# Développement de glycomimétiques antagonistes du récepteur lectine de type C, DC-SIGN : une nouvelle stratégie préventive anti-HIV

Ieva Sutkeviciute

▶ To cite this version:

Ieva Sutkeviciute. Développement de glycomimétiques antagonistes du récepteur lectine de type C, DC-SIGN : une nouvelle stratégie préventive anti-HIV. Sciences agricoles. Université de Grenoble, 2012. Français. <NNT : 2012GRENV055>. <tel-00819832>

HAL Id: tel-00819832

https://tel.archives-ouvertes.fr/tel-00819832

Submitted on 2 May 2013

**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.

UNIVERSITÉ DE GRENOBLE

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

## Ieva SUTKEVIČIŪTĖ

Thèse dirigée par **Franck FIESCHI**

préparée au sein du **Institut de Biologie Structurale – Jean-Pierre Ebel**
dans **l'École Doctorale Chimie et Sciences du Vivant**

# Développement de glycomimétiques antagonistes du recepteur lectine de type C, DC-SIGN : une nouvelle stratégie préventive anti-HIV

Thèse soutenue publiquement le **14/12/2012**,
devant le jury composé de :

**Mr, Bruce, TURNBULL**
Dr., University of Leeds, United Kingdom, Rapporteur
**Mme, Julie, BOUCKAERT**
Dr., Université des Sciences et Technologies de Lille, Rapporteur
**Mme, Anna, BERNARDI**
Pr., Universitá degli studi di Milano, Italy, Examinateur
**Mr, Olivier, RENAUDET**
Pr., Université Joseph Fourier-Grenoble, Examinateur
**Mr, Franck, FIESCHI**
Pr., Université Joseph Fourier-Grenoble, Directeur de thèse

*Université Joseph Fourier / Université Pierre Mendès France /*
*Université Stendhal / Université de Savoie / Grenoble INP*

# Preamble

This thesis was accomplished in the Institut de Biologie Structurale – Jean-Pierre Ebel, Grenoble (FRANCE), under supervision of Pr. Franck Fieschi.

Although this work is affiliated to Grenoble University Joseph Fourier and by the University regulations it should be written in French language, due to the European context of the work and my non-french nationality, it is presented in English language. However, the requirement of the title to be in French is regarded, and the title of the thesis in English is the following:

## Development of the glycomimetic antagonists of the C-type lectin receptor DC-SIGN: a new strategy of HIV infection prevention

The work presented herein is the part of the research project of the European ITN CARMUSYS (`http://www.carmusys.iiq.csic.es/`), which aims to develop inhibitors of DC-SIGN, the C-type lectin receptor usurped by a wide range of pathogens, including HIV.

The thesis project was conducted in several collaborations with the following researchers:

- ✧ Pr. Anna Bernardi, *Universita' degli Studi di Milano, Dipartimento di Chimica Organica e Industriale and CISI*, Milano, ITALY.
- ✧ Dr. F. Javier Rojo Marcos, *Instituto de Investigaciones Químicas, CSIC*, Seville, SPAIN.
- ✧ Pr. Jitka Moravcová, *Institute of Chemical Technology, Prague*, CZECH REPUBLIC.
- ✧ Dr. Pedro Manuel Nieto Mesa, *Instituto de Investigaciones Químicas, CSIC*, Seville, SPAIN.
- ✧ Dr. Jörg Weiser, *Anterio Consult & Research GMBH*, Mannheim, GERMANY.

The overall structure of ITN CARMUSYS and the description of the collaborations can be found in figure 4.1 on page 94.

The thesis is organized in four parts: **Introduction**, where the scientific background and the objectives of the thesis project are presented; **Matherials&Methods**; **Results**; and **Discussion and future perspectives**. The organization of the results in this thesis can be viewed in figure 7.1 on page 119. In the **Appendix** the articles and the relevant information is added.

The work lasted from 1st of October, 2009, untill 9th of December, 2012. During this period, I have accomplished two secondments in the context of CARMUSYS network:

- ✧ a short-term secondment (15th - 19th November, 2010) in *Anterio Consult & Research GMBH.* Mannheim (Germany) with Dr. Pietro Alfarano, a post-doctoral student in CARMUSYS. There I got familiar with the docking procedures of the small molecules to DC-SIGN binding site using Schrödinger Suite (Glide). I have also got acquainted with other applications within Schrödinger Suite (Maestro, Primex, etc.);

- a long-term secondment (26th September - 7thDecember, 2011) in *Instituto de Investigaciones Químicas*, CSIC. Seville (Spain) with a PhD student Cinzia Guzzi, Dr. Jesus Angulo and Dr. Pedro Nieto. I learnt there the basics of nuclear magnetic resonance (NMR) theory, saturation transfer difference (STD) NMR technique and was involved in STD NMR experiments as well as data analysis.

Additionally, I have supervised the long-term secondment (7th March 2011 - 19th May 2011) of a PhD student Joanna Luczkowiak. During this period, we attempted to study the interaction of pseudotyped viral particles (HIV, Ebola) with immobilized DC-SIGN using surface plasmon resonance technique. The results of these studies are not presented in this thesis.

The work was financially supported by the Marie Curie ITN FP7 project CARMUSYS (PITN-GA-2008-213592).

# Acknowledgements

This doctoral thesis would not have been possible without the help and the guidance of several people, and I would like to express my sincere gratitude to all of them.

First of all and foremost, I owe my deepest gratitude to my supervisor Pr. Franck Fieschi, who believed in me without seeing and welcomed me to his team and gave a great opportunity to pursue this scientific adventure. I am grateful for his valuable guidance and support, for being available despite his busy schedule, and for an utmost trust he showed to me during all these three years. Also a big thanks to him for finding time to translate the summaries of each chapter of this thesis to French language.

My immense gratitude goes to Dr. Michel Thépaut who patiently taught me all the tips and tricks of protein biochemistry and whose positive and witty attitudes were always inspiring. I am grateful for all of his valuable advices and support I got during these years, and for the help in all other life situations and translation of thesis summary to French language.

My sincere thanks also goes to Dr. Corinne Vivès for teaching me the secrets of molecular biology, for her goodwill and supportiveness.

I would like to express my deepest appreciation to the jury members, Pr. Anna Bernardi, Dr. Julie Bouckaert, Dr. Bruce Turnbull and Pr. Olivier Renaudet, who kindly accepted to evaluate my dissertation and thoroughly examined it. Thank you all for such a good evaluation!

I wish to thank to Isabelle Petit-Haertlein and Céline Juillan-Binard for taking care of all the four corners of the laboratory, which made the work so much easier. I am very grateful to Isabelle for her help to perform experiments when I needed it mostly and for her amity and never fading smile.

I owe a great acknowledgement to the Biacore platform of the institute, particularly to Dr. Nicole Thielens and Isabelle Bally. Without their care of the instruments, the SPR experiments would not have been possible.

I am indebted to AUC and PAOL platform of the institute, and especially to Pr. Christine Ebel and to Aline LeRoy for helping and teaching to perform AUC and DLS experiments and assistance in data treatment. I am particularly thankful to Christine for her availability and very kind help on the complicated data interpretation.

The interesting ITC data would not have been acquired without an access to the ITC instrument in CERMAV, and hereby I would like to express my sincere gratitude to Pr. Anne Imberty for giving this access and for all of her helpful advices. I would also like to thank to Dr. Aymeric Audfray for teaching to use ITC instrument and treat the data.

I thank Dr. Eric Chabrol for the continuous provision of langerin and for his useful advices.

I would like to give an important credit to Dr. Lina Šiaučiūnaitė for all the unselfish help she offered when I arrived to the institute and continued to be supportive, warm and friendly ever since.

My special thanks goes to Dr. Richard Kahn, who was always radiating good mood, for all the interesting stories and discussions on science and not, and for the company during long evenings in the laboratory and the way back home. Rest in peace, Richard…

The accomplishment of this work would not have been possible without all of the people of European ITN CARMUSYS, especially the teams of Pr. Anna Bernardi, Dr. Javier Rojo and Pr. Jitka Moravcová and their students Norbert Varga, Daniela Doknic, Renato Ribeiro-Viana and Benedetta Bertolotti, who were designing and synthesizing the molecules to be tested. I am also very grateful to CARMUSYS network for the great workshops and memorable, joyful moments shared during these meetings.

My sincere gratitude goes to Dr. Pedro M. Nieto, Dr. Jesús Angulo, Cinzia Guzzi and Juan Carlos Muñoz García for a very warm welcome to their group in IIQ, CSIC in Seville to carry out a 2.5-month long secondment. I am grateful to them all, and especially Jesús, for introducing a mysterious NMR spectroscopy world to me.

I am very thankful to Dr. Jörg Weiser and Dr. Pietro Alfarano for kindly accepting me to Anterio Consult & Research in Mannheim for the whole week. I am grateful to Pietro for his kindness and patience to teach me a very new tool to me – small ligand docking to protein receptor.

I must acknowledge the financial support given by the Marie Curie ITN FP7 project CARMUSYS (PITN-GA-2008-213592), without which my work wouldn't have been possible.

And finally and very importantly I want to give a big thanks to my family, especially to my mother Alina, my brother Adomas and my beloved husband Abed for their constant support and love.

# Abbreviations

**a.a.** amino acids
**Ac-LDL** Acetylated low density lipoprotein
**AGE** Advanced glycation end-product
**AIDS** Acquired Immunodeficiency Syndrome
**APCs** Antigen Presenting Cells
**ARV** Antiretroviral
**AUC** Analytical Ultracentrifugation

**BCG** Mycobacterium bovis Bacillus Calmette Guérin
**BG** Birbeck Granules
**cDNA** complementary DNA
**CEA** Carcinoembryonic antigen
**CLRs** C-type lectin receptors
**CM** CarboxyMethyl
**CMV** Cytomegalovirus
**CRD** Carbohydrate recognition domain
**CRP** C-reactive protein
**CTLDs** C-type lectin-like domains
**CuAAC** Cu(I) catalyzed azide–alkyne cycloaddition

**DAP** Diaminopimelic acid
**DCs** Dendritic cells
**DC-SIGN** DC-Specific ICAM3 Grabbing Non-integrin
**DLS** Dynamic Light Scattering
**DMSO** Dimethyl sulfoxide
**DNA** Deoxyribonucleic acid

**dNTP** deoxyribonucleotide triphosphates

**dsRNA** double-stranded RNA

**ECD** Extracellular domain

**ECM** Extracellular matrix

**EDC** 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride

**EDTA** Ethylenediaminetetraacetic acid

**EPO** Erythropoietin

**Fc** Flow cell

**Flt3L** Fms-like tyrosine kinase 3 ligand

**FPLC** Fast Protein Liquid Chromatography

**G$^-$** Gram-negative bacteria

**G$^+$** Gram-positive bacteria

**GalNAc** N-acetylgalactose amine

**G-CSF** Granulocyte Colony-Stimulating Factor

**Gdn-HCl** Guanidine hydrochloride

**GlcNAc** N-acetylglucose amine

**GM-CSF** Granulocyte-Macrophage Colony-Stimulating Factor

**GPI** Glycosylphosphatidyl inositol

**GM-tripeptide** N-acetyl-D-glucosaminyl-β(1,4)-N-acetylmuramyl-L-Ala-D-Glu

**GNPs** Gold nanoparticles

**HAART** Highly Active AntiRetroviral Treatment

**HCV** Hepatitis C Virus

**HEPES** 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

**HIV** Human Immunodeficiency Virus

**HRP** Horse Radish Peroxidase

**HSC** Hematopoietic Stem Cell

**HSV** Herpes Simplex Virus

**iDCs** immature DCs

**ICAM-3** InterCellular Adhesion Molecule-3

**IFN** Interferon

**IL** Interleukin

**IPTG** β-D-1-thiogalactopyranoside

**ITAMs** Immunoreceptor Tyrosine-based Activation Motifs

**ITC** Isothermal Titration Calorimetry

**ITIMs** Immunoreceptor Tyrosine-based Inhibition Motifs

**ITN** Initial Training Network

**LB** Luria Bertani

**LCs** Langerhans cells

**LDL** Low-Density Lipoproteins

**LDN-antigen** GalNAcβ1-4GlcNAc-R

**LDNF-antigen** GalNAcβ1-4(Fucα1-3)GlcNAc-R

**LPS** Lipopolysaccharide

**LSECtin** Liver and lymph node sinusoidal endothelial cell C-type lectin

**LTA** Lipoteichoic acid

**MØ** Macrophages

**ManBSA** Mannosylated Bovine Serum Albumin

**MARCO** Macrophage receptor with a collagenous structure

**MBL** Mannan-Binding Lectin

**M-CSF** Macrophage Colony-Stimulating Factor

**mDCs** mature DCs

**MDP** Muramyl Dipeptide

**MES** 2-(*N*-morpholino)ethanesulfonic acid

**MGL** Macrophage Galactose-type C-type Lectin

**MHC** Major Histocompatibility Complex

**Mo** Monocytes

**myDCs** myeloid DCs

**MWCO** Molecular Weight Cut Off

**NADPH** Nicotinamide Adenine Dinucleotide Phosphate

**Neu** Neutrophils

**NHS** *N*-hydroxysuccinimide

**NKs** Natural killer cells

**NMR** Nuclear Magnetic Resonance

**NOX** NADPH oxidases

**OAS** 2'-5'-oligoadenylate synthase

**Ox-LDL** Oxidized LDL

**PAMPs** Pathogen Associated Molecular Patterns

**PBMCs** Peripheral Blood Mononuclear Cells

**PBS** Phosphate buffered saline

**PCR** Polymerase Chain Reaction

**pDCs** plasmacytoid DCs

**PES** Polyethersulfone

**PGRPs** Peptidoglycan-recognition proteins

**PKR** Protein kinase R

**PMNs** Polymorphonuclear neutrophils

**PRRs** Pattern Recognition Receptors

**psDi** pseudomannobioside

**psTri** pseudomannotrioside

**PVDF** Polyvinylidene fluoride

**RIG-1** Retinoid acid–inducible gene–1

**RLRs** RIG-1-like receptors

**RNA** Ribonucleic acid

**ROS** Reactive Oxygen Species

**RU** Resonance/Response Units

**SAP** Serum Amyloid Protein

**SARS** Severe Acute Respiratory Syndrome-associated

**SAXS** Small Angle X-ray Scattering

**SCF** Stem Cell Factor

**SD** Standard Deviation

**SDS** Sodium Dodecyl Sulfate

**SDS-PAGE** Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

**SLS** Static light scattering

**SPR** Surface Plasmon Resonance

**STI** Sexually Transmitted Infections

**sLewis$^{X}$** sulphated Lewis$^{X}$

**ssRNA** single-stranded RNA

**TAE** Tris-acetate-EDTA buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA)

**TCR** T Cell Receptor

**TE** Tris-EDTA buffer (10 mM Tris-HCl pH 8, 1 mM EDTA)

**TEMED** Tetramethylethylenediamine

**TGF** Transforming Growth Factor

**Th** T helper cells

**TIRF** Total Internal Reflection Fluorescence

**TLRs** Toll-like receptors

**TM** Transmembrane domain

**Tn-antigen** GalNAcα1-O-Ser/Thr

**TNF** Tumor Necrosis Factor

**TPO** Thrombopoietin

**TRAPs** Transmembrane adaptor proteins

**UGRP-1** Uteroglobin-Related Protein 1

**VHMs** Vaginal HIV microbicides

**WB** Western Blot

**WNV** West Nile Virus

# Contents

Les résumés en Français . . . . . . . . . . . . . . . . . . . . . . . 27

**I. Introduction** **35**

**1. An overview of human innate immunity** **37**

1.1. The components of innate immunity . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
  1.1.1. *Anatomical and chemical-physiological barriers* . . . . . . . . . . . . . . . . . . . 38
  1.1.2. *The cells of innate immunity* . . . . . . . . . . . . . . . . . . . . . . . . . . . 39
1.2. Pathogen recognition by innate immunity . . . . . . . . . . . . . . . . . . . . . . . . . 41
  1.2.1. *The diversity of Pattern Recognition Receptors* . . . . . . . . . . . . . . . . . . 42
  1.2.2. *C-type lectin receptors as PRRs* . . . . . . . . . . . . . . . . . . . . . . . . . 45
    1.2.2.1. *The structure of C-type lectin-like domains* . . . . . . . . . . . . . . 50
    1.2.2.2. *The structural basis of sugar binding within carbohydrate recognition domains* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
1.3. Dendritic cells and adaptive immunity initiation . . . . . . . . . . . . . . . . . . . . 53

**2. Human immunodeficiency virus: the infection via sexual transmission** **57**

2.1. The mechanism of HIV entry to the host cell . . . . . . . . . . . . . . . . . . . . . . . 57
2.2. Dendritic cell implication in HIV-1 sexual transmission . . . . . . . . . . . . . . . . 59
  2.2.1. *Mechanism of HIV invasion of genital mucosa* . . . . . . . . . . . . . . . . . . 59
    2.2.1.1. *Stromal DCs in mucosa are implicated in DC-SIGN-mediated HIV infection enhancement* . . . . . . . . . . . . . . . . . . . . . . . . . 62
    2.2.1.2. *Suppressive role of epithelial LCs in HIV transmission* . . . . . . . 64
2.3. Structure and function of C-type lectin receptor DC-SIGN . . . . . . . . . . . . . . 66
  2.3.1. *DC-SIGN CRD structure and carbohydrate recognition* . . . . . . . . . . . . 67
  2.3.2. *Oligomeric structure of DC-SIGN* . . . . . . . . . . . . . . . . . . . . . . . . 69
  2.3.3. *Clustering of DC-SIGN tetramers on the surface of DCs* . . . . . . . . . . . 70
  2.3.4. *Ligand recognition by DC-SIGN* . . . . . . . . . . . . . . . . . . . . . . . . . 72
    2.3.4.1. *Endogeneous and exogeneous ligands of DC-SIGN.* . . . . . . . . . . 72
    2.3.4.2. *Ligand-induced DC-SIGN signaling* . . . . . . . . . . . . . . . . . . 73
2.4. Langerin: a specific CLR of Langerhans cells . . . . . . . . . . . . . . . . . . . . . . 76
  2.4.1. *The structure and carbohydrate recognition by langerin* . . . . . . . . . . . 77
  2.4.2. *Trimeric structure of langerin and its role in BG formation* . . . . . . . . . 79
2.5. The comparison of the CRDs of DC-SIGN and langerin . . . . . . . . . . . . . . . . 80

**3. HIV sexual transmission prevention strategy 83**

3.1. Microbicides as the means of HIV prevention . . . . . . . . . . . . . . . . . . . . . . . 83

3.2. The concept and design of glycomimetic drugs . . . . . . . . . . . . . . . . . . . . . . 86

3.3. Glycomimetics as DC-SIGN inhibitors . . . . . . . . . . . . . . . . . . . . . . . . . . . 88

3.4. Multivalent presentation of the active substances . . . . . . . . . . . . . . . . . . . . . 90

**4. The scope and objectives of this thesis 93**

**II. Materials & Methods 97**

**5. Materials 99**

5.1. Materials and equipment used for cloning . . . . . . . . . . . . . . . . . . . . . . . . . 99

5.2. Materials and equipment used for bacterial cultures . . . . . . . . . . . . . . . . . . . . 99

5.3. Materials and equipment used for protein purification and characterization . . . . . . 99

5.3.1. *SDS-PAGE* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 99

5.3.2. *Western Blot* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 100

5.3.3. *Protein purification* . . . . . . . . . . . . . . . . . . . . . . . . . 100

5.3.3.1. *Materials and equipement used for inclusion body extraction and refolding* . . . . . . . . . . . . . . . . . . . . . . . . . . . . 100

5.3.3.2. *Materials and equipement used for protein purification* . . . . . . . . 100

5.3.3.3. *Protein samples concentration* . . . . . . . . . . . . . . . . . . . . 100

5.3.4. *Surface Plasmon Resonance* . . . . . . . . . . . . . . . . . . . . . 100

5.3.5. *Equipement and materials for ITC, AUC, SLS/DLS and NMR experiments* . 100

**6. Methods 101**

6.1. Production of recombinant C-type lectin constructs . . . . . . . . . . . . . . . . . . . 101

6.1.1. *Transformation of $Ca^{2+}$-competent E.coli strains* . . . . . . . . . . . . . . . 101

6.1.2. *Production of DC-SIGN S-ECD, ECD, S-CRD and langerin ECD constructs* 101

6.1.2.1. *Over-expression and purification of DC-SIGN ECD* . . . . . . . . . 102

Over-expression. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102

Inclusion body extraction. . . . . . . . . . . . . . . . . . . . . . . . . 102

Refolding. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102

Purification. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 102

6.1.2.2. *Over-expression and purification of DC-SIGN S-ECD* . . . . . . . . 103

Over-expression. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103

Inclusion body extraction. . . . . . . . . . . . . . . . . . . . . . . . . 103

Refolding. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 103

Purification. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 104

6.1.2.3. *DC-SIGN S-CRD production* . . . . . . . . . . . . . . . . . . . . . 104

Over-expression. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 104

Inclusion body extraction and refolding. . . . . . . . . . . . . . . . . . 104

Purification. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 105

6.1.3. *Cloning of S-ECD and S-CRD constructs of LSECtin and MGL* . . . . . . . 105
    6.1.3.1. *Generation and amplification of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences* . . . . . . . . . . . . . . . . . . . . . . . 105
    6.1.3.2. *Insertion of LSECtin and MGL S-ECD or S-CRD coding cDNA sequences to pET30b vector* . . . . . . . . . . . . . . . . . . . . . . . 107
    6.1.3.3. *Insertion of LSECtin and MGL S-ECD or S-CRD coding cDNA sequences to pET20b vector* . . . . . . . . . . . . . . . . . . . . . . . 108
6.1.4. *Cloning of LSECtin S-CRD construct for targeting to periplasm* . . . . . . . . 108
    6.1.4.1. *Construction of pEt30b vector with encoded OmpA signal sequence* . 108
    6.1.4.2. *Generation of pEt30b-ompA/LSECtin S-CRD construct* . . . . . . . 109
        Preparation of LSECtin S-CRD cDNA sequences. . . . . . . . . . . . . 109
        Insertion of LSECtin S-CRD coding cDNA sequence to pEt30b-ompA vector. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 110
6.1.5. *Evaluation of LSECtin and MGL construct functionality* . . . . . . . . . . . . 110
    6.1.5.1. *The functionality of pEt20b and pEt30b constructs* . . . . . . . . . . 110
        Over-expression of proteins. . . . . . . . . . . . . . . . . . . . . . . . . 110
        Preparation of bacterial culture samples for SDS-PAGE. . . . . . . . . 111
    6.1.5.2. *The functionality of pEt30b-ompA/LSECtin S-CRD construct* . . . . 111
6.1.6. *Over-expression and purification of MGL S-ECD* . . . . . . . . . . . . . . . . 112
        Over-expression. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
        Inclusion body extraction. . . . . . . . . . . . . . . . . . . . . . . . . 112
        Refolding. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
        Purification. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
6.2. Biochemical/biophysical protein characterization and glycomimetic compound analysis assays . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 113
6.2.1. *SDS-PAGE of protein samples* . . . . . . . . . . . . . . . . . . . . . . . . . 113
6.2.2. *Protein concentration determination* . . . . . . . . . . . . . . . . . . . . . . 113
6.2.3. *Carbohydrate and glycomimetic compound preparation for SPR assays* . . . . 113
6.2.4. *The SPR competition assays* . . . . . . . . . . . . . . . . . . . . . . . . . . 113
    6.2.4.1. *Single point inhibition assay* . . . . . . . . . . . . . . . . . . . . . . 113
    6.2.4.2. *Competition assay for $IC_{50}$ determination* . . . . . . . . . . . . . . . 114
6.2.5. *The SPR direct interaction assay* . . . . . . . . . . . . . . . . . . . . . . . . 114
    6.2.5.1. *Preparation of oriented DC-SIGN surface* . . . . . . . . . . . . . . . 114
    6.2.5.2. *Establishment of regeneration conditions* . . . . . . . . . . . . . . . . 115
    6.2.5.3. *Compound interaction with oriented DC-SIGN surface* . . . . . . . . 115
6.2.6. *Other biophysical assays* . . . . . . . . . . . . . . . . . . . . . . . . . . . . 115

**III. Results** **117**

**7. The explanation of the organization of the results in this thesis.** **119**

**8. Results of preparation of recombinant lectins 123**
8.1. Preparation of different DC-SIGN constructs . . . . . . . . . . . . . . . . . . . . . . . . 123
8.1.1. *DC-SIGN ECD over-expression and purification results* . . . . . . . . . . . . 124
8.1.2. *DC-SIGN S-ECD over-expression and purification results* . . . . . . . . . . . 126
8.1.3. *DC-SIGN S-CRD refolding optimization results* . . . . . . . . . . . . . . . . 128
8.2. Generation of recombinant LSECtin and MGL S-ECD and S-CRD constructs . . . . 133
8.2.1. *General information about LSECtin and MGL* . . . . . . . . . . . . . . . . 133
LSECtin . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133
MGL . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 134
8.2.2. *Results of LSECtin and MGL S-ECD and S-CRD construct cloning into pEt20b and pEt30b vectors* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 136
8.2.2.1. *Generation of pEt20b and pEt30b constructs of LSECtin and MGL S-ECD and S-CRD* . . . . . . . . . . . . . . . . . . . . . . . . . . 136
8.2.2.2. *Results of LSECtin and MGL S-ECD and S-CRD pEt20b and pEt30b construct functionality evaluation* . . . . . . . . . . . . . . . . . . . 138
8.2.3. *Results of generation of LSECtin S-CRD construct to direct expression to periplasm* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 139
8.2.3.1. *The results of LSECtin S-CRD construct cloning* . . . . . . . . . . 139
8.2.3.2. *The results of pEt30b-ompA/LSECtin S-CRD construct functionality evaluation* . . . . . . . . . . . . . . . . . . . . . . . . . . . . 140
8.2.4. *MGL S-ECD over-expression and purification results* . . . . . . . . . . . . . 141
8.3. Concluding remarks on different lectin production . . . . . . . . . . . . . . . . . . . 145

**9. The results of the development of glycomimetic DC-SIGN antagonists 147**
9.1. Development of monovalent glycomimics . . . . . . . . . . . . . . . . . . . . . . . . . 147
9.1.1. *The SPR competition assay characteristics* . . . . . . . . . . . . . . . . . . . 147
9.1.1.1. *The experimental design of SPR competition assay* . . . . . . . . . 147
9.1.1.2. *The surface activity control during the experiment* . . . . . . . . . . 148
9.1.1.3. *The evaluation of the variation of $IC_{50}$ values* . . . . . . . . . . . . 149
9.1.1.4. *The set-up of the evaluation of compound selectivity to DC-SIGN versus langerin* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 150
9.1.2. *C-glycoside development* . . . . . . . . . . . . . . . . . . . . . . . . . . . . 153
9.1.3. *Development of fucose-based glycomimetic compounds* . . . . . . . . . . . . . 157
*Paper n$^o$ 1:* Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity *versus* Langerin . . . . . . 161
Additional studies. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 172
9.1.4. *Development of mannose-based glycomimetic compounds* . . . . . . . . . . . . 173
9.1.4.1. *Evaluation of tetravalent psTri to inhibit HIV infection in human cervicovaginal explants, and estimation of monovalent psTri selectivity to DC-SIGN versus langerin* . . . . . . . . . . . . . . . . . . . 174
*Paper n$^o$ 2:* A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models . . . . 177

9.1.4.2. *Evaluation of psDi and psTri as DC-SIGN antagonists presented in monovalent and multivalent manner using Ebola infection inhibition and SPR competition assays* . . . . . . . . . . . . . . . . . . . . . . 191

*Paper n<sup>o</sup>3:* Pseudosaccharide Functionalized Dendrimers as Potent Inhibitors of DC-SIGN Dependent Ebola Pseudotyped Viral Infection . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 195

9.1.4.3. *Structural characterization of psDi binding to DC-SIGN and evaluation of its selectivity to DC-SIGN versus langerin* . . . . . . . . . 209

*Paper n$^{o}$4:* Structure of a glycomimetic ligand in the Carbohydrate Recognition Domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design . . . . . . . . 213

9.1.4.4. *Structural characterization of psTri interaction with DC-SIGN; comparison of thermodynamic and hydrodynamic properties of psDi and psTri interaction with DC-SIGN* . . . . . . . . . . . . . . . . . . . . 241

*Paper n$^{o}$5:* Lectin clustering by a glycomimetic without any multivalent presentation: a case study in DC-SIGN antagonist development . . . . . . . . . . . . . . . . . . . . . . . . . . 245

9.1.4.5. *The bis-benzylamide derivatives of psDi to achieve better affinity and selectivity to DC-SIGN* . . . . . . . . . . . . . . . . . . . . . . . . 265

*Paper n$^{o}$6:* Selective targeting of DC-SIGN with mannose-based glycomimetics. Synthesis and interaction studies of bis-benzylamide derivatives of a pseudomannobioside . . . . . . . . . . . . . 271

The additional data not presented in the article. . . . . . . . . . . . . . 287

9.1.4.6. *The second type of psDi optimization to achieve better selectivity to DC-SIGN* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 290

9.2. The development of multivalent glycomimetic DC-SIGN antagonists . . . . . . . . . 293

9.2.1. *The structures of all tested multivalent compounds* . . . . . . . . . . . . . . . 293

9.2.2. *The results of multivalent compound analysis by SPR competition assay* . . . 297

9.2.2.1. *The evaluation of mannose-based compounds tethered to flexible multivalent scaffolds* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 297

9.2.2.2. *The evaluation of fucose-based compounds tethered to flexible multivalent scaffolds* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 300

9.2.2.3. *The evaluation of mannose-based compounds tethered to rod-like multivalent scaffolds* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 301

9.2.3. *The results of multivalent compound analysis by SPR direct interaction assay* 302

9.2.3.1. *The set-up of the direct interaction assay* . . . . . . . . . . . . . . 304

DC-SIGN surface preparation. . . . . . . . . . . . . . . . . . . . . . . . 304

DC-SIGN surface regeneration optimization and stability. . . . . . . . 306

9.2.3.2. *The studies of multivalent compound interaction with oriented DC-SIGN surface* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 308

The results of the first SPR direct interaction experiment. . . . . . . . 308

The results of the second SPR direct interaction experiment. . . . . . . 311

The results of the third SPR direct interaction experiment. . . . . . . . 315

IV. Discussion and future perspectives 325

10. The newly identified monovalent lead compounds as DC-SIGN antagonists. 327

11. The development of multivalent scaffolds. 331

12. Does the shape of the inhibition curve indicate "the mode of action" of the compound? 335

V. Bibliography 341

VI. Appendices 359

13. Supplementary information to paper n°1 361

14. Supplementary information to paper n°3 381

15. Supplementary information to paper n°4 385

16. Supplementary information to paper n°5 393

17. Supplementary information to paper n°6 397

18. The list of compounds and their names used in this thesis 415

19. The list of $IC_{50}$ values for all compounds 419

20. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by monovalent C-glycosides 429

21. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by *NV*-type compounds 433

22. All of the sensorgrams and inhibition curves for langerin inhibition by *NV*-type compounds 437

23. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by fucose-based monovalent and multivalent compounds 441

24. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h 445

25. All of the sensorgrams showing the interaction of 4h-functionalized dendrimers with DC-SIGN surfaces 453

26. The sensorgrams of double regeneration test of DC-SIGN surfaces 457

# List of Figures

1.1. The importance of the innate immunity. . . . . . . . . . . . . . . . . . . . . . . . . . 37
1.2. Different levels of innate immunity: an example of skin infection by *Staphylococcus aureus*. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
1.3. The scheme of hematopiesis. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40
1.4. Schematic representation of different groups of pattern recognition receptors with selected examples. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42
1.5. The structural diversity of CLR family. . . . . . . . . . . . . . . . . . . . . . . . . 46
1.6. A typical CRD structure (DC-SIGN CRD. . . . . . . . . . . . . . . . . . . . . . . . 51
1.7. Monosaccharide binding in Ca site 2. . . . . . . . . . . . . . . . . . . . . . . . . . 52
1.8. Colored scanning electron micrograph of a dendritic cell. . . . . . . . . . . . . . . 53
1.9. Antigen uptake, maturation and migration of dendritic cells to activate T cells. . . . 54
1.10. The cross-talk between TLR and CLR signaling shape the adaptive immunity. . . . . 55
1.11. Formation of immunological synapse. . . . . . . . . . . . . . . . . . . . . . . . . . 55

2.1. Schematic representation of HIV entry to the target cell. . . . . . . . . . . . . . . . 57
2.2. HIV-1 sexual transmission sites in women (A) and men (B). . . . . . . . . . . . . . 59
2.3. HIV invasion pathways in the female lower genital tract mucosa. . . . . . . . . . . 61
2.4. Model of HIV transfer from DCs to $CD4^+$ T cells. . . . . . . . . . . . . . . . . . . 62
2.5. The significance of DC–T-cell interactions and DC-SIGN implication in HIV-1 transmission. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
2.6. The micrograph showing an uptake of HIV-1 virions and their internalization to Birbeck granules by immature LCs. . . . . . . . . . . . . . . . . . . . . . . . . . . 64
2.7. Primary structure of human DC-SIGN. . . . . . . . . . . . . . . . . . . . . . . . . 66
2.8. Domain organization of DC-SIGN molecule. . . . . . . . . . . . . . . . . . . . . . . 67
2.9. Carbohydrate recognition by DC-SIGN CRD. . . . . . . . . . . . . . . . . . . . . . 68
2.10. Schematic representation of lectin/pathogen interaction affinity increase due to the avidity phenomenon. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 69
2.11. The structure of the extracellular part of DC-SIGN. . . . . . . . . . . . . . . . . . 70
2.12. The nanostructure of DC-SIGN microdomains. . . . . . . . . . . . . . . . . . . . . 71
2.13. Possible mechanisms of DC-SIGN microdomain formation and stability. . . . . . . 72
2.14. The physiological role of DC-SIGN and endogeneous ligands. . . . . . . . . . . . . 73
2.15. TLR signaling modulation by DC-SIGN triggering induced by mannosylated ligands bearing pathogens such as HIV-1 or *Mycobacteria*. . . . . . . . . . . . . . . . . . . 74
2.16. Electron microscopy of a Langerhans cell from human epidermis. . . . . . . . . . . 76
2.17. Primary structure of human langerin. . . . . . . . . . . . . . . . . . . . . . . . . . 77
2.18. The structure of langerin CRD. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78

2.19. The CRD of langerin in complex with carbohydrate ligands. . . . . . . . . . . . . . . 78
2.20. The trimeric organization of truncated langerin. . . . . . . . . . . . . . . . . . . . . . . . 79
2.21. Contribution of langerin to Birbeck granule formation. . . . . . . . . . . . . . . . . . . 80
2.22. The comparison of DC-SIGN and langerin CRDs. . . . . . . . . . . . . . . . . . . . . . 81
2.23. The differences of sugar-binding sites of DC-SIGN and langerin. . . . . . . . . . . . . 81
2.24. The micrograph of mucosal tissue. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82

3.1. The examples of the approved carbohydrate-based and glycomimetic drugs. . . . . . 86
3.2. Examples of non-carbohydrate ligands of DC-SIGN. . . . . . . . . . . . . . . . . . . . 89
3.3. The glycan shield of gp120. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 90
3.4. The structures of multivalent presentation platforms. . . . . . . . . . . . . . . . . . . 91

4.1. The principal architecture of ITN CARMUSYS and our group involvement. . . . . . 94

6.1. Multiple cloning site of pASK-IBA6. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 108

7.1. A schematic representation of the organization of the results in this thesis. . . . . . . 119

8.1. The DC-SIGN constructs used in this thesis. . . . . . . . . . . . . . . . . . . . . . . . 123
8.2. The SDS-PAGE analysis of DC-SIGN ECD expression. . . . . . . . . . . . . . . . . . . 124
8.3. The chromatograms of DC-SIGN ECD purification. . . . . . . . . . . . . . . . . . . . 124
8.4. The SDS-PAGE analysis of DC-SIGN ECD purification. . . . . . . . . . . . . . . . . . 125
8.5. The expression of DC-SIGN S-ECD. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 126
8.6. The chromatograms of DC-SIGN S-ECD purification. . . . . . . . . . . . . . . . . . . 126
8.7. The SDS-PAGE analysis of DC-SIGN S-ECD purification. . . . . . . . . . . . . . . . 127
8.8. The chromatograms showing elution profiles of DC-SIGN S-CRD purification on Mannan-Agarose column. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 128
8.9. The SDS-PAGE analysis of DC-SIGN S-CRD production. . . . . . . . . . . . . . . . . 129
8.10. The absorbance spectra of DC-SIGN S-CRD refolding samples. . . . . . . . . . . . . 129
8.11. The SDS-PAGE analysis of different purification stages of DC-SIGN S-CRD. . . . . . 130
8.12. The chromatograms of DC-SIGN S-CRD purification. . . . . . . . . . . . . . . . . . 130
8.13. The SDS-PAGE analysis of DC-SIGN S-CRD purification on StrepTactin column in the absence of glutathiones. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 131
8.14. The SDS-PAGE analysis of DC-SIGN S-CRD purification on Mannan-Agarose column.131
8.15. The structure of LSECtin. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133
8.16. ClustalW sequence alignment of canonical MGL and isoform 2. . . . . . . . . . . . . 134
8.17. The strategy of LSECtin and MGL S-ECD and S-CRD cloning. . . . . . . . . . . . . 136
8.18. The results of generation of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences by PCR. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137
8.19. The results of insertion of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences to pEt30b vector. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137
8.20. The results of insertion of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences to pEt20b vector. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137
8.21. The results of LSECtin and MGL S-ECD and S-CRD production by pEt30b constructs.138

