biophysical characterisation of the interaction of three glycomimetics by SPR. The SPR results obtained confirmed the tendency observed in the glycomimetic array. Moreover, this was the first time that pseudo-mannobioside based glycomimetics were tested in competition assay against DC-SIGNR and these structures may be further elaborated for the development of glycomimetics directed towards DC-SIGNR.

2) For the first time, glycomimetic ligands able to interact with dectin-2 were identified. In addition, some of them were shown to selectively interact with dectin-2 (Fig.149), opening the possibility of an unprecedented selectivity between the two CLRs towards glycomimetic compounds that may be related to the different nature of their two binding sites. To investigate these findings, neo glycoproteins decorated with those glycomimetics will be synthesised by Blanka Didak and will be tested by SPR direct interaction assays on dectin-2 and DC-SIGN surfaces.

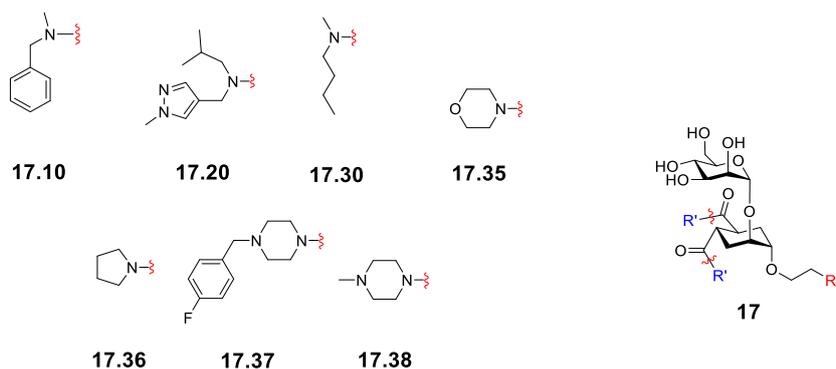


Fig.149 Structure of mannose based glycomimetics that recognise dectin-2 and not DC-SIGN.

12.3 Enhancement and development of new glycomimetics specific to DC-SIGN

Two approaches have been successfully exploited to develop a third generation of glycomimetics stable and selective towards DC-SIGN.

The **first** approach explored the substitution of the oxygen involved in the glycosidic bond, into sulphur in order to obtain a glycomimetic with increased stability to enzymatic hydrolysis. SPR competition assay revealed that S-linkage does not affect the capacity of the compound to inhibit DC-SIGN interaction with the mannosylated surface and the results are presented in publication **8.3.1.1**. In addition, assay with α -Mannosidase proved the *in vitro* resistance of S-linkage glycomimetic towards glycosidases.

In parallel, C-glycosides and one Si-glycoside were synthesized and SPR competition assay results suggested that changing the O-glycosidic nature of the disaccharidic bond does not affect the glycomimetic inhibition power. The following Figure 150 allows a comparison of the Man α 1-2Man compounds derived designed to enhance stability against glycosidases.

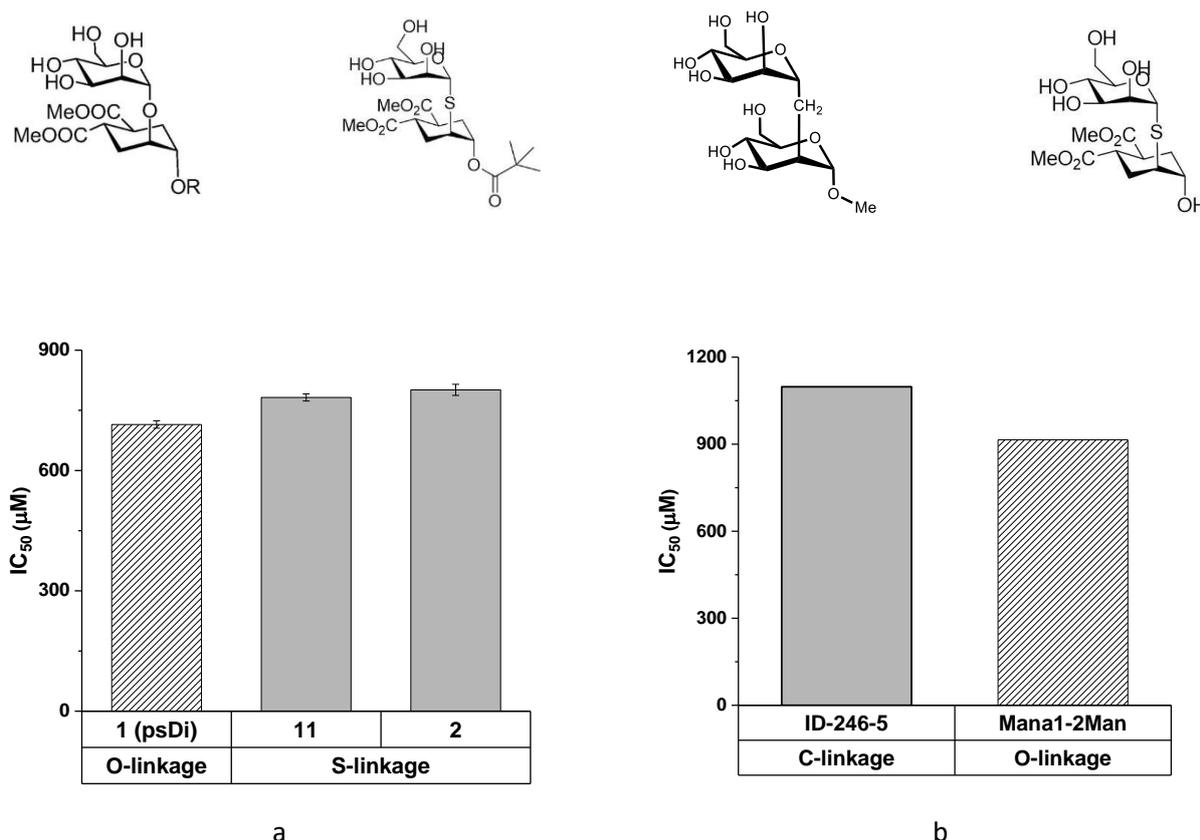


Fig.150 Structures and IC_{50} comparison of a) glycomimetics b) glycosides designed to enhance stability against glycosidases. PsDi and Man α 1-2Man are used as controls.

In the future, glycosidases hydrolysis test will have to be performed on ID246-5 and on the other C-glycosides and Si-glycoside developed, in order to identify the most resistant compound.

The **second** approach combined three different expertise: molecular docking, organic chemistry and biophysical/structural characterisation. Docking studies allowed the identification of fragments suitable for the modification of the mannose 2-hydroxyl group of psDi to create favourable interactions with the Phe313 region of DC-SIGN that was shown to be involved in the binding of natural oligomannosides. These fragments shared the presence of an ammonium ion that could enable cation- π interactions with Phe313, hydrogen bond with Ser360 and ionic interaction with Glu358.

Two synthetic strategies were developed and the 2-hydroxyl group of mannose was replaced by a nitrogen atom. The first strategy produced α -amino acid derivatives compounds. However, the SPR results suggested that the distance between the psDi scaffold and the ammonium ion pocket was not optimally covered by the α -amino framework. Therefore, the approach was abandoned.

The second approach consisted in triazole derivatives synthesis. One of those compound, **Man062** (Fig.151a) was tested in competition assay against DC-SIGN mannosylated surface interaction and revealed an interesting IC_{50} of $113.09 \pm 11.99 \mu\text{M}$ (Fig.151b). Seeking higher selectivity towards DC-SIGN, **Man62** was then implemented into **Man069** (Fig.151a), which combines **Man030** backbone structure and **Man062** modification in position C-2. Figure 148 presents **Man062-069** structures, the IC_{50} values and the inhibition curves of inhibition assays performed with DC-SIGN (square) and langerin (circle).

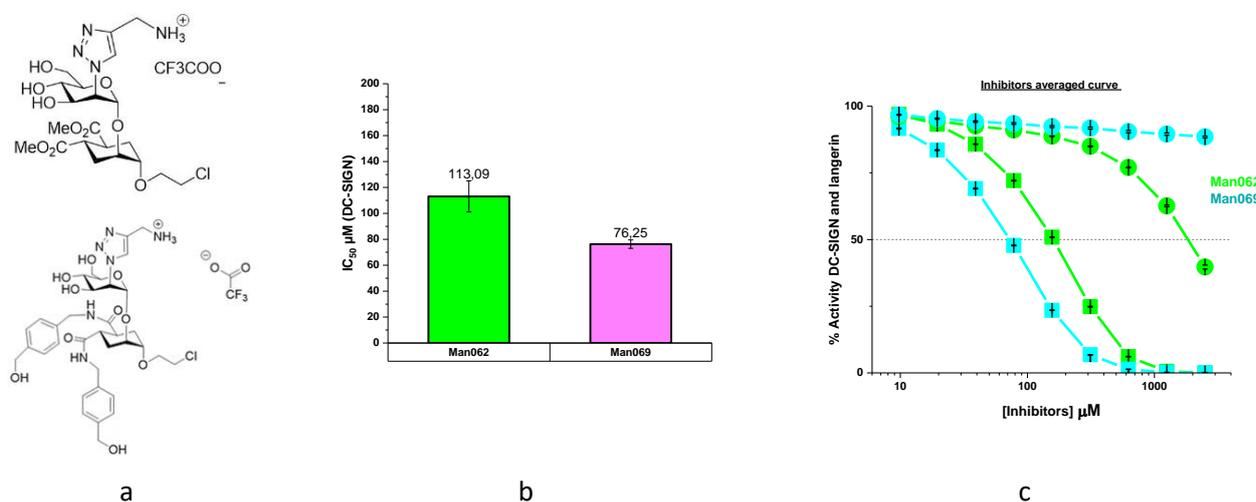


Fig.151 Man062 (light grey) and Man069 (dark grey) competition assays. a) Structures of the glycomimetics b) IC_{50} obtained from the competition assay with DC-SIGN and c) Inhibition curves from competition assays with DC-SIGN and langerin.

Finally, **Man069** was successfully co-crystallised with DC-SIGN CRD. The X-ray structure perfectly confirmed the three interactions postulated by the docking studies. In addition, the ligand:protein stoichiometry was unexpectedly showed to be 2:2. In fact, one **Man069** molecule interacts with Phe313, Ser360 and with Glu358, while a second **Man069** is interacting with Arg312 of the same CRD. In order to understand if this

2:2 stoichiometry is an artefact of the crystallisation process and assess whether it reflect the actual binding mode with one **Man069** ligand able to bridge two CRDs also in solution, AUC experiments will be performed using both DC-SIGN CRD and ECD. The ITC experiments that have already been performed with the tetrameric ECD revealed a 1:1 stoichiometry of the interaction. However, if two molecules of **Man069** are actually bridging two ECDs the final stoichiometry would still appear as 1:1 because the ITC technique measured the heat released after binding of **Man069** to DC-SIGN. AUC technique will discriminate species in solution and reveal the corresponding masses. If two proteins are bridged by two **Man069**, the mass would correspond to a dimer.

Moreover, **Man069** proved to be a valid candidate to be loaded onto multivalent scaffold.

Finally, another triazole derivative was synthesized, namely **Man065**, which contained a pyridine. SPR competition assay against DC-SIGN revealed an IC_{50} of $145.4 \pm 1.9 \mu M$. We suppose that pyridine can only interact with the Phe313, since at the pH 8 used the pyridine cannot make any H bond. A future control will be performed replacing the pyridine by a simple aromatic cycle. Finally, co-crystallisation assay of **Man065** in complex with DC-SIGN CRD will be performed.

12.4 Study of mannose clusters targeting CLRs as a future tool to trigger anticancer immune response

The final goal of this project is the design of a highly defined molecule using CLR ligand to target APCs for various applications, including **cancer vaccination**. In this context, the multivalent presentation of the glycan is crucial and it is important to study how it will influence interaction with the CLR. We focused on DC-SIGN and Langerin and the common high affinity ligand for both, high mannose (Man9). Man9 has been deconstructed into Man1, Man2 (in 3 different conformations) and Man3. These ligands were loaded onto different however well defined peptidic backbones, with different multivalency. The rationale is to understand the relationship between multivalency and/or mannoside structures.

In this case, the purpose was not to differentially target DC-SIGN and langerin it could be the case for HIV-1 inhibition strategies (**chapter 8.3 and 9.3**). Here, these two CLRs can be both exploited to target APCs in the context, for example, of melanoma cancer, since both lectins are found in epidermal tissue. Nevertheless, glycomimetics developed by the group on Anna Bernardi that showed an improved affinity/selectivity towards for DC-SIGN compared to the natural glycans could also be considered in the future.

A K_{dapp} value between 0.9 and 7 μM was obtained for all of the hexavalent clusters in interaction with DC-SIGN and langerin surfaces, showing the importance of multivalency. Compared to the ELISA assays performed by the group of Yvette van Kooyk that could only anticipate a “vague” interaction enhancement triggered by multivalency our SPR characterisation was able to precisely characterise this gain. In addition, preliminary biological tests seem to indicate that *in cellulo* the same tendencies as observed with the SPR data are reproduced. Further analysis are ongoing.

Of course, since the final molecule will also possess, in addition to a CLR ligand, a TLR ligand and an antigenic peptide, new SPR tests will have to verify that CLR targeting will be maintained.

12.5 Final conclusions

The aim of the IMMUNOSHAPE network is to develop lead structures for highly selective glycan based multivalent immunotherapeutics. This ambitious goal requires a high-level expertise in different fields. Being both at the origin of the CLR production and handling the SPR technique that has been intensely used in the network, I had the chance to occupy a central position and participated to many projects.

1. I have achieved the production of several human CLRs and effectively tested them using a panel of glycan/glycomimetic arrays as screening technology. These screening allowed the identification of interesting and, in some cases, unexpected ligand/CLR interactions that will be further analysed.
2. I successfully characterised through biophysical techniques the interactions of glycomimetics designed to selectively bind one specific CLR. Besides, the crystal structure obtained between DC-SIGN and its ultimate optimised ligand corroborated all our computational-based speculations and constituted a very pleasant achievement to this long optimisation.
3. I studied multivalency both at the ligand and protein level:
 - Multivalent scaffolds were individuated as potential candidates for APC targeting in the context of cancer vaccination;
 - Artificial tetrameric proteins termed TETRALEC were developed and structural/functional characterisations were performed. Moreover, I had the possibility during this study to identify a new CLR/pathogen interaction.

In conclusion, interesting potential candidates have been identified and will be explored in the future to address the long-term goal of this multidisciplinary network: to tailor the immune system response.

Part V.
Bibliography

- [1] J. R. Rainville, M. Tsyglakova, and G. E. Hodes, "Deciphering sex differences in the immune system and depression," *Front. Neuroendocrinol.*, Dec. 2017.
- [2] K. M. Yatim and F. G. Lakkis, "A Brief Journey through the Immune System," *Clin. J. Am. Soc. Nephrol.*, vol. 10, no. 7, pp. 1274–1281, Jul. 2015.
- [3] C. A. Janeway and R. Medzhitov, "Innate Immune Recognition," *Annu. Rev. Immunol.*, vol. 20, no. 1, pp. 197–216, Apr. 2002.
- [4] M. D. Turner, B. Nedjai, T. Hurst, and D. J. Pennington, "Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease," *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1843, no. 11, pp. 2563–2582, Nov. 2014.
- [5] H. Herwald and A. Egesten, "The Origin of a Paradigm," *J. Innate Immun.*, vol. 8, no. 3, pp. 221–222, 2016.
- [6] J. M. M. den Haan, R. Arens, and M. C. van Zelm, "The activation of the adaptive immune system: Cross-talk between antigen-presenting cells, T cells and B cells," *Immunol. Lett.*, vol. 162, no. 2, pp. 103–112, Dec. 2014.
- [7] B. Geering, C. Stoeckle, S. Conus, and H.-U. Simon, "Living and dying for inflammation: neutrophils, eosinophils, basophils," *Trends Immunol.*, vol. 34, no. 8, pp. 398–409, Aug. 2013.
- [8] C. Hajjar *et al.*, "The NOX Family of Proteins Is Also Present in Bacteria," *mBio*, vol. 8, no. 6, Dec. 2017.
- [9] S. R. J., L. John, K. R. S., S. Vivek, M. Michael, and C. M. J., "Macrophages: A review of their role in wound healing and their therapeutic use," *Wound Repair Regen.*, vol. 24, no. 4, pp. 613–629, Jul. 2016.
- [10] G. Holleran *et al.*, "The Innate and Adaptive Immune System as Targets for Biologic Therapies in Inflammatory Bowel Disease," *Int. J. Mol. Sci.*, vol. 18, no. 10, p. 2020, Oct. 2017.
- [11] W. E. Paul, "Bridging Innate and Adaptive Immunity," *Cell*, vol. 147, no. 6, pp. 1212–1215, 2011
- [12] S. Maldonado and P. Fitzgerald-Bocarsly, "Antifungal Activity of Plasmacytoid Dendritic Cells and the Impact of Chronic HIV Infection," *Frontiers in Immunology*, vol. 8, p. 1705, 2017.
- [13] D. Dieke, S. D. A., V. Rieneke, K. Yvette, G. T. D., and H. J. M. M., "Targeting C-type lectin receptors: a high-carbohydrate diet for dendritic cells to improve cancer vaccines," *J. Leukoc. Biol.*, vol. 102, no. 4, pp. 1017–1034, Oct. 2017.
- [14] C. Qu, N.-S. Brinck-Jensen, M. Zang, and K. Chen, "Monocyte-derived dendritic cells: targets as potent antigen-presenting cells for the design of vaccines against infectious diseases," *Int. J. Infect. Dis.*, vol. 19, pp. 2013
- [15] C. Qian and X. Cao, "Dendritic cells in the regulation of immunity and inflammation," *Semin. Immunol.*, vol. 35, pp. 3–11, Feb. 2018.
- [16] L. Chen and D. B. Flies, "Molecular mechanisms of T cell co-stimulation and co-inhibition," *Nat. Rev. Immunol.*, vol. 13, no. 4, pp. 227–242, Apr. 2013.
- [17] F. Sallusto, J. Geginat, and A. Lanzavecchia, "Central Memory and Effector Memory T Cell Subsets: Function, Generation, and Maintenance," *Annu. Rev. Immunol.*, vol. 22, no. 1, pp. 745–763, Mar. 2004.
- [18] J. C. Hoving, G. J. Wilson, and G. D. Brown, "Signalling C-Type lectin receptors, microbial recognition and immunity," *Cell. Microbiol.*, vol. 16, no. 2, pp. 185–194, Feb. 2014.
- [19] J. Luban, "Innate immune sensing of HIV-1 by dendritic cells," *Cell Host Microbe*, vol. 12, no. 4, pp.

408–418, Oct. 2012.

- [20] A. Rubartelli and M. T. Lotze, “Inside, outside, upside down: damage-associated molecular-pattern molecules (DAMPs) and redox,” *Trends Immunol.*, vol. 28, no. 10, pp. 429–436, 2007
- [21] T. P. Monie and T. P. Monie, “A Snapshot of the Innate Immune System,” in *The Innate Immune System*, Elsevier, 2017, pp. 1–40.
- [22] P. G. Vallés, A. G. Lorenzo, V. Bocanegra, and R. Vallés, “Acute kidney injury: what part do toll-like receptors play?,” *Int. J. Nephrol. Renovasc. Dis.*, vol. 7, pp. 241–251, Jun. 2014.
- [23] S. M. Levitz and D. T. Golenbock, “Beyond empiricism: Informing vaccine development through innate immunity research,” *Cell*, vol. 148, no. 6, pp. 1284–1292, Mar. 2012.
- [24] L. Zaffaroni and F. Peri, “Recent advances on Toll-like receptor 4 modulation: new therapeutic perspectives,” *Future Med. Chem.*, vol. 10, no. 4, pp. 461–476, 2018.
- [25] T. Kawai and S. Akira, “The roles of TLRs, RLRs and NLRs in pathogen recognition,” *Int. Immunol.*, vol. 21, no. 4, pp. 317–337, Apr. 2009.
- [26] T. B. H. Geijtenbeek and S. I. Gringhuis, “Signalling through C-type lectin receptors: shaping immune responses,” *Nat. Rev. Immunol.*, vol. 9, p. 465, Jul. 2009.
- [27] V. S. J. G. J. J. and K. Yvette, “Dendritic cells and C-type lectin receptors: coupling innate to adaptive immune responses,” *Immunol. Cell Biol.*, vol. 86, no. 7, pp. 580–587, Aug. 2008.
- [28] D. Solís *et al.*, “A guide into glycosciences: How chemistry, biochemistry and biology cooperate to crack the sugar code,” *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1850, no. 1, pp. 186–235, Jan. 2015.
- [29] W. C. BOYD and E. SHAPLEIGH, “Specific Precipitating Activity of Plant Agglutinins (Lectins),” *Science (80-)*, vol. 119, no. 3091, p. 419 LP-419, Mar. 1954.
- [30] M. R. Cummings RD, “C-Type Lectins,” in *Essentials of Glycobiology [Internet]. 3rd edition.*, 2017.
- [31] Y. Feng *et al.*, “Role of selectins and their ligands in human implantation stage,” *Glycobiology*, vol. 27, no. 5, pp. 385–391, May 2017.
- [32] A. S. Powlesland, E. M. Ward, S. K. Sadhu, Y. Guo, M. E. Taylor, and K. Drickamer, “Widely divergent biochemical properties of the complete set of mouse DC-SIGN-related proteins,” *J. Biol. Chem.*, vol. 281, no. 29, pp. 20440–9, Jul. 2006.
- [33] M. R. Leach and D. B. Williams, “Calnexin and Calreticulin, Molecular Chaperones of the Endoplasmic Reticulum BT - Calreticulin: Second Edition,” P. Eggleton and M. Michalak, Eds. Boston, MA: Springer US, 2003, pp. 49–62.
- [34] N. M. Dahms and M. K. Hancock, “P-type lectins,” *Biochim. Biophys. Acta - Gen. Subj.*, vol. 1572, no. 2–3, pp. 317–340, Sep. 2002.
- [35] Z. A. N and G. J. E, “The C-type lectin-like domain superfamily,” *FEBS J.*, vol. 272, no. 24, pp. 6179–6217, Dec. 2005.
- [36] K. Neumann *et al.*, “Clec12a Is an Inhibitory Receptor for Uric Acid Crystals that Regulates Inflammation in Response to Cell Death,” *Immunity*, vol. 40, no. 3, pp. 389–399, 2014
- [37] J. Brown *et al.*, “Structure of the fungal β -glucan-binding immune receptor dectin-1: Implications for function,” *Protein Sci.*, vol. 16, no. 6, pp. 1042–1052, Jun. 2007.
- [38] K. Drickamer and M. E. Taylor, “Recent insights into structures and functions of C-type lectins in the immune system,” *Curr. Opin. Struct. Biol.*, vol. 34, pp. 26–34, Oct. 2015.

- [39] A. Jonas *et al.*, "Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN," *Angew. Chemie Int. Ed.*, vol. 56, no. 25, pp. 7292–7296, May 2017.
- [40] Z. Xiaohui, B. D. F., and M. V. T., "Molecular Basis of the Dynamic Strength of the Sialyl Lewis X—Selectin Interaction," *ChemPhysChem*, vol. 5, no. 2, pp. 175–182, Feb. 2004.
- [41] J. T. Monteiro and B. Lepenies, "Myeloid C-Type Lectin Receptors in Viral Recognition and Antiviral Immunity," *Viruses*, vol. 9, no. 3, p. 59, Mar. 2017.
- [42] G. D. Brown, J. A. Willment, and L. Whitehead, "C-type lectins in immunity and homeostasis," *Nat. Rev. Immunol.*, 2018.
- [43] D. Sancho and C. Reis e Sousa, "Signaling by myeloid C-type lectin receptors in immunity and homeostasis," *Annu. Rev. Immunol.*, vol. 30, pp. 491–529, Jan. 2012.
- [44] O. Gross *et al.*, "Syk kinase signalling couples to the Nlrp3 inflammasome for anti-fungal host defence," *Nature*, vol. 459, p. 433, Apr. 2009.
- [45] M. J. Robinson, D. Sancho, E. C. Slack, S. LeibundGut-Landmann, and C. R. e Sousa, "Myeloid C-type lectins in innate immunity," *Nat. Immunol.*, vol. 7, p. 1258, Dec. 2006.
- [46] J. S. Bonifacino and E. C. Dell; Angelica, "Molecular Bases for the Recognition of Tyrosine-based Sorting Signals," *J. Cell Biol.*, vol. 145, no. 5, p. 923 LP-926, May 1999.
- [47] L. Shen, M. L. Lang, and W. F. Wade, "The ins and outs of getting in: structures and signals that enhance BCR or Fc receptor-mediated antigen presentation," *Immunopharmacology*, vol. 49, no. 3, pp. 227–240, Sep. 2000.
- [48] A. Engering *et al.*, "The Dendritic Cell-Specific Adhesion Receptor DC-SIGN Internalizes Antigen for Presentation to T Cells," *J. Immunol.*, vol. 168, no. 5, p. 2118 LP-2126, Mar. 2002.
- [49] T. SX, M. DL, R. JP, M. Blagojevic, and N. JR, "Epithelial discrimination of commensal and pathogenic *Candida albicans*," *Oral Dis.*, vol. 22, no. S1, pp. 114–119, Apr. 2016.
- [50] S. I. Gringhuis, J. den Dunnen, M. Litjens, B. van het Hof, Y. van Kooyk, and T. B. H. Geijtenbeek, "C-Type Lectin DC-SIGN Modulates Toll-like Receptor Signaling via Raf-1 Kinase-Dependent Acetylation of Transcription Factor NF-kappaB," *Immunity*, vol. 26, no. 5, pp. 605–616, 2007.
- [51] van den B. L. M., G. S. I., and G. T. B.H., "An evolutionary perspective on C-type lectins in infection and immunity," *Ann. N. Y. Acad. Sci.*, vol. 1253, no. 1, pp. 149–158, Jan. 2012.
- [52] N. Masamichi, I. Akemi, K. Yu, M. Naoki, Y. Kazuo, and Y. Yoshiki, "Crystal structures of carbohydrate recognition domain of blood dendritic cell antigen-2 (BDCA2) reveal a common domain-swapped dimer," *Proteins Struct. Funct. Bioinforma.*, vol. 82, no. 7, pp. 1512–1518, Jun. 2014.
- [53] E. Riboldi *et al.*, "Human C-type Lectin Domain Family 4, Member C (CLEC4C/BDCA-2/CD303) Is a Receptor for Asialo-galactosyl-oligosaccharides," *J. Biol. Chem.*, vol. 286, no. 41, pp. 35329–35333, Oct. 2011.
- [54] S. A. F. Jégouzo, H. Feinberg, T. Dungarwalla, K. Drickamer, W. I. Weis, and M. E. Taylor, "A Novel Mechanism for Binding of Galactose-terminated Glycans by the C-type Carbohydrate Recognition Domain in Blood Dendritic Cell Antigen 2," *J. Biol. Chem.*, vol. 290, no. 27, pp. 16759–16771, Jul. 2015.
- [55] N. Kanazawa, K. Tashiro, and Y. Miyachi, "Signaling and immune regulatory role of the dendritic cell immunoreceptor (DCIR) family lectins: DCIR, DCAR, dectin-2 and BDCA-2," *Immunobiology*, vol. 209, no. 1–2, pp. 179–190, Aug. 2004.
- [56] B. Reizis, A. Bunin, H. S. Ghosh, K. L. Lewis, and V. Sisirak, "Plasmacytoid Dendritic Cells: Recent

Progress and Open Questions," *Annu. Rev. Immunol.*, vol. 29, pp. 163–183, 2011.

- [57] G. Tabarani *et al.*, "DC-SIGN Neck Domain Is a pH-sensor Controlling Oligomerization: SAXS AND HYDRODYNAMIC STUDIES OF EXTRACELLULAR DOMAIN," *J. Biol. Chem.*, vol. 284, no. 32, pp. 21229–21240, Aug. 2009.
- [58] T. B. H. Geijtenbeek *et al.*, "Identification of DC-SIGN, a Novel Dendritic Cell Specific ICAM-3 Receptor that Supports Primary Immune Responses," *Cell*, vol. 100, no. 5, pp. 575–585, 2000
- [59] J. J. Garcia-Vallejo and Y. van Kooyk, "The physiological role of DC-SIGN: A tale of mice and men," *Trends Immunol.*, vol. 34, no. 10, pp. 482–486, 2013
- [60] T. B. H. Geijtenbeek, S. J. van Vliet, G. C. F. van Duijnhoven, C. G. Figdor, and Y. van Kooyk, "DC-SIGN, a Dendritic Cell-Specific HIV-1 Receptor Present in Placenta That Infects T Cells In Trans—A Review," *Placenta*, vol. 22, pp. S19–S23, Apr. 2001.
- [61] A. L. Smith, L. Ganesh, K. Leung, J. Jongstra-Bilen, J. Jongstra, and G. J. Nabel, "Leukocyte-specific protein 1 interacts with DC-SIGN and mediates transport of HIV to the proteasome in dendritic cells," *J. Exp. Med.*, vol. 204, no. 2, p. 421 LP-430, Feb. 2007.
- [62] E. J. SOILLEUX, "DC-SIGN (dendritic cell-specific ICAM-grabbing non-integrin) and DC-SIGN-related (DC-SIGNR): friend or foe?," *Clin. Sci.*, vol. 104, no. 4, p. 437 LP-446, Apr. 2003.
- [63] H. Feinberg, D. A. Mitchell, K. Drickamer, and W. I. Weis, "Structural Basis for Selective Recognition of Oligosaccharides by DC-SIGN and DC-SIGNR," *Science (80-.)*, vol. 294, no. 5549, p. 2163 LP-2166, Dec. 2001.
- [64] D. A. Mitchell, A. J. Fadden, and K. Drickamer, "A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands.," *J. Biol. Chem.*, vol. 276, no. 31, pp. 28939–45, Aug. 2001.
- [65] Y. Guo *et al.*, "Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR," *Nat. Struct. & Mol. Biol.*, vol. 11, p. 591, Jun. 2004.
- [66] H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer, and W. I. Weis, "Extended neck regions stabilize tetramers of the receptors DC-SIGN and DC-SIGNR.," *J. Biol. Chem.*, vol. 280, no. 2, pp. 1327–35, Jan. 2005.
- [67] Y. Guo *et al.*, "Dissecting Multivalent Lectin–Carbohydrate Recognition Using Polyvalent Multifunctional Glycan-Quantum Dots," *J. Am. Chem. Soc.*, vol. 139, no. 34, pp. 11833–11844, Aug. 2017.
- [68] R. Celerino da Silva, L. Segat, and S. Crovella, "Role of DC-SIGN and L-SIGN receptors in HIV-1 vertical transmission," *Hum. Immunol.*, vol. 72, no. 4, pp. 305–311, Apr. 2011.
- [69] K. E. A., V. G. K. P. J. M., G. T. B. H., and V. K. Yvette, "Distinct functions of DC-SIGN and its homologues L-SIGN (DC-SIGNR) and mSIGNR1 in pathogen recognition and immune regulation," *Cell. Microbiol.*, vol. 7, no. 2, pp. 157–165, Dec. 2004.
- [70] F. Grünebach, M. M. Weck, J. Reichert, and P. Brossart, "Molecular and functional characterization of human Dectin-1," *Exp. Hematol.*, vol. 30, no. 11, pp. 1309–1315, 2002
- [71] K. A. Heyl *et al.*, "Dectin-1 Is Expressed in Human Lung and Mediates the Proinflammatory Immune Response to Nontypeable Haemophilus influenzae," *MBio*, vol. 5, no. 5, pp. e01492-14, Oct. 2014.
- [72] B. N. Gantner, R. M. Simmons, S. J. Canavera, S. Akira, and D. M. Underhill, "Collaborative Induction of Inflammatory Responses by Dectin-1 and Toll-like Receptor 2," *J. Exp. Med.*, vol. 197, no. 9, pp. 1107–1117, May 2003.

- [73] J. R. Nerren and M. H. Kogut, "The selective Dectin-1 agonist, curdlan, induces an oxidative burst response in chicken heterophils and peripheral blood mononuclear cells," *Vet. Immunol. Immunopathol.*, vol. 127, no. 1–2, pp. 162–166, Jan. 2009.
- [74] B. Ferwerda *et al.*, "Human Dectin-1 Deficiency and Mucocutaneous Fungal Infections," *N. Engl. J. Med.*, vol. 361, no. 18, pp. 1760–1767, Oct. 2009.
- [75] L. M. Graham and G. D. Brown, "The Dectin-2 family of C-type lectins in immunity and homeostasis," *Cytokine*, vol. 48, no. 1–2, pp. 148–155, Oct. 2009.
- [76] K. Sato *et al.*, "Dectin-2 is a pattern recognition receptor for fungi that couples with the Fc receptor gamma chain to induce innate immune responses," *J. Biol. Chem.*, vol. 281, no. 50, pp. 38854–66, Dec. 2006.
- [77] T. Ishikawa *et al.*, "Identification of Distinct Ligands for the C-type Lectin Receptors Mincle and Dectin-2 in the Pathogenic Fungus *Malassezia*," *Cell Host Microbe*, vol. 13, no. 4, pp. 477–488, Apr. 2018.
- [78] H. Feinberg, S. A. F. Jégouzo, M. J. Rex, K. Drickamer, W. I. Weis, and M. E. Taylor, "Mechanism of pathogen recognition by human dectin-2," *J. Biol. Chem.*, vol. 292, no. 32, pp. 13402–13414, Aug. 2017.
- [79] L. Bi *et al.*, "CARD9 mediates dectin-2-induced I κ B kinase ubiquitination leading to activation of NF- κ B in response to stimulation by the hyphal form of *Candida albicans*," *J. Biol. Chem.*, vol. 285, no. 34, pp. 25969–77, Aug. 2010.
- [80] J. C. Muñoz-García *et al.*, "Langerin–Heparin Interaction: Two Binding Sites for Small and Large Ligands As Revealed by a Combination of NMR Spectroscopy and Cross-Linking Mapping Experiments," *J. Am. Chem. Soc.*, vol. 137, no. 12, pp. 4100–4110, Apr. 2015.
- [81] M. Thépaut *et al.*, "Structural Studies of Langerin and Birbeck Granule: A Macromolecular Organization Model," *Biochemistry*, vol. 48, no. 12, pp. 2684–2698, Mar. 2009.
- [82] E. Chabrol *et al.*, "Glycosaminoglycans Are Interactants of Langerin: Comparison with gp120 Highlights an Unexpected Calcium-Independent Binding Mode," *PLoS One*, vol. 7, no. 11, p. e50722, Nov. 2012.
- [83] L. de Witte *et al.*, "Langerin is a natural barrier to HIV-1 transmission by Langerhans cells," *Nat. Med.*, vol. 13, p. 367, Mar. 2007.
- [84] W. Liu *et al.*, "Characterization of a novel C-type lectin-like gene, LSEctin: demonstration of carbohydrate binding and expression in sinusoidal endothelial cells of liver and lymph node," *J. Biol. Chem.*, vol. 279, no. 18, pp. 18748–58, Apr. 2004.
- [85] F. Zhang, S. Ren, and Y. Zuo, "DC-SIGN, DC-SIGNR and LSEctin: C-Type Lectins for Infection," *Int. Rev. Immunol.*, vol. 33, no. 1, pp. 54–66, Jan. 2014.
- [86] A. Dominguez-Soto *et al.*, "The DC-SIGN-related lectin LSEctin mediates antigen capture and pathogen binding by human myeloid cells," *Blood*, vol. 109, no. 12, p. 5337 LP-5345, Jun. 2007.
- [87] D. Zhao *et al.*, "The Myeloid LSEctin Is a DAP12-Coupled Receptor That Is Crucial for Inflammatory Response Induced by Ebola Virus Glycoprotein," *PLoS Pathog.*, vol. 12, no. 3, p. e1005487, Mar. 2016.
- [88] A. S. Powlesland *et al.*, "A novel mechanism for LSEctin binding to Ebola virus surface glycoprotein through truncated glycans," *J. Biol. Chem.*, vol. 283, no. 1, pp. 593–602, Jan. 2008.
- [89] T. Gramberg *et al.*, "Interactions of LSEctin and DC-SIGN/DC-SIGNR with viral ligands: Differential pH dependence, internalization and virion binding," *Virology*, vol. 373, no. 1, pp. 189–201, Mar. 2008.

- [90] A. Furukawa *et al.*, “Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 43, pp. 17438–17443, Oct. 2013.
- [91] Y. Miyake *et al.*, “C-type Lectin MCL Is an FcγR-Coupled Receptor that Mediates the Adjuvanticity of Mycobacterial Cord Factor,” *Immunity*, vol. 38, no. 5, pp. 1050–1062, 2013
- [92] S. J. Williams, “Sensing Lipids with Mincle: Structure and Function,” *Front. Immunol.*, vol. 8, p. 1662, Nov. 2017.
- [93] F. Behler *et al.*, “Macrophage-inducible C-type lectin Mincle-expressing dendritic cells contribute to control of splenic Mycobacterium bovis BCG infection in mice.,” *Infect. Immun.*, vol. 83, no. 1, pp. 184–96, Jan. 2015.
- [94] R. Kiyotake, M. Oh-hora, E. Ishikawa, T. Miyamoto, T. Ishibashi, and S. Yamasaki, “Human Mincle Binds to Cholesterol Crystals and Triggers Innate Immune Responses,” *J. Biol. Chem.*, vol. 290, no. 42, pp. 25322–25332, Oct. 2015.
- [95] N. D. S. Rambaruth, S. A. F. Jégouzo, H. Marlor, M. E. Taylor, and K. Drickamer, “Mouse Mincle: Characterization as a Model for Human Mincle and Evolutionary Implications,” *Molecules*, vol. 20, no. 4, pp. 6670–6682, Apr. 2015.
- [96] H. Feinberg *et al.*, “Mechanism for Recognition of an Unusual Mycobacterial Glycolipid by the Macrophage Receptor Mincle,” *J. Biol. Chem.*, vol. 288, no. 40, pp. 28457–28465, Oct. 2013.
- [97] C. D. Braganza, T. Teunissen, M. S. M. Timmer, and B. L. Stocker, “Identification and Biological Activity of Synthetic Macrophage Inducible C-Type Lectin Ligands,” *Front. Immunol.*, vol. 8, p. 1940, Jan. 2017.
- [98] R. L. Schnaar, “Glycobiology simplified: diverse roles of glycan recognition in inflammation,” *J. Leukoc. Biol.*, vol. 99, no. 6, pp. 825–838, Jun. 2016.
- [99] G. A. Rabinovich, Y. van Kooyk, and B. A. Cobb, “Glycobiology of immune responses,” *Ann. N. Y. Acad. Sci.*, vol. 1253, p. 10.1111/j.1749-6632.2012.06492.x, Apr. 2012.
- [100] J. E. Turnbull and R. A. Field, “Emerging glycomics technologies,” *Nat. Chem. Biol.*, vol. 3, p. 74, Feb. 2007.
- [101] H. Ghazarian, B. Itoni, and S. B. Oppenheimer, “A glycobiology review: carbohydrates, lectins, and implications in cancer therapeutics,” *Acta Histochem.*, vol. 113, no. 3, pp. 236–247, May 2011.
- [102] F. Schwarz and M. Saebi “Mechanisms and principles of N-linked protein glycosylation,” *Curr. Opin. Struct. Biol.*, vol. 21, no. 5, pp. 576–582, Oct. 2011.
- [103] F. Higel, A. Seidl, F. Sörgel, and W. Friess, “N-glycosylation heterogeneity and the influence on structure, function and pharmacokinetics of monoclonal antibodies and Fc fusion proteins,” *Eur. J. Pharm. Biopharm.*, vol. 100, pp. 94–100, 2016.
- [104] H.-J. Gabius, *How to Crack the Sugar Code*, vol. 63. 2017.
- [105] H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, and D. Solís, “From lectin structure to functional glycomics: principles of the sugar code,” *Trends Biochem. Sci.*, vol. 36, no. 6, pp. 298–313, 2011
- [106] P. Gemeiner *et al.*, “Lectinomics: II. A highway to biomedical/clinical diagnostics,” *Biotechnol. Adv.*, vol. 27, no. 1, pp. 1–15, Jan. 2009.
- [107] P. H. Seeberger, “Glycan arrays and other tools produced by automated glycan assembly,” *Perspect. Sci.*, vol. 11, pp. 11–17, Jan. 2017.
- [108] M. Anderluh, “DC-SIGN Antagonists – A Paradigm of C-Type Lectin Binding Inhibition,” C.-F. Chang,

Ed. Rijeka: InTech, 2012, p. Ch. 7.

- [109] T. B. H. Geijtenbeek *et al.*, "DC-SIGN, a Dendritic Cell Specific HIV-1-Binding Protein that Enhances trans-Infection of T Cells," *Cell*, vol. 100, no. 5, pp. 587–597, Mar. 2000.
- [110] A. Tamburrini *et al.*, "Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist," *Bioorg. Med. Chem.*, vol. 25, no. 19, pp. 5142–5147, Oct. 2017.
- [111] V. Porkolab *et al.*, "Rational-Differential Design of Highly Specific Glycomimetic Ligands: Targeting DC-SIGN and Excluding Langerin Recognition," *ACS Chem. Biol.*, vol. 13, no. 3, pp. 600–608, Mar. 2018.
- [112] R. José J. *et al.*, "1,2-Mannobioside Mimic: Synthesis, DC-SIGN Interaction by NMR and Docking, and Antiviral Activity," *ChemMedChem*, vol. 2, no. 7, pp. 1030–1036, Jun. 2007.
- [113] E.-C. Wamhoff *et al.*, "19F NMR-Guided Design of Glycomimetic Langerin Ligands," *ACS Chem. Biol.*, vol. 11, no. 9, pp. 2407–2413, Sep. 2016.
- [114] M. J. Borrok and L. L. Kiessling, "Non-Carbohydrate Inhibitors of the Lectin DC-SIGN," *J. Am. Chem. Soc.*, vol. 129, no. 42, pp. 12780–12785, Oct. 2007.
- [115] K. C. A. Garber, K. Wangkanont, E. E. Carlson, and L. L. Kiessling, "A general glycomimetic strategy yields non-carbohydrate inhibitors of DC-SIGN," *Chem. Commun.*, vol. 46, no. 36, pp. 6747–6749, 2010.
- [116] S. L. Mangold, L. R. Prost, and L. L. Kiessling, "Quinoxalinone Inhibitors of the Lectin DC-SIGN," *Chem. Sci.*, vol. 3, no. 3, pp. 772–777, Mar. 2012.
- [117] R. C. *et al.*, "Principles of Glycan Recognition," in *Essentials of Glycobiology [Internet]. 3rd edition.*, 2017.
- [118] J. M. Antos *et al.*, "Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity," *J. Am. Chem. Soc.*, vol. 131, no. 31, pp. 10800–10801, Aug. 2009.
- [119] B. W. J., D. A. H., C. C. J., R. A. K., S. C. C., and A. K. Ravi, "Molecular features of the sortase enzyme family," *FEBS J.*, vol. 282, no. 11, pp. 2097–2114, Apr. 2015.
- [120] D. A. Levary, R. Parthasarathy, E. T. Boder, and M. E. Ackerman, "Protein-Protein Fusion Catalyzed by Sortase A," *PLoS One*, vol. 6, no. 4, p. e18342, Apr. 2011.
- [121] T. Sijbrandij, N. Cukkemane, K. Nazmi, E. C. I. Veerman, and F. J. Bikker, "Sortase A as a Tool to Functionalize Surfaces," *Bioconjug. Chem.*, vol. 24, no. 5, pp. 828–831, May 2013.
- [122] S. R. Taylor ME, Drickamer K, "Discovery and Classification of Glycan-Binding Proteins," in *Essentials of Glycobiology [Internet]. 3rd edition.*, 2017.
- [123] Y. C. Lee and R. T. Lee, "Carbohydrate-Protein Interactions: Basis of Glycobiology," *Acc. Chem. Res.*, vol. 28, no. 8, pp. 321–327, Aug. 1995.
- [124] L. L. Kiessling, J. E. Gestwicki, and L. E. Strong, "Synthetic multivalent ligands in the exploration of cell-surface interactions," *Curr. Opin. Chem. Biol.*, vol. 4, no. 6, pp. 696–703, Dec. 2000.
- [125] T. Johannssen and B. Lepenies, "Glycan-Based Cell Targeting To Modulate Immune Responses," *Trends Biotechnol.*, vol. 35, no. 4, pp. 334–346, 2017
- [126] T. Georges *et al.*, "Mannose hyperbranched dendritic polymers interact with clustered organization of DC-SIGN and inhibit gp120 binding," *FEBS Lett.*, vol. 580, no. 10, pp. 2402–2408, Mar. 2006.
- [127] A. Berzi *et al.*, "Pseudo-Mannosylated DC-SIGN Ligands as Immunomodulants," *Sci. Rep.*, vol. 6, p. 35373, Oct. 2016.

- [128] D. A. Mitchell *et al.*, "Manipulation of cytokine secretion in human dendritic cells using glycopolymers with picomolar affinity for DC-SIGN" *Chem. Sci.*, vol. 8, no. 10, pp. 6974–6980, Oct. 2017.
- [129] J. Luczkowiak *et al.*, "Glycofullerenes Inhibit Viral Infection," *Biomacromolecules*, vol. 14, no. 2, pp. 431–437, Feb. 2013.
- [130] D. Arosio *et al.*, "Effective Targeting of DC-SIGN by α -Fucosylamide Functionalized Gold Nanoparticles," *Bioconjug. Chem.*, vol. 25, no. 12, pp. 2244–2251, Dec. 2014.
- [131] D. Safari *et al.*, "Gold nanoparticles as carriers for a synthetic Streptococcus pneumoniae type 14 conjugate vaccine," *Nanomedicine*, vol. 7, no. 5, pp. 651–662, May 2012.
- [132] F. Alberto, C. F. Javier, and J. Jesús, "Recent Developments in Synthetic Carbohydrate-Based Diagnostics, Vaccines, and Therapeutics," *Chem. – A Eur. J.*, vol. 21, no. 30, pp. 10616–10628, Jun. 2015.
- [133] M. S. Itano *et al.*, "Super-Resolution Imaging of C-Type Lectin and Influenza Hemagglutinin Nanodomains on Plasma Membranes Using Blink Microscopy," *Biophys. J.*, vol. 102, no. 7, pp. 1534–1542, Apr. 2012.
- [134] K. Brzezicka, U. Vogel, S. Serna, T. Johannssen, B. Lepenies, and N.-C. Reichardt, "Influence of Core β -1,2-Xylosylation on Glycoprotein Recognition by Murine C-type Lectin Receptors and Its Impact on Dendritic Cell Targeting," *ACS Chem. Biol.*, vol. 11, no. 8, pp. 2347–2356, Aug. 2016.
- [135] O. H. Hashim, J. J. Jayapalan, and C.-S. Lee, "Lectins: an effective tool for screening of potential cancer biomarkers," *PeerJ*, vol. 5, p. e3784, Jun. 2017.
- [136] K. Nakajima *et al.*, "Establishment of new predictive markers for distant recurrence of colorectal cancer using lectin microarray analysis," *Cancer Med.*, vol. 4, no. 2, pp. 293–302, Feb. 2015.
- [137] I. Gudelj *et al.*, "Low galactosylation of IgG associates with higher risk for future diagnosis of rheumatoid arthritis during 10 years of follow-up," *Biochim. Biophys. Acta - Mol. Basis Dis.*, vol. 1864, no. 6, pp. 2034–2039, 2018.
- [138] K. Jaroslav, Š. Juraj, G. Peter, K. Tibor, and T. Jan, "Glycan and lectin microarrays for glycomics and medicinal applications," *Med. Res. Rev.*, vol. 30, no. 2, pp. 394–418, Jan. 2010.
- [139] F. Kamena *et al.*, "Synthetic GPI array to study antitoxic malaria response," *Nat. Chem. Biol.*, vol. 4, p. 238, Mar. 2008.
- [140] J. Stevens, O. Blixt, J. C. Paulson, and I. A. Wilson, "Glycan microarray technologies: tools to survey host specificity of influenza viruses," *Nat. Rev. Microbiol.*, vol. 4, p. 857, Oct. 2006.
- [141] A. Zoran *et al.*, "Internalization and Accumulation in Dendritic Cells of a Small pH-Activatable Glycomimetic Fluorescent Probe as Revealed by Spectral Detection," *ChemBioChem*, vol. 16, no. 18, pp. 2660–2667, Nov. 2015.
- [142] C. Bies, C.-M. Lehr, and J. F. Woodley, "Lectin-mediated drug targeting: history and applications," *Adv. Drug Deliv. Rev.*, vol. 56, no. 4, pp. 425–435, Mar. 2004.
- [143] H.-Y. Yu and J. H. Lin, *Intracellular delivery of membrane-impermeable hydrophilic molecules to a hepatoblastoma cell line by asialoglycoprotein-labeled liposomes*, vol. 99. 2000.
- [144] D. H. Dube and C. R. Bertozzi, "Glycans in cancer and inflammation — potential for therapeutics and diagnostics," *Nat. Rev. Drug Discov.*, vol. 4, p. 477, Jun. 2005.
- [145] C. A. Aarnoudse, J. J. G. Vallejo, E. Saeland, and Y. van Kooyk, "Recognition of tumor glycans by antigen-presenting cells," *Curr. Opin. Immunol.*, vol. 18, no. 1, pp. 105–111, Feb. 2006.

- [146] P. J. Tacke, I. J. M. de Vries, R. Torensma, and C. G. Figdor, "Dendritic-cell immunotherapy: from ex vivo loading to in vivo targeting," *Nat. Rev. Immunol.*, vol. 7, p. 790, Oct. 2007.
- [147] P. J. Tacke *et al.*, "Effective induction of naive and recall T-cell responses by targeting antigen to human dendritic cells via a humanized anti-DC-SIGN antibody," *Blood*, vol. 106, no. 4, p. 1278 LP-1285, Aug. 2005.
- [148] I. Moraes and M. Archer, "Methods for the Successful Crystallization of Membrane Proteins BT - Structural Proteomics: High-Throughput Methods," R. J. Owens, Ed. New York, NY: Springer New York, 2015, pp. 211–230.
- [149] K. Brzezicka, B. Echeverria, S. Serna, A. van Diepen, C. H. Hokke, and N.-C. Reichardt, "Synthesis and Microarray-Assisted Binding Studies of Core Xylose and Fucose Containing N-Glycans," *ACS Chem. Biol.*, vol. 10, no. 5, pp. 1290–1302, May 2015.
- [150] "Notizen: Radiative Decay of Non Radiative Surface Plasmons Excited by Light," *Zeitschrift für Naturforschung A*, vol. 23, p. 2135, 1968.
- [151] R. B. M. Schasfoort, "Chapter 1 Introduction to Surface Plasmon Resonance," in *Handbook of Surface Plasmon Resonance (2)*, The Royal Society of Chemistry, 2017, pp. 1–26.
- [152] H. H. Nguyen, J. Park, S. Kang, and M. Kim, "Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications," *Sensors (Basel)*, vol. 15, no. 5, pp. 10481–10510, May 2015.
- [153] P. Schuck and H. Zhao, "The Role of Mass Transport Limitation and Surface Heterogeneity in the Biophysical Characterization of Macromolecular Binding Processes by SPR Biosensing," *Methods Mol. Biol.*, vol. 627, pp. 15–54, 2010.
- [154] S. Cecioni, A. Imbert, and S. Vidal, "Glycomimetics versus Multivalent Glycoconjugates for the Design of High Affinity Lectin Ligands," *Chem. Rev.*, vol. 115, no. 1, pp. 525–561, Jan. 2015.
- [155] W. B. Turnbull and A. H. Daranas, "On the Value of c : Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry?," *J. Am. Chem. Soc.*, vol. 125, no. 48, pp. 14859–14866, Dec. 2003.
- [156] W. D. J., F. M. A., W. M. E., and T. W. Bruce, "Efficient N-Terminal Labeling of Proteins by Use of Sortase," *Angew. Chemie Int. Ed.*, vol. 51, no. 37, pp. 9377–9380, Sep. 2012.
- [157] R. Chen, W. Schmidmayr, C. Krämer, U. Chen-Schmeisser, and U. Henning, "Primary structure of major outer membrane protein II (ompA protein) of *Escherichia coli* K-12," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 77, no. 8, pp. 4592–4596, Aug. 1980.
- [158] W. Guo, L. González-Candelas, and P. E. Kolattukudy, "Cloning of a novel constitutively expressed pectate lyase gene pelB from *Fusarium solani* f. sp. pisi (*Nectria haematococca*, mating type VI) and characterization of the gene product expressed in *Pichia pastoris*," *J. Bacteriol.*, vol. 177, no. 24, pp. 7070–7077, Dec. 1995.
- [159] C. F. Schierle, M. Berkmen, D. Huber, C. Kumamoto, D. Boyd, and J. Beckwith, "The DsbA Signal Sequence Directs Efficient, Cotranslational Export of Passenger Proteins to the *Escherichia coli* Periplasm via the Signal Recognition Particle Pathway," *J. Bacteriol.*, vol. 185, no. 19, pp. 5706–5713, Oct. 2003.
- [160] E. Chabrol *et al.*, "Alteration of the Langerin Oligomerization State Affects Birbeck Granule Formation," *Biophys. J.*, vol. 108, no. 3, pp. 666–677, Feb. 2015.
- [161] M. Schlapschy, S. Grimm, and A. Skerra, "A system for concomitant overexpression of four periplasmic folding catalysts to improve secretory protein production in *Escherichia coli*," *Protein Eng. Des. Sel.*, vol. 19, no. 8, pp. 385–390, Aug. 2006.
- [162] A. Picciocchi *et al.*, "C-Terminal Engineering of CXCL12 and CCL5 Chemokines: Functional

Characterization by Electrophysiological Recordings,” *PLoS One*, vol. 9, no. 1, p. e87394, Jan. 2014.

- [163] V. Norbert *et al.*, “Selective Targeting of Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Nonintegrin (DC-SIGN) with Mannose-Based Glycomimetics: Synthesis and Interaction Studies of Bis(benzylamide) Derivatives of a Pseudomannobioside,” *Chem. – A Eur. J.*, vol. 19, no. 15, pp. 4786–4797, Feb. 2013.
- [164] H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger, and W. I. Weis, “Multiple modes of binding enhance the affinity of DC-SIGN for high mannose N-linked glycans found on viral glycoproteins,” *J. Biol. Chem.*, vol. 282, no. 6, pp. 4202–9, Feb. 2007.
- [165] T. B. H. Geijtenbeek *et al.*, “Identification of different binding sites in the dendritic cell-specific receptor DC-SIGN for intercellular adhesion molecule 3 and HIV-1,” *J. Biol. Chem.*, vol. 277, no. 13, pp. 11314–20, Mar. 2002.
- [166] D. L. Flaviana *et al.*, “The Lipid A from *Rhodopseudomonas palustris* Strain BisA53 LPS Possesses a Unique Structure and Low Immunostimulant Properties,” *Chem. – A Eur. J.*, vol. 23, no. 15, pp. 3637–3647, Dec. 2016.
- [167] A. B. Biter, A. H. de la Peña, R. Thapar, J. Z. Lin, and K. J. Phillips, “DSF Guided Refolding As A Novel Method Of Protein Production,” *Sci. Rep.*, vol. 6, p. 18906, Jan. 2016.
- [168] E. Boeri Erba and C. Petosa, “The emerging role of native mass spectrometry in characterizing the structure and dynamics of macromolecular complexes,” *Protein Sci.*, vol. 24, no. 8, pp. 1176–1192, Aug. 2015.

Part VI.
Annexes

pASK6-ΔOmpA-Strep-GGG-Dectin1 ECD

Gène : human Dectin 1 Numéro genebank AY026769 Type A	Plasmide pASK-IBA6 bacterial expression	Resistance Ampicilline
Date 25/06/15	Boîte Immunoshape II F3	Clone 1
Concentration 34.1 ng/µl		A_{260/280} 1.83
Caractéristiques de la construction 		
Objectifs : Expression cytoplasmique / Fonctionnalisable via une sortase / Gène synthétique, optimisé pour l'expression bactérienne		
Nom VIVES Corinne	Référence cahier Labo Immunoshape p84	Référence séquençage COL15-1FB8
Clonage : Déletion de la séquence OmpA de la construction pAsk6-OmpA-Strep-GGG-Dectin1 ECD, grâce aux oligos SupOmpA Pet R. ATG		
Séquence nucléique annotée : 		
Taille gène 588 pb	Taille construction 3710 pb	
Séquence protéique annotée : 		
Avant coupure Xa 195 AA	MW = 21903.4Da	e =51825 M ⁺ .cm ⁻¹
Après coupure Xa 180 AA	MW = 20118.4 Da	e =46325 M ⁺ .cm ⁻¹

pASK6-ΔOmpA-Strep-GGG-Dectin2 ECD

Gène : human Dectin 2 Numéro genebank AY365135 Type A	Plasmide pASK-IBA6 bacterial expression	Resistance Ampicilline
Date 25/06/15	Boîte Immunoshape II F4	Clone 1
Concentration 41 ng/µl		A_{260/280} 1.81
Caractéristiques de la construction 		
Objectifs : Expression cytoplasmique / Fonctionnalisable via une sortase / Gène synthétique, optimisé pour l'expression bactérienne		
Nom FARHAT Dayana VIVES Corinne	Référence cahier Labo Immunoshape p 84	Référence séquençage COL15-1FB8
Clonage : Déletion de la séquence OmpA de la construction pAsk6-OmpA-Strep-GGG-Dectin2 ECD, grâce aux oligos SupOmpA Pet R. ATG		
Séquence nucléique annotée : 		
Taille gène 561 pb	Taille construction 3683 pb	
Séquence protéique annotée : 		
Avant coupure Xa 186 AA	MW = 21404.8 Da	e =65930 M ⁺ .cm ⁻¹
Après coupure Xa 171 AA	MW = 19619.8 Da	e =60430 M ⁺ .cm ⁻¹

pASK6-OmpA-His-GGG-Dectin1 ECD

Gène : human Dectin 1 Numéro genebank AY026769 Type A		Plasmide pASK-IBA6 bacterial expression		Resistance Ampicilline
Date 16/06/15	Boîte Immunoshape II E3	Clone 1	Concentration 113.8 ng/ μ l	A₂₆₀/280 1.86
Caractéristiques de la construction				
Objectifs : Expression périplasmique / Gène synthétique, optimisé pour l'expression bactérienne				
Nom FARHAT Dayana VIVES Corinne		Référence cahier Labo Immunoshape p 73		Référence séquençage COL15-1E1W
Clonage : Mutation Strep/His avec oligos Strep/His et screen par mutation du site NheI				
Séquence nucléique annotée :				
<p>ATGAAAAGAGATATGGGATGGAGTGGGATGGGTTGGCTTGGCTAGCGTAGCGAGGCGGCTGTAGCAACAGCGGTTCTAACACCTCGGAGAACCGGTTACTTCC ATATCATCATATCGAAGGGCGGGGGCGGCTGAGCAACAGCGGTTCTAACACCTCGGAGAACCGGTTACTTCC TGAGCGTGAACAGAAAGCAGACGCAACCACTGAGCAGCGGAAAGATTCTGCTACTCTAGCAACCGGTTACTTCC TGAAAACACGGGTGTGCTCTTCCATCGCCGGGAAATTGGATTATCTAGGAAAATCTCTGCTAGCAACCGGTTACTTCC CATGAGCCTGAACCTCTGGAGGTTCTAAAGGCCAGTGGCTGGCGCTCTACTCTGCTGAAGATCGATTC AGCAAGAACTGGCTTCTGTGAAACAGGTATCTCCCAAGCGGGAATACTCTTCTGGATGGCGCTCCCGGC CGAAACGAAGTTCGGTGGCTGGGAGAGGCTCTACCTTCTCTAACCTGTTTCAGATTTCAGATTTCGATACACCGG GACGAGAAAACCGCCGCAACTGTGTGGATTCTATGTTCTGTTTCTATCTATGACGACCTGTCTGTACCC GTCTATTCCATCTCGGAGAAAAAATCTCCATGTGA</p>				
Taille gène avec OmpA 642 pb		Taille construction 3764 pb		
Séquence protéique annotée :				
<p>MRSKTAIAVAVALAGFAIVAAQASHHHHHHIEGRGGGNSGNTLENGYFLSRNKENHSQPTQ SSIEDSVPTTKAVKTTGVLSSFPNPNIWIYKESCYLFSMLNSWPDGSKRCQWQLGSLNKIDSS NELGFVKQSSQPDNSFWIGLSRPFQTEVPWLWEDGSTFSSNLFQIRTTAQENPSPNCVWIH VSVIYDQLCSVPYSICEKFKPSM-</p>				
Avec OmpA 213 AA		MW = 23,583.4 Da		pI = 7.67
Sans OmpA 192 AA		MW = 21,554.9 Da		pI = 6.77
Après coupure Xa 180 AA		MW = 20,118.4 Da		pI = 6.14

pASK6-OmpA-His-GGG-Dectin2 ECD

Gène : human Dectin 2 Numéro genebank AY365135 Type A		Plasmide pASK-IBA6 bacterial expression		Resistance Ampicilline
Date 16/06/15	Boîte Immunoshape II E4	Clone 2	Concentration 61.3 ng/ μ l	A₂₆₀/280 1.87
Caractéristiques de la construction				
Objectifs : Expression périplasmique / Gène synthétique, optimisé pour l'expression bactérienne				
Nom FARHAT Dayana VIVES Corinne		Référence cahier Labo Immunoshape p73		Référence séquençage COL15-1E1W
Clonage : Mutation Strep/His avec oligos Strep/His et screen par mutation du site NheI				
Séquence nucléique annotée :				
<p>ATGAAAAGAGATATGGGATGGAGTGGGATGGGTTGGCTTGGCTAGCGTAGCGAGGCGGCTGTAGCAACAGCGGTTCTAACACCTCGGAGAACCGGTTACTTCC ATATCATCATATCGAAGGGCGGGGGCGGCTGAGCAACAGCGGTTCTAACACCTCGGAGAACCGGTTACTTCC AACTGCTAGCTTACCACTCTCCCTGACTTTCACAGGAGGTTACTAAGAGTGGCTGCTGATCGGTTGTGTCCAG TTCTTGGAAAATCTTCGGTAGCTCTTCTACTTCTCAACAGCAGGAAAGAAAGAACTCTGCTGCAAACTCGGACGAGA CTGGTGGAAATCTGGCTGCTGCTGCTTCTTCTCAACAGCAGGAAAGAAAGAACTCTGCTGCAAACTCGGACGCTGA CGAGACTTCTCTACTTCTTGGGCTGCTGACCGCGGCAAGGAATAATACTGGCAGTGGATGGATATAAAACCGG TATGAAAAGAACTTCGCTTTCGACCTGGGAAACCACTCTCTGGGAAACAGCTGGCGCTCCATTTGTTATT TGAACACGAGCTGGCTGGGCTGGAAACGAGTAATCTGGGAAACCGGCTGGTAACCTCCATCTGGGAAATGAACAAA ATCTATCTGTGA</p>				
Taille gène avec OmpA 615 pb		Taille construction 3737 pb		
Séquence protéique annotée :				
<p>MRSKTAIAVAVALAGFAIVAAQASHHHHHHIEGRGGGTYHFYTGEGKRLSELHSHSSLTCFSE GTKVPAWGCFFASWKSFGSFCYFISSEKQNCVEMGAHLVVFNTEAEQNFIVQQLNE SFSYFLGLSDPQGNNNWQWIDKTPYEKNYRFWHLGEPNHSAEQCASIVFWKPTGAWGNVYI GETRNNISCEMINKIVL-</p>				
Avec OmpA 204 AA		MW = 23,084.8 Da		pI = 6.44
Sans OmpA 183 AA		MW = 21,056.3 Da		pI = 6.17
Après coupure Xa 171 AA		MW = 19,619.8 Da		pI = 5.66

pET30b-His-GGG-Mincle CRD

Gène : human Mincle Numéro genebank AB024718		Plasmide pET30b bacterial expression	Resistance Kanamycine
Date 01/10/15		Clone 1	A_{260/280} 1.97
Boîte Immuno III A6		Concentration 45.2 ng/µl	
Caractéristiques de la construction			
			
Objectifs : Expression cytoplasmique / Fonctionnalisable via une sortase / Gène synthétique, optimisé pour l'expression bactérienne			
Nom VIVES	Référence cahier Labo Immuno III p112	Référence séquençage COL15-INFN	
Clonage : Clonage par mutagenèse (déletion du Neck) à partir de la construction pET30-His-GGG-Mincle ECD avec les oligos Mincle CRD F et R (attention non optimisés)			
Séquence nucléique annotée :			
<p>ATC ATCGAAGGGGGGCGGGGGGAAATCGCGATTGAAGTGGCAATATTTAGCT CCAAGCTACTTCTCCAGGAGCAATACAGCTGGGCACTGAGCTGAAAACCTTCGGCTATGGGTGCGCA CCTGGTTGTTATTAACCTCCAGAGAAGAAAGGAAATTTCTGTGCTACAAAAACCAAAATGGGTGATTTCTCAT CGGCGTCTCTGACAGGGTGGTGAAGGCGAGTGGAGTGGTGGAGGGTACCGGCTGACCAAACTCTGCTCTTTT TGGAAATTTGGTGAACCGAATAACATCGGAGCGGCTGGAAGAATGTGGGACCATGGGTGATTTCTAACCGGGG CAAAATTTGAAGAGAGTGAACCTCTTCTGAAACTATTTTTCGCATCTCGGAAATGGTGGGATTAACCGGCTGAC AAAAGCAAAATCTCTTGA</p> <p>En rouge : une cystéine qui aurait pu former un pont disulfure a été remplacé par une sérine En minuscule les bases qui ont été modifiées parce que les oligos ont été basés sur les séquences non optimisées En vert des bases qui auraient dues être changées par l'oligo mais qui restent comme sur la séquence optimisée.</p>			
Taille gène 471 pb		Taille construction 5682 pb	
Séquence protéique annotée :			
<p>M ATHTHTHTEGRGGGNQPLNWEYFQSSCYFFETDITISWALSIKNGSAMGAHLVINSQEEQEF LSYKPKMREFFLGLSDQVYVQWVVDGTLTKLSLFWVDGVEPNNIATLEDCAIMRDSNP RQWVNDVTCFLNFRICEMVGINPLNKGKSL*</p> <p>Avant coupure Xa 156 AA MW= 17911.1 Da ϵ 39335 M¹.cm⁻¹ ϵ 38960 M¹.cm⁻¹ pl= 5.74</p> <p>Après coupure Xa 145 AA MW= 16501.5 Da ϵ 39335 M¹.cm⁻¹ ϵ 38960 M¹.cm⁻¹ pl= 4.77</p>			

pET30b-Strep-GGG-BDCA2 ECD

Gène : human BDCA2 Numéro genebank AF325459		Plasmide pET30b bacterial expression	Resistance Kanamycine
Date 08/09/15		Clone 4	A_{260/280} 1.97
Boîte Immuno shape II D1		Concentration 29.1 ng/µl	
Caractéristiques de la construction			
			
Objectifs : Expression cytoplasmique / Fonctionnalisable via une sortase / Gène synthétique, optimisé pour l'expression bactérienne			
Nom FARHAT Dayana GOITA Aissata	Référence cahier Labo Clonage 30/04/15 p 64 Immuno shape FARHAT Dayana p42 GOITA Aissata p54	Référence séquençage COL15-1AIL, 06/05/15	
Clonage : XbaI/HindIII Ligation dans pET30b du fragment XbaI/HindIII de la construction pASK7plus-Strep-GGG-BDCA2 ECD			
Séquence nucléique annotée :			
<p>ATC ATCGAGCACCGCAATTTGGAAGAAATTCGAAAGGGGGGCGGGGGGCATAAATCTGACTCAAGAGCTCAA GGCTTCATAACTGGCCAAATACAGCAGTACACACCTCTCTGACTTGTGTGATGGAAGTAAAGATGAGGAGCTGGTCT TGTTCGCACTCGTGGACTCTTCCAGACTCTTCTGCTTCTGCTTCACTGACAGCCGGTATGCACTTGGACCAAAAGCCGAA AACTCTGGTGAATGGTCTGATCTGTGTAATCAACACCCGCTGAAGAGCAGGATTTCAATCCAGAACTGAAAGCACA TCCTCTACTTCTGCTGAGCAGTCCGGGGGGGGCTGCAATGGCAATGGGCTTAAACCTGTAACACCCGCTATAATGAAATGTA ACCTTTGGCACTCCGGGAGCAAGCACTGCGAAGGATTTGGCGATTTAACTTTGCTAGGAGGAGAAATGGGCTGG AAGGACATCACTCCACGCTTCCGCAAAATCCATTTGCCAAATGAAAAAGATCTATATC TGA</p>			
Taille gène 567 pb		Taille construction 5647 pb	
Séquence protéique annotée :			
<p>M ATVSHQPTETEGRRGGGHNFMVSKYTKRLSKI REYQQYHPSITCVMEGKDIEDWSSCGPTPWTSFOSS CYFTSGMSQWTKSQKNCYSVMGADLVYINTRREQDFIQNLKRNSSYFLGLSDPGRRRHVQWVDQTPY NENVTFVHSGEPNNLDERCAINFRSSEEWGWNDDHCHVPOKSIKMKKIYI- Avant coupure Xa 188 AA MW= 22046.8 Da ϵ-60430 M¹.cm⁻¹ pl= 7.6</p> <p>Après coupure Xa 173 AA MW= 20261.8 Da ϵ-54930 M¹.cm⁻¹ pl= 7.72</p>			

pEt30b-Strep-GGG-Mincle ECD

Gène : human Mincle Numéro genebank AB024718		Plasmide pEt30b bacterial expression		Resistance Kanamycine
Date 06/01/17 Silvia		Boîte Immunoshape II D6	Clone 2	Concentration A _{260/280} 1.86
Caractéristiques de la construction				
Objectifs : Expression cytoplasmique / Fonctionnalisable via une sortase / Gène synthétique, optimisé pour l'expression bactérienne				
Nom FARHAT Dayana GOITA Aissata		Référence cahier Labo Clonage 30/04/15 p 64 Immunoshape FARHAT Dayana p 42 GOITA Aissata p 54		Référence séquençage COL15-LAIL 06/05/15
Clonage : XbaI/HindIII Ligation dans pEt30b du fragment XbaI/HindIII de la construction pASK7plus-Strep-GGG-Mincle ECD				
Séquence nucléique annotée : ATG ATCGAGGCAACCGCAATTCGAAABAAATCGAAGGGGGGACCCGGTGTGTAGTGCATTCGTAATCTTC GAGAGTTGGGATGAGAGAAAGTTCCAGCTGCTTGAAGACTTCAAGAGCTGAGCTGTTCACACTAGCTAGAGGTTCTGTAA AAATCTGCCCGCTGACTGGGAATATTCACGTCAGCTTCTTCTCGACGGACACTATCAGCTGGGCACTGAGCCT GAAAAAGTTTCGGCTATGAGGTGCCACCTGTTGTTATTAACTCCAAAGAAAGAGAAATTCCTCTACAAAAACCAA AATGGCTGACTTCTTCAATCGCCCTCTGACACAGCTCGGAAAGCCAGTGGCTAGAGCTACCCCGCTGACCAAAATC TCTGCTTTTTGGGATTTGGTGAACGAATAACATCGGACCTGGAAAGATTGTGGCAATTTGTGGCAATTTCTTAACGGGG CCAAAATTGGAAGGACCTGACTGCTTTCTGAACTATTTTCGCAATTTTCGCAAAATGTGGCAATTAACCGCTCAAGAAAGGCA ATCTCTGTA				
Taille gène 597 pb		Taille construction 5677 pb		
Séquence protéique annotée : M MSVSHFQFEKIEGRGGGTRCVVTRFRFQTDKPKQLPENTFELSCYNYGSGVKNCPLNWEYFOSS CYFSTDTISWALSLKNSAMGAHLVINSQEQEFLSYKPKMREFFHGLSDQVVEGQWQWVDGTLT KSLSFWDVGEPNNIATLEDCATMRDSSNRQNWVDVTCFLNYFRICEMVGINPLNKGKSL-				
Avant coupure Xa 198 AA		MW = 22724.6 Da		pi = 5.26
Après coupure Xa 183 AA		MW = 20939.6 Da		pi = 5.06

Supporting Information Paper.1

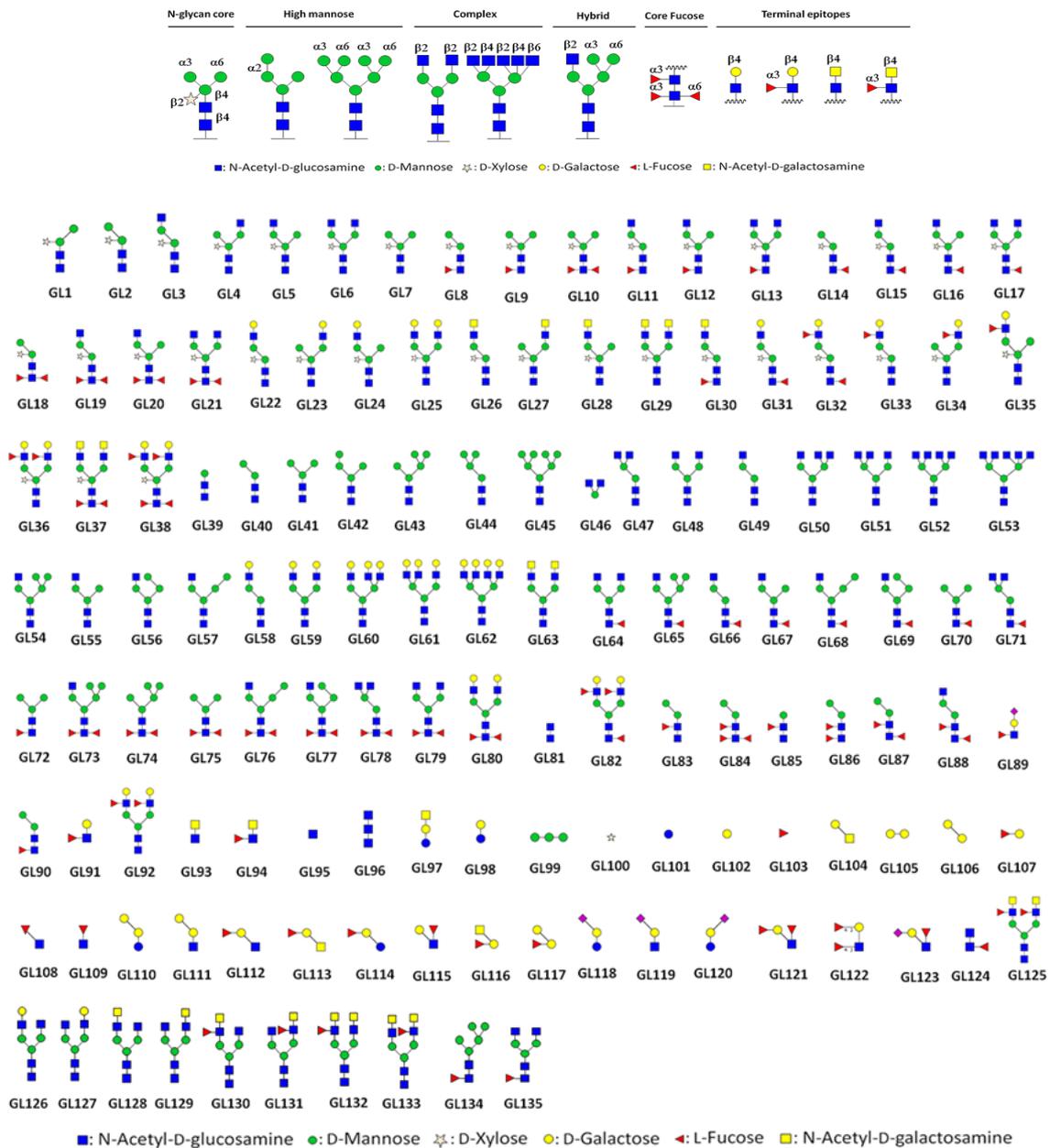
TETRALEC, Artificial Tetrameric Lectins: a tool to screen ligand and pathogen interaction

Silvia Achilli¹, Joao Monteiro², Sonia Serna³, Michel Thepaut¹, Aline Leroy¹, Christine Ebel¹, Bernd Lepenies², Niels Reichardt³, Franck Fieschi¹, Corinne Vivès¹

¹Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

²University of Veterinary Medicine Hannover, Immunology Unit & Research Center for Emerging Infections and Zoonoses (RIZ),

³Glycotechnology laboratory, CIC biomaGUNE, Paseo Miramón 182, 20014, Donostia/San Sebastián, Spain



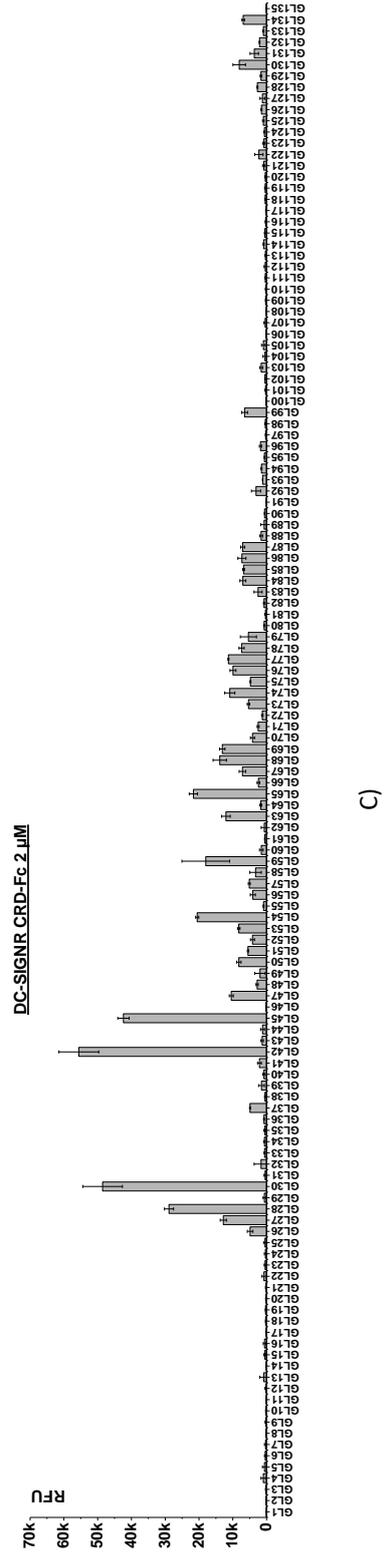
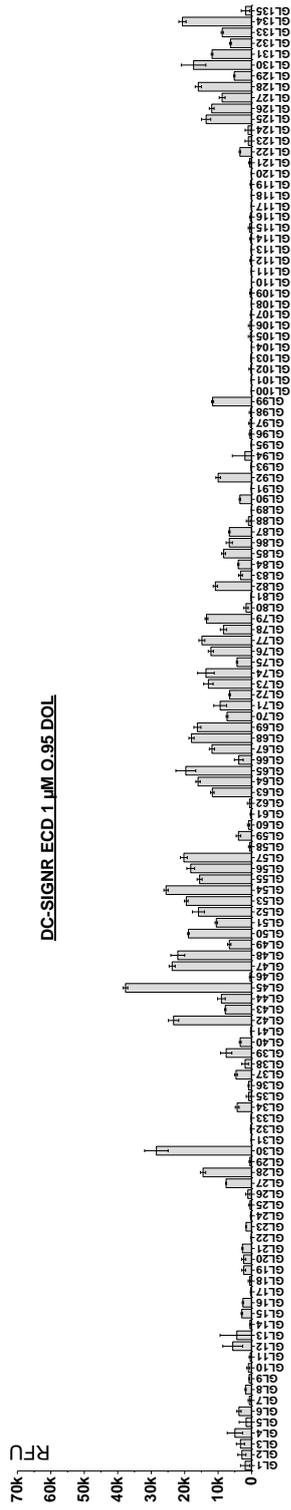
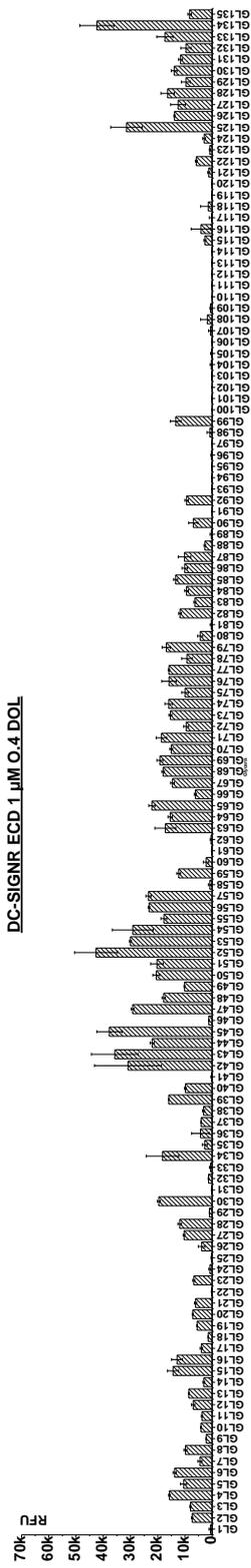
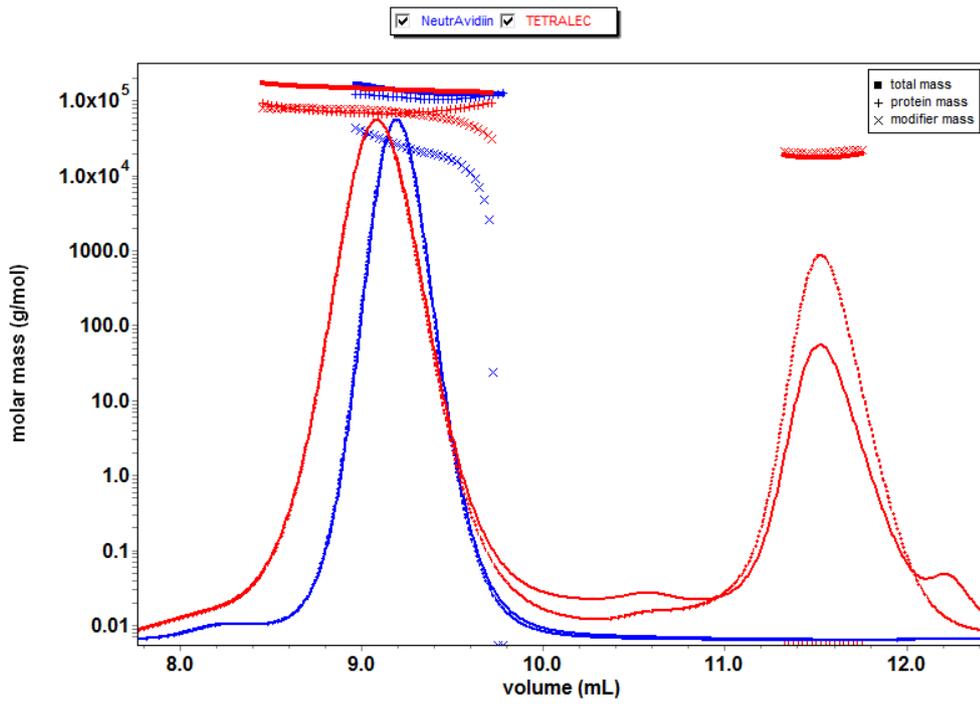
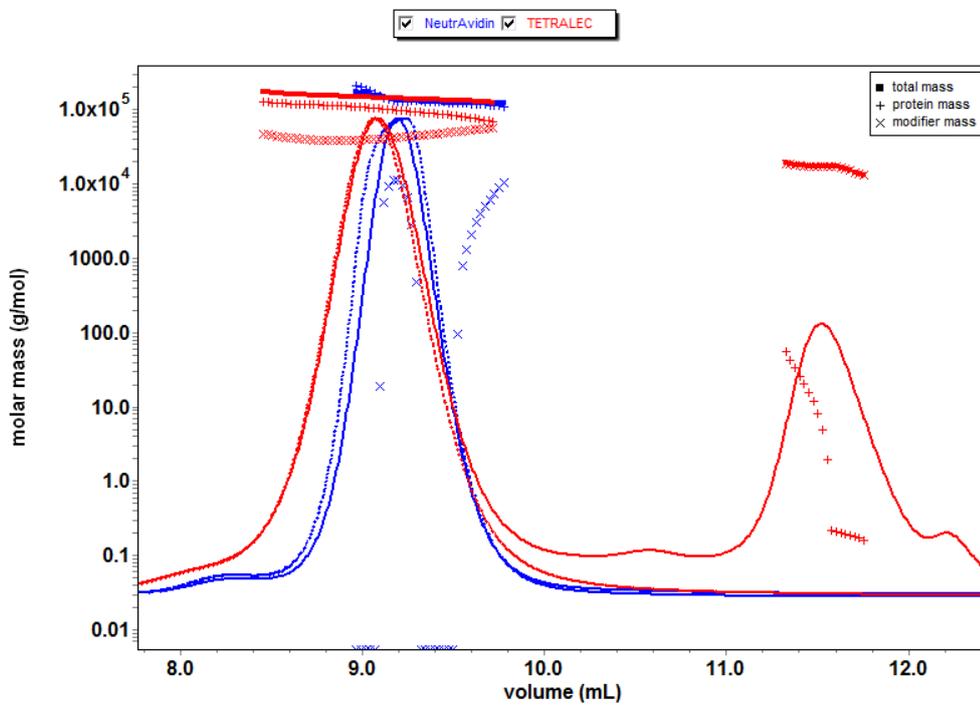


Fig. Incubations on glycan array A) DC-SIGNR ECD 0.4 DOL B) DC-SIGNR ECD 0.95 DOL C) DC-SIGNR CRD-Fc

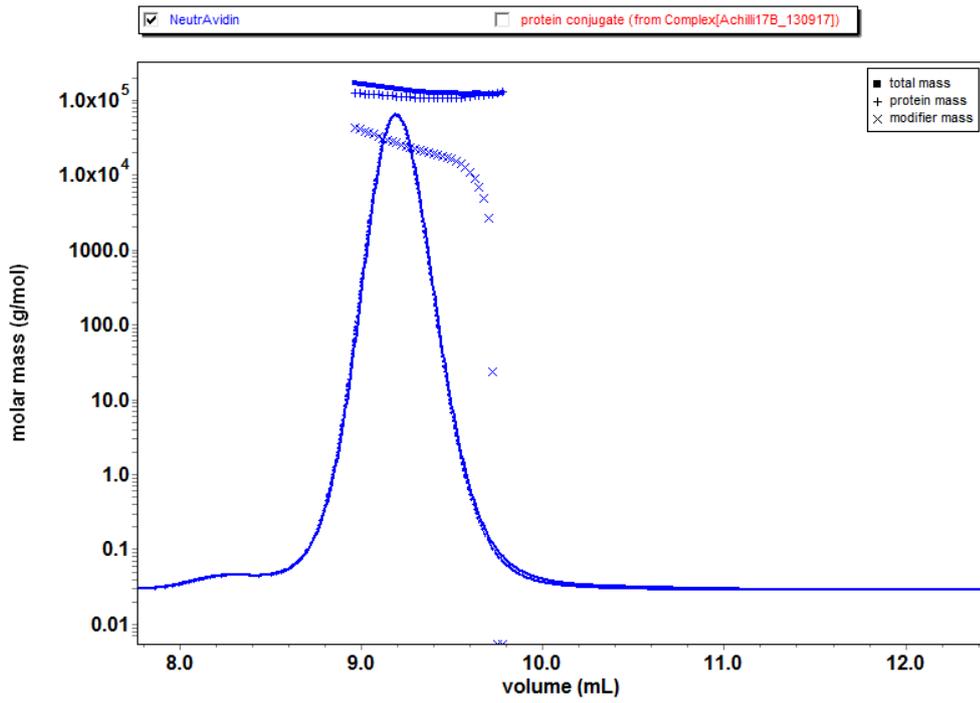
NeutrAvidin +TETRALEC 280 nm



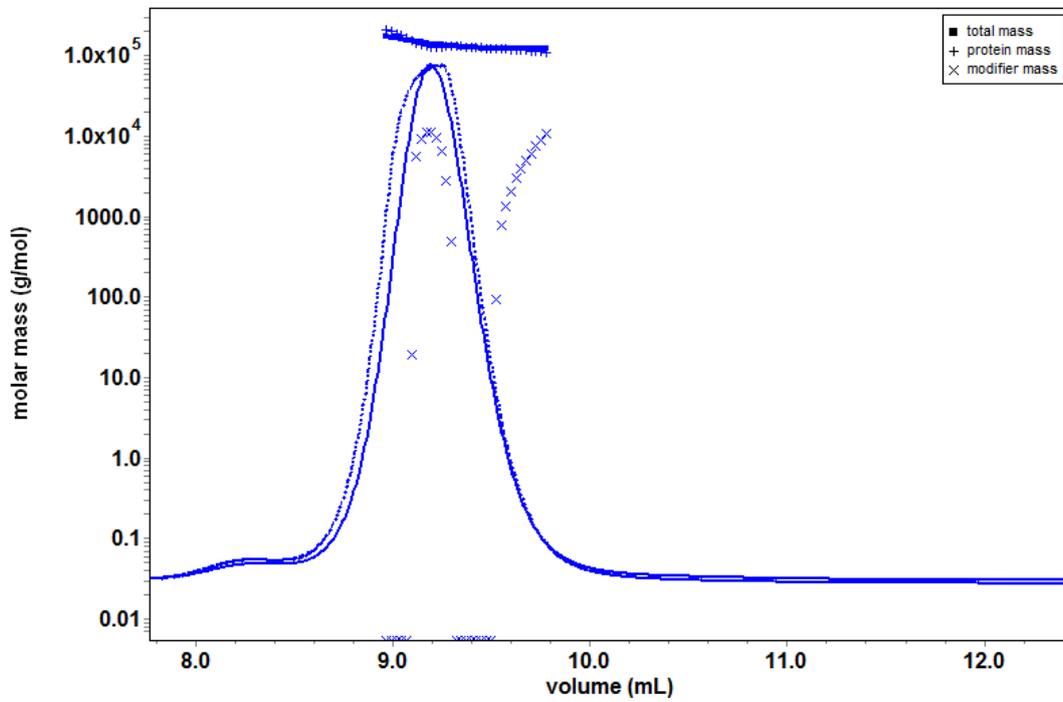
NeutrAvidin +TETRALEC 550 nm



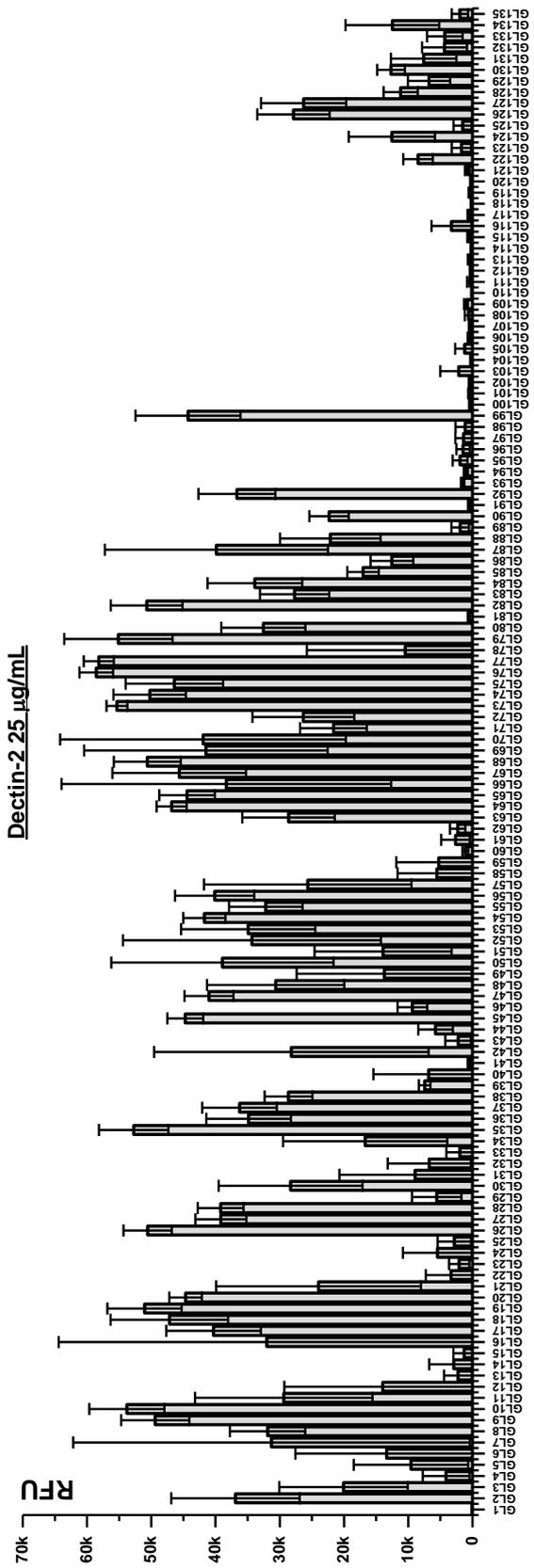
NeutrAvidin 280 nm



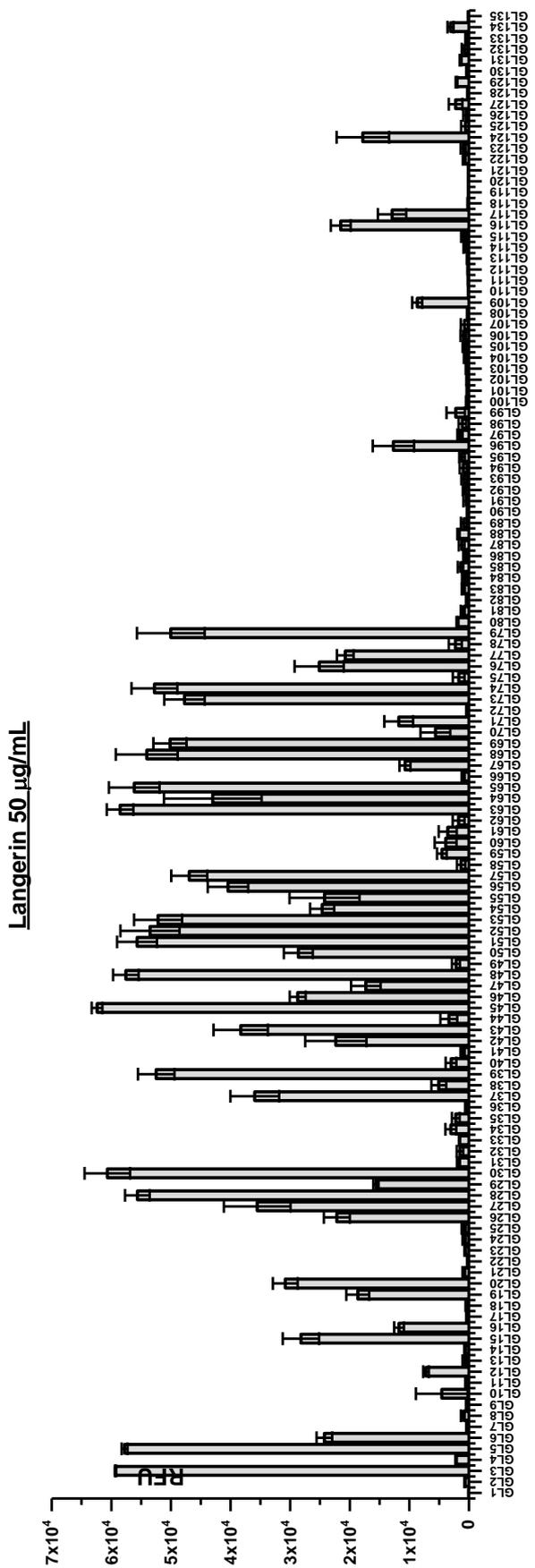
NeutrAvidin 550 nm



Annexes Chapter 9



Glycans



Glycans

Supporting Information Paper.2

Chemoenzymatic Synthesis of N-glycan Positional Isomers and Evidence for Branch Selective Binding by Monoclonal Antibodies and Human C-type Lectin Receptors

Begoña Echeverría, Sonia Serna, Silvia Achilli, Corinne Vivès, Julie Pham, Michel Thépaut, Cornelis H. Hokke, Franck Fieschi, Niels-Christian Reichardt

Chemical Synthesis

Materials. Chemicals were purchased from Sigma-Aldrich or Acros Organics and were used without further purification. All organic solvents were dried over activated 4 Å or 3 Å molecular sieves. Thin layer chromatography was carried out using Merck aluminum sheets silica gel 60 F254 and visualized by UV irradiation (254 nm) or by staining with vanillin solution. Uridine 5'-diphosphogalactose disodium salt (UDP-Gal), uridine 5'-diphospho-2-acetamido-2-deoxy- α -D-galactosamine disodium salt (UDP-GalNAc) and guanosine 5'-diphospho- β -L-fucose sodium salt (GDP-Fuc) were purchased from Carbosynth. All aqueous solutions were prepared from nanopure water produced with a Diamond UV water purification system (Branstead International).

Enzyme expression. pET30a-GalT1 (bovine milk β -1,4-galactosyltransferase) plasmid was kindly provided by Dr. Peter Both from Manchester University. Plasmid was transformed into *E.coli* BL21 StarTM(DE3) One Shot[®] strain from InvitrogenTM and GalT1 enzyme expressed under IPTG induction as previously described.⁴ pET30a-GalT1 plasmid was used as template to generate double mutant (DM) GalT1 (C342T&Y289L mutant)⁵ with QuikChange Site-Directed Mutagenesis kit (Agilent Technologies, CA, USA) according to manufacturer instructions. *In vitro* folding of GalT1 and GalT1 DM from inclusion bodies was achieved following the procedure described by Boeggeman *et al.*⁶ The synthetic gene coding for optimized CeFUT6 sequence lacking the transmembrane domain was assembled by Genscript (Piscataway, NJ, USA) and subcloned in pPICZalphaB vector from Invitrogen. pPICZalphaB-CeFUT6 vector was linearized with *SacI*, subsequently transformed into *Pichia pastoris* X-33 by electroporation. Protein expression and purification was performed as previously described.⁷

Instrumentation. Microwave irradiation was performed on Biotage Initiator monomode oven, (Biotage AB). Hydrogenation reactions were performed in continuous-flow hydrogenation reactor H-Cube from ThalesNano Nanotechnology Inc. Purifications of compounds were performed on: SampliQ high performance graphitized carbon cartridges (1 mL) from Agilent Technologies, C18 Sep-Pak Cartridges (1 mL) from Waters (Milford), flash chromatography using Merck 62 Å 230–400 mesh silica gel or on a Biotage SP4

⁴ Beloqui, A.; Calvo, J.; Serna, S.; Yan, S.; Wilson, I. B. H.; Martin-Lomas, M.; Reichardt, N. C. *Angew. Chemie - Int. Ed.* **2013**, *52*, 7477–7481.

⁵ Ramakrishnan B.; Qasba P. K. *J. Biol. Chem.*, **2002**, *277*, 20833–20839.

⁶ Boeggeman, E. E.; Ramakrishnan, B.; Qasba, P. K. *Protein Expr. Purif.* **2003**, *30*, 219-229.

⁷ Yan, S.; Serna, S.; Reichardt, N.-C.; Paschinger, K.; Wilson, I. B. H. *J. Biol. Chem.*, **2013**, *288*, 21015–21028.

automated flash chromatography system, (Biotage AB) employing prepacked silica cartridges. All aqueous solutions were prepared from nanopure water produced with a Diamond UV water purification system (Branstead International). Pooled glycan containing fractions were lyophilized on an ALPHA-2-4 LSC freeze-dryer from Christ. All organic solvents were concentrated using rotary evaporation. NMR spectra were acquired on a Bruker 500 MHz spectrometer. Chemical shifts were reported in ppm (δ) and referenced to the residual signal of the solvent used (MeOD 4.87 ppm; CDCl₃ 7.26 ppm; D₂O 4.79 ppm). Splitting patterns are designated as s, singlet; d, doublet; t, triplet; m, multiplet. Coupling constants (J) are reported in Hz. High-resolution mass spectra were acquired on a Waters LCT Premier XE instrument, (Waters) equipped with a standard ESI source by direct injection. The instrument was operated with a capillary voltage of 1.0 kV and a cone voltage of 200 V. Cone and desolvation gas flow were set to 50 and 600 L/h, respectively; source and desolvation temperatures were 100 °C. MALDI-TOF mass analyses were performed on an Ultraflextreme III time-of-flight mass spectrometer equipped with a pulsed N₂ laser (337 nm) and controlled by FlexControl 3.3 software (Bruker Daltonics).

HPLC methods:

Compound mixtures obtained after enzymatic elongation were separated by preparative HPLC on a Waters preparative HPLC including a Waters 600 Controller equipped with Waters In-line Degasser AF, Waters 2998 Photodiode Array Detector and Waters Fraction Collector III. Tetrabenzylated derivatives were separated on a Phenomenex Gemini RP C18 10x250mm column with 5 μ m particle size while the pentabenzylated compounds were separated on a Waters XBridge C18 10x100mm column with 5 μ m particle size. The samples were eluted with a flow rate of 4 mL/min, using a volume of injection of 200 μ L and a maximum concentration 20 mg/mL of glycan in H₂O: ACN 8:2.

For purifications one of the following gradients specified in individual experimental protocols was employed

Gradient A: mobile phase (A) Ammonium formate 20 mM / (B) Acetonitrile; gradient: 0-7 at min 65% A, 7-15 min to 60% A, 15-17 at 60% A, 17-18 min to 20% A and 18-21 min at 20% A.

Gradient B: mobile phase (A) 0.1% formic acid in water / (B) Acetonitrile; gradient: 0 min at 75% A, 0-1 min to 70% A, 1-15 min at 70% A, 15-30 min to 65% A, 30-35 min to 20% and 35-40 min at 20% A.

Gradient C: mobile phase (A) Ammonium formate 20 mM / (B) Acetonitrile; gradient: 0 min at 70% A, 0-1 min to 65% A, 1-5 min at 65% A, 5-14 min to 40% A, 14-15 min at 40% A, 15-17 min to 20% and 17-20 min at 20%.

Gradient D: eluents (A) Ammonium formate 20 mM / (B) Acetonitrile; gradient: 0 min at 70% A, 0-1 min to 65% A, 1-5 min at 65% A, 5-14 min to 60% A, 14-15 min at 60% A, 15-17 min to 20% A and 17-20 min 20% A.

Gradient E: eluents (A) Ammonium formate 20 mM / (B) Acetonitrile; gradient: 0 min at 70% A, 0-6 min to 65% A, 6-16 min at 65% A, 16-18 min to 20% A and 18-20 min 20% A.

5-Azidopentyl 3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido- β -D-glucopyranosyl-(1 \rightarrow 2)-O-(3,4,6-tri-O-acetyl- α -D-mannopyranosyl)-(1 \rightarrow 3)-2-O-acetyl-4,6-O-benzylidene- β -D-mannopyranosyl-(1 \rightarrow 4)-O-3,6-di-O-benzyl-2-deoxy-2-phthalimido- β -D-gucopyranosyl-(1 \rightarrow 4)-O-3,6-di-O-benzyl-2-deoxy-2-phthalimido- β -D-

gucopyranoside 5. A solution of **3**⁸ (160 mg, 0.117 mmol) and **4**⁹ (126 mg, 0.141 mmol, 1.2eq) in dry CH₂Cl₂ with molecular sieves was stirred at room temperature for 1h. To this mixture, TMSOTf (3 μL, 0.017 mmol, 15 %) was added and stirred until TLC showed complete conversion of the starting material (1 h). The reaction was quenched by adding triethylamine (20 μL), filtered through a plug of Celite® and the filtrate was concentrated. The crude was purified by flash chromatography obtaining **5** (185 mg, 76%). Rf 0.13 (hexane:EtOAc 1:1); $[\alpha]_D^{20} = -10.5$ (c=0.5, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.92 – 7.61 (m, 12H, Ar), 7.54 (t, *J* = 7.6 Hz, 2H, Ar), 7.45 – 7.36 (m, 4H, Ar), 7.36 – 7.22 (m, 8H, Ar), 6.97 (m, *J* = 17.1, 5.1, 2.3 Hz, 7H, Ar), 6.83 – 6.74 (m, 3H, Ar), 5.49 – 5.42 (m, 2H, H-3E, CHPh), 5.24 (d, *J* = 7.9 Hz, 1H, H-1B), 5.19 (d, *J* = 3.4 Hz, 1H, H-2C), 5.02 (t, *J* = 10.1 Hz, 1H, H-4D), 4.97 (t, *J* = 9.7 Hz, 1H, H-4E), 4.94 – 4.91 (m, 2H, H-1D, H-1A), 4.91 – 4.83 (m, 3H, H-1E, H-3D, CH₂ Bn), 4.81 (d, *J* = 12.4 Hz, 1H, CH₂ Bn), 4.65 (d, *J* = 12.0 Hz, 1H, CH₂ Bn), 4.53 (s, 1H, H-1C), 4.51 (d, *J* = 12.8 Hz, 1H, CH₂ Bn), 4.47 (s, 2H, 2x CH₂ Bn), 4.37 (d, *J* = 12.1 Hz, 1H, CH₂ Bn), 4.34 (d, *J* = 12.1 Hz, 1H, CH₂ Bn), 4.28 – 4.05 (m, 8H, H-2E, H-2A, H-2B, H-6C, H3A, H-3B, H-4A, H-4B), 4.01 (dd, *J* = 3.1, 1.8 Hz, 1H, H-2D), 3.95 (dd, *J* = 12.4, 3.6 Hz, 1H, H-6aE), 3.83 (dt, *J* = 10.0, 3.7 Hz, 1H, H-5D), 3.77 – 3.62 (m, 6H, H-6bE, H-6aD, H-6bD, H-4C, H-6aB, CH₂O), 3.57 (m, *J* = 9.6, 8.3, 3.1 Hz, 2H, H-3C, H-6bB), 3.54 – 3.45 (m, 2H, H-6bC, H-6aA), 3.39 (dd, *J* = 11.0, 3.8 Hz, 1H, H-6bA), 3.33 – 3.27 (m, 1H, H-5A), 3.27 – 3.21 (m, 1H, CH₂O), 3.20 – 3.15 (m, 1H, H-5B), 3.01 (td, *J* = 9.7, 4.9 Hz, 1H, H-5C), 2.91 – 2.79 (m, 2H, CH₂N₃), 2.14 (s, 3H, CH₃ Ac), 2.13 – 2.09 (m, 1H, H-5E), 2.04 (s, 3H, CH₃ Ac), 2.03 (s, 3H, CH₃ Ac), 1.99 (s, 3H, CH₃ Ac), 1.97 (s, 3H, CH₃ Ac), 1.87 (s, 3H, CH₃ Ac), 1.85 (s, 3H, CH₃ Ac), 1.42 – 1.22 (m, 4H, CH₂), 1.13 – 0.98 (m, 2H, CH₂). ¹³C NMR(CDCl₃, 126 MHz): δ (ppm) 170.6, 170.6, 170.5, 170.2, 170.1, 169.5, 169.2, 168.6, 168.5, 168.0, 167.9, 167.7, 167.6, 138.8, 138.7, 138.4, 137.9, 137.4, 134.3, 134.1, 133.9, 133.7, 131.8, 131.7, 131.5, 130.2, 129.0, 128.8, 128.6, 128.3, 128.1, 128.1, 127.9, 127.7, 127.6, 127.3, 127.0, 123.7, 123.7, 123.3, 123.2, 102.4, 98.5, 98.2, 98.0, 97.2, 95.8, 78.8, 78.2, 76.5, 76.1, 75.3, 74.6, 74.5, 74.5, 74.3, 73.5, 72.9, 72.8, 71.1, 70.6, 70.5, 69.4, 68.9, 68.6, 68.4, 68.3, 67.5, 66.1, 65.5, 62.9, 61.1, 56.6, 55.8, 54.1, 51.1, 28.7, 28.3, 23.1, 20.8, 20.7, 20.6, 20.5. HRMS (ESI): *m/z* calcd C₁₀₈H₁₁₂N₆NaO₃₆: 2091.7010 [M+Na]⁺, found 2091.7109.