8.22. The results of LSECtin and MGL S-ECD and S-CRD production by pEt20b constructs. 138
8.23. The results of design of pEt30b vector containing OmpA signal peptide coding sequence. 139
8.24. The PCR results of LSECtin S-CRD coding cDNA that included ompA-annealing sequences (455 bp) amplification. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 140
8.25. The functionality of pEt30b-ompA/LSECtin S-CRD construct to express LSECtin S-CRD. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 140
8.26. The SDS-PAGE analysis of MGL S-ECD expression. . . . . . . . . . . . . . . . . . . . 141
8.27. The chromatogram of MGL S-ECD elution by desthiobiotin from StrepTactin column. 141
8.28. The chromatograms of MGL S-ECD injection to 125 mL Superose12 column. . . . . 142
8.29. The chromatogram of MGL S-ECD elution profile on immobilized galactose column. 143
8.30. The SDS-PAGE analysis of MGL S-ECD purification steps. . . . . . . . . . . . . . 143

9.1. The SPR competition assay. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 148
9.2. The ManBSA surface titration with DC-SIGN ECD. . . . . . . . . . . . . . . . . . . 149
9.3. The raw sensorgams representing titration of difference surfaces with langerin ECD. . 150
9.4. The sensorgrams comparing langerin interaction with ManBSA and reference surfaces. 151
9.5. The langerin inhibition dependence on ManBSA surface density. . . . . . . . . . . . 151
9.6. The structures of C-glycosides. . . . . . . . . . . . . . . . . . . . . . . . . . . 153
9.7. DC-SIGN inhibition activities of D-mannose and L-fucose C-glycosidic monosaccharides compared to their natural counterparts. . . . . . . . . . . . . . . . . . . . . 154
9.8. DC-SIGN inhibition activities of D-mannose C-glycosidic pseudo-disaccharides compared to D-mannose, Manα1-2Man and Manα1-3Man. . . . . . . . . . . . . . . . . . 154
9.9. DC-SIGN inhibition activities of L-fucose C-glycosidic pseudo-disaccharides compared to L-fucose and Lewis$^X$. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 155
9.10. Sensorgrams illustrating the effect of compound ZL2(L) for DC-SIGN ECD binding to CM-dextran and ManBSA surfaces. . . . . . . . . . . . . . . . . . . . . . . . . 156
9.11. The structures of previously designed fucose-based compounds ***2a*** and ***17***. . . . . . 157
9.12. The structures of newly synthesized fucose-based compounds. . . . . . . . . . . . . 157
9.13. The structures and activities of compound **10b** and its derivatives. . . . . . . . . . 172
9.14. The structures of previously designed mannose-based glycomimics psDi and psTri. . 173
9.15. The structures of the scaffolds of tetravalent dendron and multivalent Boltorn type dendrimer. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 191
9.16. psDi binding within DC-SIGN CRD. . . . . . . . . . . . . . . . . . . . . . . . . 265
9.17. The structure of new psDi-based compound series. . . . . . . . . . . . . . . . . . 265
9.18. Schematic explanation of 1D STD NMR experiment. . . . . . . . . . . . . . . . . . 266
9.19. An illustration of the determination of a binding epitope of a small molecule. . . . . 267
9.20. The comparison of the selectivity gains of bis-benzylamides and psTri with respect to psDi as a reference compound. . . . . . . . . . . . . . . . . . . . . . . . . . . . 287
9.21. The structure of compound **11**. . . . . . . . . . . . . . . . . . . . . . . . . . . 287
9.22. 1D $^1$H NMR spectra of **4h** and psDi comparing the signals of the key protons in the two molecules. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 288
9.23. 1D $^1$H NMR spectra of **4h** at 10°C and 25°C. . . . . . . . . . . . . . . . . . . . 289
9.24. 1D 1H NMR spectra of **4h** in presence of different DMSO concentrations, at 25°C. . 289

9.25. The structures of psDi derivatives with different substituents at position 6 of non-reducing mannose compared to **4h** structure. . . . . . . . . . . . . . . . . . . . . . . 290
9.26. The activities of new psDi derivatives to inhibit DC-SIGN and langerin. . . . . . . . 290
9.27. The comparison of the selectivity gains with respect to psDi as a reference compound. 291
9.28. The principal scheme of CuAAC reaction. . . . . . . . . . . . . . . . . . . . . . . . 293
9.29. The structures of different multivalent flexible scaffolds generated by CuAAC reactions. 294
9.30. The structure of rod-like scaffold and the control compounds ***R1*** and ***R2***. . . . . . . 295
9.31. DC-SIGN inhibition activities of multivalent manoose-based compounds on flexible scaffolds. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 298
9.32. DC-SIGN inhibition activity improvement factors for the multivalent compounds on flexible scaffolds with respect to corresponding monovalent ligands. . . . . . . . . . . 299
9.33. The structures of L-fucose with β-alanine/ethylene glycol linkers. . . . . . . . . . . . 300
9.34. DC-SIGN inhibition activities of multivalent fucose-based compounds on flexible scaffolds. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 300
9.35. DC-SIGN inhibition activities of different hexavalent scaffolds with **4h**. . . . . . . . . 302
9.36. The structures of psDi and psTri functionalized Boltorn type dendrons and dendrimers that were analyzed in SPR direct interaction assay. . . . . . . . . . . . . . 303
9.37. A schematic illustration of the SPR direct interaction assay. . . . . . . . . . . . . . . 304
9.38. DC-SIGN S-ECD affinity to StrepTactin. . . . . . . . . . . . . . . . . . . . . . . . . 305
9.39. The activity of oriented DC-SIGN surface. . . . . . . . . . . . . . . . . . . . . . . . 305
9.40. The sensorgrams showing the immobilization of DC-SIGN constructs on reactivated dextran/StrepTactin surface. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 306
9.41. Reference surface corrected overlaid sensorgrams showing the inability to regenerate DC-SIGN surface with EDTA. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 307
9.42. An example of reference surface corrected overlaid sensorgrams showing improved compound detachment from DC-SIGN surface. . . . . . . . . . . . . . . . . . . . . . 307
9.43. The evaluation of DC-SIGN surface stability. . . . . . . . . . . . . . . . . . . . . . . 308
9.44. The interaction of tetravalent dendrons with oriented DC-SIGN surf ace. . . . . . . . 309
9.45. The interaction of G3(pseudodi)$_{32}$ and G3(pseudotri)$_{32}$ dendrimers with oriented DC-SIGN surface. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 309
9.46. Reference surface corrected sensorgrams showing the binding of flexible dendrimers with mannose-based compounds to DC-SIGN surface. . . . . . . . . . . . . . . . . 312
9.47. DC-SIGN surface titration with flexible dendrimers. . . . . . . . . . . . . . . . . . . 313
9.48. The example sensorgrams showing a stronger interaction of the flexible dendrimer with StrepTactin than DC-SIGN surface. . . . . . . . . . . . . . . . . . . . . . . . . 314
9.49. The experimental set-up of SPR direct interaction assay to determine if the compound binds to 2 CRDs within the same DC-SIGN tetramer. . . . . . . . . . . . . . . . . 315
9.50. The example sensorgram DC-SIGN S-ECD covalent capture followed by surface rinsing to remove non-covalently bound protein. . . . . . . . . . . . . . . . . . . . . . . 316
9.51. The results of mass transfer test for different density DC-SIGN S-ECD surfaces. . . . 317
9.52. The interaction of ManBSA with DC-SIGN surfaces of different densities. . . . . . . 318

9.53. The results of the interaction of **4h** functionalized dendrimers with DC-SIGN surfaces of different density. . . . . . . 319
9.54. The results of the interaction of **4h** functionalized rod-like and related dendrimers with DC-SIGN surfaces of different density. . . . . . . 320
9.55. The results of the interaction of tetravalent fucosylated dendrimer with DC-SIGN surfaces of different density. . . . . . . 320
9.56. A relative comparison of $EC_{50}$ values obtained for **4h** functionalized series of dendrimers. . . . . . . 322
9.57. The results of the test of DC-SIGN S-ECD surface regeneration after multivalent compound binding. . . . . . . 323

10.1. The new identified glycomimetic leads. . . . . . . 327
10.2. A perspective mannose-based compound with combined features of the lead **4h** and compound NV243. . . . . . . 328
10.3. The anticipated structural basis for the improvement of the selectivity to DC-SIGN *vs* langerin by the perspective mannose-based compound. . . . . . . 329
10.4. The comparison of DC-SIGN inhibition activities of the corresponding multivalent compounds with C-glycosidic L-fucose and C-glycosidic D-mannose. . . . . . . 330

11.1. The comparison of DC-SIGN inhibition activities of the corresponding multivalent compounds with O-glycosidic L-fucose and D-mannose. . . . . . . 331
11.2. The comparison of DC-SIGN inhibition activities of all glycomimetic compounds on tetravalent dendrimer scaffolds. . . . . . . 332
11.3. The comparison of DC-SIGN inhibition activities of tetravalent psDi and psTri dendrimers and dendrons. . . . . . . 332

12.1. The comparison of three inhibition curves with different slopes, for which $IC_{50}$ values are the same. . . . . . . 335
12.2. Possible mechanisms of how compounds inhibit DC-SIGN in the SPR competition assay. . . . . . . 336
12.3. The analysis of all the tested compounds in terms of the slope of the inhibition curve. 337

# List of Tables

1.1. Examples of selected PRRs and their ligands. . . . . . . . . . . . . . . . . . . . . . . . 43
1.1. Examples of selected PRRs and their ligands. . . . . . . . . . . . . . . . . . . . . . . . 44
1.1. Examples of selected PRRs and their ligands. . . . . . . . . . . . . . . . . . . . . . . . 45
1.2. Selected examples of CLRs that function as PRRs. . . . . . . . . . . . . . . . . . . . . 47
1.2. Selected examples of CLRs that function as PRRs. . . . . . . . . . . . . . . . . . . . . 48
1.2. Selected examples of CLRs that function as PRRs. . . . . . . . . . . . . . . . . . . . . 49

2.1. Contribution of mucosal HIV invasion sites to global HIV infections in adults. . . . . 60

3.1. The examples of different type vaginal microbicides. . . . . . . . . . . . . . . . . . . 84
3.2. Some examples of the published monovalent mannose-based DC-SIGN glycomimetics. 88
3.3. Some examples of the published glycomimetic fucose-based DC-SIGN ligands. . . . . 89
3.4. Some multivalent DC-SIGN inhibitors and their potencies. . . . . . . . . . . . . . . . 92

6.1. Preparation of primers for generation of lectin constructs. . . . . . . . . . . . . . . . 106
6.2. Preparation of LSECtin and MGL S-ECD or S-CRD cDNA sequences and pET30b digestion mixtures. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 107
6.3. Preparation of NdeI/XhoI digestion mixtures for the test of positive cDNA insertion. 107
6.4. Preparation of primers for LSECtin S-CRD cDNA sequence amplification. . . . . . . 109

7.1. The list of papers presented in this thesis. . . . . . . . . . . . . . . . . . . . . . . . . 120

9.1. Evaluation of variation of $IC_{50}$ values. . . . . . . . . . . . . . . . . . . . . . . . . . . 149
9.2. Mannose-based ligands conjugated to the multivalent scaffolds and the corresponding names of the compounds. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 296
9.3. The list of fucose-based ligands conjugated to multivalent scaffolds and the corresponding names of the compounds. . . . . . . . . . . . . . . . . . . . . . . . . . . . 297
9.4. The $IC_{50}$ ± SD values (µM) obtained for multivalent mannose-based compounds on flexible scaffolds. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 298
9.5. The $IC_{50}$ ± SD values (µM) obtained for multivalent fucose-based compounds on flexible scaffolds. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 301
9.6. The $IC_{50}$ obtained for rod-like and other hexavalent scaffolds. . . . . . . . . . . . . . 302
9.7. The comparison of calculated $R_{\mathrm{max}}$ values with the experimentally observed maximal binding responses of Boltorn type dendrimers and dendrons functionalized with psDi or psTri. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 310
9.8. The calculated apparent $K_{\mathrm{D}}$ values for psDi and psTri functionalized $3^{\mathrm{rd}}$ generation Boltorn type dendrimers. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 311

9.9. The comparison of calculated $R_{\text{max}}$ values with the experimentally observed maximal binding responses of the flexible dendrimers. . . . . . . . . . . . . . . . . . . . . . . . . 313
9.10. The comparison of calculated $R_{\text{max}}$ values with the experimentally observed maximal binding responses for the tested dendrimers. . . . . . . . . . . . . . . . . . . . . . . . . 321
9.11. The list of multivalent compounds, which were tested for ability to be removed from DC-SIGN surface by EDTA. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 323

12.1. The determination of upper and lower A2 limits for the “normal” compound. . . . . 336
12.2. The summary of all observed and assumed characteristics of all multivalent compounds. 338

# Les résumés en Français

## Introduction

L'objet central de ce travail de recherche est DC-SIGN (Dendritic Cell Specific ICAM-3 Grabbing Non-integrin) un récepteur lectine de type C exprimée sur les cellules dendritiques et impliqués dans les processus infectieux de nombreux agents pathogènes, dont le virus de l'immunodéficience humaine (VIH). Les cellules dendritiques (DC) sont un élément très important de l'immunité innée. Elles ont un rôle de sentinelles au sein des tissus périphériques pour détecter les agents pathogènes envahisseurs, les capturer et présenter des antigènes dérivés de ceux-ci aux lymphocytes T pour l'induction de réponses immunitaires adaptatives.

Bien que DC-SIGN soit exploitée par de nombreux agents pathogènes comme le virus Ebola, la dengue et d'autres virus, mycobactéries de la tuberculose et d'autres pathogènes dangereux, ce travail de thèse se concentre sur les étapes de l'infection à VIH dépendantes de DC-SIGN.

Compte tenu de ces enjeux, l'introduction de cette thèse est divisée en trois chapitres, dont chacun décrit brièvement ces points : l'immunité innée, la transmission sexuelle du VIH et la prévention de la transmission sexuelle du VIH, tandis que le quatrième chapitre présente la portée et les objectifs de ce travail de recherche.

### Chapitre 1. Un aperçu de l'immunité innée de l'homme

L'immunité innée a un rôle primordial, car c'est la première ligne de défense de l'organisme contre les agents pathogènes et elle active dans un second temps l'immunité adaptative. Ce chapitre décrit brièvement les éléments les plus importants de l'immunité innée, qu'ils soient d'ordre anatomique, chimique, ou physiologique (température, pH, facteurs solubles). Il décrira également les acteurs cellulaires impliqués.

La peau et les muqueuses représentent un obstacle physique pour l'invasion des pathogènes. Elles sont également protégées par des substances antimicrobiennes sécrétées par les kératinocytes ou des cellules épithéliales spécialisées, et ces substances comprennent les mucines, des enzymes et des peptides antimicrobiens. En outre, la peau et les muqueuses sont habités par des microorganismes non pathogènes, de la flore normale ou microbiote commensal, qui sont en concurrence avec les pathogènes pour les nutriments et les sites de fixation et, de cette façon, participent à la défense de l'organisme. La température, plus basse à la surface de la peau, le pH acide de la peau et des muqueuses et d'autres facteurs solubles constituent une barrière chimique physiologique.

Les cellules de l'immunité innée fournissent une défense plus spécifique contre les infections. Ces cellules, chacune ayant un rôle spécifique, 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 par certaines cytokines, à partir de précurseurs cellulaire commun de la moelle osseuse, les cellule souche hématopoïétique (CSH).

Les CDs sont constitués d'un groupe diversifié de cellules présentatrices d'antigène profes- sionnelles (APC). Cette diversité résultent d'une combinaison de plusieurs facteurs tels que leurs origine, leur état de différenciation, et leur localisation tissulaire. Les CDs sont particulièrement importantes car elles constituent un lien entre l'immunité innée et adaptative - leur rôle principal est d'induire l'immunité adaptative. Les CDs immatures passent en revue les tissus périphériques et y capturent les micro-organismes étrangers. Cela conduit à la présentation de l'antigène dérivée des pathogènes à leur surface conjointement à la migration vers les organes lymphoïdes où les CDs rencontreront des cellules T naïves et les activer.

L'identification des micro-organismes étrangers par l'immunité innée est basée sur la reconnais- sance des structures moléculaires conservés propres aux micro-organismes et absentes chez l'hôte. 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 exercent leur fonction par la phagocytose des pathogènes, l'activation des voies de signalisation pro-inflammatoires, l'opsonisation, l'activation de cascades du complément et de la coagulation, et l'induction de l'apoptose des celluels infectées. Les PRRs peuvent être solubles, exprimées à la surface cellulaire ou intra-cellulaire. Les récepteurs lectine de type-C (CLRs) sont l'une des familles de PRsR. Les CLRs reconnaissent des groupements glucidiques spécifiques présent à la surface de ses cibles. Cette reconnaissance est dépendantes d'ions $Ca^{2+}$ présents dans le site actif.

## Chapitre 2. Virus de l'immunodéficience humaine (VIH): l'infection par transmission sexuelle

La transmission sexuelle par le biais des muqueuses génitales et rectales est responsable de la grande majorité des infections à VIH dans le monde. Les différentes voies d'invasion du VIH dans les muqueuses et le mécanisme de sa transmission puis de l'infection de sa cible cellulaire majeure, les cellules T CD4 +, sont décrite dans ce chapitre.

L'infection à VIH par transmission sexuelle commence au niveau des muqueuses génitale (ou rectal) où les virions du VIH détournent les CDs stromales, comme des chevaux de Troie, pour at- teindre les lymphocytes T CD4+. Cependant, un autre sous-type de CDs, les cellules de Langerhans (CL) résidant dans la couche supérieure de la muqueuse épithéliale, a été découvert pour supprimer l'infection par le VIH.

Il a été démontré que les récepteurs, de type CLR, impliqués dans la capture du VIH, que soit par les CDs stromales du derme ou les cellules de langerhans de l'épiderme, sont étroitement liés (DC-SIGN et langerine, respectivement). Mais la capture du VIH par DC-SIGN résulte dans une promotion de l'infection, alors que la reconnaissance par langérine conduit à une protection vis à vis de l'infection (élimination de virus). Ces deux récepteurs interagissent avec le virus au travers des nombreuses glycosylations présentes à sa surface sur la protéine d'enveloppe gp120. Une description détaillée de la structure et de la fonction de ces deux CLRs, ainsi que leur comparaison, est présentée dans ce chapitre.

## Chapitre 3. VIH stratégie prévention de la transmission sexuelle

Plus de 40 millions de personnes dans le monde sont infectées par le VIH. Malheureusement, il n'existe pas de remède pour l'infection à VIH à ce jour et le traitement de long-terme disponible, le HAART (traitement antirétroviral hautement actif, il combine plusieurs substances qui interfèrent à différentes étapes du cycle viral), est assez cher et accessible que pour une faible population des personnes infectées. Ainsi, le développement de moyens de prévention de l'infection est d'une importance capitale, car ce serait le moyen le plus efficace pour enrayer l'épidémie.

Bien que la transmission sexuelle par la muqueuse génitale est la voie dominante de l'infection pour les deux sexes, les femmes semblent être plus sensibles à ce type d'invasion virale dans les tissus vaginaux et utérins. Les microbicides vaginaux anti-VIH (VHMs), des substances destinées à bloquer les événements précoces de l'infection à VIH, sont une alternative prometteuse pour une prévention saine et efficace de la transmission sexuelle du VIH. De plus, les VHMs représentent un nouvel outil de prévention particulièrement attractif pour les femmes. En effet, ils permettraient une autonomie de celles-ci dans leur protection. En effet, dans les régions à haut risque de certains pays en développement les femmes ne sont pas toujours en mesure de réduire leurs risques de contamination, en contraignant leur partenaires, pour des raisons sociales, culturelles et économiques. Les types, le mode d'action et des exemples de VHMs sont détaillées dans ce chapitre.

Dans la recherche de nouveaux modes de prévention du VIH, DC-SIGN a été identifié comme une cible thérapeutique potentielle en raison de son rôle important dans la potentialisation de l'infection à VIH et la diffusion dans les muqueuses. Puisque DC-SIGN est un récepteur lectine qui reconnaît des monosaccharides mannose et fucose dans les structures oligosaccharidiques cibles, la stratégie de synthèses de glycomimétiques correspondant pourrait être utilisée pour concevoir des antagonistes de DC-SIGN.

Les glycomimétiques peuvent être définis comme des entités chimiques qui imitent les glucides bioactifs et se lient aux protéines cibles, principalement des lectines ou des enzymes modifiant des sucres, interférant ainsi avec le processus ciblés. Les substances médicamenteuses glycomimétique présentent plusieurs avantages par rapport aux ligands glucidiques naturels: elles peuvent être conçues pour avoir une plus grande affinité et spécificité pour la protéine cible ainsi qu'une demi-vie plasmatique plus longue (au travers d'une résistance renforcée aux glycosidases). Cette stratégie de conception de glycomimétique et les exemples de réussite sont décrits dans ce chapitre.

DC-SIGN reconnaît des monosaccharides de type mannose et fucose. Ainsi la principale stratégie de développement de glycomimétique spécifique de DC-SIGN est de conserver le mannose ou le fucose comme ancre de reconnaissance primaire, tandis que d'autres parties de la molécule peut être conçu et optimisé pour atteindre la meilleure affinité, et apporter de la spécificité. Les travaux préalables disponibles à propos des tentatives de production de ligands dirigées contre DC-SIGN sont détaillées dans ce chapitre.

La conception de ligands monovalents pour DC-SIGN est la première étape dans le développe- ment des antagonistes. Cependant, dans les systèmes biologiques, les interactions entre DC-SIGN avec ses ligands, en particulier les PAMPs comme la glycoprotéine d'enveloppe gp120 du VIH, sont de nature multivalente. Cela conduit, au travers d'un phénomène d'avidité, à une affinité globale bien supérieur à l'interaction primaire d'un sucre avec un site actif de DC-SIGN. Par conséquent, les ligand monovalent glycomimétiques de DC-SIGN doivent être au final présentées en plusieurs exemplaires sur une plate-forme moléculaire donnée, afin d'obtenir une substance qui est capable de rivaliser, par multivalence, avec les interactions DC-SIGN/pathogen. Les différentes stratégies de développement de telles plateformes multivalentes et des exemples préalables, dirigés contre DC-SIGN, sont décrits dans ce chapitre.

## Chapitre 4. La portée et les objectifs de cette thèse

L'objectif principal de cette thèse est de développer des antagonistes sélectifs de DC-SIGN, et sa réalisation implique les tâches suivantes:

- ✧ Criblage de ligand monovalent par test de compétition SPR pour:
  - identifier les meilleurs ligands de DC-SIGN;
  - évaluer le rapport de sélectivité pour DC-SIGN vis-à-vis de la langerine.
- ✧ Caractérisation structurale des modes d'interactions avec DC-SIGN des composés sélectionnés par cristallographie aux rayons X.
- ✧ Caractérisation hydrodynamique/thermodynamique de l'interaction de DC-SIGN avec les composés d'intérêt.
- ✧ Criblage des composés multivalent par des analyses SPR:
  - mise au point/optimisation des tests;
  - détermination du pouvoir inhibiteur des dendrimers vis à vis de DC-SIGN (test de com- pétition);
  - l'évaluation de l'interaction directe des composés avec DC-SIGN.

En plus de ces tâches, une partie de mon travail a été la surexpression et la purification des constructions ECD et CRD de DC-SIGN, nécessaires pour les approches SPR et les tests et co-cristallisation mais également pour fournir nos partenaires en protéines pour leurs propres expériences.

En outre, la caractérisation de la sélectivité des composés étant une des préoccupations de cette thèse, une autre tâche a été le développement de la production d'autres lectines de type C, à savoir la LSECtin et MGL.

## Chapitre 5. Matériels

Tous les produits chimiques, biologiques et les équipements utilisés dans ce travail sont détaillés dans ce chapitre.

## Chapitre 6. Méthodes

Ce chapitre décrit en détail toutes les procédures expérimentales utilisées dans ce travail de recherche.

## Chapitre 7. L'explication de l'organisation des résultats dans cette thèse

Ce chapitre décrit l'organisation du travail au sein de cette thèse et dans le contexte du réseau européen (résumé dans la figure 7.1).

## Chapitre 8. Préparation de lectines recombinantes

Ce chapitre contient les résultats de la prorduction de différentes lectines. Pour commencer, la production de différentes constructions DC-SIGN, qui ont été utilisés pour l'évaluation composé glycomimétique structurelle et fonctionnelle, sont présentés dans la section 8.1.

La production de DC-SIGN ECD a été réalisé sans obstacle majeur, et des rendements con- sidérables de protéines purifiées fonctionnelles ont été obtenus (habituellement 50-80 mg de 1 L de culture cellulaire). Étonnamment, l'ajout de StrepTag II à extrémité N-terminale de DC-SIGN ECD entraîne une nettement diminution du rendement de protéine purifiée principalement en raison de fortes précipitations pendant les étapes de repliement. Néanmoins, seules de faibles quantités de cette construction sont été nécessaires pour la préparation de la surface orientée pour les dosage en interaction directe par SPR.

Les efforts d'optimisation des rendements de production de DC-SIGN S-CRD n'ont pas aboutis. Cependant, la raison de la faiblesse des rendements a été révélé: lors des étapes de concentrations, la protéine forme des ponts disulfures non fonctionnels.

En dehors de DC-SIGN, différentes constructions de deux autres lectines de type C, à savoir LSECtin et MGL, ont été réalisé dans le but d'étendre les outils d'analyse de la sélectivité des composés. Les résultats du clonage et de l'évaluation de l'expression des divers des constructions sont présentées dans la section 8.2.

Quatre nouvelles constructions lectine de type C ont été développés pour l'utilisation future dans le laboratoire: LSECtin S-CRD et S-ECD ainsi que MGL S-CRD et S-ECD. A l'exception de LSECtin S-CRD, toutes les constructions se sont révélées capables d'exprimer efficacement les protéines correspondantes. Néanmoins, LSECtin S-CRD construction est également produit mais le niveau d'expression est plus faible.

La production de MGL S-ECD a été étudiée un peu plus loin, et les premiers essais de purifi- cation de cette construction ont été réalisés. Bien que des résultats prometteurs ont été obtenus, la procédure doit être optimisé.

## Chapitre 9. Les résultats de la mise au point de glycomimétique antagonistes de DC-SIGN.

Ce chapitre décrit la mise au point de composés glycomimétique, et est également divisée en deux sous-parties principales, dont l'une décrit le développement de glycomimétiques monovalents, et dans la seconde les résultats de la mise au point de composés multivalents sont présentés.

Le développement de composé monovalent a été menée en trois axe parallèles:

- ✧ L'équipe du Pr. Jitka Moravcová a développé les C-glycosides et les résultats de leur évaluation sont présentés au paragraphe 9.1.2.
- ✧ L'équipe du Pr. Anna équipe de Bernardi a travaillé sur deux concepts, à savoir des glycomimétiques à base de fucose base et à base de mannose, et la plupart des résultats de leur développement sont présentés sous la forme d'articles répertoriés dans le tableau 7.1.

Les résultats du développement de glycomimétiques à base de fucose sont présentés dans l'article **n°1**. Le troisième axe de développement de composé glycomimétique, à base de mannose, était plus élaboré. Il commence par la caractérisation fonctionnelle de deux composés glycomimétique à base de mannose, pseudomannobioside (psDi) et pseudomannotrioside (psTri). De ces deux composés, psTri s'est avéré avoir une pouvoir inhibiteur bien meilleur que le psDi. Il a été donc été utilisé pour préparer un dendron tétravalent et testés pour sa capacité à inhiber l'infection en trans des lymphocytes T au VIH. Plus tard, les capacités de ce composé à inhiber la transmission du VIH dans des explants du col de l'utérus ont pu également être étudiée et les résultats sont présentés dans l'article **n°2**.

Les dendrimères 32-valents de psDi et psTri ont également été produit et testé en comparant les activités avec les psDi et psTri mono-et tétravalents dans des tests de compétition par SPR et des test d'inhibition de l'infection par le virus Ebola. Les résultats sont présentés dans l'article **n°3**. Ces études ont indiqués des divergences entre les caractéristiques structurelles et fonctionnelles de psTri. Ainsi, les formes monovalentes de psDi et psTri ont une différence d'un ordre de grandeur dans leurs activités en faveur de psTri, mais cet avantage est perdu lorsque l'on comparent les composés multivalents. Ainsi, l'interaction des deux composés monovalents, psDi et psTri, a été caractérisé en détail: les structures aux rayons X de ces deux composés en complexe avec DC-SIGN CRD ont été résolus, les propriétés structurales de l'interaction en solution a été étudiée par RMN, et les complexes ont été caractérisés par d'autres techniques biophysiques (ITC, ASC, DLS). Les résultats de ces études sont présentés dans deux documents, les article **n°4** et **n°5** (encore au stade de manuscript).

Au final, ces études ont disqualifié le psTri, mais ont révélé que le psDi possédait toutes les qualités souhaitables pour un composé candidat au développement. Ainsi, le développement de glycomimétique à base de mannose s'est poursuivi en mettant l'accent sur l'optimisation du psDi. Deux série de composés, de types différents, dérivé du psDi ont été synthétisés et évalués pour l'inhibition de DC-SIGN et leur sélectivité par rapport à la langérine. Les résultats de l'évaluation de l'une des deux séries (bis-benzylamides) sont présentés dans le document **n°6**, et l'analyse des autres séries (NV-type) est décrite dans le sous-paragraphe 9.1.4.6.

Parallèlement à l'élaboration de glycomimétiques monovalents, différents échafaudages polyvalents ont été conçues et synthétisées par l'équipe du Dr Javier Rojo. Ensuite, les glycomimétiques monovalents choisis précédemment ont été attachés à ces échafaudages polyvalents et testé comme antagonistes de DC-SIGN. Les résultats de l'évaluation des composé polyvalents correspondant par SPR, tests par compétition ou et d'interaction directes, sont présentés aux paragraphes 9.2.2. et 9.2.3., respectivement.

## Chapitre 10. Les composés monovalents nouvellement identifiés antagonistes de DC-SIGN

Les résultats pour l'ensemble des composés testés monovalents glycomimétiques sont résumés et comparés dans ce chapitre. Une proposition d'un composé optimal est proposé sur la base des résultats obtenus.

## Chapitre 11. Le développement des échafaudages multivalents

Les résultats pour l'ensemble des composés multivalents conçus sont résumés et comparés dans ce chapitre.

## Chapitre 12. Est-ce que la forme de la courbe d'inhibition indique "le mode d'action" du composé?

Une hypothèse sur le sens de la pente de la courbe d'inhibition est proposée, appliquée à l'ensemble des composés testés et discutée. Les moyens pour prouver cette hypothèse sont suggérés.

# Part I.

# Introduction

# 1. An overview of human innate immunity

## 1.1. The components of innate immunity

The immune system has evolved to an amazing defense mechanism against the invasion of a broad variety of pathogens and parasites. It is possessed by the majority of the living organisms including such simple beings as unicellular bacteria or invertebrates, where it exists in a rudimental form [1, 2]. In higher vertebrates, including humans, it has evolved to an extremely sophisticated system that has the features of specificity and memory, and which is implemented by the two components – innate and adaptive immunities. The adaptive immunity is highly specific with a unique property of memory and it requires stimulation with an antigen prior to the response, whereas, innate immunity is less specific and has no memory. However, innate immunity has a crucial role (fig. 1.1), as it is the first line defense of the organism against the invading pathogens and its responses activate the adaptive immunity. The importance of innate immunity is also highlighted by the fact that the defects of its components are relatively rare and almost always lethal [3, 4].

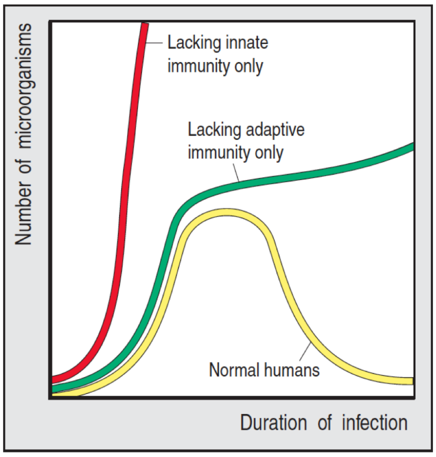


**Figure 1.1.:** The importance of the innate immunity.

In humans with healthy immune system the infection is almost completely cleared out of the body due to the synergistic action of innate and adaptive immunities (yellow line). If adaptive immunity is impaired, the infection is initially stopped but can't be cleared out and persists (green line), but if innate immunity is compromised, the infection becomes uncontrolled since there are no innate immunity responses generated to activate the adaptive immunity [4].

The innate immunity comprises several levels: anatomic, chemical-physiological (temperature, pH, soluble factors), cellular, and inflammatory barriers (fig. 1.2), which are closely related and work together to provide the effective protection.

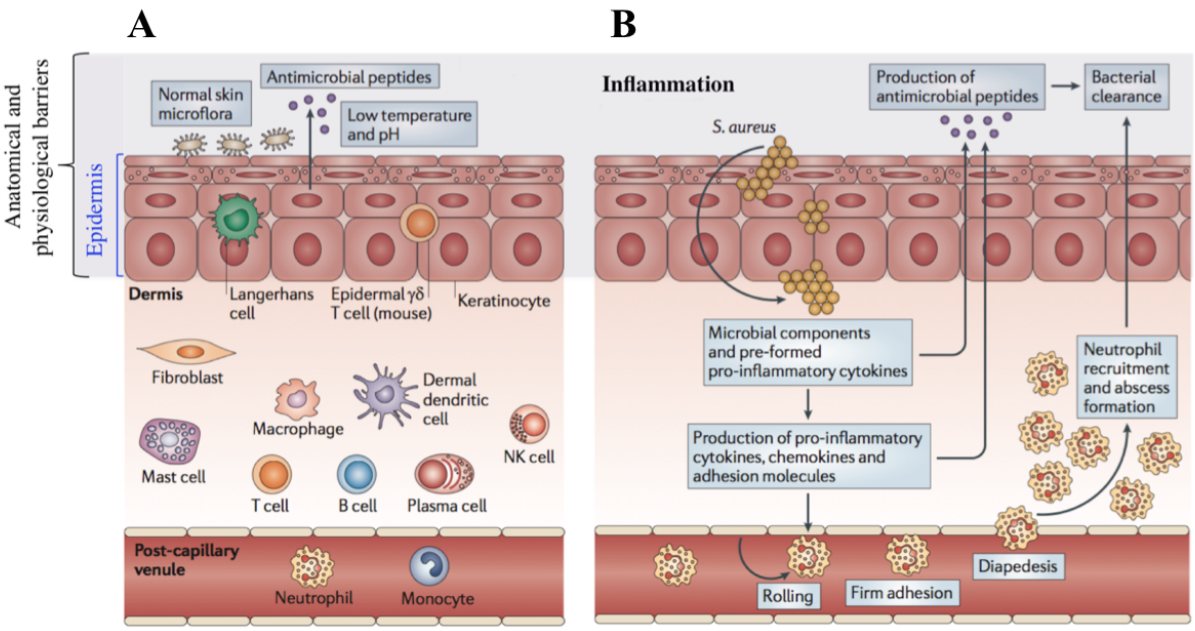

A
B
Anatomical and physiological barriers
Epidermis
Normal skin microflora
Antimicrobial peptides
Low temperature and pH
Dermis
Langerhans cell
Epidermal γδ T cell (mouse)
Keratinocyte
Fibroblast
Macrophage
Dermal dendritic cell
Mast cell
NK cell
T cell
B cell
Plasma cell
Post-capillary venule
Neutrophil
Monocyte
Inflammation
S. aureus
Production of antimicrobial peptides
Bacterial clearance
Microbial components and pre-formed pro-inflammatory cytokines
Production of pro-inflammatory cytokines, chemokines and adhesion molecules
Neutrophil recruitment and abscess formation
Rolling
Firm adhesion
Diapedesis

**Figure 1.2.:** Different levels of innate immunity: an example of skin infection by *Staphylococcus aureus*.

**A,** An intact skin protects the body from pathogen invasion: the tightly packed epidermis with the uppermost corneal layer of dead keratinocytes provides a physical barrier; normal skin microflora covers the skin and thus fills the niches preventing pathogen attachment and growth, which is also inhibited by low temperature and pH at the skin surface; antimicrobial peptides, secreted to the surface by underlying keratinocytes, provide an additional defense. The innate immunity cells in dermis survey their environment for invaded pathogens. **B**, Once this physical barrier is breached, an inflammation occurs: epidermal and dermal immune cells produce pro-inflammatory cytokines, chemokines and adhesion molecules, which recruit more neutrophils from blood stream to control the infection by formation of the abscess. Pro-inflammatory cytokines also induce the production of antimicrobial peptides. (Adapted from [5])

The anatomical barrier includes the skin and mucosa, and provides a physical obstacle for the pathogen invasion. Its importance is obvious as in the case of the loss of the integrity of the body's internal epithelia, infection is a major cause of mortality and morbidity [6].