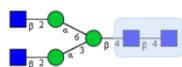
5-azidopentyl 3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→2)-3,4,6-tri-O-acetyl-α-D-mannopyranosyl-(1→3)2-O-acetyl-β-D-mannopyranosyl-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl-(1→4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido-β-D-glucopyranoside 6. To a solution of **5** (185 mg, 0.089 mmol) in CH₂Cl₂ (1 mL) at 0°C, ethanethiol (33 μL, 0.445 mmol, 5 eq) and boron trifluoride diethyl etherate (2 μL, 0.018 mmol 20%) were added. After 2h at room temperature, triethylamine was added. The mixture was concentrated and purified by flash chromatography (hexane: EtOAc, 3:1) obtaining the title compound **6** (140 mg, 79%). Rf 0.1 (hexane:EtOAc 1:2); $[\alpha]_D^{20} = +0.9$ (c= 0.5, CHCl₃); ¹H NMR (500 MHz, CDCl₃) δ 7.95 – 7.47 (m, 12H, Ar), 7.34 – 7.20 (m, 9H, Ar), 7.20 – 7.11 (m, 1H, Ar), 7.03 – 6.92 (m, 7H, Ar), 6.81 – 6.70 (m, 3H, Ar), 5.72 (dd, *J* = 10.8, 9.0 Hz, 1H, H-3E), 5.35 (d, *J* = 8.5 Hz, 1H, H-1E), 5.23 (d, *J* = 7.9 Hz, 1H, H-1B), 5.17 – 5.10 (m, 3H, H-2C, H-4D, H-4E), 4.94 (d, *J* = 2.0 Hz, 1H, H-1D), 4.91 (d, *J* = 7.3 Hz, 1H, H-1A), 4.89 – 4.82 (m, 3H, H-3D, 2x CH₂ Bn), 4.59 (d, *J* = 12.1 Hz, 1H, CH₂ Bn), 4.53 (s, 1H, H-1C), 4.52 – 4.44 (m, 3H, 3x CH₂ Bn), 4.42 – 4.35 (m, 3H, H-2E, CH₂ Bn), 4.28 (dd, *J* = 12.3, 4.8 Hz, 1H, H-6aE), 4.26 – 4.03 (m, 8H, H-2B, H-2A, H-3A, H-3B, H-4A, H-4B, H-2D, H-6bE), 3.86 – 3.71 (m, 5H, H-6aD, H-6bD, H-4C, H-5D, H-5E), 3.71 – 3.58 (m, 3H, CH₂O, H6aB, H6aC), 3.58 – 3.48 (m, 3H, H6bC, H6bB, H6aA), 3.39 (dd, *J* = 11.1, 3.8 Hz, 1H, H-6bA), 3.33 (dd, *J* = 9.4, 3.5 Hz, 1H, H-3C), 3.29 (dd, *J* = 9.1, 2.0 Hz, 1H, H-5A), 3.27 – 3.21 (m, 1H, CH₂O), 3.18 (d, *J* = 9.8 Hz, 1H, H-5B), 2.98 (dt, *J* = 8.8, 4.2 Hz, 1H, H-5C), 2.92 – 2.81 (m, 2H, CH₂N₃), 2.80 (s, 1H, OH), 2.10 (s, 6H, 2x CH₃ Ac), 2.02 (s, 3H, CH₃ Ac), 2.01 (s, 3H, CH₃ Ac), 2.00 (s, 3H, CH₃ Ac), 1.97 (s, 3H, CH₃ Ac), 1.84 (s, 3H, CH₃ Ac), 1.41 – 1.25 (m, 4H, CH₂), 1.10 – 1.00 (m, 2H, CH₂). ¹³C NMR(CDCl₃, 126 MHz): δ

⁸ Serna S., Kardak B., Reichardt N., Martin-Lomas M., *Tetrahedron Asymmetry*, **2009**, *20*, 851-856

⁹ Unverzagt, C.; Eller, S.; Mezzato, S.; Schuberth, R. *Chem. Eur. J.* **2007**, *14*, 1304-1311.

(ppm) 170.9, 170.7, 170.7, 170.2, 170.1, 169.5, 169.5, 168.5, 168.1, 168.0, 168.0, 167.9, 167.7, 138.7, 138.6, 138.4, 137.8, 134.4, 134.2, 134.0, 133.7, 131.8, 131.4, 128.7, 128.3, 128.3, 128.2, 128.1, 128.1, 127.9, 127.5, 127.4, 127.3, 126.9, 123.8, 123.7, 123.2, 123.2, 98.4, 98.1, 97.7, 97.2, 97.2, 77.6, 76.6, 76.1, 75.2, 74.7, 74.6, 74.5, 74.5, 74.4, 73.3, 72.8, 72.0, 70.7, 70.5, 69.9, 69.0, 69.0, 68.9, 68.6, 68.3, 67.3, 65.5, 62.5, 62.2, 62.0, 56.5, 55.8, 54.4, 51.1, 28.7, 28.3, 23.1, 20.9, 20.9, 20.8, 20.7, 20.6, 20.5. HRMS (ESI): m/z : calcd $C_{95}H_{107}N_5NaO_{36}$: 2003.6697 $[M+Na]^+$, found 2003.6598.

5-azidopentyl 3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido- β -D-glucopyranosyl-(1 \rightarrow 2)-3,4,6-tri-O-acetyl- α -D-mannopyranosyl)-(1 \rightarrow 3)-[3,4,6-tri-O-acetyl-2-deoxy-2-phthalimido- β -D-glucopyranosyl-(1 \rightarrow 2)-3,4,6-tri-O-acetyl- α -D-mannopyranosyl-(1 \rightarrow 6)]-2-O-acetyl- β -D-mannopyranosyl-(1 \rightarrow 4)-3,6-di-O-benzyl-2-deoxy-2-phthalimido- β -D-glucopyranoside 7. A solution of **6** (140 mg, 0.071 mmol) and **4**² (80 mg, 0.092 mmol, 1.3 eq) in dry CH_2Cl_2 (14 mL) with molecular sieves was stirred at room temperature for 1h. The mixture was cooled down to $-40^\circ C$, TMSOTf (2 μ L, 15%) was added and stirred at this temperature until TLC showed complete conversion of the starting material (1 h). The reaction was quenched by adding triethylamine (5 μ L), filtered through a plug of Celite[®] and the filtrate was concentrated. The crude was purified by flash chromatography and preparative plate obtaining **7** (100 mg, 52%). Rf 0.28 (hexane:acetone 1:1); $[\alpha]_D^{20} = +2.7$ (c=0.5, $CHCl_3$); 1H NMR (500 MHz, $CDCl_3$) δ 7.92 – 7.53 (m, 16H, Ar), 7.29 – 7.21 (m, 8H, Ar), 7.19 – 7.11 (m, 2H, Ar), 7.01 – 6.90 (m, 4H, Ar), 6.88 – 6.77 (m, 3H, Ar), 6.77 – 6.70 (m, 3H, Ar), 5.70 (dd, $J = 10.8, 9.1$ Hz, 1H, H-3E), 5.66 (dd, $J = 10.9, 9.2$ Hz, 1H, H-3E'), 5.41 (d, $J = 8.5$ Hz, 1H, H-1E), 5.26 – 5.05 (m, 7H, H1E', H-1B, H4D, H-4D', H-3E, H-3E', H-2C), 4.96 (dd, $J = 10.1, 3.3$ Hz, 1H, H-3D'), 4.93 – 4.79 (m, 4H, H-1A, H-1D, H-3D, CH_2 Bn), 4.73 (d, $J = 12.5$ Hz, 1H, CH_2 Bn), 4.61 (d, $J = 12.1$ Hz, 1H, CH_2 Bn), 4.54 (s, 1H, H-1C), 4.50 – 4.25 (m, 10H, 5x CH_2 Bn, H-1D', H-2E, H-2E', H-6aE, H-2D), 4.24 – 4.01 (m, 9H, H6aE', H-6bE, H2D', H3A, H-3B, H-4A, H-4B, H-2A, H-2B), 3.91 – 3.56 (m, 12H, H-6bE', H-6aD, H-6bD, H-6aD', H-6bD', H-5E, H-5D, H-5D', H-4C, H-6aC, H-6aB, CH_2O), 3.55 – 3.41 (m, 3H, H-6bB, H-6aA, H-5E'), 3.40 – 3.19 (m, 5H, H-6bA, CH_2O , H-5A, H-6bC H-3C), 3.17 (dt, $J = 9.9, 2.3$ Hz, 1H, H-5B), 3.12 – 3.00 (m, 1H, H-5C), 2.92 – 2.76 (m, 2H, CH_2N_3), 2.11 (s, 3H, CH_3 Ac), 2.07 (s, 3H, CH_3 Ac), 2.02 (s, 3H, CH_3 Ac), 2.02 (s, 3H, CH_3 Ac), 2.01 (s, 3H, CH_3 Ac), 1.98 (s, 15H, 3x CH_3 Ac), 1.93 (s, 3H, CH_3 Ac), 1.84 (s, 6H, 2x CH_3 Ac), 1.39 – 1.25 (m, 4H, 2x CH_2), 1.10 – 0.99 (m, 2H, CH_2). ^{13}C NMR($CDCl_3$, 126 MHz): δ (ppm) 171.0, 170.8, 170.8, 170.7, 170.4, 170.3, 170.2, 169.5, 169.4, 168.3, 167.6, 138.8, 138.7, 138.4, 138.0, 134.5, 134.1, 133.8, 133.7, 131.8, 131.7, 131.5, 131.4, 129.1, 128.7, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.5, 127.3, 127.0, 123.7, 123.6, 123.2, 99.0, 98.1, 97.9, 97.3, 97.2, 97.1, 78.2, 78.0, 77.0, 76.7, 75.9, 74.6, 74.5, 74.5, 74.4, 73.3, 72.8, 71.8, 71.7, 70.7, 70.7, 70.4, 70.0, 69.4, 69.1, 68.9, 68.5, 68.2, 68.2, 67.3, 65.7, 65.4, 62.5, 62.4, 61.8, 61.6, 56.6, 55.8, 54.5, 51.2, 28.7, 28.3, 23.1, 20.9, 20.9, 20.8, 20.7, 20.7, 20.5. HRMS (ESI): m/z : calcd $C_{133}H_{143}N_7O_{53}Na$: 2708.8602 $[M+Na]^+$, found 2708.8569.



5-azidopentyl 2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside 1. A solution of **7** (32 mg, 11.91 μ mol) in *n*-butanol: ethylene diamine (4:1, 500 μ L) was heated at $120^\circ C$ (3x30 minutes) under microwave irradiation. The mixture was concentrated, co-evaporated with toluene and ethanol and dried under high vacuum overnight. The crude was dissolved in pyridine (1 mL), cooled to $0^\circ C$ and Ac_2O (0.5 mL) and DMAP (1 mg) were added. After overnight reaction at room temperature, the mixture was concentrated and purified by column chromatography in EtOAc:MeOH 95:5, obtaining the crude peracetylated compound. The peracetylated

compound is dissolved in MeOH (2 mL) and 0.5M NaOMe in MeOH was added (50 μ L). After 4 h under reflux, the mixture was cooled down to room temperature and Amberlite® IR120(H) was added until neutral pH. The resulting solution was filtrated and concentrated to dryness obtaining **1** (17.81 mg, 85%). ^1H NMR (500 MHz, MeOD) δ 7.43 – 7.12 (m, 20H, Ph), 5.07 (s, 1H, H-1D), 5.00 (t, J = 12.2 Hz, 2H, CH_2 Bn), 4.79 (d, J = 1.8 Hz, 1H, H1D'), 4.75 (d, J = 12.1 Hz, 1H, CH_2 Bn), 4.71 – 4.54 (m, 5H, 3x CH_2 Bn, H-1C, H-1B), 4.50 – 4.41 (m, 3H, 2x CH_2 Bn, H-1E), 4.39 (d, J = 8.0 Hz, 1H, H-1A), 4.31 (d, J = 8.3 Hz, 1H, H-1E'), 4.11 (d, J = 3.2 Hz, 1H, H-2C), 4.08 (dd, J = 3.2, 1.7 Hz, 1H, H-2D), 4.03 – 3.95 (m, 2H, H-4B, H-4A), 3.93 – 3.53 (m, 27H, H-6aC, H-6bC, H-6aA, H-6bA, H-2A, H-2B, H-4C, H-6aD, H-6bD, H-6aD', H-6bD', H-6aE, H-6bE, H-6aE', H-6bE', H-6aB, H-6bB, CH_2O , H-3D, H-3D', H-5D, H-2D', H-3B, H-3A, H-5D', H-2E, H-2E'), 3.52 – 3.40 (m, 7H, H-4D, H-4D', H-3E, H-3E', H-5A, H-3C, CH_2O), 3.37 – 3.30 (m, 3H, H-4E, H-4E', H-5B), 3.29 – 3.22 (m, 3H, H-5E, CH_2N_3), 3.19 – 3.13 (m, 2H, H-5C, H-5E'), 1.99 (s, 3H, CH_3 Ac), 1.98 (s, 3H, CH_3 Ac), 1.85 (s, 3H, CH_3 Ac), 1.83 (s, 3H, CH_3 Ac), 1.62 – 1.52 (m, 4H, CH_2), 1.47 – 1.37 (m, 2H, CH_2). ^{13}C NMR from HSQC experiment (126MHz, MeOD) δ 127.92, 127.85, 127.7, 127.3, 126.8, 101.2 (C-1A), 100.5 (C-1E'), 100.23 (C-1E), 100.15 (C-1B), 100.1 (C-1C), 99.7 (C-1D), 97.3 (C-1D'), 81.5, 81.0, 80.4, 77.6, 77.1, 76.7, 76.20, 76.15, 75.8, 75.6, 75.0, 74.9, 73.92, 73.86, 73.85, 73.8, 73.2, 72.94, 72.88, 70.5, 70.27, 70.27, 70.2, 68.9, 68.1, 67.9, 65.9, 65.8, 62.0, 61.8, 61.2, 61.1, 55.84, 55.82, 55.7, 54.6, 51.1, 28.5, 23.0, 21.9, 21.6, 21.5. HRMS (MALDI): m/z : calcd $\text{C}_{83}\text{H}_{117}\text{N}_7\text{O}_{36}\text{Na}$: 1810.7432 $[\text{M}+\text{Na}]^+$, found 1810.7581.

β -1,4-galactosylation of 1: A solution (1.65 mL) of **1** (6.09 mg, 3.41 μ mol), uridine 5'-diphospho- α -D-galactose disodium salt UDP-Gal 20 (179 μ L, 3.58 μ mol, 1.05 eq), bovine serum albumin BSA (1 mg), bovine milk β -1,4-galactosyltransferase (100 mU), MnCl_2 (2 mM) and Hepes buffer (50mM, pH=7.4) was incubated at 37°C overnight. The resulting mixture was heated at 95°C for 5 min to precipitate the enzyme. After centrifugation, the supernatant was purified by semipreparative HPLC (C18 10x250 mm 5 μ m, ammonium formate 20 mM:ACN gradient A) and the collected fractions were evaporated and freeze-dried obtaining 1.53 mg (0.726 μ mol, 21%) of compound **10**, 1.20 mg (0.613 μ mol, 18%) of compound **9**, 1.55 mg (0.793 μ mol, 23%) of compound **8** and 1.35 mg (0.754 μ mol, 22%) of compound **1**.



5-azidopentyl

β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-

glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl)-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **8. After β -1,4-galactosylation of **1** (6.09 mg, 3.41 μ mol) and HPLC purification of the crude using gradient A, **8** (1.55 mg, 0.793 μ mol, 23%) was obtained. ^1H NMR (500 MHz, MeOD) δ 7.63 – 7.02 (m, 20H, Ar), 5.07 (s, 1H, H-1D), 5.00 (t, J = 12.1 Hz, 2H, 2x CH_2 Bn), 4.79 (d, J = 1.7 Hz, 1H, H-1D'), 4.75 (d, J = 12.1 Hz, 1H, CH_2 Bn), 4.71 – 4.53 (m, 5H, 3x CH_2 Bn, H-1B, H-1C), 4.50 – 4.33 (m, 5H, 2x CH_2 Bn, H-1E, H-1A, H-1F), 4.30 (d, J = 8.3 Hz, 1H, H-1E'), 4.11 (d, J = 3.2 Hz, 1H, H-2C), 4.08 (dd, J = 3.4, 1.7 Hz, 1H, H-2D), 4.03 – 3.94 (m, 2H, H-4A, H-4B), 3.94 – 3.37 (m, 40H), 3.20 – 3.11 (m, 2H, H-5C, H-5E'), 1.99 (s, 3H, CH_3 Ac), 1.98 (s, 3H, CH_3 Ac), 1.85 (s, 3H, CH_3 Ac), 1.83 (s, 3H, CH_3 Ac), 1.62 – 1.52 (m, 4H, 2x CH_2 linker), 1.47 – 1.37 (m, 2H, CH_2 linker). ^{13}C NMR (from HSQC experiment 126MHz, MeOD) δ 128.0, 127.9, 127.72, 127.70, 127.3, 127.3, 126.7, 103.7, 101.3, 100.4, 100.1, 100.0, 99.6, 97.3, 81.6, 81.0, 80.4, 79.4, 77.7, 77.0, 76.3, 75.9, 75.8, 75.7, 75.4, 75.0, 74.8, 73.93, 73.89, 73.79, 73.76, 73.5, 73.2, 73.1, 72.9, 72.2, 71.2, 70.32, 70.26, 70.20, 70.18, 68.8, 68.4, 68.2, 68.1, 65.8, 65.7, 62.2, 61.9, 61.5, 61.1, 60.3, 55.9, 55.2, 54.6, 51.1, 28.5, 23.0, 21.9, 21.8. HRMS (MALDI): m/z : calcd $\text{C}_{89}\text{H}_{127}\text{N}_7\text{O}_{41}\text{Na}$: 1972.7960 $[\text{M}+\text{Na}]^+$, found 1972.8055.**

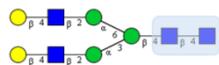


5-azidopentyl

β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-

glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-

glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside 9. After β -1,4-galactosylation of **1** (6.09 mg, 3.41 μ mol) and HPLC purification of the crude using gradient A, **9** (1.20 mg, 0.613 μ mol, 18%) was obtained. ^1H NMR (500 MHz, MeOD) δ 9.00 – 8.69 (m, 20H, Ar), 6.63 (d, J = 1.7 Hz, 1H, H-1D), 6.56 (t, J = 12.2 Hz, 2H, 2xCH₂ Bn), 6.30 (d, J = 12.1 Hz, 1H, CH₂ Bn), 6.26 – 6.10 (m, 5H, 2xCH₂ Bn, H-1D', H-1B, H-1C), 6.06 – 5.91 (m, 5H, 2xCH₂ Bn, H-1E, H-1A, H-1F'), 5.89 (d, J = 8.1 Hz, 1H, H-1E'), 5.67 (d, J = 3.1 Hz, 1H, H-2C), 5.64 (dd, J = 3.2, 1.7 Hz, 1H, H-1D), 5.59 – 5.50 (m, 2H, H-4B, H-4A), 5.50 – 4.96 (m, 41H), 4.76 – 4.68 (m, 1H, H-5C), 3.54 (d, J = 1.4 Hz, 6H, 2xAc), 3.39 (d, J = 5.4 Hz, 6H, 2xAc), 3.12 (m, 4H, 2xCH₂ linker), 2.98 (q, J = 7.6, 7.2 Hz, 2H, CH₂ linker). ^{13}C NMR (from HSQC experiment 126MHz, MeOD) δ 128.0, 127.9, 127.83, 127.75, 127.3, 126.7, 103.7, 101.2, 100.3, 100.12, 100.10, 99.7, 97.1, 81.6, 80.9, 80.5, 79.1, 77.3, 77.1, 76.6, 76.5, 75.73, 75.69, 75.4, 75.0, 74.9, 73.91, 73.90, 73.89, 73.8, 73.5, 73.2, 72.9, 72.9, 72.1, 71.3, 70.7, 70.3, 70.21, 70.18, 68.8, 68.5, 68.3, 68.2, 67.8, 65.9, 65.6, 55.8, 55.7, 55.3, 54.6, 51.1, 28.5, 23.0, 22.0, 21.8. HRMS (MALDI): m/z : calcd C₈₉H₁₂₇N₇O₄₁Na: 1972.7960 [M+Na]⁺, found 1972.7864.



5-azidopentyl

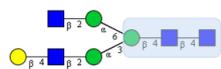
β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-

glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-

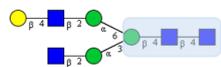
glucopyranoside 10. After β -1,4-galactosylation of **1** (6.09 mg, 3.41 μ mol) and HPLC purification of the crude using gradient A, **10** (1.53 mg, 0.725 μ mol, 21%) was obtained. ^1H NMR (500 MHz, D₂O) δ 7.70 – 7.13 (m, 20H), 5.10 (s, 1H), 4.92 (d, J = 12.2 Hz, 1H), 4.85 (d, J = 12.3 Hz, 1H), 4.69 (t, J = 12.2 Hz, 2H), 4.64 – 4.58 (m, 2H), 4.57 – 4.52 (m, 2H), 4.50 – 4.39 (m, 5H), 4.37 (d, J = 11.7 Hz, 1H), 4.24 (d, J = 7.7 Hz, 1H), 4.20 (d, J = 3.2 Hz, 1H), 4.08 (d, J = 3.1 Hz, 1H), 4.02 – 3.90 (m, 5H), 3.90 – 3.48 (m, 36H), 3.45 (t, J = 9.5 Hz, 1H), 3.39 – 3.33 (m, 1H), 3.29 (t, J = 6.9 Hz, 3H), 3.19 – 3.10 (m, 1H), 2.05 (s, 3H), 2.03 (s, 3H), 1.78 (s, 3H), 1.78 (s, 3H), 1.61 – 1.50 (m, 4H), 1.40 – 1.31 (m, 3H). ^{13}C NMR (from HSQC experiment 126MHz, MeOD) δ 128.8, 128.7, 128.5, 128.2, 102.8, 100.8, 100.1, 99.6, 99.5, 99.4, 96.4, 80.9, 80.3, 80.1, 78.5, 78.0, 76.8, 76.5, 76.3, 75.3, 74.7, 74.3, 74.2, 73.9, 73.84, 73.76, 73.7, 73.3, 73.04, 72.98, 72.6, 72.5, 71.8, 70.7, 70.3, 70.1, 69.7, 68.6, 68.2, 68.1, 67.9, 67.32, 67.27, 65.4, 65.3, 65.3, 61.5, 61.5, 61.0, 59.9, 59.6, 54.9, 54.8, 54.8, 54.6, 51.1, 27.9, 22.4, 22.3, 22.1. HRMS (MALDI): m/z : calcd C₉₅H₁₃₇N₇O₄₆Na: 2134.8488 [M+Na]⁺, found 2134.8472.

β -1,4-galactosylation of 2: A solution (0.5 mL) of **2** (2.13 mg, 1.19 μ mol), uridine 5'-diphospho- α -D-galactose disodium salt UDP-Gal 20 mM (72 μ L, 1.44 μ mol, 1.21 eq), bovine serum albumin BSA (1 mg), bovine milk β -1,4-galactosyltransferase(100 mU), MnCl₂ (2 mM) and Hepes buffer (50 mM, pH=7.4) was incubated at 37°C for 24h. The resulting mixture was heated at 95 °C for 5 min to precipitate the enzyme. After centrifugation, the supernatant was purified by HPLC (C18 10x100 mm, ACN: H₂O 0.1% formic acid, gradient B) and the collected fractions were evaporated and freeze-dried obtaining 720 μ g (0.327 μ mol,

27%) of **13**, 0.302 mg (0.148 μmol , 12%) of **12**, 0.496 mg (0.243 μmol , 20%) of **11** and 0.0508 mg (0.0268 μmol , 2.3%) of compound **2**



5-azidopentyl [(β -D-galactopyranosyl-(1 \rightarrow 4)-2-amino-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl) -(1 \rightarrow 6)]-[2-amino-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)]-2-benzyl- β -D-mannopyranosyl-(1 \rightarrow 4)-2-amino-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **11**. After β -1,4-galactosylation of **2** (2.13 mg, 1.19 μmol) and HPLC purification of the crude using gradient B, **11** (496 μg , 0.243 μmol , 20%) was obtained. ^1H NMR (500 MHz, MeOD) δ 7.42 – 7.11 (m, 25H, Ar), 5.19 (d, J = 1.9 Hz, 1H, H-1D), 5.06 – 4.97 (m, 2H, 2xCH₂ Bn), 4.73 (s, 1H, H-1C), 4.71 (d, J = 1.5 Hz, 1H, H-1D), 4.67 – 4.53 (m, 5H, m, 5H, 4xCH₂ Bn, H-1B), 4.44 (m, 3H, 2xCH₂ Bn, H-1E), 4.39 (d, J = 8.2 Hz, 1H, H-1A), 4.36 (d, J = 7.6 Hz, 1H, H-1F), 4.18 (d, J = 8.3 Hz, 1H, H-E'), 4.11 (dd, J = 3.3, 1.7 Hz, 1H, H-2D), 4.05 – 3.95 (m, 4H, H-4B, H-4A, H-4C, H-2C), 3.95 – 3.88 (m, 3H, H-6Ca, H-6Ea, H-6Da), 3.88 – 3.43 (m, 36H), 3.42 – 3.33 (m, 3H, H-4E', H-3E', H-5E), 3.26 (t, J = 6.8 Hz, 2H, CH₂N₃), 3.22 – 3.14 (m, 2H, H-5B, H-5C), 3.06 – 3.00 (m, 1H, H-5E'), 2.00 (s, 3H, CH₃ Ac), 1.98 (s, 3H, CH₃ Ac), 1.83 (bs, 6H, 2xCH₃ Ac), 1.63 – 1.52 (m, 4H, 2xCH₂ linker), 1.48 – 1.38 (m, 2H, CH₂ linker). ^{13}C NMR (from HSQC experiment 126MHz, MeOD) δ : 129.0, 128.8, 128.5, 128.2, 128.0, 104.8(C-1F), 102.4(C-1A), 101.8(C-1E'), 101.5(C-1C), 101.3(C-1B), 101.1(C-1E'), 100.3(C-1D), 98.5(C-1D'), 81.9, 81.4, 80.5, 79.9, 78.9, 78.0, 77.7, 77.3, 76.8, 76.7, 76.3, 76.2, 76.1, 75.8, 75.4, 75.3, 75.0, 74.9, 74.6, 74.1, 73.9, 73.3, 72.4, 71.4, 71.2, 70.2, 69.9, 69.7, 69.3, 68.7, 68.3, 67.0, 63.1, 62.8, 62.2, 61.9, 61.5, 56.9, 56.2, 55.6, 52.2, 29.7, 29.4, 24.2, 23.1, 22.9. HRMS (MALDI): m/z : calcd C₉₆H₁₃₃N₇NaO₄₁: 2062.8430 [M+Na]⁺, found 2062.8416.



5-azidopentyl [(2-amino-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl) -(1 \rightarrow 6)]-[(β -D-galactopyranosyl-(1 \rightarrow 4)-2-amino-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl)-(1 \rightarrow 3)]-2-benzyl- β -D-mannopyranosyl-(1 \rightarrow 4)-2-amino-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **12**. After β -1,4-galactosylation of **2** (2.13 mg, 1.19 μmol) and HPLC purification of the crude using gradient B, **12** (302 μg , 0.148 μmol , 12%) was obtained. ^1H NMR (500 MHz, MeOD) δ 7.48 – 7.01 (m, 25H, Ar), 5.18 (d, J = 1.9 Hz, 1H, H-1D), 5.01 (t, J = 12.7 Hz, 2H, 2x CH₂ Bn), 4.74 (bs, 1H, H-1C), 4.72 (bs, 1H, H-1D), 4.67 – 4.52 (m, 5H, 4x CH₂ Bn, H-1B), 4.46 – 4.41 (m, 3H, 2xCH₂ Bn, H-1E), 4.39 (d, J = 8.2 Hz, 1H, H-1A), 4.36 (d, J = 7.7 Hz, 1H, H-1F'), 4.20 (d, J = 8.2 Hz, 1H, H-1E'), 4.11 (dd, J = 3.4, 1.7 Hz, 1H, H-2D), 4.05 – 3.88 (m, 6H, H-4B, H-4A, H-4C, H-2C, H-6Ca, H-6Da), 3.88 – 3.41 (m, 38H), 3.28 – 3.22 (m, 3H, H-4E, CH₂N₃), 3.21 – 3.15 (m, 2H, H-5B, H-5C), 3.15 – 3.09 (m, 1H, H-5E'), 2.00 (s, 3H, CH₃ Ac), 1.98 (s, 3H, CH₃ Ac), 1.83 (s, 3H, CH₃ Ac), 1.82 (s, 3H, CH₃ Ac), 1.63 – 1.53 (m, 4H, 2xCH₂ linker), 1.48 – 1.38 (m, 2H, CH₂ linker). ^{13}C NMR (from HSQC experiment 126MHz, MeOD) δ : 129.0, 128.8, 128.7, 128.6, 128.2, 104.7 (C-1F'), 102.2(C-1A), 101.5(C-1C, C-1E'), 101.3 (C-1E), 101.2(C-1B), 100.5 (C-1D), 98.3(C-1D'), 81.5, 81.9, 80.1, 79.7, 78.8, 78.1, 77.9, 77.7, 76.8, 76.7, 76.1, 76.1, 76.0, 75.7, 75.3, 75.2, 75.0, 74.9, 74.7, 74.4, 74.0, 73.9, 72.2, 71.6, 71.5, 71.3, 70.2, 70.1, 69.5, 69.3, 68.8, 68.2, 67.0, 63.2, 62.9, 62.4, 62.4, 61.0, 57.0, 56.5, 56.3, 55.7, 52.2, 29.9, 29.6, 24.4, 23.1, 22.9. HRMS (MALDI): m/z : calcd C₉₆H₁₃₃N₇NaO₄₁: 2062.843 [M+Na]⁺, found 2062.8367.



5-azidopentyl [(β-D-galactopyranosyl-(1→4)-2-amino-2-deoxy-β-D-glucopyranosyl-

(1→2)-α-D-mannopyranosyl)-(1→6)]-[2-amino-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl-(1→3)]-2-benzyl-β-D-mannopyranosyl-(1→4)-2-amino-3,6-di-O-benzyl-2-deoxy-β-D-glucopyranosyl-(1→4)-2-amino-3,6-di-O-benzyl-2-deoxy-β-D-glucopyranoside **13.** After β-1,4-galactosylation of **2** (2.13 mg, 1.19 μmol) and HPLC purification of the crude using gradient B, **13** (720 μg, 0.327 μmol, 27%) was obtained. ¹H NMR (500 MHz, MeOD) δ 7.43 – 7.09 (m, 25H, Ar), 5.19 (s, 1H, H-1D), 5.01 (t, *J* = 13.0 Hz, 2H, 2xCH₂ Bn), 4.73 (s, 1H, H-1C), 4.71 (s, 1H, H-1D'), 4.67 – 4.52 (m, 5H, 4xCH₂ Bn, H-1B), 4.48 – 4.41 (m, 3H, 2xCH₂ Bn, H-1E), 4.41 – 4.34 (m, 3H, H-1A, H-1F, H-1F'), 4.20 (d, *J* = 8.3 Hz, 1H, H-E'), 4.14 – 4.08 (m, 1H, H-2D), 4.05 – 3.88 (m, 7H, H-4A, H-4B, H-2D', H-2C), 3.88 – 3.43 (m, 44H), 3.39 (s, 1H, H-5E'), 3.26 (t, *J* = 6.8 Hz, 2H, CH₂N₃), 3.21 – 3.15 (m, 2H, H-5B, H-5C), 3.15 – 3.09 (m, 1H, H-5E'), 1.99 (s, 3H, CH₃ Ac), 1.98 (s, 3H, CH₃ Ac), 1.83 (s, 3H, CH₃ Ac), 1.82 (s, 3H, CH₃ Ac), 1.63 – 1.52 (m, 4H, 2xCH₂ linker), 1.48 – 1.38 (m, 2H, CH₂ linker). ¹³C NMR (from HSQC experiment 126MHz, MeOD) δ: 129.1, 128.7, 128.5, 128.5, 128.1, 128.1, 104.7(C-1F, C-1F'), 102.3(C-1A), 101.4(C-1C, C-1E'), 101.2(C-1B), 101.0(C-1E), 100.3(C-1D'), 98.2, 81.8, 81.7, 79.9, 79.8, 78.7, 78.0, 77.7, 76.8, 76.7, 76.3, 76.0, 76.0, 75.7, 75.3, 75.0, 74.9, 74.1, 73.9, 73.2, 72.3, 71.2, 70.1, 70.0, 69.8, 69.3, 68.8, 68.2, 67.0, 63.0, 62.8, 62.4, 62.2, 61.4, 61.1, 56.4, 56.1, 56.1, 55.7, 29.6, 29.5, 24.1, 23.0, 22.8. HRMS (MALDI): *m/z*: calcd C₁₀₂H₁₄₃N₇NaO₄₆: 2224.8958 [M+Na]⁺, found 2224.8909.

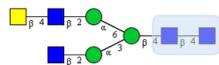
β-1,4-galactosamination of 1: A solution (1 mL) of **1** (5.77 mg, 3.22 μmol), uridine 5'-diphospho-2-acetamido-2-deoxy-α-D-galactosamine disodium salt UDP-GalNAc 20mM (178 μL, 3.22 μmol, 1.1 eq), double mutant β-1,4-galactosyltransferase (400 μL), MnCl₂ (10 mM) in HEPES buffer (50mM, pH=7.4) was incubated at 37°C for 44h. The resulting mixture was heated at 95°C for 5 min to precipitate the enzyme. After centrifugation the supernatant was purified by semi-preparative HPLC (ACN:H₂O, C18 10x250 mm, 5μm, gradient C). The fractions were concentrated and freeze-dried, obtaining 1.89 mg (0.861 μmol, 26%) of compound **16**, 1.67 mg (0.837 μmol, 26 %) of compound **15**, 1.48 mg (0.742 μmol, 23%) of compound **14** and 1.118 mg (0.625 μmol, 19%) of compound **1**



5-azidopentyl 2-acetamido-2-deoxy-β-D-galactopyranosyl-(1→4)-2-acetamido-2-

deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl-(1→3)-[2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→2)-α-D-mannopyranosyl-(1→6)]-β-D-mannopyranosyl-(1→4)-2-acetamido-3,6-di-O-benzyl-2-deoxy-β-D-glucopyranosyl-(1→4)- 2-acetamido-3,6-di-O-benzyl-2-deoxy-β-D-glucopyranoside **14.** After β-1,4-galactosamination of **1** (5.77 mg, 3.22 μmol) and HPLC purification of the crude using gradient C, **14** (1.48 mg, 0.742 μmol, 23%) was obtained. ¹H NMR (500 MHz, MeOD) δ 7.43 – 7.14 (m, 20H, Ph), 5.07 (s, 1H, H-1D), 5.04 – 4.96 (m, 2H, 2x CH₂ Bn), 4.79 (d, *J* = 1.8 Hz, 1H, H-1D'), 4.74 (d, *J* = 12.1 Hz, 1H, CH₂ Bn), 4.70 – 4.55 (m, 5H, 3x CH₂ Bn, H-1B, H-1C), 4.50 – 4.37 (m, 5H, 2x CH₂ Bn, H-1F, H-1E, H-1A), 4.31 (d, *J* = 8.3 Hz, 1H, H-1E'), 4.10 (d, *J* = 3.1 Hz, 1H, H-2C), 4.07 (dd, *J* = 3.4, 1.7 Hz, 1H, H-2D), 4.02 – 3.94 (m, 3H, H-2F, H-4B, H-4A), 3.93 – 3.41 (m, 40H), 3.25 (t, *J* = 6.8 Hz, 2H, CH₂N₃), 3.19 – 3.12 (m, 2H, H-5C, H-5E'), 2.01 (s, 3H, CH₃ Ac), 1.99 (s, 3H, CH₃ Ac), 1.97 (s, 3H, CH₃ Ac), 1.85 (s, 3H, CH₃ Ac), 1.83 (s, 3H, CH₃ Ac), 1.61 – 1.53 (m, 4H, 2xCH₂), 1.47 – 1.37 (m, 2H, CH₂). ¹³C NMR from HSQC experiment (126 MHz, MeOD) δ 127.8, 127.70, 127.65, 127.3, 126.7, 102.1 (C-1F), 101.1 (C-1A), 100.6, 100.1, 100.0, 99.4 (C-1D), 97.3 (C-1D'), 81.6, 80.9, 80.4, 80.0, 77.65, 76.9, 76.34, 76.0, 75.9, 75.8, 75.3, 75.1, 75.0, 74.8, 73.85, 73.82, 73.81, 73.79, 73.7, 73.3, 72.90, 72.89, 71.5, 70.3, 70.1, 68.9, 68.4, 68.3, 68.2, 68.0, 65.9, 65.8, 62.1, 61.4, 60.74, 60.66, 60.3,

55.9, 55.8, 54.8, 54.6, 52.9 (C-2F), 51.1, 28.5, 23.1, 21.94, 21.86, 21.72. HRMS (MALDI): m/z : calcd $C_{91}H_{130}N_8O_{41}Na$: 2013.8226 $[M+Na]^+$, found 2013.8300.

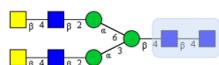


5-azidopentyl

2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-

mannopyranosyl-(1 \rightarrow 3)- [2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)- 2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside

15. After β -1,4-galactosamination of **1** (5.77 mg, 3.22 μ mol) and HPLC purification of the crude using gradient C, **15** (1.668 mg, 0.837 μ mol, 26 %) was obtained. 1H NMR (500 MHz, MeOD) δ 7.45 – 7.10 (m, 20H, Ph), 5.07 (d, J = 1.5 Hz, 1H, H-1D), 5.01 – 4.96 (m, 2H, 2x CH_2 Bn), 4.78 (d, J = 1.0 Hz, 1H, H-1D'), 4.74 (d, J = 12.3 Hz, 1H, CH_2 Bn), 4.69 – 4.54 (m, 5H, 3x CH_2 Bn, H-1B, H-1C), 4.49 – 4.41 (m, 4H, 2x CH_2 Bn, H-1E, H-1F'), 4.39 (d, J = 8.2 Hz, 1H, H-1A), 4.23 (d, J = 8.2 Hz, 1H, H-1E'), 4.11 (d, J = 3.1 Hz, 1H, H-2C), 4.08 (dd, J = 3.4, 1.7 Hz, 1H, H-2D), 4.04 – 3.95 (m, 3H, H-2F', H-4B, H-4A), 3.94 – 3.42 (m, 40H), 3.28 – 3.22 (m, 4H, CH_2N_3 , H-5B, H-5E), 3.20 – 3.14 (m, 1H, H-5C), 3.11 – 3.05 (m, 1H, H-5E'), 2.04 (s, 3H, CH_3 Ac), 1.98 (s, 6H, 2x CH_3 Ac), 1.83 (s, 6H, 2x CH_3 Ac), 1.62 – 1.53 (m, 4H, 2x CH_2), 1.48 – 1.37 (m, 2H, CH_2). ^{13}C NMR from HSQC experiment (126 MHz, MeOD) δ 127.81, 127.77, 127.5, 127.3, 126.7, 102.1 (C-1F), 101.2 (C-1A), 100.3, 100.22, 100.15, 99.7 (C-1D), 97.1 (C-1D'), 81.6 (C-3C), 80.9, 80.4, 79.7, 77.2, 77.1, 76.6, 75.8, 75.7, 75.4, 75.0, 74.9, 74.6, 74.0, 73.83, 73.80, 73.77, 73.2, 72.92, 72.85, 71.7, 70.5, 70.3, 70.2, 68.8, 68.6, 68.3, 68.2, 67.9, 65.79, 65.76, 62.0, 61.5, 61.3, 61.0, 59.9, 55.82, 55.79, 54.9, 54.6, 53.0 (C-2F'), 51.0, 28.5, 23.1, 21.96, 21.78, 21.75. HRMS (MALDI): m/z : calcd $C_{91}H_{130}N_8O_{41}Na$: 2013.8226 $[M+Na]^+$, found 2013.8292.

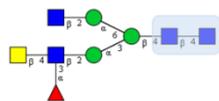


5-azidopentyl

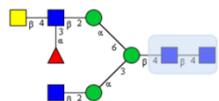
2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-

deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)- 2-

16. After β -1,4-galactosamination of **1** (5.77 mg, 3.22 μ mol) and HPLC purification of the crude using gradient C, **16** (1.89 mg, 0.861 μ mol, 26%) was obtained. 1H NMR (500 MHz, MeOD) δ 7.49 – 7.07 (m, 20H, Ph), 5.08 (s, 1H, H-1D), 4.98 (d, J = 12.2 Hz, 2H, 2x CH_2 Bn), 4.80 (s, 1H, H-1D), 4.74 (d, J = 12.2 Hz, 1H, CH_2 Bn), 4.69 – 4.54 (m, 5H, 3x CH_2 Bn, H-1B, H-1C), 4.49 – 4.41 (m, 5H, CH_2 Bn, H-1F, H-1F', H-1E), 4.39 (d, J = 8.1 Hz, 1H, H-1A), 4.25 (d, J = 8.2 Hz, 1H, H-1E'), 4.10 (d, J = 3.1 Hz, 1H, H-2C), 4.07 (dd, J = 3.3, 1.7 Hz, 1H, H-2D), 4.05 – 3.94 (m, 4H, H-2F, H-2F', H-4B, H-4A), 3.94 – 3.43 (m, 46H), 3.36 – 3.32 (m, 1H, H-5E), 3.29 – 3.22 (m, 3H, CH_2N_3 , H-5B), 3.20 – 3.14 (m, 1H, H-5C), 3.13 – 3.05 (m, 1H, H-5E'), 2.03 (s, 3H, CH_3 Ac), 2.01 (s, 3H, CH_3 Ac), 1.98 (s, 3H, CH_3 Ac), 1.97 (s, 3H, CH_3 Ac), 1.83 (d, J = 1.4 Hz, 6H 2x CH_3 Ac), 1.62 – 1.51 (m, 4H, 2x CH_2), 1.49 – 1.37 (m, 2H, CH_2). ^{13}C NMR from HSQC experiment (126 MHz, MeOD) δ 127.76, 127.74, 127.68, 127.3, 126.9, 102.2, 101.1, 100.2, 100.1, 99.9, 99.4, 97.0, 81.5, 81.0, 80.4, 79.9, 77.2, 76.8, 76.6, 75.8, 75.40, 75.04, 74.99, 74.8, 74.6, 73.86, 73.85, 73.8, 73.2, 73.1, 72.9, 72.0, 71.7, 70.4, 70.2, 70.0, 68.9, 68.6, 68.3, 68.2, 67.9, 65.9, 65.7, 62.1, 61.2, 61.0, 60.1, 55.9, 54.84, 54.81, 54.5, 52.9, 52.8, 51.1, 28.56, 23.0, 22.0, 21.80, 21.77. HRMS (MALDI): m/z : calcd $C_{99}H_{143}N_9O_{46}Na$: 2216.9020 $[M+Na]^+$, found 2216.9104.



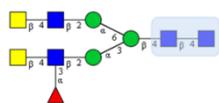
5-azidopentyl α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **17.** A solution (0.5 mL) of **14** (1.61 mg, 0.808 μ mol), guanosine 5'-diphospho- β -L-fucose sodium salt GDP-Fucose 20 mM (118.5 μ L, 2.35 μ mol, 2.9 eq), α -1,3-fucosyltransferase CeFUT6 (110 μ L), MnCl₂ (20 mM) in MES 80mM buffer pH 6.5 was incubated at room temperature for 48h. The resulting mixture was heated at 95°C for 5 min to precipitate the enzyme. After centrifugation the supernatant was purified by semipreparative HPLC (C18 10x250 mm, 5 μ m, gradient D, Ammonium formate 20 mM:ACN) and the collected fractions were evaporated and freeze-dried obtaining 1.04 mg of **17** (60%). ¹H NMR (500 MHz, MeOD) δ 7.43 – 7.12 (m, 20H, Ph), 5.07 (t, J = 2.9 Hz, 2H, H-1G, H-1D), 5.00 (t, J = 11.5 Hz, 2H, 2x CH₂ Bn), 4.80 – 4.72 (m, 3H, CH₂ Bn, H-5G, H-1D'), 4.70 – 4.54 (m, 5H, 3x CH₂ Bn, H-1B, H-1C), 4.50 – 4.41 (m, 4H, 2x CH₂ Bn, H-1E, H-1F), 4.39 (d, J = 8.2 Hz, 1H, H-1A), 4.31 (d, J = 8.3 Hz, 1H, H-1E'), 4.13 – 4.09 (m, 1H, H-2C), 4.05 (dd, J = 3.4, 1.7 Hz, 1H, H-2D), 4.03 – 3.51 (m, 43H), 3.51 – 3.40 (m, 6H), 3.25 (t, J = 6.9 Hz, 3H, CH₂N₃, H-5B), 3.19 – 3.12 (m, 2H, H-5', H-5C), 1.99 (s, 3H, CH₃ Ac), 1.97 (d, J = 1.2 Hz, 6H, 2xCH₃ Ac), 1.85 (s, 3H, CH₃ Ac), 1.83 (s, 3H, CH₃ Ac), 1.60 – 1.54 (m, 4H, 2xCH₂), 1.42 (td, J = 8.0, 7.6, 5.5 Hz, 2H, CH₂), 1.27 (d, J = 6.5 Hz, 3H, H-6G). ¹³C NMR from HSQC experiment (126 MHz, MeOD) δ 129.3, 129.1, 128.8, 128.6, 128.3, 102.5, 101.8, 101.7, 101.3, 101.3, 100.8, 100.7, 99.9, 98.6, 82.7, 82.1, 81.6, 78.8, 77.9, 77.4, 77.1, 77.0, 76.9, 76.6, 76.2, 76.1, 75.2, 75.2, 75.07, 75.06, 74.9, 74.44, 74.42, 74.13, 74.12, 73.6, 72.8, 71.6, 71.49, 71.47, 71.1, 71.0, 70.4, 69.9, 69.77, 69.75, 69.4, 67.7, 67.0, 66.9, 63.4, 63.0, 62.3, 62.0, 61.5, 57.1, 57.0, 56.8, 55.8, 54.2, 52.3, 49.0, 29.7, 24.3, 23.1, 22.9, 16.5. HRMS (MALDI): m/z : calcd C₉₇H₁₄₀N₈NaO₄₅: 2159.8805 [M+Na]⁺, found 2159.8953.



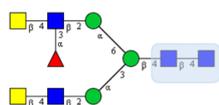
5-azidopentyl 2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)- { α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **18.** A solution (0.5 mL) of **15** (1.60 mg, 0.803 μ mol), guanosine 5'-diphospho- β -L-fucose sodium salt GDP-Fucose 20 mM (118.5 μ L, 2.35 μ mol, 2.9 eq), α -1,3-fucosyltransferase CeFUT6 (110 μ L), MnCl₂ (20 mM) in MES 80mM buffer pH 6.5 was incubated at room temperature for 48h. The resulting mixture was heated at 95°C for 5 min to precipitate the enzyme. After centrifugation, the supernatant was purified by semi-preparative HPLC (C18 10x250 mm, 5 μ m, ammonium formate 20 mM:ACN, gradient D) and the collected fractions were evaporated and freeze-dried obtaining 1.03 mg of **18** (60%). ¹H NMR (500 MHz, MeOD) δ 7.43 – 7.13 (m, 20H, Ph), 5.10 – 5.04 (m, 2H, H-1D, H-1G'), 5.00 (t, J = 11.9 Hz, 2H, CH₂ Bn), 4.81 – 4.72 (m, 3H, H-5G', CH₂ Bn, H-1D'), 4.69 – 4.55 (m, 5H, 3x CH₂ Bn, H-1B, H-1C), 4.49 – 4.41 (m, 4H, 2x CH₂ Bn, H-1E, H-1F'), 4.39 (d, J = 8.1 Hz, 1H, H-1A), 4.28 (d, J = 8.0 Hz, 1H, H-1E'), 4.11 (d, J = 3.1 Hz, 1H, H-2C), 4.08 (dd, J = 3.5, 1.7 Hz, 1H, H-2D), 4.04 – 3.95 (m, 3H, H-4B, H-4A, H-2F'), 3.94 – 3.40 (m, 46H), 3.29 – 3.23 (m, 4H, CH₂N₃, H-5B, H-5E), 3.19 – 3.15 (m, 1H, H-5C), 3.15 – 3.10 (m, 1H, H-5E'), 2.01 (s, 3H, CH₃ Ac), 1.98 (m, 6H, CH₃ Ac), 1.83 (m, 6H, CH₃ Ac), 1.61 – 1.54 (m, 4H, 2x CH₂), 1.46 – 1.38 (m, 2H, CH₂), 1.30 (d, J = 6.6 Hz, 3H, H-6G'). ¹³C NMR from HSQC experiment (126 MHz, MeOD) δ 129.2, 128.94, 128.85, 128.4,

127.8, 102.3, 101.7, 101.4, 101.3, 101.2, 101.0, 100.8, 99.7, 98.2, 82.8, 82.2, 81.5, 78.4, 78.3, 77.69, 77.67, 76.84, 76.82, 76.5, 76.2, 76.1, 76.0, 75.07, 75.05, 75.0, 74.9, 74.2, 74.04, 74.01, 73.97, 73.5, 72.7, 71.7, 71.5, 71.4, 70.9, 70.0, 69.6, 69.4, 69.11, 69.10, 67.6, 66.94, 66.93, 63.2, 62.9, 62.6, 62.43, 62.37, 61.1, 57.0, 56.8, 56.6, 55.8, 54.1, 52.2, 49.0, 29.6, 24.2, 23.1, 22.92, 22.87, 22.86, 16.5. HRMS (MALDI): m/z : calcd $C_{97}H_{140}N_8NaO_{45}$: 2159.8805 $[M+Na]^+$, found 2159.8882.

α -1,3-fucosylation of 16: A solution (1 mL) of **16** (1.05 mg, 0.476 μ mol), guanosine 5'-diphospho- β -L-fucose sodium salt GDP-Fucose 20 mM (51.5 μ L, 1.03 μ mol, 2.2 eq), α -1,3-fucosyltransferase CeFUT6 (60 μ L), $MnCl_2$ (20 mM) in MES 80mM buffer pH 6.5 was incubated at room temperature for 48h. The resulting mixture was heated at 95°C for 5 min to precipitate the enzyme. After centrifugation, the supernatant was purified by semi-preparative HPLC (C18 10x250 mm 5 μ m, ammonium formate 20 mM:ACN, gradient E) and the collected fractions were evaporated and freeze-dried obtaining 118 μ g of **21** (0.047 μ mol, 10%), 300 μ g of **19** (0.128 μ mol, 27%) and 44 μ g of **20** (0.018 μ mol 4%) and 0.148 mg (0.674 μ mol, 14%)



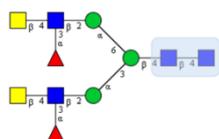
5-azidopentyl α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **19.** After α -1,3-fucosylation of **16** (1.05 mg, 0.476 μ mol) and HPLC purification of the crude using gradient E, 300 μ g (0.128 μ mol, 27%) of **19** were obtained. 1H NMR (500 MHz, MeOD) δ 7.42 – 7.11 (m, 20H, Ph), 5.06 (d, J = 3.4 Hz, 2H, H-1G, H-1D), 4.98 (d, J = 12.2 Hz, 2H, 2x CH_2 Bn), 4.80 – 4.72 (m, 3H, CH_2 Bn, H-5G, H-1D'), 4.70 – 4.54 (m, 5H, 3x CH_2 Bn, H-1B, H-1C), 4.49 – 4.41 (m, 5H, 2x CH_2 Bn, H-1E, H-1F, H-1F'), 4.39 (d, J = 8.1 Hz, 1H, H-1A), 4.23 (d, J = 8.2 Hz, 1H, H-1E'), 4.11 (d, J = 3.0 Hz, 1H, H-2C), 4.07 – 3.39 (m, 57H), 3.25 (t, J = 6.9 Hz, 3H, CH_2N_3 , H-5B), 3.17 (td, J = 4.0, 2.4 Hz, 1H, H-5C), 3.07 (d, J = 8.9 Hz, 1H, H-5E'), 2.04 (m, 3H, CH_3 Ac), 2.00 – 1.95 (m, 9H, CH_3 Ac), 1.83 (m, 6H, CH_3 Ac), 1.62 – 1.53 (m, 4H, 2x CH_2), 1.47 – 1.37 (m, 2H, CH_2), 1.27 (d, J = 6.5 Hz, 3H, H-6G). ^{13}C NMR from HSQC experiment (126 MHz, MeOD) δ 129.1, 129.0, 128.9, 128.7, 128.5, 103.3, 102.2, 102.0, 101.4, 101.3, 100.9, 100.3, 99.6, 98.1, 82.6, 82.1, 81.5, 80.8, 78.4, 78.1, 77.7, 77.3, 76.9, 76.8, 76.6, 76.5, 76.1, 75.8, 75.8, 75.0, 74.93, 74.90, 74.87, 74.4, 74.3, 74.0, 73.3, 72.7, 71.4, 70.8, 69.6, 69.4, 69.30, 69.25, 68.8, 67.6, 66.8, 66.7, 63.3, 62.8, 62.3, 61.3, 61.2, 61.0, 56.8, 56.6, 55.9, 55.8, 54.1, 53.9, 52.1, 29.6, 24.1, 22.98, 22.97, 22.9, 16.3. HRMS (MALDI): m/z : calcd $C_{105}H_{153}N_9NaO_{50}$: 2362.9598 $[M+Na]^+$, found 2362.9724.



5-azidopentyl 2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)- { α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **20.** After α -1,3-fucosylation of **16** (1.05 mg, 0.476 μ mol) and HPLC purification of the crude using gradient E, 44 μ g (4%) of

20 were obtained. ^1H NMR (500 MHz, MeOD) δ 7.44 – 7.13 (m, 20H, Ph), 5.09 – 5.04 (m, 2H, H-1D, H-1G'), 4.99 (dd, $J = 12.2, 10.2$ Hz, 2H, 2x CH_2 Bn), 4.81 – 4.71 (m, 3H, H-5G', H-1D, CH_2 Bn), 4.69 – 4.54 (m, 5H, H-1B, H-1C, 3x CH_2 Bn), 4.49 – 4.37 (m, 6H, H-1F, H-1F', H-1E, H-1A, 2x CH_2 Bn), 4.28 (d, $J = 8.0$ Hz, 1H, H-1E'), 4.10 (d, $J = 3.1$ Hz, 1H, H-2C), 4.07 (dd, $J = 3.5, 1.7$ Hz, 1H, H-2D), 4.03 – 3.94 (m, 4H, H-4A, H-4B, H-2F, H-2F'), 3.94 – 3.40 (m, 52H), 3.28 – 3.23 (m, 3H, CH_2N_3 , H-5B), 3.18 – 3.15 (m, 1H, H-5C), 3.15 – 3.10 (m, 1H, H-5E'), 2.02 – 1.99 (m, 6H, 2x CH_3 Ac), 1.99 – 1.96 (m, 6H, 2x CH_3 Ac), 1.86 – 1.79 (m, 6H, 2x CH_3 Ac), 1.62 – 1.53 (m, 4H, 2x CH_2), 1.47 – 1.38 (m, 2H, CH_2), 1.30 (d, $J = 6.5$ Hz, 3H, H-6G'). ^{13}C NMR from HSQC experiment (126 MHz, MeOD) δ 129.22, 129.21, 129.0, 128.8, 128.5, 128.0, 103.3, 102.3, 101.9, 101.3, 101.2, 101.0, 100.6, 99.9, 98.2, 82.7, 82.1, 81.6, 81.1, 78.5, 78.1, 77.7, 77.0, 76.93, 76.92, 76.6, 76.5, 76.3, 76.00, 75.96, 75.03, 75.02, 75.00, 74.4, 74.2, 74.02, 74.00, 73.47, 73.46, 72.9, 72.65, 72.65, 71.5, 71.4, 71.2, 70.9, 69.9, 69.8, 69.5, 69.3, 69.1, 67.8, 67.0, 66.9, 63.3, 63.0, 62.5, 62.2, 62.1, 61.5, 61.2, 57.0, 56.6, 56.0, 55.7, 54.0, 52.2, 29.6, 24.2, 23.1, 22.9, 22.9, 16.5. HRMS (MALDI): m/z : calcd $\text{C}_{105}\text{H}_{153}\text{N}_9\text{NaO}_{50}$: 2362.9598 $[\text{M}+\text{Na}]^+$, found 2362.9724.

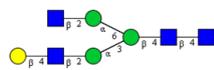
A solution (0.5 mL) of **18** (1 mg, 0.468 μmol), uridine 5'-diphospho-2-acetamido-2-deoxy- α -D-galactosamine disodium salt UDP-GalNAc 20mM (35 μL , 0.702 μmol , 1.5 eq), double mutant β -1,4-galactosyltransferase (50 μL), MnCl_2 (10 mM) in HEPES buffer (50mM, pH=7.4) was incubated at 37°C for 24h. The resulting mixture was heated at 95°C for 5 min to precipitate the enzyme. After centrifugation the supernatant was purified by semi-preparative HPLC (ACN:H₂O, C18 10x250 mm 5 μm , gradient C). The fractions were concentrated and freeze-dried, obtaining 664 μg (0.284 μmol , 60%) of compound **20**.



5-azidopentyl α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-{ α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)}- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranoside **21.** After α -1,3-fucosylation of **16** (1.05 mg, 0.476 μmol) and HPLC purification of the crude using gradient E, the collected fractions were evaporated and freeze-dried obtaining 0.118 mg, (0.047 μmol , 10%) of **21**. ^1H NMR (500 MHz, MeOD) δ 9.18 – 8.58 (m, 20H, Ph), 6.64 – 6.60 (m, 3H, H-1D, H-1G, H-1G'), 6.55 (t, $J = 11.6$ Hz, 2H, 2x CH_2), 6.37 – 6.27 (m, 5H, H-5G, H-5G', H-1D', CH_2), 6.25 – 6.10 (m, 5H, H-1B, H-1C, 3x CH_2), 6.05 – 5.96 (m, 5H, H-1F, H-1F', H-1E, 2x CH_2), 5.95 (d, $J = 8.1$ Hz, 1H, H-1A), 5.84 (d, $J = 8.1$ Hz, 1H, H-1E'), 5.67 (d, $J = 3.1$ Hz, 1H, H-2C), 5.61 (dd, $J = 3.4, 1.7$ Hz, 1H, H-2D), 5.60 – 4.95 (m, 56H), 4.81 (t, $J = 6.8$ Hz, 3H, H-5B, CH_2N_3), 4.75 – 4.72 (m, 1H, H-5C), 4.71 – 4.66 (m, 1H, H-5E'), 3.57 (s, 3H, CH_3 Ac), 3.53 (s, 3H, CH_3 Ac), 3.54 (s, 6H, 2x CH_3 Ac), 3.38 (s, 3H, CH_3 Ac), 3.39 (s, 3H, CH_3 Ac), 3.17 – 3.10 (m, 4H, 2x CH_2), 3.03 – 2.94 (m, 2H, CH_2), 2.86 (d, $J = 6.6$ Hz, 3H, H-6G'), 2.83 (d, $J = 6.5$ Hz, 3H, H-6G). ^{13}C NMR from HSQC experiment (126 MHz, MeOD) δ 129.2, 129.1, 129.0, 128.9, 128.4, 127.9, 102.3, 102.0, 101.4, 101.1, 100.94, 100.93, 100.5, 99.7, 98.2, 82.7, 82.2, 81.5, 78.5, 78.0, 77.7, 77.4, 77.03, 77.00, 76.6, 76.5, 76.02, 75.99, 75.2, 75.0, 74.9, 74.33, 74.27, 74.0, 73.6, 72.7, 71.5, 71.0, 69.81, 69.75, 69.7, 69.4, 69.1, 67.7, 67.0, 66.8, 63.2, 63.0, 62.6, 62.5, 61.3, 61.2, 57.0, 56.7, 56.5, 55.8, 54.1, 52.2, 29.6, 24.2, 23.1, 22.9, 22.8, 16.5. HRMS (MALDI): m/z calcd $\text{C}_{111}\text{H}_{163}\text{N}_9\text{NaO}_{54}\text{Na}$: 2509.0178 $[\text{M}+\text{Na}]^+$, found 2509.0291.

General hydrogenation procedure: Glycan (0.5-1 mg) is dissolved in 0.5 mL of MeOH containing trifluoroacetic acid (0.1%). The mixture was hydrogenated on a H-Cube hydrogenation apparatus, using

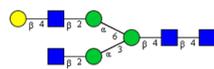
10%Pd/C as catalyst, MeOH containing 0.1%TFA as mobile phase at 2 mL/min, at 30°C and full H₂. The collected fraction was concentrated to dryness and purify with SampliQ high performance graphitized carbon cartridges eluting the compound in a mixture of H₂O:ACN 8:2.



5-aminopentyl

β -D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-

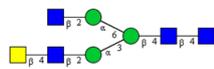
glucopyranosyl-(1→2)- α -D-mannopyranosyl-(1→3)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1→2)- α -D-mannopyranosyl-(1→6)]- β -D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-glucopyranoside 22. The general hydrogenation procedure was applied to 1.55 mg (0.793 μ mol) of **8** and 756 μ g of **22** (61% yield) were obtained. ¹H NMR (500 MHz, D₂O) δ 5.12 (s, 1H), 4.92 (s, 1H), 4.62 – 4.54 (m, 3H), 4.52 – 4.45 (m, 2H), 4.25 (d, *J* = 1.6 Hz, 1H), 4.19 (dd, *J* = 3.1, 1.5 Hz, 1H), 4.11 (dd, *J* = 3.1, 1.3 Hz, 1H), 4.02 – 3.40 (m, 47H), 2.99 (t, *J* = 7.7 Hz, 2H), 2.08 (s, 3H), 2.05 (s, 6H), 2.03 (s, 3H), 1.72 – 1.63 (m, 2H), 1.63 – 1.55 (m, 2H), 1.44 – 1.36 (m, 2H). ¹³C NMR from HSQC experiment (126 MHz, D₂O) δ 102.8, 101.4, 101.1, 100.6, 99.62, 99.58, 99.5, 97.0, 80.4, 79.6, 79.3, 78.5, 76.4, 76.3, 75.8, 75.3, 74.7, 74.5, 74.2, 73.5, 73.2, 72.8, 72.4, 72.0, 71.0, 70.2, 70.1, 69.8, 69.4, 68.5, 67.4, 66.0, 61.63, 61.58, 60.8, 60.1, 60.0, 55.0, 54.8, 39.3, 28.1, 26.5, 22.3, 22.0. HRMS (MALDI): *m/z* calcd for C₆₁H₁₀₅N₅O₄₁Na: 1586.6177, found 1586.6118.



5-aminopentyl

β -D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-

glucopyranosyl-(1→2)- α -D-mannopyranosyl-(1→6)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1→2)- α -D-mannopyranosyl-(1→3)]- β -D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-glucopyranoside 23. The general hydrogenation procedure was applied to 1.20 mg (0.793 μ mol) of **9** and 782 μ g of **24** (81% yield) were obtained. ¹H NMR (500 MHz, D₂O) δ 5.12 (s, 1H), 4.93 (s, 1H), 4.64 – 4.53 (m, 3H), 4.49 (t, *J* = 8.0 Hz, 2H), 4.25 (d, *J* = 2.6 Hz, 1H), 4.21 – 4.17 (m, 1H), 4.11 (dd, *J* = 3.1, 1.3 Hz, 1H), 4.03 – 3.40 (m, 47H), 2.99 (t, *J* = 7.7 Hz, 2H), 2.09 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H), 2.03 (s, 3H), 1.72 – 1.63 (m, 2H), 1.63 – 1.55 (m, 2H), 1.45 – 1.34 (m, 2H). ¹³C NMR from HSQC experiment (126 MHz, D₂O) δ 102.9, 101.3, 101.1, 100.4, 99.6, 99.5, 97.0, 80.5, 79.3, 78.8, 76.4, 76.2, 75.8, 75.2, 74.38, 74.37, 74.35, 73.5, 73.2, 72.8, 72.3, 72.0, 71.0, 70.2, 70.0, 69.8, 69.6, 68.5, 67.3, 65.9, 65.6, 61.63, 60.8, 60.0, 59.9, 55.1, 54.8, 39.3, 28.0, 26.3, 22.2, 22.1. HRMS (MALDI): *m/z* calcd for C₆₁H₁₀₅N₅O₄₁Na: 1586.6177, found 1586.6042.



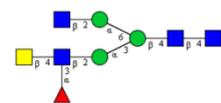
5-aminopentyl 2-acetamido-2-deoxy- β -D-galactopyranosyl-(1→4)-2-acetamido-2-

deoxy- β -D-glucopyranosyl-(1→2)- α -D-mannopyranosyl-(1→3)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1→2)- α -D-mannopyranosyl-(1→6)]- β -D-mannopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy- β -D-glucopyranoside 24. The general hydrogenation procedure was applied to 786 μ g (0.39 μ mol) of compound **14** and 506 μ g of **24** (80% of yield) were obtained. ¹H NMR (500 MHz, D₂O) δ 5.12 (s, 1H), 4.93 (s, 1H), 4.61 (d, *J* = 7.5 Hz, 1H), 4.57 (d, *J* = 7.7 Hz, 2H), 4.51 (dd, *J* = 10.7, 8.1 Hz, 2H), 4.26 (s, 1H), 4.19 (d, *J* = 3.5 Hz, 1H), 4.12 (d, *J* = 3.0 Hz, 1H), 4.00 – 3.41 (m, 47H), 2.98 (t, *J* = 7.7 Hz, 2H), 2.09 (s, 3H), 2.08 (s, 3H), 2.06 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H), 1.72 – 1.63 (m, 2H), 1.63 – 1.56 (m, 2H), 1.45 – 1.36 (m, 2H). ¹³C NMR from HSQC experiment (126 MHz, D₂O) δ 101.8, 101.4, 101.1, 100.4, 99.54, 99.48, 97.0, 80.4, 79.5, 79.3, 76.4, 76.3, 75.8, 75.3, 74.5, 74.3, 73.5, 73.4, 72.8, 72.4, 72.0, 70.6, 70.20, 70.17, 69.8, 69.5, 67.5, 67.3, 66.0, 65.8, 61.6, 60.9, 60.12, 60.0, 55.1, 54.9, 54.7,

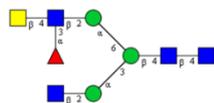
52.4, 39.3, 28.1, 26.5, 22.2, 22.1. HRMS (MALDI): m/z : calcd $C_{63}H_{108}N_6O_{41}Na$: 1627.6443 $[M+Na]^+$, found 1627.6384.



5-aminopentyl 2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)- 2-acetamido-2-deoxy- β -D-glucopyranoside 25. The general hydrogenation procedure was applied to 806 μ g (0.405 μ mol) of compound **15** and 540 μ g (0.336 μ mol) of **25** (83% of yield) were obtained. 1H NMR (500 MHz, D_2O) δ 5.13 (s, 1H), 4.93 (s, 1H), 4.61 (d, J = 8.1 Hz, 1H), 4.57 (dd, J = 8.2, 3.5 Hz, 2H), 4.53 (d, J = 8.5 Hz, 1H), 4.50 (d, J = 7.6 Hz, 1H), 4.26 (s, 1H), 4.20 (dd, J = 3.4, 1.6 Hz, 1H), 4.14 – 4.09 (m, 1H), 4.00 – 3.42 (m, 47H), 2.98 (t, J = 7.7 Hz, 1H), 2.10 (s, 3H), 2.09 (s, 2H), 2.07 (s, 3H), 2.05 (s, 2H), 2.04 (s, 3H), 1.72 – 1.63 (m, 2H), 1.63 – 1.56 (m, 2H), 1.46 – 1.37 (m, 2H). ^{13}C NMR from HSQC experiment (126 MHz, D_2O) δ 101.8, 101.4, 101.2, 100.4, 99.59, 99.55, 98.7, 97.0, 80.5, 79.6, 79.4, 76.5, 76.3, 75.8, 75.3, 74.5, 74.3, 73.6, 72.8, 72.4, 70.7, 70.1, 70.1, 70.0, 69.54, 69.52, 67.4, 67.3, 65.9, 65.7, 61.5, 61.0, 60.4, 60.0, 55.1, 55.0, 54.7, 52.5, 39.4, 28.1, 26.5, 22.3, 22.1. HRMS (MALDI): m/z : calcd $C_{63}H_{108}N_6O_{41}Na$: 1627.6443 $[M+Na]^+$, found 1627.6412.

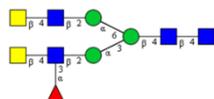


5-aminopentyl α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)- 2-acetamido-2-deoxy- β -D-glucopyranoside 26. The general hydrogenation procedure was applied to 988 μ g (0.462 μ mol) of **17** and 762 μ g (94% of yield) of **26** were obtained. 1H NMR (500 MHz, D_2O) δ 5.13 (d, J = 4.0 Hz, 1H), 5.10 (s, 1H), 4.92 (s, 1H), 4.91 – 4.83 (m, 1H), 4.60 (d, J = 7.5 Hz, 1H), 4.58 – 4.52 (m, 2H), 4.49 (d, J = 7.4 Hz, 1H), 4.45 (d, J = 8.4 Hz, 1H), 4.25 (s, 1H), 4.17 (s, 1H), 4.11 (s, 1H), 4.01 – 3.40 (m, 50H), 2.97 (t, J = 7.7 Hz, 3H), 2.08 (s, 3H), 2.05 (s, 3H), 2.04 (s, 6H), 2.03 (s, 3H), 1.71 – 1.62 (m, 2H), 1.62 – 1.55 (m, 2H), 1.40 (d, J = 7.8 Hz, 3H), 1.26 (d, J = 6.5 Hz, 3H). ^{13}C NMR from HSQC experiment (126 MHz, D_2O) δ 101.4, 101.0, 100.7, 100.4, 99.6, 99.4, 98.4, 97.0, 80.3, 79.5, 79.4, 76.3, 76.2, 75.8, 74.93, 74.87, 74.5, 74.2, 73.6, 73.3, 72.8, 72.4, 71.9, 70.8, 70.2, 70.2, 69.9, 69.6, 69.3, 67.7, 67.4, 67.3, 66.9, 65.8, 65.7, 61.6, 61.6, 60.4, 60.3, 60.0, 55.5, 55.1, 55.0, 52.4, 39.3, 28.0, 26.5, 22.3, 22.1, 15.3. HRMS (MALDI): m/z : calcd $C_{69}H_{118}N_6O_{45}Na$: 1773.7022 $[M+Na]^+$, found 1773.6988.



5-aminopentyl α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)-[2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)- 2-acetamido-2-deoxy- β -D-glucopyranoside 27. The general hydrogenation procedure was applied to 1.15 mg (0.536 μ mol) of **18** and 862 μ g (91% of yield) of **27** were obtained. 1H NMR (500 MHz, D_2O) δ 5.15 (d, J = 4.0 Hz, 1H), 5.13 (s, 1H), 4.94 – 4.86 (m, 2H), 4.61 (d, J = 7.8 Hz, 1H), 4.57 (d, J = 8.4 Hz, 2H), 4.51 (d, J = 7.5 Hz, 1H), 4.47 (d, J = 8.4 Hz, 1H), 4.26 (s, 1H), 4.20 (d, J = 3.4 Hz, 1H), 4.09 (d, J = 3.5 Hz, 1H), 4.03 – 3.44 (m, 5H), 2.98 (t, J = 7.7 Hz, 2H), 2.08 (s, 3H), 2.07

(s, 3H), 2.06 (s, 3H), 2.05 (s, 3H), 2.04 (s, 3H), 1.72 – 1.64 (m, 2H), 1.64 – 1.57 (m, 2H), 1.45 – 1.36 (m, 2H), 1.28 (d, $J = 6.5$ Hz, 3H). ^{13}C NMR from HSQC experiment (126 MHz, D_2O) δ 101.4, 101.0, 100.7, 100.5, 99.6, 99.4, 98.4, 97.0, 80.4, 79.6, 79.3, 76.4, 76.2, 75.8, 74.8, 74.3, 73.6, 73.31, 73.27, 72.72, 72.69, 72.5, 72.0, 72.0, 71.8, 70.6, 70.24, 70.17, 69.7, 69.5, 69.1, 67.4, 67.3, 67.2, 66.9, 65.9, 65.8, 61.7, 61.4, 60.2, 60.1, 55.6, 55.1, 55.0, 52.4, 39.4, 28.1, 26.5, 22.3, 22.1. HRMS (MALDI): m/z : calcd $\text{C}_{69}\text{H}_{118}\text{N}_6\text{O}_{45}\text{Na}$: 1773.7022 $[\text{M}+\text{Na}]^+$, found 1773.6975.

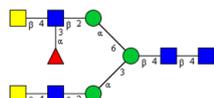


5-aminopentyl

α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-

galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-

2-acetamido-2-deoxy- β -D-glucopyranoside **28**. The general hydrogenation procedure was applied to 1.96 mg (0.837 μmol) of **19** and 1.5 mg (90%) of **28** were obtained. ^1H NMR (500 MHz, D_2O) δ 5.14 (d, $J = 4.0$ Hz, 1H), 5.11 (s, 1H), 4.93 (s, 1H), 4.91 – 4.84 (m, 1H), 4.61 (d, $J = 8.0$ Hz, 1H), 4.59 – 4.49 (m, 4H), 4.46 (d, $J = 8.4$ Hz, 1H), 4.26 (s, 1H), 4.18 (dd, $J = 3.5, 1.5$ Hz, 1H), 4.11 (s, 1H), 4.02 – 3.56 (m, 51H), 3.54 – 3.46 (m, 5H), 3.00 (t, $J = 7.7$ Hz, 2H), 2.10 (s, 3H), 2.09 (s, 3H), 2.05 (d, $J = 1.6$ Hz, 9H), 2.04 (s, 3H), 1.73 – 1.64 (m, 2H), 1.64 – 1.55 (m, 2H), 1.47 – 1.37 (m, 1H), 1.28 (d, $J = 6.5$ Hz, 3H). ^{13}C NMR from HSQC experiment (126 MHz, D_2O) δ 101.8, 101.3, 101.0, 100.7, 100.5, 99.5, 99.4, 98.4, 97.1, 80.4, 79.4, 79.3, 76.32, 76.31, 75.4, 74.9, 74.8, 74.38, 74.37, 74.2, 73.5, 73.2, 72.8, 72.4, 72.1, 70.6, 70.18, 70.16, 69.5, 69.0, 67.7, 67.6, 67.4, 67.3, 66.9, 65.7, 61.7, 61.7, 61.0, 60.0, 59.9, 55.7, 55.0, 54.7, 52.4, 39.3, 28.1, 26.3, 22.2, 22.1, 15.3. HRMS (MALDI): m/z : calcd $\text{C}_{69}\text{H}_{118}\text{N}_6\text{O}_{45}\text{Na}$: 1976.7815 $[\text{M}+\text{Na}]^+$, found 1976.7864.



5-aminopentyl

α -L-fucopyranosyl-(1 \rightarrow 3)-[2-acetamido-2-deoxy- β -D-

galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 6)-[2-acetamido-2-deoxy- β -D-galactopyranosyl-(1 \rightarrow 4)]-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 2)- α -D-mannopyranosyl-(1 \rightarrow 3)]- β -D-mannopyranosyl-(1 \rightarrow 4)-2-acetamido-2-deoxy- β -D-glucopyranosyl-(1 \rightarrow 4)-

2-acetamido-2-deoxy- β -D-glucopyranoside **29**. The general hydrogenation procedure was applied to 810 μg (0.346 μmol) of **20**, and 534 μg (79%) of **29** were obtained. ^1H NMR (500 MHz, D_2O) δ 5.15 (d, $J = 4.0$ Hz, 1H), 5.12 (s, 1H), 4.93 – 4.85 (m, 2H), 4.83 (s, 1H), 4.61 (d, $J = 7.7$ Hz, 1H), 4.59 – 4.48 (m, 4H), 4.47 (d, $J = 8.3$ Hz, 1H), 4.26 (s, 1H), 4.20 – 4.18 (m, 1H), 4.09 (d, $J = 2.7$ Hz, 1H), 4.02 – 3.57 (m, 51H), 3.54 – 3.46 (m, 5H), 3.00 (t, $J = 7.6$ Hz, 2H), 2.10 (s, 3H), 2.08 (s, 3H), 2.06 (s, 6H), 2.05 (s, 3H), 2.04 (s, 3H), 1.72 – 1.64 (m, 2H), 1.64 – 1.56 (m, 2H), 1.45 – 1.37 (m, 2H), 1.28 (d, $J = 6.4$ Hz, 3H). ^{13}C NMR from HSQC experimente δ 101.6, 101.4, 100.8, 100.7, 100.5, 99.5, 99.4, 98.4, 97.0, 80.3, 79.32, 79.29, 76.4, 76.2, 75.4, 74.7, 74.6, 74.6, 74.2, 73.6, 73.3, 72.6, 72.5, 72.11, 72.05, 70.8, 70.3, 70.2, 69.6, 69.1, 67.7, 67.6, 67.3, 67.2, 67.0, 65.7, 62.8, 61.7, 61.5, 61.0, 60.0, 59.8, 55.7, 55.3, 55.1, 54.6, 52.6, 52.4, 39.3, 28.0, 26.4, 22.2, 22.0, 15.3. HRMS (MALDI): m/z : calcd $\text{C}_{69}\text{H}_{118}\text{N}_6\text{O}_{45}\text{Na}$: 1976.7815 $[\text{M}+\text{Na}]^+$, found 1976.7827.

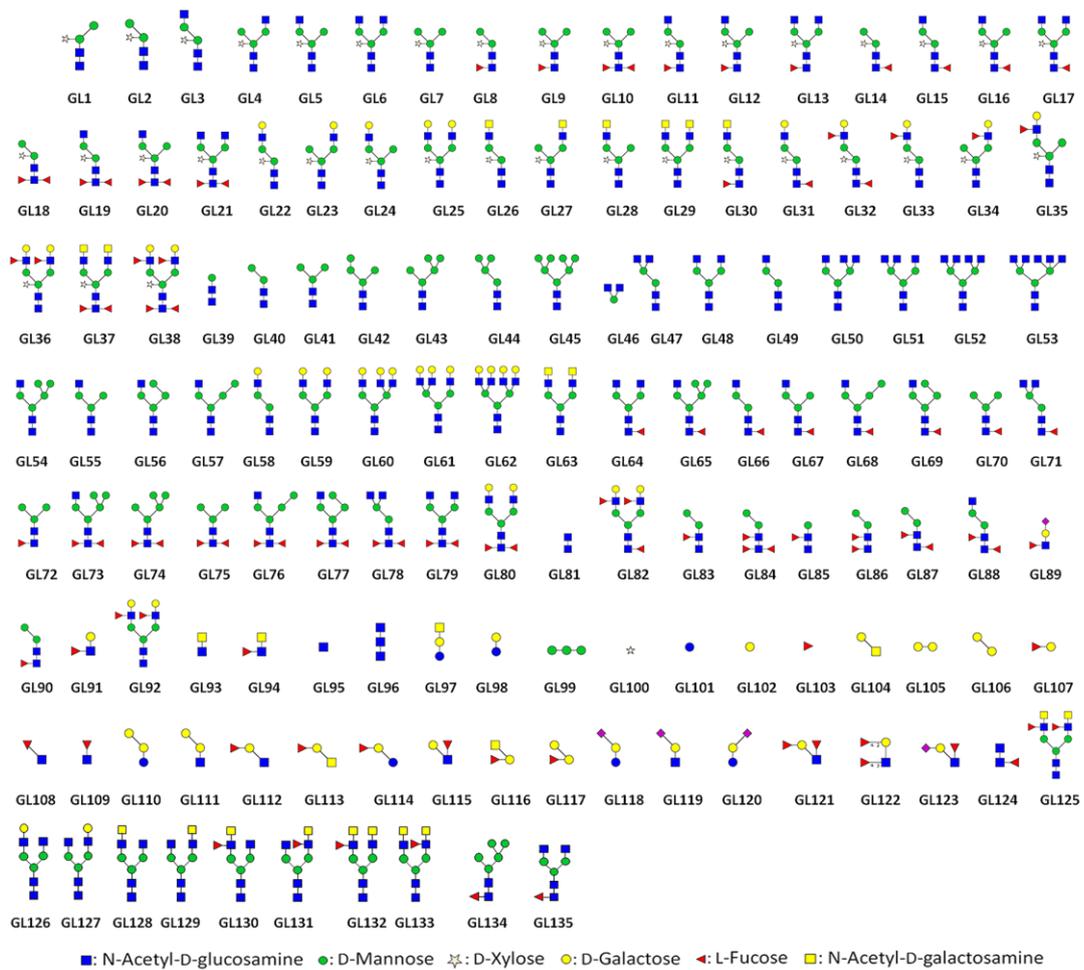


Figure S1. Glycan structures included on microarrays

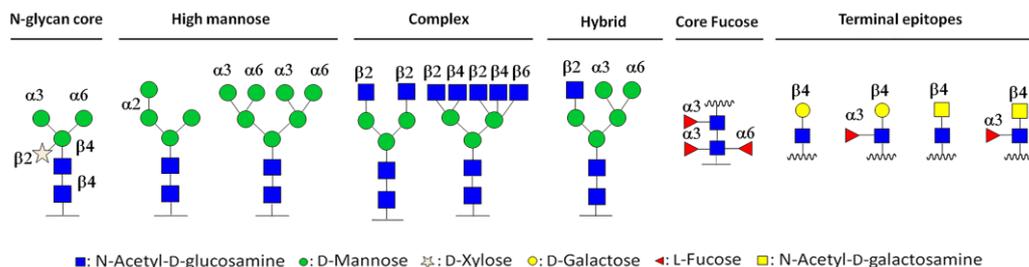
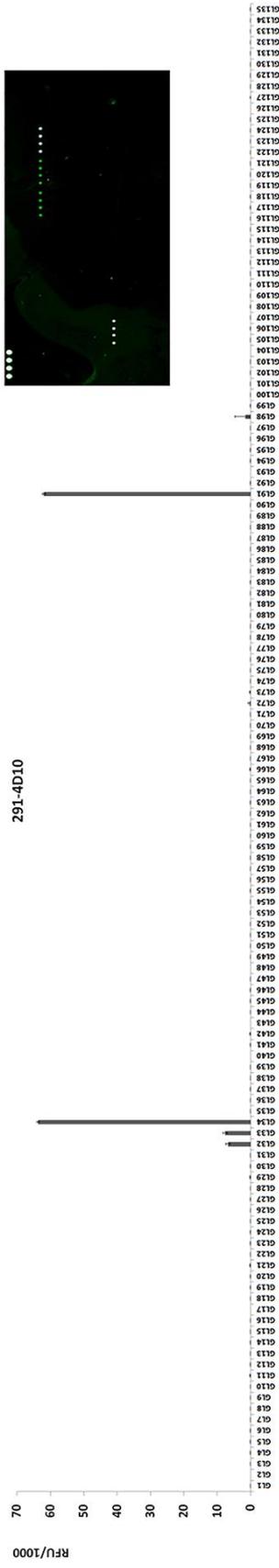
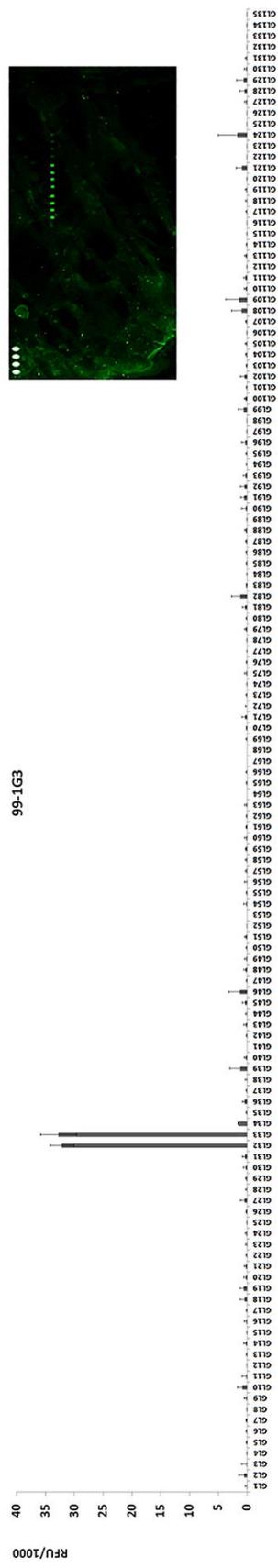
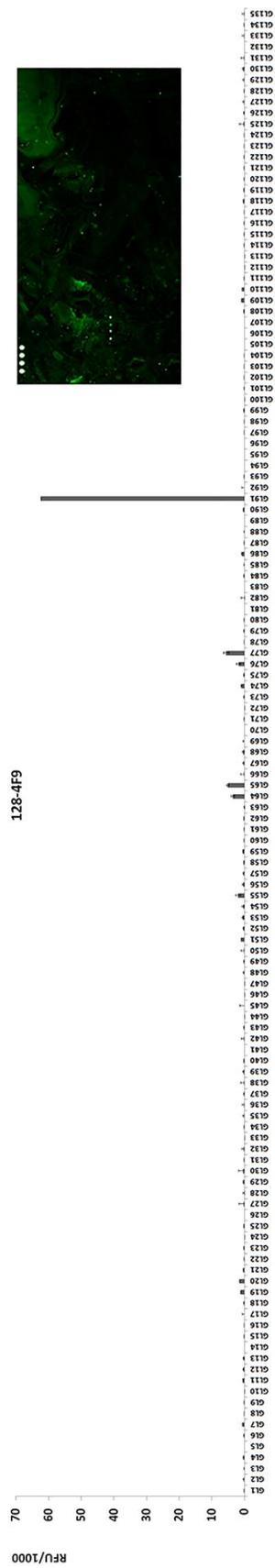


Figure S2. Glycosidic bond nature for N-glycan structures on microarrays.

Table S1. Additional glycan structures included on microarrays.

GL46	GlcNAc β 1-2(GlcNAc β 1-4)Man α -sp	
GL81	GlcNAc β 1-4GlcNAc β -sp	
GL89	Neu5Ac α 2-3Gal β 1-4(Fuc α 1-3)GlcNAc β -sp	SialylLeX
GL91	Gal β 1-4(Fuc α 1-3)GlcNAc β -sp	LeX
GL93	GalNAc β 1-4GlcNAc β -sp	LDN
GL94	GalNAc β 1-4(Fuc α 1-3)GlcNAc β -sp	
GL95	GlcNAc β -sp	
GL96	GlcNAc β 1-4GlcNAc β 1-4GlcNAc β -sp	
GL97	GalNAc β 1-4Gal β 1-4GlcNAc β -sp	GG3
GL98	Gal β 1-4Glc β -sp	lactose
GL99	Man α 1-2Man α 1-2Man α -sp	
GL100	Xyl β -sp	
GL101	Glc β -sp	
GL102	Gal β -sp	
GL103	Fuc α -sp	
GL104	Gala1-3GalNAca-sp	
GL105	Gal α 1-2Gal β -sp	
GL106	Gal α 1-3Gal β -sp	
GL107	Fuc α 1-2Gal β -sp	
GL108	Fuc α 1-3GlcNAc β -sp	
GL109	Fuc α 1-4GlcNAc β -sp	
GL110	Gal α 1-3Gal β 1-4Glc β -sp ⁴	
GL111	Gal α 1-3Gal β 1-4GlcNAc β -sp	
GL112	Fuc α 1-2Gal β 1-3GlcNAc β -sp	Le ^d (H type 1)
GL113	Fuc α 1-2Gal β 1-3GalNAc α -sp	H (type 3)
GL114	Fuc α 1-2Gal β 1-4Glc β -sp	H (type 6)
GL115	Gal β 1-3(Fuc α 1-4)GlcNAc β -sp	Le ^a
GL116	GalNAc α 1-3(Fuc α 1-2)Gal β -sp	A _{tri}
GL117	Gal α 1-3(Fuc α 1-2)Gal β -sp	B _{tri}
GL118	Neu5Ac α 2-3Gal β 1-4Glc β -sp	3'SL
GL119	Neu5Ac α 2-3Gal β 1-4GlcNAc β -sp	3'SLN
GL120	Neu5Ac α 2-6Gal β 1-4GlcNAc β -sp	6'SL
GL121	Fuc α 1-2Gal β 1-3(Fuc α 1-4)GlcNAc β -sp	Le ^b
GL122	Fuc α 1-2Gal β 1-3(Fuc α 1-3)GlcNAc β -sp	Le ^y
GL123	Neu5Ac α 2-3Gal β 1-3(Fuc α 1-4)GlcNAc β -sp	SialLe ^a
GL124	GlcNAc β 1-4(Fuc α 1-6)GlcNAc β -sp	FucGlcNAc ₂



Fluorescence images and bar charts showing the mean RFU values for four spots with the SD of the mean. after incubation with Alexa Fluor[®]-555 anti-mouse IgM. Each histogram represents the mean RFU values for four spots with the SD of the mean.

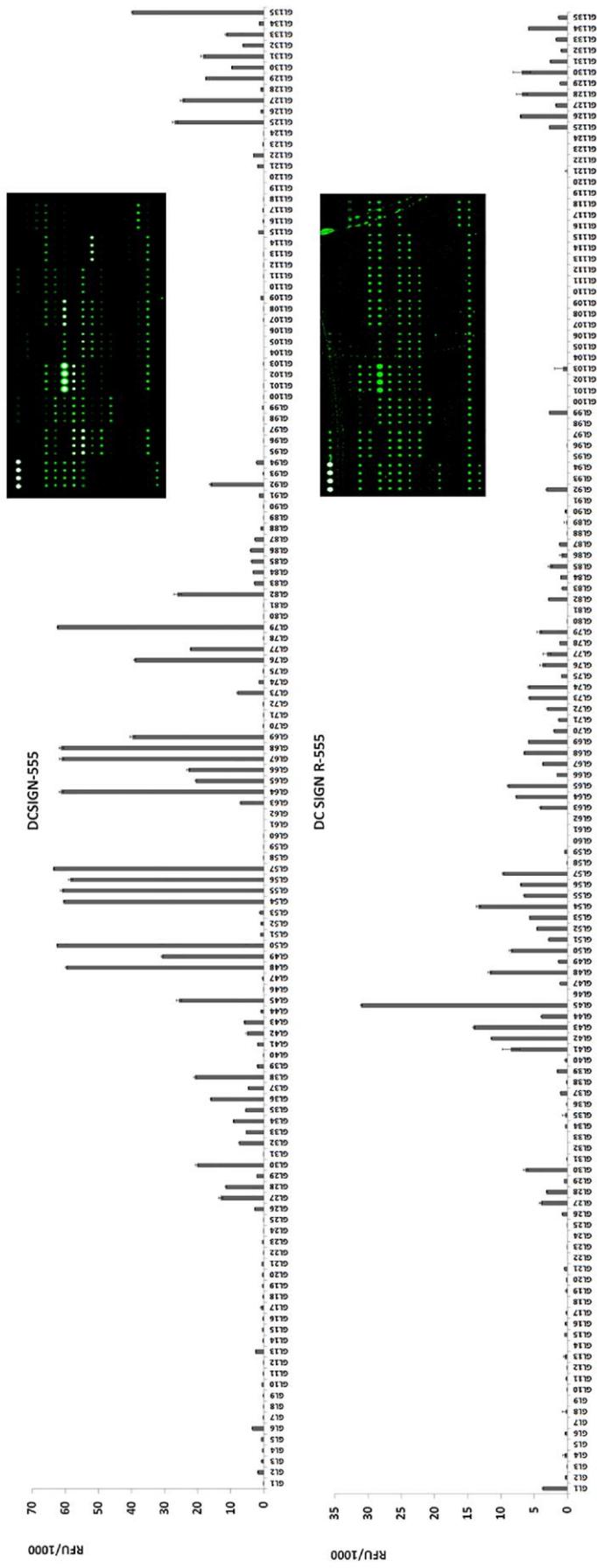


Figure S3. Glycan microarray binding profile of DC SIGN and DC SIGN-R. Each histogram represents the mean RFU values for four spots with the SD of the mean.

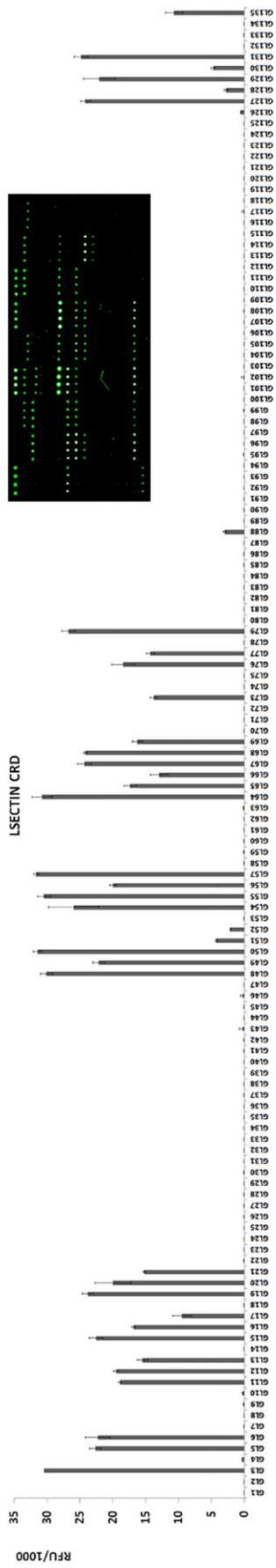


Figure S3. Glycan microarray binding profile of LSEctin CRD. Each histogram represents the mean RFU values for four spots with the SD of the mean.

Compound names used in thesis and the corresponding initial names of the same compounds Paper 2

Compound name in thesis	Corresponding initial name	Compound name in thesis	Corresponding initial name
GL1	KB03_21	GL41	SS5_67
GL2	KB01_96	GL42	BE02_93
GL3	KB03_19	GL43	SS5_62
GL4	KB02_87	GL44	SS5_68
GL5	KB02_110	GL45	SS5_72
GL6	KB02_90	GL46	SS5_69
GL7	KB03_83	GL47	SS5_63
GL8	KB03_42	GL48	NG4_010
GL9	KB03_89	GL49	NG2_091
GL10	KB03_99	GL50	NG2_092
GL11	KB03_41	GL51	SS5_102
GL12	KB03_64	GL52	SS6_88
GL13	KB03_43	GL53	BE3_57_2
GL14	KB03_69	GL54	SS5_99
GL15	KB03_60	GL55	NG3_010
GL16	KB03_61	GL56	NG3_110
GL17	KB03_62	GL57	NG4_061
GL18	KB03_79	GL58	GNG02_091
GL19	KB03_66	GL59	SS6_94
GL20	KB03_67	GL60	SS8_50
GL21	KB03_78_95	GL61	SS8_51
GL22	KB03_28	GL62	SS8_52
GL23	KB03_107	GL63	KB03_104
GL24	KB03_108	GL64	SS6_105
GL25	KB03_87	GL65	SS7_6
GL26	KB03_71	GL66	SS8_10
GL27	KB03_81	GL67	SS8_33
GL28	KB03_84	GL68	SS8_34
GL29	KB03_85	GL69	SS8_35
GL30	KB03_86	GL70	SS8_39
GL31	KB03_93	GL71	SS8_36
GL32	KB03_103	GL72	SS8_29
GL33	KB03_77	GL73	SS8_32
GL34	KB03_80	GL74	SS8_40
GL35	KB03_76	GL75	SS8_41
GL36	KB03_91	GL76	SS8_42
GL37	KB03_102	GL77	SS8_43
GL38	KB03_110	GL78	SS8_44
GL39	SS5_73	GL79	SS8_45
GL40	SS5_64	GL80	KB03_111

Compound name in thesis	Corresponding initial name	Compound name in thesis	Corresponding initial name
GL81	BE2_19	GL124	O-725
GL82	KB03-113	GL125	BE6_60
GL83	SS8_53	GL126	BE6_70
GL84	SS8_17	GL127	BE6_72
GL85	SS7_114	GL128	BE6_52
GL86	SS8_47	GL129	BE6_49
GL87	SS8_14	GL130	BE6_80
GL88	SS8_13	GL131	BE6_82
GL89	SS7_72	GL132	BE6_58
GL90	SS8_28	GL133	BE6_56
GL91	SS7_12	GL134	SS8_30
GL92	SS8_48	GL135	SS6_101
GL93	SS7_46		
GL94	SS7_49		
GL95	SS7_64		
GL96	BE2_13		
GL97	GG3		
GL98	lactose		
GL99	man3		
GL100	xylose		
GL101	glucose		
GL102	O-23		
GL103	O-27		
GL104	O-53		
GL105	O-55		
GL106	O-88		
GL107	O-91		
GL108	O-49		
GL109	O-50		
GL110	O-10a		
GL111	O-70		
GL112	O-42H1		
GL113	O-59H3		
GL114	O-940H6		
GL115	O-40Lea		
GL116	O-85Atri		
GL117	O-86Btri		
GL118	O-60		
GL119	O-36		
GL120	O-63a		
GL121	O-41Leb		
GL122	O-45Ley		
GL123	O-SiaLea		

Supplementary Information Paper.3

On-chip screening of a glycomimetic library with C-type lectins reveals structural features responsible for preferential binding of dectin-2 over DC-SIGN/R and langerin

Laura Medve,¹ Silvia Achilli,² Sonia Serna,³ Fabio Zuccotto,⁴ Norbert Varga,¹ Michel Thépaut,² Monica Civera,¹ Corinne Vivès,² Franck Fieschi,² Niels Reichardt,^{3,5} Anna Bernardi¹

¹ Dipartimento di Chimica, Università degli Studi di Milano, Via Golgi 19, 20133 Milano, Italy

² Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

³ Glycotechnology laboratory, CIC biomaGUNE, Paseo Miramón 182, 20014, Donostia/San Sebastián, Spain

⁴ University of Dundee, Dundee, United Kingdom

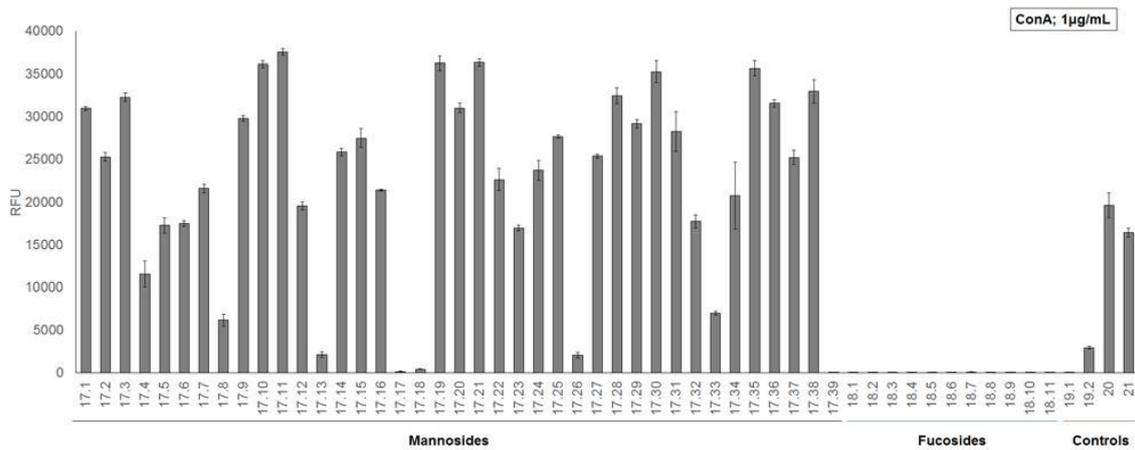
⁵ CIBER-BBN, 20014 Donostia-San Sebastián, Spain

Content Table

<i>Validation of the microarray with plant, and fungal lectins-Figures SI-1, SI-2</i>	<i>p. 2</i>
<i>Materials and Instrumentation</i>	<i>p. 3</i>
<i>SI-Table 1 - MALDI-TOF m/z values before and after conjugation with BCN-amine</i>	<i>p. 5</i>
<i>Heat-map for CLR screenings (Figure SI-3)</i>	<i>p. 7</i>
<i>Protein preparation and labeling</i>	<i>p. 8</i>
<i>Surface Plasmon Resonance (SPR) analysis</i>	<i>p. 10</i>
<i>Figure SI-4 – IC₅₀ of red bar ligands in Figure 3A (from ref 11c)</i>	
<i>Figure SI-5 – Inhibition data for 12.15 towards langerin</i>	
<i>Sensorgrams for the titration of 12.11, 12.15, 12.19 and 12.27</i>	
<i>Characterization of the ligands</i>	<i>p. 15</i>
<i>¹H- and ¹³C-NMR spectra of new compounds</i>	<i>p. 31</i>

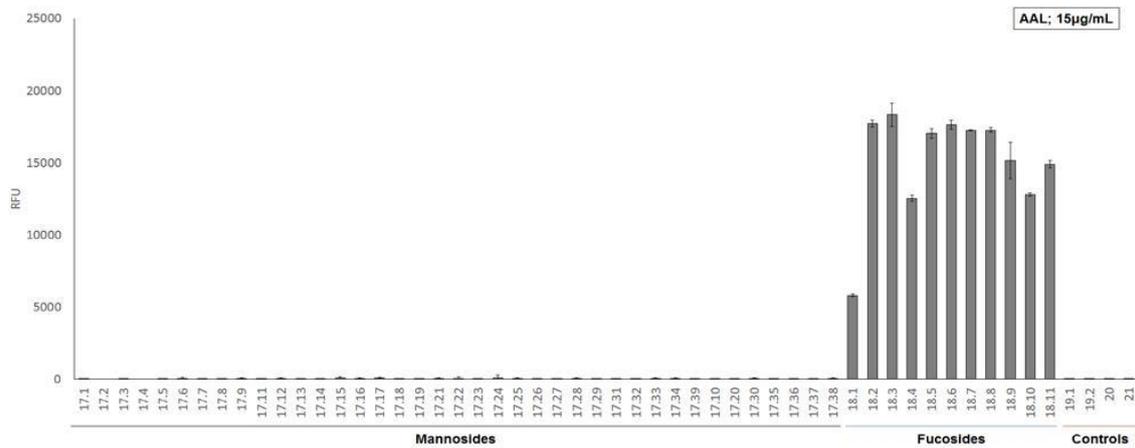
Validation of the microarray with plant and fungal lectins

SI-1 – Mannose specific plant Lectin (Con-A)



Con-A glycomimetic ligand binding profile *RFU – Relative Fluorescence Unit. Each histogram represents the average RFU values for four spots with the standard deviation of the mean.

SI-2 – Fucose specific fungal Lectin (AAL)



AAL glycomimetic ligand binding profile *RFU – Relative Fluorescence Unit. Each histogram represents the average RFU values for four spots with the standard deviation of the mean.

Materials and Instrumentation

Materials

Chemicals were purchased from Sigma-Aldrich or Acros Organics, and specific amines from Key Organics, Crea-Chim, Vitas M Labs, Life Chemicals, Alinda Chemicals, Chem Bridge or Enamine BB (source indicated at the bisamide derivatives' characterization) and were used without further purification. All reaction solvents were dried over activated 4Å or 3Å molecular sieves. Thin layer chromatography was carried out using 60 F₂₅₄ TLC plates and visualized by UV irradiation (254 nm) or by staining with cerium molybdate, potassium permanganate or ninhydrin solution. *Canavalia ensiformis* lectin (ConA) and *Aleuria aurantia* lectin (AAL) were purchased from VectorLabs. These lectins were labeled with Alexa Fluor® 555 NHS Ester (Succinimidyl Ester) (Thermo Fisher Scientific). Human CLRs were prepared and labeled as described below (*Protein preparation and labeling*)

Protein incubations were performed using 8 Well ProPlate™ Slide Modules from Electron Microscopy Sciences, Hatfield, UK.

Instrumentation

Purifications of the compounds were performed on direct phase silica gel or on an Isolera™ Prime Advanced Automated Flash Purification system (Biotage AB, Uppsala Sweden) employing prepacked Biotage® SNAP KP-C18-HS or Biotage® SNAP KP-Sil cartridges. Aqueous solutions were prepared from nanopure water produced with a Diamond UV water purification system (Branstead International, IA, Madrid, Spain).

The final ligands were lyophilized on an ALPHA-2-4 LSC freeze-dryer from Christ, Osterode, Germany. All organic solvents were concentrated using rotary evaporation.

¹H and ¹³C spectra were acquired on Bruker 400 MHz spectrometer. Chemical shifts were reported in ppm (δ) and referenced to the residual signal of the solvent used (CD₃OD 4.87 ppm; CDCl₃ 7.26 ppm; D₂O 4.79 ppm). Splitting patterns are designated as s, singlet; d, doublet; t, triplet; m, multiplet. Coupling constants (*J*) are reported in Hz.

High resolution mass spectra were acquired on a Waters SYNAPT G2 Si ESI QToF instrument, (Waters, Manchester, UK) and processed with Waters MassLynx software. The instrument was operated with a capillary voltage of 1.0 – 3.0 kV; sample infusion flow rate: 5 μ l/min; source temperature: 100°C; sampling cone: 24.0000V; desolvation temperature: 150°C. Desolvation gas flow was set to 500.0 L/hr; and the nebulizer gas flow to 7.0 Bar. MALDI-ToF mass measurements were performed on an Ultraflexextreme III time-of-flight mass spectrometer equipped with a pulsed Nd:YAG laser (λ 355 nm) and controlled by FlexControl 3.3 software (Bruker Daltonics, Bremen, Germany). The acquisitions (total of 2000-3000) were carried out in positive or negative reflector ion extraction of 450 ns and laser frequency of 1000 Hz. Laser fluence was set up to 60-80 % and the *m/z* range was chosen according to the mass of the sample. The accumulated spectra were processed using the Bruker software FlexAnalysis 3.3.

Microarrays were printed employing a robotic non-contact piezoelectric SciFLEXARRAYER spotter S11 (Scienion, Berlin, Germany) on NHS activated glass slides - Nexterion® Slide H - were purchased from Schott AG, Mainz, Germany. Fluorescence measurements were performed in an Agilent G2565BA microarray scanner system (Agilent Technologies, Santa Clara, USA) with 10 μ m resolution, using 2 lasers (532 nm, 633

nm). Quantification of fluorescence was achieved by ScanArray[®] Express software from Perkin Elmer, Shelton, USA. Quantification of spots was performed using an adaptive circle with minimum diameter of 50 μm and maximum diameter of 200 μm . Histograms represents average of mean fluorescence of spots with local background subtraction.

SI-Table 1 - MALDI-TOF m/z values before and after conjugation with BCN-amine

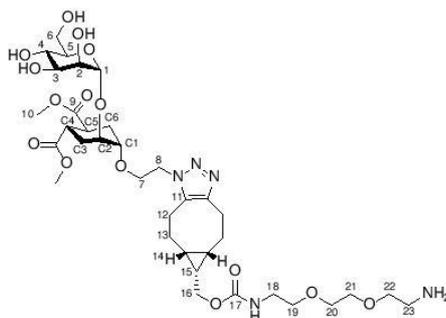
Ligand (before conjugation)	MS calculated [M+Na] ⁺	MS (MALDI-TOF)	Ligand (after conjugation)	MS (MALDI-TOF)
12.1	696.286	696.564	17.1	1020.981
12.2	486.169	486.330	17.2	810.749
12.3	696.286	696.580	17.3	1020.976
12.4	696.286	696.576	17.4	1020.969
12.5	696.286	696.561	17.5	1020.956
12.6	756.306	756.654	17.6	1081.046
12.7	752.276	752.580	17.7	1076.973
12.8	745.218*	745.0625*	17.8	1044.960
12.9	636.264	636.504	17.9	960.908
12.10	664.295	664.555	17.10	988.952
12.11	638.254	638.508	17.11	962.912
12.12	752.348	752.778	17.12	1077.195
12.13	788.327	788.689	17.13	1113.082
12.14	732.267	732.661	17.14	1057.071
12.15	732.267	732.643	17.15	1057.054
12.16	540.264	540.585	17.16	865.058
12.17	800.239	800.774	17.17	1125.236
12.18	720.286	720.609	17.18	1045.005
12.19	696.286	696.586	17.19	1020.992
12.20	756.401	756.741	17.20	1081.135
12.21	796.339	796.710	17.21	1121.105
12.22	780.251	780.600	17.22	1104.993
12.23	780.251	780.604	17.23	1104.999
12.24	536.233	536.435	17.24	860.847
12.25	706.338	706.646	17.25	1031.055
12.26	700.244	700.559	17.26	1024.960
12.27	688.295	688.601	17.27	1013.009
12.28	536.233	536.457	17.28	860.880
12.29	786.323	786.718	17.29	1111.120
12.30	596.327	596.587	17.30	921.001
12.31	668.254	668.554	17.31	992.966
12.32	742.317	742.661	17.32	1067.061
12.33	760.252	760.608	17.33	1085.012
12.34	660.252	660.542	17.34	984.942
12.35	596.254	596.580	17.35	921.00
12.36	564.264	564.546	17.36	888.979
12.37	810.361	810.815	17.37	1135.218
12.38	622.317	622.625	17.38	947.049
12.39	724.317	724.688	17.39	1049.093
15.1	470.175	470.134	18.1	794.258
15.2	680.290	680.185	18.2	1004.246
15.3	764.256	764.131	18.3	1088.303

15.4	652.259	652.160	18.4	976.320
15.5	726.332	726.177	18.5	1050.353
15.6	580.259	580.182	18.6	904.330
15.7	622.260	622.173	18.7	946.326
15.8	520.238	520.177	18.8	844.316
15.9	606.332	606.235	18.9	930.388
15.10	736.280	736.158	18.10	1032.297
15.11	708.249	708.129	18.11	1032.297

*MS (ESI-TOF) calculated for $[C_{32}H_{37}N_5O_{13}Na_2]^+ = [M-2H^+ + 2Na]^+$

Conjugate **17.2** was fully characterized for analytical purpose and the characterization is reported below

17.2



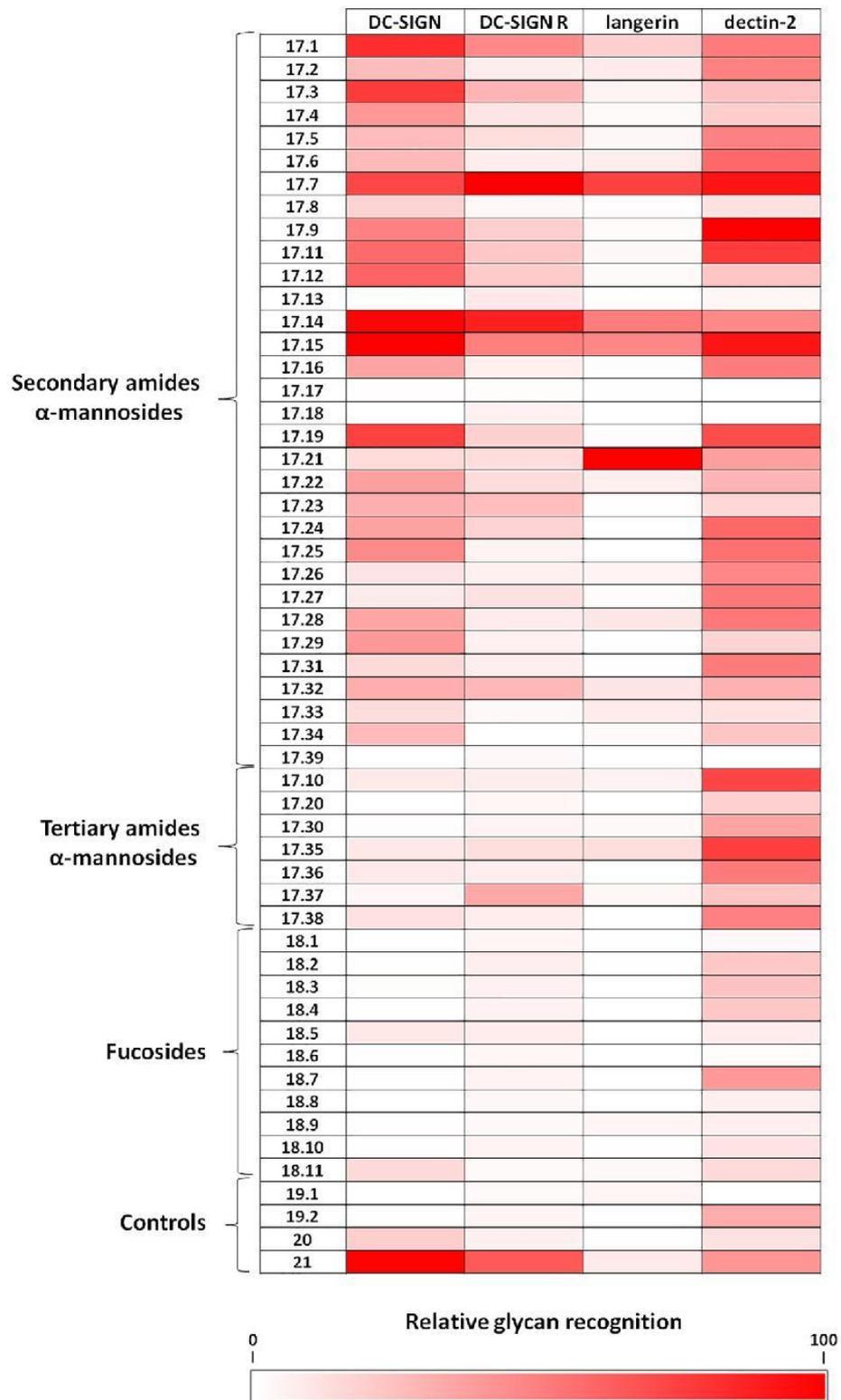
Prepared according to the general conditions used for SPAAC-couplings (Experimental Section).