### 1.1.1. *Anatomical and chemical-physiological barriers*

The skin is composed of two layers, epidermis and dermis. The epidermis consists of several layers of tightly packed cells, where the outer layer contains dead cells and waterproof protein keratin, and most of the pathogens are not capable to cross such a barrier. The underlying dermis contains sebaceous glands that produce oily secretion called sebum consisting of lactic acid and fatty acids providing the acidic pH, which inhibits the growth of many pathogens. The structure and composition of the skin makes it a formidable physical barrier, whereas mucosa does not possess as strong physical features, but it has other properties that compensate the lack of rigidity. Mucosa is composed of tight epithelial and the underlying connective tissue layers, and lines the conjunctiva, the gastrointestinal, respiratory, and urogenital tracts. These linings possess numerous non-specific defense mechanisms such as mucous secretions (tears, saliva) that wash away the microorganisms, or viscous mucus, which entraps the potential pathogens, and specialized organelles of lung epithelial cells, i.e. cilia, that expel the pathogens by its movement.

The skin and mucosal surfaces are also protected by antimicrobial substances secreted by keratinocytes or specialized epithelial cells. These substances include mucins (the glycoproteins that prevent attachment and entry of the microbes), antimicrobial enzymes and peptides. Lysozyme and phospholipase $A_2$ are antimicrobial enzymes secreted in mucosa and responsible for the destruction of microbial cell walls and membranes, respectively. Antimicrobial peptides in mammals include defensins, cathelicidins and histatins. Defensins are small amphipatic peptides possessing a positively charged region separated from hydrophobic region. It is suggested that they function by binding to the pathogen membranes through their positive region due to electrostatic attraction, followed by insertion of the hydrophobic region into the membrane, which results in pore formation and pathogen destruction. Defensins are secreted by keratinocytes in epidermis, the cells of tongue, respiratory and urogenital tracts, and by the Paneth cells of the gut into the gut lumen. Defensins are also secreted within the tissues by phagocytes, and they constitute one of the components of primary granules of the neutrophils. Cathelicidins are constitutively produced by neutrophils, macrophages, and by keratinocytes in the skin and epithelial cells in the lungs and intestine in response to infection. Their mechanism of action is similar to that of defensins. Histatins are produced by the parotid, sublingual, and submandibular glands in the oral cavity. They are short histidine-rich cationic peptides that are active against pathogenic fungi such as *Cryptococcus neoformans* and *Candida albicans* [6].

Both skin and mucosa are inhabited by non-pathogenic microorganisms, the normal flora or commensal microbiota, that compete with the pathogens for the nutrients and attachment sites, and this way defend the organism. Besides, the microbiota may produce antimicrobial substances as lactic acid or antimicrobial peptides bacteriocins [3, 6].

In the absence of wounding or disruption, despite all these defense mechanisms of the body, some pathogens can infect by specifically adhering to and colonizing epithelial surfaces. Some pathogens can also use surface molecules on the epithelial cells as footholds to invade the cells or get into the underlying tissues. Once pathogens invade the body, they face an immediate response by humoral innate immunity elements: the aforementioned antimicrobial peptides and a circulating plasma protein complex, a complement system, which targets pathogens for lysis and for phagocytosis by innate immunity cells (e.g. macrophages) [3, 6].

### 1.1.2. *The cells of innate immunity*

Apart from non-specific or broadly specific counter-action of physiological component of the innate immunity, a more specific response is carried out by the innate immunity cells: macrophages, neutrophils, basophils, eosinophils, mast cells, natural killer cells (NKs), and dendritic cells (DCs). Moreover, the tissue intrusion by the microbes triggers inflammatory responses, which recruit more effector cells and molecules from the bloodstream to the infection site and serve to destroy, dilute or wall off both the injurious agents and the injured tissue [7].

All cells of immune system, together with other blood cells, are believed to develop from a common precursor in the bone marrow, a Hematopoietic Stem Cell (HSC), under the stimulation by certain cytokines (fig. 1.3). The differentiation of the generated progenitor cells may be either completed in the bone marrow, or the cells may mature in several stages after they leave bone marrow [8].

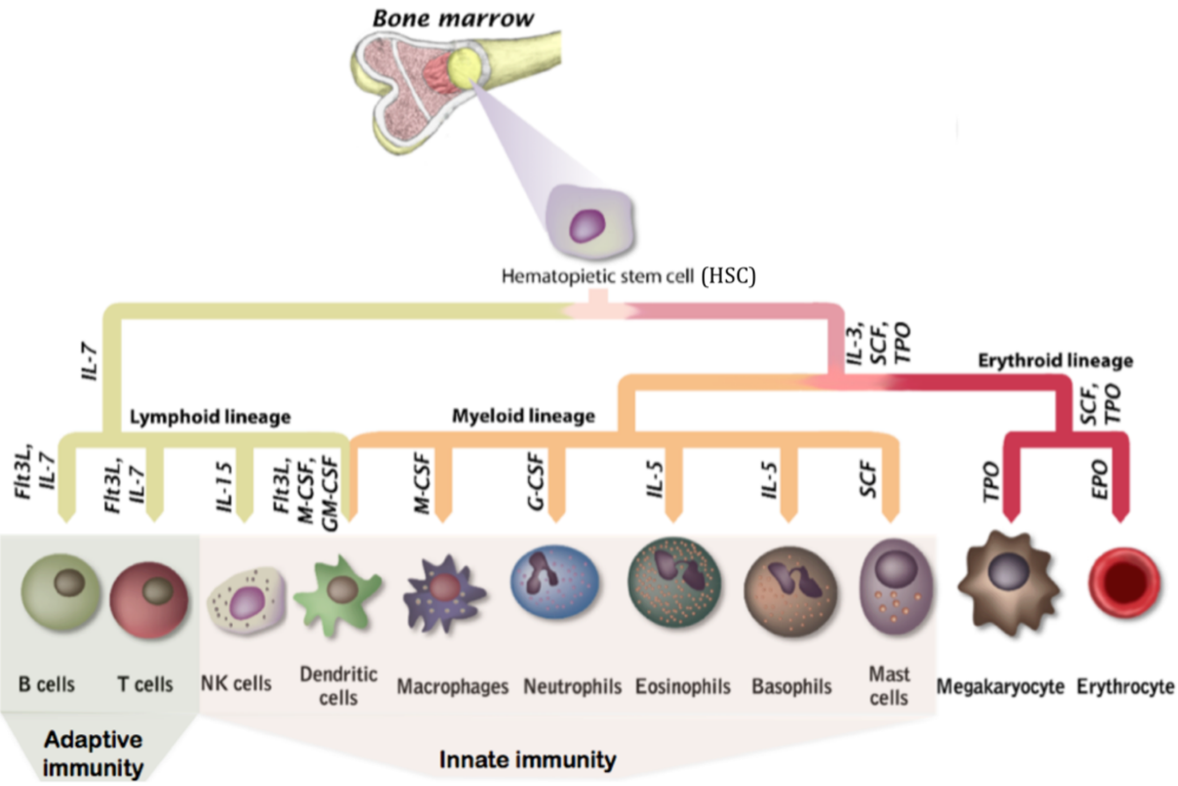

Bone marrow
Hematopietic stem cell (HSC)
IL-7
IL-3, SCF, TPO
Erythroid lineage
SCF, TPO
Lymphoid lineage
Myeloid lineage
Flt3L, IL-7
Flt3L, IL-7
IL-15
Flt3L, M-CSF, GM-CSF
M-CSF
G-CSF
IL-5
IL-5
SCF
TPO
EPO
B cells
T cells
NK cells
Dendritic cells
Macrophages
Neutrophils
Eosinophils
Basophils
Mast cells
Megakaryocyte
Erythrocyte
Adaptive immunity
Innate immunity

**Figure 1.3.:** The scheme of hematopiesis.
The blood cells are believed to be derived from a common hematopietic stem cell (HSC) via committed progenitors (not shown) that give rise to the erythroid, myeloid and lymphoid lineages. The main hematopoietins necessary for proliferation and differentiation of different lineages and cell types are indicated: G-CSF – Granulocyte Colony-Stimulating Factor (induces the differentiation of neutrophils, which are also known as granulocytes); M-CSF – Macrophage Colony-Stimulating Factor (induces the differentiation of macrophages); GM-CSF – Granulocyte-Macrophage Colony-Stimulating Factor (induces differentiation of both granulocytes (neutrophils, basophils, eosinophils and mast cells) and macrophages). IL-3, IL-5, IL-7 and IL-15 are interleukins; Flt3L is Fms-like tyrosine kinase 3 ligand; TPO is thrombopoietin; EPO is erythropoietin; SCF is Stem Cell Factor, which also plays a part in the differentiation of non-hematopoietic cells. (Adapted from [8])

Basophils, mast cells and eosinophils have a special role in the protection of epithelial surfaces, especially the mucosa of the gastrointestinal, respiratory and urogenital tracts, against the multicellular parasites, such as helminthes. They all are characterized by cytoplasmic granules containing inflammatory and cytotoxic mediators, thus they are referred to as granulocytes. Mast cells have a sentinel role and reside in mucosal and connective tissues, while basophils and eosinophils are circulating cells recruited from the bloodstream. They can recognize microorganisms either directly or through activation by complement or lymphocyte products. These granulocytes operate by releasing the contents of their granules to the exterior upon activation, thereby either creating an environment hostile to invading organisms or directly killing the parasites (in the case of eosinophils). Moreover, basophils and mast cells release molecules including histamine that are clinically important as the mediators of allergic and pathological inflammatory responses [8, 9].

The phagocytic cells of innate immunity comprise neutrophils, macrophages and immature DCs (iDCs). Neutrophils are the front-line effector cells of innate immunity: once the differentiation is complete, they circulate in a bloodstream for a few hours and after entering the tissues, they can ingest infectious microorganisms and kill them by microbicidal substances stored in specialized vesicles as well as by reactive oxygen species (ROS) generated by NOX family NADPH oxidases

[10]. These vesicles appear as granules when stained, therefore neutrophils have a name of granulocytes. They are also called polymorphonuclear leukocytes because of their characteristic feature of multilobed nucleus. Neutrophils are short-lived and die after two or three days, while macrophages are long-lived innate immunity cells that carry out immune surveillance in the tissues and play an important part in tissue maintenance. They differentiate from blood circulating monocytes as they leave the bloodstream. Like neutrophils, they ingest and destroy microorganisms. Macrophages also participate as effector cells in adaptive immune responses when activated by T lymphocytes or by antibodies secreted by B cells. Both neutrophils and macrophages release inflammatory cytokines.

Macrophages and most iDCs are sentinel cells residing in tissues, and depending on the tissue where they reside, their receptors and functional properties can vary. Apart from microbial invaders, both of these cell types sample tissues for normal tissue debris, as they possess a variety of scavenger receptors, specific for molecules characteristic to cells that have undergone apoptosis.

DCs form a link between innate and adaptive immunity, as their principal function is to induce the adaptive immunity. They may develop from both lymphoid and myeloid lineages, and a myeloid progenitor of DCs and macrophages is known as a monocyte. For a period of days to weeks immature DCs survey the peripheral tissues and operate as phagocytes to internalize the foreign microorganisms. This leads to differentiation into mature DCs (mDCs) that have the antigens presented on their surface and are no longer phagocytic. They leave the peripheral tissues to migrate into the lymphoid organs where they will encounter circulating naïve T cells and activate them. mDCs are characterized by the long branches, for which they are named: these enable to make contacts with several T cells simultaneously [8].

## 1.2. Pathogen recognition by innate immunity

The innate immune system functions using at least two recognition strategies: it is capable to distinguish microbial *non-self* from *missing self*. These recognition events occur thanks to a range of receptors that all cells of the innate immune system are equipped with, and through which various signals can be triggered by cytokines, conserved components of microorganisms, complement components and antibodies produced by B lymphocytes. The recognition of missing self is based on the recognition of molecules expressed only on healthy uninfected host cells, which leads to inhibition of innate immunity response, and the expression of these molecules is lost once the cells get infected, i.e. the missing self emerges. The missing self recognition plays an important role in the function of NK cells and complement [8, 9].

The recognition of non-self is based on the recognition of conserved molecular structures that are unique to microorganisms and that are not produced by the host. These structures are called Pathogen Associated Molecular Patterns (PAMPs) and include lipopolysaccharide (LPS) of gram-negative ($G^-$) bacteria, lipoteichoic acid (LTA) of gram-positive bacteria ($G^+$), peptidoglycans, lipoproteins generated by palmitoylation of the N-terminal cysteines of many bacterial cell wall proteins, lipoarabinomannan of mycobacteria, double-stranded RNA (dsRNA) and single-stranded RNA (ssRNA) produced by most viruses during the infection cycle, and β-glucans and mannans in fungal cell walls [9].

### 1.2.1. *The diversity of Pattern Recognition Receptors*

The receptors of the innate immune system that recognize PAMPs are called Pattern Recognition Receptors (PRRs). The main functions of PRRs include pathogen-induced phagocytosis, activation of pro-inflammatory signaling pathways, opsonization, activation of complement and coagulation cascades, and induction of apoptosis [11].

All PRRs can be defined as humoral or cell associated, and the latter ones can be subdivided to intracellular and cell surface molecules (fig. 1.4). The selected examples of PRRs and their ligands are listed in table 1.1.

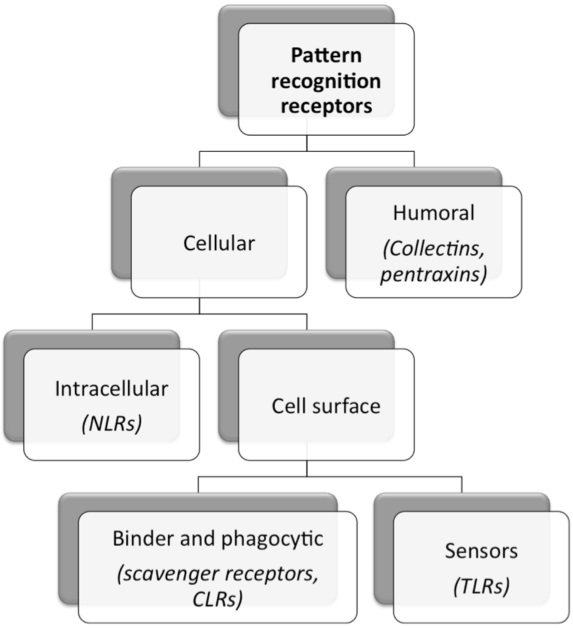


**Figure 1.4.:** Schematic representation of different groups of pattern recognition receptors with selected examples.

(Adapted from [12])

Secreted, or humoral, PRRs may activate complement, opsonize microbial cells to facilitate their phagocytosis, and in some cases function as accessory proteins for PAMP recognition by transmembrane receptors like Toll-like receptors (TLRs). Examples of humoral PRRs include the mannan-binding lectin (MBL), C-reactive protein (CRP), serum amyloid protein (SAP) and peptidoglycan-recognition proteins (PGRPs) [12, 11].

Cell surface PRRs may be either phagocytic/endocytic or sensor in nature. The sensor receptors do not bind or internalize ligand directly, but recognize PAMPs and induce pro-inflammatory signaling cascades, which lead to various antimicrobial effector responses. They include TLRs, but also many intracellular PRRs are sensing molecules [12]. Examples of intracellularly functioning sensors include NOD-like receptors (NLRs), protein kinase R (PKR), 2'-5'-oligoadenylate synthase (OAS)/RNaseL, and retinoid acid–inducible gene–1 (RIG-1)-like receptors (RLRs).

Phagocytic receptors bind and internalize ligands directly in a temperature dependent, saturable and inhibitable ligand binding manner, characteristic to classical receptors. These receptors include scavenger and C-type lectin receptors (CLRs). Scavenger receptors play several roles: they are important in uptake and clearance of degenerate components, such as modified host molecules and apoptotic cells, and they bind and internalize microorganisms and their products [13]. CLRs are a diverse family of receptors with the ability to bind to carbohydrate moieties and with variable physiological functions from cell adhesion to pattern recognition [14].

The cellular and humoral arms of the innate immune system collaborate and maintain host defense.

**Table 1.1.**: Examples of selected PRRs and their ligands.

| **Receptor family** | **Localization** | **Cell types** | **Receptor** | **Ligands** | **Ligand origin** |
|---|---|---|---|---|---|
| Collectin | Humoral | - | MBL | Terminal mannose residues | Bacterial surfaces |
| Pentraxins | Humoral | - | CRP | Phosphorylcholine | Bacterial surfaces |
| | | | SAP | Phosphorylcholine | Bacterial surfaces |
| Scavenger receptors | Cell surface | MØ; DCs; certain endothelial cells | SR-A | LPS and LTA | Microbial cell walls |
| | | | | Unidentified protein ligand in serum; activated B cells | Endogenous |
| | | | | $\beta$ amyloid protein; apoptotic cells; Ox-LDL and Ac-LDL; AGE modified proteins | Modified self |
| | | | MARCO | LPS | Microbial cell walls |
| | | | | UGRP-1 | Endogenous |
| | | | | Ac-LDL | Modified self |
| | | | SRCL-1 | $G^+$, $G^-$ bacteria, yeast | Microbes |
| | | | | T and Tn antigen | Endogenous |
| | | | | Ox-LDL | Modified self |
| TLRs | Ubiquitous | Ubiquitous | TLR1 | Triacyl lipopeptides | Mycobacteria |
| | Cell surface | Myeloid cells; mast cells; NKs; DCs; $\alpha\beta$ and $\gamma\delta$ T cells | TLR2 | Envelope proteins | Virus |
| | | | | GPI-linked proteins | Trypanosomes |
| | | | | Lipoproteins | Mycobacteria |
| | | | | LTA | $G^+$ bacteria |
| | | | | Peptidoglycans | $G^+$ bacteria |
| | | | | Zymosan | Fungi |
| | Intracellular (endosomal) | DCs; NKs | TLR3 | dsRNA | Viruses |

**Table 1.1.:** Examples of selected PRRs and their ligands.

| **Receptor family** | **Localization** | **Cell types** | **Receptor** | **Ligands** | **Ligand origin** |
|---|---|---|---|---|---|
| TLRs | Cell surface | Mo; mast cells; Neu; γδ T cells; Golgi in gut epithelial cells | TLR4 | Fusion protein | Respiratory syncytial virus |
| | | | | Glycoinositol phospholipids | Fungi |
| | | | | LPS | G$^-$ bacteria |
| | | | | Mannan | Fungi |
| | Cell surface | Epithelial cells; NKs; Mo; DCs | TLR5 | Flagellin | Bacteria |
| | Cell surface | Myeloid cells; mast cells; B cells | TLR6 | Diacyl lipopeptides | Mycobacteria |
| | | | | LTA | G$^+$ bacteria |
| | | | | Zymosan | Fungi |
| | Intracellular (endosomal) | DCs; B cells; eosinophils | TLR7 | ssRNA | Viruses |
| | Intracellular (endosomal) | NKs; T cells; myeloid cells | TLR8 | ssRNA | Viruses |
| | Intracellular (endosomal) | DCs; B cells; surface of tonsillar cells | TLR9 | CpG-containing DNA | Bacteria, protozoa, virus |
| | | | | Herpesvirus DNA | Viruses |
| | Cell surface | DCs; B cells | TLR10 | Unknown | Unknown |
| NLRs | Intracellular | Lymphocytes, MØ, DCs, epithelial and mesothelial cells | Nod1 | GM-tripeptide<br>*meso*-lanthionine, meso-DAP | *Helicobacter pylori*<br>*Shigella flexneri* |
| | | | | γ-D-Glu-DAP | *Listeria monocytogenes*<br>*Campylobacter jejuni* |
| | | | | D-lactyl-L-Ala-γ-Glu-meso-DAP-Gly | Enteropathogenic *Escherichia coli* |
| | | | | heptanolyl-γ-Glu-meso-DAP-Ala | *Chlamydia pneumoniae*<br>*Pseudomonas aeruginosa*<br>*Bacillus* spp. |
| | | | Nod2 | MDP | *Streptococcus pneumoniae* |
| | | | | MurNAc-L-Ala-γ-D-Glu-L-Lys | *Listeria monocytogenes*<br>*Mycobacterium tuberculosis*<br>*Salmonella typhimurium*<br>*Staphylococcus aureus*<br>*Shigella flexneri* |

**Table 1.1.:** Examples of selected PRRs and their ligands.

| **Receptor family** | **Localization** | **Cell types** | **Receptor** | **Ligands** | **Ligand origin** |
|---|---|---|---|---|---|
| NLRs | | | Nlrc4 | flagellin | *Salmonella typhimurium*<br>*Legionella pneumoniae*<br>*Pseudomonas aeruginosa* |
| | | | Nlrp1b | anthrax lethal toxin | *Bacillus anthracis* |
| | | | Nlrp3 | Bacterial and viral RNA, viral DNA, uric acid crystals, LPS, LTA, MDP, silica, asbestos | |
| CLRs | *See table 1.2* | | | | |

Abbreviations: Ac-LDL, acetylated low density lipoprotein; AGE, advanced glycation end-product; DAP, diaminopimelic acid; GM-tripeptide, N-acetyl-D-glucosaminyl-β(1,4)-N-acetylmuramyl-L-Ala-D-Glu; GPI, Glycosylphosphatidylinositol; LDL, Low-Density Lipoproteins; LTA, lipoteichoic acid; MØ, macrophages; MDP, muramyl dipeptide; Mo, monocytes; Neu, neutrophils; Ox-LDL, oxidized LDL; UGRP-1, Uteroglobin-Related Protein 1. Information collected from [11, 12, 15, 16].

### 1.2.2. *C-type lectin receptors as PRRs*

The C-type lectin family comprises a large group of Metazoan proteins that contain C-type lectin-like domains (CTLDs). Although originally CTLDs were identified as the structures that bind carbohydrates in a $Ca^{2+}$-dependent manner (thereof the term C-type), not all the members of this family recognize carbohydrates and not all need $Ca^{2+}$ for ligand binding. Therefore, CTLDs are referred as “C-type lectin-like domains” [17].

The mammalian C-type lectin receptors (CLRs) are divided into 17 types based on their phylogenetic relationships and domain structures (fig. 1.5). Most of the CLR family members function as adhesion receptors, and only CLRs of type II, V and VI are present mostly on myeloid lineage immune cells and function as PRRs, while type III CLRs are soluble PRRs [18].

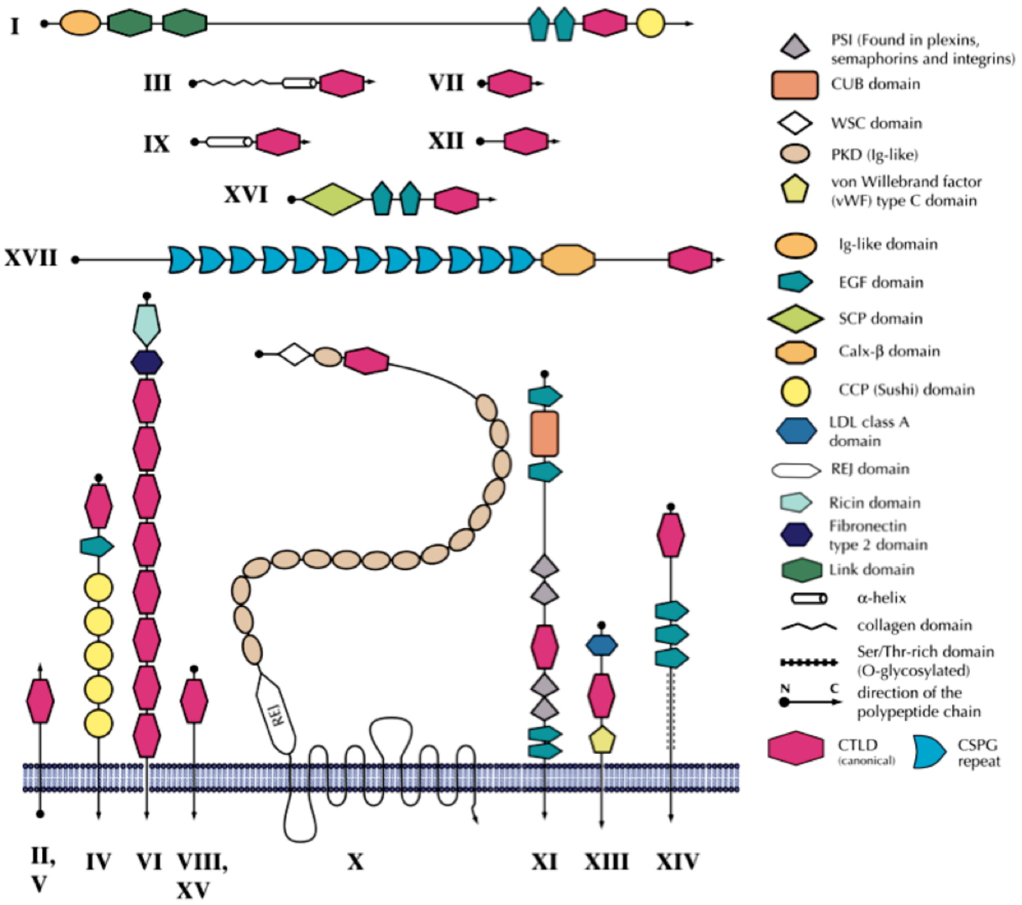

I
III
VII
IX
XII
XVI
XVII
REJ
II, V
IV
VI
VIII, XV
X
XI
XIII
XIV
PSI (Found in plexins, semaphorins and integrins)
CUB domain
WSC domain
PKD (Ig-like)
von Willebrand factor (vWF) type C domain
Ig-like domain
EGF domain
SCP domain
Calx-β domain
CCP (Sushi) domain
LDL class A domain
REJ domain
Ricin domain
Fibronectin type 2 domain
Link domain
α-helix
collagen domain
Ser/Thr-rich domain (O-glycosylated)
N
C
direction of the polypeptide chain
CTLD (canonical)
CSPG repeat

**Figure 1.5.:** The structural diversity of CLR family.

CLR types are marked by roman numbers: I – lecticans, II – asialoglycoprotein receptor (ASGR) DC receptor group, III – collectins, IV – selectins, V – NK receptors, VI – multi-CTLD endocytic receptors (macrophage mannose receptor group), VII – Reg proteins, VIII – chondrolectin group, IX – tetranectin group, X – polycystin 1, XI – attractin, XII – EMBP (eosinophil major basic protein), XIII –DGCR2 (the product of DiGeorge syndrome critical region gene 2), XIV – thrombomodulin group, XV – Bimlec, XVI – SEEC (soluble protein containing SCP, EGF, EGF and CTLD domains), XVII – CBCP (Calx-β and CTLD containing protein). (Adapted from [17])

CLRs that function as PRRs are mostly expressed by different DC subsets (table 1.2). They bind pathogens through the recognition of mannose, fucose, glucan and other carbohydrate structures. The combination of CLRs on DCs enables the recognition of most classes of human pathogens. Pathogen recognition by CLRs leads to its internalization, degradation and subsequent antigen presentation [19].

**Table 1.2.:** Selected examples of CLRs that function as PRRs.

| Type | CLR | Expression | Recognized glycans | Glycans from pathogen | Endogenous ligands |
|---|---|---|---|---|---|
| **II** (Ca$^{2+}$-dependent CRD) | **DC-SIGN**<br>(CLEC4L, CD209) | myDCs | High mannose and fucose (Lewis$^{X}$, Lewis$^{Y}$, Lewis$^{A}$, Lewis$^{B}$) | **Bacteria:**<br>*Mycobacterium tuberculosis*; *Mycobacterium leprae*; BCG; *Lactobacilli* spp.; *Streptococcus pneumoniae*; *Leptospira interrogans*; *Helicobacter pylori*<br>**Viruses:**<br>HIV-1; measles virus; Dengue virus; HCV; CMV; SARS coronavirus; HSV; H5N1; WNV; Ebola virus and other filoviruses; phleboviruses<br>**Fungi:**<br>*Candida albicans*; *Aspergillus fumigatus*<br>**Protozoa:**<br>*Leishmania* spp.<br>**Other:**<br>Tick *Ixodes scapularis* saliva protein Salp15; peanut allergen Ara h1; *Schistosoma mansoni* soluble egg antigens | ICAM-2; ICAM-3; CEACAM-1; Mac-1; CEA |
| | **DC-SIGNR**<br>(CLEC4M, CD299) | Endothelial cells in lymph-node sinuses; liver sinusoidal endothelial cells | Mannosylated glycans, high mannose N-glycans | HIV-1; Ebola; *Schistosoma mansoni* | ICAM-1; ICAM-2; ICAM-3 |

**Table 1.2.:** Selected examples of CLRs that function as PRRs.

| Type | CLR | Expression | Recognized glycans | Glycans from pathogen | Endogenous ligands |
|---|---|---|---|---|---|
| **II** (Ca$^{2+}$-dependent CRD) | **Langerin** (CLEC4K, CD207) | LCs; dermal DC subset | High mannose, fucose (Lewis$^{Y}$, Lewis$^{B}$), GlcNAc, β-glucans, sulphated sugars (heparin) | **Bacteria:**<br>*Mycobacterium leprae*<br>**Fungi:**<br>*Candida*; *Saccharomyces cerevisiae*; *Malassezia furfur*<br>**Viruses:**<br>HSV;<br>Measles virus;<br>HIV-1 | Type I pro-collagen |
|  | **MGL** (CLEC10A, CD301) | myDCs; MØ | Terminal GalNAc (Tn, LDN and LDNF antigens) | Filoviruses; Influenza virus; *Schistosoma mansoni* | CD45; gangliosides; MUC-1 |
|  | **Dectin-2** (CLEC6A) | myDCs; pDCs; Mo; MØ; B cells; Neu | High mannose | **Bacteria:**<br>*Mycobacterium tuberculosis*<br>**Fungi:**<br>*Candida albicans*;<br>*Trichophyton rubrum*;<br>*Aspergillus fumigatus*;<br>*Microsporum audounii*;<br>*Paracoccoides brasiliensis*<br>**Allergens:**<br>House dust mite *Dermatophagoides pteronyssinus* allergens | Unknown |
|  | **Mincle** (CLEC4E) | myDCs; Mo; MØ | α-mannose | *Malassezia* spp.; *Mycobacteria*; *Candida* | Damaged cells |
|  | **DCIR** (CLEC4A) | myDCs; pDCs; Mo; MØ; B cells; Neu | Unknown | HIV-1 | Unknown |
|  | **BDCA2** (CLEC4C, CD303) | pDCs; Mo; MØ; Neu | Unknown | Unknown | Unknown |

**Table 1.2.:** Selected examples of CLRs that function as PRRs.

| **Type** | **CLR** | **Expression** | **Recognized glycans** | **Glycans from pathogen** | **Endogenous ligands** |
|---|---|---|---|---|---|
| **V** (Ca$^{2+}$-independent CRD) | **Dectin-1**<br>(CLEC7A) | myDCs; Mo; MØ; B cells | β-1,3-glucan | **Bacteria:**<br>*Mycobacterium tuberculosis*<br>*Mycobacterium abscessus*<br>**Fungi:**<br>*Candida albicans*;<br>*Aspergillus fumigatus*;<br>*Pneumocystis carinii*;<br>*Penicillium marneffei*;<br>*Coccidioides posadasii*;<br>*Histoplasma capsulatum* | Ligand on T cells |
| | **MICL**<br>(CLEC12A) | myDCs; Mo; MØ; Neu | Unknown | Unknown | Unknown |
| | **CLEC2**<br>(CLEC1B) | Platelets, peripheral blood Neu | Unknown | HIV-1; Snake venom rhodocytin; podoplanin | Unknown |
| **VI** (Ca$^{2+}$-dependent CRD) | Mannose receptor (CLEC13D, CD206) | myDCs; MØ | High mannose, fucose, sulphated sugars (sLewis$^{X}$) | **Bacteria**:<br>*Mycobacterium tuberculosis*;<br>*Mycobacterium kansasii*;<br>*Francisella tularensis*;<br>*Klebsiella pneumoniae*;<br>*Streptococcus pneumoniae*<br>**Viruses**:<br>HIV-1; Dengue<br>**Fungi**:<br>Candida albicans;<br>Cryptococcus neoformans;<br>*Pneumocystis carinii*<br>**Protozoa**:<br>*Leishmania* spp.<br>Allergens | L-selectin; MUC-1 |
| | **DEC-205**<br>(CLEC13B, CD205) | myDCs | Unknown | Unknown | Unknown |

Abbreviations: BCG, *Mycobacterium bovis* Bacillus Calmette Guérin; CEA, carcinoembryonic antigen; CMV, cytomegalovirus; GalNAc, N-acetylgalactose amine; GlcNAc, N-acetylglucose amine; HCV, hepatitis C virus; HIV, human immunodeficiency virus; HSV, herpes simplex virus; LCs, Langerhans cells; LDN antigen, GalNAcβ1-4GlcNAc-R; LDNF antigen, GalNAcβ1-4(Fucα1-3)GlcNAc-R; MGL, macrophage galactose-type C-type lectin; myDCs, myeloid DCs; pDCs, plasmacytoid DCs; SARS, severe acute respiratory syndrome-associated; sLewis$^{X}$, sulphated Lewis$^{X}$; Tn antigen, GalNAcα1-O-Ser/Thr; WNV, West Nile virus. Information collected from [19, 20, 21, 22, 23, 24, 25].

The CLRs can be immune activating or inhibitory depending on their ability to associate with certain signaling molecules or the presence of specific motifs in their cytoplasmic tails. Most of the type II CLRs are predicted to be activating as in their transmembrane regions they have a positively charged residue which allows association with adaptor proteins. The activating CLRs may harbour the immunoreceptor tyrosine-based activation motifs (ITAMs), which constitute two YxxI/L motifs separated by 6-12 amino acid spacers (YxxI/Lx(6-12)YxxI/L). Upon ligand binding, clustering of CLRs occurs and ITAMs are phosphorylated, which initiates a downstream signaling cascade eventually leading to activation of various cellular responses. Besides, there are CLRs that bear ITAM-like motifs (N-terminal tyrosin in YxxxL/I) in their cytoplasmic tail, for ex. dectin-1. The activating CLRs include dectin-2, DCAR, BDCA2, Mincle and DC-SIGN [20].

The inhibitory CLRs themselves possess the immunoreceptor tyrosine-based inhibition motifs (ITIMs: I/V/L/SxYxxI/L/V) in their cytoplasmic tails, and in this case the ligand binding to CLR followed by phosphorylation of ITIM initiates signaling cascades that culminate at the inhibition of cellular activation. The examples of inhibitory CLRs include DCIR and MICL. There also exist CLRs that harbour ITIMs but mediate cellular activation [20].

#### 1.2.2.1. *The structure of C-type lectin-like domains*

The common feature of all CLRs is that they possess at least one CTLD – a compact globular structure with a characteristic fold designated "C-type lectin-like fold" that is unusual to any other known proteins. For the majority of CLRs that function as PRRs, the CTLDs actually bind sugars, usually in $Ca^{2+}$-dependent manner, and therefore this domain is commonly called a "carbohydrate recognition domain" (CRD) rather than CTLD.

All CRDs possess a characteristic "double-loop" fold (fig. 1.6). The whole domain can be regarded as a loop with two flanking α helices (α1 and α2) and two antiparallel β-sheets: N- and C-terminal β strands β1 and β5 constitute the basal β-sheet, and the top β-sheet is formed by strands β2, β3, and β4. The long loop region enters and exits the core domain at the same location, and is involved in $Ca^{2+}$-dependent carbohydrate binding, and for some CRDs in domain-swapping dimerization. Four highly conserved cysteins form two disulphide bridges at the bases of the loops: C1-C4 bridge links α1 and β5, and C2-C3 bridges β3 strand and a loop upstream the β5 strand. Another highly conserved sequence feature of the CRDs is the "WIGL" motif located within β2 strand and believed to stabilize the core of the domain.

The long loop region among different CRDs varies, and those that possess it are designated "canonical", while those that lack it are called compact. The presence or absence of a short extension at N-terminus, a β1´-hairpin, further subdivides CRDs to long or short forms, respectively. Two additional cysteins at the beginning of CRDs sequence are characteristic for the long form CRDs. The corresponding disulphide bridge (C0-C0´) stabilizes the β-hairpin.

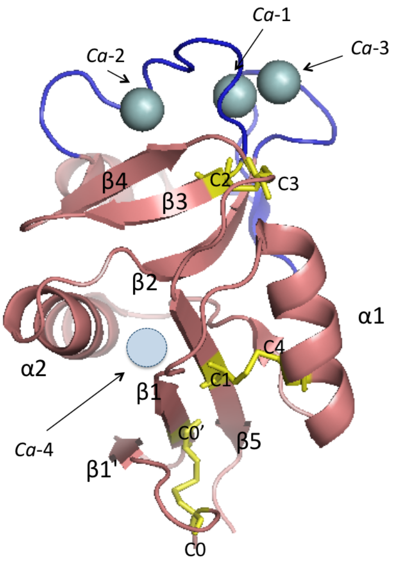


**Figure 1.6.:** A typical CRD structure (DC-SIGN CRD.
The long loop is shown in blue, disulphide bridges in yellow sticks. The three Ca-binding sites present in this lectin presented as cyan spheres, and the location of the fourth Ca-binding site (absent in this particular structure) is shown as a cyan circle. (pdb:1k9i; adapted from [17])

There may be up to four Ca-binding sites in the CRDs, and their occupancy depends on the sequence of a particular CRD and on crystallization conditions. Sites 1, 2, and 3 are located within the long loop, while the fourth site participates in the salt bridge formation between helix α2 and β1/β5 sheet [17].

#### 1.2.2.2. *The structural basis of sugar binding within carbohydrate recognition domains*

The sugar binding in the CRDs occurs at $Ca^{2+}$-binding site 2, and both carbonyl sidechains coordinating calcium and $Ca^{2+}$ itself are involved in sugar binding. $Ca^{2+}$-coordination at this site is provided by carbonyl sidechains mainly within two characteristic motifs. One of these motifs, EPN or QPD, resides in a long loop region and defines the monosaccharide binding specificity. The second one, a WND motif, is contributed by β4 strand. Additionally, a carbonyl sidechain provided by the residue preceding the second conserved cysteine, also participates in $Ca^{2+}$ coordination at site 2. The schematic representation of $Ca^{2+}$ coordination and hexose binding is depicted in figure 1.7A.

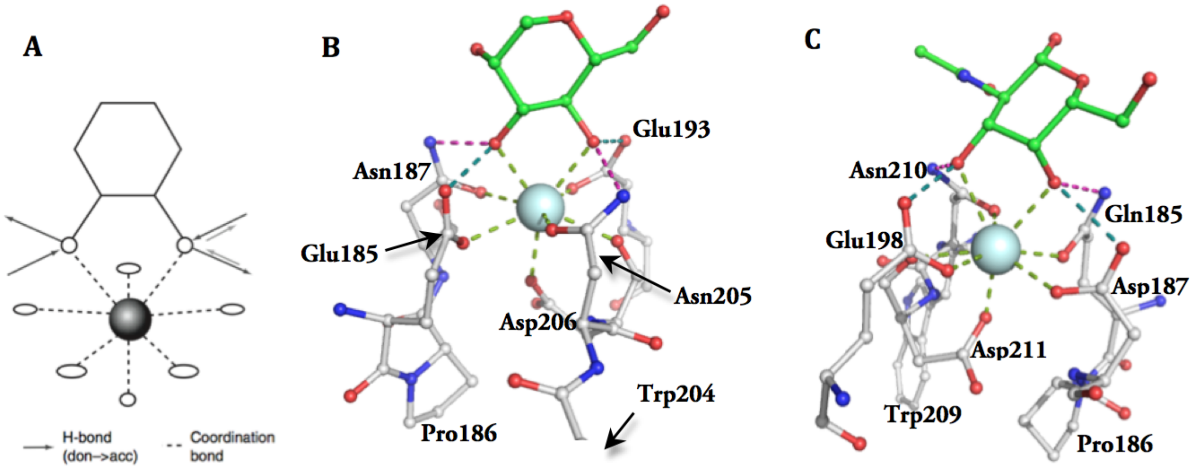


**Figure 1.7.:** Monosaccharide binding in Ca site 2.

A, schematic representation of $Ca^{2+}$ coordination and hexose binding: a grey sphere represents $Ca^{2+}$, empty circles and ovals show the oxygens, black arrows show the direction of H-bonds in mannose group CRDs, and light grey arrows mark changed directions in galactose-specific CRDs. **B**, mannose residue bound to MBP-A CRD (pdb:2msb). **C**, engineered for galactose-type specificity CRD of MBP-A complexed with GalNAc (MBP-A mutant QPDWGHV ([26], pdb:1bcj). Coordination bonds are light green; H-bonds where sugar hydroxyl acts as acceptor or donor are marked as pink or cyan dashed lines, respectively. (Adapted from [17].)

The overall network of the hydrogen-bond donors and acceptors in the site determines the binding orientation of the carbohydrate and also which hydroxyls of the carbohydrate it can accept, i.e. the monosaccharide specificity. The EPN motif has a configuration, which accommodates mannose group monosaccharides (fig. 1.7B), while QPD motif determines specificity for galactose group monosaccharides (fig. 1.7C). In both of these motifs the cis-configuration of the two carbonyl sidechains separated by proline is crucial for $Ca^{2+}$ coordination and sugar binding. Besides the H-bond network imposed constraints, other structural elements in the binding sites introduce selectivity to particular ligands within the mannose or galactose groups.

The other three $Ca^{2+}$ binding sites play the structural stabilization role, as removal of $Ca^{2+}$ increases susceptibility to proteolysis and changes physical properties of the domain. $Ca^{2+}$ binding site 2 is also important for structural stability of the domain. It has been shown that pH-induced loss of $Ca^{2+}$ causes the destabilization of the loops, which has an important physiological role for CRDs of endocytic receptors as internalization of ligand-bound receptor to acidic lysosomes and consequent $Ca^{2+}$ loss leads to the release of the ligand for further processing, while receptor is recycled to the cell surface [17].

## 1.3. Dendritic cells and adaptive immunity initiation

Dendritic cells comprise a diverse group of professional antigen presenting cells (APCs), which share a particular morphology, i.e. long surface membrane extensions called dendrites (fig. 1.8), and have a common feature of being potent stimulators of T cells to induce the adaptive immunity. The diversity of DCs is a result of a combination of several features including their origin, differentiation state, and anatomic location. The different subsets are identified by expression of surface markers, but the distinction between them is often debatable. Nevertheless, DCs can be broadly divided to myeloid (myDCs) and plasmacytoid DCs (pDCs). The latter originate from lymphoid progenitor, they resemble plasma cells, hence they are called “plasmacytoid”. myDCs are derived from monocytes and differentiated from peripheral blood mononuclear cells (PBMCs). Other DC subpopulations include Langerhan’s cells (LCs) residing in epidermis of the skin, and interstitial DCs (iDCs) found in dermis of the skin and stromal compartments of mucosal tissues. While myDCs are very potent at pathogen uptake and antigen presentation, the lack of CLRs and Fc receptors on pDCs together with weak macropinocytosis suggest different function of pDCs, which may involve self-antigen presentation and/or it is restricted to viral antigen presentation. Besides, pDCs are the major source of type I interferons (IFN) supporting their role in antiviral immunity [14, 27].

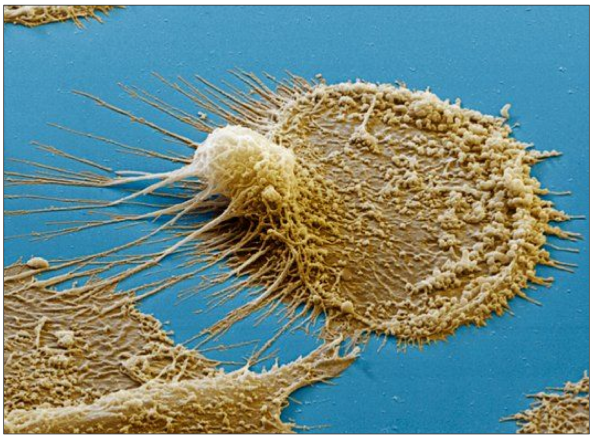


**Figure 1.8.:** Colored scanning electron micrograph of a dendritic cell.
From `http://www.sciencephoto.com/media/305587/enlarge`.

All DCs express PRRs on their surface including TLRs and CLRs, but each subset has its characteristic differential expression of PRRs, therefore distinct subsets will be reactive to a certain group of PAMPs present in particular pathogens (examples in tables 1.1 and 1.2). The diversity of these receptors on DCs enables them to uptake multiple signals from their surroundings and finally to transfer them for induction of appropriate adaptive immune responses [14, 28].

The immature skin and mucosal DCs form a dense network of resident cells, which sample their environment for foreign microorganisms. Upon pathogen recognition or in response to cytokines such as TNF-α, IL-4, GM-CSF, the maturation of DCs is triggered. This involves down-regulation of endocytic receptors, up-regulation of maturation markers (e.g. costimulatory molecules such as CD40, CD80 and CD86), raise in the levels of major histocompatibility complex (MHC) class II expression, loss of adhesiveness and induction of DC migration to lymphoid organs to present the processed antigens for T cell activation (fig. 1.9). Receptor-triggered DC responses also include expression of pro-inflammatory (TNF-α, IL-1, IL-6) or anti-inflammatory (IL-10) cytokines and differential production of chemokines, which results in recruitment of different T-cell subsets at the site of infection [14].

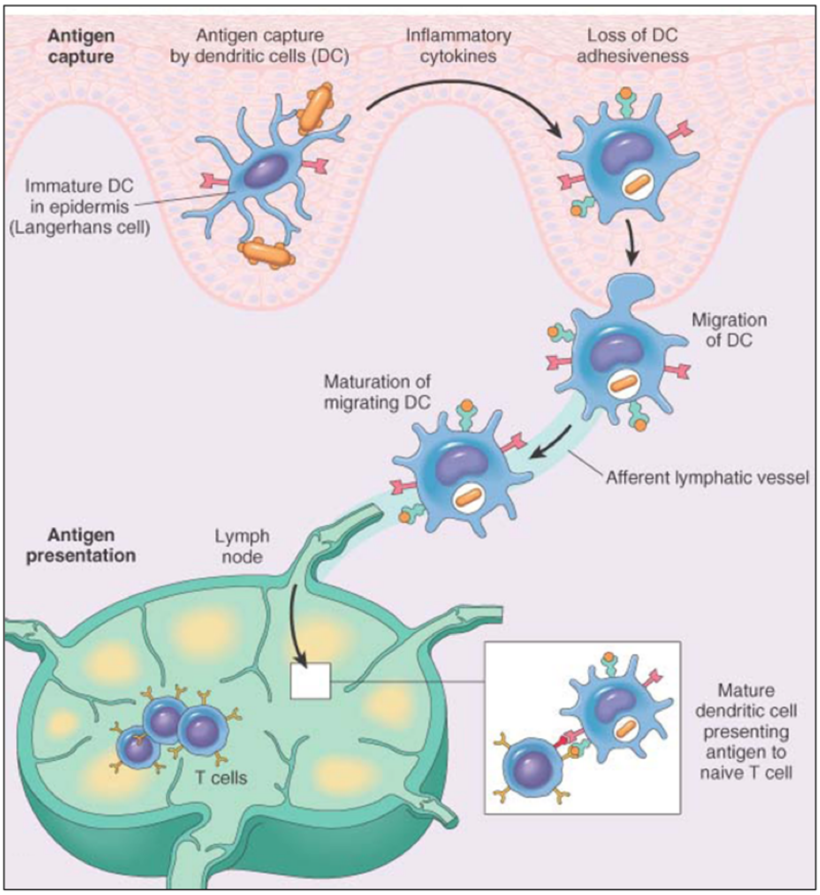

Antigen capture
Antigen capture by dendritic cells (DC)
Inflammatory cytokines
Loss of DC adhesiveness
Immature DC in epidermis (Langerhans cell)
Migration of DC
Maturation of migrating DC
Afferent lymphatic vessel
Antigen presentation
Lymph node
T cells
Mature dendritic cell presenting antigen to naive T cell

**Figure 1.9.:** Antigen uptake, maturation and migration of dendritic cells to activate T cells. From `http://medicinembbs.blogspot.fr/2011/03/process-of-immunity-images.html`.

Depending on which PRRs are triggered by pathogens, different signaling pathways are initiated in DCs. Generally, DC triggering through TLRs leads to DC activation, while PAMP recognition by CLRs result in pathogen uptake, processing, antigen presentation on MHC class I or II molecules, but also may involve modulation of DC activation [14]. On the other hand, the maturation of myDCs may be inhibited by immunosuppressive cytokines such as IL-10 and TGF-β, also by corticosteroids, cyclosporine A, and 1α,25-dihydroxyvitamin D. Such DCs may induce regulatory T cells, which maintain tolerance to self-antigens. Thus DCs have a dual function: they either boost the immune system, or dampen it leading to tolerance and maintenance of the immune homeostasis [28].

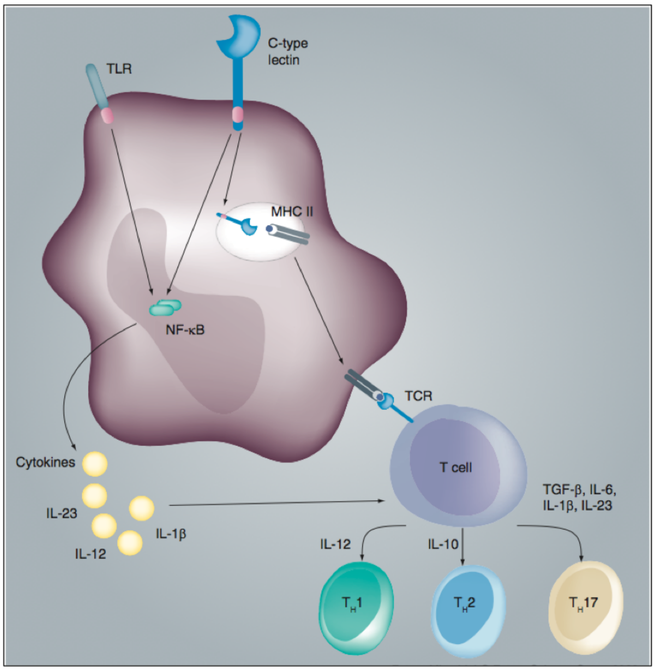

C-type
lectin
TLR
MHC II
NF-κB
TCR
T cell
Cytokines
IL-23
IL-12
IL-1β
TGF-β, IL-6,
IL-1β, IL-23
IL-12
IL-10
$T_H1$
$T_H2$
$T_H17$

**Figure 1.10.:** The cross-talk between TLR and CLR signaling shape the adaptive immunity. (From [29])

The differentially activated DCs induce naïve CD4$^+$ T cell differentiation into distinct T helper cells: Th1, Th2 or Th17 (fig. 1.10). These T helper cells are responsible for fighting different types of pathogens. Th1 produce interferon-γ (IFN-γ), which activates macrophages and cytotoxic T cells to fight intracellular pathogens. Th2 cells secrete interleukin-4 (IL-4), IL-5, and IL-13, which are responsible for the activation of B cells and humoral immune responses against extracellular pathogens. Antifungal and antibacterial immunity is contributed by Th17 cells secreting IL-17, which mobilizes phagocytes [18].

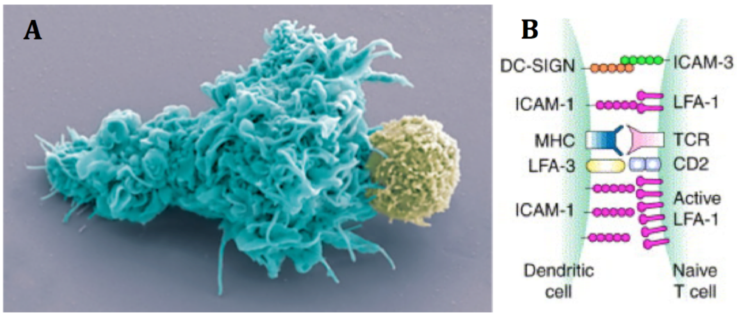

A
B
DC-SIGN
ICAM-3
ICAM-1
LFA-1
MHC
TCR
LFA-3
CD2
ICAM-1
Active
LFA-1
Dendritic
cell
Naive
T cell

**Figure 1.11.:** Formation of immunological synapse.
**A**, Field emission scanning electron microscope image of a human dendritic cell (blue pseudo-color) in close interaction with a lymphocyte (yellow pseudo-color). This contact may lead to the creation of an immunological synapse (from `http://microscopic.co/`) **B**, The receptors participating in immunological synapse formation (adapted from [30]).

The information about invading pathogens or abnormal self is transferred from DCs to naïve T lymphocytes through immunological synapse – a specialized cell–cell adhesive junction characterized by stability and directed secretion. The formation of immunological synapse is initiated by adhesive contacts between the two cells that are mediated by a number of adhesion receptors (fig. 1.11 A, B), which facilitate to form T cell receptor (TCR) and antigen-bound MHC class I or II complexes for antigen scanning by T cells and signaling to induce appropriate adaptive immune responses [31].

# 2. Human immunodeficiency virus: the infection via sexual transmission

## 2.1. The mechanism of HIV entry to the host cell

The host cell infection by HIV starts by the entry of the virus into the cell. The major target cells of HIV are $CD4^+$ T lymphocytes, which also express chemokine receptors CCR5 and CXCR4 that are exploited by HIV to enter the cells, hence also called HIV coreceptors. The virus entry starts with the adhesion of the virion to the target cell and ends with the fusion of the cell and viral membranes followed by delivery of the viral core into the cytoplasm (reviewed in [32]).

Attachment of the virion (fig. 2.1(1)) can be relatively nonspecific, for instance, HIV envelope protein (Env, the trimer of gp120 and gp41 heterodimers, where gp41 initially is hidden) can interact with negatively charged cell-surface heparan sulfate proteoglycans [33]. More specific adhesion includes interactions between Env and α4β7 integrin [34, 35] or CLRs such as DC-SIGN [36]. Either way of adhesion most likely brings Env into close proximity with the host CD4 and a coreceptor, leading to the fusion of viral and target cell membranes (fig. 2.1) [37].

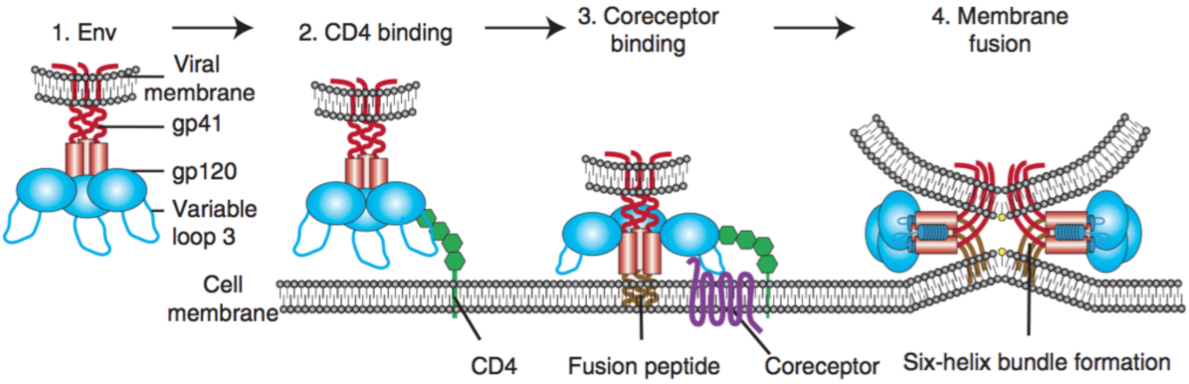


**Figure 2.1.**: Schematic representation of HIV entry to the target cell.
**(1)** HIV first attaches to the host cell via Env binding to host cell surface structures. Then binding to CD4 receptor occurs **(2)**, which causes conformational changes in Env, allowing coreceptor binding, which is mediated in part by the V3 loop of Env **(3)**. Binding to CD4 also exposes gp41 subunit and thus initiates the membrane fusion process as the fusion peptide of gp41 inserts into the target membrane, followed by six-helix bundle formation and complete membrane fusion **(4)**. (From [32])

Env binding to its primary receptor CD4 (fig. 2.1(2)), a member of the immunoglobulin superfamily that enhances T-cell receptor (TCR)-mediated signaling, is absolutely required for the infection. This binding event induces rearrangements in gp120 subunit of Env, which ultimately result in V3 loop repositioning and bridging sheet exposure that are essential for coreceptor engagement (reviewed in [32]). Subsequent binding to the coreceptor (fig. 2.1(3)), CCR5 or CXCR4 depending on the virus strain R5 or R4, triggers the membrane fusion potential of Env, and usually is followed by the “surfing” of the virus particle to the site where productive membrane fusion may occur. It is thought that HIV might usurp the host cell machinery to reach cell surface sites where membrane fusion can occur [38]. Besides, HIV may need to be endocytosed by the host cell for productive membrane fusion to occur [39].

Coreceptor-bound Env undergoes conformational changes that expose the hydrophobic gp41 fusion peptide (fig. 2.1(3)), which then inserts into the host cell membrane and folds to form six-helix bundle (fig. 2.1(4)). The latter is the driving force that brings the opposing membranes into close proximity, resulting in the formation of a fusion pore (reviewed in [40]).

Once the virus enters the cell, it can start its replication and productive infection, which ultimately lead to the depletion of $CD4^+$ T lymphocytes in the body, and thus the acquired immunodeficiency syndrome (AIDS).

Although the major target of HIV is $CD4^+$ T cells, earlier studies have shown that DCs are crucial for HIV-1 infection enhancement and dissemination in mucosa, in the case of sexual HIV transmission, which leads to the productive infection of the $CD4^+$ T cells and the burst of the disease [41][42].

## 2.2. Dendritic cell implication in HIV-1 sexual transmission

### 2.2.1. *Mechanism of HIV invasion of genital mucosa*

Sexual transmission through genital and rectal mucosa accounts for the vast majority of HIV-1 infections worldwide. The viral invasion in women occurs mostly through the non-keratinized squamous epithelium of the vagina and ectocervix, and also through the single-layer columnar epithelium of the endocervix (figure 2.2A). Sexual transmission in men most frequently takes place through the inner foreskin and the penile urethra as a consequence of penile–vaginal or penile–anal intercourse (figure 2.2B). Both women and men are infected following the receptive anal intercourse [43].

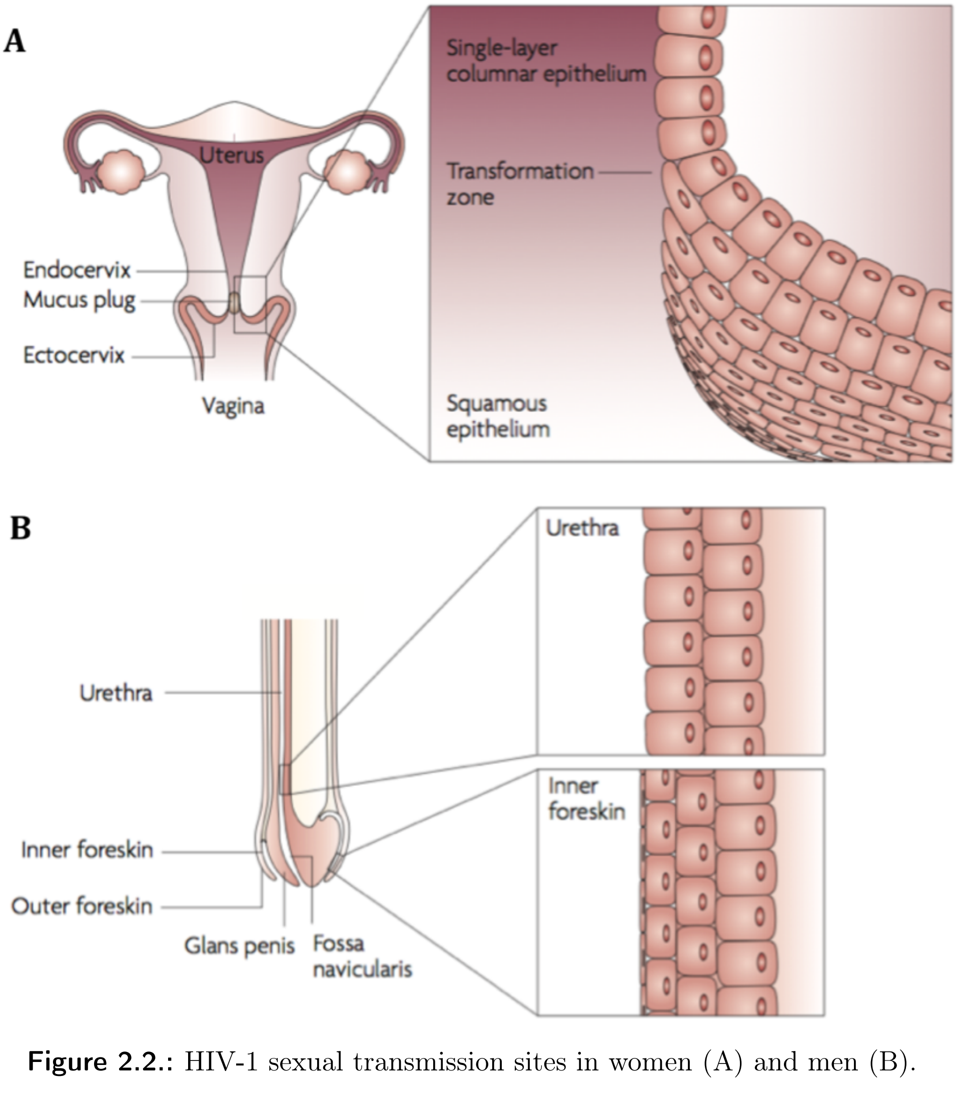

A
Single-layer columnar epithelium
Uterus
Transformation zone
Endocervix
Mucus plug
Ectocervix
Vagina
Squamous epithelium
B
Urethra
Urethra
Inner foreskin
Inner foreskin
Outer foreskin
Glans penis
Fossa navicularis

**Figure 2.2.:** HIV-1 sexual transmission sites in women (A) and men (B).
(From [43])

Although sexual transmission through genital mucosa is the dominating route of infection for both genders, women appear to be more susceptible to this type of viral invasion through cervicovaginal tissues (table 2.1) [44]. Therefore, HIV sexual transmission mechanisms in female genital mucosa are reviewed below in more detail.

**Table 2.1.:** Contribution of mucosal HIV invasion sites to global HIV infections in adults.

| **HIV invasion site** | **Anatomic sublocation** | **Type of epithelium** | **Transmission medium** | **Transmission probability per exposure event** | **Estimated contribution to HIV cases worldwide, $n$ (millions)** |
|---|---|---|---|---|---|
| Female genital tract | Vagina | Squamous, non-keratinized | Semen | 1 in 200 to 1 in 2000 | 12.6 |
| | Ectocervix | Squamous, non-keratinized | | | |
| | Endocervix | Columnar, single layer | | | |
| | Other | Various epithelia | | | |
| Male genital tract | Inner foreskin | Squamous, poorly keratinized | Cervicovaginal and rectal secretions and desquamations | 1 in 700 to 1 in 3000 | 10.2 |
| | Penile urethra | Columnar, stratified | | | |
| | Other | Various epithelia | | | |
| Intestinal tract | Rectum | Columnar, single layer | Semen | 1 in 20 to 1 in 300 | 3.9 |
| | Upper gastrointestinal tract | Various epithelia | | 1 in 2500 | 1.5 |

Human vagina and ectocervix are covered by non-keratinized squamous epithelium, which can be abraded by shearing during sexual intercourse. The subepithelial layer of the genital mucosa is a very favourable environment for HIV replication, since it is enriched with DCs, macrophages, and T cells that expresses CD4, CCR5, and in lesser quantities, CXCR4 receptors, which make them vulnerable to HIV infection. There are several pathways for HIV virions, either free or infected donor cell-associated, to traverse the surface of female genital mucosa. Once the virus traverses the genital mucosa it may access the vulnerable target cells in the basal epithelium and the underlying stroma, i.e. LCs in epithelium, stromal DCs located beneath epithelium, T cells and macrophages in stroma [43].

As illustrated in figure 2.3, HIV virions or HIV-infected donor cells can be trapped in the mucus (fig. 2.3A), which then may lead to HIV-infected donor cell attachment to the luminal surface of the mucosa and the attached cell may start to secret virions (fig. 2.3B). The virions can further penetrate between epithelial cells (fig. 2.3C), where the residing LCs may capture the virions (fig. 2.3D) and internalize them to endocytic compartments (fig. 2.3E). Equally, the penetrating virions may fuse with intraepithelial $CD4^{+}$ T lymphocytes and cause the productive infection (fig. 2.3F). The penetrating virions may also follow the transcytosis through epithelial cells (fig. 2.3G) and result to the productive infection of the basal epithelial cells (fig. 2.3H) or internalization into their endocytic compartments (fig. 2.3I). If not trapped to the mucus, the infected donor cells (fig. 2.3J) and free virions (fig. 2.3K) may migrate along the abrasions into the vulnerable mucosal stroma.

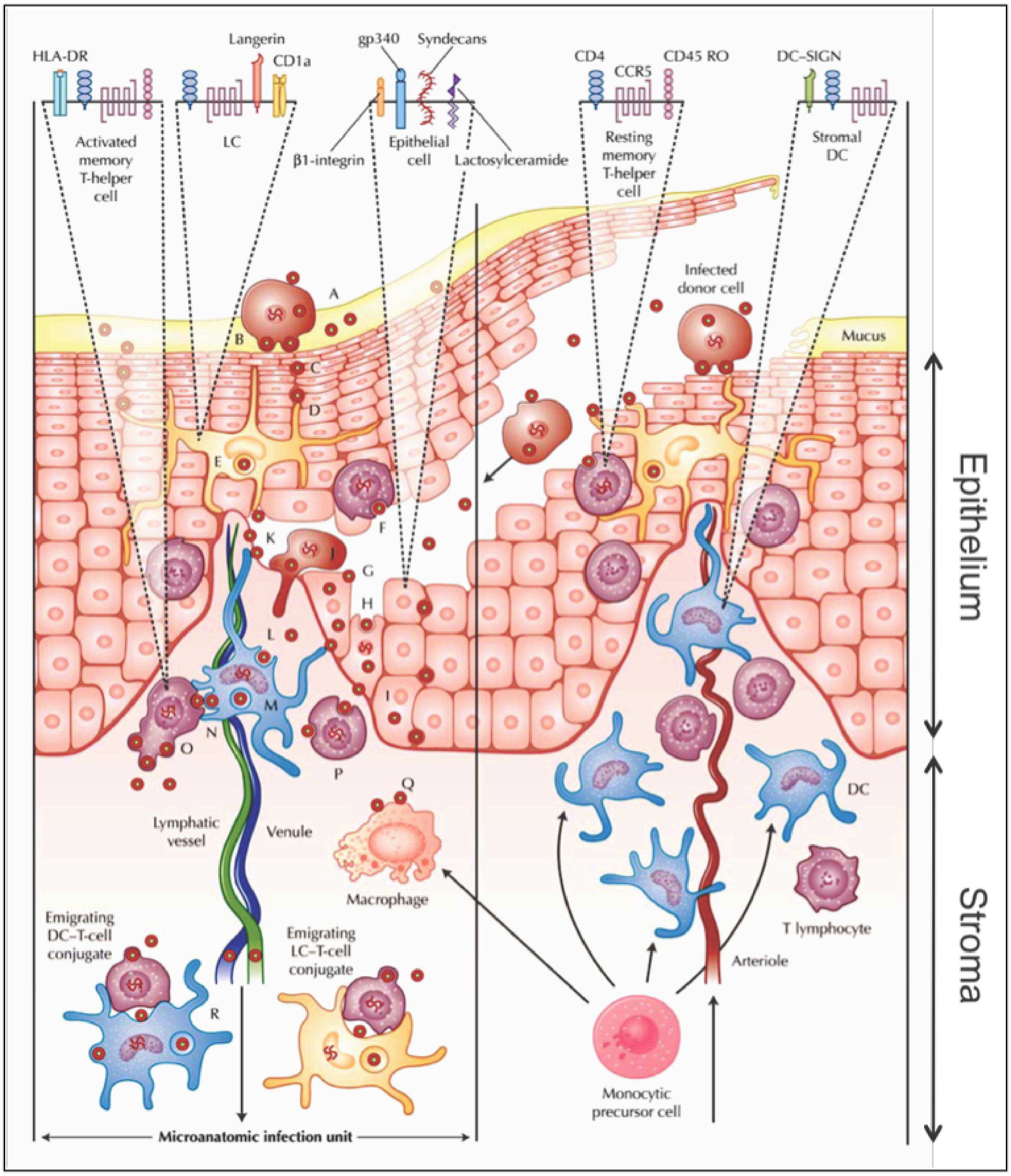

HLA-DR
Langerin
CD1a
gp340
Syndecans
CD4
CCR5
CD45 RO
DC-SIGN
Activated memory T-helper cell
LC
β1-integrin
Epithelial cell
Lactosylceramide
Resting memory T-helper cell
Stromal DC
Infected donor cell
Mucus
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
Epithelium
Stroma
DC
Lymphatic vessel
Venule
Macrophage
T lymphocyte
Arteriole
Emigrating DC–T-cell conjugate
Emigrating LC–T-cell conjugate
Monocytic precursor cell
Microanatomic infection unit

**Figure 2.3.:** HIV invasion pathways in the female lower genital tract mucosa.
The stromal papilla on the right illustrates a capillary through which the blood cells are supplied to the mucosa where the monocytes differentiate to DCs or macrophages. The left side depicts an uncovered papilla, a possible site of infection, resulting from a sharing of the outer epithelium during sexual intercourse. On the top of the figure, HIV receptors and some phenotypic cell receptors are shown. (From[44])

Once HIV virons reach the mucosal stroma, they can cause the productive infection of stromal DCs (fig. 2.3L) or the virions can be internalized to endocytic compartments of DCs (fig. 2.3M). The infected stromal DCs can transfer virus to CD4$^{+}$ T cells across the infectious synapse (fig. 2.3N), which then results to a massive productive infection of mucosal CD4$^{+}$ T cells (fig. 2.3O) and a productive infection of resting mucosal CD4$^{+}$ memory T cells (fig. 2.3P). Productively infected

$CD4^+$ T cells and DCs or their conjugates can migrate to submucosa and the draining lymphatic and venous microvessels, and thus disseminate the virus (fig. 2.3R). HIV virions may also bind and possibly infect the stromal macrophages (fig. 2.3Q) [44].

#### **2.2.1.1.** *Stromal DCs in mucosa are implicated in DC-SIGN-mediated HIV infection enhancement*

DCs play particularly important role in HIV-1 infection enhancement and dissemination in epithelia since the virus hijacks them to achieve the productive infection of the $CD4^+$ T cells and the burst of the disease [41, 42]. The stromal DCs that express a C-type lectin DC-SIGN (DC-Specific ICAM3 Grabbing Non-integrin) and reside in mucosa of vagina and ectocervix, have been repeatedly reported to be exploited by HIV-1 to enhance its infectivity of T cells. Different mechanisms of how DCs augment *de novo* infection of T cells were evaluated and suggested by several research groups [43]. Electron microscopy 3D structural studies of the infectious synapse have shed light on this mechanism (fig. 2.4): DCs engulf the surrounding extracellular environment, trapping virions in a surface-accessible but protected compartment. Then $CD4^+$ T cells may contact DCs by extending their membrane protrusions (filopodia, cytoneme (fig. 2.5)), enriched for CD4 and coreceptor, into the invaginated DC compartments that contain bound virions, allowing efficient virion transmission from DCs to $CD4^+$T cells [45].