¹H NMR (400 MHz, D₂O): δ = 4.95 (s, 1H, H₁), 4.55 – 4.47 (m, 2H, H₈), 4.19 (d, 2H, H₁₆, J_{16-15} = 7.6 Hz), 4.02 – 3.54 (m, 21H, C₂, H₂, H_{7a,b}, C₁, H₃, H₄, H₁₀, H₁₉, H₂₀, H₂₁, H₂₂), 3.14 – 2.97 (m, 4H, H₁₈, H₂₃), 2.91 – 2.75 (m, 2H, C₄, C₅), 2.12 – 2.32 (m, 4H, C₃, C₆), 2.09 – 1.84 (m, 2H, H_{12a}), 1.74 – 1.51 (m, 4H, H_{12b}, H_{13a}), 1.47 – 1.20 (m, 3H, H₁₅, H_{13b}), 1.15 – 0.94 (m, m, 2H, H₁₄)

¹³C NMR (100 MHz, D₂O): δ = 177.4, 177.0 (C₉); 158.8 (C₁₇); 144.9 (C_{11a}); 136.0 (C_{11b}); 98.5 (C₁); 73.6, 73.4 (C₅, C₃); 70.7, 70.4 (C₁, C₂, C₂); 69.5, 69.3, 69.2 (C₁₉, C₂₀, C₂₁ C₂₂); 66.7 (C₄); 66.6 (C₇); 63.7 (C₆); 60.9 (C₁₆); 52.6, 52.5 (C₁₀); 47.7 (C₈); 39.9, 39.5 (C₂₃, C₁₈); 38.7, 38.6 (C₄, C₅); 28.5 (C₁₂, C₁₃); 26.6, 26.3 (C₃, C₆); 20.7 (C₁₂, C₁₃); 19.7 (C₁₅); 17.0 (C₁₄)

MS (MALDI-TOF) calculated for $[C_{35}H_{57}N_5O_{15}Na]^+$: 810.375; found: 810.749

Figure SI-3- Heat map for CLR screenings



Protein preparation and labeling.

Cloning

Standard pUC57 plasmids containing optimized synthetic human genes encoding human DC-SIGNR ECD (amino acids 78-399) and dectin-2 ECD (amino acids 46-209), designed for the efficient production in *E. coli*, were manufactured by GeneCust Europe (Luxembourg). PCR amplification using suitable primers and restriction enzyme digestion were used to sub-clone DC-SIGNR ECD into the pET30-b (Novagen) between the NdeI and HindIII restriction sites and N-terminally strep tagged dectin-2 ECD into pET30-b (Novagen) between the XbaI and HindIII ones. The sequencing of each construction was done by Genewiz (Takeley, UK).

Protein Expression and Purification

DC-SIGN extracellular domain (DC-SIGN ECD) and langerin extracellular domain (langerin ECD) constructs were produced and purified as previously described.^{1,2}

DC-SIGNR ECD and dectin-2 ECD was expressed in *E. coli* BL21(DE3) in 1 liter of LB medium supplemented with 50 µg/ml kanamycin at 37 °C. Expression was induced by addition of 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG) when the culture had reached an $A_{600\text{ nm}}$ of 0.8 and maintained for 3h. The protein was expressed in the cytoplasm as inclusion bodies. Cells were harvested by a 20-min centrifugation at 5000 g at 4 °C. The pellet was resuspended in 30 mL of a solution containing 150 mM NaCl, 25 mM Tris-HCl, pH 8 and one anti-protease mixture tablet (Complete EDTA free, Roche). Cells were disrupted by sonication and cell debris eliminated by centrifugation at 100,000 g for 45 min at 4 °C in a Beckman 45Ti rotor. The pellet was solubilized in 30 mL of 6 M guanidine-HCl containing 25 mM Tris-HCl pH 8, 150 mM NaCl and 0,01% beta-mercaptoethanol. The mixture was centrifuged at 100,000g for 45 min at 4°C and the supernatant was diluted 5-fold, by slow addition with stirring, with 1.25 M NaCl, 25 mM CaCl₂ and 25 mM or 200 mM Tris-HCl pH 8 for DC-SIGNR and dectin-2 ECD respectively. The diluted mixture was dialyzed against 10 volumes of 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mM CaCl₂ (buffer A) with 3 buffer changes. After dialysis, insoluble precipitate was removed by centrifugation at 100,000g for 1h at 4°C.

The supernatant containing DC-SIGNR ECD was loaded on Mannan agarose column (Sigma) for purification by affinity chromatography equilibrated with buffer A. After loading, DC-SIGNR ECD was tightly bound to the column and eluted in the same buffer without CaCl₂ but supplemented with 1 mM EDTA (buffer B). This step was followed by SEC (Size Exclusion Chromatography) using a Superose 6 column (GE Healthcare) equilibrated with buffer A. Fractions were analyzed by SDS-PAGE (12%) and DC-SIGNR ECD containing fractions were pooled and concentrated by ultrafiltration (YM10 membrane from Amicon).

The supernatant containing the Strep tagged dectin-2 ECD was loaded onto a StrepTrap HP column (GE Healthcare) at 4°C. Unbound proteins were washed away with buffer A before dectin-2 ECD was eluted with buffer C (150 mM NaCl, 25 mM Tris-HCl, pH 8, 4 mM CaCl₂, 2.5 mM D-desthiobiotin). Eluted fractions were analyzed by SDS-PAGE (15%) and the dectin-2 ECD containing fractions were pooled and concentrated by ultrafiltration (YM10 membrane from Amicon).

Each protein construct was checked by N-terminal amino acid sequencing and mass spectrometry.

Cy3 Labeling of DC-SIGN ECD, langerin ECD, DC-SIGNR ECD and dectin-2 ECD

1mg of DC-SIGN, langerin, and DC-SIGNR ECDs and 0.25 mg of dectin-2 ECD in 25 mM HEPES pH 7,25, 4 mM CaCl₂ were prepared. 1 µL of 10 mg/ml Cy3-NHS ester (Gene Copoeia) were added per 0.3 mg of protein for

the ECDs of DC-SIGN, langerin, DC-SIGNR, while 3 μL of Cy3-NHS were added per 0.3 mg of dectin-2 ECD. The reactions were gently shaken at RT for 2 h and then at 4°C for 4h. Excess dye was removed by two dialysis (3.5k Z-lyser from Thermo Scientific) of 3h against 25 mM Tris pH8, 150 mM NaCl, 4 mM CaCl_2 . The amount of attached Cy3 was estimated spectrophotometrically based on the dye molar extinction coefficient ($\epsilon_{555}=150\,000\text{ cm}^{-1}\text{M}^{-1}$) and protein extinction coefficients (DC-SIGN ECD $\epsilon_{280}=70400\text{cm}^{-1}\text{M}^{-1}$; langerin ECD $\epsilon_{280}=56170$, DC-SIGNR ECD $\epsilon_{280}=60890.\text{M}^{-1}.\text{cm}^{-1}$, dectin-2 ECD $\epsilon_{280}=65930\text{ cm}^{-1}\text{M}^{-1}$).

The obtained degree of labeling (DOL) was 0.95 for both DC-SIGN ECD and DC-SIGNR ECD, 0.75 for langerin ECD and 0.42 for dectin-2 ECD.

1. G.Tabarani, et al. DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain *J Biol Chem* **2009**, *284*, 21229-21240
2. M. Thepaut, et al. Structural studies of langerin and Birbeck granule: a macromolecular organization model *Biochemistry*, **2009**, *48*,: 2684-2698

Surface Plasmon Resonance (SPR) analysis

The SPR experiments were performed on a BIAcore T200 using a CM3 sensor chip. Flow cells were activated as previously described.¹ Flow cell 1 was functionalized with BSA, blocked with ethanolamine and subsequently used as a control surface. Flow cells 2 and 3 were treated with BSA-Man α 1-3[Man α 1-6]Man (Dextra) (60 μ g/ml) in 10 mM NaOAc pH 4 to reach different binding densities and blocked with ethanolamine. The final densities on flow cells 2 and 3 were 2126 and 2085 RU, respectively. The affinity of the various compounds for DC-SIGN ECD were evaluated via an established inhibition assay² in which DC-SIGN ECD was injected at 20 μ M alone or in the presence of increasing concentration of inhibitors (ranging from 0 to 5 mM). Injections were performed at 5 μ L/min using 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl₂, 0,05% P20 surfactant as running buffer. The surface was regenerated by the injection of 50 mM EDTA.

1. F. Halary, A. Amara, H. Lortat-Jacob *et al. Immunity* **2002**, *17*, 653
2. M. Andreini, D. Doknic, I. Sutkeviciute *et al. Org Biomol Chem.* **2011**, *9*, 5778

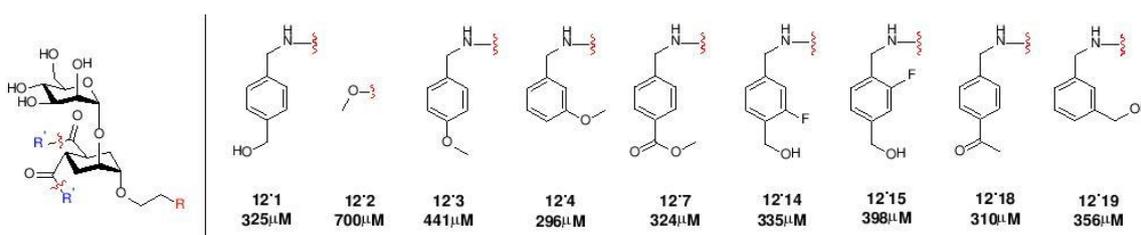


Figure SI-4. Structure of the red-bar ligands in Figure 3A. The IC₅₀ values measured by SPR in *ref 11c* for the parent recognition element **12.n** are shown below the structures

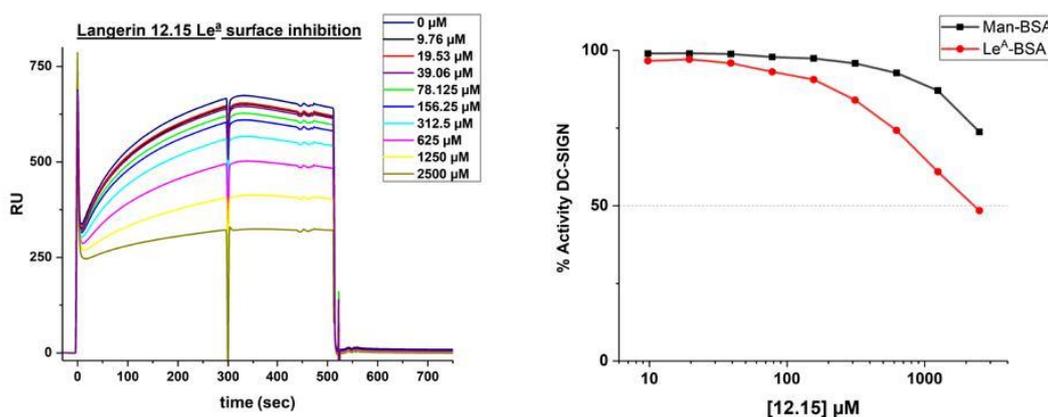
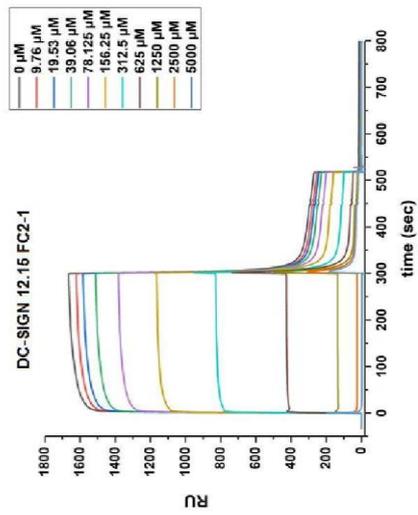
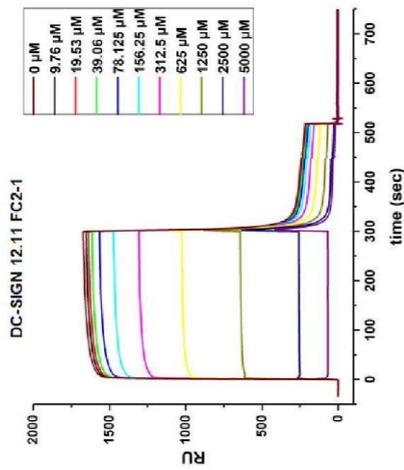
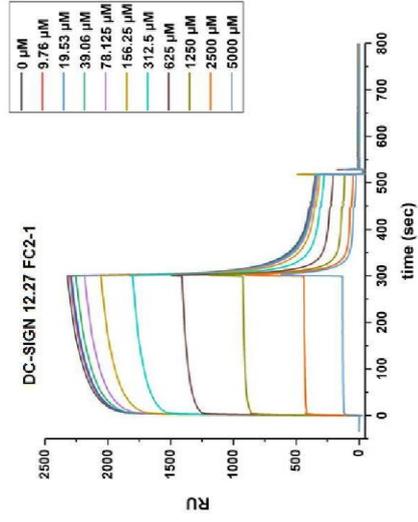
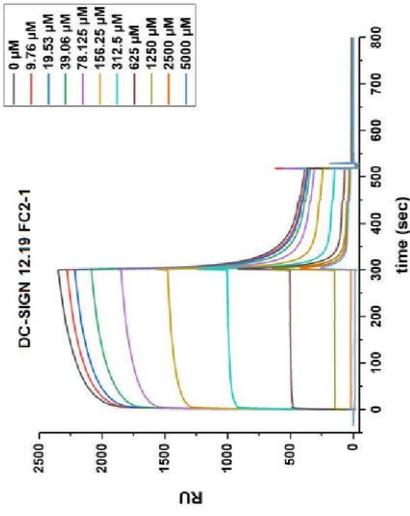
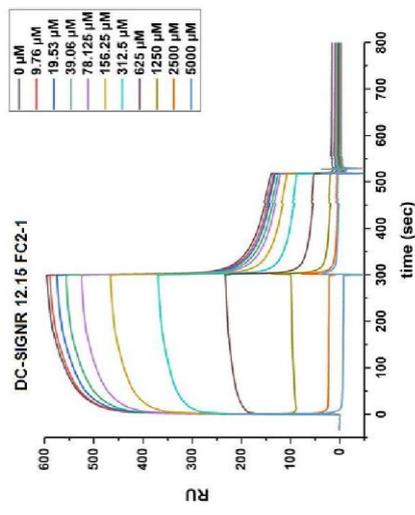
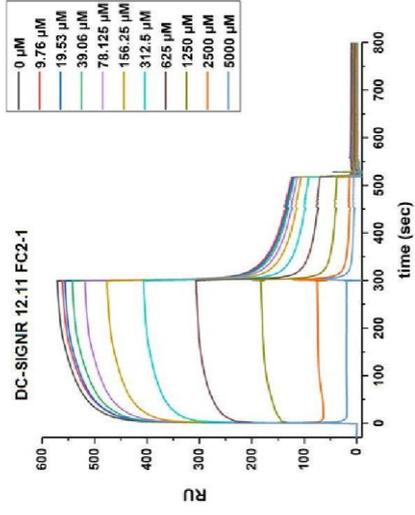
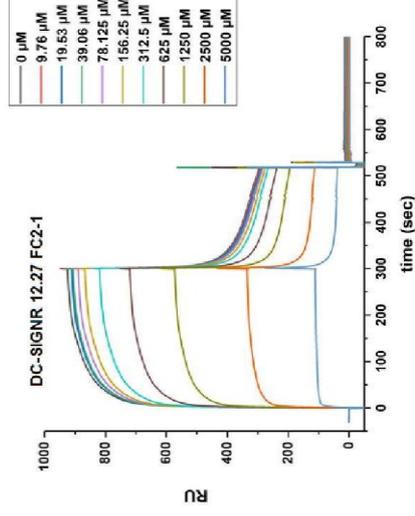
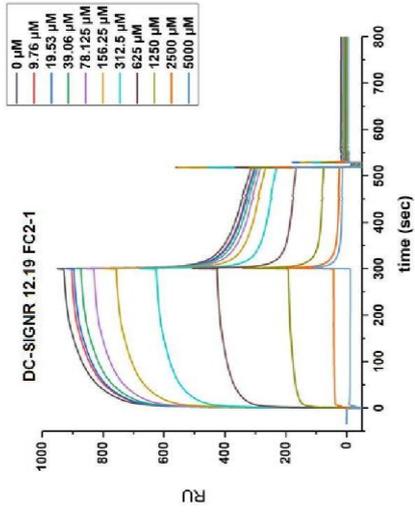


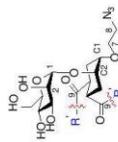
Figure SI-5. Left panel: SPR competition assay of **12.15** toward the langerin ECD/Lea-BSA interaction. Right panel: Comparison of inhibitory potency of **12.15** towards langerin interaction with Man-BSA or Lea-BSA.





Characterization of the ligands

Ligands 12.1-39



The numbering used for the NMR spectra is shown in the pictures. The unconventional numbering of the cyclohexane ring is used in analogy to sugar numbering.

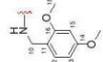
12.5



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF overnight, with the addition of 3.2 mol equiv. Et₃N.

Yield: 65% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 7.26 - 7.20 (t, 2H, H₁₃, J_{13,14} = 7.5 Hz, J_{13,12} = 7.9 Hz), 7.19 - 7.14 (d, 2H, H₁₅, J_{15,14} = 7.4 Hz), 6.97 - 6.91 (d, 2H, H₁₂, J_{12,13} = 7.9 Hz), 6.91 - 6.84 (t, 2H, H₁₄, J_{14,15} = 7.4 Hz, J_{14,13} = 7.5 Hz), 4.95 (d, 1H, H₁, J_{1,2} = 1.4 Hz), 4.35 (dd, 2H, H_{10a}, J_{10a,10b} = 4.3 Hz, J_{10a,10c} = 15.3 Hz), 4.23 (dd, 2H, H_{10b}, J_{10b,10c} = 3.9 Hz, J_{10b,10a} = 15.3 Hz), 4.07 - 4.03 (m, 1H, C₃), 3.92 (dd, 1H, H₂, J_{2,3} = 1.4 Hz, J_{2,1} = 3.1 Hz), 3.89 - 3.84 (m, 7H, H_{6a}, H₇), 3.83 - 3.65 (m, 5H, C₄, H_{6b}, H_{7a,b}, H₈), 3.62 - 3.57 (m, 2H, H₄, H₅), 3.43 - 3.37 (m, 2H, H_{8a,b}), 2.99 - 2.84 (m, 2H, C₆, C₅), 2.05 - 1.88 (m, 4H, C₉, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 177.8, 177.6 (C₁); 159.5 (C₁₀); 130.3, 130.3 (C₁); 130.1, 130.0 (C₁); 128.3 (C₁); 122.3, 122.3 (C₁₄); 112.1, 112.1 (C₁₂); 101.1 (C₁); 77.4 (C₃); 76.4 (C₃); 73.4 (C₃); 73.2 (C₁); 73.1 (C₁); 70.6 (C₃); 69.6 (C₃); 63.9 (C₆); 56.6 (C₇); 52.8 (C₃); 42.9, 42.8 (C₁₀); 40.2, 40.1 (C₆, C₅); 30.5, 29.6 (C₆, C₅)
 MS (MALDI-TOF) calculated for [C₂₃H₂₉N₂O₇Na]⁺: 696.285; found: 696.561

12.6



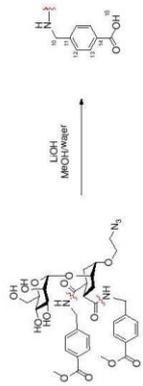
Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF overnight, with the addition of 3.2 mol equiv. Et₃N.

Yield: 60% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 7.02 (d, 2H, H₁₂, J_{12,13} = 8.3 Hz), 6.48 (d, 2H, H₁₅, J_{15,13} = 2.2 Hz), 6.38 (dt, 2H, H₁₃, J_{13,15} = 2.2 Hz, J_{13,12} = 8.3 Hz), 4.91 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.26 (dd, 2H, H_{10a}, J_{10a,10b} =

15

15.0 Hz), 4.15 (dd, 2H, H_{10b}, J_{10b,10a} < 1.0 Hz, J_{10b,10c} = 15.0 Hz), 4.01 (m, 1H, C₂), 3.89 - 3.86 (m, 2H, H₂, H₆), 3.79 (s, 6H, H₁₈), 3.73 (s, 6H, H₁₇), 3.86 - 3.60 (m, 5H, C₁, H_{6a}, H_{7a,b}, H₈), 3.60 - 3.50 (m, 2H, H₄, H₅), 3.39 - 3.32 (m, 2H, H_{8a,b}), 2.91 - 2.76 (m, 2H, C₄, C₅), 2.00 - 1.81 (m, 4H, C₉, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 177.6, 177.5 (C₁); 172.6 (C₁₀); 160.5 (C₁₄); 131.1, 131.0 (C₁₃); 120.6 (C₁₁); 106.1, 106.0 (C₂); 101.0 (C₁); 100.0 (C₁₅); 77.4 (C₃); 76.4 (C₃); 73.4 (C₃); 73.2 (C₁); 73.1 (C₁); 70.0 (C₁); 69.6 (C₄); 63.9 (C₆); 56.7, 56.6 (C₇, C₈); 42.9, 42.8 (C₁₀); 39.9, 39.8 (C₆, C₅); 30.5, 29.5 (C₆, C₅)
 MS (MALDI-TOF) calculated for [C₂₈H₃₇N₂O₁₁Na]⁺: 756.306; found: 756.654

12.8



127

128

12.7 (14.1 mg, 0.02 mmol, 1 equiv.) was dissolved in dry MeOH and 3 equiv. aqueous LiOH solution was added to the mixture. Upon completion, the reaction was concentrated *in vacuo*.

Yield: quantitative.

¹H NMR (400 MHz, CD₃OD): δ = 10.04 (s, 2H, H₁₂), 7.82 (d, 4H, H₁₃, J_{13,12} = 7.4 Hz), 7.2 (d, 4H, H₂, J_{2,3} = 7.4 Hz), 4.99 (s, 1H, H₁), 4.60 (dd, 2H, H_{10a}, J_{10a,10b} < 1 Hz, J_{10a,10c} = 16.1 Hz), 4.38 (dd, 2H, H_{10b}, J_{10b,10c} < 1 Hz, J_{10b,10a} = 16.1 Hz), 4.08 - 4.04 (m, 1H, C₂), 3.95 - 3.92 (m, 1H, H₂), 3.88 - 3.54 (m, 8H, H_{6a,b}, C₁, H_{7a,b}, H₈, H₉), 3.47 - 3.34 (m, 2H, H_{8a,b}), 3.07 - 2.91 (m, 2H, C₄, C₅), 2.08 - 1.91 (m, 4H, C₉, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 178.3, 178.0 (C₁); 176.6 (C₁₀); 146.9 (C₁₄); 137.8 (C₁₁); 130.6 (C₁₃); 129.2 (C₁₂); 101.1 (C₁); 77.3 (C₃); 76.3 (C₃); 73.4 (C₃); 73.3 (C₂); 73.2 (C₁); 70.1 (C₁); 69.6 (C₄); 63.7 (C₆); 52.9 (C₆); 44.4 (C₆); 42.7, 42.6 (C₆, C₅); 31.6, 30.7 (C₆, C₅)
 MS (ESI-TOF) calculated for [C₂₃H₂₇N₂O₁₁Na]⁺: 745.2183; found: 745.0625 [M-2H⁺+2Na⁺]

12.9



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

Yield: 88% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.29 - 7.16 (m, 10H, H₁₂, H₁₃, H₁₄), 4.93 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.32 - 4.22 (m, 4H, H_{6a,b}), 4.04 - 4.00 (m, 1H, C₂), 3.90 - 3.86 (m, 1H, H₂), 3.86 - 3.80 (m, 1H, H_{6a}), 3.80 - 3.62 (m, 5H, C₁, H_{6b}, H_{7a,b}, H₈), 3.60 - 3.51 (m, 2H, H₄, H₅), 3.43 - 3.30 (m, 2H, H_{8a,b}), 2.98 - 2.83 (m, 2H, C₄, C₅), 2.01 - 1.86 (m, 4H, C₉, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 177.9, 177.7 (C₁); 140.9 (C₁); 130.3, 129.3, 129.2 (C₁₂, C₁₃, C₁₄); 101.1 (C₁); 77.3 (C₃); 76.4 (C₃); 73.4 (C₃); 73.3 (C₂); 73.2 (C₁); 70.0 (C₁); 69.6 (C₄); 63.9 (C₆); 52.8 (C₆); 44.8, 44.7 (C₆); 42.8, 42.6 (C₆, C₅); 30.6, 29.8 (C₆, C₅)
 MS (MALDI-TOF) calculated for [C₂₈H₃₇N₂O₁₁Na]⁺: 636.264; found: 636.504

16

12.10

Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 2 hours.

Yield: 74% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.42 – 7.16 (m, 10H, H₁₂, H₁₃, H₁₄), 4.98 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.82 – 4.65 (m, 4H, H_{10a,b}), 4.15 – 4.10 (m, 1H, C₃), 4.07 – 3.98 (m, 2H, H₂, H_{6a}), 3.89 – 3.44 (m, 5H, C₁, H_{6b}, H_{7a,b}, H₈), 3.43 – 3.33 (m, 2H, H₇, H₈), 3.33 – 3.27 (m, 2H, H_{10a,b}), 3.27 – 3.13 (m, 6H, H₁₅), 2.91 – 2.83 (m, 2H, C₄, C₅), 2.11 – 1.76 (m, 4H, C₂, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 178.4, 178.1 (C₉), 139.2 (C₁₁), 130.5, 129.4, 129.1 (C₂, C₃, C₁₀), 101.6 (C₁); 77.4 (C₄); 76.6 (C₅); 73.5 (C₆); 73.4 (C₇); 69.6 (C₈); 64.0 (C₁₂); 53.0, 52.8 (C₁₃, C₁₄); 49.9 (C₁₅); 39.3, 36.3 (C₁₆, C₁₇); 30.3, 29.1 (C₁₈, C₁₉)

MS (HRMS) calculated for [C₂₂H₂₈N₂O₂Na]⁺: 664.2958; found: 664.3009

12.11

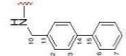
Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 4 hours.

Yield: 54% over two steps

¹H NMR (400 MHz, D₂O): δ = 8.37 (br s, 4H, H₂₇, H₁₃), 7.71 (d, 2H, H₁₀, J_{10,11} = J_{10,12} = 7.6 Hz), 7.39 (br s, 2H, H₁₂), 5.02 (d, 1H, H₁, J_{1,2} = 1.6 Hz), 4.43 – 4.24 (m, 4H, H₁₀), 4.10 – 4.05 (m, 1H, C₃), 4.04 – 4.00 (m, 1H, H₂), 3.91 – 3.85 (m, 7H, H₁₀, H_{7a,b}, H₈, H₅, C₁), 3.48 (t, 2H, H_{6a,b}, J = 4.9 Hz), 2.94 – 2.83 (m, 2H, C₄, C₅), 2.10 – 1.85 (m, 4H, C₂, C₆)

¹³C NMR (HSQC, 100 MHz, CD₃OD): δ = 177.4, 177.1 (C₉); 147.2 (C₁₁); 147.0 (C₁₀); 145.9 (C₁); 140.1 (C₂); 127.2 (C₃); 101.9 (C₁); 76.9 (C₄); 74.0 (C₅); 73.8 (C₆); 72.5 (C₇); 71.8 (C₈); 69.0 (C₁₂); 63.3 (C₁₃); 52.8 (C₁₄); 43.8, 43.6 (C₁₅, C₁₆); 42.7 (C₁₇); 30.2, 30.0 (C₁₈, C₁₉)

MS (MALDI-TOF) calculated for [C₂₆H₃₂N₂O₂Na]⁺: 638.254; found: 638.508

12.13

Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF overnight.

Yield: 59% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.60 (t, 8H, H₁₉, H₁₀, J_{10,11} = J_{10,12} = 7.4 Hz), 7.48 (t, 4H, H₁₂, J_{12,13} = 7.4 Hz), 7.40 (d, 6H, H₁₇, H₁₈, J_{17,18} = 7.8 Hz), 4.98 – 4.96 (d, 1H, H₁, J_{1,2} = 1.3 Hz), 4.40 (dd, 2H, H₁₀, J_{10,11} = 4.0 Hz, J_{10,12} = 15.3 Hz), 4.30 (dd, 2H, H₁₀, J_{10,11} < 1 Hz, J_{10,12} = 15.3 Hz), 4.09 – 4.05 (m, 1H, C₃), 3.94 – 3.90 (m, 1H, H₂), 3.90 – 3.84 (m, 1H, H_{6a}), 3.84 – 3.63 (m, 5H, C₁, H_{6b}, H_{7a,b}, H₈), 3.62 – 3.56 (m, 2H, H₄, H₅), 3.47 – 3.37 (m, 2H, H_{10a,b}), 3.05 – 2.90 (m, 2H, C₄, C₅), 2.09 – 1.91 (m, 4H, C₂, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 178.1, 177.9 (C₉); 143.1, 142.2 (C₁₀, C₁₁); 140.1 (C₁); 130.8 (C₂); 129.8 (C₁₇, C₁₈), 129.2, 129.0, 128.8 (C₃, C₁₆), 101.2 (C₁); 77.4 (C₄); 76.4 (C₅); 73.5 (C₆); 73.3 (C₇); 73.3 (C₈); 70.1 (C₁₂); 69.7 (C₁₃); 64.0 (C₁₄); 44.6 (C₁₅); 42.9, 42.8 (C₁₆, C₁₇); 30.7, 29.9 (C₁₈, C₁₉)

MS (MALDI-TOF) calculated for [C₂₂H₂₇N₂O₂Na]⁺: 788.327; found: 788.689

12.16

Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

¹H NMR (400 MHz, CD₃OD): δ = 4.98 (d, 1H, H₁), 4.10 – 4.06 (m, 1H, C₃), 3.94 (s, 1H, H₂), 3.87 (s, 2H, H_{6a,b}), 3.83 – 3.64 (m, 6H, C₁, H_{7a,b}, H₈, H₄, H₅, H₁₀), 3.46 – 3.39 (m, 2H, H_{6a,b}), 3.20 – 3.04 (m, 4H, H₁₀), 2.91 – 2.79 (m, 2H, C₄, C₅), 2.05 – 1.84 (m, 4H, C₂, C₆), 1.57 – 1.46 (sex, 4H, H₁₁, J_{11,12} = 7.3 Hz), 0.97 – 0.89 (t, 6H, H₁₂, J₁₂ = 7.3 Hz)

¹³C NMR (100 MHz, CD₃OD): δ = 176.9 (C₉); 101.3 (C₁); 77.4 (C₄); 76.5 (C₅); 73.4 (C₆); 73.2 (C₇, C₈); 70.1 (C₁₂); 69.8 (C₁₃); 63.9 (C₁₄); 53.0 (C₁₅); 33.2 (C₁₆); 33.0 (C₁₇, C₁₈, C₁₉); 30.7, 29.8 (C₂₀, C₂₁); 24.5 (C₂₂); 12.7 (C₂₃)

MS (MALDI-TOF) calculated for [C₂₂H₂₉N₂O₂Na]⁺: 540.264; found: 540.585

12.17

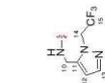
Amine purchased from LifeChemicals. Coupling reaction performed in DMF overnight.

¹H NMR (400 MHz, CD₃OD): δ = 8.04 – 7.97 (m, 4H, H₁₆, H₁₇), 7.63 – 7.15 (m, 10H, H₁₂, H₁₃, H₁₈), 4.93 – 4.86 (m, 1H, H₁), 4.42 – 4.27 (dd, 4H, H_{10a,b}), 4.05 (s, 1H, C₃), 3.94 (s, 1H, H₂), 3.86 – 3.84 (m, 2H, H_{6a,b}), 3.78 – 3.65 (m, 4H, H₇, C₄, H_{7a,b}), 3.60 – 3.50 (m, 2H, H₄, H₅), 3.40 – 3.35 (m, 2H, H_{10a,b}), 2.94 – 2.79 (m, 2H, C₄, C₅), 1.93 – 1.80 (m, 4H, C₂, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 177.8 (C₉); 143.2 (C₁₁); 138.4 (C₁₀); 137.8 (C₁); 130.7, 130.0, 129.8, 129.5, 129.0, 127.5, 124.9 (C₂, C₃, C₁₆, C₁₇, C₁₈); 101.0 (C₁); 77.3 (C₄); 76.3 (C₅); 73.4 (C₆); 73.2 (C₇); 73.0 (C₈); 70.0 (C₁₂); 69.6 (C₁₃); 63.8 (C₁₄); 43.2 (C₁₅); 42.5 (C₁₆, C₁₇); 30.4, 29.6 (C₁₈, C₁₉)

MS (ESI) calculated for [C₂₄H₂₉N₂O₂Na]⁺: 800.239; found: 800.774

12.20



12.22

Amine purchased from Crea-Chim. Coupling reaction performed in DMF overnight, with the addition of 3.2 mol equiv. Et₃N.

Yield: 73% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 7.41 (t, 2H, H₁₃, J_{13,12} = 1.7 Hz), 6.26 (d, 2H, H₁₂, J_{12,13} = 1.7 Hz), 5.03 – 4.89 (m, 5H, H₁, H₂), 4.36 – 4.30 (m, 4H, H₁₀), 4.03 – 3.98 (m, 1H, C₃), 3.88 – 3.85 (m, 1H, H₂), 3.85 – 3.79 (m, 1H, H_{6a}), 3.79 – 3.49 (m, 5H, H₅, H_{6b}, H_{7a,b}, C₁), 3.42 – 3.31 (m, 4H, H₄, H₅, H_{6a,b}), 2.92 – 2.78 (m, 2H, C₆, C₇), 1.98 – 1.81 (m, 4H, C₅, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 178.1, 177.9 (C₉); 144.4 (C₁₁); 142.1 (C₁₃); 125.8 (C₁₀, J_{10,F} = 279.3 Hz); 108.9 (C₁₂); 101.2 (C₁); 77.2 (C₂); 76.4 (C₃); 73.5 (C₄); 73.2 (C₅); 70.1 (C₇); 69.6 (C₄); 63.9 (C₆); 52.9 (C₈); 51.8, 51.5, 51.1 (C₁₀); 42.4, 42.3 (C_{6a}, C_{6b}); 35.3 (C_{10b}); 30.5, 29.7 (C_{5a}, C_{5b})
 MS (MALDI-TOF) calculated for [C₂₈H₃₇F₃N₃O₃Na]⁺: 780.251; found: 780.600

12.23



Amine purchased from Vitas M Labs. Coupling reaction performed in DMF in 2 days, with the addition of 3.2 mol equiv. Et₃N.

Yield: 65% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 6.57 (s, 2H, H₁₂), 4.95 (d, 1H, H₁, J_{1,2} = 1.4 Hz), 4.34 – 4.19 (m, 4H, H₁₀), 4.05 – 4.01 (m, 1H, C₃), 3.91 (s, 6H, H₁₅), 3.89 – 3.52 (m, 7H, H₂, H₅, H_{6a,b}, H_{7a,b}, C₁), 3.44 – 3.52 (m, 4H, H₄, H₅, H_{6a,b}), 2.97 – 2.82 (m, 2H, C₆, C₇), 2.02 – 1.86 (m, 4H, C₅, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 178.1, 177.9 (C₉); 151.6 (C₁₃); 134.2 (C₁₁); 122.3 (C₁₀, J_{10,F} = 269.1 Hz); 108.1 (C₁₂); 101.1 (C₁); 77.3 (C₂); 76.4 (C₃); 73.5 (C₄); 73.1 (C₅); 70.1 (C₇); 69.6 (C₄); 63.9 (C₆); 52.9 (C₈); 42.6, 42.5 (C_{6a}, C_{6b}); 39.1 (C₁₀); 38.3 (C_{10b}); 30.5, 29.7 (C_{5a}, C_{5b})
 MS (MALDI-TOF) calculated for [C₂₈H₃₇F₆N₃O₃Na]⁺: 780.251; found: 780.604

12.24



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

Yield: 89% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 4.92 (d, 1H, H₁, J_{1,2} = 1.6 Hz), 4.01 – 3.97 (m, 1H, C₃), 3.89 – 3.86 (m, 1H, H₂), 3.86 – 3.80 (m, 1H, H_{6a}), 3.80 – 3.50 (m, 5H, C₁, H_{6b}, H₅, H₆, H_{7a,b}), 3.43 – 3.31 (m, 2H, H_{6a,b}), 2.79 – 2.64 (m, 2H, C₄, C₅), 2.61 – 2.52 (sept d, 2H, H₁₀, J = 1.0 Hz), 1.94 – 1.76 (m, 4H, C₆, C₇), 0.71 – 0.63 (m, 4H, H_{11a}), 0.47 – 0.40 (m, 4H, H_{11b})

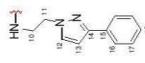
20



Amine purchased from Chem Bridge. Coupling reaction performed in DMF in 2 days, with the addition of 3.2 mol equiv. Et₃N.

Yield: 56% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 7.75 – 7.27 (m, 4H, H₁₅, H₁₆), 4.91 (s, 1H, H₁), 4.78 – 4.21 (m, 8H, H₁₀, H₁₃), 4.08 – 3.98 (m, 1H, C₃), 3.91 – 2.83 (m, 19H, H₂, H_{6a,b}, C₁, H_{7a,b}, H₅, H₆, H₅, H_{6a,b}, H₁₇, C₆, C₇), 2.18 – 1.73 (m, 6H, C₅, C₆, H₁₁), 1.01 – 0.92 (m, 6H, H_{12a}), 0.88 – 0.77 (m, 6H, H_{12b})
¹³C NMR (100 MHz, CD₃OD): δ = 178.1, 177.7 (C₉); 140.7, 140.0 (C₁₃); 132.5, 132.3 (C₁₀); 120.3, 120.2 (C₃); 101.9 (C₁); 77.4 (C₂); 76.6 (C₃); 73.9 (C₄); 73.6 (C₅); 73.4 (C₆); 70.6 (C₇); 69.7 (C₄); 64.0 (C₆); 57.2, 54.4 (C_{6a}, C_{6b}); 53.1 (C₈); 44.3, 44.2, 43.1, 42.9 (C_{6a}, C₁₃); 39.9, 39.7, 38.6, 39.1 (C₇); 31.2, 30.4 (C₆, C₇); 30.1, 28.8 (C₁); 21.7, 21.5, 21.4, 21.2 (C₅)
 MS (MALDI-TOF) calculated for [C₅₄H₅₅N₄O₃Na]⁺: 756.401; found: 756.741

12.21



Amine purchased from Vitas M Labs. Coupling reaction performed in DMF overnight, with the addition of 3.2 mol equiv. Et₃N.

Yield: 62% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 7.76 – 7.72 (m, 4H, H₁₆), 7.59 (dd, 2H, H₁₂, J_{12,13} = 2.2 Hz, J = 5.3 Hz), 7.39 – 7.33 (m, 4H, H₁₇), 7.29 – 7.23 (m, 2H, H₁₀), 6.60 (t, 2H, H₁₃, J_{13,12} = 2.2 Hz), 4.93 – 4.91 (d, 1H, H₁, J_{1,2} = 1.6 Hz), 4.30 – 4.13 (m, 4H, H₁₀), 3.93 – 3.88 (m, 1H, C₃), 3.87 – 3.77 (m, 2H, H₅, H_{6a}), 3.71 – 3.42 (m, 11H, H₅, H₆, H₅, H_{6a,b}, H_{11a}, C₁, H_{6b}), 3.34 – 3.21 (H_{6a,b}), 2.82 – 2.68 (m, 2H, C₆, C₇), 1.95 – 1.68 (m, 4H, C₅, C₈)
¹³C NMR (100 MHz, CD₃OD): δ = 178.5, 178.3 (C₉); 154.2, 154.1 (C₁₃); 153.6, 153.6 (C₁₀); 134.4, 134.2 (C₃); 130.5 (C₇); 129.6 (C₁₀); 127.5 (C₁₀); 104.9, 104.8 (C₁₃); 101.1 (C₁); 77.3 (C₂); 76.3 (C₃); 73.4 (C₄); 73.3 (C₅); 73.2 (C₆); 69.9 (C₇); 69.6 (C₄); 63.8 (C₆); 52.8, 52.8 (C₆, C₁₀); 42.7, 42.1 (C_{6a}, C_{6b}); 41.7 (C₁₁); 30.6, 30.0 (C_{5a}, C_{5b})
 MS (MALDI-TOF) calculated for [C₅₀H₄₇N₄O₃Na]⁺: 796.339; found: 796.710

19

¹³C NMR (100 MHz, CD₃OD): δ = 179.5, 179.3 (C₃); 101.2 (C₁); 77.3 (C₅); 76.4 (C₆); 73.5 (C₂); 73.3 (C₂); 73.2 (C₄); 70.0 (C₇); 69.6 (C₈); 63.9 (C₉); 52.9 (C₁₀); 42.5, 42.4 (C_{6a}, C_{8a}); 30.3, 29.5 (C_{3a}, C_{10a}); 24.1 (C_{10b}), 7.4, 7.3 (C₁).

MS (MALDI-TOF) calculated for [C₂₂H₂₀N₂O₃Na]⁺: 536.233; found: 536.435

12.25



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

Yield: 87% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.94 (d, 1H, H₁, J_{1,2} = 1.4 Hz), 4.04 – 3.99 (m, 1H, C₃), 3.92 – 3.88 (m, 1H, H₂), 3.87 – 3.81 (m, 1H, H_{6a}), 3.81 – 3.62 (m, 5H, C₅, H_{6b}, H₅, H_{7a,b}), 3.61 – 3.53 (m, 2H, H₄, H₈), 3.45 (t, 4H, H₁₀, J_{6,7} = 7.1 Hz), 3.41 – 3.35 (m, 2H, H_{9a,b}), 3.29 – 3.22 (m, 4H, H₁₂), 3.11 (t, 4H, H₁₀, J_{10,11} = 6.8 Hz), 2.87 – 2.72 (m, 2H, C₉), 2.38 (t, 4H, H₁₀, J_{10,11} = 8 Hz), 2.04 (q, 4H, H₁₃), 1.98 – 1.85 (m, 4H, C₉), 1.72 – 1.62 (m, 4H, H₁₁)

¹³C NMR (100 MHz, CD₃OD): δ = 178.7 (C₉); 178.0, 177.8 (C₁₃); 101.2 (C₁); 77.3 (C₅); 76.4 (C₆); 73.4 (C₂); 73.3 (C₂); 69.7 (C₃); 63.9 (C₈); 52.9 (C₁₀); 49.6 (C_{10b}); 42.8, 42.7 (C_{6a}, C_{8a}); 42.1 (C₁₂); 38.5 (C_{10a}); 32.9 (C₁₂); 30.7, 29.8 (C_{3a}, C_{10a}); 28.8 (C₁₁); 19.3 (C₁₅)

MS (MALDI-TOF) calculated for [C₂₈H₃₀N₂O₄Na]⁺: 706.338; found: 706.646

12.26



Amine purchased from Sigma-Aldrich. Coupling reaction performed in DMF overnight, with the addition of 3.2 mol equiv. Et₃N.

Yield: 70% over two steps – Unstable compound

¹H NMR (400 MHz, CD₃OD): δ = 8.09 – 7.91 (m, 2H, H₁₀), 7.45 (q, 2H, H₁₃, J_{13,12} = 8.2 Hz), 6.75 – 6.61 (m, 2H, H₁₂), 5.00 – 4.65 (m, 5H, H₁, H₁₀), 4.23 – 4.13 (m, 1H, C₃), 4.13 – 3.06 (m, 11H, H₂, H_{6a,b}, C₅, H₅, H_{7a,b}, H₈, H₅, H_{9a,b}), 3.06 – 2.73 (m, 2H, C₉, C₈), 2.10 – 1.71 (m, 4H, C₇, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 178.4 (C₉); 164.9, 164.0 (C₆, C₁₃); 141.0 (C₁); 131.6, 130.6 (C_{6a}, C_{13a}); 116.7 (C₁); 100.3 (C₁); 76.6 (C₅); 76.3 (C₅); 73.5 (C₂); 73.3 (C₂); 73.2 (C₂); 69.7 (C₃); 69.1 (C₄); 63.5 (C₈); 52.1 (C₁₀); 45.1, 44.2 (C_{10b}); 42.5 (C_{6a}); 42.2 (C_{8a}); 31.4, 29.9 (C_{3a}, C_{10a})

MS (MALDI-TOF) calculated for [C₂₈H₃₀N₂O₄Na]⁺: 700.244; found: 700.559

Unstable – slow decomposition in solution.

12.27



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

Yield: 79% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.22 – 7.15 (m, 4H, H₁₀), 7.15 – 7.09 (m, 4H, H₁₃), 4.91 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.53 – 4.44 (m, 2H, H₁₀), 4.02 – 3.98 (m, 1H, C₃), 3.88 – 3.85 (m, 1H, H₂), 3.86 – 3.79 (m, 1H, H_{6a}), 3.79 – 3.60 (m, 5H, C₅, H_{6b}, H₅, H_{7a,b}), 3.59 – 3.50 (m, 2H, H₄, H₈), 3.42 – 3.31 (m, 2H, H_{9a,b}), 3.22 (quint, 4H, H_{11a}, J = 7.8 Hz), 2.86 – 2.71 (m, 6H, C₉, C₇, H_{11b}), 1.96 – 1.79 (m, 4H, C₉, C₈)

¹³C NMR (100 MHz, CD₃OD): δ = 177.9, 177.7 (C₉); 143.1, 143.0 (C₁); 128.6 (C₆); 126.4 (C₆); 101.1 (C₁); 77.3 (C₅); 76.4 (C₆); 73.4 (C₂); 73.2 (C₂); 73.1 (C₂); 70.0 (C₇); 69.6 (C₈); 63.9 (C₉); 52.8 (C₁₀); 52.7 (C₁₀); 42.7, 42.5 (C_{6a}, C_{8a}); 41.3, 41.2, 41.1 (C₁); 30.3, 29.4 (C_{3a}, C_{10a})

MS (MALDI-TOF) calculated for [C₂₈H₃₀N₂O₃Na]⁺: 688.295; found: 688.601

12.28



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

Yield: 90% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 5.84 – 5.71 (app oct, 2H, H₁₃, J = 5.4 Hz), 5.17 (dd, 2H, H_{13a}, J = 17.1 Hz, J = 1.5 Hz), 5.08 (dt, 2H, H_{13b}, J = 10.3 Hz, J = 1.3 Hz), 4.94 (d, 1H, H₁, J_{1,2} = 1.4 Hz), 4.06 – 3.99 (m, 1H, C₃), 3.92 – 3.77 (m, 7H, H₂, H_{6a,b}, C₅, H_{7a,b}, H₈), 3.77 – 3.63 (m, 2H, H₁₀), 3.61 – 3.50 (m, 2H, H₄, H₅), 3.44 – 3.31 (m, 2H, H_{9a,b}), 3.05 – 2.73 (m, 2H, C₉, C₇), 2.00 – 1.80 (m, 4H, C₉, C₈)

¹³C NMR (100 MHz, CD₃OD): δ = 177.9, 177.7 (C₉); 136.4 (C₁); 116.8 (C₁); 101.1 (C₁); 77.3 (C₅); 76.4 (C₆); 73.4 (C₂); 73.3, 73.2 (C₂); 70.0 (C₇); 69.6 (C₈); 63.9 (C₈); 52.9 (C₁₀); 43.4 (C₁₀); 42.8, 42.7 (C_{6a}, C_{8a}); 30.6, 29.7 (C_{3a}, C_{10a})

MS (ESI) calculated for [C₂₇H₂₈N₂O₃Na]⁺: 536.233; found: 536.457

12.29



Amine purchased from Enamine BB. Coupling reaction performed in DMF in two days, with the addition of 3.2 mol equiv. Et₃N.

Yield: 60% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.96 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.05 – 4.01 (m, 1H, C₃), 3.94 – 3.91 (m, 1H, H₂), 3.89 – 3.83 (m, 1H, H_{6a}), 3.83 – 3.64 (m, 5H, H₅, H_{6b}, H_{7a,b}, C₅), 3.64 – 3.53 (m, 4H, H₄, H₅, H₁₀), 3.47 – 3.34 (m, 2H, H_{9a,b}), 3.05 (q, 4H, H₁₃, J_{13,12} = 9.9 Hz), J = 19.6 Hz), 2.99 – 2.90 (m, 4H, H_{12a}), 2.86 – 2.72 (m, 2H, C₉, C₈)

2.46 (td, 4H, $H_{2,9}$, $J_{2,13} = 9.9$ Hz, $J_{2,11} = 3.0$ Hz), 1.98 – 1.85 (m, 4H, C_9 , C_6), 1.80 (t, 4H, $H_{11,11'}$, $J = 15.2$ Hz), 1.59 – 1.46 (m, 4H, H_{10})
¹³C NMR (100 MHz, CD_3OD): $\delta = 177.3$, 177.1 (C_6), 128.0 (C_{10} , $J_{10,11} = 277.7$ Hz), 101.2 (C_1); 77.4 (C_1); 76.4 (C_1); 73.4 (C_2); 73.3 (C_1); 73.2 (C_{10}); 69.6 (C_1); 63.9 (C_1); 59.7 (C_{13} , $J_{13,14} = 30.6$ Hz); 55.0, 54.9 (C_2); 52.9 (C_1); 48.3, 48.2 (C_{10}); 42.7, 42.6 (C_{10} , C_{13}); 33.6, 33.5, 33.4, 33.4 (C_{11}); 30.6, 29.7 (C_{10} , C_{13})
MS (MALDI-TOF) calculated for $[C_{38}H_{47}F_3N_3O_3Na]^+$: 786.323; found: 786.718

12.30



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 2 hours.
 Yield: 75% over two steps
¹H NMR (400 MHz, CD_3OD): $\delta = 4.99$ – 4.92 (m, 1H, H_1), 4.12 – 4.01 (m, 1H, C_1), 3.90 – 3.81 (m, 2H, H_2 , H_{10}), 3.81 – 3.62 (m, 5H, C_1 , H_{8b} , H_3 , $H_{7a,b}$), 3.62 – 3.53 (m, 2H, H_4 , H_5), 3.53 – 3.20 (m, 9H, $H_{8a,b}$, H_{10a} , H_{11}), 3.16 – 3.05 (m, 2H, C_4 , C_5), 2.87 (s, 3H, H_{10b}), 1.99 – 1.80 (m, 4H, C_9 , C_6), 1.80 – 1.21 (m, 8H, H_{12} , H_{13}), 0.94 (t, 6H, H_{14} , $J_{14,15} = 7.3$ Hz)
¹³C NMR (100 MHz, CD_3OD): $\delta = 177.9$, 177.5 (C_1); 101.4 (C_1); 77.5 (C_3); 76.6 (C_3); 73.8 (C_2); 73.5 (C_4); 73.3 (C_2); 70.6 (C_1); 69.6 (C_1); 64.0 (C_6); 53.0 (C_6); 52.0 (C_{11}); 39.1 (C_{10}); 36.8 (C_6); 36.6 (C_{10}); 35.2 (C_{10}); 32.8, 31.3 (C_2); 30.2 (C_{11}); 29.0 (C_{10}); 21.8 (C_{11}); 15.1 (C_{14})
MS (HRMS) calculated for $[C_{38}H_{47}N_3O_3Na]^+$: 596.3271; found: 596.3281

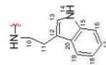
12.31



Amine purchased from Sigma-Aldrich. Coupling reaction performed in DMF in 4 hours.
 Yield: 47% over two steps
¹H NMR (400 MHz, CD_3OD): $\delta = 7.43$ (d, 4H, H_{12} , $J_{12,13} = 7.0$ Hz), 7.40 – 7.32 (m, 6H, H_{13} , H_{14}), 4.92 (d, 1H, H_1 , $J_{1,2} = 1.4$ Hz), 4.85 – 4.76 (m, 4H, H_{10}), 4.04 – 4.00 (m, 1H, C_1), 3.90 – 3.83 (m, 2H, H_2 , H_3 , H_{10}), 3.79 – 3.49 (m, 7H, H_2 , H_4 , H_5 , H_{8a} , $H_{7a,b}$, C_1), 3.45 – 3.35 (m, 2H, $H_{8a,b}$), 2.80 – 2.64 (m, 2H, C_6 , C_5), 2.01 – 1.74 (m, 4H, C_9 , C_6)
¹³C NMR (100 MHz, CD_3OD): $\delta = 176.8$, 176.6 (C_1); 138.01 (C_{11}); 131.5, 131.4 (C_{12}); 130.6, 130.4 (C_{13} , C_{14}); 101.3 (C_1); 80.1 (C_{10}); 77.2 (C_1); 76.5 (C_1); 73.5 (C_{11}); 73.4 (C_2); 73.1 (C_2); 69.7 (C_1); 64.0 (C_6); 52.9 (C_1); 39.4, 39.4 (C_{10} , C_6); 30.5, 29.7 (C_{10} , C_6)
MS (MALDI-TOF) calculated for $[C_{38}H_{49}N_3O_3Na]^+$: 668.254; found: 668.554

23

12.32



Amine purchased from Sigma-Aldrich. Coupling reaction performed in DMF in two days, with the addition of 3.2 mol equiv. Et_3N .
 Yield: 55% over two steps
¹H NMR (400 MHz, CD_3OD): $\delta = 8.61$ – 8.56 (m, 2H, H_{19}), 7.55 (d, 2H, H_{19} , $J_{19,18} = 7.7$ Hz), 7.32 (d, 2H, H_{16} , $J_{16,17} = 7.7$ Hz), 7.12 – 7.04 (m, 4H, H_{13} , H_{12}), 7.02 – 6.95 (m, 2H, H_{18}), 4.93 (d, 1H, H_1 , $J_{1,2} = 1.4$ Hz), 4.02 – 3.98 (m, 1H, C_1), 3.92 – 3.84 (m, 2H, H_2 , H_{10}), 3.78 – 3.53 (m, C_1 , H_{8b} , $H_{7a,b}$, H_3 , H_4 , H_5), 3.53 – 3.35 (m, 6H, H_{11} , $H_{8a,b}$), 2.92 (oct, 4H, H_{10} , $J = 7.4$ Hz), 2.84 – 2.70 (m, 2H, C_9 , C_6), 1.97 – 1.74 (m, 4H, C_9 , C_6)
¹³C NMR (100 MHz, CD_3OD): $\delta = 177.9$ (C_1); 139.1 (C_{13}); 139.8 (C_{10}); 124.5, 123.2 (C_{16} , C_{17}); 120.5 (C_{19}); 120.2 (C_{19}); 113.8 (C_{12}); 113.2 (C_{10}); 101.3 (C_1); 77.5 (C_1); 76.5 (C_1); 73.6 (C_2); 73.4 (C_2); 70.1 (C_1); 69.8 (C_1); 64.0 (C_6); 53.0 (C_6); 43.1, 42.9 (C_{10} , C_{13}); 42.2 (C_{11}); 30.4, 29.4 (C_{10} , C_6); 27.2 (C_{10})
MS (MALDI-TOF) calculated for $[C_{38}H_{49}N_3O_3Na]^+$: 742.317; found: 742.661

12.33



Amine purchased from Key Organics. Coupling reaction performed in DMF in two days, with the addition of 3.2 mol equiv. Et_3N .
 Yield: 25% over two steps
¹H NMR (400 MHz, CD_3OD): $\delta = 4.94$ (d, 1H, H_1 , $J_{1,2} = 1.4$ Hz), 4.13 – 4.04 (m, 2H, H_{10}), 4.04 – 4.00 (m, 1H, C_1), 3.93 – 3.88 (m, 2H, H_2 , H_{10}), 3.80 – 3.64 (m, 5H, H_{8b} , $H_{7a,b}$, C_1), 3.62 – 3.54 (m, 2H, H_4 , H_5), 3.48 – 3.33 (m, 2H, $H_{8a,b}$), 2.91 – 2.76 (m, 4H, C_9 , C_6 , H_{10a}), 2.72 – 2.65 (m, 8H, H_{15} , H_{16}), 2.61 – 2.52 (m, 1H, H_{10a}), 2.48 – 2.39 (m, 1H, H_{10a}), 2.00 – 1.90 (m, 4H, C_9 , C_6)
¹³C NMR (100 MHz, CD_3OD): $\delta = 177.6$, 177.5 (C_1); 145.5, 145.5 (C_{13}); 115.6 (C_{16}); 101.2 (C_1); 77.4 (C_1); 76.4 (C_1); 73.5 (C_1); 73.3 (C_2); 73.2 (C_2); 70.1 (C_1); 69.6 (C_6); 63.9 (C_6); 52.9 (C_6); 47.9, 47.8 (C_{10}); 42.7, 42.7 (C_{10} , C_6); 41.3 (C_{15} , C_{16}); 30.5, 30.3 (C_{10} , C_6); 26.2 (C_{11})
MS (MALDI-TOF) calculated for $[C_{38}H_{49}N_3O_3Na]^+$: 760.252; found: 760.608

24

12.34



Amine purchased from Alinda Chemicals. Coupling reaction performed in DMF in two days, with the addition of 3.2 mol equiv. Et₃N.