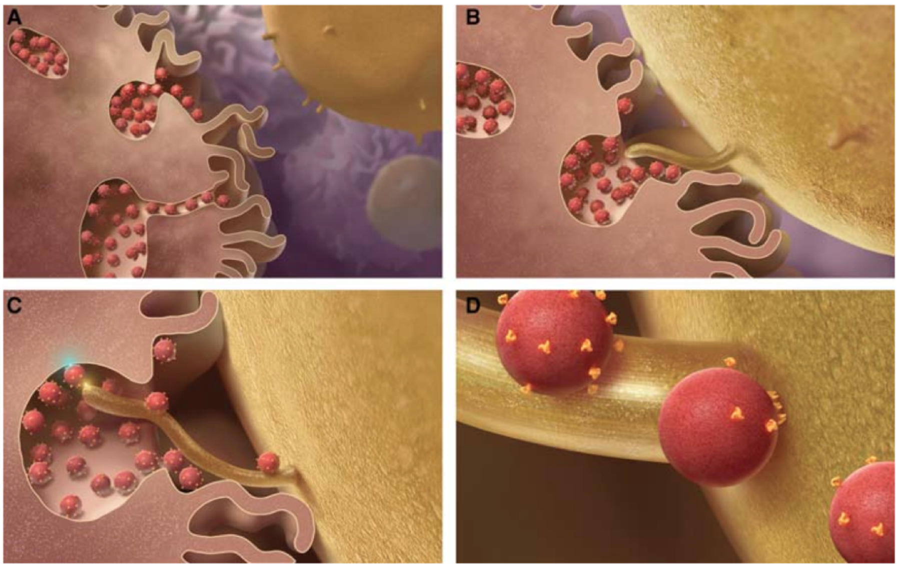


**Figure 2.4.:** Model of HIV transfer from DCs to $CD4^+$ T cells.
**A**, HIV virions are trapped and concentrated in surface-accessible trypsin-resistand compartments of DCs. **B**, DCs and $CD4^+$T cells containing membrane protrusions, come into contact. **C**, $CD4^+$T cell protrusions invade the virus-containing compartments and efficiently bind HIV virions. **D**, HIV virions migrate toward T cell body to initiate infection. (From [45])

The above described *trans* infection mechanism via infectious synapse is only one of the possible trans infection pathways. There are several parallel mechanisms of this process (fig. 2.5), and it must be noticed that DC-SIGN has a very important role in DC-mediated HIV infection enhancement [36]. Following the virion binding to DC-SIGN, HIV can be endocytosed into DCs, and subsequently the intact virions can be stored in multivesicular bodies (fig. 2.5(1)) [36, 46]. HIV can also be stored as an integrated provirus following productive infection of DCs (fig. 2.5(2)). Likewise, HIV can be stored as DC-SIGN-bound virions on the cell surface and protected from degradation (fig. 2.5(3)) [47]. During the formation of the infectious synapse either case leads to accumulation of intact virions on DC side while HIV receptors (CD4, CCR5) are presented on CD4$^+$ T cell side. Such a situation greatly facilitates HIV-1 passage from DCs to T cells. Indeed, it has been demonstrated that blocking DC-SIGN prevents HIV-1 binding and subsequent *trans* infection of CD4$^+$ T cells [48, 49, 50].

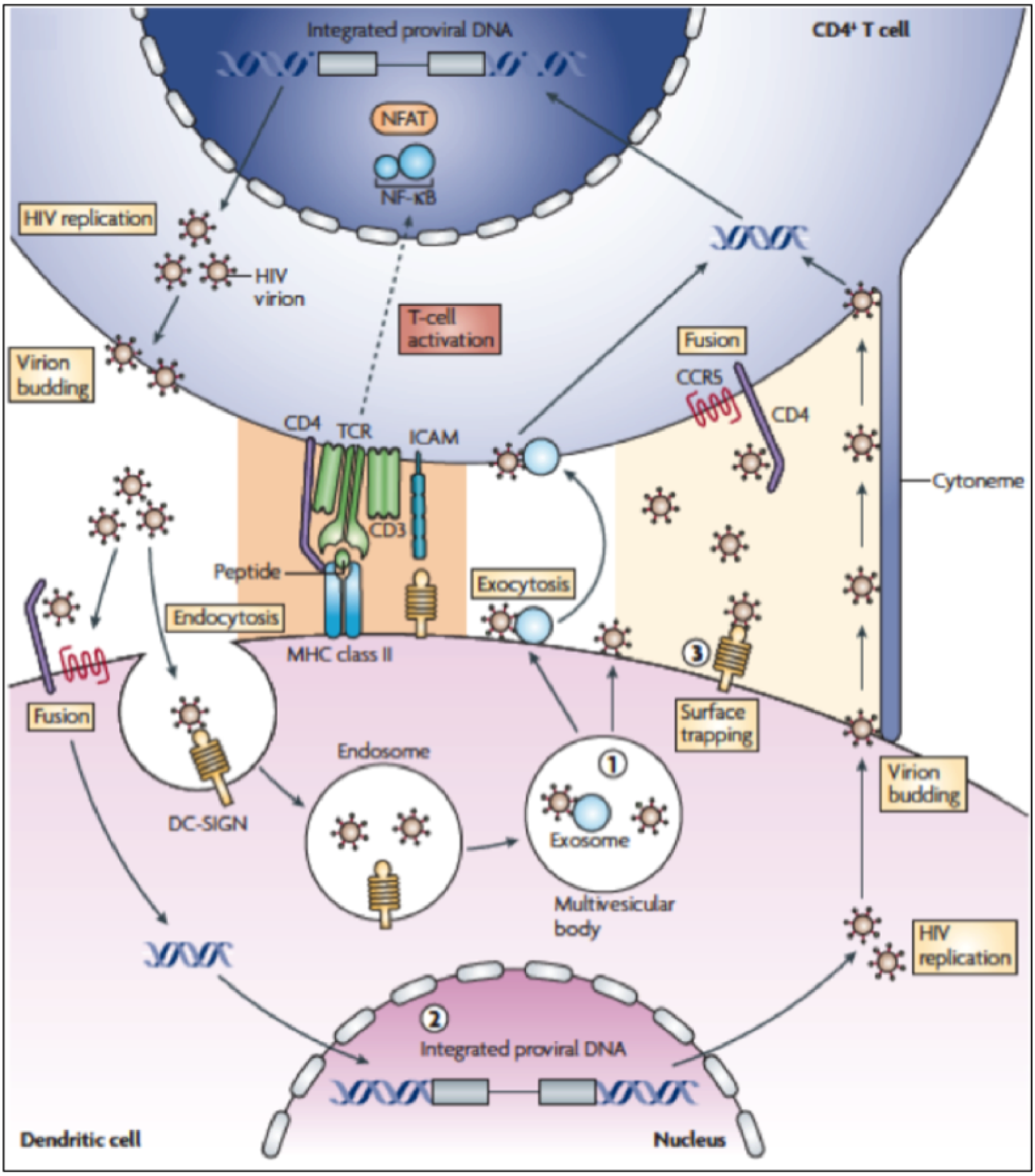

Integrated proviral DNA
CD4⁺ T cell
NFAT
NF-κB
HIV replication
HIV virion
T-cell activation
Virion budding
Fusion
CCR5
CD4
CD4
TCR
ICAM
CD3
Peptide
MHC class II
Endocytosis
Exocytosis
Cytoneme
Fusion
DC-SIGN
Endosome
①
Exosome
Multivesicular body
③
Surface trapping
Virion budding
HIV replication
②
Integrated proviral DNA
Dendritic cell
Nucleus

**Figure 2.5.:** The significance of DC–T-cell interactions and DC-SIGN implication in HIV-1 transmission.

(From [43])

While *trans* infection is responsible for the early stage of infection (24 h after HIV exposure), there is a different pathway of DC-SIGN-bound HIV transmission to T lymphocytes. This pathway is involved in long-term HIV transfer (72 h after exposure) and it occurs as an infection in *cis* of DCs by transfer of DC-SIGN-bound virus to canonical HIV entry receptors, CD4 and CCR5, which leads to productive infection of DCs and in turn presentation of increased viral load to the T cells [51, 52]. Furthermore, DC-SIGN is not only hijacked by virus to enter the DCs, but HIV also exploits DC-SIGN signaling for the recruitment of host transcription-elongation factors to produce viral transcription-elongation factor Tat, which is essential for further transcription of viral genes, and without DC-SIGN signaling only short abortive mRNAs are produced [53].

#### 2.2.1.2. *Suppressive role of epithelial LCs in HIV transmission*

By contrast, another subset of dendritic cells, the LCs, plays an important preventive role in HIV infection process. LCs express HIV receptors including CD4, CCR5 and C-type lectin langerin [54, 55, 56]. It has been shown that epidermal LCs expressing langerin efficiently bind HIV virions (fig. 2.6A), which in turn are directed to Birbeck granules (BG) for degradation (fig. 2.6B) [57]. This shows an important protective role of langerin in HIV invasion process.

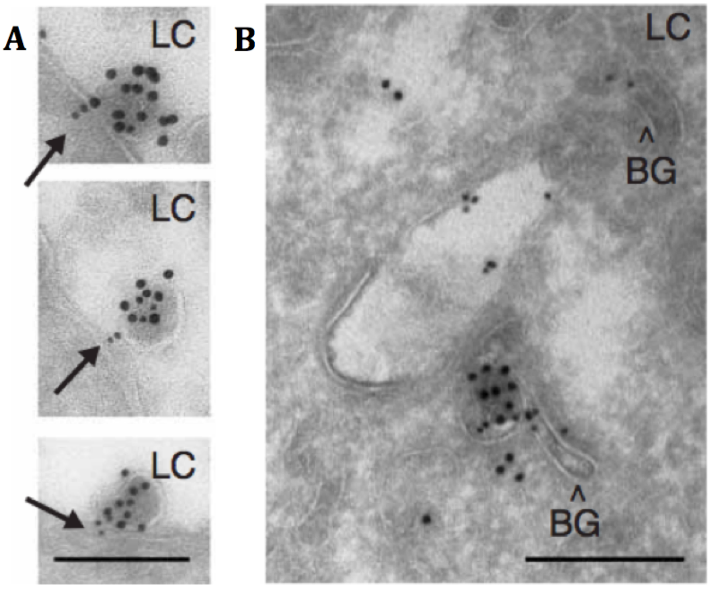


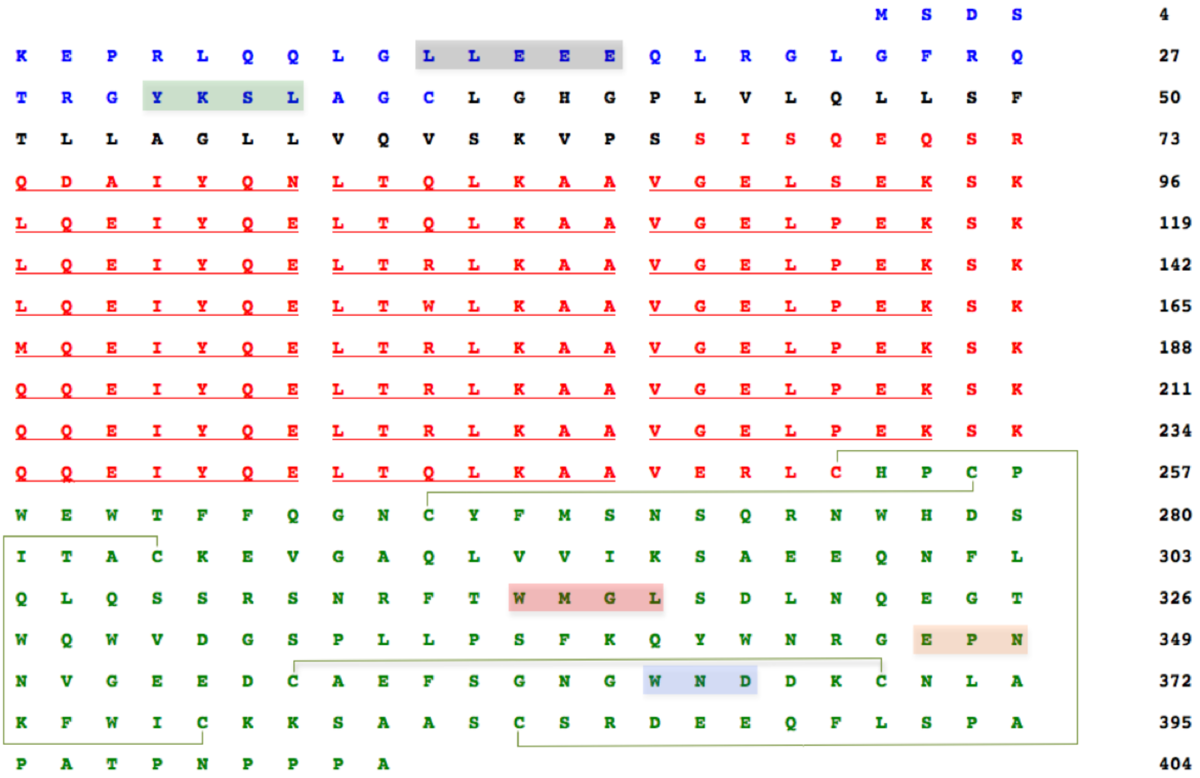
**Figure 2.6.:** The micrograph showing an uptake of HIV-1 virions and their internalization to Birbeck granules by immature LCs.

**A**, HIV-1 capture by langerin (arrows) expressed on LCs. **B**, HIV-1 internalization into Birbeck granules (arrowheads). HIV-1 p24 capsid and langerin staining are labeled by 15 nm and 10 nm gold, respectively; scale bar 100 nm. (From [57])

The importance of langerin for the protective function of LCs can be seen very well in the examples where the impaired presence/expression of langerin subverts this function of LCs. Firstly, pre-incubation of LCs with anti-langerin antibody 10E2 or mannan as well as high viral loads resulting in langerin saturation abrogated the LCs' antiviral function against HIV as the virus was transmitted to T cells [57]. The coinfection of mucosa with HSV-2 and HIV-1 increases the susceptibility to the latter infection because HSV-2 decreases the surface concentration of langerin due to competition with HIV-1 for binding to langerin or through TLR-3 induced signaling, which down-regulates expression of langerin [58]. The recent study showed that vaginal LCs may have low or no expression of langerin, and thus they are susceptible to HIV infection [59]. Hence two CLRs, DC-SIGN and langerin, despite the high homology of their CRDs and the overlap of ligand specificities [60] have very different roles in HIV invasion: while DC-SIGN promotes HIV dissemination and infection, langerin helps to prevent HIV invasion.

## 2.3. Structure and function of C-type lectin receptor DC-SIGN

DC-SIGN is a type II membrane protein with a single CRD, i.e. a type II C-type lectin. It was first cloned and identified as a C-type lectin binding to HIV glycoprotein gp120 by Curtis *et al* in 1992 [61] and later rediscovered and renamed by Geijtenbeek *et al* in 2000 when screening a library of DC-specific monoclonal antibodies that inhibit binding to ICAM-3 (InterCellular Adhesion Molecule-3), the T cells activating adhesion molecule [62]. DC-SIGN is present at low levels on blood monocytes and as they leave blood to peripheral tissues and differentiate to DCs the expression of this lectin is extremely increased. Finally, following DC maturation and migration to secondary lymphoid tissues, DC-SIGN levels are down-regulated [62]. DC-SIGN expression is restricted to certain DC subsets, i.e. dermal and stromal DCs, and it is not expressed on follicular DCs or on skin-resident LCs. Other cell types that express DC-SIGN include macrophages [63, 64, 65, 66] and activated B cells [67]. Moreover, there are reports showing the presence of soluble forms of DC-SIGN resulting from alternative splicing [68, 69, 70].



| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | | | | | M | S | D | S | 4 |
| K | E | P | R | L | Q | Q | L | G | L | L | E | E | E | Q | L | R | G | L | G | F | R | Q | 27 |
| T | R | G | Y | K | S | L | A | G | C | L | G | H | G | P | L | V | L | Q | L | L | S | F | 50 |
| T | L | L | A | G | L | L | V | Q | V | S | K | V | P | S | S | I | S | Q | E | Q | S | R | 73 |
| Q | D | A | I | Y | Q | N | L | T | Q | L | K | A | A | V | G | E | L | S | E | K | S | K | 96 |
| L | Q | E | I | Y | Q | E | L | T | Q | L | K | A | A | V | G | E | L | P | E | K | S | K | 119 |
| L | Q | E | I | Y | Q | E | L | T | R | L | K | A | A | V | G | E | L | P | E | K | S | K | 142 |
| L | Q | E | I | Y | Q | E | L | T | W | L | K | A | A | V | G | E | L | P | E | K | S | K | 165 |
| M | Q | E | I | Y | Q | E | L | T | R | L | K | A | A | V | G | E | L | P | E | K | S | K | 188 |
| Q | Q | E | I | Y | Q | E | L | T | R | L | K | A | A | V | G | E | L | P | E | K | S | K | 211 |
| Q | Q | E | I | Y | Q | E | L | T | R | L | K | A | A | V | G | E | L | P | E | K | S | K | 234 |
| Q | Q | E | I | Y | Q | E | L | T | Q | L | K | A | A | V | E | R | L | C | H | P | C | P | 257 |
| W | E | W | T | F | F | Q | G | N | C | Y | F | M | S | N | S | Q | R | N | W | H | D | S | 280 |
| I | T | A | C | K | E | V | G | A | Q | L | V | V | I | K | S | A | E | E | Q | N | F | L | 303 |
| Q | L | Q | S | S | R | S | N | R | F | T | W | M | G | L | S | D | L | N | Q | E | G | T | 326 |
| W | Q | W | V | D | G | S | P | L | L | P | S | F | K | Q | Y | W | N | R | G | E | P | N | 349 |
| N | V | G | E | E | D | C | A | E | F | S | G | N | G | W | N | D | D | K | C | N | L | A | 372 |
| K | F | W | I | C | K | K | S | A | A | S | C | S | R | D | E | E | Q | F | L | S | P | A | 395 |
| P | A | T | P | N | P | P | P | A | | | | | | | | | | | | | | | 404 |

**Figure 2.7.:** Primary structure of human DC-SIGN.
The sequences of cytosolic, transmembrane, neck region, and CRD domains are in blue, black, red and green letters. The internalization motif and incomplete ITAM are highlighted in grey and green, and the conserved motifs defining sugar specificity and $Ca^{2+}$ coordination are in orange and blue. The third conserved motif, WMGL instead of WIGL in case of DC-SIGN, is highlighted in red. The amino acid repeats (heptades) contributing to coiled-coil formation are individually underlined, and disulphide bridges are depicted by light green lines.

The gene encoding DC-SIGN (CD209) is located on human chromosome 19p13.2-3, contains 7 exons and 6 introns, and is closely related to other CLR genes such as DC-SIGNR and CD23. The gene encodes 404 amino acids containing protein (fig. 2.7) of 44 kDa in molecular weight [71].

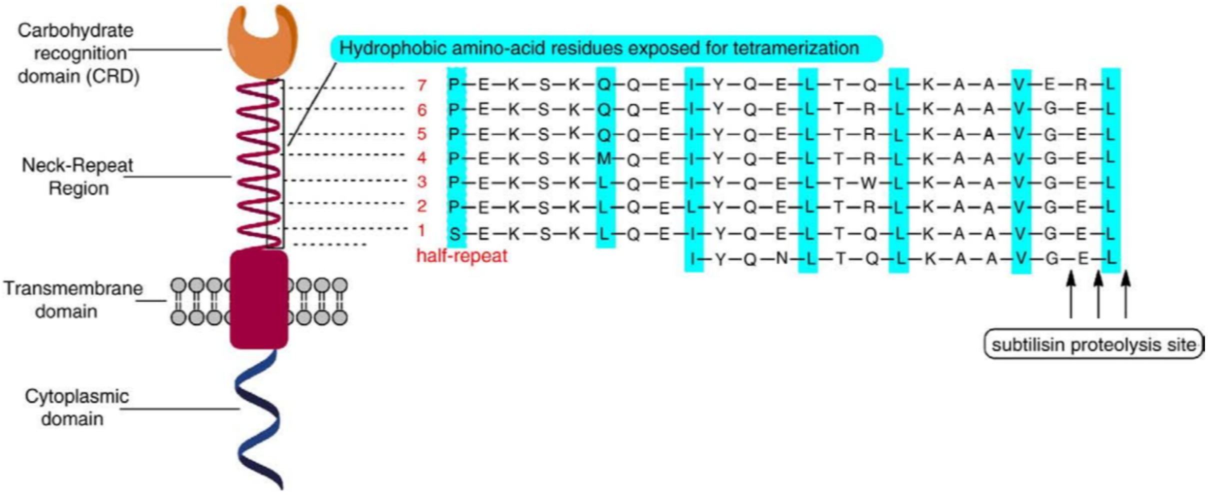

Carbohydrate recognition domain (CRD)
Hydrophobic amino-acid residues exposed for tetramerization
Neck-Repeat Region
7 P—E—K—S—K—Q—Q—E—I—Y—Q—E—L—T—Q—L—K—A—A—V—E—R—L
6 P—E—K—S—K—Q—Q—E—I—Y—Q—E—L—T—R—L—K—A—A—V—G—E—L
5 P—E—K—S—K—Q—Q—E—I—Y—Q—E—L—T—R—L—K—A—A—V—G—E—L
4 P—E—K—S—K—M—Q—E—I—Y—Q—E—L—T—R—L—K—A—A—V—G—E—L
3 P—E—K—S—K—L—Q—E—I—Y—Q—E—L—T—W—L—K—A—A—V—G—E—L
2 P—E—K—S—K—L—Q—E—L—Y—Q—E—L—T—R—L—K—A—A—V—G—E—L
1 S—E—K—S—K—L—Q—E—I—Y—Q—E—L—T—Q—L—K—A—A—V—G—E—L
half-repeat
I—Y—Q—N—L—T—Q—L—K—A—A—V—G—E—L
Transmembrane domain
subtilisin proteolysis site
Cytoplasmic domain

**Figure 2.8.:** Domain organization of DC-SIGN molecule.

(From [72])

DC-SIGN is composed of four domains: the N-terminal cytoplasmic tail, transmembrane domain followed by the extracellular part composed of a neck region, which contains 7 and a half repeats of 23 amino acid residues, and the C-terminal carbohydrate recognition domain (fig. 2.7 and 2.8) [61].

### 2.3.1. *DC-SIGN CRD structure and carbohydrate recognition*

The $Ca^{2+}$-dependent CRD of DC-SIGN (fig. 2.9A) has a typical long-form C-type lectin fold sustained by four disulphide bridges, contains three occupied $Ca^{2+}$-binding sites (1, 2, and 3) and EPN motif, thus the specificity for mannose group sugars. It binds mannose residues through equatorial 3-OH and 4-OH groups that are coordinated to $Ca^{2+}$ at site 2, and the overall binding is maintained by the hydrogen bond network formed between distinct sugar groups and amino acid residues in the protein, which can be mediated or not by water molecules (fig. 2.9B). The interaction with high mannose ligands is not limited to terminal mannose residues and rather involves internal residues that stretch along the extended binding site (fig. 2.9C) [73].

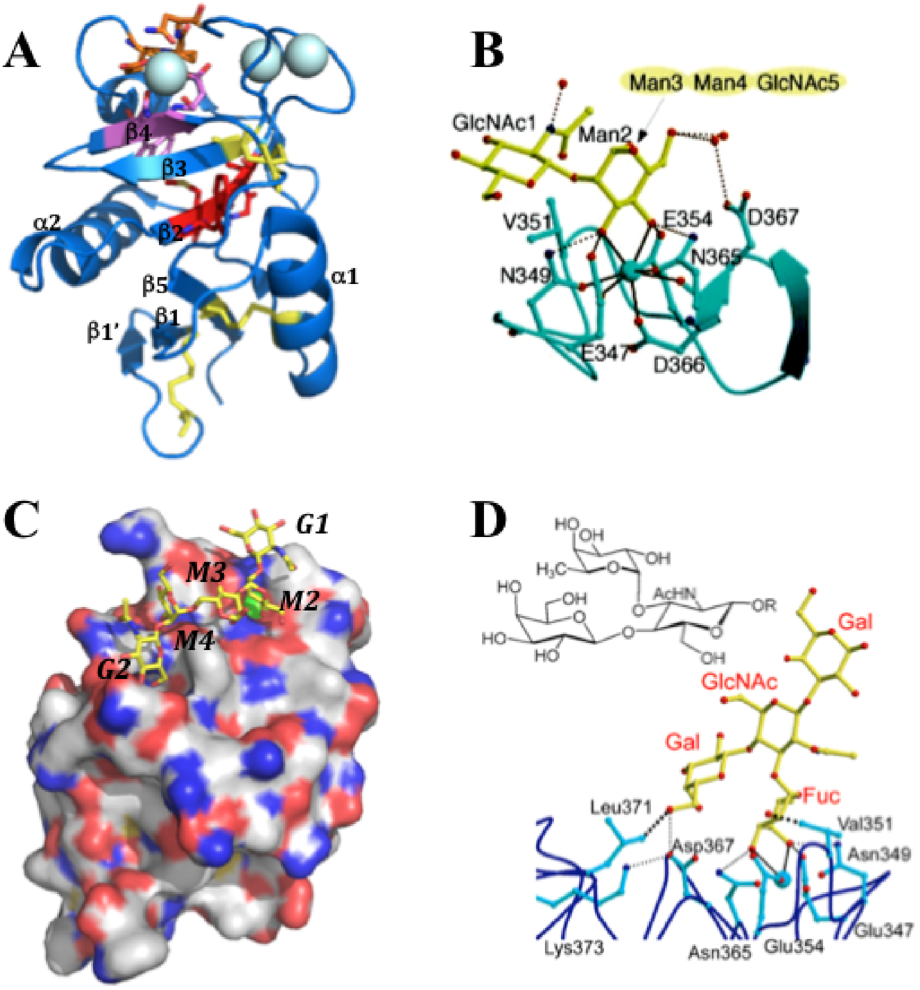

A
β4
β3
α2
β2
β5
α1
β1'
β1
B
Man3 Man4 GlcNAc5
GlcNAc1
Man2
V351
E354
D367
N365
N349
E347
D366
C
G1
M3
M2
M4
G2
D
HO
OH
OH
H₃C
O
HO
OH
AcHN
OR
HO
HO
OH
Gal
GlcNAc
Gal
Fuc
Leu371
Val351
Asp367
Asn349
Glu347
Lys373
Asn365
Glu354

**Figure 2.9.:** Carbohydrate recognition by DC-SIGN CRD.

**A**, the structure of DC-SIGN CRD: EPN, WND and WMGL motifs are shown in orange, pink and red stick representation; disulphide bridges depicted in yellow sticks and large cyan spheres are the three $Ca^{2+}$. **B**, binding of mannose residue in the CRD, the coordination bonds to $Ca^{2+}$ (cyan sphere) are black solid lines and the H-bonds are represented as black dotted lines. **C**, Surface representation of DC-SIGN CRD bound to $GlcNAc_2Man_3$ (oxygens and nitrogens are in red and blue, carbons of protein and of oligosaccharide are in white and yellow, $Ca^{2+}$ is green, $GlcNAc_2Man_3$ is in stick representation). **D**, Lewis$^X$ trisaccharide (upper panel) and a structure of lacto-N-fucopentaose bound to DC-SIGN CRD (lower panel). (**A** and **B** are from [73]; **C** is pdb:1k9i; **D** is from [74])

DC-SIGN binds N-acetylglucosamine (GlcNAc) also through equatorial 3-OH and 4-OH groups [73], and recognizes fucosylated structures, including Lewis-type blood group antigens [74]. In this case, fucose residue is bound in Ca-binding site 2 via equatorial 3-OH and axial 4-OH groups, and additional tight van de Waals contacts are formed between 2-OH group of fucose and Val351 residue of the protein (fig. 2.9D). While GlcNAc residue of Lewis$^X$ points away from the protein, the terminal galactose residue interacts with the CRD at a secondary binding site through H-bonds and van der Waals contacts with protein side chains. The observed orientation of Lewis$^X$ bound to DC-SIGN CRD suggests that other fucose-containing oligosaccharides may bind DC-SIGN primarily through fucose in $Ca^{2+}$-binding site 2 and make additional contacts with the protein at other sites [74].

### 2.3.2. *Oligomeric structure of DC-SIGN*

The affinity of the sugars for DC-SIGN CRD is low ($K_\mathrm{D}$ in mM range). However, the oligomerization of the lectin to a tetrameric form afforded by the neck domain [75] and the heavy glycosylation of its ligands gives a way for high affinity ($K_\mathrm{D}$ in nM range) interactions due to the avidity phenomenon (fig. 2.10). The other functions that can be attributed to the neck include the projection of the CRDs further away from the cell membrane for appropriate ligand scanning and binding. The certain positioning of the CRDs due to oligomerization provides additional means of the specificity to the ligands [76].

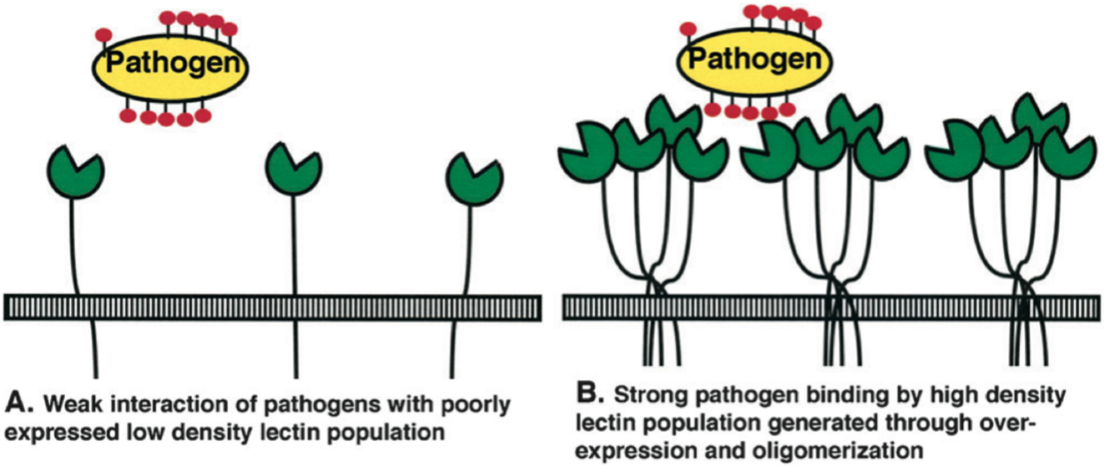

Pathogen
Pathogen
A. Weak interaction of pathogens with poorly expressed low density lectin population
B. Strong pathogen binding by high density lectin population generated through over-expression and oligomerization

**Figure 2.10.:** Schematic representation of lectin/pathogen interaction affinity increase due to the avidity phenomenon.

**A**, Low density of monomeric lectin results in weak interactions. **B**, Lectin binding site clustering due to oligomerization and overexpression leads to strong pathogen binding. (From [77])

Oligomerization of DC-SIGN to tetramers is driven by intermolecular hydrophobic interactions of amino acid residues of the heptad repeats (fig. 2.7 and 2.8) in the neck domains and supported by the lateral salt bridges (fig. 2.11A): in the heptads, represented as *abcdefg*, *a* and *d* are hydrophobic, and *e* and *g* are charged residues [78].

The studies of truncated DC-SIGN forms with variable length of neck region have shown that at least 6 repeats are needed for the tetramerization, the 5.5 repeats result an equilibrium between tetramers and dimers, and a neck of only 2 repeats leads to equilibrium between dimers and monomers [79]. Moreover, the studies by the group, which I later joined to carry out my PhD, have demonstrated that the state and stability of oligomerization is pH-dependent: the decrease of pH induces dissociation of the tetramers into monomers with an approximate mid-point of transition at about pH 5.9. It was also observed that the decrease of pH from mild alkaline (pH 7.9) to slight acidic (pH 5.9) induces the change of the shape of DC-SIGN tetramers from elongated to more compact [80].

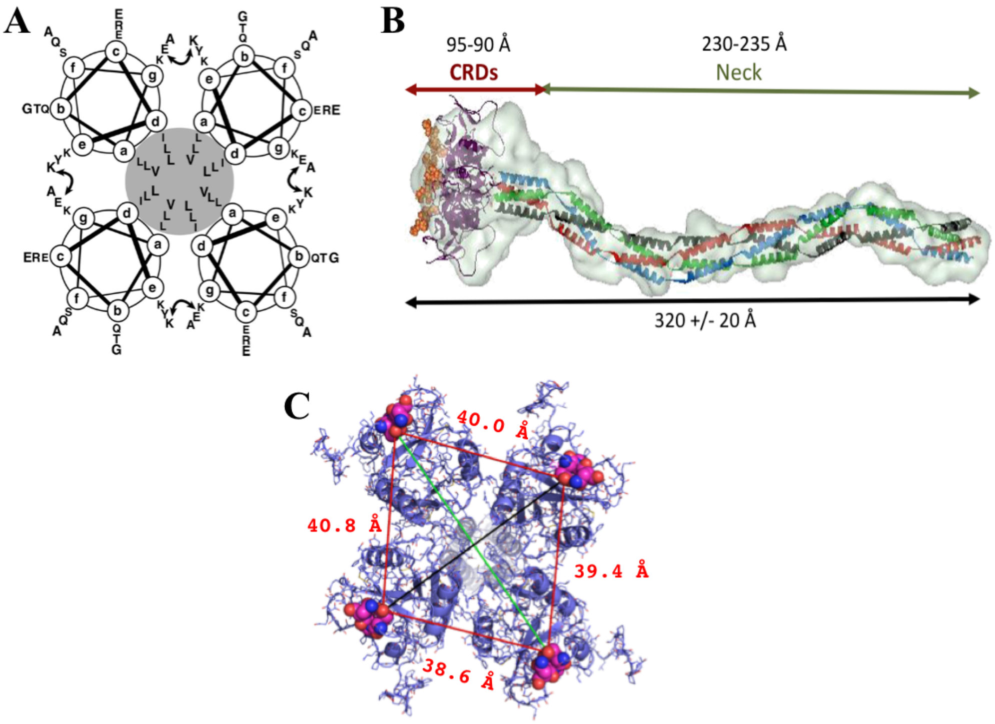


**Figure 2.11.:** The structure of the extracellular part of DC-SIGN.
**A**, a schematic representation of the coiled-coil interactions in the neck domain: each of the four circles represents a top-view of an α-helix formed by two 23 amino acid repeats, the hydrophobic core is highlighted in grey, and the curved double arrow represents the putative interchain salt bridge between the glutamic acid (position $g$) of the first heptad and the lysine residue (position $e$) of the second one. **B**, SAXS envelop of DC-SIGN extracellular domain. **C**, superimposition of DC-SIGN CRDs (pdb:1k9i) into SAXS envelop: the view from the top. The distances between vicinal sugar-binding sites are indicated in red (measurement between the same oxygen atom of Gln147 within EPN motif, depicted as spheres), and diagonal distances in black and green are 51.1 and 61.0 Å, respectively. (Adapted from [80])

The overall length of the extracellular part of the lectin was found to be around 32 nm with a certain degree of flexibility observed when binding to the ligand (the length decreases to 28 nm) [81]. This plasticity of the tetramers is likely to be furnished by a flexible CRD-neck junction [79], and allows to adapt to the glycan structures on the ligands, thus enabling all CRDs to interact. The global shape of the extracellular part of DC-SIGN revealed by small angle X-ray scattering (SAXS) measurements revealed and confirmed the presumed elongated shape of the lectin with an approximate 40 Å spacing between vicinal carbohydrate binding sites (fig. 2.11B and C) [80].

### 2.3.3. *Clustering of DC-SIGN tetramers on the surface of DCs*

Regarding the organization of DC-SIGN molecules at the cell membrane level, the lectin is clustered into microdomains in lipid rafts of the plasma membrane of DCs, and also when ectopically expressed in fibroblasts and other cell types. This clustering occurs even in the absence of the ligands, and the formed microdomains range from $<$200 nm to 1.5 μm as revealed by different techniques [82, 83, 84]. Moreover, they are remarkably stable in terms of location and diffusion of the receptor within the microdomain, as well as exchange between the microdomain and its surroundings in the timescale of several minutes [85, 86]. The recent studies of the nanostructure of DC-SIGN microdomains by super-resolution imaging technique, Blink Microscopy, indicated that DC-SIGN is organized in small ≈80 nm nanodomains (fig. 2.12), which are randomly distributed on the plasma membrane of DCs and contain as few as 1-3 tetramers as a lower limit, suggesting that other proteins and lipids occupy these nanodomains [87].

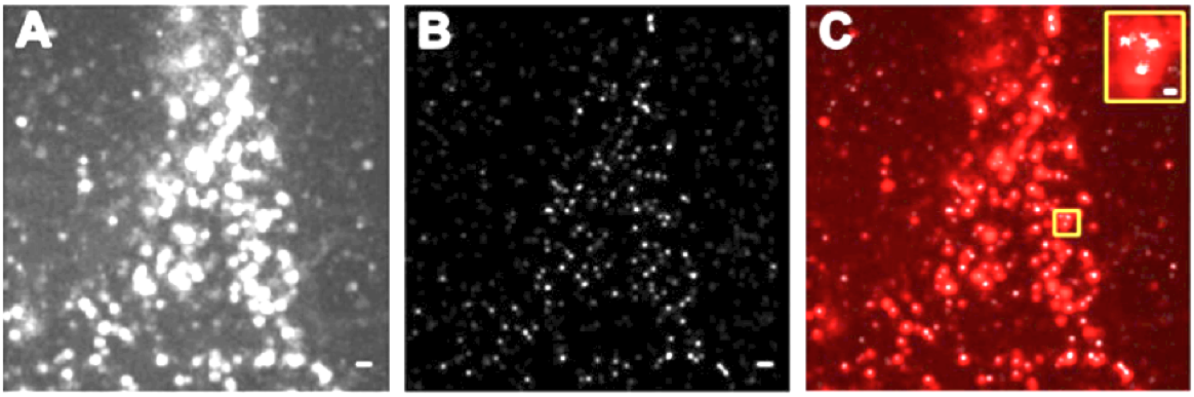


**Figure 2.12.:** The nanostructure of DC-SIGN microdomains.
**A**, DC-SIGN expression on fixed DC revealed by diffraction limited total internal reflection fluorescence (TIRF) illumination. **B**, Super-resolution Blink image of DC-SIGN expression in the same region as in **A**. **C**, superimposition of TIRF image (in red) and Blink image (in white). Scale bars, 500 nm; scale bar inset, 100 nm. (From [87])

The mechanism of DC-SIGN clustering on the membrane was investigated using truncated forms of DC-SIGN and the influence of distinct DC-SIGN parts for the formation and stability of microdomains was examined [88]. These studies indicated that the driving force of microdomain formation is the saccharide binding to CRD, since CRD-depleted DC-SIGN or mutants of a single amino acid required for $Ca^{2+}$-coordination at sugar binding site showed a diffused expression over all the membrane with no visible microdomains and high mobility. The neck domain appeared to be responsible for the stability of the microdomains, as in case of DC-SIGN depleted of tandem repeats the microdomains were dynamic, often moving on membrane from one position to another, and DC-SIGN within these domains was able to exchange with other DC-SIGN molecules from the surrounding membrane. By contrast, the removal of whole cytoplasmic domain and N-glycosylation site at Asn80 had no significant effect on microdomain formation and stability suggesting that the clustering mechanism is not dependent on cytoskeletal structures, despite the observation that DC-SIGN constitutively binds to scaffold proteins LSP1, KSR1 and CNK [89].

The finding that CRD is essential for the clustering of DC-SIGN on the membrane suggests several mechanisms: (1) CRD directly binds to polysaccharides of the extracellular matrix (ECM), e.g. glycosaminoglycans or glycosyl moieties of ECM proteins (fig 2.13A); (2) CRD directly binds to transmembrane proteoglycans that link to the ECM (fig 2.13B); (3) glycosylated transmembrane adaptor proteins (TRAPs), which are directly or indirectly linked to the membrane-apposed cytoskeleton and have cis interactions with DC-SIGN, serve to stabilize the microdomains (fig 2.13C).

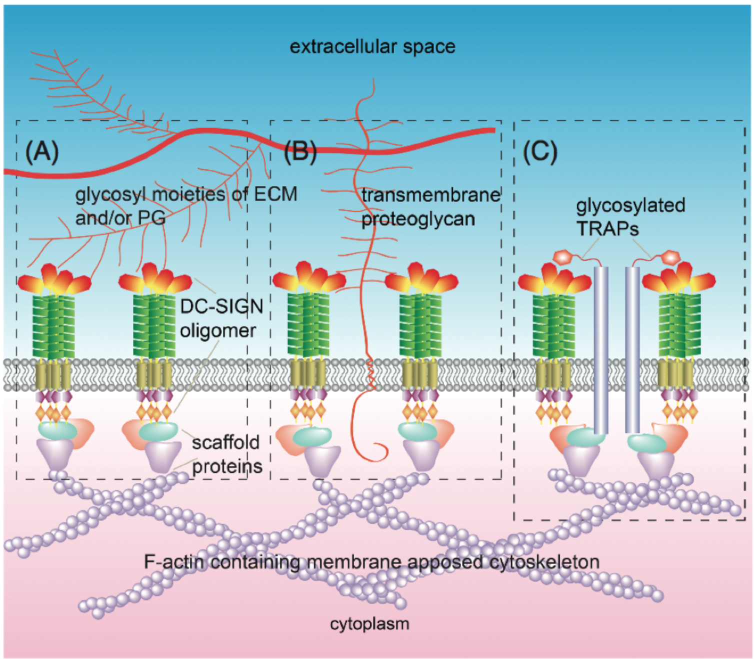

extracellular space
(A)
(B)
(C)
glycosyl moieties of ECM and/or PG
transmembrane proteoglycan
glycosylated TRAPs
DC-SIGN oligomer
scaffold proteins
F-actin containing membrane apposed cytoskeleton
cytoplasm

**Figure 2.13.:** Possible mechanisms of DC-SIGN microdomain formation and stability. Microdomain formation facilitated and stabilized through DC-SIGN CRDs interactions with ECM components (**A**), transmembrane proteoglycans that are linked to ECM components (**B**), and/or glycosylated TRAPs, having cis interactions with DC-SIGN, link the receptor to the cytoskeleton (**C**). (From [88])

Such clustering of DC-SIGN into microdomains further contributes to the avidity of DC-SIGN interactions with its ligands. Furthermore, it has been proposed that such clustering may modulate DC-SIGN specificity by favoring carbohydrates with a certain density and spacing [90].

## 2.3.4. *Ligand recognition by DC-SIGN*

### 2.3.4.1. *Endogeneous and exogeneous ligands of DC-SIGN.*

DC-SIGN has been discovered as a C-type lectin of DCs that binds to the T cells adhesion molecules ICAM-3, and this interaction is believed to mediate transient, antigen nonspecific adhesion of DCs with T cells (fig. 2.14A), leading to effective T cell receptor engagement and screening of the MHC-peptide complexes [62]. DC-SIGN binds ICAM-3 through the recognition of the N-glycans of high mannose-type oligosaccharides and, depending on the cell population, Lewis$^X$ residues [91]. ICAM-2 that is expressed on endothelial cells was identified as another endogenous counterstructure of DC-SIGN [92]. The interaction of these two molecules resists shear stresses and is involved in DC adhesion, tethering, and rolling along the ICAM-2 expressing surfaces (fig. 2.14A). DC-SIGN binding to ICAM-2 was shown to be strictly glycan-specific and Lewis$^Y$ structures on ICAM-2 were identified as the glycans possibly mediating this interaction [93]. Additionally to the adhesion molecules ICAM-2 and ICAM-3, DC-SIGN binds to Lewis$^X$ expressed on β2-integrin Mac-1 and CEACAM1 on neutrophils (PMNs), and thus drives the interaction between DCs and neutrophils (fig. 2.14B) [94, 95], and by binding to Lewis$^X$ and Lewis$^Y$ sugars recognizes carcinoembryonic antigen (CEA) overexpressed on majority of colorectal cancer cells [96]. DC-SIGN also recognizes human semen clusterin and this binding interferes HIV capture by DCs [97].

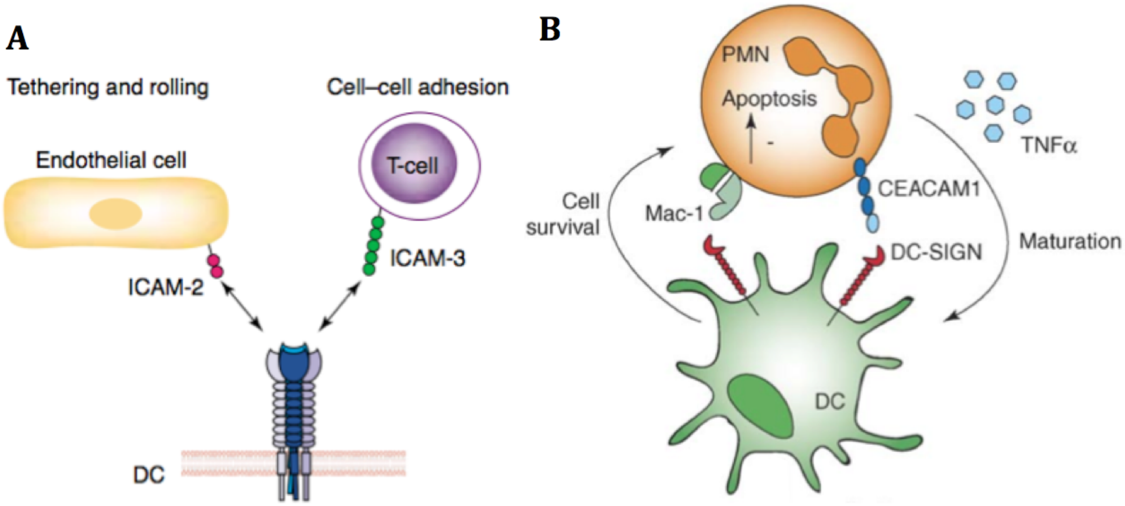


**Figure 2.14.:** The physiological role of DC-SIGN and endogeneous ligands.
**A**, DC-SIGN promotes transmigration of DCs from the blood vessels into lymphoid tissues by interacting with ICAM-2 expressed on endothelial, and mediates immunological synapse initiation by binding to ICAM-3 on T cells (from [98]). **B**, DC-SIGN binds Mac-1 and CEACAM on PMNs and induces PMN survival, while PMNs can induce DC maturation (from [99]).

DC-SIGN also functions as a PRR since it is capable to recognize the glycans expressed on various pathogens. It has been shown that DC-SIGN is implicated in antigen processing and presentation to T cells [100]. Initially, DC-SIGN was found to bind HIV-1 through interaction with envelope glycoprotein gp120, but soon it appeared that this lectin also captures many other dangerous pathogens (table 1.2). Unfortunately, this recognition does not lead to infection clearance, but rather the pathogens exploit DC-SIGN to escape immunity and facilitate their invasion and infective processes.

#### 2.3.4.2. *Ligand-induced DC-SIGN signaling*

The ligand binding to DC-SIGN triggers endocytosis and cellular signaling, which requires simultaneous TLR signaling and is dependent on the nature of the ligand. The cytoplasmic domain of DC-SIGN contains di-leucine (LL) and triacidic cluster (EEE) and other internalization motifs, which direct the DC-SIGN-bound ligands into late lysosomes or MHC class II positive endosomes [100, 101]. However, this process is ligand-dependent since HIV-1 virions bound to DC-SIGN are directed to early endosomes [46]. DC-SIGN-mediated internalization has been shown to occur via clathrin-coated pits and requires membrane cholesterol and dynamin [102], but it is worth to note that ligands, as antibodies, binding to DC-SIGN neck region induce clathrin-independent internalization that directs ligands to early endosomes [103].

The cytoplasmic tail of DC-SIGN also contains an incomplete ITAM indicating the lectin's signaling capability [104]. Indeed, the final outcome of DC-SIGN triggering is the gene transcription by NF-ϰB (reviewed in [19, 29]), despite that DC-SIGN is considered ITAM/ITIM independent CLR. NF-ϰB is an important transcription factor composed of several subunits and activated by several classes of PRRs following pathogen recognition. It regulates the expression of genes encoding costimulatory molecules, cytokines and chemokines. However, the proper induction of gene transcription requires NF-ϰB translocation to nucleus, and this cannot be accomplished by triggering DC-SIGN alone – the prior activation of NF-ϰB by TLR signaling is necessary. Thus, DC-SIGN modulates the NF-ϰB activity and the cross-talk between DC-SIGN and TLR signaling gives the final immune response.

Mannose-containing ligand binding to DC-SIGN induces the downstream effector kinases, which further leads to Ser/Thr protein kinase RAF1 activation (fig 2.15). The activated RAF1 then phosphorylates p65 subunit of NF-ϰB, which makes it a target for subsequent acetylation and finally increases its activity (prolongs its nuclear activity, enhances transcriptional rate and increases binding affinity to DNA) on *Il6*, *Il8*, *Il12a* and *Il12b* promotors, and particularly on *Il10* promotor, which strongly increases anti-inflammatory IL-10 production. Although RAF1 activation is followed by DC-SIGN triggering and does not require TLR signaling, p65 cannot be activated by DC-SIGN alone – DC-SIGN signaling only modulates p65 activation when it has been already translocated to nucleus (reviewed in [29, 18]).

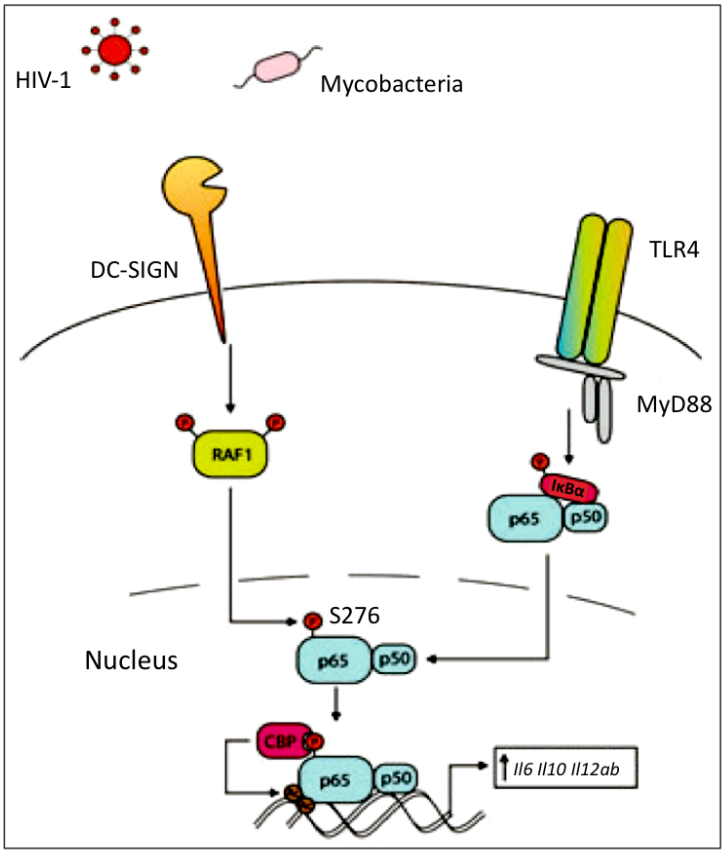


**Figure 2.15.:** TLR signaling modulation by DC-SIGN triggering induced by mannosylated ligands bearing pathogens such as HIV-1 or *Mycobacteria*.
(From [18])

There is very limited knowledge about DC-SIGN signaling in response to fucose-containing pathogen binding, and it comes only from Lewis$^{Y}$ positive and negative *H. pylori* infection studies and examinations of immune responses to soluble egg antigens of the parasite Schistosoma mansoni (reviewed in [29]). Only Lewis$^{Y+}$ *H. pylori* interacts with DC-SIGN, which induces high up-regulation of only anti-inflammatory IL-10 production. Such a response is significantly different from responses to Lewis$^{Y-}$ *H. pylori* strain as well as to mannose-containing pathogens that induce up-regulation of both pro-inflammatory cytokines such as IL-6 and IL-12, and anti-inflammatory IL-10.

Thus the pathogens evolved to exploit DC-SIGN signaling for their infection promotion. The cross-talk between DC-SIGN and TLR signaling also may explain why DC-SIGN recognition of endogenous ligands does not lead to DC maturation and cytokine production: since there is no simultaneous activation of PRRs by pathogens to induce NF-$\varkappa$B activation, DC-SIGN triggering alone does not induce immune responses [18].

## 2.4. Langerin: a specific CLR of Langerhans cells

As DC-SIGN, langerin is a type II C-type lectin, almost exclusively expressed on epidermal LCs, but also present on dermal CD103$^+$ DCs and lymph node resident CD8$^+$ DCs [105, 106, 107]. LCs is a subset of DCs, whose role is not completely clear. Although initially they were assumed to function as APCs, like dermal or stromal DCs [108], the accumulating evidence of their fail to present antigens from various viruses and parasites to T cells activation and the observation that LCs induce T regulatory cells supports the presumption that these cells may have an immunosuppressive, tolerogenic role [22].

A characteristic hallmark of LCs is the Birbeck granules (fig. 2.16), and langerin has been shown to be the main component and responsible for the formation of these tennis racquet or rod-shaped membranous structures [105].

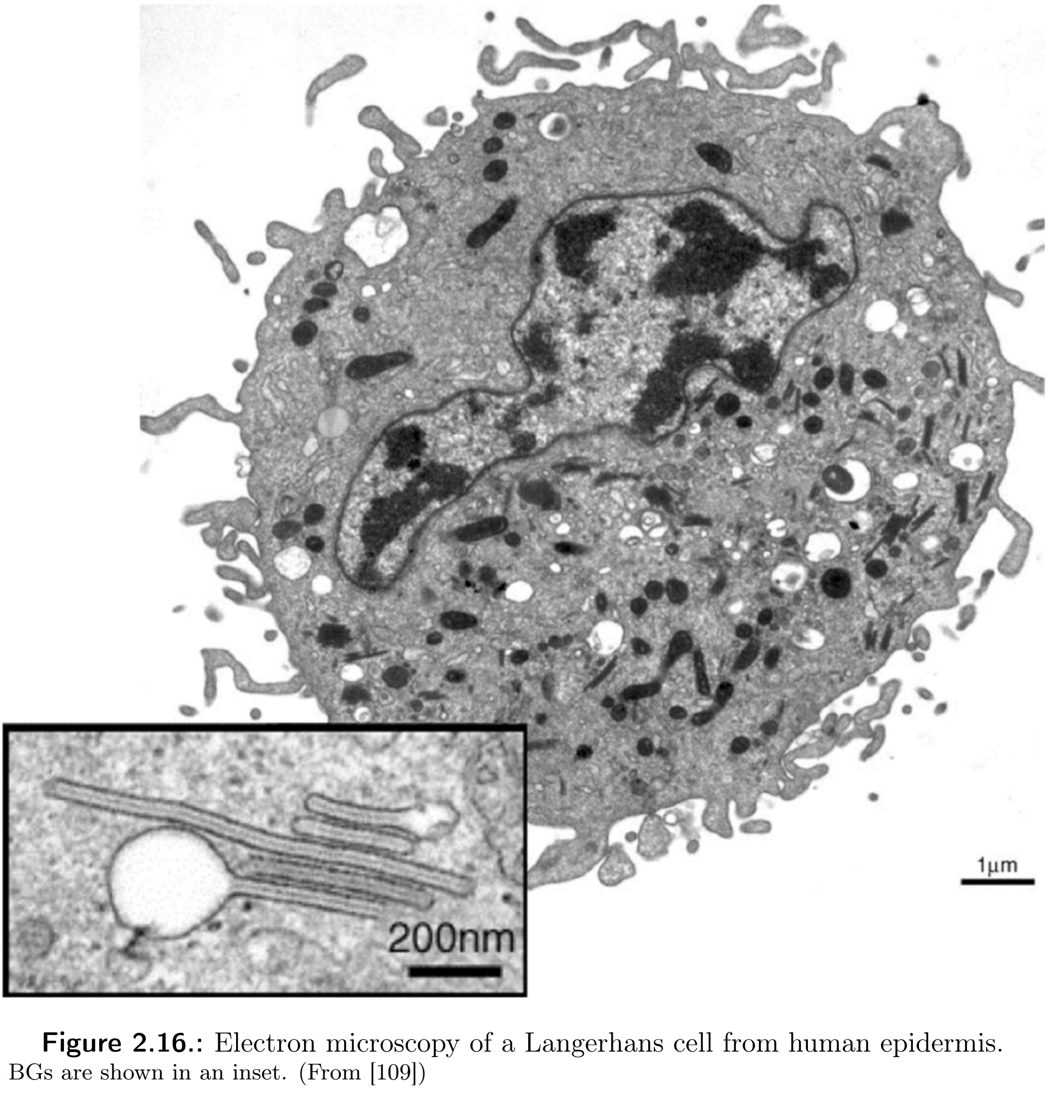


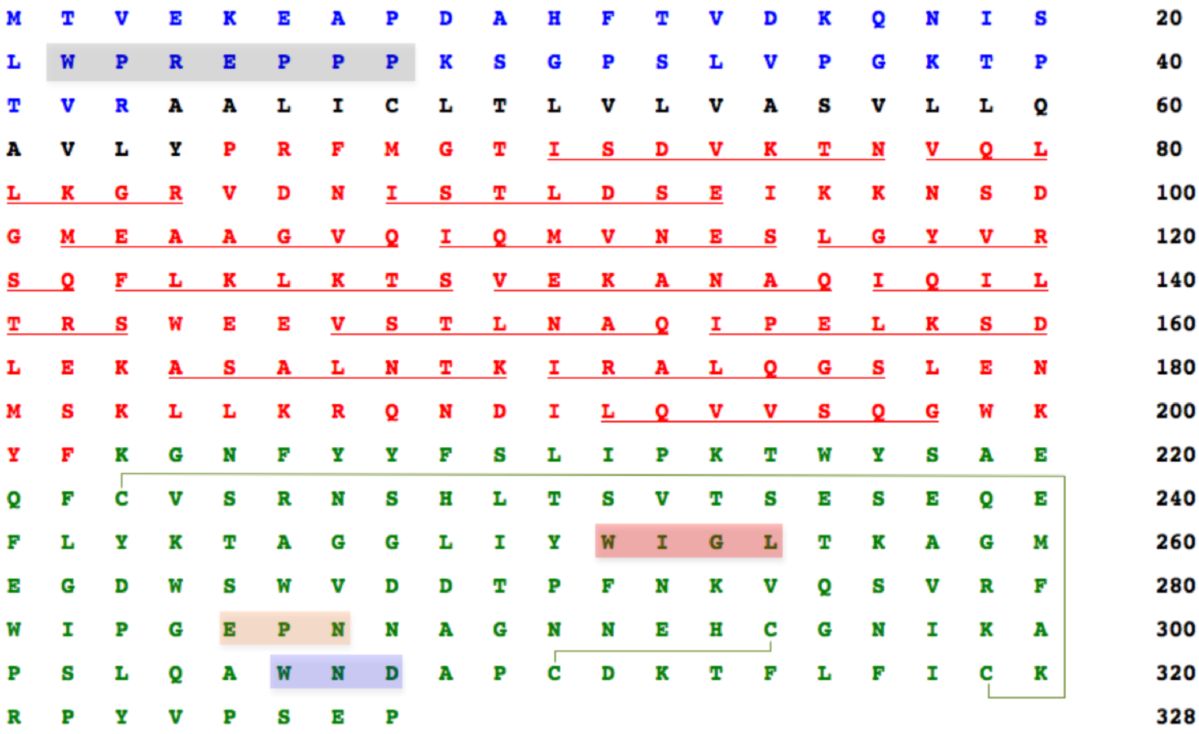
**Figure 2.16.:** Electron microscopy of a Langerhans cell from human epidermis. Typical BGs are shown in an inset. (From [109])

Langerin has been identified as a major pathogen-capturing endocytic receptor of LCs (table 1.2) [110, 21]. It binds pathogens through the recognition of high-mannose structures [111] present on viral envelopes, mannan and β-glucan structures on fungi [21].

As already described in sub-subsection 2.2.1.2, langerin is particularly important in protection against HIV infection since it captures HIV-1 virions by binding to envelope glycoprotein gp120, and this leads to their degradation in BGs [57]. The anti-fungal function of LCs is suggested as well, with a central protective role exerted also by langerin [21].

### 2.4.1. *The structure and carbohydrate recognition by langerin*

Langerin is a type II transmembrane receptor (328 amino acids, 37.5 kDa) with an overall molecular organization similar to DC-SIGN and other CLRs: it consists of N-terminal cytoplasmic domain, transmembrane domain, neck region, and a C-terminal C-type CRD (fig. 2.17) [105]. The cytoplasmic domain of langerin contains a proline-rich signaling motif (WPREPPP), which could function as a docking site for signal transduction proteins, and indeed, it was demonstrated to be important for langerin intracellular targeting [112]. Nonetheless, the role of langerin signaling is still obscure [18].



| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M | T | V | E | K | E | A | P | D | A | H | F | T | V | D | K | Q | N | I | S | 20 |
| L | W | P | R | E | P | P | P | K | S | G | P | S | L | V | P | G | K | T | P | 40 |
| T | V | R | A | A | L | I | C | L | T | L | V | L | V | A | S | V | L | L | Q | 60 |
| A | V | L | Y | P | R | F | M | G | T | I | S | D | V | K | T | N | V | Q | L | 80 |
| L | K | G | R | V | D | N | I | S | T | L | D | S | E | I | K | K | N | S | D | 100 |
| G | M | E | A | A | G | V | Q | I | Q | M | V | N | E | S | L | G | Y | V | R | 120 |
| S | Q | F | L | K | L | K | T | S | V | E | K | A | N | A | Q | I | Q | I | L | 140 |
| T | R | S | W | E | E | V | S | T | L | N | A | Q | I | P | E | L | K | S | D | 160 |
| L | E | K | A | S | A | L | N | T | K | I | R | A | L | Q | G | S | L | E | N | 180 |
| M | S | K | L | L | K | R | Q | N | D | I | L | Q | V | V | S | Q | G | W | K | 200 |
| Y | F | K | G | N | F | Y | Y | F | S | L | I | P | K | T | W | Y | S | A | E | 220 |
| Q | F | C | V | S | R | N | S | H | L | T | S | V | T | S | E | S | E | Q | E | 240 |
| F | L | Y | K | T | A | G | G | L | I | Y | W | I | G | L | T | K | A | G | M | 260 |
| E | G | D | W | S | W | V | D | D | T | P | F | N | K | V | Q | S | V | R | F | 280 |
| W | I | P | G | E | P | N | N | A | G | N | N | E | H | C | G | N | I | K | A | 300 |
| P | S | L | Q | A | W | N | D | A | P | C | D | K | T | F | L | F | I | C | K | 320 |
| R | P | Y | V | P | S | E | P | | | | | | | | | | | | | 328 |

**Figure 2.17.:** Primary structure of human langerin.
The sequences of cytosolic, transmembrane, neck region, and CRD domains are in blue, black, red and green letters.The potential proline-rich signaling motif is highlighted in grey, and the conserved motifs defining sugar specificity and Ca$^{2+}$-coordination are in orange and blue. The conserved WIGL motif is in red. The putative heptads contributing to coiled-coil formation are individually underlined, and the disulphide bridges are depicted by light green lines.

The CRD of langerin (fig. 2.18) contains EPN motif indicating its specificity for mannose type sugars. A glycan array analysis using truncated trimeric langerin, has revealed several types of carbohydrates as ligands of langerin [113]. As expected, langerin bound high-mannose N-linked oligosaccharides, however, the only fucose-based ligand of langerin was blood group B antigen (Galα1–3(Fucα1–2)Gal). Moreover, a peculiar property of langerin to bind glycans terminating in 6-sulfated galactose was also found [113] that explained the ability of langerin to recognize keratan sulphate [114].

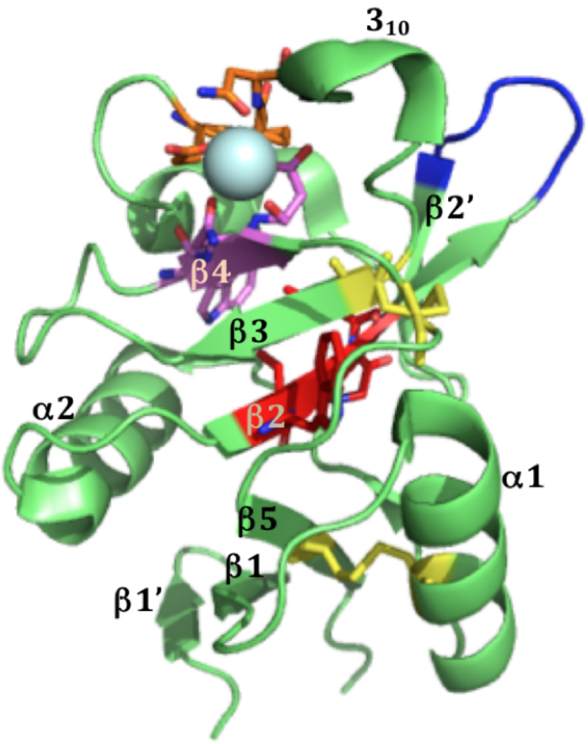


**Figure 2.18.:** The structure of langerin CRD.
EPN, WND and WIGL motifs are shown in orange, pink and red stick representation. Disulphide bridges depicted in yellow sticks and a cyan sphere is the $Ca^{2+}$. The flexible loop comprising residues 258-262 is shown in blue.

Structural studies of ligand recognition by langerin revealed that it binds high-mannose structures through non-reducing terminal mannoses and internal Manα1-2Man (fig. 2.19A), which has multiple binding modes in the conventional $Ca^{2+}$-dependent sugar-binding site. The binding mode of this non-reducing mannose residue is virtually the same as mannose monosaccharide (pdb:3P7G), except that it is 180° flipped by the axis perpendicular to C3-C4 bond. In contrast to DC-SIGN, langerin has only a small extended binding site that contacts other sugar residues in high-mannose oligosaccharides, and better binding affinity to these oligosaccharides compared to mannose may be a statistical effect of having multiple available mannose residues, as is in case of DC-SIGN [115].

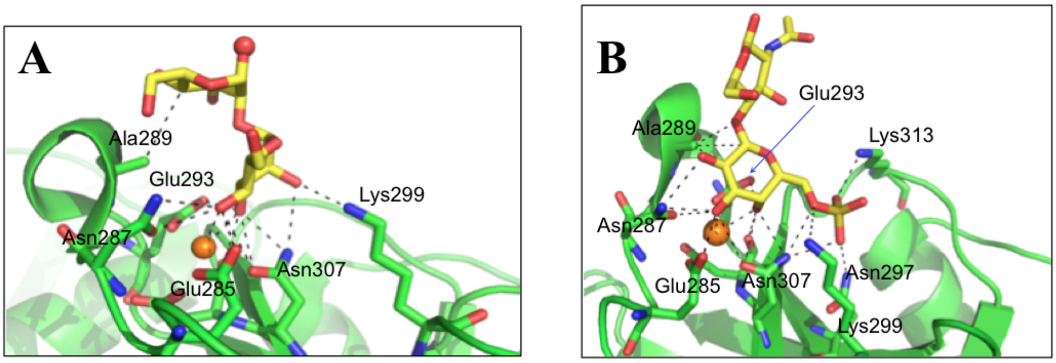


**Figure 2.19.:** The CRD of langerin in complex with carbohydrate ligands.
**A**, Manα1-2Man in a likely preferred and favored binding mode in the context of longer oligosaccharide with non-reducing mannose coordinating to $Ca^{2+}$ and oxygen atom that would be linked to the oligosaccharide shown in red sphere (pdb: 3P5F). **B**, $6SO_4$–Galβ1–4GlcNAc (pdb:3P5I). The $Ca^{2+}$ ions are shown as orange spheres.

It was also demonstrated that langerin is unique among C-type CRDs since apart of binding mannose/fucose type sugars, it is also able to interact with galactose-based saccharides: it recognizes 6SO4–Gal residues that are present in keratan sulfate as $6SO_4$–Galβ1–4GlcNAc repeating unit[114]. In crystal structure of langerin CRD in complex with $6SO_4$–Galβ1–4GlcNAc (fig. 2.19B) the galactose residue was found to make coordination to $Ca^{2+}$ by axial 4-OH and equatorial 3-OH groups, and the galactose residue packs against Ala289. Finally, the $SO_4$ group makes salt bridges with Lys299 and Lys313 [115].

### 2.4.2. *Trimeric structure of langerin and its role in BG formation*

The neck region of langerin is composed of the series of heptads (fig. 2.17), although these repeats are not as obvious as the tandem 23 amino acid repeats in the neck of DC-SIGN. These heptads are responsible for the oligomerization of langerin into trimers (fig. 2.20A) [111, 113].

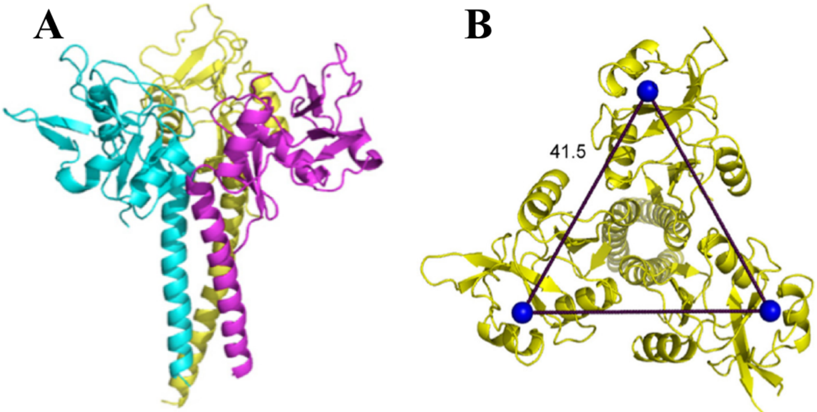


**Figure 2.20.:** The trimeric organization of truncated langerin.
**A**, The side view of the trimer. **B**, The top view of the trimer; the spacing shown between CRDs is the average of the distances in Å. (From [113])

The trimeric langerin is rather a rigid unit with the CRDs in fixed positions as can be seen in the crystal structure of the truncated trimeric langerin: the CRDs make multiple contacts with the neck region and this leads to the arrangement of primary sugar-binding sites at fixed positions separated by a distance of 42 Å (fig. 2.20B). The only significant flexibility in the CRD was observed at the loop region comprising residues 258-262 (fig. 2.18), and results most likely due to the absence of auxiliary $Ca^{2+}$ sites that are present in many other CRDs. This kind of fixed organization of sugar-binding sites may restrict ligand recognition by langerin [113].

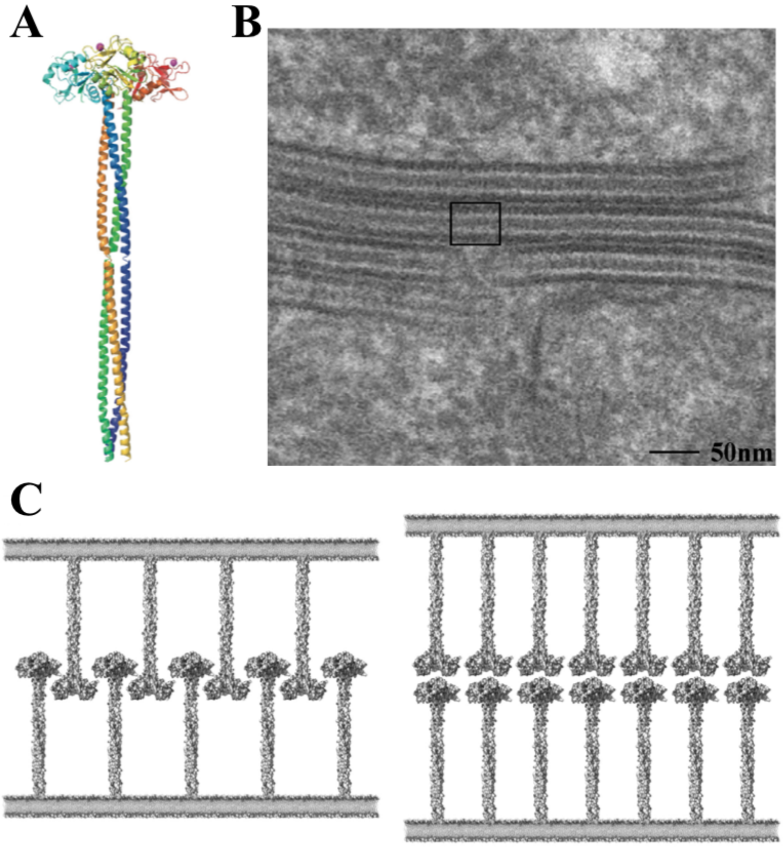
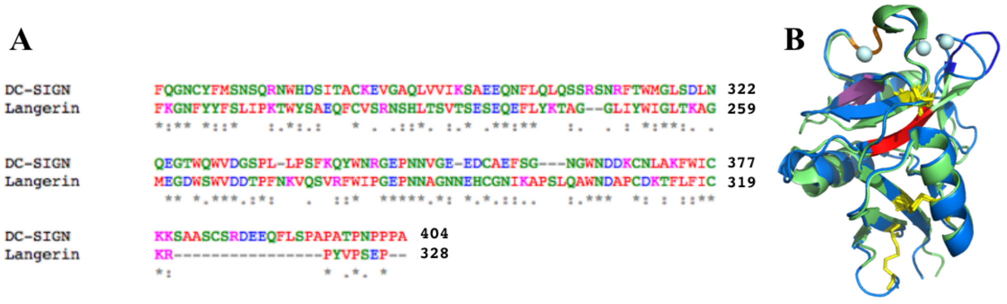


**Figure 2.21.:** Contribution of langerin to Birbeck granule formation.
**A**, The model of langerin extracellular domain. **B**, Stacked Birbeck granules of LC. **C**, two suggested macromolecular organization models of BG stacking with langerin as a major component and the driving force. (From [112])

It has been demonstrated that the CRD of langerin is essential for BG formation [116, 112]. Based on hydrodynamic properties of extracellular domain (ECD) of langerin as a fairly elongated protein, and sequence analysis, the group, which I later joined, has built a model of the whole ECD (fig. 2.21A). This model, together with electron microscopy data on BGs, was used to construct a hypothesis of a possible molecular mechanism of BG formation driven by langerin (fig. 2.21B) [112]. However, the plausibility of such membrane zipping mechanisms and molecular details of the process still need to be confirmed.