Yield: 40% over two steps (incomplete amide coupling – monoamide methyl ester)

¹H NMR (400 MHz, CD₃OD): δ = 4.97 (s, 1H, H₁), 4.08 – 4.01 (m, 1H, C₁), 3.92 – 3.84 (m, 2H, H₂, H₆), 3.84 – 3.49 (m, 12H, C₁, H₃, H₄, H₅, H_{6a}, H_{7a,b}, H₁₁, H₁₃), 3.43 – 3.35 (m, 6H, H_{6a,b}, H₁₂), 3.27 – 3.12 (m, 2H, H₁₀), 3.08 – 2.91 (m, 1H, C_{4 or 5}), 2.86 – 2.70 (m, 1H, C_{5 or 4}), 2.20 – 1.80 (m, 4H, C₃, C₆), 1.80 – 1.72 (m, 4H, H₁₃), 1.72 – 1.63 (m, 4H, H₁₄)

¹³C NMR (100 MHz, CD₃OD) (monoamide substituent in position C4 or C5): δ = 178.5 (C₁), 177.9, 177.6 (C₂), 102.2, 101.4 (C₁), 77.4, 77.0 (C₃), 76.5 (C₄), 73.5 (C₅), 73.4 (C₆), 73.3 (C₇), 73.3 (C₈), 73.3 (C₉), 70.2, 70.1 (C₁₀), 69.6 (C₁₁), 63.9 (C₁₂), 53.1 (C₁₃), 52.9 (C₁₄), 50.7 (C₁₀), 49.9 (C₁₂), 42.4, 42.2, 41.1, 40.9 (C_{6a}, C_{6b}), 36.1 (C₁₁), 31.6 (C₁₃), 30.6, 29.9, 29.8, 28.9 (C_{6a}, C_{6b}), 28.8 (C₁₄)

MS (MALDI-TOF) calculated for [C₂₁H₃₃N₂O₂SNa]⁺: 660.252; found: 660.542

12.35



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 2 hours.

Yield: 43% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.92 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.09 – 4.05 (m, 1H, C₁), 3.90 – 3.31 (m, 29H, C₁, C₃, C₅, H₂, H₃, H₄, H₅, H_{6a,b}, H_{7a,b}, H_{8a,b}, H₁₀, H₁₁), 1.97 – 1.77 (m, 4H, C₃, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 176.6, 176.3 (C₁), 101.7 (C₁), 77.3 (C₃), 76.6 (C₂), 73.4 (C₄), 73.4 (C₅), 73.3 (C₆), 70.8 (C₇), 76.6 (C₈), 68.7, 68.6 (C₉), 63.9 (C₁₀), 53.0 (C₁₁), 44.3, 48.3 (C₁₀), 38.4, 38.3 (C_{6a}, C_{6b}), 30.8, 29.5 (C_{3a}, C_{3c})

MS (HRMS) calculated for [C₂₁H₃₃N₂O₁Na]⁺: 596.2544; found: 596.2544

12.36



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 2 hours.

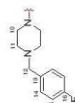
Yield: 80% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.95 (d, 1H, H₁, J_{1,2} = 1.5 Hz), 4.09 – 4.05 (m, 1H, C₁), 3.86 – 3.81 (m, 2H, H₂, H_{6a}), 3.78 – 3.31 (m, 17H, C₁, H₃, H₄, H₅, H_{6b}, H_{7a,b}, H_{8a,b}, H₁₀), 3.28 – 3.10 (m, 2H, C₃, C₆), 2.02 – 1.78 (m, 12H, C₃, C₅, H₁₁)

¹³C NMR (100 MHz, CD₃OD): δ = 176.5, 176.3 (C₁), 101.3 (C₁), 77.4 (C₃), 76.5 (C₂), 73.5 (C₄), 73.4 (C₅), 73.1 (C₆), 70.6 (C₇), 69.6 (C₈), 63.9 (C₉), 53.0 (C₁₀), 48.7, 48.5, 47.8 (C₁₀), 40.7, 40.6 (C_{6a}, C_{6b}), 29.9, 28.7 (C₃), 27.8, 27.7 (C₃, C₆)

MS (MALDI-TOF) calculated for [C₂₄H₃₉N₂O₁Na]⁺: 564.264; found: 564.546

12.37



Amine purchased from Sigma-Aldrich. Coupling reaction performed in DMF, overnight.

Yield: 54% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.44 – 7.33 (H₁₃), 7.12 – 7.04 (H₁₄), 4.92 (s, 1H, H₁), 4.11 – 4.06 (m, 1H, C₁), 3.89 – 3.33 (m, 25H, C₁, C₃, C₅, H₂, H₃, H₄, H₅, H_{6a,b}, H_{7a,b}, H_{8a,b}, H₁₀, H₁₂), 2.69 – 2.40 (m, 8H, H₁₁), 1.98 – 1.77 (m, 4H, C₃, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 176.3, 176.1 (C₁), 164.7 (C₁₀, J_{C,F} = 245.9 Hz), 134.7, 134.5 (C₁₃), 133.3 (C₁₄), 117.1, 116.9 (C₁₁), 101.6 (C₁), 77.3 (C₃), 76.6 (C₂), 73.4 (C₄), 73.3 (C₅), 70.7 (C₆), 69.6 (C₇), 64.0 (C₈), 63.4 (C₉), 54.8, 54.7, 54.5, 54.4 (C₁₀), 47.3, 43.4 (C₁₀), 38.4, 38.4 (C_{6a}, C_{6b}), 30.9, 29.6 (C_{3a}, C_{3c})

MS (MALDI-TOF) calculated for [C₂₈H₃₅F₂N₂O₂Na]⁺: 810.361; found: 810.815

12.38



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 2 hours.

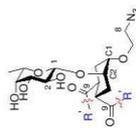
Yield: 75% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.95 (d, 1H, H₁, J_{1,2} = 1.4 Hz), 4.11 – 4.07 (m, 1H, C₁), 3.89 – 3.51 (m, 17H, C₁, H₂, H₃, H₄, H₅, H_{6a,b}, H_{7a,b}, H₁₀), 3.46 – 3.32 (m, 4H, C₃, C₅, H_{8a,b}), 2.94 – 2.59 (m, 8H, H₁₁), 2.54 (d, 6H, H₁₂), J = 8.8 Hz), 2.05 – 1.77 (m, 4H, C₃, C₆)

¹³C NMR (100 MHz, CD₃OD): δ = 176.5, 176.2 (C₁), 101.7 (C₁), 77.3 (C₃), 76.7 (C₂), 73.5 (C₄), 73.4 (C₅), 73.3 (C₆), 69.6 (C₇), 64.0 (C₈), 56.1, 55.8 (C₉), 53.0 (C₁₀), 46.4, 46.2 (C_{10a,b}), 46.0, 45.8 (C₁₂), 42.5, 42.4 (C_{10a,b}), 38.5, 38.5 (C_{6a}, C_{6b}), 30.8, 29.6 (C_{3a}, C_{3c})

MS (MALDI-TOF) calculated for [C₂₆H₄₃N₂O₂Na]⁺: 622.317; found: 622.625

Ligands 15.1-11



The numbering used for the NMR spectra is shown in the pictures. The unconventional numbering of the cyclohexane ring is used in analogy to sugar numbering.

15.1



Yield: 77% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.34 – 4.27 (m, 1H, H₁), 4.00 – 3.93 (m, 1H, C₂), 3.80 – 3.71 (m, 2H, C₃, H_{2a}), 3.69 – 3.55 (m, 9H, H_{7b}, H_{7c}, H_{7d}, H_{7e}, H_{7f}), 3.50 – 3.42 (m, 2H, H₂, H₃), 3.35 (t, 2H, H_{8a,b}, J_{8a,b} = 4.9 Hz), 3.07 (td, 1H, C₄, J = 12.5 Hz, J = 3.7 Hz), 2.93 (td, 1H, C₅, J = 12.5 Hz, J = 3.7 Hz), 2.17 – 1.06 (m, 2H, C_{6a}, C_{6b}), 1.91 (td, 1H, C_{6c}, J = 13.6 Hz, J = 2.4 Hz), 1.82 (td, 1H, C_{6d}, J = 13.6 Hz, J = 2.4 Hz), 1.25 (d, 1H, H₆, J_{6,5} = 6.4 Hz)

¹³C NMR (100 MHz, CD₃OD): δ = 178.2, 178.0 (C₃), 104.8 (C₁), 77.2 (C₂), 75.9 (C₄), 75.1 (C₅), 73.4 (C₆), 73.1 (C₇), 72.8 (C₈), 70.2 (C₉); 53.2, 53.1 (C₁₀), 52.9 (C₁₁), 41.0 (C₁₂, C₁₃), 29.5, 29.2 (C₁₄, C₁₅); 17.7 (C₁₆)

MS (MALDI-TOF) calculated for [C₁₆H₁₉N₂O₂Na]⁺: 470.175; found: 470.134

15.2



Amine prepared according to ref. 11c. Coupling reaction performed in THF overnight.

Yield: 68% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.25 (dd, 4H, H_{13a,b}, J = 8.0 Hz, J = 2.9 Hz), 7.19 (dd, 4H, H_{13a,b}, J = 8.0 Hz, J = 2.9 Hz), 4.53 (d, 4H, H_{15a,b}, J = 3.2 Hz), 4.31 (s, 1H, H₁), 4.32 – 4.26 (m, 4H, H_{10a,b}), 4.02 – 3.98 (m, 1H, C₁), 3.79 – 3.72 (m, 1H, H_{2a}), 3.72 – 3.69 (m, 1H, H₂), 3.65 – 3.56 (m, 3H, H_{7b}, H_{7c}), 3.52 – 3.42 (m, 2H, H₄, H₅), 3.42 – 3.30 (m, 2H, H_{8a,b}), 2.98 – 2.86 (m, 2H, C₆), 2.03 – 1.83 (m, 4H, C₃, C₄), 1.24 (d, 3H, H₆, J_{6,5} = 1.24)

¹³C NMR (100 MHz, CD₃OD): δ = 178.4, 177.8 (C₃), 142.5, 142.4 (C₁), 139.9, 139.8 (C₁₁), 129.3, 129.2, 129.0, 129.0 (C₂, C₁₃), 103.7 (C₁), 77.6 (C₂), 74.4 (C₄), 73.8 (C₅), 73.3 (C₆), 72.8 (C₇), 69.9 (C₈), 65.8 (C₉); 52.9 (C₁₀), 44.6, 44.5 (C₁₀), 42.3 (C₁₀, C₁₀), 30.0, 29.9 (C₁₁, C₁₂), 17.6 (C₁)

MS (HRMS) calculated for [C₂₁H₂₄N₂O₂Na]⁺: 680.2908; found: 680.2918

15.3



Amine purchased from Crea-Chim. Coupling reaction performed in THF overnight, with the addition of 3.2 mol equiv. Et₃N.

Yield: 36% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.50 (q, 2H, H₁₃, J_{13,12} = 1.8 Hz), 6.31 (d, 2H, H₁₂), 5.05 – 4.89 (m, 4H, H_{10a}), 4.46 – 4.35 (m, 4H, H_{10b}), 4.34 (d, 1H, H₁, J_{1,2} = 7.0 Hz), 4.05 – 4.00 (m, 1H, C₂), 3.83 – 3.75 (m, 1H, H_{10c}), 3.75 – 3.71 (m, 1H, C₁), 3.97 – 3.59 (m, 3H, H_{7b}, H_{7c}), 3.57 – 3.46 (m, 2H, H₂, H₃), 3.46 – 3.34 (m, 2H, H_{8a,b}), 2.98 – 2.86 (m, 2H, C₆, C₃), 2.01 – 1.81 (m, 4H, C₄, C₅), 1.27 (d, 1H, H₆, J_{6,5} = 6.4 Hz)

¹³C NMR (100 MHz, CD₃OD): δ = 178.6, 178.2 (C₃), 144.3 (C₁), 142.1 (C₁₁), 125.8 (C₁₂, J_{12,11} = 278.8 Hz); 108.9 (C₂), 103.8 (C₂), 77.5 (C₂), 75.9 (C₂), 74.4 (C₄), 73.9 (C₅), 73.3 (C₆), 72.9 (C₇), 69.9 (C₈), 52.9 (C₉), 51.8, 51.5 (C₁₀), 42.0, 41.9 (C₁₀, C₁₀), 35.4 (C₁₀), 30.0, 29.9 (C₁₁, C₁₂), 17.6 (C₁)

MS (HRMS) calculated for [C₂₁H₂₇N₂O₂Na]⁺: 764.2567; found: 764.2571

28

15.4



Amine purchased from Sigma-Aldrich. Coupling reaction performed in DMF overnight, with the addition of 3.2 mol equiv. Et₃N.

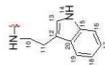
Yield: 55% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.45 – 7.29 (m, 10H, H₁₃, H₁₃, H₁₃), 4.88 – 4.73 (m, 4H, H₁₀), 4.28 (d, 1H, H₁, J_{1,2} = 7.0 Hz), 4.01 – 3.94 (m, 1H, C₂), 3.79 – 3.66 (m, 2H, H_{7b}, C₁), 3.66 – 3.56 (m, 3H, H_{7b}, H_{7c}), 3.53 – 3.43 (m, 2H, H₂, H₃), 3.42 – 3.32 (m, 2H, H_{8a,b}), 2.78 – 2.63 (m, 2H, C₆, C₃), 2.00 – 1.74 (m, 4H, C₄, C₅, C₆), 1.26 (d, 1H, H₆, J_{6,5} = 6.4 Hz)

¹³C NMR (100 MHz, CD₃OD): δ = 174.7, 174.4 (C₃), 137.8, 137.7 (C₁₁), 131.4 (C₁₁), 130.5, 130.4, 130.3 (C₂, C₁₂), 103.6 (C₁), 80.1, 80.0 (C₁₀), 77.4 (C₂), 75.9 (C₄), 74.1 (C₂), 73.9 (C₅), 73.7 (C₅), 72.9 (C₅), 69.9 (C₇), 52.9 (C₈), 39.0, 38.9 (C₆, C₆), 29.8, 29.7 (C₃, C₃), 17.6 (C₁)

MS (MALDI-TOF) calculated for [C₁₆H₁₉N₂O₂Na]⁺: 652.259; found: 652.160

15.5



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF overnight.

Yield: 61% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 8.82 (s, 2H, H₁₀), 7.55 (t, 2H, H₁₀, J_{10,9} = 7.4 Hz), 7.32 (d, 2H, H₁₀, J_{10,7} = 8.0 Hz), 7.07 (t, 4H, H₁₀, H₁₀, J = 7.4 Hz), 6.98 (t, 2H, H₁₀, J = 7.4 Hz), 4.26 (d, 1H, H₁, J_{1,2} = 7.3 Hz), 4.00 – 3.96 (m, 1H, C₂), 3.74 – 3.34 (m, 13H, H_{7b}, C₁, H₂, H₃, H₄, H₅, H_{8a,b}, H₁₁), 3.00 – 2.84 (m, 4H, H₁₀), 2.84 – 2.74 (m, 2H, C₆, C₃), 1.96 – 1.74 (m, 4H, C₄, C₅), 1.28 (d, 1H, H₆, J_{6,5} = 6.28 Hz)

¹³C NMR (100 MHz, CD₃OD): δ = 178.4, 177.9 (C₃), 139.1 (C₁₁), 129.8, 129.7 (C₁₀), 124.6, 124.5 (C₂), 123.2 (C₁), 120.5 (C₁₂), 120.3, 120.2 (C₁₀), 114.1, 114.0 (C₁₀), 103.7 (C₁), 77.7 (C₁), 75.9 (C₄), 74.4 (C₄), 73.9 (C₅), 73.4 (C₅), 73.9 (C₅), 69.9 (C₇), 52.9 (C₈), 42.6, 42.5 (C₆, C₆), 42.2, 42.1 (C₁), 29.9, 29.8 (C₃, C₃), 27.3, 27.2 (C₁), 17.7 (C₁)

MS (MALDI-TOF) calculated for [C₂₀H₂₆N₂O₂Na]⁺: 726.332; found: 726.177

15.6



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF overnight.

Yield: 67% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 4.32 (d, 1H, H₁, J_{1,2} = 7.0 Hz), 4.10 – 4.05 (m, 1H, C₂), 3.79 – 3.33 (m, 2H, H₁, H₂, H₃, H₄, H₅, H_{8a,b}, H₁₀, H₁₁, C₃), 2.00 – 1.73 (m, 4H, C₄, C₅), 1.27 (d, 1H, H₆, J_{6,5} = 6.5 Hz)

28

¹³C NMR (100 MHz, CD₃OD): δ = 177.0, 176.5 (C₃); 103.9 (C₁); 77.2 (C₁₁); 76.0 (C₆); 73.9 (C₂); 73.0 (C₇); 72.8 (C₅); 70.6 (C₄); 69.0, 68.7, 68.7, 68.6 (C₁₁); 53.0 (C₉); 48.4, 44.4 (C₁₀); 38.2, 38.1 (C₈, C₁₀); 30.2, 29.2 (C₃, C₆); 17.7 (C₁)
MS (MALDI-TOF) calculated for [C₂₄H₂₈N₂O₁₀Na]⁺: 580.259; found: 580.182

15.7



Amine purchased from Sigma-Aldrich. Coupling reaction performed in DMF overnight.

Yield: 57% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 8.35 (br s, 2H, H₁₂), 8.29 (t, 2H, H₁₃, J = 4.0 Hz), 7.65 – 7.58 (m, 2H, H₁₄), 7.24 (dt, 2H, H₁₅, J = 7.8 Hz, J = 4.7 Hz); 4.34 – 4.18 (m, 5H, H₁, H₁₀); 3.97 – 3.90 (m, 1H, C₂); 3.75 – 3.62 (m, 2H, H_{7a}, C₁); 3.59 – 3.49 (m, 3H, H_{7b}, H₆, H₁); 3.47 – 3.36 (m, 2H, H₇, H₁); 3.36 – 3.21 (m, 2H, H_{8a,b}); 2.93 – 2.79 (m, 2H, C₆, C₃); 1.94 – 1.74 (m, 4H, C₅, C₆); 1.22 – 1.13 (d, 1H, H₆, J_{6,5} = 6.4 Hz)
¹³C NMR (100 MHz, CD₃OD): δ = 178.8, 178.3 (C₂); 150.2, 149.6 (C₂, C₁₃); 138.3 (C₁₄); 137.7 (C₁); 126.1 (C₉); 103.9 (C₃); 77.7 (C₁); 76.0 (C₄); 74.6 (C₂); 73.9 (C₅); 73.4 (C₁); 72.9 (C₃); 70.1 (C₁); 53.0 (C₆); 42.4, 42.3 (C₈, C₅); 42.3 (C₁₀); 30.2 (C₃, C₆); 17.7 (C₁)
MS (HRMS) calculated for [C₂₈H₃₇N₂O₁₀Na]⁺: 622.2601; found: 622.2617

15.8



Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 1 hour.

Yield: 75% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 4.32 (d, 1H, H₁, J_{1,2} = 7.0 Hz), 4.09 – 4.04 (m, 1H, C₂); 3.81 – 3.37 (m, 11H, H₂, H₃, H₄, H₅, H_{6a,b}, H_{7a,b}, C₅, C₆, C₃); 2.61 – 2.52 (m, 6H, H₁₀); 1.98 – 1.71 (m, 4H, C₇, C₈); 1.26 (d, 1H, H₆, J_{6,5} = 6.4 Hz); 0.72 – 0.61 (m, 4H, H_{11a}); 0.47 – 0.38 (m, 4H, H_{11b})
¹³C NMR (100 MHz, CD₃OD): δ = 179.1, 178.6 (C₂); 102.9 (C₁); 76.7 (C₁₁); 75.0 (C₆); 73.5 (C₂); 73.0 (C₃); 72.4 (C₇); 71.9 (C₅); 69.0 (C₇); 52.0 (C₈); 41.2, 41.1 (C_{6a}, C_{6b}); 28.9, 28.8 (C₃, C₈); 23.2, 23.1 (C₉); 16.8 (C₂); 6.6, 6.5, 6.4 (C₁₁)
MS (MALDI-TOF) calculated for [C₂₇H₃₅N₂O₈Na]⁺: 520.238; found: 520.177

15.9



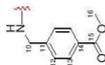
Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 6 hours.

Yield: 68% over two steps
¹H NMR (400 MHz, CD₃OD): δ = 4.32 (d, 1H, H₁, J_{1,2} = 7.0 Hz), 4.09–4.04 (m, 1H, C₂); 3.81 – 3.37 (m, 11H, H₂, H₃, H₄, H₅, H_{6a,b}, H_{7a,b}, C₅, C₆, C₃); 3.60 – 2.27 (m, 8H, H₁₀, H₁₁); 1.98 – 1.71 (m, 4H, H₃, H₆); 1.26 (d, 1H, H₆, J_{6,5} = 6.4 Hz)

29

¹³C NMR (100 MHz, CD₃OD): δ = 176.8, 176.3 (C₃); 103.9 (C₁); 77.7 (C₁₁); 76.0 (C₆); 73.8 (C₂); 73.7 (C₅); 73.0 (C₇); 72.8 (C₅); 70.6 (C₄); 57.1, 56.9, 56.5, 56.4 (C₁₁); 53.0 (C₉); 50.7 (C_{2a}); 47.3, 47.2 (C_{11a}); 46.8 (C_{10a}); 43.4 (C_{11b}); 38.3, 38.2 (C_{6a}, C₆); 30.2, 29.2 (C₃, C₆); 17.7 (C₁)
MS (MALDI-TOF) calculated for [C₂₈H₃₆N₂O₁₀Na]⁺: 606.332; found: 606.235

15.10

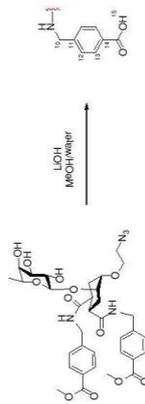


Amine purchased from Sigma-Aldrich. Coupling reaction performed in THF in 6 hours.

Yield: 66% over two steps

¹H NMR (400 MHz, CD₃OD): δ = 7.82 (dd, 4H, H₁₃, J_{13,12} = 8.4 Hz, J = 5.4 Hz); 7.19 (dd, 4H, H₁₂, J_{12,13} = 8.4 Hz, J = 5.4 Hz); 4.35 (d, 1H, H₁, J_{1,2} = 7.3 Hz); 4.61 – 4.24 (m, 4H, H₁₀); 4.07 – 4.02 (m, 1H, C₂); 3.86 (d, 6H, H₁₆, J = 3.4 Hz); 3.78 – 3.74 (m, 2H, H_{7a}, C₁); 3.69 – 3.58 (m, 3H, H_{7b}, H₆, H₅); 3.56 – 3.46 (m, 2H, H₂, H₃); 3.46 – 3.28 (m, 2H, H_{8a,b}); 3.05 – 2.96 (m, 2H, C₄, C₃); 2.06 – 1.89 (m, 4H, C₅, C₆); 1.27 (d, 1H, H₆, J_{6,5} = 6.5 Hz)
¹³C NMR (100 MHz, CD₃OD): δ = 176.4 (C₂); 175.9 (C₁₃); 129.3, 129.2 (C₁₁); 128.5, 128.4 (C₁₄); 126.6 (C₁); 101.5 (C₇); 75.4 (C₃); 73.6 (C₂); 72.3 (C₂); 71.6 (C₃); 71.1 (C₂); 70.6 (C₂); 76.7 (C₇); 51.1 (C₁₆); 50.6 (C₈); 42.1 (C₁₀); 40.1, 40.0 (C_{6a}, C_{6b}); 27.9, 27.8 (C_{3a}, C_{3b}); 15.4 (C₁)
MS (MALDI-TOF) calculated for [C₃₁H₄₃N₂O₁₁Na]⁺: 736.280; found: 736.158

15.11



15.10 (14.1 mg, 0.02 mmol, 1 equiv.) was dissolved in dry MeOH and 3 equiv. aqueous LiOH solution was added to the mixture. Upon completion, the reaction was concentrated *in vacuo*.

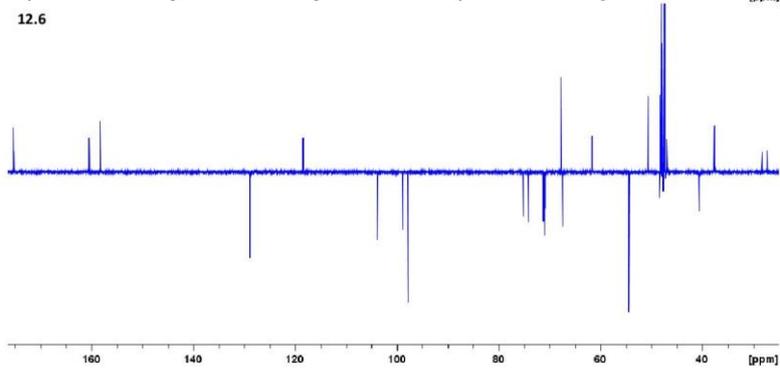
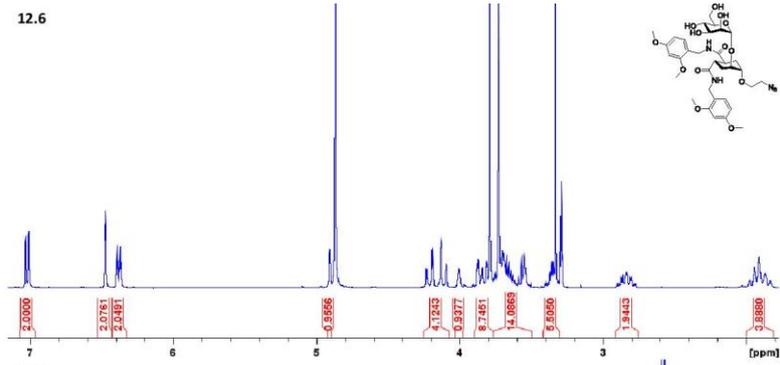
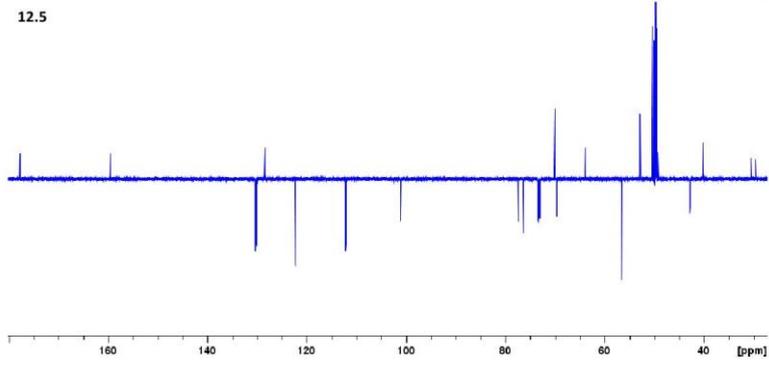
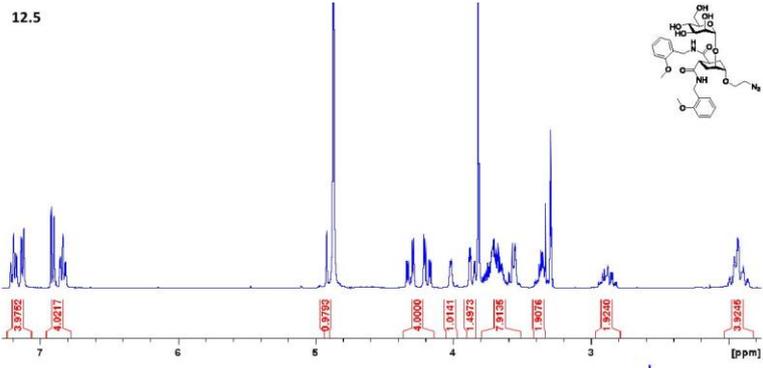
Yield: 95%

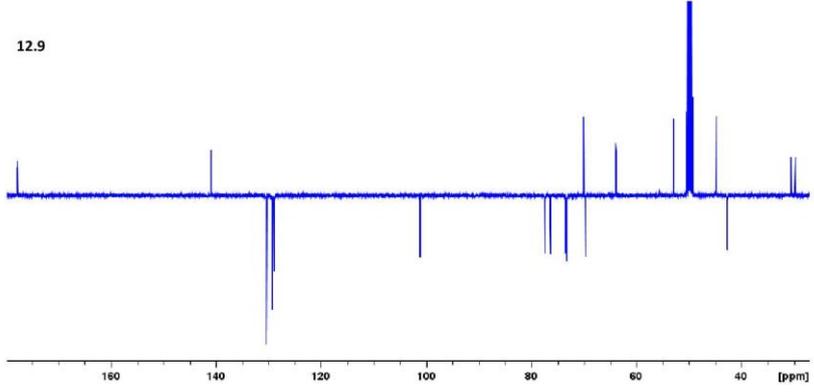
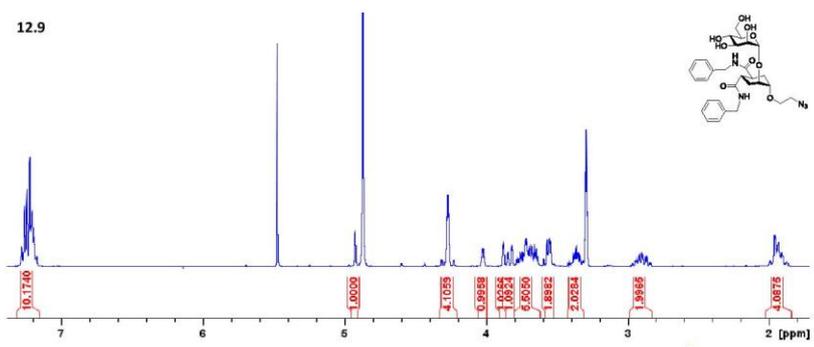
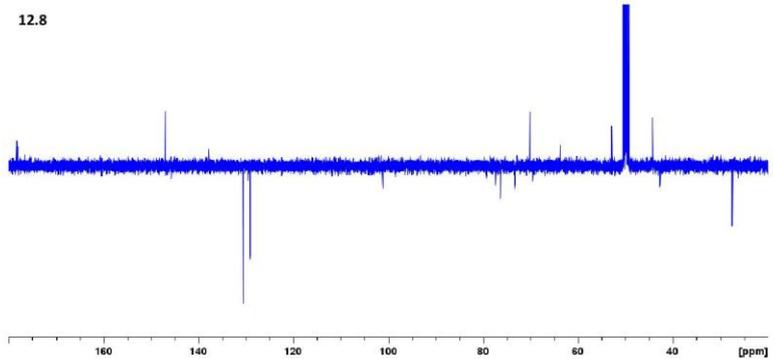
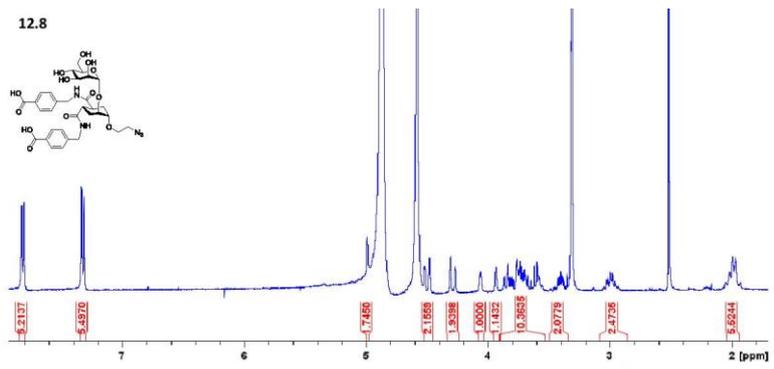
¹H NMR (400 MHz, CD₃OD): δ = 8.00 – 7.79 (br s, 4H, H₁₃); 7.27 (t, 4H, H₁₂, J = 7.3 Hz); 4.49 – 4.19 (m, 4H, H₁₀); 4.36 – 4.31 (d, 1H, H₁, J_{1,2} = 7.3 Hz); 4.07 – 3.98 (m, 1H, C₂); 3.83 – 3.70 (m, 2H, H_{7a}, C₁); 3.70 – 3.56 (m, 3H, H_{7b}, H₆, H₅); 3.56 – 3.44 (m, 2H, H₂, H₃); 3.44 – 3.31 (m, 2H, H_{8a,b}); 3.04 – 2.91 (m, 2H, C₄, C₃); 2.07 – 1.85 (m, 4H, C₅, C₆); 1.25 (d, 1H, H₆, J_{6,5} = 6.4 Hz)

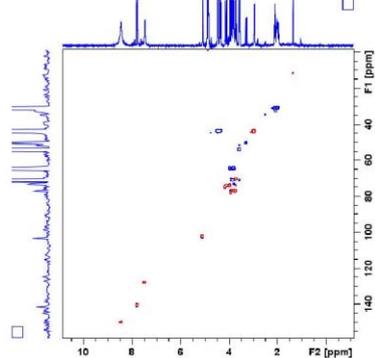
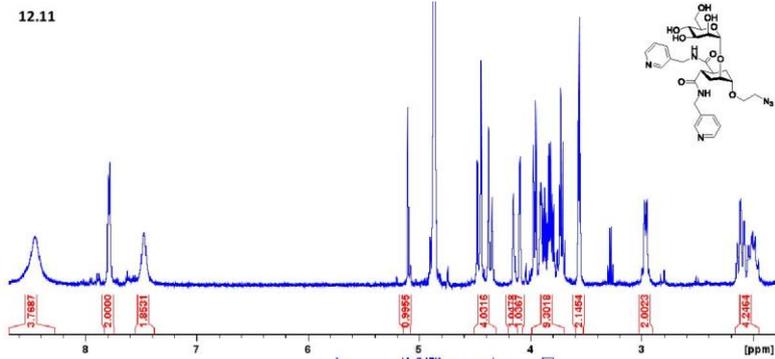
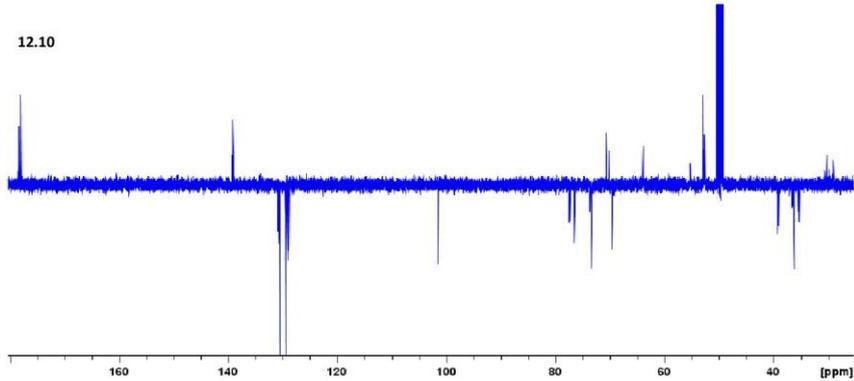
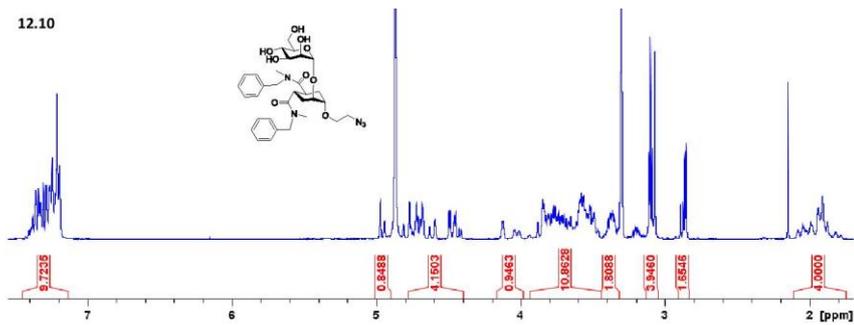
¹³C NMR (100 MHz, CD₃OD): δ = 178.6, 178.1 (C₂); 175.8 (C₁₃); 145.2, 145.1 (C₁₁); 131.9 (C₁₄); 130.6 (C₁); 128.9 (C₇); 103.8 (C₃); 77.6 (C₃); 75.9 (C₂); 74.5 (C₂); 73.9 (C₂); 72.9 (C₃); 70.0 (C₇); 52.9 (C₈); 44.6 (C₁₀); 42.3, 42.2 (C_{6a}, C_{6b}); 30.1 (C_{3a}, C_{3b}); 17.1 (C₁)

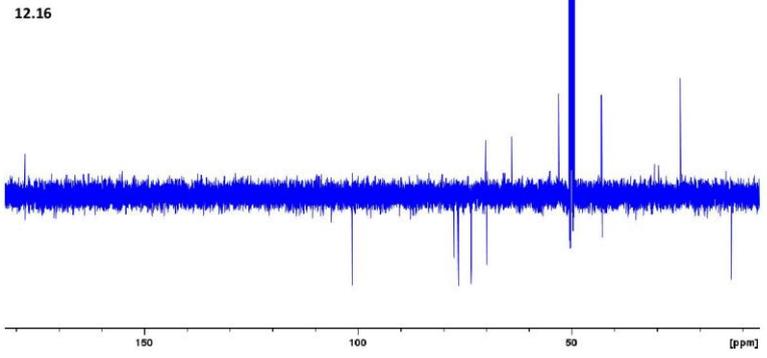
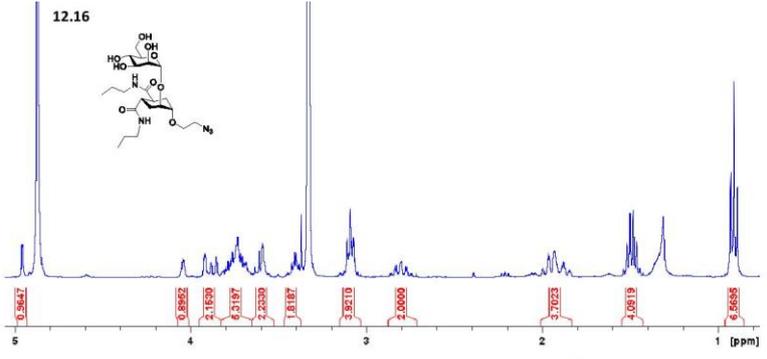
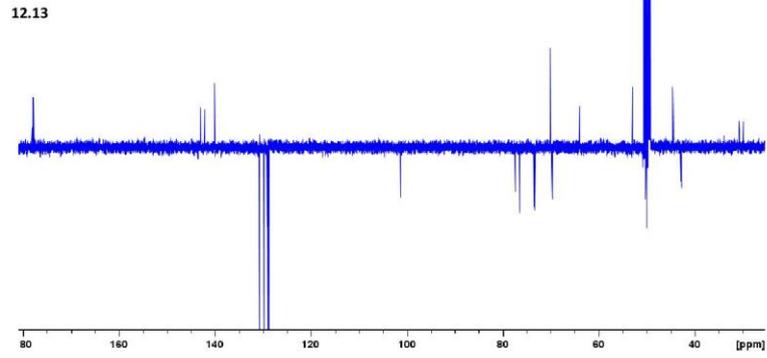
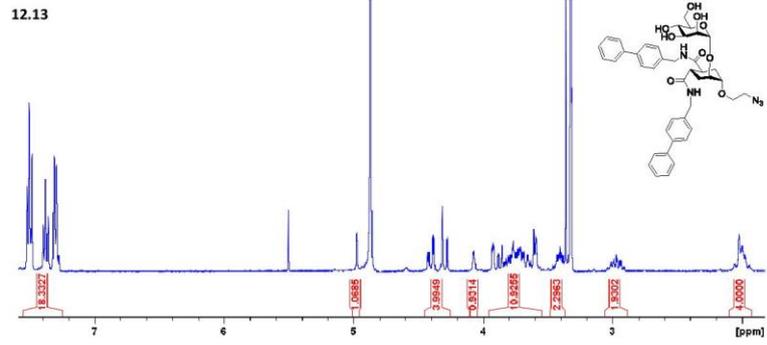
MS (MALDI-TOF) calculated for [C₃₁H₄₃N₂O₁₁Na]⁺: 708.249; found: 708.129

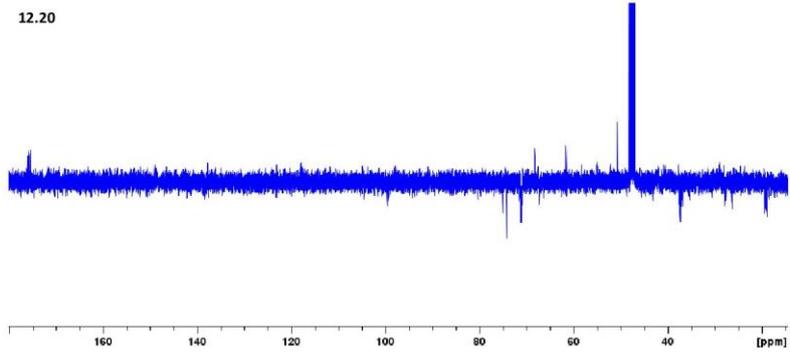
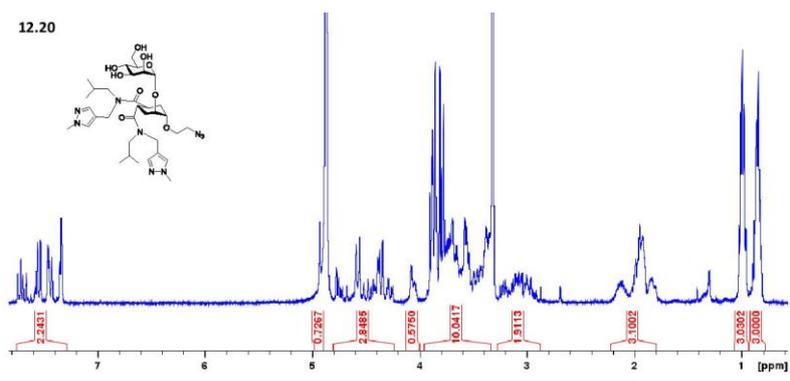
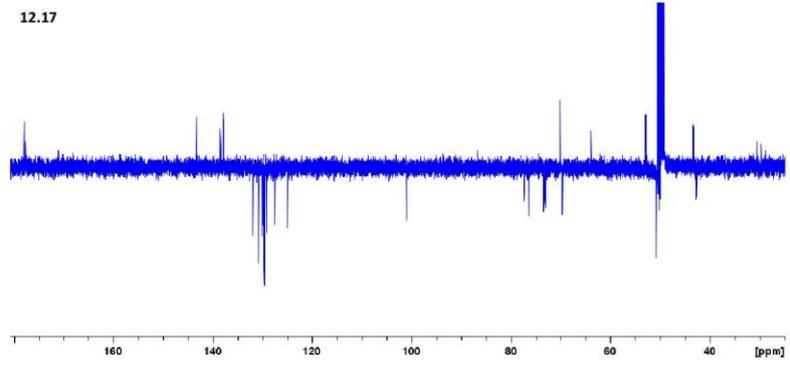
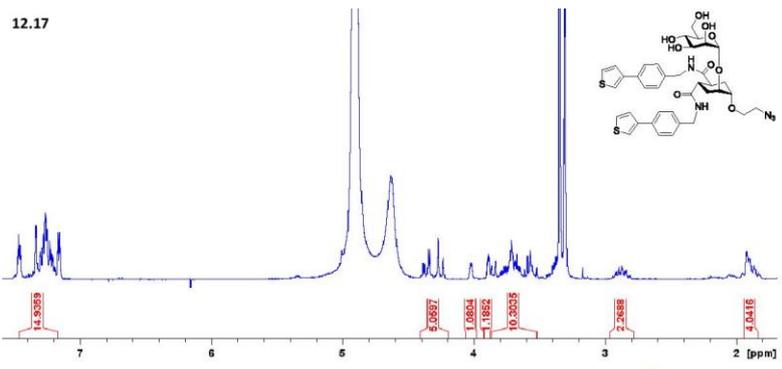
30

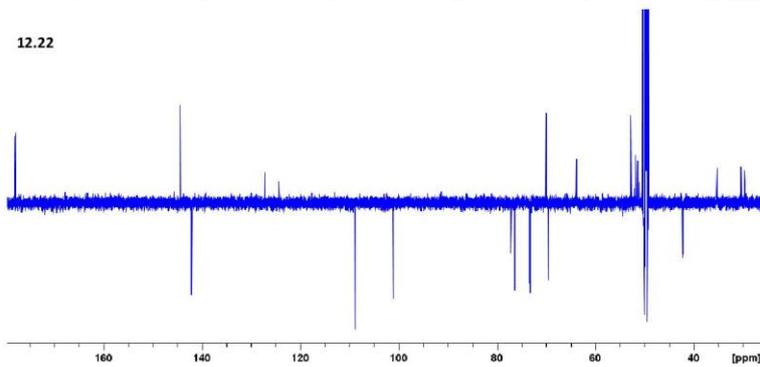
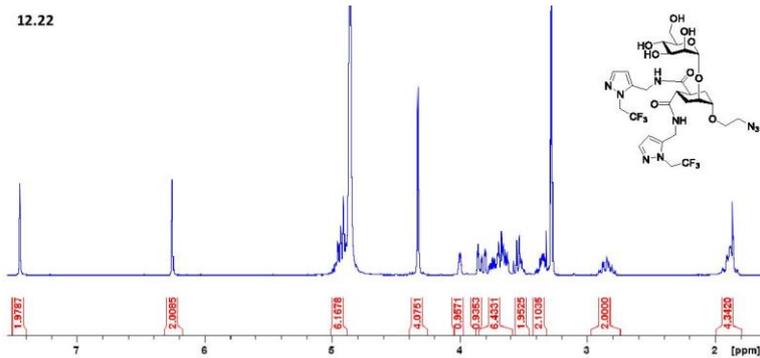
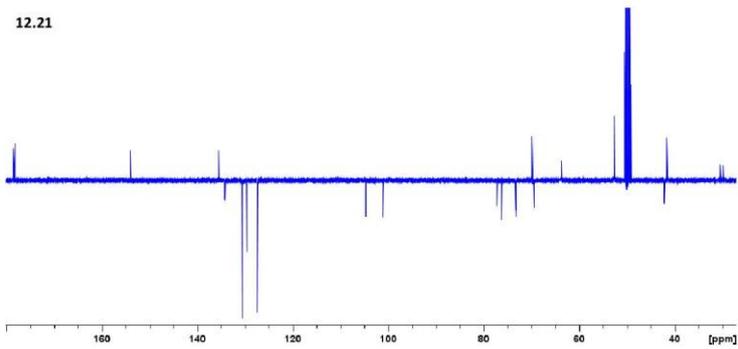
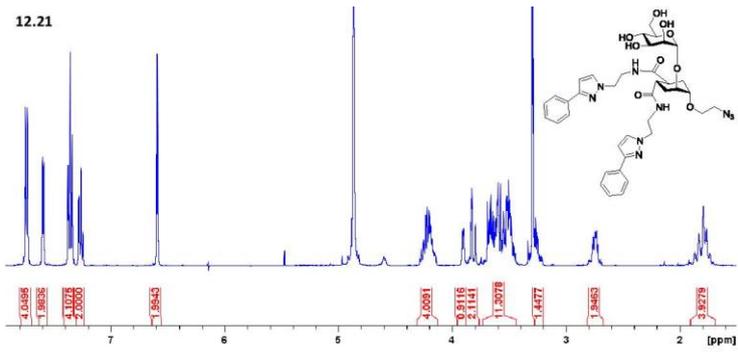




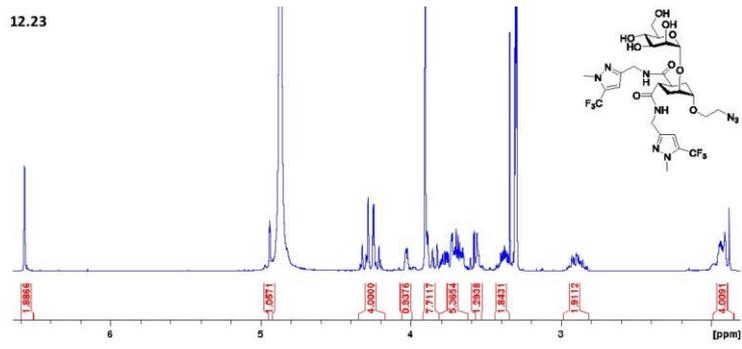




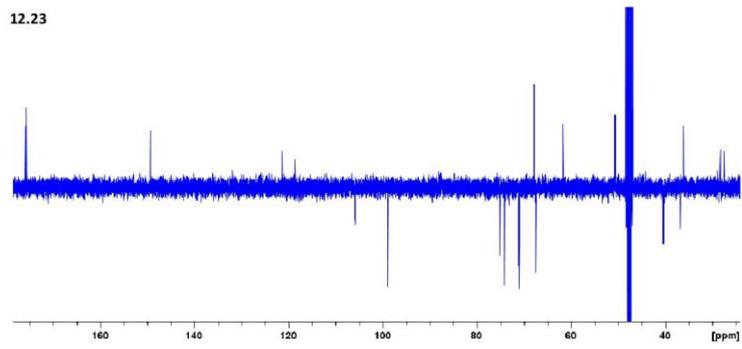




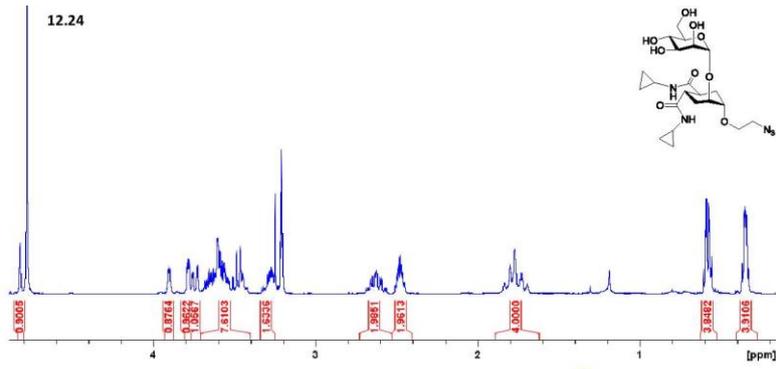
12.23



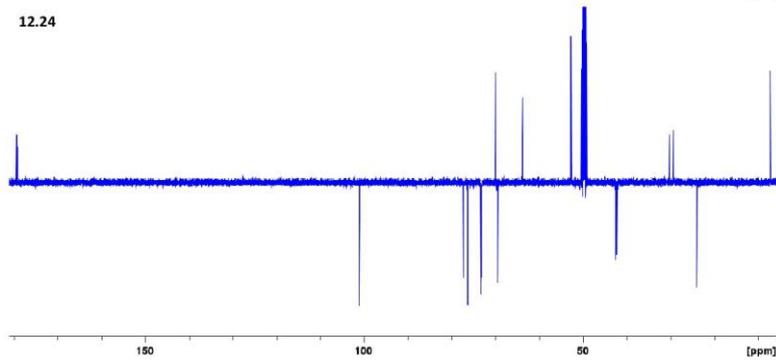
12.23

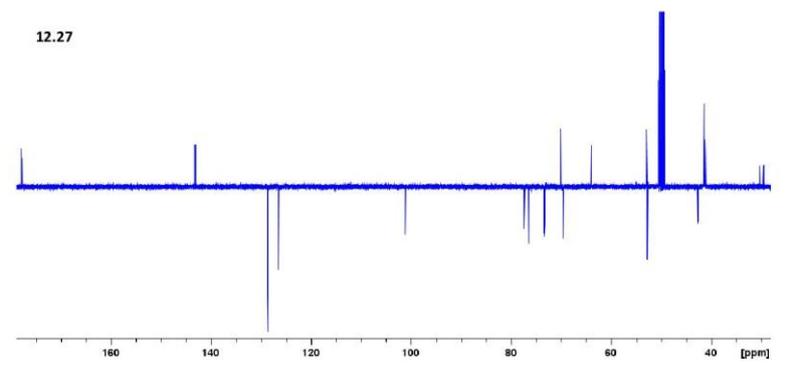
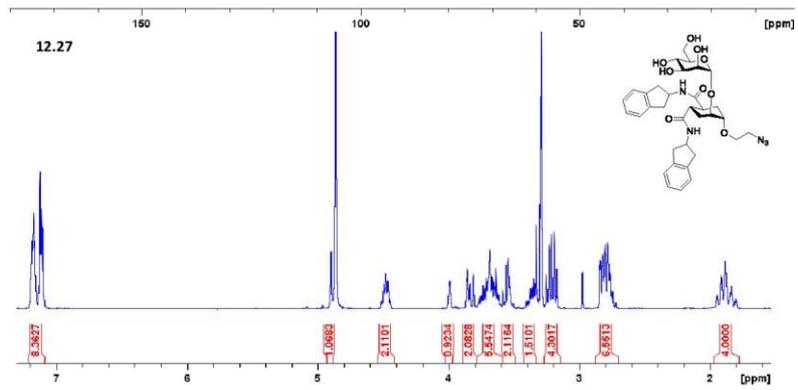
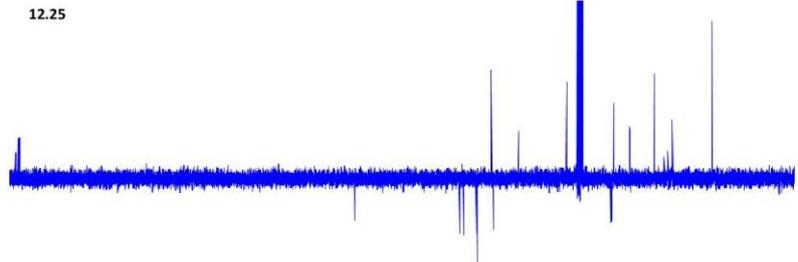
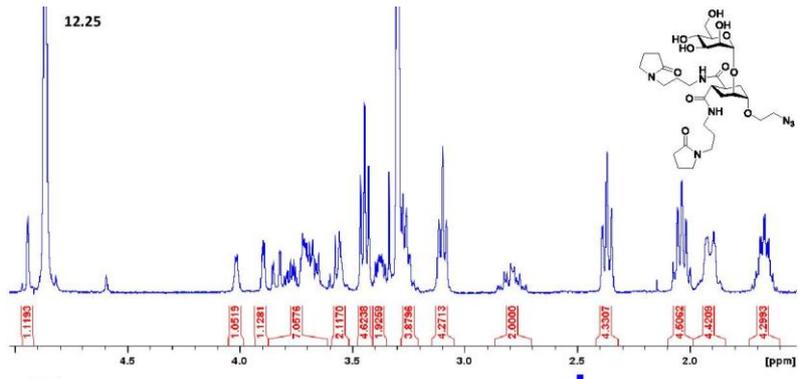