## 2.5. The comparison of the CRDs of DC-SIGN and langerin

Both DC-SIGN and langerin are lectins of type II with C-type CRDs. The sequences and the overall structures of their CRDs are close (fig. 2.22) and correspond to the classical CRD fold (fig. 1.6). However, several important differences can be tracked in these two CRDs.



```
DC-SIGN         FQGNCYFMSNSQRNWHDSITACKEVGAQLVVIKSAEEQNFLQLQSSRSNRFTWMGLSDLN 322
Langerin        FKGNFYYFSLIPKTWYSAEQFCVSRNSHLTSVTSESEQEFLYKTAG--GLIYWIGLTKAG 259
                *:** *::*   :.*:.:   * . .::*. :.* .**:**   :.  . : *:**:. .

DC-SIGN         QEGTWQWVDGSPL-LPSFKQYWNRGEPNNVGE-EDCAEFSG---NGWNDDKCNLAKFWIC 377
Langerin        MEGDWSWVDDTPFNKVQSVRFWIPGEPNNAGNNEHCGNIKAPSLQAWNDAPCDKTFLFIC 319
                 ** *.***.:*:   .  ::*  *****.*: *.*.::..   :.***  *: : ::**

DC-SIGN         KKSAASCSRDEEQFLSPAPATPNPPPA 404
Langerin        KR----------------PYVPSEP-- 328
                *:                * .*. *
```

**Figure 2.22.:** The comparison of DC-SIGN and langerin CRDs.
**A**, An alignment of the sequences of DC-SIGN and langerin CRDs by ClustalW (amino acid color codes: red – hydrophobic, green – polar, pink – positively charged, blue – negatively charged). **B**, structural alignment of DC-SIGN and langerin CRDs (the color coding is the same as in fig. 2.9 and 2.18).

Langerin has several structural elements that are not present in DC-SIGN CRD, including a $3_{10}$ helix close to sugar-binding site, and an additional β2´ strand (fig. 2.18). The whole CRD structure of langerin is supported by two conserved disulphide bridges, in contrast to four S-S bonds in DC-SIGN. Unlike DC-SIGN, langerin CRD has only one $Ca^{2+}$ ion at Ca site 2 (conventional sugar-binding site), and the lack of other $Ca^{2+}$ ions in sites 1 and 3 might be the reason for the high flexibility of the β2-β2´ loop comprising residues 258-262, which also leads to the formation of a large groove specific to the langerin structure.

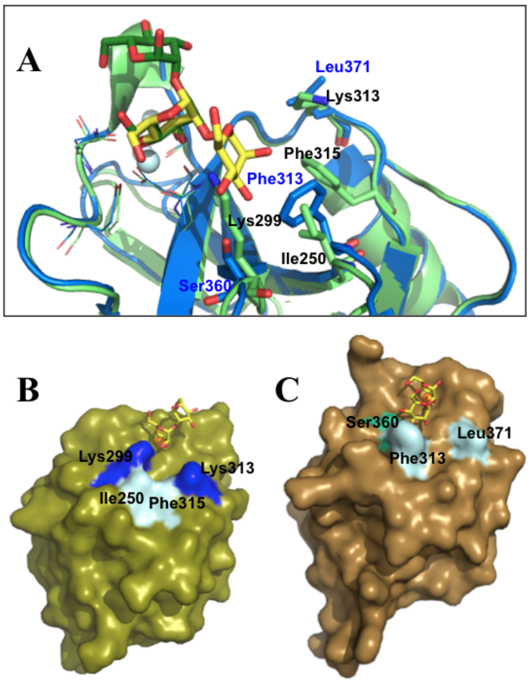


**Figure 2.23.:** The differences of sugar-binding sites of DC-SIGN and langerin.
**A**, The alignment of DC-SIGN (blue) (pdb:2IT5) and langerin (light green) (pdb:3P5F) CRDs bound to Manα1-2Man, shown in yellow (for DC-SIGN) and green (for langerin) sticks; the side chains of EPN and WND motifs are shown in lines, and the most important different amino acid residues required for sugar binding are highlighted by stick representation; the blue and black labels correspond to DC-SIGN and langerin side chains, respectively; light cyan spheres are $Ca^{2+}$ ions. **B** and **C** are surface representations of langerin and DC-SIGN CRDs, respectively, complexed with Manα1-2Man; the important side chains are highlighted; green sphere is $Ca^{2+}$.

However, the most relevant difference of these two lectins is in their sugar-binding site topologies. A unique feature of langerin CRD is the presence of two lysine residues (Lys299 and Lys313) in sugar binding-site, which provide the binding site the basic character that allows langerin to accommodate sulfated sugars in its binding site. On the other hand, while Phe313 is an important side chain in DC-SIGN sugar-binding site as it forms stacking interactions with the sugar ring (fig. 2.23A) [117, 73], in langerin Phe315 positioning is not suitable to participate in such sugar binding (fig. 2.23A). Overall topology of langerin CRD composes only a small binding site strongly constrained by Lys299 (fig. 2.23B), while DC-SIGN has a potential to adapt more extended oligosaccharides with a widely open binding site (fig. 2.23C and 2.9C).

The CRD-neck junctions of DC-SIGN and langerin are also different: while the flexibility in DC-SIGN is retained, langerin trimers present CRDs in a rather rigid manner [113]. This likely leads to different way of ligand recognition of the two lectins.

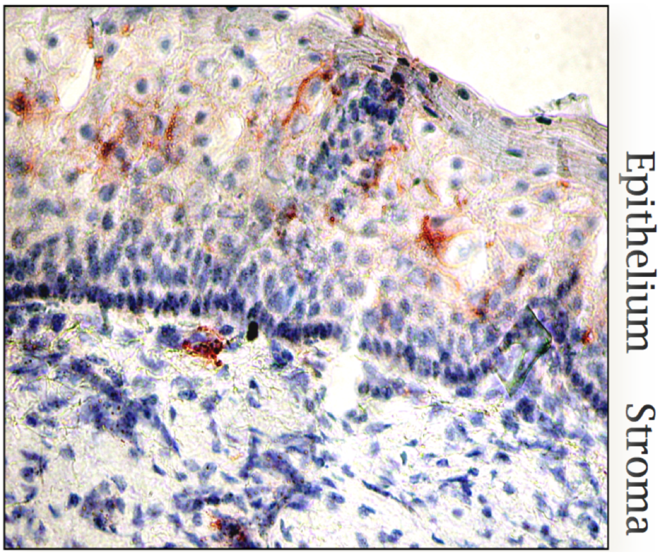


**Figure 2.24.:** The micrograph of mucosal tissue.
Langerin is visualized in red, cell nuclei are in blue. (Personal communication from Dr. Jenny Valladeau-Guilemond and Dr. Colette Dezutter-Dambuyant, Centre Léon Bérard – UMR INSERM 1052-CNRS 5286, Lyon, France)

Finally, the localization of these two lectins in mucosa is different. Langerin is expressed on LCs that reside in epithelia (fig. 2.24), while DC-SIGN expressing stromal DCs survey the underlying stroma. Thus LCs with langerin are the very first barrier that pathogens encounter invading the mucosal tissues.

# 3. HIV sexual transmission prevention strategy

## 3.1. Microbicides as the means of HIV prevention

More than 40 million people around the world are infected with HIV, with an estimate of 15,000-20,000 new infections each day [118]. However, there is no cure for HIV infection so far and the existing life-long treatment, HAART (highly active antiretroviral treatment, a combination of several substances that target to interfere with various stages in HIV life cycle), is rather expensive and affordable to only a limited population of those infected. Moreover, HAART has numerous adverse effects, some of which are even life threatening (reviewed in [119, 120]). This might impose high physical and psychological burden, and part of the patients were discontinuing the therapy within the first year [121]. Also HAART may lead to drug-induced virus resistance. Thus, the development of the means of infection prevention is of the utmost importance as this would be the most effective way to halt the epidemic.

Unfortunately, up to date there is no effective anti-HIV vaccine developed, despite extensive research efforts for more than 15 years. The only exception was RV144 trial, also called ‘Thai trial’, which gave only modest results of 31.2% vaccine efficacy [122].

The sexual transmission of HIV is a dominating route of infection (table 2.1). Women globally comprise over 50% of all people living with HIV and they appear to be at higher risk for infection transmission through cervicovaginal tissues (table 2.1) [44]. Moreover, in sub-Saharan Africa and the Caribbean women make up 59% and 53% of the HIV infected population, respectively [118]. Although the simplest and indeed effective means of prevention is a proper and consistent use of condoms, the possible social, cultural, ethical and economical factors may limit their use. Therefore, alternative means should be considered.

Vaginal HIV microbicides (VHMs) are a promising alternative for safe and effective prevention methods from sexually transmitted HIV. Furthermore, they are particularly important for women since such microbicides would provide the way of independent self protection, as women are not always able to lower their risks because of social, cultural and economical reasons in the high-risk regions of the developing countries [123].

The concept of HIV prevention that would be controlled by women themselves by using a topical vaginally applied “virucide” was introduced in 1990 by Zena A. Stein [124] and gave the starting point for the development of VHMs.

VHMs are substances designed to inhibit or block early events in HIV infection. According to their mode of action, the VHMs can be broadly classified to five groups (table 3.1, information collected from [125, 126, 127]): surfactants, vaginal milieu protectors, inhibitors of entry, fusion and reverse transcriptase [125].

Surfactants and vaginal milieu protectors are rather non-specific protection means, while HIV entry together with HIV reverse transcriptase inhibitors must be designed as specifically interacting substances. Furthermore, HIV entry inhibitors can be designed to target either viral envelope or host surface molecules implicated in the virus entry process.

To be suitable for the use as intravaginal topical microbicides, the active substances and their formulation, apart from their efficacy, must meet the following criteria (reviewed in [128]): they must be safe and non-toxic with no inflammatory responses, resist to seminal fluid, vaginal pH and body temperature, generally acceptable by women (i.e. easy to use, comfortable, not expensive), and available to produce at large-scale and affordable cost. VHMs can be formulated as gels/creams, rings, foams, and films, and the mode of formulation is one of the major factors for the adherence to a potential VHM.

**Table 3.1.:** The examples of different type vaginal microbicides.

| Microbicide type | Mode of action | Examples |
|---|---|---|
| Surfactants/membrane disruptors | Disrupt membranes non-specifically and thus provide contraceptive activity against a wide range of potential sexually transmitted infections (STI). | • Nonoxynol-9 (N-9);<br>• C31G (Savvy);<br>• Sodium lauryl sulfate (Invisible Condom). |
| Vaginal milieu protectors | Maintain, restore, or enhance the natural protective mechanisms, i.e. the acidic pH naturally maintained by *lactobacilli* (pH between 4.0 and 5.8 inactivates HIV), within vaginal canal. | • Carbopol 974P (BufferGel);<br>• Acidform (Amphora);<br>• "Probiotics": natural or bioengineered *lactobacilli*. |
| Entry inhibitors | Anionic polymers:<br>Negative charge causes interaction with HIV's viral envelope proteins (gp120 of R4 strain in particular) and interferes with attachment to $CD4^+$ T cells. | • Naphthalene sulfonate (PRO2000);<br>• Carrageenan (Carraguard/ PC-515);<br>• Cellulose sulfate (Ushercell);<br>• Cellulose acetate phthalate (CAP);<br>• Dendrimers: SPL7013 (Vivagel). |
| | CCR5 blockers<br>Block CCR5 coreceptor and interfere with attachment of HIV to host cells | • PSC-RANTES;<br>• CMPD167. |
| Fusion inhibitors | Prevention/inhibition of viral-host cell membrane fusion | • C-52L (inhibits gp41-mediated viral—cell fusion);<br>• Algal lectins (prevents viral-host cell fusion by binding high mannose residues in the HIV envelope). |
| Reverse transcriptase inhibitors | Interfere with HIV reverse transcriptase | • Tenofovir (PMPA, nucleotide analogue);<br>• TMC120 (non-nucleoside reverse transcriptase inhibitor (NNRTI));<br>• UC781 (NNRTI). |

Disappointingly, for many of the clinically tested VHMs (cellulose sulphate, SAVVY, Carraguard, PRO 2000 and BufferGel) there was no evidence found on reduction of HIV acquisition (reviewed in [129]), some even showed toxicity (N-9, CAP) [125]. The most successful VHM candidate, the antiretroviral (ARV) drug tenofovir based gel, has provided clinical data to support its efficacy in CAPRISA 004 trial [130]; nonetheless, the second trial that enrolled 5000 women was stopped early due to low likelihood of showing a protective effect (reviewed in [129]). There are other compounds that are undergoing Phase II or III clinical trials [125], but in conclusion, up to date there is no VHM substance to have successfully passed the Phase III.

The ARVs used for HIV therapy are increasingly investigated for the use as microbicides [131]. Moreover, in the search of the new ways of HIV-1 prevention, the human C-type lectin DC-SIGN was identified as a potential therapeutic target because of its important role in HIV infection enhancement and dissemination in the mucosal site (see section 2.2). There are several groups working on the design of the compounds, mostly glycomimetic and some non-carbohydrate, that have the ability to bind to DC-SIGN blocking its interaction with HIV (reviewed in [132]).

It is also important to notice that the discovery of efficient and safe DC-SIGN antagonists would be beneficial not only for protection from HIV infection. This lectin is usurped by numerous pathogens, including Ebola, Dengue viruses, *Mycobacteria tuberculosis* and others (table 1.2), to enhance their infective processes and survive host immunity, causing chronic diseases. This makes it an interesting target for therapeutical intervention, and DC-SIGN antagonists could be applicable as anti-infective agents against multiple pathogens [132].

## 3.2. The concept and design of glycomimetic drugs

Glycomimetics can be defined as chemical entities that mimic the bioactive carbohydrates and bind to target proteins, mostly lectins or sugar-modifying enzymes, thus interfering various pathological processes. The glycomimetic drug substances present several advantages over the natural carbohydrate ligands: they can be designed to have higher affinity and specificity for the target protein as well as longer plasma half-life (i.e. resistance to glycosidases). Moreover, these substances may be designed to have a lower polarity compared to carbohydrates, which would improve their pharmacokinetic properties [133, 132].

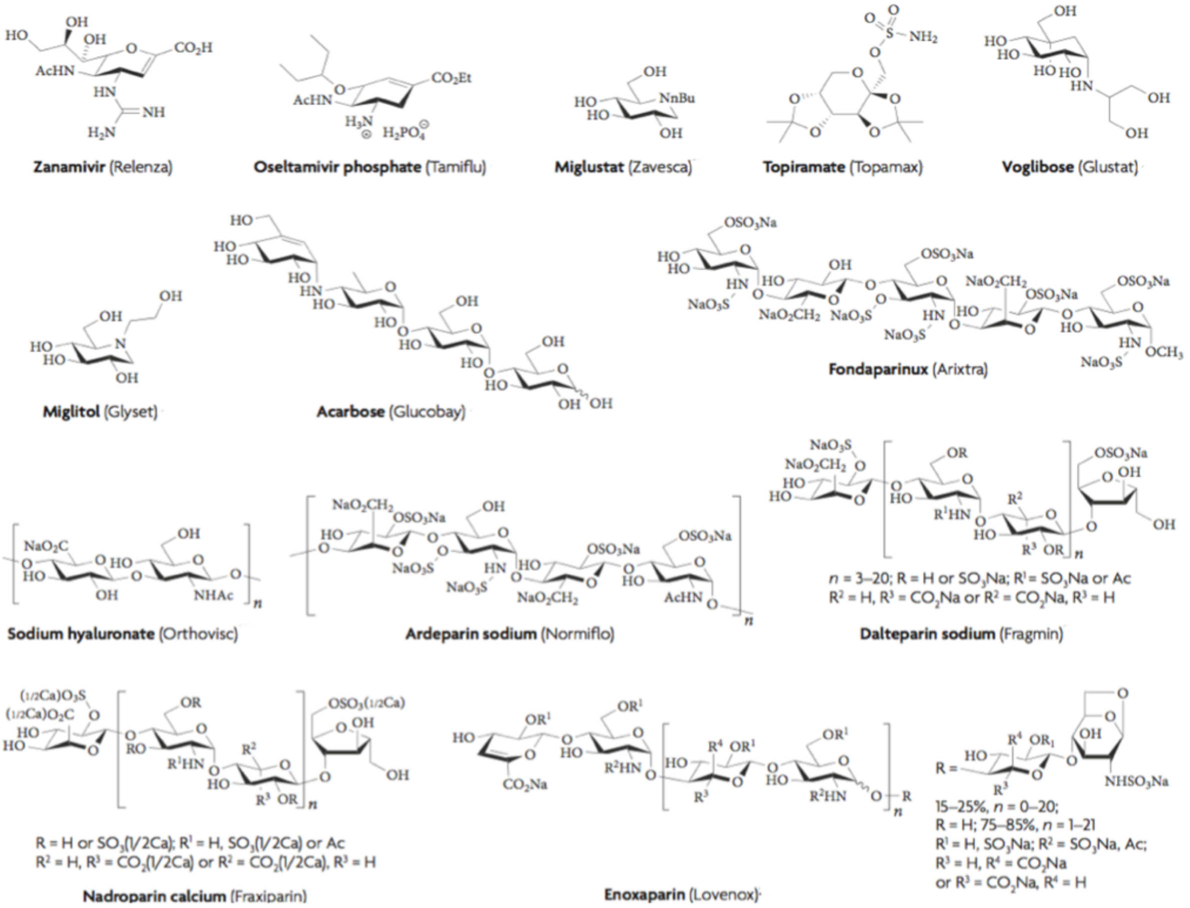

**Zanamivir** (Relenza)
**Oseltamivir phosphate** (Tamiflu)
**Miglustat** (Zavesca)
**Topiramate** (Topamax)
**Voglibose** (Glustat)
**Miglitol** (Glyset)
**Acarbose** (Glucobay)
**Fondaparinux** (Arixtra)
**Sodium hyaluronate** (Orthovisc)
**Ardeparin sodium** (Normiflo)
n = 3–20; R = H or SO3Na; R¹ = SO3Na or Ac
R² = H, R³ = CO2Na or R² = CO2Na, R³ = H
**Dalteparin sodium** (Fragmin)
R = H or SO3(1/2Ca); R¹ = H, SO3(1/2Ca) or Ac
R² = H, R³ = CO2(1/2Ca) or R² = CO2(1/2Ca), R³ = H
**Nadroparin calcium** (Fraxiparin)
**Enoxaparin** (Lovenox)
15–25%, n = 0–20;
R = H; 75–85%, n = 1–21
R¹ = H, SO3Na; R² = SO3Na, Ac;
R³ = H, R⁴ = CO2Na
or R³ = CO2Na, R⁴ = H

**Figure 3.1.:** The examples of the approved carbohydrate-based and glycomimetic drugs. Voglibose, miglitol and acarbose are the drugs for diabetes treatment (glycosidase inhibitors that prevent the digestion of carbohydrates). Zanamivir and oseltamivir are active substances for the prevention and treatment of influenza virus infections. Fondaparinux, dalteparin, ardeparin, nardoparin and enoxaparin, which are sulphated glycosaminoglycans, function as anticoagulants for the treatment of thrombosis. Miglustat, topiramate and sodium hyaluronate are used to treat Gaucher's disease, epilepsy and osteoarthritis, respectively. Trade names are in the brackets (from [133])

The glycomimetics form a new class of therapeutics and currently they cover only a small area of the multitude of available drugs. There are several examples of already developed and commercially available carbohydrate-derived drugs (fig. 3.1). Among them, a successful anti-influenza prodrug oseltamivir (Tamiflu), which is metabolically converted to a viral neuraminidase inhibitor, is a paradigm of a glycomimetic drug: it was developed from a carbohydrate lead by systematical modifications of the molecule to finally achieve a prodrug [133].

Generally, the rational design of a glycomimetic drug requires the explicit knowledge of the structure–activity relationships, i.e. the molecular basis of the interaction between the carbohydrate epitope and its target protein receptor. To achieve that, the carbohydrate in complex with its receptor can be studied by empirical techniques such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography [134].

The affinity of the designed compound can be further improved by optimizing those positions of the molecule that are crucial for the interaction. The knowledge of the bioactive conformation can be also very useful: the design of glycomimetics with a pre-organized favorable conformation in solution leads to reduced entropy costs upon binding, thus increased affinity [134].

The next step would be the optimization of the compound's valency depending on the nature of the biological interaction. In biological systems, the interactions between lectins and their carbohydrate ligands are usually oligovalent due to oligomerisation of the lectin receptor, its clustering on the cell surface or presentation of carbohydrate epitopes in a clustered manner or all of these factors working simultaneously. Hence, in order to attain an efficient compound that would be capable to compete such interactions, the strategies of multivalent presentation must be considered and optimized [133]. There are many different platforms being used for the design of glycoclusters, and the examples include glycodendrimers [135], gold glyconanoparticles [136], fullerene sugar balls [137], cyclodextrin-based [138], calyx[$n$]arene-based [139], cyclopeptide-based glycoclusters [140] and other scaffolds (reviewed in [141]).

The above described stages of the glycomimetic drug design outlines the principal workflow and objectives of the European initial training network (ITN) CARMUSYS, which aims to develop DC-SIGN antagonists with the purpose to block early infections of HIV as well as other pathogens. I joined this network to carry out the PhD studies, and thus this workflow will be reflected later throughout this thesis manuscript.

The final step in the rational design of the glycomimetic drug substance is the improvement of the pharmacokinetic/pharmacodynamic properties of the molecule as well as optimization for the chosen way of administration [134].

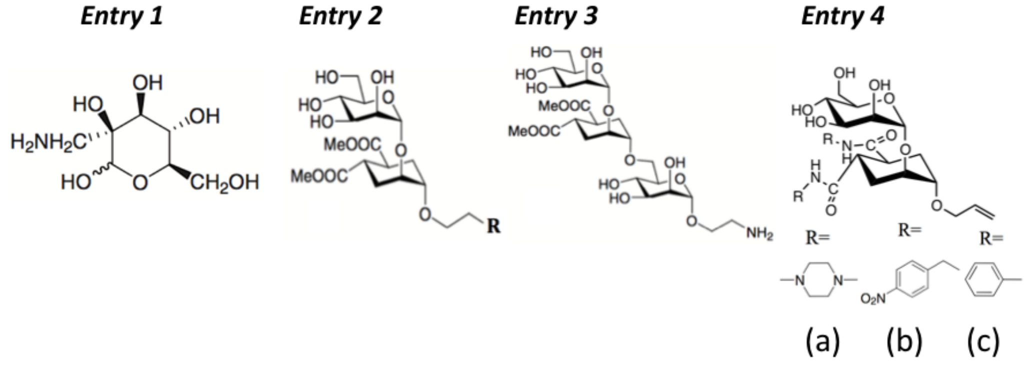
## 3.3. Glycomimetics as DC-SIGN inhibitors

The strategy of glycomimetics is a rational approach for the development of DC-SIGN antagonists, since this protein is a lectin receptor recognizing mannose or fucose-based sugars. Therefore, the main strategy of the development of glycomimetic DC-SIGN ligands is to retain mannose or fucose monosaccharide as an anchor, while other parts of the molecule can be designed and optimized to attain the better affinity, specificity and selectivity. Indeed, a handful of the glycomimetic DC-SIGN ligands (tables 3.2 and 3.3) bearing both mannose or fucose anchors has been developed by several research groups (reviewed in [128, 132]), and some of these compounds have successfully passed cell-based *in vitro* tests for their anti-viral activity.

**Table 3.2.:** Some examples of the published monovalent mannose-based DC-SIGN glycomimetics.

| Entry | Potency | Assay | Ref. |
|---|---|---|---|
| 1 | $K_\mathrm{I}$ 350 µM | SPR competition, gp120 coated surface | [142] |
| 2 | R=$NH_2$: $IC_{50}$ 620 µM | Inhibition of Ebola envelope-pseudotyped virus infection | [143] |
|  | R=$N_3$: $IC_{50}$ 1005 µM | SPR competition, mannosylated bovine serum albumin (ManBSA) coated surface | [49] |
| 3 | $IC_{50}$ 125 µM | SPR competition, ManBSA coated surface | [49] |
|  | $IC_{50}$ 80 µM | Inhibition of HIV *trans* infection of $CD4^+$ T cells |  |
| 4 | (a) $IC_{50}$ 6.9 µM | Inhibition of DC-SIGN-mediated adhesion to mannan coated plates | [144] |
|  | (b) $IC_{50}$ 12.8 µM |  |  |
|  | (c) $IC_{50}$ 12.5 µM |  |  |

***Structures:***



The strategy to design fucose-based mimics (table 3.3) is generally based on Lewis$^\mathrm{X}$ trisaccharide ($IC_{50}$ 0.8 mM in SPR competition assay with ManBSA surface [145]) (fig. 2.9D): the fucose anchor is retained while changing the nature of the linker (GlcNAc in Lewis$^\mathrm{X}$) and the second moiety (galactose in Lewis$^\mathrm{X}$) that is interacting in DC-SIGN CRD [145].

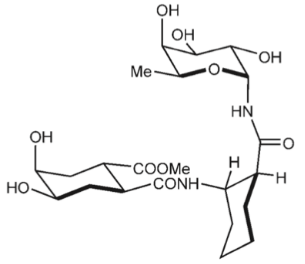
The majority of these mannose and fucose-based glycomimics (entries 2-4 in table 3.2, and both compounds in table 3.3) were designed by Pr. Anna Bernardi's group, who is currently our collaborator in the ITN CARMUSYS, and they were functionally characterized (except compound 4, table 3.2) by the group of Pr. Franck Fieschi, which I later joined.

**Table 3.3.:** Some examples of the published glycomimetic fucose-based DC-SIGN ligands.

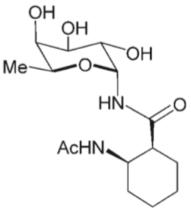
| Entry | Potency | Assay | Ref. |
|---|---|---|---|
| **1** | $IC_{50}$ 350 µM | SPR competition, ManBSA coated surface | [145] |
| **2** | $IC_{50}$ 500 µM | SPR competition, ManBSA coated surface | [145] |

***Structures:***

**Entry 1**



**Entry 2**



Another already applied approach for DC-SIGN ligand discovery is the screening of the non-carbohydrate molecules that would have necessary features of the carbohydrates that bind DC-SIGN. In such a pursuit, shikimic acid was used as a scaffold, and out of 192 synthesized its derivatives a reasonable DC-SIGN binder was identified (fig. 3.2A), which was suggested as a starting point for further optimization [146]. Furthermore, the analogue of this compound coupled to BSA (fig. 3.2B) for multivalent presentation was found to induce cellular signaling via JNK pathway [147].

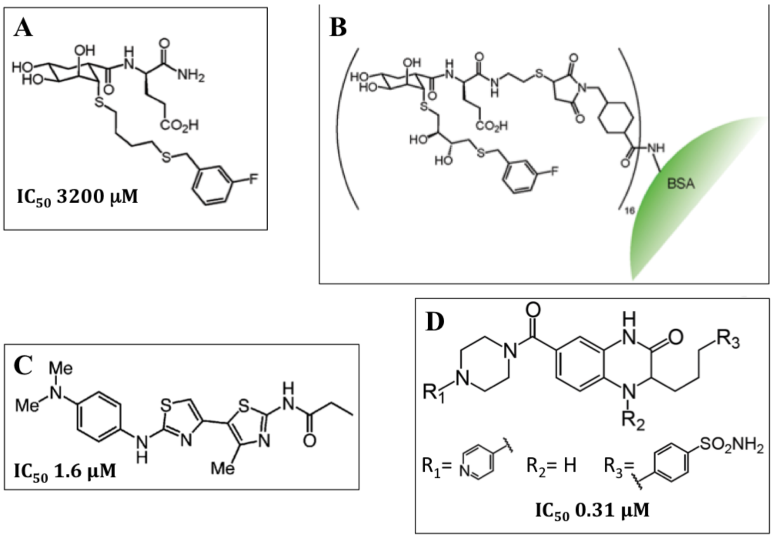


**Figure 3.2.:** Examples of non-carbohydrate ligands of DC-SIGN.

**A**, Shikimic acid derivative [146]. **B**, The analogue of compound in panel **A**, coupled to ManBSA [147]. **C**, A compound identified by high-throughput screening of two libraries [148]. **D**, quinoxalinone derivative [149]. **A**, **C** and **D**: $IC_{50}$ determined by competition assay of ManBSA binding to immobilized DC-SIGN ECD.

The other examples of non-carbohydrate DC-SIGN inhibitors include the small molecules ($IC_{50}$ values of 1.6-10 µM) identified by high-throughput fluorescence-based competition assay from two commercial libraries comprising about 36 000 compounds (fig. 3.2C) [148], and the more recent potent inhibitors ($IC_{50}$ values of 0.31-329 µM) derived from a collection of quinoxalinones (fig. 3.2D)

[149].

It is very important to note, that the selectivity issue is essential for the development of DC-SIGN antagonists. As it has been described in sub-subsection 2.2.1.2, the closely related lectin langerin has a protective role against HIV infection, and its location in mucosal surfaces makes it the very first barrier for the invading virus. Therefore, the compounds must be designed with strict selectivity to DC-SIGN and not langerin so that they wouldn't interfere with this protective role of langerin.

## 3.4. Multivalent presentation of the active substances

In biological systems, the interactions of DC-SIGN with its ligands, especially PAMPs, are of multivalent nature, which results in high overall affinity through the avidity effects. There are several factors that build up this multivalency. Firstly, DC-SIGN is expressed on cell surfaces in a tetrameric form and these tetramers are further clustered into microdomains in lipid rafts (see subsections 2.3.2 and 2.3.3), thus the binding sites are presented in a concentrated manner at the cell surface. Secondly, the pathogens usually present their surface glycans also in a clustered way. For instance, the high mannose type glycans of HIV envelope glycoprotein gp120 (fig. 3.3A) appear to be clustered resulting in an unusual density (fig. 3.3B) [150].

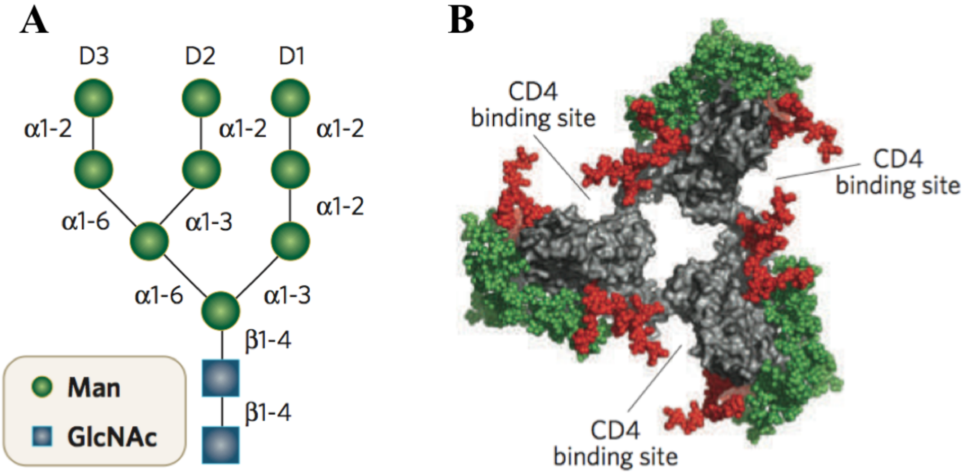


**Figure 3.3.:** The glycan shield of gp120.
**A**, Schematic representation of gp120 N-glycan undecasaccharide showing the structure, glycosidic linkages and identity of the D1, D2 and D3 termini of the A, B and C arms. **B**, molecular model of gp120 with the point of view from the viral membrane: the glycosylation surface for oligomannose glycans, found on the densely glycosylated outer domain, is shown in green, and in red is the glycosylation surface for complex sugars, which are distributed on the more exposed receptor-binding sites and hypervariable loops. (Adapted from [150])

To cope with such multivalent interactions, the design of DC-SIGN inhibitors must include the strategies of multimeric presentation of the developed monovalent ligands. The use of gold nanoparticles (GNPs) for ligand clustering is one of the possible approaches to attain multivalency of DC-SIGN inhibitors. Such mannose glyconanoparticles (manno-GNPs) are water-soluble nanoclusters with a three-dimensional display of about 100 ligands and a globular shape [151]. This strategy was used to cluster at variable density various oligomannosides that are the structural motifs of the gp120 undecasaccharide $Man_9(GlcNAc)_2$ (fig. 3.3A and 3.4A,B), and yielded very potent multivalent DC-SIGN inhibitors (table 3.4) (reviewed in [152]): Manα1-2Man based manno-GNP had four orders of magnitude higher potency than its monovalent counterpart [151]. These manno-GNPs (*tetra*-manno-GNP, fig. 3.4A) were found to be internalized into early endosomes by immature DCs in a CLR-dependent (50%) manner [153]. However, the *in vivo* application of GNPs might be limited because of potential toxicity of gold accumulation.

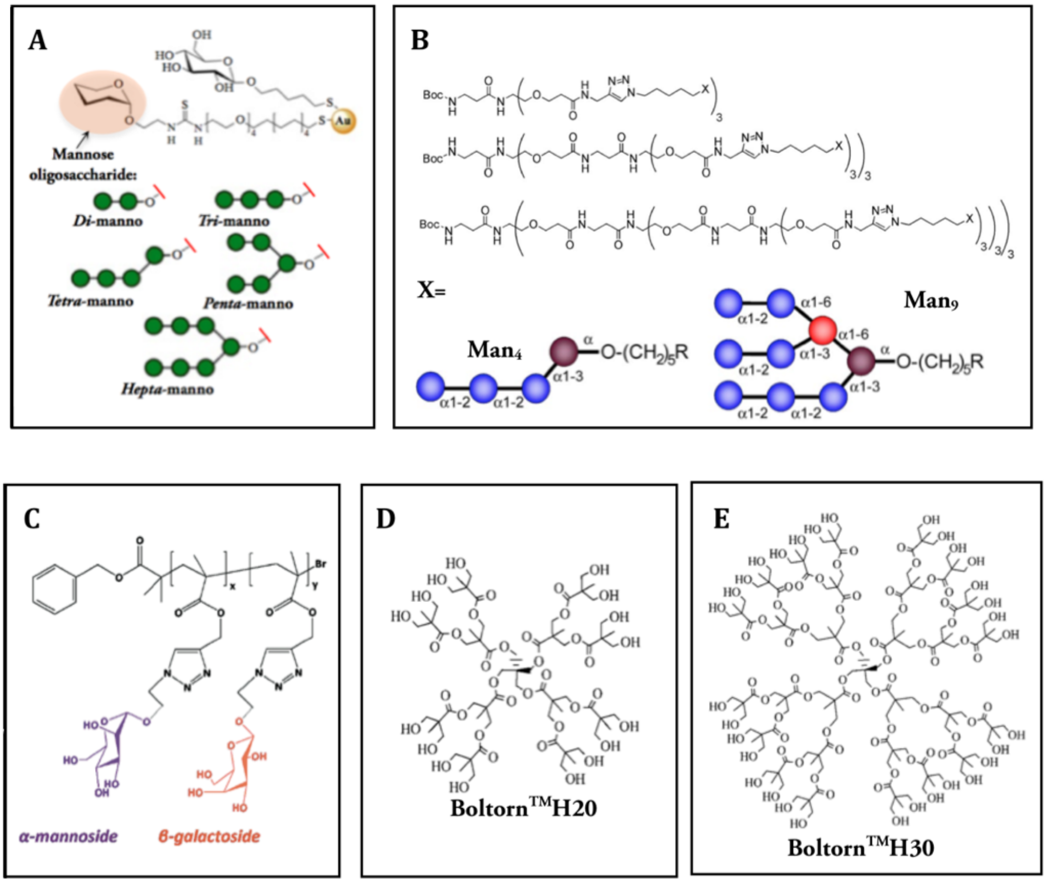


**Figure 3.4.:** The structures of multivalent presentation platforms.
**A**, manno-GNPs bearing different oligosaccharides; a glucose conjugate is used to control the density of oligomannosides on the manno-GNPs [154]. **B**, alkynyl dendrimer scaffolds of 1G (upper panel), 2G (middle panel) and 3G (lower panel) [155]. **C**, a glycopolymer with variable α-mannose and β-galactose ratios [156]. **D** and **E**, Boltorn type dendritic polymers of 2nd generation (2G, Boltorn™H20) in **D**, and 3rd generation (3G, Boltorn™H30) in **E** [135].