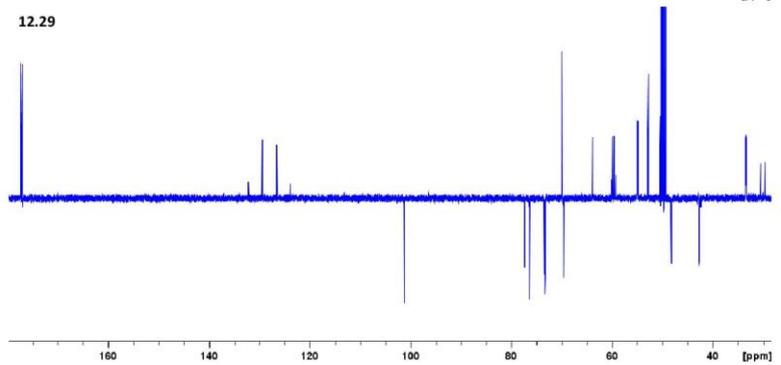
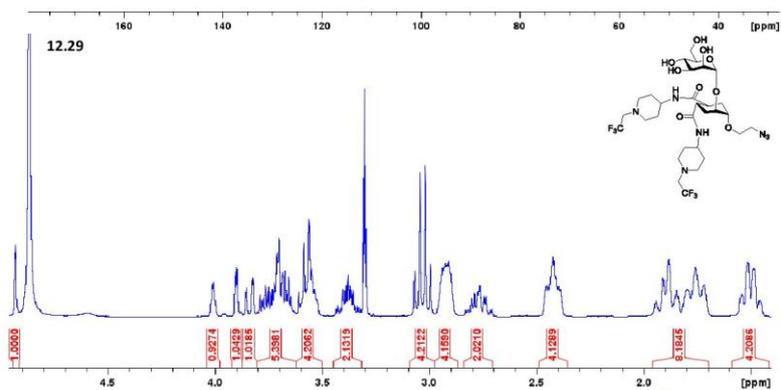
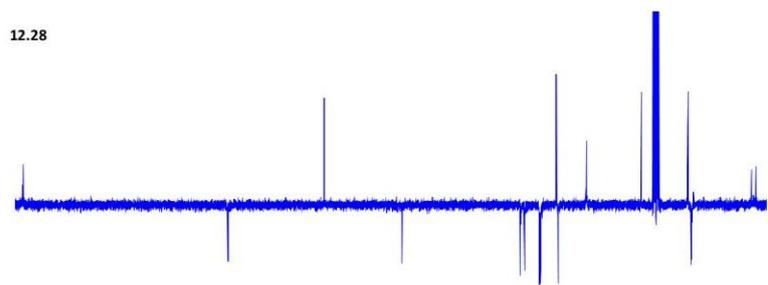
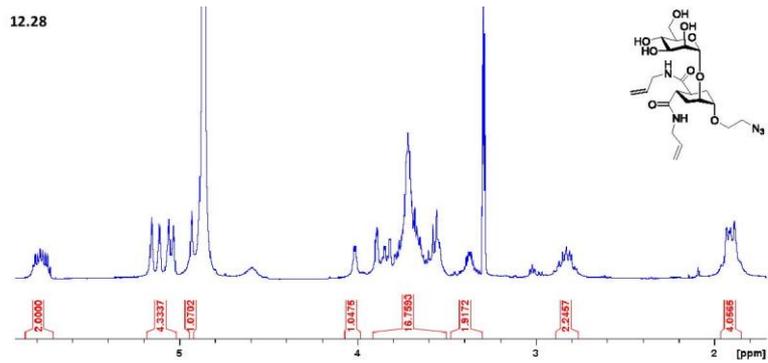
12.24

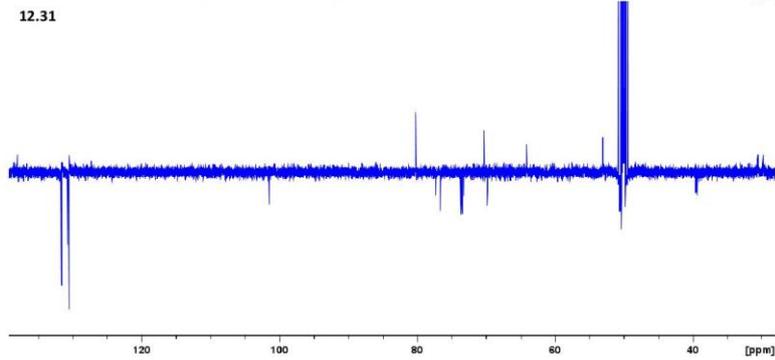
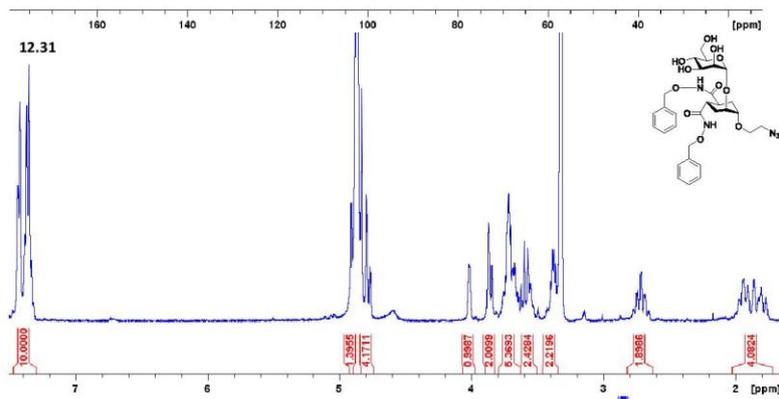
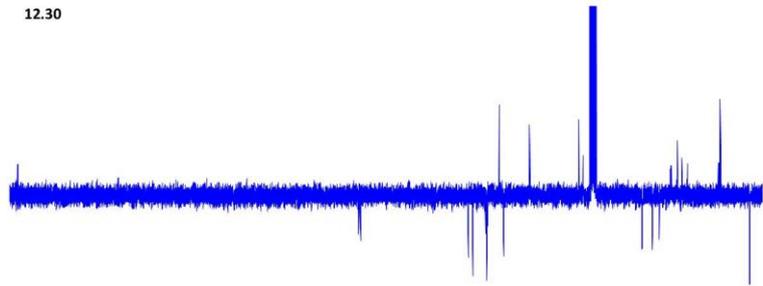
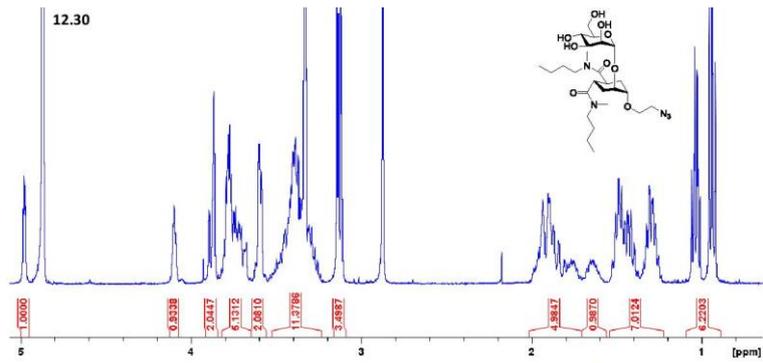


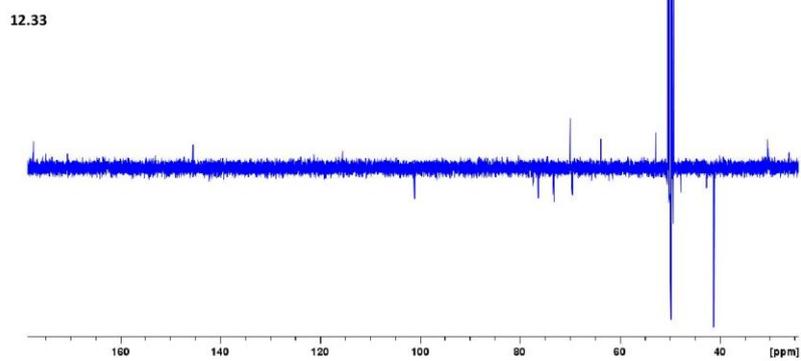
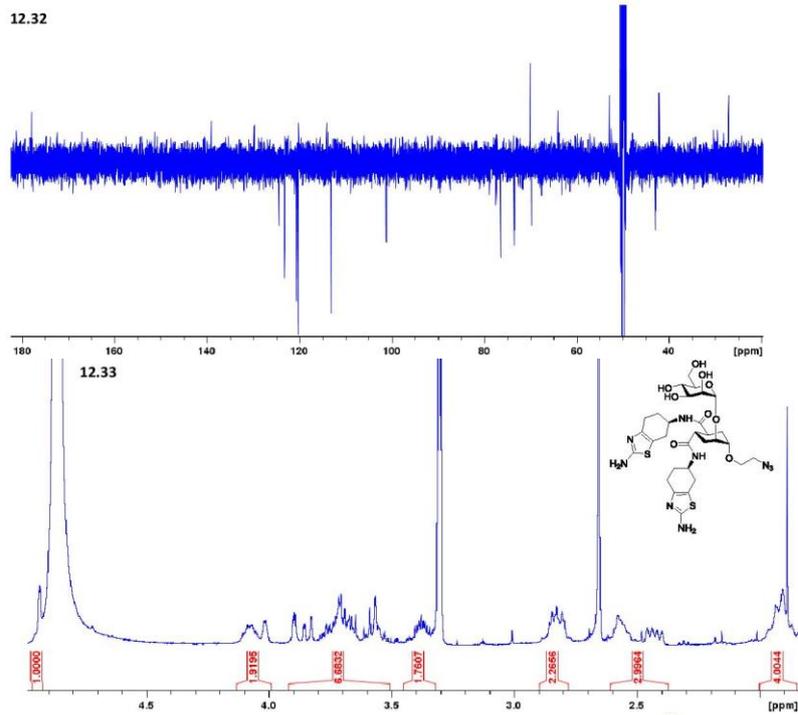
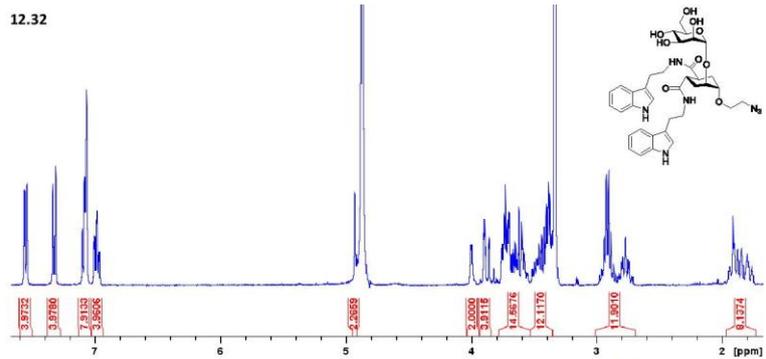
12.24

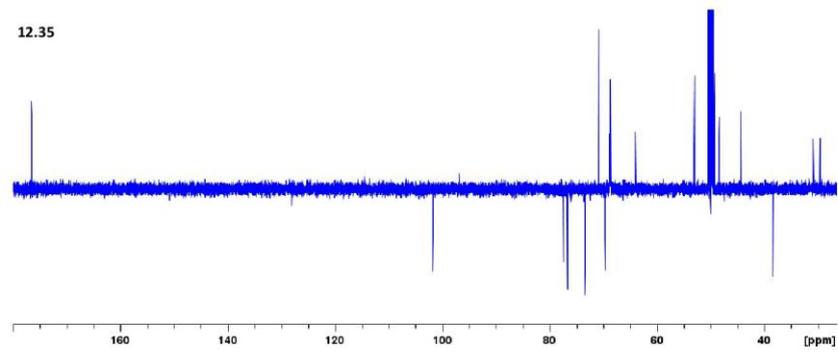
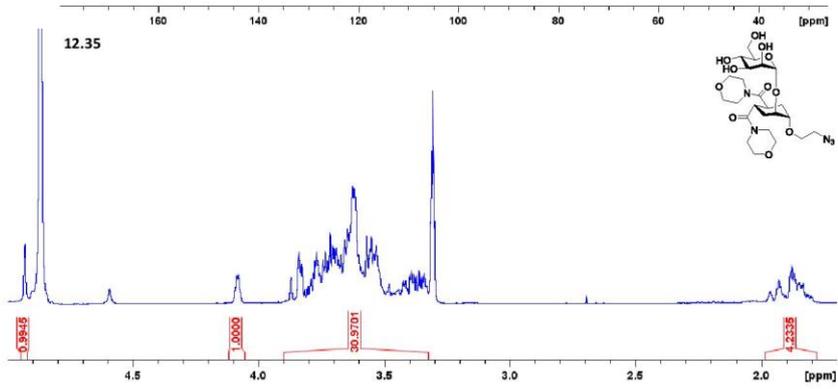
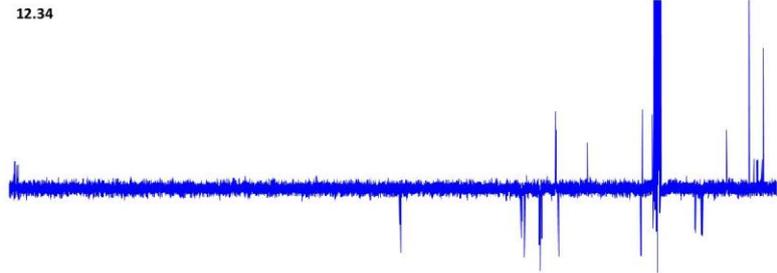
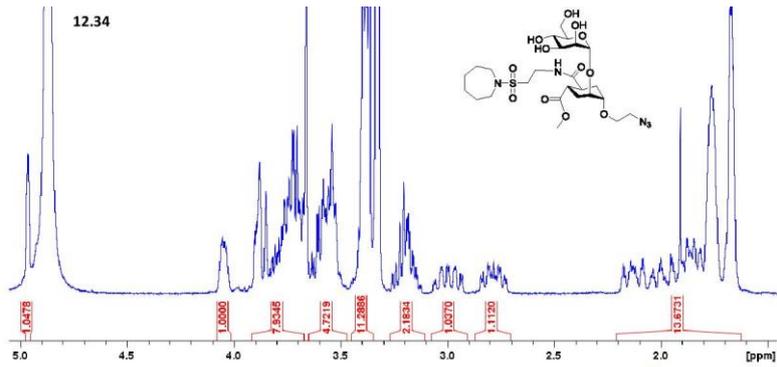


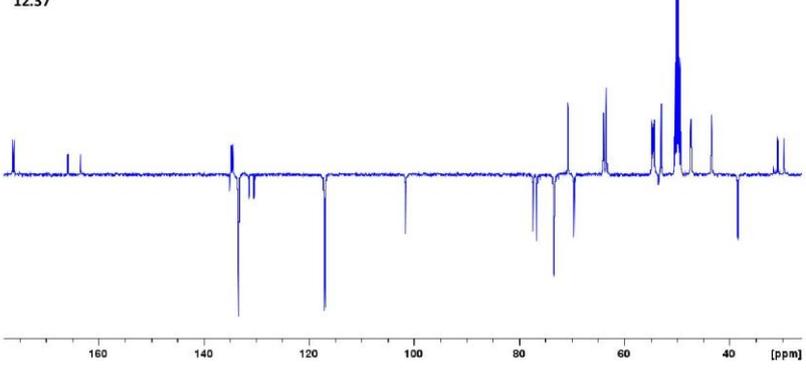
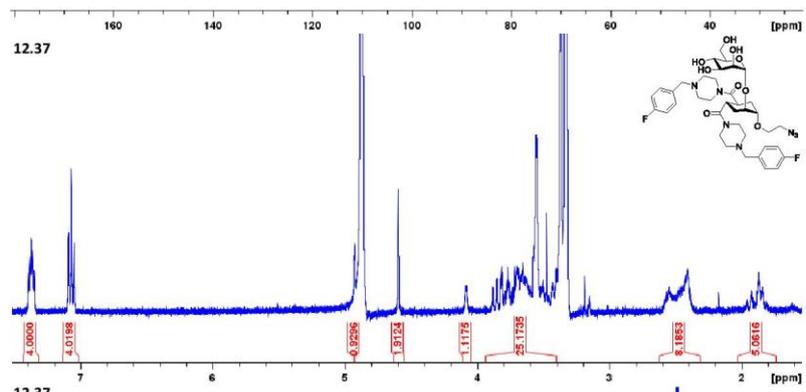
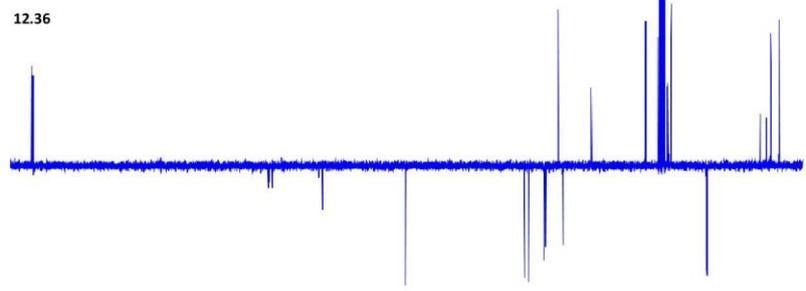
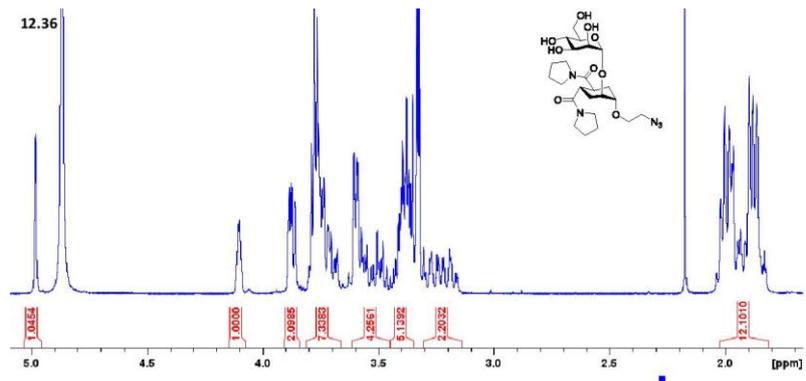


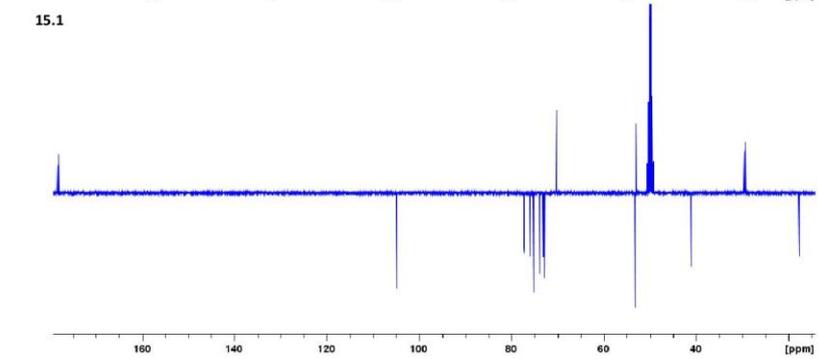
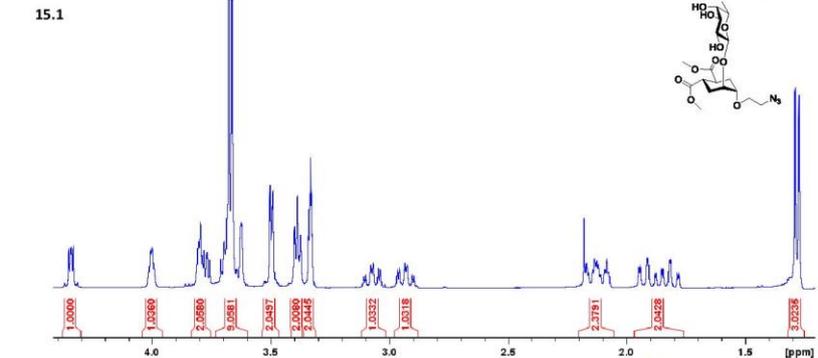
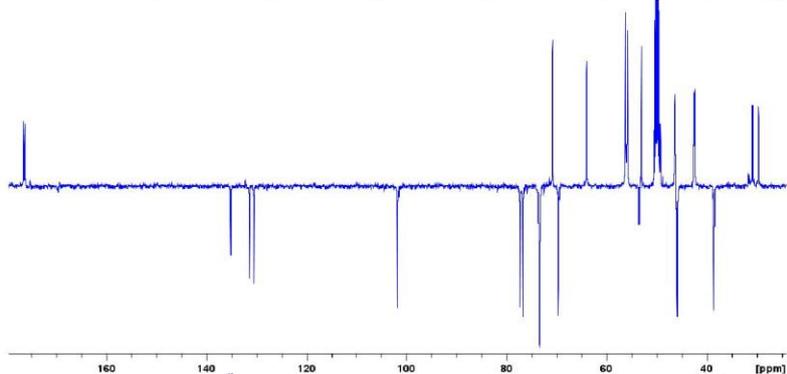
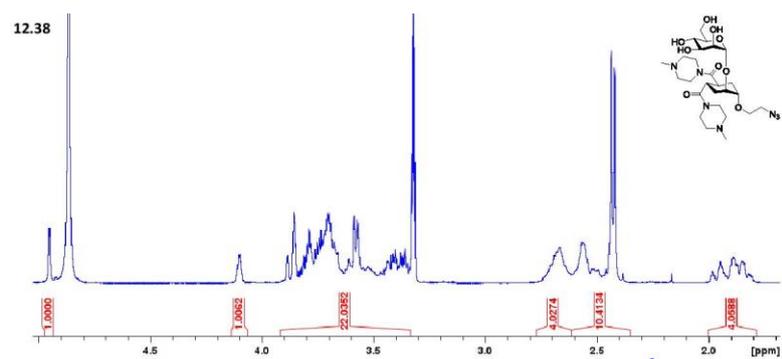


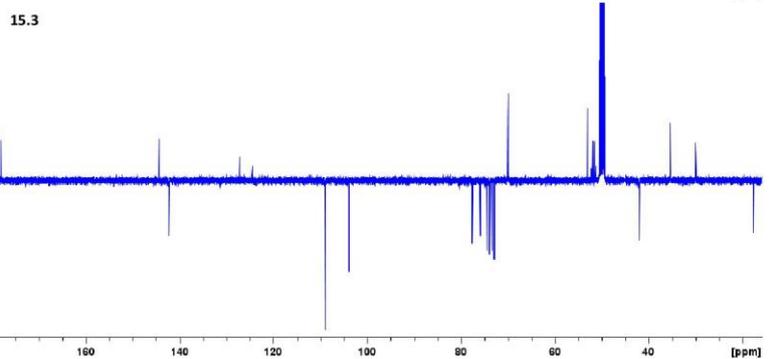
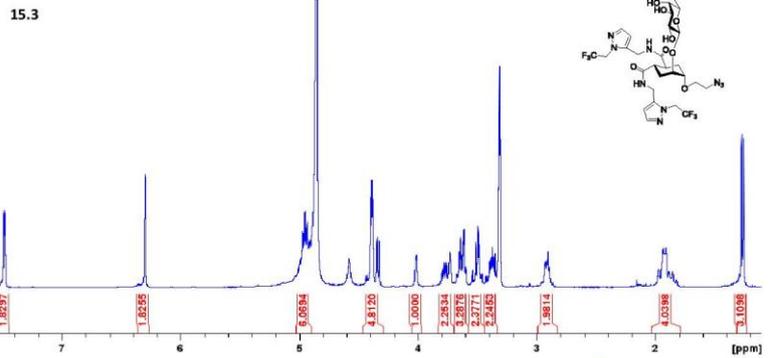
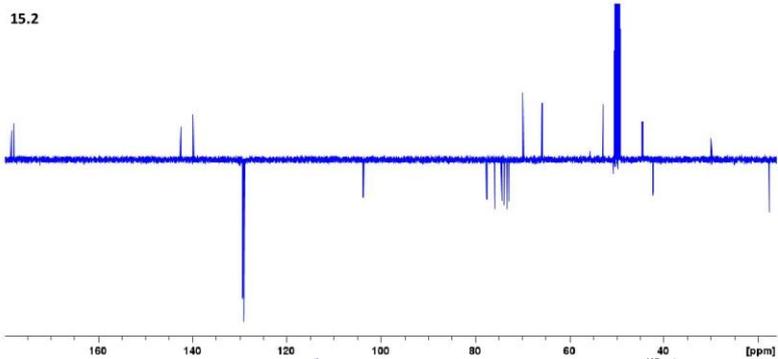
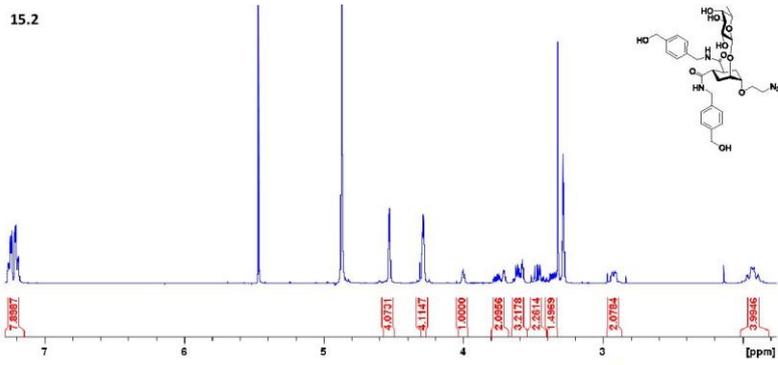


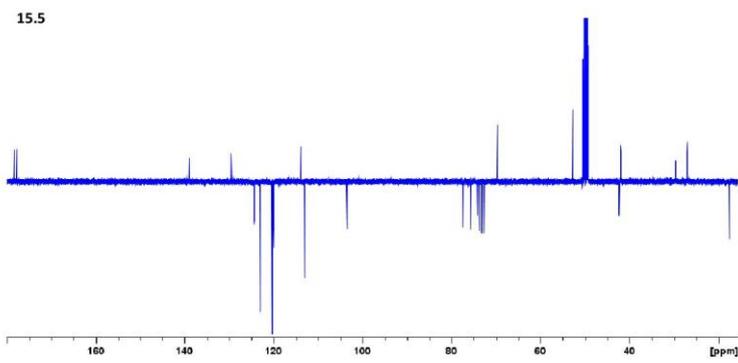
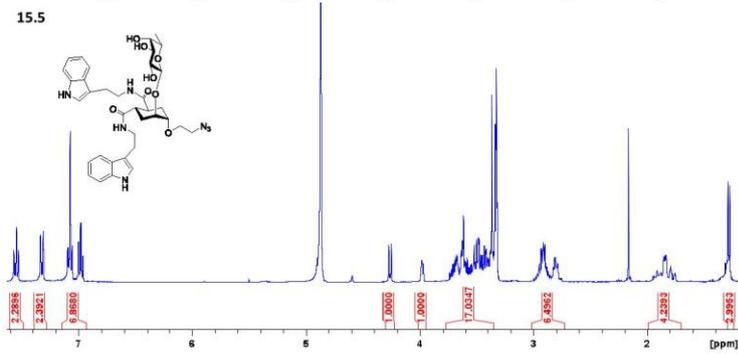
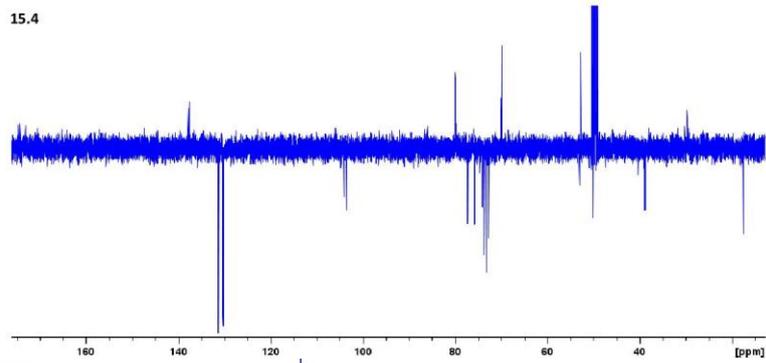
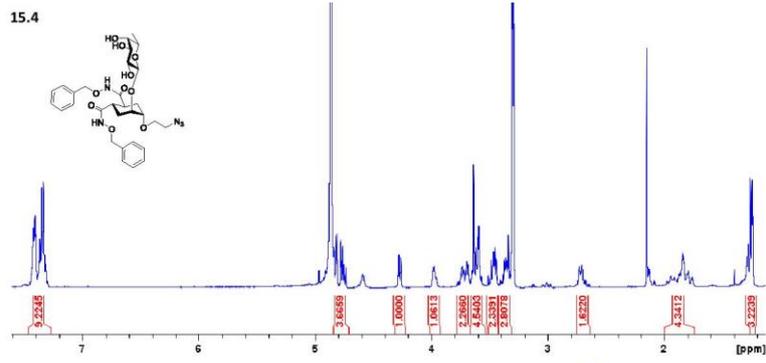


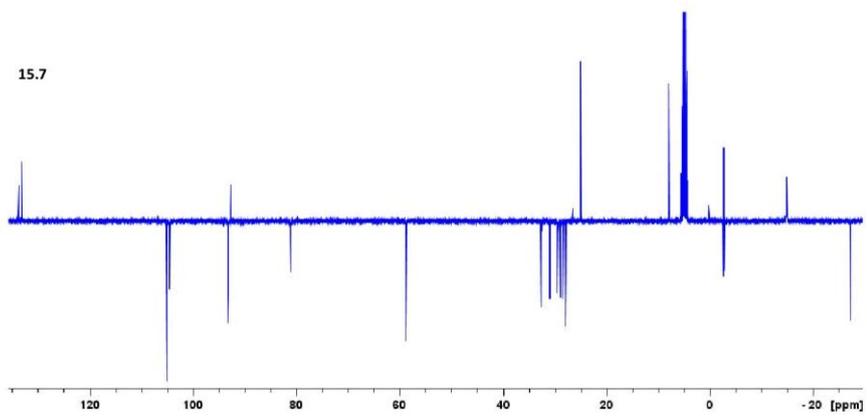
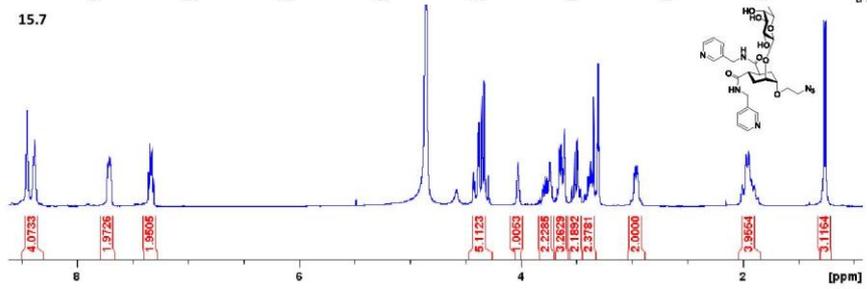
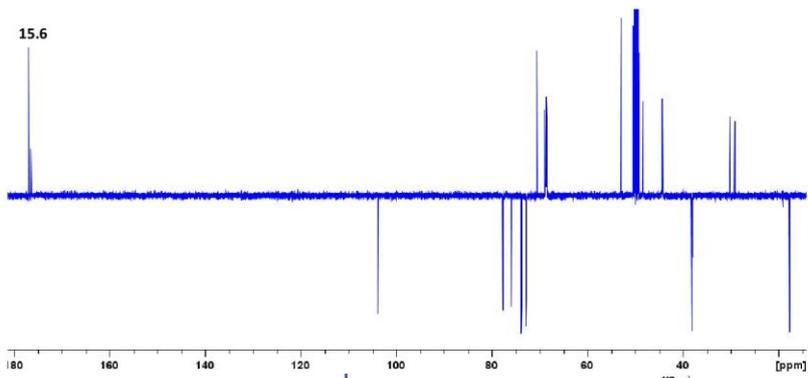
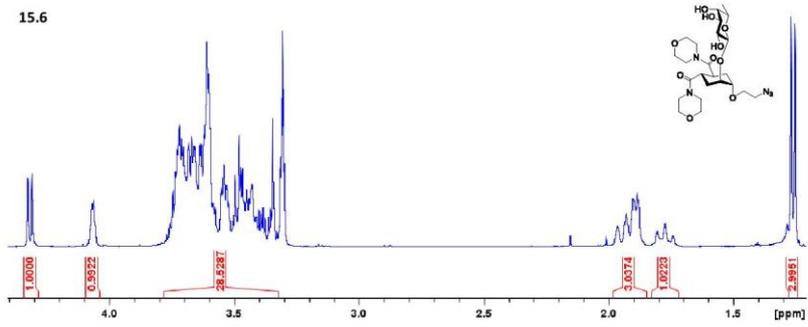


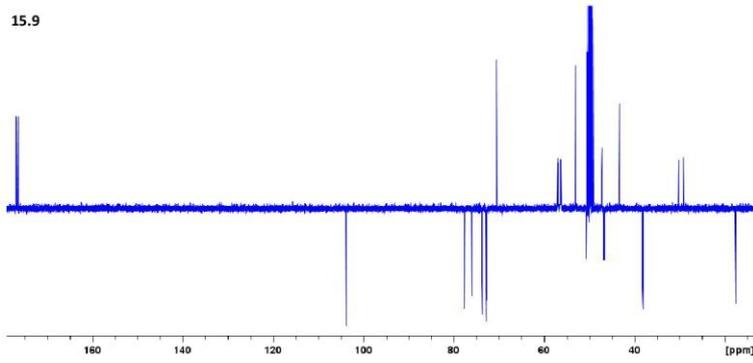
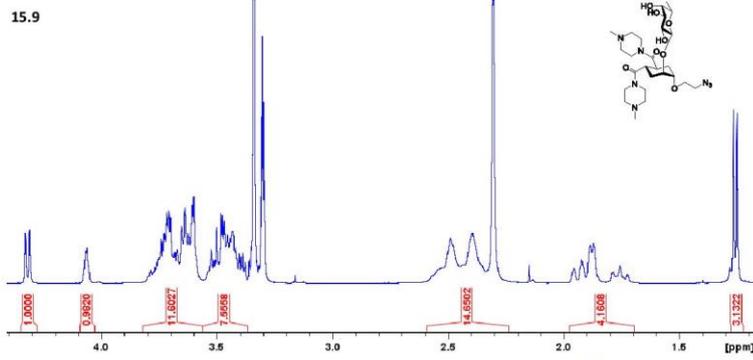
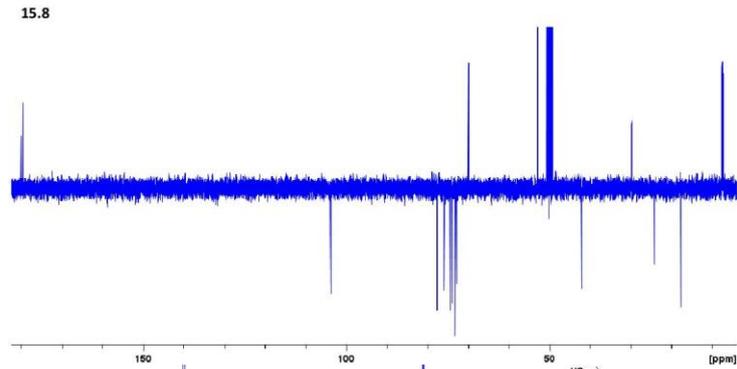
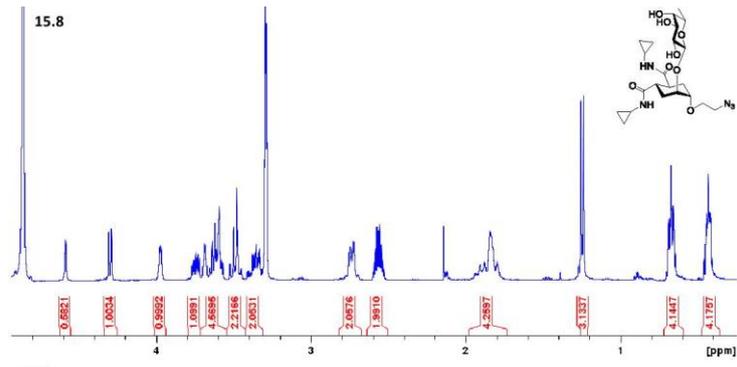


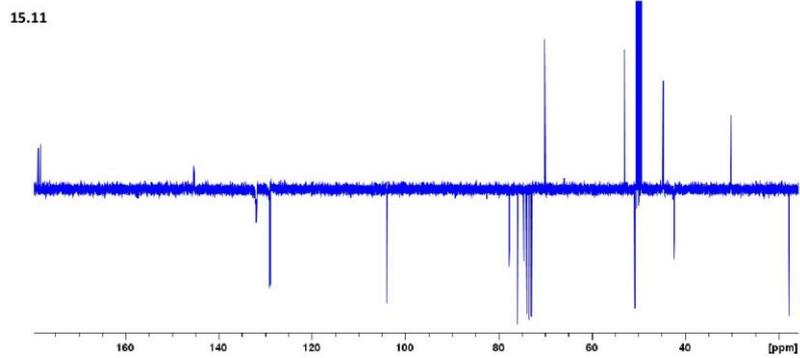
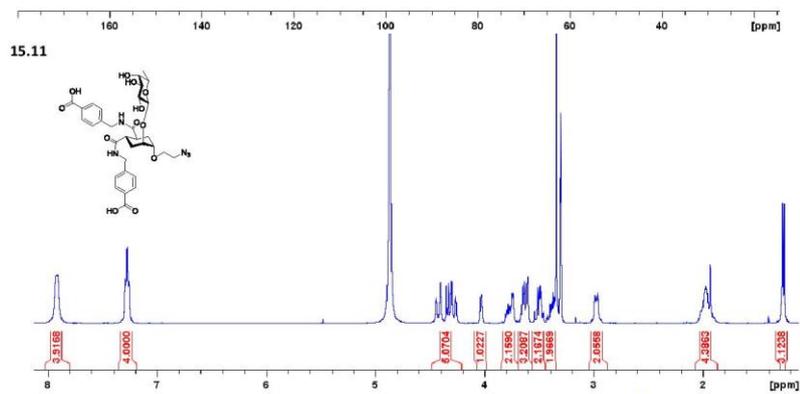
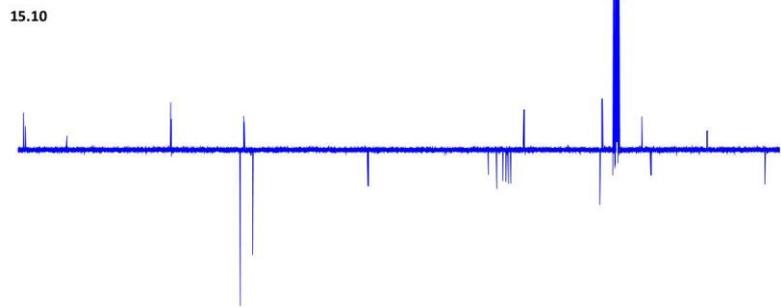
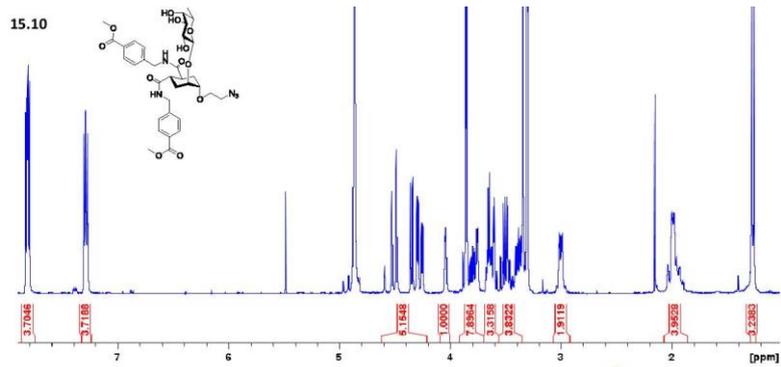


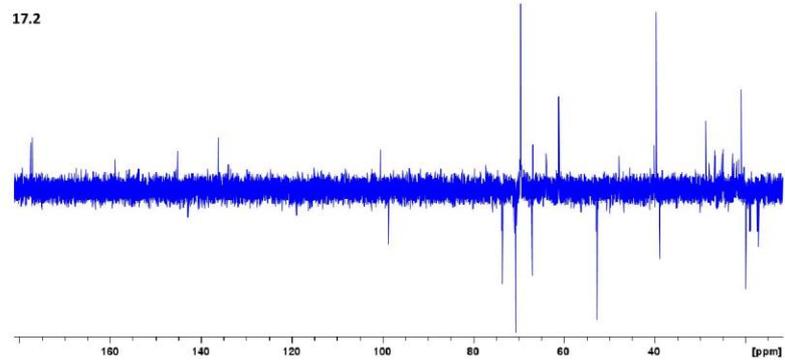
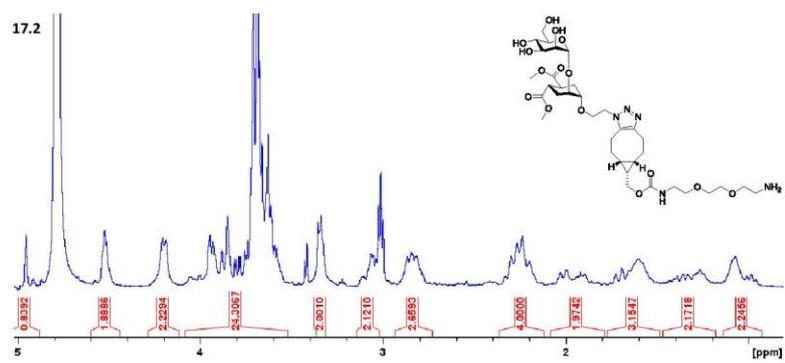










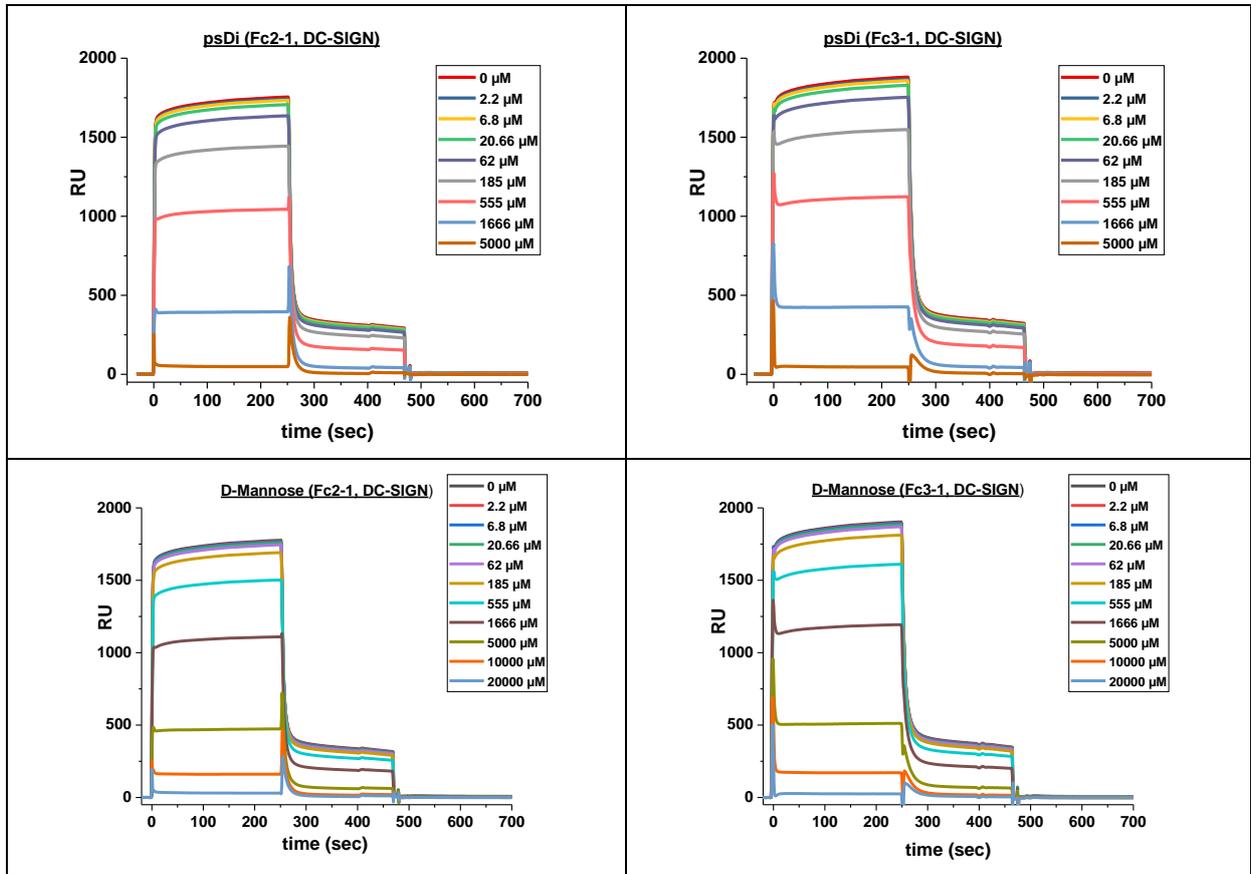


Compound names used in thesis and the corresponding initial names of the same compounds Paper 3

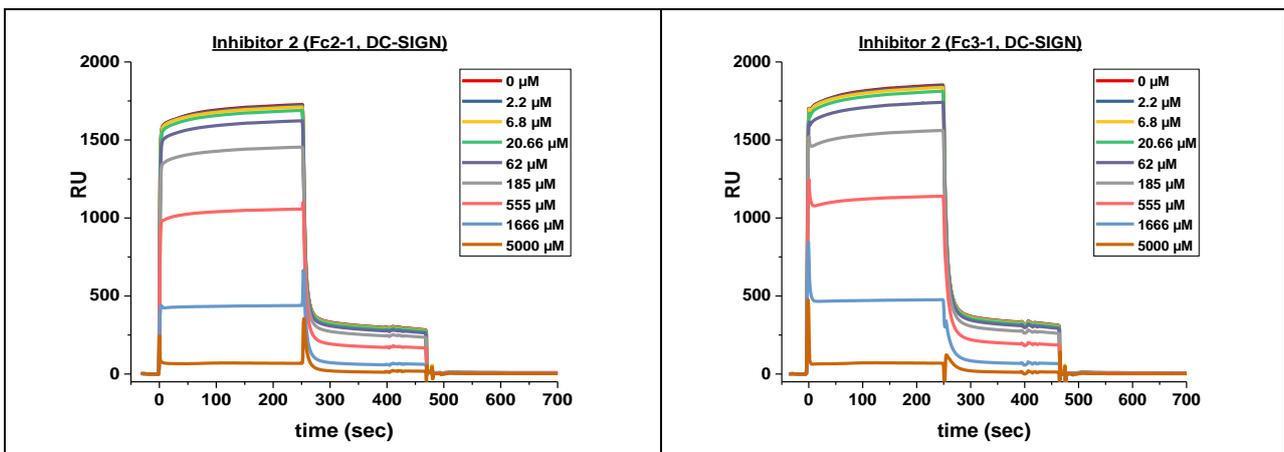
Compound name in thesis	Corresponding initial name	Compound name in thesis	Corresponding initial name
17.1	ISh001	18.1	ISh049
17.2	ISh002	18.2	ISh050
17.3	ISh003	18.3	ISh051
17.4	ISh004	18.4	ISh052
17.5	ISh005	18.5	ISh053
17.6	ISh006	18.6	ISh054
17.7	ISh007	18.7	ISh055
17.8	ISh008	18.8	ISh056
17.9	ISh009	18.9	ISh057
17.10	ISh010	18.10	ISh058
17.11	ISh011	18.11	ISh059
17.12	ISh012	19.1	AZBCN
17.13	ISh013	19.2	ISh044
17.14	ISh014	20	ISh045
17.15	ISh015	21	ISh046
17.16	ISh016		
17.17	ISh017		
17.18	ISh018		
17.19	ISh019		
17.20	ISh020		
17.21	ISh021		
17.22	ISh022		
17.23	ISh023		
17.24	ISh024		
17.25	ISh025		
17.26	ISh026		
17.27	ISh027		
17.28	ISh028		
17.29	ISh029		
17.30	ISh030		
17.31	ISh031		
17.32	ISh032		
17.33	ISh033		
17.34	ISh034		
17.35	ISh035		
17.36	ISh036		
17.37	ISh037		
17.38	ISh038		
17.40	ISh039		
17.1	ISh040		

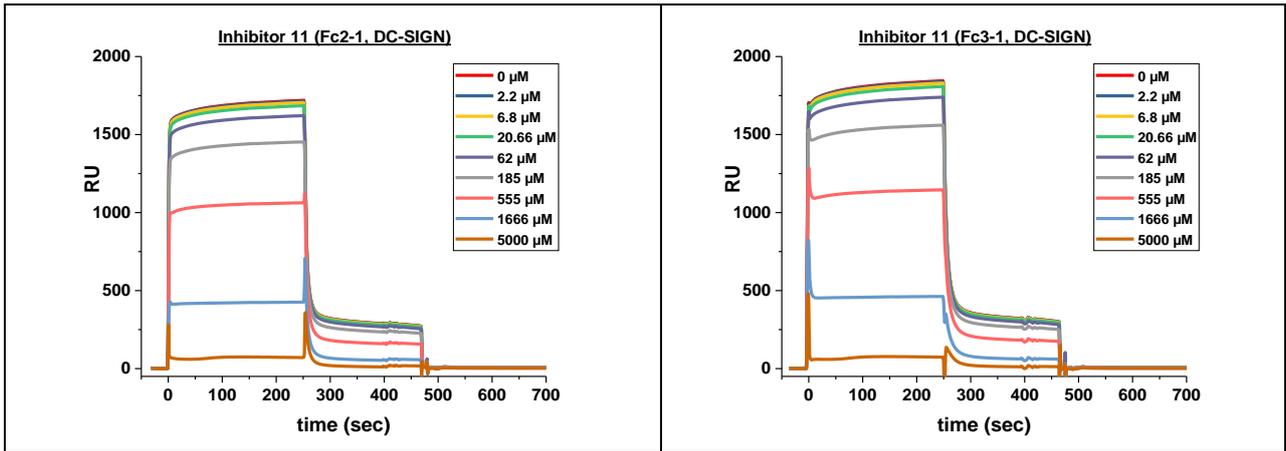
ANNEXES Chapter 10

Control

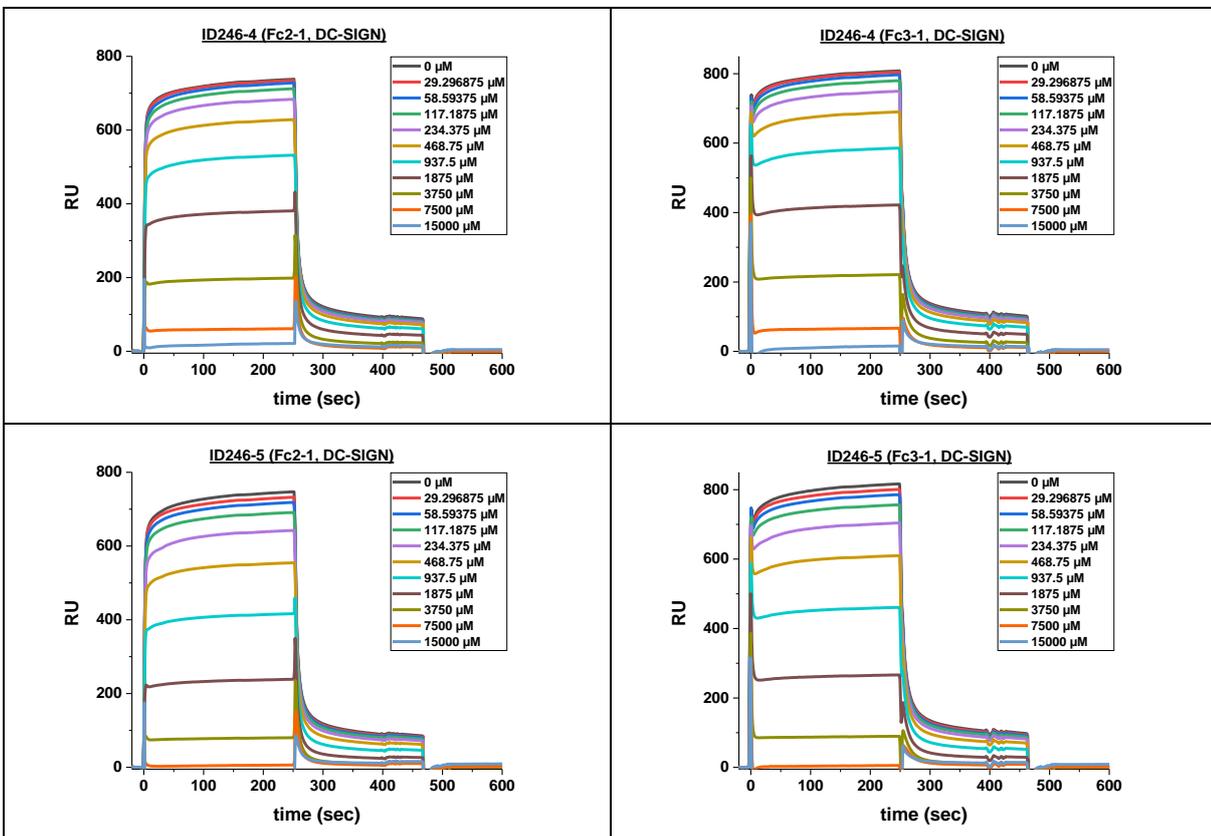


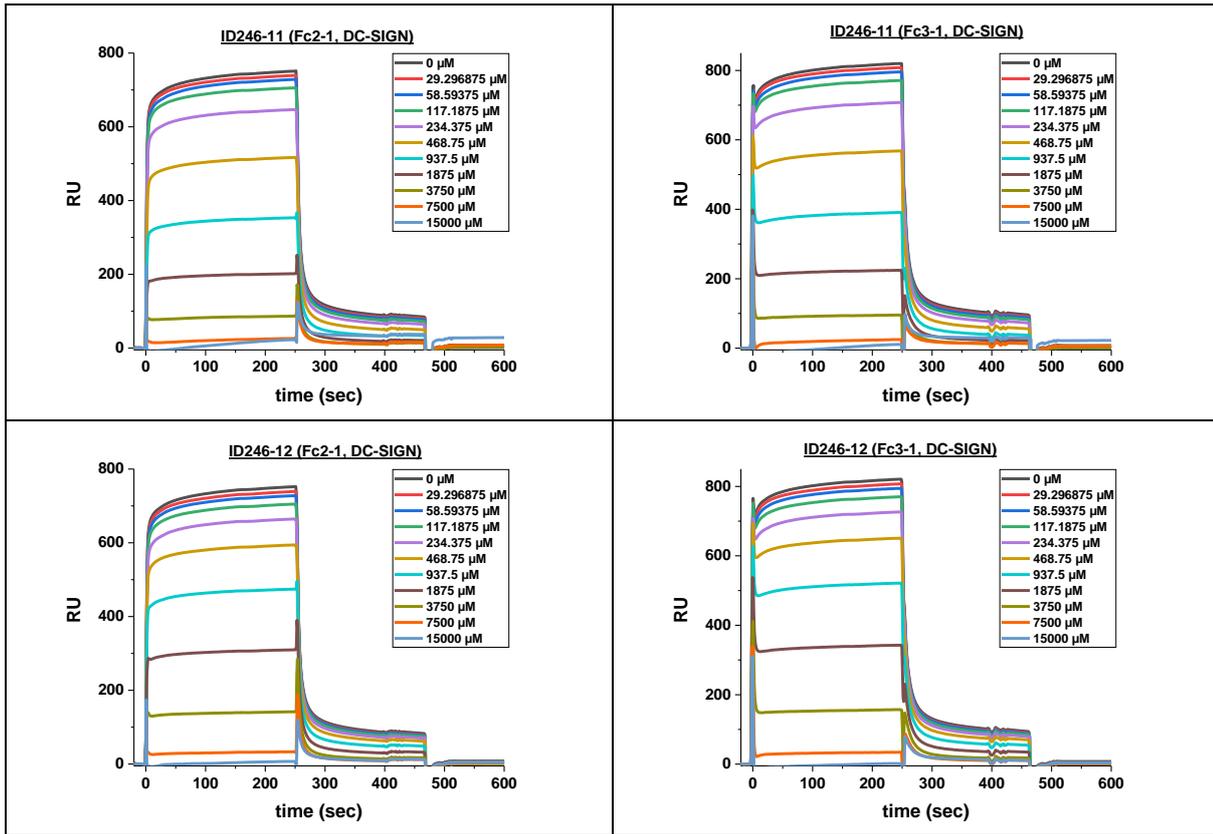
S-linkage



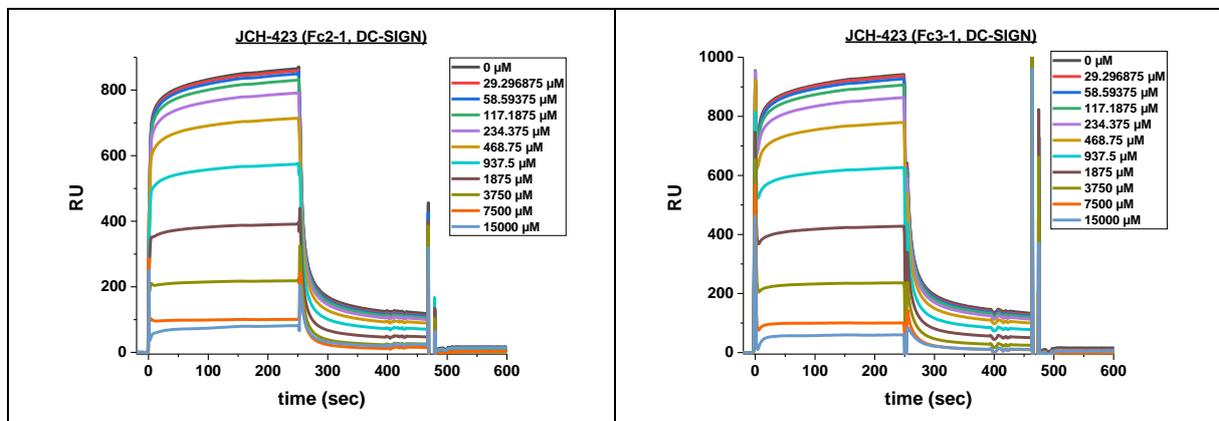


C-linkage (glycosides)

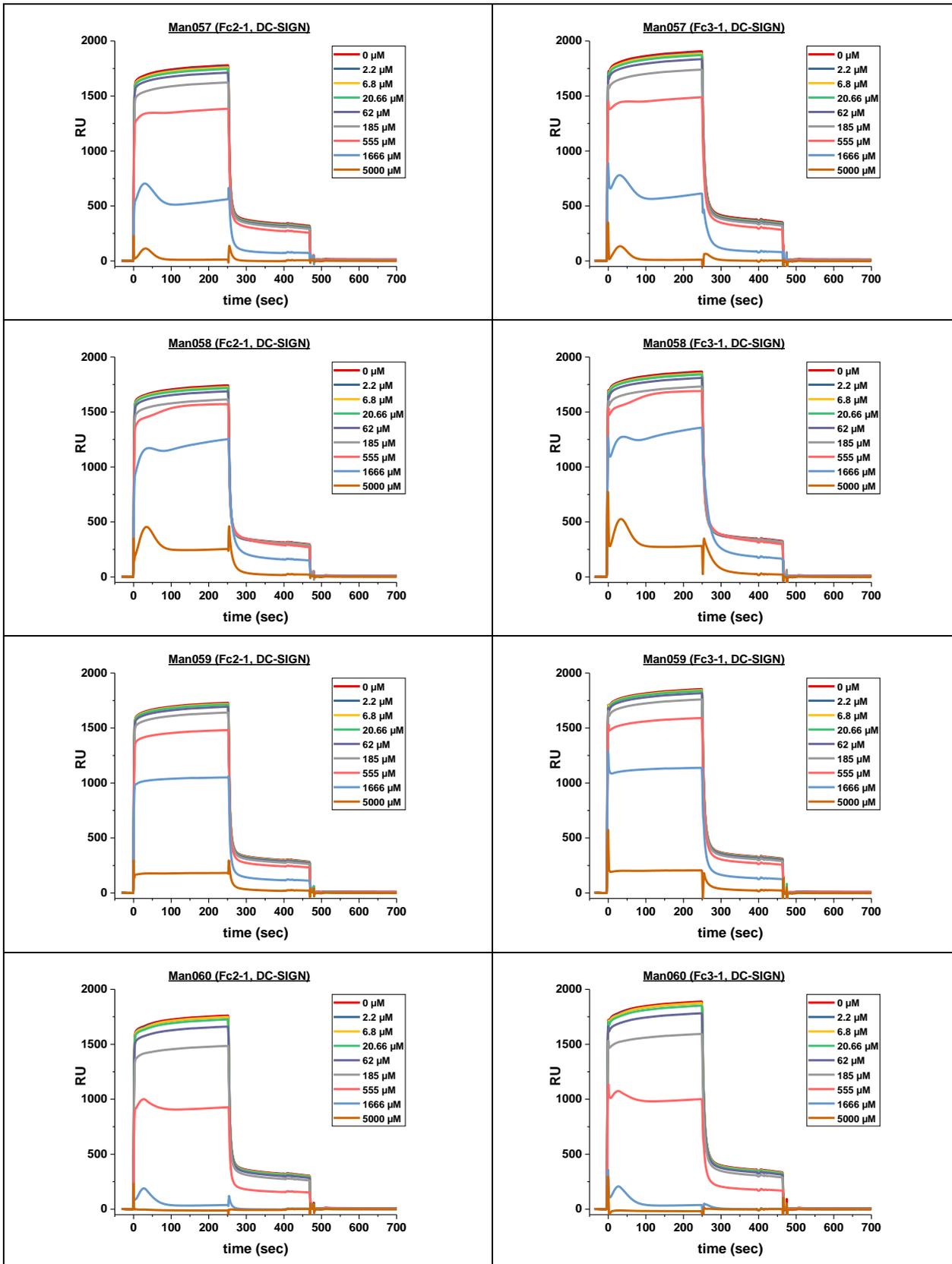


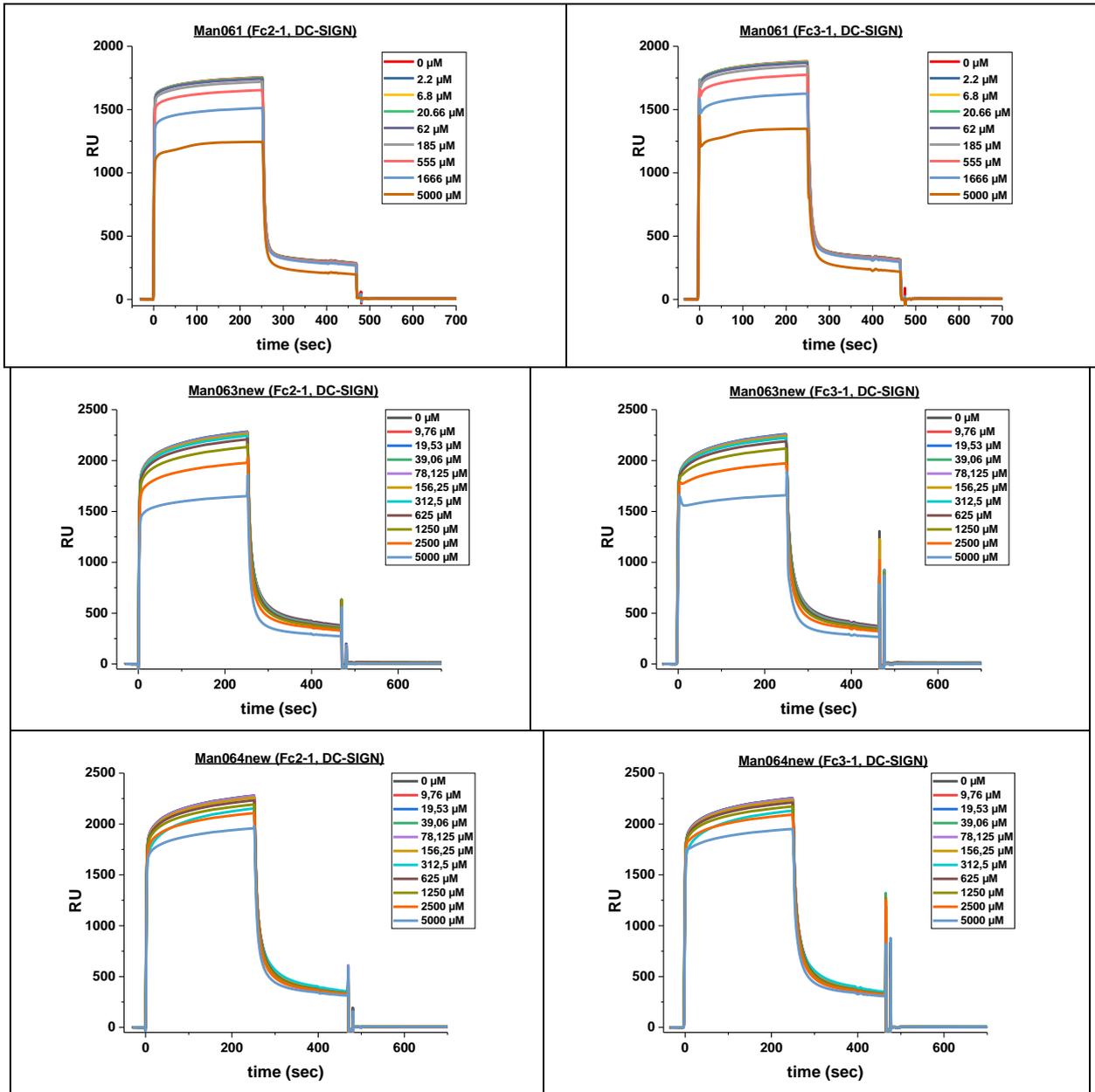


Si-linkage (glycosides)

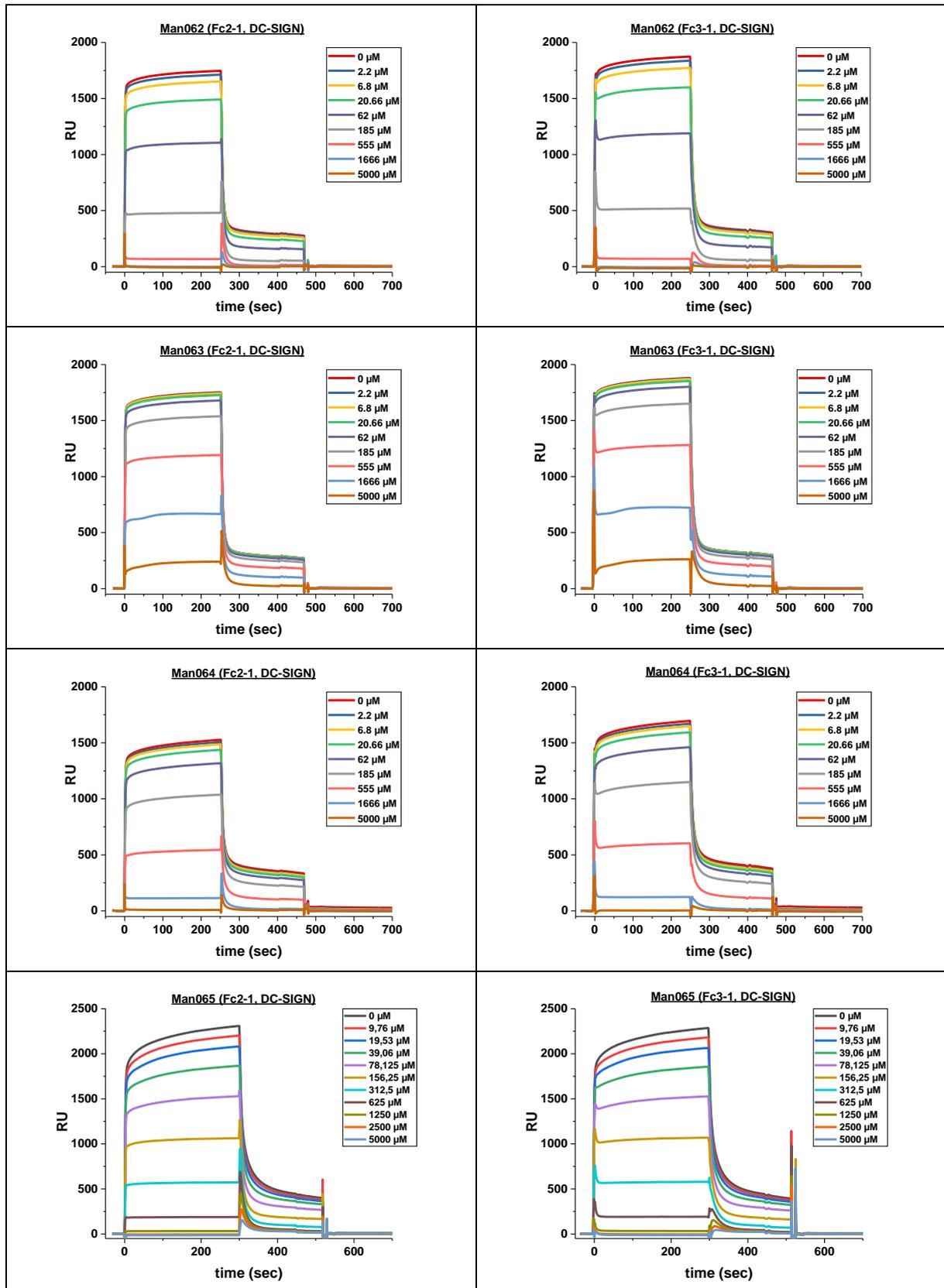


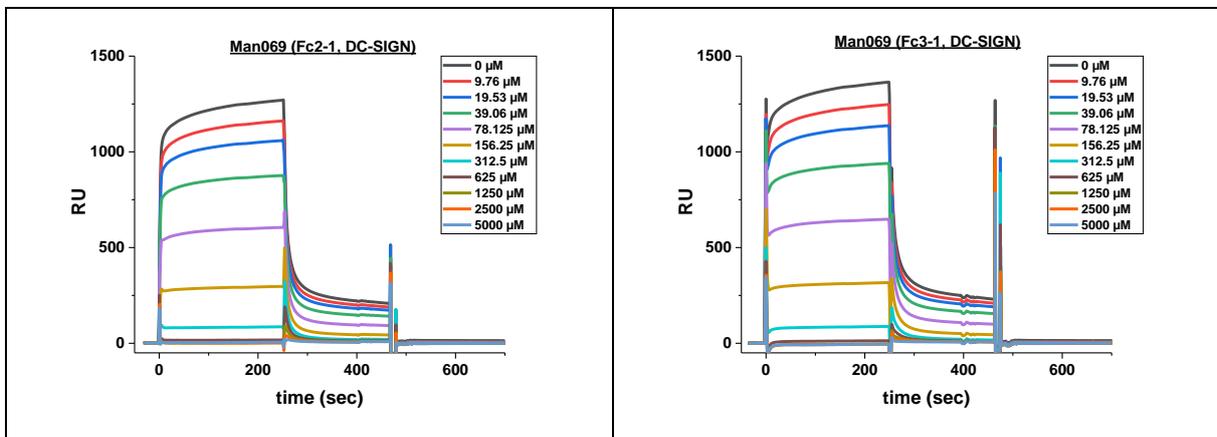
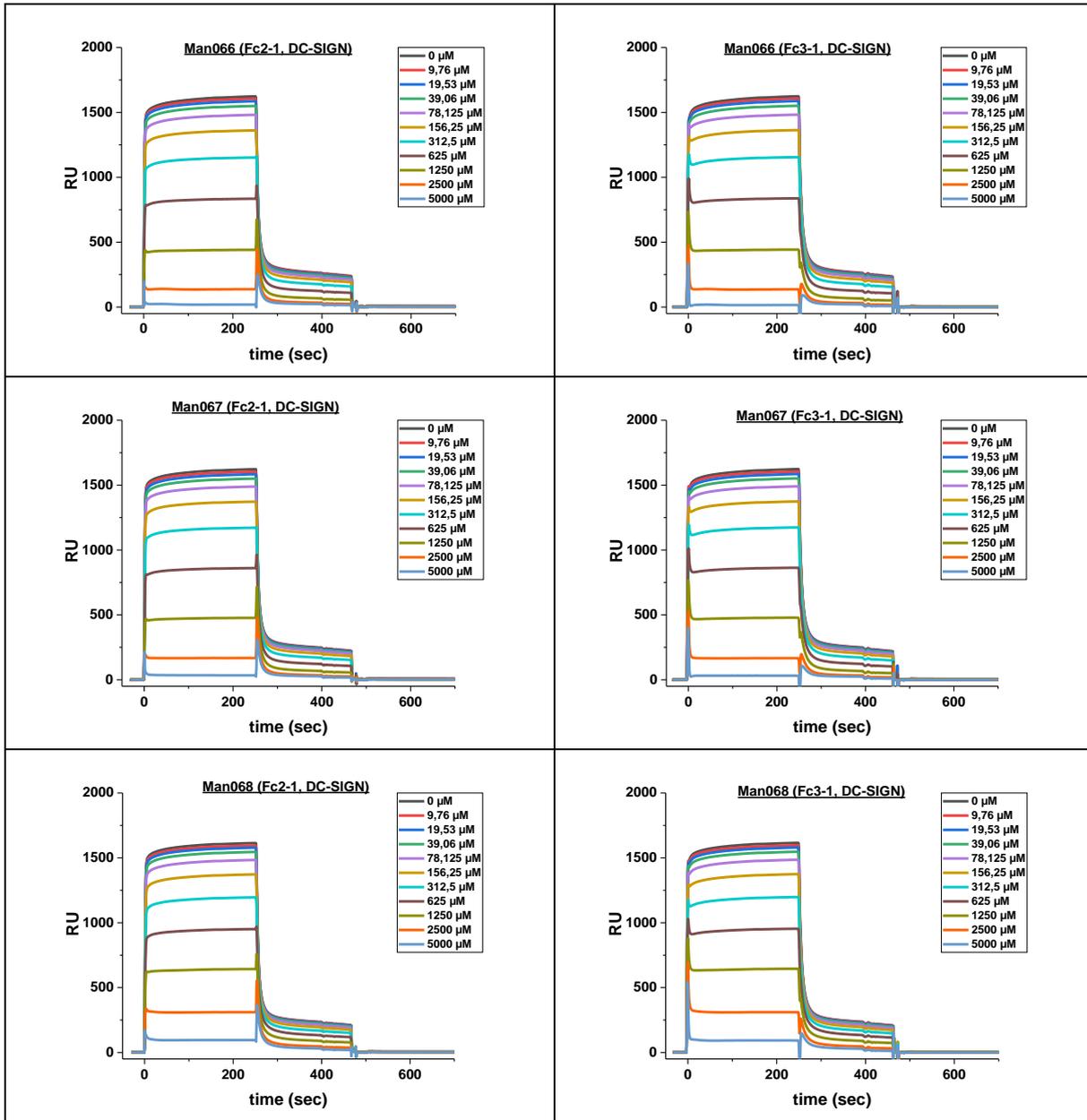
Amino derivatives

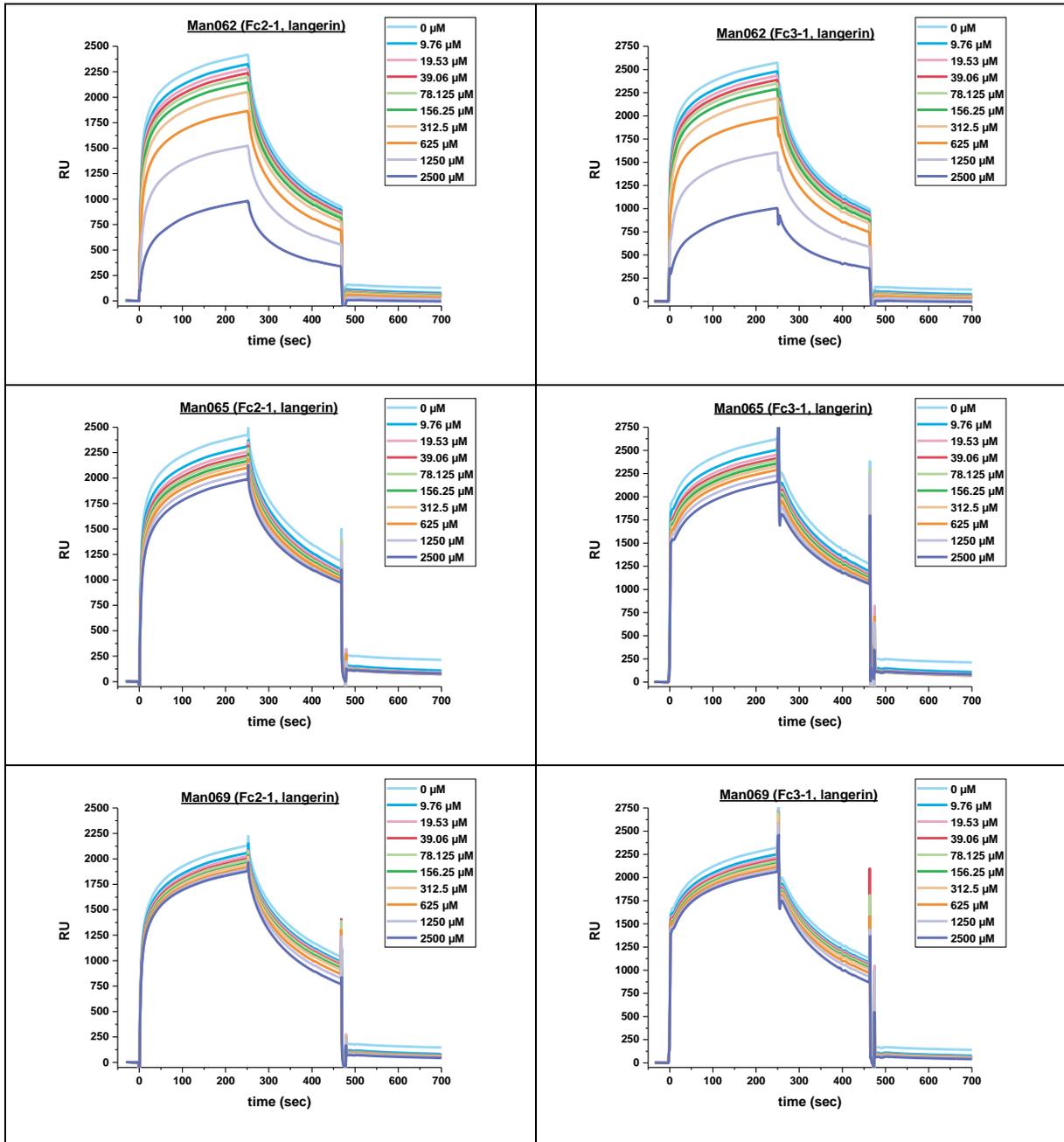




Triazole derivatives







All acquired IC₅₀/K_d values from chapter 10

<u>Type of Compound</u>	<u>Compound</u>	<u>IC₅₀ μM</u>
<i>Control</i>	psDi	714.7 ± 9,2
	D-Mannose	2500
<i>S-linkage</i>	2	801,1 ± 13,79
	11	782.13 ± 8,82
<i>C-linkage (glycosides)</i>	ID246-4	885.2±1.1
	ID246-5	1439.5±15.5
	ID246-11	885.2±1.1
	ID246-12	1099±4
<i>Si- linkage (glycosides)</i>	JCH-423	1556.5±37.5
<i>Amino derivatives</i>	Man057	1143±15.94
	Man058	3011±55
	Man059	2274±34
	Man061	x
	Man063 NEW	>9000
	Man064 NEW	x
<i>Triazole derivatives</i>	Man062	113.09 ± 11.99
	Man064	145.4 ± 1.9
	Man065	339 ± 6.3 μM
	Man066	
	Man067	
	Man068	

	<u>IC₅₀ μM (SPR)</u>	<u>K_D μM (ITC)</u>	<u>K_{Dapp} μM (SPR)</u>
Man069	76.25 ± 3.25	52.08 ± 1.32	50

Supplementary Information

to

Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist.

Alice Tamburrini,^[a] Silvia Achilli,^[b] Francesca Vasile,^[a] Sara Sattin,^[a] Corinne Vivès,^[b] Cinzia Colombo,^[a] Franck Fieschi,^[b] Anna Bernardi,^[a] *

[a] Università degli Studi di Milano, Dipartimento di Chimica, via Golgi 19, 20133 Milano, Italy

[b] Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France

* Corresponding author: anna.bernardi@unimi.it

Content:

<i>DC-SIGN binding inhibition studies: SPR sensograms</i>	p 2
<i>Calculated interproton distances in 1 and 2</i>	p 3
<i>¹H and ¹³C NMR spectra of new compounds</i>	p 4

Supplementary Figures

Figure S1: SPR sensograms of compound 1 (R=CH₂CH₂N₃).

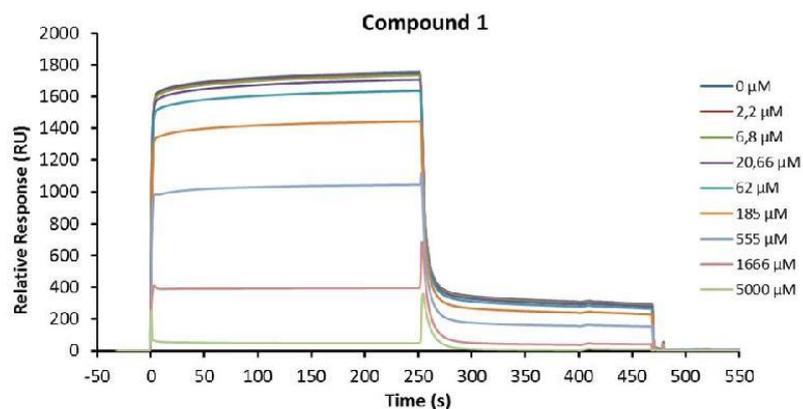


Figure S2: SPR sensograms of compound 2 (R=H).

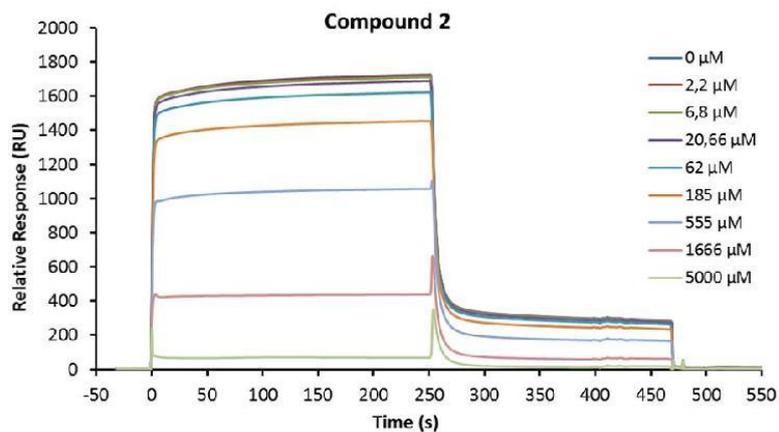
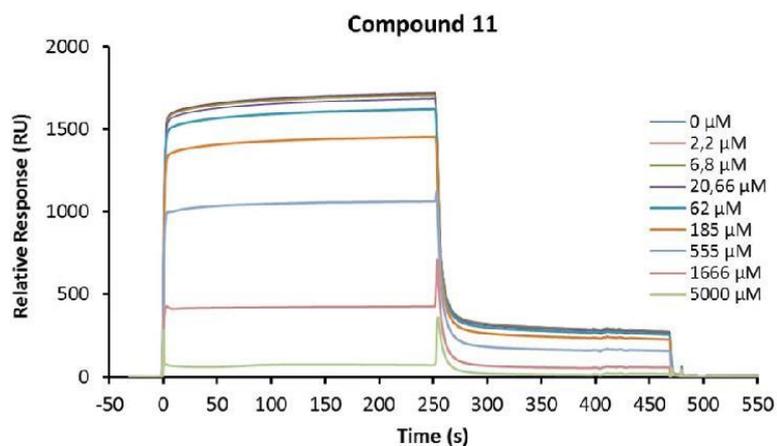
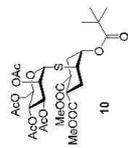


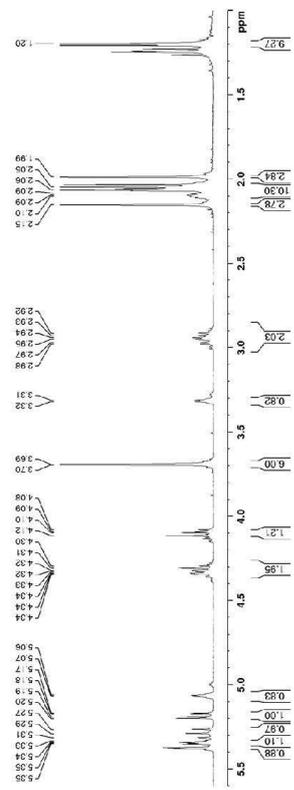
Figure S3: SPR sensograms of compound 11.



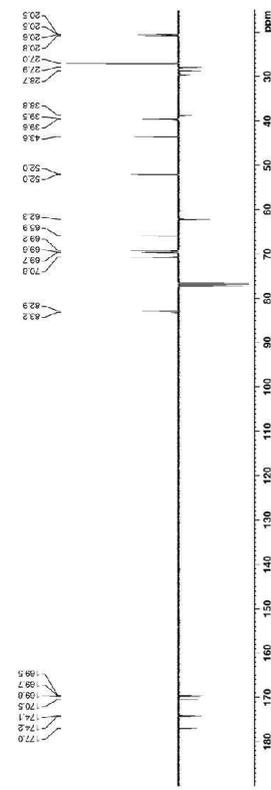
Compound 10



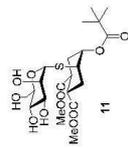
¹H NMR (400 MHz, CDCl₃):



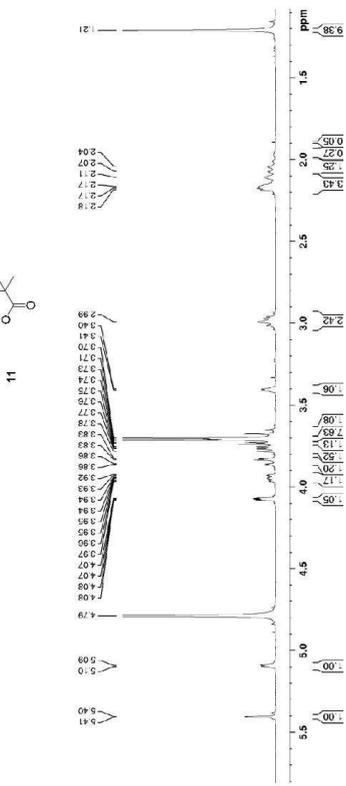
¹³C NMR (100 MHz, CDCl₃):



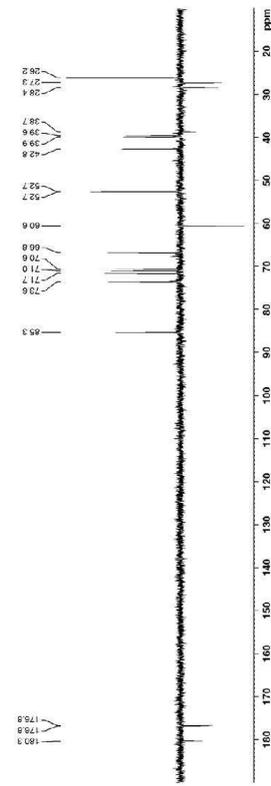
Compound 11



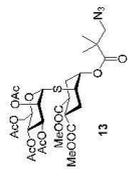
¹H NMR (400 MHz, D₂O):



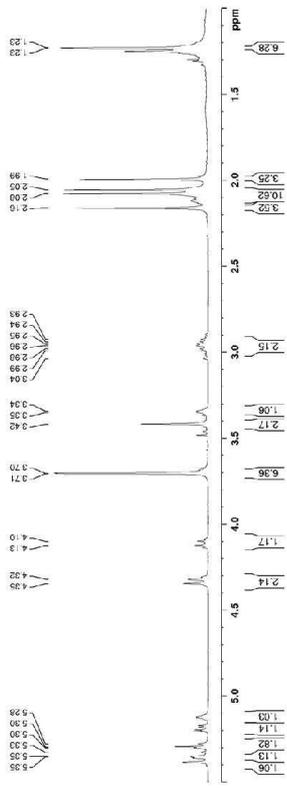
¹³C NMR (100 MHz, D₂O):



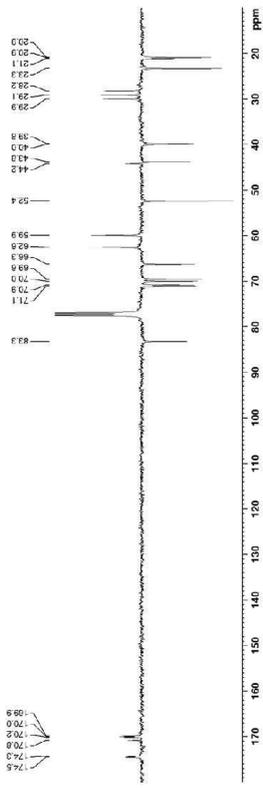
Compound 13



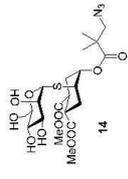
¹H NMR (400 MHz, CDCl₃):



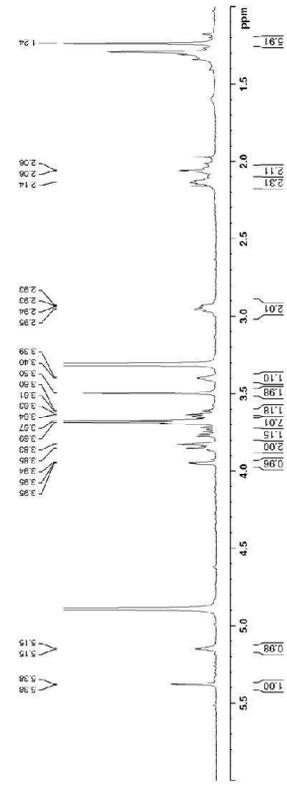
¹³C NMR (100 MHz, CDCl₃):



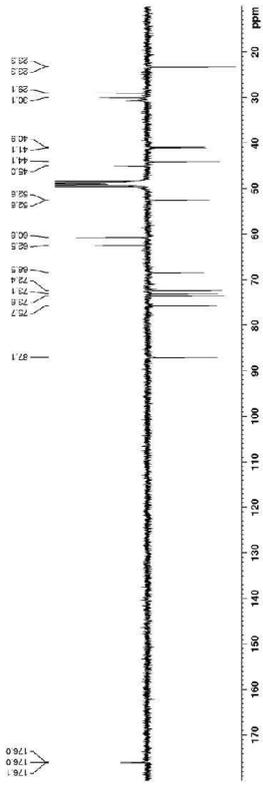
Compound 14



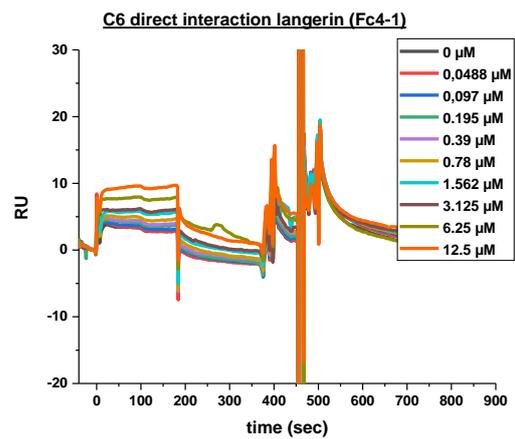
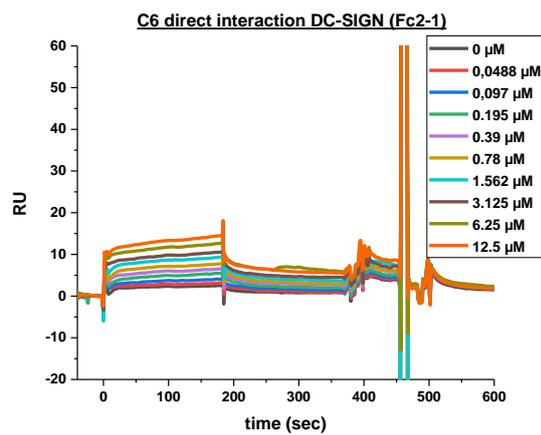
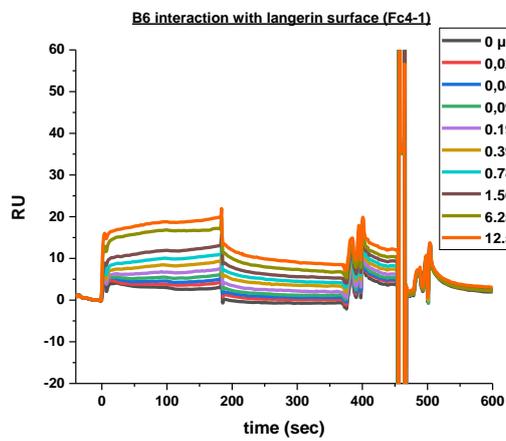
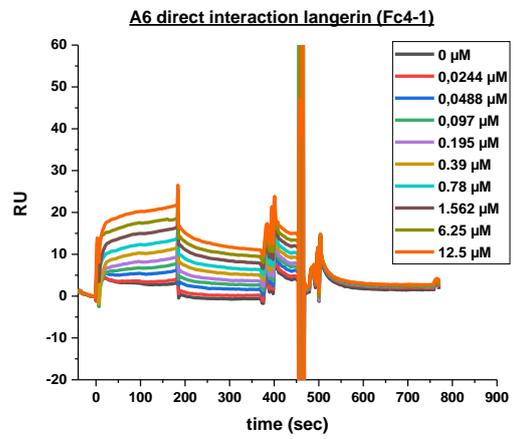
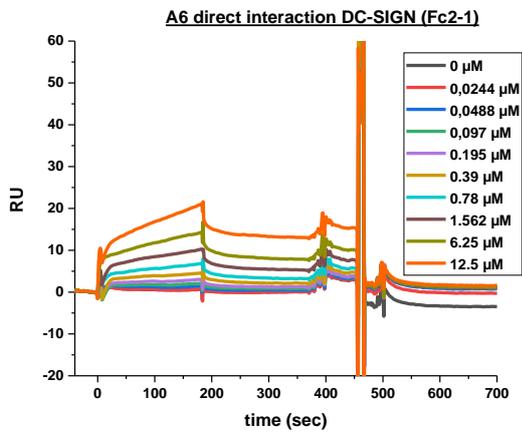
¹H NMR (400 MHz, CD₃OD):

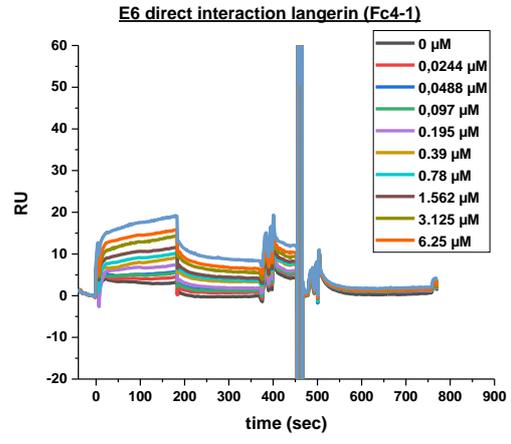
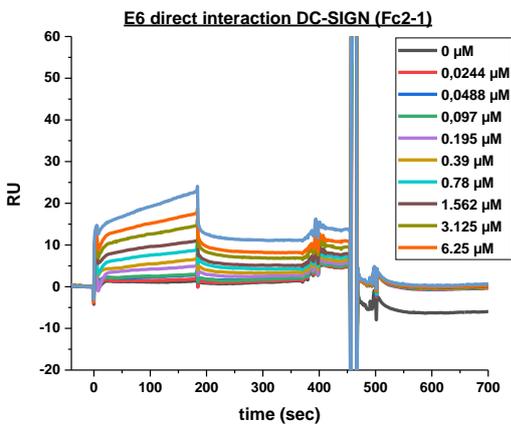
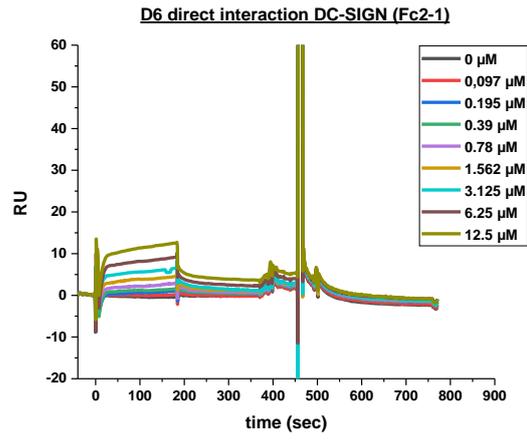
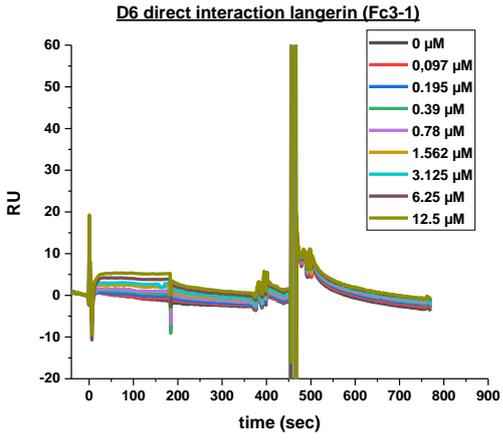


¹³C NMR (100 MHz, CD₃OD):



ANNEXES Chapter 11





Oral Communications

Axe I2HP (Immunity and host-pathogen interactions, Grenoble (IBS), *September 2015*)

“Development of Selective Carbohydrate Immunomodulators Targeting C-Type Lectin Receptors on Antigen Presenting Cells”

IMMUNOSHAPE meeting, Potsdam (DE), *November 2015*

“Recombinant production of human C-type lectins: characterization of ligand binding profile, selectivity and structure of complex with ligands”

IMMUNOSHAPE meeting, Grenoble (IBS), *June 2016*

“A sweet talk: C-type Lectin Receptors and their interaction with selective glycans”

IMMUNOSHAPE meeting Bilbao (ES), *October 2016*

“Surface Plasmon Resonance World and CLRs”

IMMUNOSHAPE meeting, Amsterdam (NL), *February 2017*

“Recombinant production of human C-type lectin: “Characterization of ligand binding profile, selectivity and structure of complex with ligands”

Journée BLI-SPR, Grenoble, *September 2017*

“Affinity and Avidity improvement of C-type lectin receptor ligands through the use of different SPR approaches”

GlycoAlps morning, Grenoble (CERMAV), *September 2017*

“Oriented SPR surface to study Cluster Targeting C-type lectin receptors: towards tool for an anticancer immune response activation”

IMMUNOSHAPE meeting, Orléans (FR), *September 2017*

“Affinity and Avidity improvement of C-type lectin receptor ligands: different SPR approaches”

27es Journées du groupe Français des Glycosciences (GFG), Nouan le Fuzelier (FR), *May 2018*

“Rational Optimization of glycomimetics towards DC-SIGN”

Posters

The Oxford Chemical immunology Conference, Oxford (UK), *April 2016*

Silvia Achilli, Blanka Didak, Corinne Vivès, Michel Thépaut, Ludovic Landemarre, Franck Fieschi

“C-type lectins receptor (CLR) arrays to screen immunocompatibility and reactivity of biological sample.”

26es Journées du groupe Français des Glycosciences » (GFG), Aussois (FR), *May 2016*

Silvia Achilli, Blanka Didak, Corinne Vivès, Michel Thépaut, Ludovic Landemarre, Franck Fieschi

“C-type lectins receptor (CLR) arrays to screen immunocompatibility and reactivity of biological sample.”

IBS day, Grenoble, *September 2016*

Silvia Achilli, Blanka Didak, Corinne Vivès, Michel Thépaut, Ludovic Landemarre, Franck Fieschi

“C-type lectins receptor (CLR) arrays to screen immunocompatibility and reactivity of biological sample.”

19th European Carbohydrate Symposium EUROCARB, Barcelona (ES), *July 2017* (+flash presentation)

Silvia Achilli, Corinne Vivès, Michel Thépaut, Sonia Serna, Niels Reichardt, Joao T. Monteiro, Bernd Lepenies and Franck Fieschi.

“TETRALEC, Artificial Tetrameric Lectins, a tool for multivalency enhancement.”

Publications

N°	Status	Title	Section
N°1	In preparation	“TETRALEC, Artificial Tetrameric Lectins: a Tool to Screen Ligand and Pathogen Interactions” S. Achilli et al (2018)	8.1.3
N°2	Published	“Chemoenzymatic Synthesis of N-glycan Positional Isomers and Evidence for Branch Selective Binding by Monoclonal Antibodies and Human C-type Lectins Receptors” B. Echeverria, S. Serna, S. Achilli et al, ACS ChemBiol (2018)	8.2.1.1
N°3	Published	“L. Medve, S.Achilli et al, Submi On-chip screening of a glycomimetic library with C-type lectins reveals structural features responsible for preferential binding of dectin-2 over DC-SIGN/R and langerin” L. Medve, S.Achilli et al, Chemistry (2018)	8.2.2.1
N°4	Published	“Facile access to pseudo-thio-1,2-dimannoside, a new glycomimetic DC-SIGN antagonist” A. Tamburrini, S. Achilli et al, Bioorg Med Chem (2017)	8.3.1.1

Abstract

C-type Lectin Receptors (CLRs) are carbohydrate-binding proteins mainly expressed on Antigen Presenting Cells (APCs), including dendritic Cells (DCs), the sentinel of the innate immune system. They recognize pathogens or damaged cells by interacting with glycan features and the encounter between the CLR and its ligand constitutes a necessary step for the activation of the adaptive immune system. This crucial role played by CLRs in the balance of immune responses offers to CLR-glycan interactions pharmaceutical applications. The long-term objective of the research project in which this PhD is included is to use these CLRs as modulators in order to tailor the immune system responses. To do so, neoglyco-conjugates selective to each individual CLR have to be developed.

Nine different CLRs were produced in this work: BDCA2, DC-SIGN, DC-SIGNR, dectin-1, dectin-2, langerin, LSECtin, MCL and Mincle. Several approaches have been explored in parallel for CLR production, ranking from bacterial periplasmic targeting, aiming to express soluble and functional protein, to inclusion bodies production into the bacterial cytoplasm, with subsequent protein refolding. Our collection of CLRs were used to screen glycan and glycomimetic arrays, highlighting context-dependent binding and identifying natural ligands or glycomimetics selective to each CLRs. Thus, several CLRs were surprisingly able to differentiate between positional isomers of a given N-Glycan, which opens new questions regarding the biological significance. Moreover, glycomimetics with a selectivity towards dectin-2 over DC-SIGN, DC-SIGNR and langerin CLRs have been identified.

To guide the choice of the glycomimetics and estimate their optimisation, diverse biophysical studies were performed to evaluate the strength and specificity of the interaction. This enabled the development of an ultimate ligand selective towards DC-SIGN. A co-crystallised structure of the protein with this ligand revealed an interesting binding mode that also opens new questions.

Simultaneously to monovalent ligand optimization, a first step towards the design of a highly defined molecule for cancer vaccination by CLR targeting was made. SPR results revealed potential candidates to exploit and preliminary biological assays were performed. Finally, a strategy for tetrameric lectin engineering as been explored, termed TETRALEC. This tool for screening and lectin characterization, has been obtained with one the lectin of the study, DC-SIGNR, by a site-specific labelling of the lectin. The TETRALEC complex was structurally characterised and functional assays were performed on glycan array and pathogen cells.

Keywords: *C-type lectins, screening, glycomimetics, multivalency, avidity*

Résumé

Les lectines de type C (CLRs) sont des récepteurs impliqués dans la reconnaissance d'oligosaccharides et principalement exprimés à la surface des cellules présentatrices d'antigène (APCs) et notamment des cellules dendritiques (DCs), véritable sentinelle de notre système immunitaire. Elles sont impliquées dans la reconnaissance de motifs spécifiques exprimés à la surface d'agents pathogènes et sont capables de stimuler le système immunitaire afin de déclencher une réponse adaptée. Ce rôle crucial joué par les CLRs dans l'équilibre de la réponse immunitaire confère aux interactions CLR/glycane des perspectives d'applications pharmaceutiques. L'objectif à long-terme du projet de recherche dans lequel cette thèse s'intègre consiste à utiliser ces CLRs pour moduler les réponses du système immunitaire. A cette fin, des néoglycoconjugués spécifiques de chaque CLR doivent être développés. Au cours de cette thèse, 9 CLRs ont été produits BDCA2, DC-SIGN, DC-SIGNR, dectin-1, dectin-2, langerin, LSECtin, MCL and Mincle. Différentes stratégies de production ont été testées en parallèle, incluant des techniques d'adressage au périplasm en vue d'obtenir des protéines solubles et fonctionnelles et de l'expression cytoplasmique, sous forme de corps d'inclusion suivie d'étapes de renaturation qui s'est révélé la plus efficace au final. Une stratégie permettant de construire des tétramères artificiels de CLRs, appelés TETRALEC, a été mise au point. Cet outil permettant le criblage et la caractérisation des lectines a été obtenu avec DC-SIGNR par un marquage spécifique de la lectine. Le complexe TETRALEC a été caractérisé au niveau structural et des tests fonctionnels ont été menés sur des puces à glycanes et des cellules pathogènes. La série de CLRs que nous avons produites a été utilisée pour cribler des puces à glycanes et à glycomimétiques. Ces études nous ont permis de mettre en évidence des interactions dépendantes de l'environnement du glycane et d'identifier de nouveaux glycanes ou glycomimétiques spécifiques de certains CLRs. En effet, de manière étonnante, plusieurs des CLRs testés sont capables, pour un glycane donné, de discriminer des isomères de position ouvrant ainsi de nouveaux questionnements sur la signification biologique de cette sélectivité. De plus des glycomimétiques reconnaissant préférentiellement dectin-2 par rapport à DC-SIGN, DC-SIGNR et langerin ont été identifiés. Le choix des glycomimétiques et l'évaluation des étapes de leur optimisation ont été permis par diverses études biophysiques qui ont quantifié la force et la spécificité des interactions. Ceci a permis le développement d'un ligand optimisé sélectif de DC-SIGN. La co-cristallisation de la protéine avec ce ligand a révélé un intéressant mode de liaison qui amène également de nouvelles questions. Simultanément à l'optimisation de ligands monovalents, un premier pas a été réalisé vers la conception d'une molécule pour permettre une vaccination contre le cancer médiée par les CLRs. Les résultats de SPR ont identifié des candidats potentiellement intéressants et des tests biologiques préliminaires ont été réalisés.

Mot-clés: *Lectines de type C, criblage, glycomimétisme, multivalence, avidité*