The alkynyl dendrimers (fig. 3.4B) with conjugated complex oligomannoses were shown to be excellent inhibitors of DC-SIGN (table 3.4) [155]. However, their applicability as therapeutical agents is limited due to high complexity of the used oligomannoses [128].

Glycopolymers with varying α-mannose and β-galactose ratios (fig. 3.4C) have been also proved to efficiently bind DC-SIGN with a favor to mannose-saturated polymers ($IC_{50}$ 37 nM for polymer with 100% mannose content compared to 1453 nM for 25%) [156]. The major problem for the *in vivo* use of such substances would probably be caused by their relatively high viscosity, a common feature of linear hydrophilic polymers, hindering their applicability as drug substances.

**Table 3.4.:** Some multivalent DC-SIGN inhibitors and their potencies.

| Entry | Platform | Ligand | Potency | Assay | Ref. |
|---|---|---|---|---|---|
| **1** | Boltorn$^{TM}$H30 (valence = 32) | Mannose | $IC_{50}$ 337 nM | Inhibition of DC-SIGN-mediated Ebola infection | [157] |
| | | | $IC_{50}$ 50 µM | SPR competition, gp120 coated surface | [158] |
| **2** | GNP (valence = 22) | *Tetra*-manno, 50% density | $IC_{50}$ ≈0.3 nM | Inhibition of DC-SIGN-mediated HIV *trans* infection | [154] |
| | GNP (valence = 56) | *Di*-manno, 50% density | $IC_{50}$ ≈2.0 nM | | |
| **3** | 2G alkynyl dendrimer (valence = 9) | $Man_4$ | $IC_{50}$ 160 nM | Inhibition of DC-SIGN-Fc binding to $Man_4$ coated glass slide | [155] |
| | | | $IC_{50}$ 20 nM | gp120/DC-SIGN-Fc ELISA | |
| | | $Man_9$ | $IC_{50}$ 26 nM | Inhibition of DC-SIGN-Fc binding to $Man_4$ coated glass slide | |
| | | | $IC_{50}$ 8 nM | gp120/DC-SIGN-Fc ELISA | |
| **4** | Boltorn type dendron (valence = 4) | Mannose | $IC_{50}$ 50 µM | Inhibition of DC-SIGN-mediated HIV *trans* infection | [49] |
| | | Pseudo mannotrioside (table 3.2, entry 3) | $IC_{50}$ 5 µM | | |

One of the common platforms for the multivalent presentation is offered by dendrimers, the non-toxic macromolecules with suitable rheological properties, defined globular shape that provides concentrated multiple attachment points to tether monovalent compounds [135]. Several groups have been working on development of such glycodendritic structures. Among them Dr. Javier Rojo group, our current collaborator in ITN CARMUSYS, suggested such glycodendritic structures anchoring multiple mannose residues conjugated to Boltorn type dendrimers (fig. 3.4D,E) as DC-SIGN inhibitors. They were tested in our group, and indeed, showed good inhibitory potency (table 3.4) [158]. It is noteworthy that these glycodendrimers were among the first ones demonstrated to inhibit DC-SIGN interaction with HIV gp120 glycoprotein, i.e. shown as potential HIV early infection inhibitors. The later collaborative studies of our group have shown that even four-valent Boltorn type dendrons were efficient inhibitors of DC-SIGN in HIV *trans* infection (table 3.4) [49].

# 4. The scope and objectives of this thesis

The research work on the development of DC-SIGN inhibitors has started in the team of Pr. Franck Fieschi (later on I refer to it as “our group”) several years before I joined it to carry out PhD studies, and has been run in collaboration with two chemist groups: the Italian team led by Pr. Anna Bernardi (University of Milan) and the Spanish team directed by Dr. Javier Rojo (Instituto de Investigaciones Químicas, CSIC; University of Sevilla).

During this period before my PhD started, our group has established and developed the overexpression and purification protocols of several constructs of recombinant human lectins DC-SIGN and Langerin. These lectins are used for the structural and functional studies of the glycomimetic compounds carried out by our team as well as by our collaborators. Our group has also developed Surface Plasmon Resonance (SPR) competition assay for evaluation of the interaction of the glycomimetic inhibitors with DC-SIGN, which allows efficient identification of the lead compounds. Our partners in Spain focus on the development of polyvalent scaffolds that would present different multivalency as well as different geometries and spacing, and to which the selected monovalent leads can be attached. As a start, the Boltorn type dendrimers carrying multiple mannose residues, were challenged as DC-SIGN/gp120 inhibitors [158]. This work was among the very first attempts to inhibit DC-SIGN/gp120 interaction using glycodendritic structures, and indeed, demonstrated that inhibitory power of such glycodendrimers is much higher than the monovalent D-mannose, which gave an encouragement to pursue more potent multivalent inhibitors.

On the other side, Pr. A. Bernardi's team has been working on the design and synthesis of monovalent mannose and fucose-based glycomimetics. Before my PhD, they have developed two mannose-based glycomimetics, namely pseudomannobioside and pseudomannotrioside (table 3.2, entries 2 and 3), and two Lewis$^X$ trisaccharide based mimics (table 3.3), and they were assessed as DC-SIGN inhibitors by our group and other partners [143, 49, 145].

The research work of the project continued and evolved into a successful European ITN CARMUSYS. Many other research groups with high-level expertise in different fields have joined the project including carbohydrate chemists from Oxford (Pr. Benjamin Davis group, University of Oxford), Berlin (Pr. Peter Seeberger group, Max Planck Institute of Colloids and Interfaces) and Prague (Pr. Jitka Moravcová group, Institute of Chemical Technology), cell biologists and immunologist from Amsterdam (Pr. Yvette van Kooyk group, VU University Medical Centre, DC4U), Madrid (Dr. Rafael Delgado group, Laboratorio de Microbiología Molecular, Hospital Universitario 12 de Octubre) and Paris (Dr. Ali Amara group, Inserm, Institute Pasteur), nuclear magnetic resonance (NMR) experts from Seville (Dr. Pedro Manuel Nieto Mesa group, Instituto de Investigaciones Químicas, CSIC, University of Sevilla), and computational chemists from Mannheim (Dr. Jörg Weiser group, Anterio consult & research). I have joined the network in October 2009, i.e. nine months after it has started, at the stage when the second generation fucose-based monovalent compound series had to be evaluated, and this PhD work is an integral part of the general workflow of the ITN CARMUSYS shown in figure 4.1.

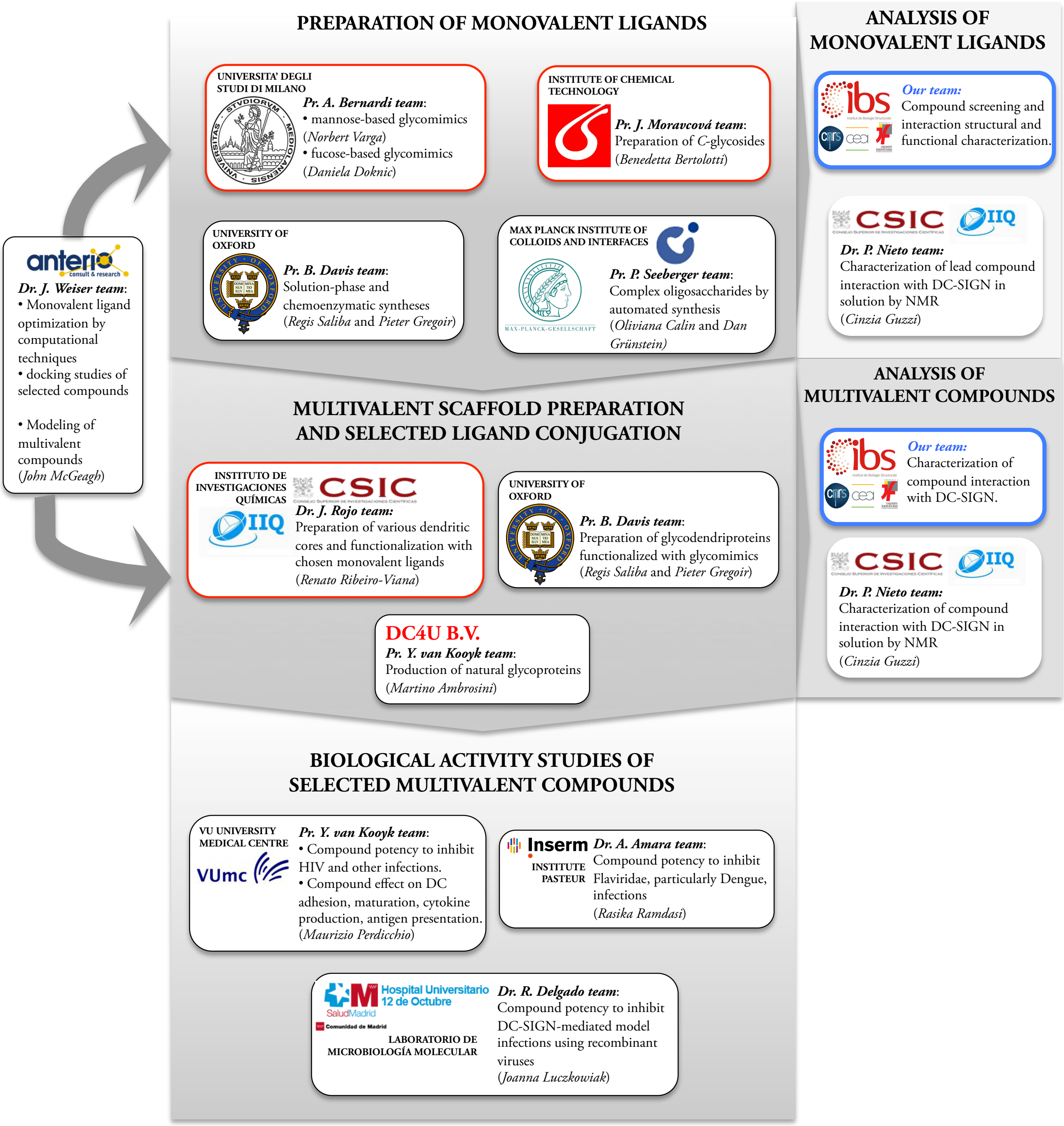

PREPARATION OF MONOVALENT LIGANDS
UNIVERSITA' DEGLI STUDI DI MILANO
Pr. A. Bernardi team:
• mannose-based glycomimics (Norbert Varga)
• fucose-based glycomimics (Daniela Doknic)
INSTITUTE OF CHEMICAL TECHNOLOGY
Pr. J. Moravcová team: Preparation of C-glycosides (Benedetta Bertolotti)
UNIVERSITY OF OXFORD
Pr. B. Davis team: Solution-phase and chemoenzymatic syntheses (Regis Saliba and Pieter Gregoir)
MAX PLANCK INSTITUTE OF COLLOIDS AND INTERFACES
Pr. P. Seeberger team: Complex oligosaccharides by automated synthesis (Oliviana Calin and Dan Grünstein)
ANALYSIS OF MONOVALENT LIGANDS
ibs
Our team: Compound screening and interaction structural and functional characterization.
CSIC IIQ
Dr. P. Nieto team: Characterization of lead compound interaction with DC-SIGN in solution by NMR (Cinzia Guzzi)
anterio consult & research
Dr. J. Weiser team:
• Monovalent ligand optimization by computational techniques
• docking studies of selected compounds
• Modeling of multivalent compounds
(John McGeagh)
MULTIVALENT SCAFFOLD PREPARATION AND SELECTED LIGAND CONJUGATION
INSTITUTO DE INVESTIGACIONES QUÍMICAS
CSIC IIQ
Dr. J. Rojo team: Preparation of various dendritic cores and functionalization with chosen monovalent ligands (Renato Ribeiro-Viana)
UNIVERSITY OF OXFORD
Pr. B. Davis team: Preparation of glycodendriproteins functionalized with glycomimics (Regis Saliba and Pieter Gregoir)
DC4U B.V.
Pr. Y. van Kooyk team: Production of natural glycoproteins (Martino Ambrosini)
ANALYSIS OF MULTIVALENT COMPOUNDS
ibs
Our team: Characterization of compound interaction with DC-SIGN.
CSIC IIQ
Dr. P. Nieto team: Characterization of compound interaction with DC-SIGN in solution by NMR (Cinzia Guzzi)
BIOLOGICAL ACTIVITY STUDIES OF SELECTED MULTIVALENT COMPOUNDS
VU UNIVERSITY MEDICAL CENTRE
VUmc
Pr. Y. van Kooyk team:
• Compound potency to inhibit HIV and other infections.
• Compound effect on DC adhesion, maturation, cytokine production, antigen presentation. (Maurizio Perdicchio)
Inserm INSTITUTE PASTEUR
Dr. A. Amara team: Compound potency to inhibit Flaviridae, particularly Dengue, infections (Rasika Ramdasi)
Hospital Universitario 12 de Octubre
SaludMadrid Comunidad de Madrid
LABORATORIO DE MICROBIOLOGÍA MOLECULAR
Dr. R. Delgado team: Compound potency to inhibit DC-SIGN-mediated model infections using recombinant viruses (Joanna Luczkowiak)

**Figure 4.1.:** The principal architecture of ITN CARMUSYS and our group involvement. The partners, with whom we have the most relevant collaboration, are highlighted by red frames; our team highlighted by a blue frame. The early-stage or experienced researchers (i.e. PhD or post-doc students) hired by CARMUSYS and working on the project in corresponding groups are written in *italic* in the brackets.

This multidisciplinary network functions on the basis of collaboration between different groups with a common final goal to design new tools for the development of DC-SIGN targeting antiviral drugs, with HIV infection being of particular interest. The principal workflow consists of several stages (fig. 4.1).

The first stage involves the monovalent glycomimetic ligand design and optimization: the chemist groups work on the design of the ligands with the computational aid and advices coming from Anterio consult & research. The series of synthesized compounds then are screened for their activity to bind DC-SIGN and selectivity versus langerin, and this work is mainly done in our team. The selected hits are structurally characterized for their interaction with DC-SIGN in order to get structural information for the further rational optimization of the compounds. This work is done also by our team (X-ray crystallography) and by Dr. P. Nieto team (NMR studies). At this stage, Anterio consult & research also has its part in providing the results of docking studies as a working model for the NMR team to facilitate the experimental result interpretation. On the other hand, the X-ray structures from our team also serve for the interpretation of NMR data as well as development of the docking protocols.

Simultaneously to the monovalent ligand development, various polyvalent scaffolds are designed in the team led by Dr. J. Rojo with computational support from Anterio consult & research. The suitability of these scaffolds, with mannose residues conjugated to them, is evaluated in our team by SPR tests. Then the selected ligands are mounted on chosen scaffolds and again tested as DC-SIGN binders in SPR tests.

The final part of the project is the evaluation of the selected multivalent compounds in biological infection inhibition assays.

Hence, the main objective of this thesis is to develop the selective DC-SIGN antagonists, and its accomplishment involves the following main tasks:

- ✧ Monovalent ligand screening by SPR competition assay to:
  - determine the best DC-SIGN binders;
  - evaluate selectivity to DC-SIGN versus langerin.
- ✧ Structural characterization of lead compound binding to DC-SIGN by X-ray crystallography.
- ✧ Hydrodynamic/thermodynamic characterization of DC-SIGN interaction with compounds of interest.
- ✧ Multivalent compound screening by SPR assays:
  - set-up/optimization of assays;
  - determination of compound potency to inhibit DC-SIGN (competition assay);
  - evaluation of direct interaction of compounds with DC-SIGN.

Along these tasks, the part of my work is over-expression and purification of DC-SIGN ECD and CRD constructs, required for SPR and other tests and co-crystallization as well as to provide other partners for their experiments.

Additionally, since the compound selectivity studies are part of this thesis, one more task is the development of production of other C-type lectins, namely LSECtin and MGL.

# Part II.

# Materials & Methods

# 5. Materials

## 5.1. Materials and equipment used for cloning

The enzymes NdeI, XhoI XbaI, HindIII, BsaI and Pfu DNA polymerase were from **Fermentas-Thermo Fisher Scientific Inc**. The enzyme DpnI was used from QuikChange kit, **Agilent Technologies**.

The buffers *O*, *R*, *Tango*, *10X Pfu*, the DNA size marker “Mass Ruler DNA Ladder Mix ready-to-use”, the dNTPs, and the “6X Orange DNA Loading Dye” were from **Fermentas-Thermo Fisher Scientific Inc**. Agarose and UltraPure$^{TM}$ agarose were from **Roche** and **Invitrogen**, respectively. Ethidium bromide was from **Fluka**.

The “Mastercycler gradient” PCR instrument from **eppendorf** and the “Mini Horizontal Gel Electrophoresis System” from **Major Science** were used.

The kits “QIAquick PCR Purification”, “QIAprep Spin”, and “QIAquick gel purification” were from **QIAGEN**. DNA Ligation Kits were from **Roche** and **Fermentas-Thermo Fisher Scientific Inc**. “QuikChange Lightning: Site Directed Mutagengesis Kit” was from **Agilent Technologies**.

The vectors pEt20b and pEt30b were from **Novagen**, and vector pASK-IBA6 was from **IBA**.

## 5.2. Materials and equipment used for bacterial cultures

*E.coli* strains BL21(DE3) and Rosetta2(DE3) were from **Novagen**, and the TOP10 strain was from **Invitrogen**.

LB (Lennox) broth and LB-Agar were from **AthenaES** and **Sigma**, respectively. Ampicillin, kanamycin and IPTG were from **Euromedex**. Anhydrotetracycline was from **Riedel-de Haën**.

## 5.3. Materials and equipment used for protein purification and characterization

### 5.3.1. *SDS-PAGE*

Bis-acrylamide (30%), SDS (20%) and TEMED were from **Euromedex**. β-mercaptoethanol was from **Carl Roth**. Acetic acid, ammonium persulphate and Brilliant blue R were from **Sigma**. Ethanol was from **Carlo Erba Reagents**. PageRuler Unstained and Prestained Protein Ladders were from **Fermentas-Thermo Fisher Scientific Inc**.

Electrophoresis system and its accessories were from **Bio-Rad**.

### 5.3.2. *Western Blot*

Mouse monoclonal anti-StrepTagII antibody from **QIAGEN** and peroxidase-labeled anti-mouse IgG antibody from **KPL-Eurobio** were used. The SDS-PAGE gels were blotted on PVDF membrane from **Bio-Rad**, and the blots were revealed using SIGMAFAST$^{TM}$ 3,3'-Diaminobenzidine staining kit from **Sigma**.

The transfer system and its accessories were from **Bio-Rad**.

### 5.3.3. *Protein purification*

#### 5.3.3.1. *Materials and equipement used for inclusion body extraction and refolding*

The bacterial cells were disrupted using BRANSON Digital Sonifier® from **Emerson Electric Co**. The Optima$^{TM}$LE-80K ultracentrifuge with Ti45 rotor was from **Beckman Coulter**.

Gdn-HCl and Triton X-100 were from **Euromedex** and **Anatrace®**, respectively. Reduced and oxidized glutathiones, urea and L-Arginine were from **Sigma**. "Complete EDTA-free protease inhibitor cocktail" was from **Roche**.

The dialyses were carried out using Spectra/Por® dialysis membrane (3.5 kDa MWCO, 9.3 mL/cm) from **Spectrum Laboratories**.

#### 5.3.3.2. *Materials and equipement used for protein purification*

The Mannan-Agarose (20 mL) and immobilized D-galactose (5 mL) were from **Sigma** and **Pierce**, respectively. StrepTactin (5 mL) and Superose12 (125 mL) columns were from **GE Healthcare**. D-desthiobiotin was from **IBA.**

The purification were performed using ÄktaFPLC system from **GE Healthcare**.

#### 5.3.3.3. *Protein samples concentration*

VIVASPIN concentrators with PES membrane (10 kDa MWCO) from **Sartorius** were used to concentrate proteins.

### 5.3.4. *Surface Plasmon Resonance*

The SPR experiments were conducted using the Biacore 3000 and Biacore X instruments from **Biacore-GE Healthcare**. Sensor chips, amine coupling kit (includes EDC, NHS and ethanolamine), HBS-P buffer (0.01 M HEPES pH 7.4, 0.15 M NaCl, 0.005% Surfactant P20), and 10% P20 were **from Biacore-GE Healthcare**.

L-fucose, D-mannose, ethylen glycol and DMSO were from **Sigma**. Lewis$^{X}$, α1-3(α1-6)-mannotriose BSA, Man α1-2 Man and Man α1-3 Man were from **Dextra Laboratories Ltd**.

### 5.3.5. *Equipement and materials for ITC, AUC, SLS/DLS and NMR experiments*

Equipment and materials for ITC, AUC and SLS/DLS experiments are described in article **n°5**. Equipment and materials for NMR experiments are described in article **n°6**.

---

All other used reagents were from **Sigma** and **Euromedex**.

# 6. Methods

## 6.1. Production of recombinant C-type lectin constructs

### 6.1.1. *Transformation of $Ca^{2+}$-competent E.coli strains*

Three $Ca^{2+}$-competent *E.coli* strains were used in this work, namely TOP10, BL21(DE3) and BL21 derivative strain Rosetta2(DE3), which are engineered for vector propagation (TOP10) and efficient recombinant protein over-expression.

The protocol of $Ca^{2+}$-competent bacterial cell transformation consists of the following steps:

1. On ice, add 1 µL of plasmid DNA to 200 µL of calcium competent cells, gently shake.
2. Incubate on ice for 30 min.
3. Heat-shock: 45 s in 42°C water, put back onto ice.
4. Add 1 mL of Luria-Bertani (LB) broth and incubate at 37°C for 1 hour.
5. Centrifuge at 5000xg for 5 min (room temperature), discard supernatant leaving a small volume in the tube.
6. Resuspend pellet in a residual volume of supernatant.
7. Plate the resuspended cells on Petri dishes with LB-Agar containing respective antibiotic and incubate at 37°C overnight.
8. Select one clone from a plate and inoculate it to 200 mL of LB with respective antibiotic.
9. Cultivate at 37°C overnight to produce the preculture.

The prepared preculture is used to start the culture for protein over-expression (in case of BL21(DE3) or Rosetta2(DE3) cells) or to purify the vector from the cells (in case of TOP10 cells). All below described plasmid purifications were performed using QIAprep Spin kit according to manufacturer's instructions, and the concentration of plasmid DNA was determined by measuring $A_{260}$ of the sample and knowing that 50 µg/mL double-stranded DNA has $A_{260} = 1$.

### 6.1.2. *Production of DC-SIGN S-ECD, ECD, S-CRD and langerin ECD constructs*

The cloning of DC-SIGN extracellular domain comprising amino acids 66-404 and containing N-terminal StrepTag II (S-ECD) has been done previously and is described in PhD thesis of Dr. Georges Tabarani, which can be found at `http://www.sudoc.fr/137188102`.

The cloning, over-expression and purification of DC-SIGN ECD and S-CRD constructs has been also set up previously and is published by Tabarani et al [80] and described in submitted article by Thépaut et al (article **n°4**, see table 7.1), respectively.

The over-expression and purification of DC-SIGN ECD, DC-SIGN S-ECD and DC-SIGN S-CRD are described in sub-subsections 6.1.2.1, 6.1.2.2 and 6.1.2.3. The extra-purification of DC-SIGN ECD is described in supplementary information of paper manuscript **n°5** (see table 7.1). Langerin ECD constructs were contributed by Eric Chabrol, a PhD student in our group.

### 6.1.2.1. *Over-expression and purification of DC-SIGN ECD*

**Over-expression.** The *E.coli* strain BL21(DE3) was used for over-expression of DC-SIGN ECD. The cells were transformed with pEt30b/DC-SIGN ECD plasmid according to protocol in subsection 6.1.1. The culture for over-expression of DC-SIGN ECD was started by inoculation of 50 mL of overnight preculture to 1 L of LB broth with 50 µg/mL kanamycin. The cells were grown at 37°C for 2 h, 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 overnight. The cells were harvested by centrifugation at 5000xg for 30 min at 4°C, resuspended in 25 mM Tris-HCl pH 8 buffer and again centrifuged.

**Inclusion body extraction.** The pellet of the cells of 1 L culture was resuspended in 30 mL of the buffer consisting of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and containing 1 tablet of "complete EDTA-free" protease inhibitor cocktail. The cells were disrupted by freezing at -20°C and thawing followed by sonication for two rounds at 90% amplitude for total time of 2 min for each round using 2 s long sonications and 10 s long pauses in-between and keeping cells in ice. The inclusion bodies together with cell debris were collected by centrifugation at 100 000xg for 30 min at 4°C. 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, 2 M urea and 1% Triton-X100, centrifuged at 100 000xg for 1 h at 4°C, then rinsed by resuspending them 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.

**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 1 h at 4°C to eliminate insoluble part. Solubilized DC-SIGN ECD (1.9 mg/mL, 135 mL) was refolded by flash dilution of 5x to a buffer consisting of 1.25 M NaCl, 25 mM Tris-HCl pH 8 and 25 mM $CaCl_2$, at 4°C. The resulting protein solution was dialyzed overnight against 3960 mL of 25 mM Tris-HCl pH 8 buffer, at 4°C. Three more dialyses (of 3 h each at 4°C) against 4.5 L of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ continued and finished refolding. The precipitates were eliminated by centrifugation at 100 000xg for 1 h at 4°C.

**Purification.** The purification of refolded DC-SIGN ECD was performed in two steps: by an affinity chromatography followed by gel-filtration. All purifications were carried out at 4°C.

Firstly, the refolded DC-SIGN ECD solution (1450 mL, 0.163 mg/mL) was loaded to Mannan-agarose column (20 mL) prior equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$

buffer. After washing the column with the same buffer (100 mL), a step elution of the bound protein was done using the same buffer but with $CaCl_2$ replaced by 1 mM EDTA (70 mL). The flow rate of 2.5 mL/min was maintained during the purification. The whole procedure was repeated twice using the flow-through of the prior purifications. To the eluted protein, $CaCl_2$ was added to final concentration of 10 mM and the protein was concentrated to about 3 mg/mL using VIVASPIN concentrators.

In order to remove EDTA and recharge the protein with $Ca^{2+}$, the gel-filtration was performed. At this step, the concentrated and centrifuged (10 min, 20800xg, 4°C) protein was repeatedly injected (about 2 mL each time) to Superose-12 column equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and eluted with the same buffer (150 mL). The flow rate of 1 mL/min was maintained during the procedure. Eluted protein fractions pooled and concentrated to 9 mg/mL, 0.5 mL aliquots prepared and frozen for storage at -80°C.

#### 6.1.2.2. *Over-expression and purification of DC-SIGN S-ECD*

**Over-expression.** The *E.coli* strain Rosetta2(DE3) (the strain that was previously established as efficient for over-expression of this construct) was used for over-expression of DC-SIGN S-ECD. The cells were transformed with pEt20b/DC-SIGN S-ECD plasmid according to protocol in subsection 6.1.1. The culture for over-expression of DC-SIGN SECD was started by inoculation of 100 mL of overnight preculture to 2 L of LB broth with 0.1 mg/mL ampicillin. The cells were grown at 37°C for 2 h, 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 3 h. The cells were harvested by centrifugation at 5000xg for 30 min at 4°C, resuspended in 25 mM Tris-HCl pH 8 buffer and again centrifuged.

**Inclusion body extraction.** The pellet of the cells of 2 L culture was resuspended in 30 mL of the buffer consisting of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and containing 1 tablet of "complete EDTA-free" protease inhibitor cocktail. The cells were disrupted by freezing at -20°C and thawing followed by sonication for two rounds at 90% amplitude for total time of 2 min for each round using 2 s long sonications and 10 s long pauses in-between and keeping cells in ice. The inclusion bodies together with cell debris were collected by centrifugation at 100 000xg for 30 min at 4°C. 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, 2 M urea and 1% Triton-X100, centrifuged at 100 000xg for 1 h at 4°C, then rinsed by resuspending them 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.

**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 1 h at 4°C to eliminate insoluble part. Solubilized DC-SIGN S-ECD (30 mL, 0.57 mg/mL) was refolded by drop-by-drop dilution to 120 mL of a buffer consisting of 1.25 M NaCl, 25 mM Tris-HCl pH 8 and 25 mM $CaCl_2$, at 4°C. The resulting protein solution was dialyzed overnight against 880 mL of 25 mM Tris-HCl pH 8 buffer, at 4°C. Three more dialyses (of 3 h each at 4°C) against 4.5 L of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ continued and finished refolding. The precipitates were eliminated by centrifugation at 100 000xg for 1 h at 4°C.

**Purification.** The purification of refolded DC-SIGN S-ECD was performed by two steps of affinity chromatography followed by gel-filtration. All purifications were carried out at 4°C.

Firstly, the refolded DC-SIGN S-ECD solution (260 mL, 0.093 mg/mL) was loaded to 5 mL StrepTactin column prior equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer. After washing the column with the same buffer (100 mL), a step elution of the bound protein was done using the same buffer supplemented with 2.5 mM desthiobiotin (70 mL). The flow rate of 2.5 mL/min was maintained during all procedure. The eluted protein was concentrated.

At the second purification step, the concentrated protein was centrifuged for 10 min at 20800xg, 4°C, to eliminate any insoluble parts, and 1.5 mL (7.6 mg/mL) was injected to Mannan-Agarose column equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$. After washing the column with the same buffer (80 mL), a step elution of the bound protein was done with 60 mL of buffer consisting of 25 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA. The flow rate of 2.5 mL/min was maintained throughout whole procedure. The eluted protein was concentrated.

In order to remove EDTA and recharge the protein with $Ca^{2+}$, a gel-filtration was performed. At this step, the concentrated and centrifuged (10 min, 20800xg, 4°C) protein was injected (1 mL, 7.59 mg/mL) to Superose-12 column equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and eluted with the same buffer (150 mL). The flow rate of 1 mL/min was maintained during the procedure. Eluted protein fractions pooled and concentrated to 15.08 mg/mL, aliquoted to 20 µL and frozen for storage at -80°C.

### 6.1.2.3. *DC-SIGN S-CRD production*

**Over-expression.** The *E.coli* strain BL21(DE3) was used for over-expression of DC-SIGN S-CRD (amino acids 250-404) construct. The cells were transformed with pASK-IBA6/DC-SIGN S-CRD plasmid according to protocol in subsection 6.1.1. The culture for over-expression of DC-SIGN SECD was started by inoculation of 25 mL of overnight preculture to 0.5 L of LB broth with 0.1 mg/mL ampicillin. The cells were grown at 37°C until $OD_{600}$ of 1.11 was reached, then the protein over-expression was induced by addition of anhydrotetracycline to final concentration of 0.2 µg/mL and the culture was continued at 37°C overnight. The cells were harvested by centrifugation at 5000xg for 30 min at 4°C, resuspended in 25 mM Tris-HCl pH 8 buffer and again centrifuged.

**Inclusion body extraction and refolding.** Inclusion bodies were extracted by selective centrifugation steps and solubilized in 6 M Gdn-HCl buffer as already described for DC-SIGN ECD and S-ECD. At this step, the screening for refolding conditions was performed. For this reason 0.2 mL protein solution (0.64 mg/mL) was diluted to 0.8 mL of the following buffers:

1. Buffer-1: 1.25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM $CaCl_2$.

2. Buffer-2: 1.25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM $CaCl_2$, 1 mM oxidized glutathione, 10 mM reduced glutathione.

3. Buffer-3: 1.25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM $CaCl_2$, 0.16 M L-Arginine.

4. Buffer-4: 1.25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM $CaCl_2$, 0.16 M L-Arginine, 1 mM oxidized glutathione, 10 mM reduced glutathione.

The UV absorbance spectra (240-500 nm) were recorded for each sample using the corresponding buffer as a blank, and compared with the protein solution 5x diluted in inclusion body solubilization buffer as a control sample to monitor the baseline.

Buffer-4 was chosen suitable and the protein (150 mL, 0.64 mg/mL) was refolded by drop-by-drop dilution to 600 mL of this buffer. First overnight dialysis was performed against 4.4 L of 25 mM Tris-HCl pH 8 buffer. Two more dialyses for 3 hours each one were performed against 4.5 L of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer.

**Purification.** Then the protein solution was centrifuged and the supernatant was loaded to Strep-Tactin column prior equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer, supplemented with oxidized and reduced glutathiones at 1 mM and 10 mM. After washing the column with the same buffer (300 mL), a step elution of the bound protein was done using the same buffer supplemented with 2.5 mM desthiobiotin (70 mL). The flow rate of 2.5 mL/min was maintained during all procedure at 4°C. The eluted protein was concentrated.

Mannan-Agarose column was used for the second step of purification. The concentrated elution peak from StrepTactin was repeatedly injected (about 2 mL) to Mannan-Agarose column prior equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer. After 60 mL of the buffer was flown through the column, a step elution with EDTA (elution buffer: 25 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA) followed to eliminate the bound protein from the column. The flow rate of 2.5 mL/min was maintained during all procedure at room temperature.

### 6.1.3. *Cloning of S-ECD and S-CRD constructs of LSECtin and MGL*

#### 6.1.3.1. *Generation and amplification of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences*

The plasmid pcDNA3.1 containing the cDNA encoding full-length LSECtin (liver and lymph node sinusoidal endothelial cell C-type lectin) was kindly donated by Dr. Angel L. Corbí, and the pRc/CMV plasmid containing the cDNA encoding full-length MGL (macrophage galactose-type C-type lectin) isoform 2 was granted by Pr. Yvette van Kooyk.

The cDNA sequences coding S-CRD and S-ECD constructs of LSECtin and MGL were generated by PCR using corresponding plasmids as matrices and the following synthetic oligonucleotides as primers (ordered from COGENICS) that included NdeI, XhoI cleavage sites and StrepTag II coding sequence followed by a sequence coding for factor Xa cleavage site to be placed in N-terminus of the protein:

✧ LSECtin S-CRD:

  - Forward *lsec-crd-f.* 5´-GGA GAT ATA CAT ATG GCT AGC TGG AGC CAC CCG CAG TTC GAA AAA ATC GAA GGG CGC AAC AAC TCC TGC GAG CCG TGC-3´;
  - Reverse *lsec-crd-r.* 5´-GTG GTG GTG CTC GAG TCA GCA GTT GTG CCT TTT CTC-3´.

- LSECtin S-ECD:
  - Forward *lsec-ecd-f.* 5´-GGA GAT ATA CAT ATG GCT AGC TGG AGC CAC CCG CAG TTC GAA AAA ATC GAA GGG CGC AAG GCC TCC ACG GAG CGC GCG-3´.
- MGL S-CRD:
  - Forward *mgl-crd-f.* 5´-GGA GAT ATA CAT ATG GCT AGC TGG AGC CAC CCG CAG TTC GAA AAA ATC GAA GGG CGC GAA GGG ACC TGC TGC CCT GTC-3´;
  - Reverse *mgl-crd-r.* 5´-GTG GTG GTG CTC GAG TCA GTG ACT CTC CTG GCT GGT-3´.
- MGL S-ECD:
  - Forward *mgl-ecd-f.* 5´-GGA GAT ATA CAT ATG GCT AGC TGG AGC CAC CCG CAG TTC GAA AAA ATC GAA GGG CGC CAA AAT TCC AAA TTT CAG AGG-3´.

The received oligonucleotides were dissolved in 50 µL of TE (10 mM Tris-HCl pH 8, 1 mM EDTA) buffer and further diluted for PCR as shown in table 6.1.

**Table 6.1.:** Preparation of primers for generation of lectin constructs.

| Name | Quantity, µg | Prepared stock conc., µg/µL | Dilution factor to obtain 100 ng/µL | Volumes for dilution in *water* (µL) |
|---|---|---|---|---|
| *lsec-crd-f.* | 112 | 2.24 | 22.4 | 2.23 + *47.7* |
| *lsec-crd-r.* | 148 | 2.96 | 29.6 | 1.7 + *48.3* |
| *lsec-ecd-f.* | 99 | 1.98 | 19.8 | 2.52 + *47.4* |
| *mgl-crd-f.* | 141 | 2.82 | 28.2 | 1.8 + *48* |
| *mgl-crd-r.* | 144 | 2.88 | 28.8 | 1.7 + *48.4* |
| *mgl-ecd-f.* | 131 | 2.62 | 26.2 | 1.9 + *48* |

The mixtures for PCR were prepared as follows: 2.5 µL of corresponding diluted primers (forward+reverse), 1 µL of matrix (corresponding plasmid containing lectin coding cDNA), 2 µL dNTP, 1 µL Pfu DNA polymerase; 10 µL 10X Pfu buffer, and 81 µL sterile water (autoclaved). The reaction proceeded as follows (25 cycles of steps 2-4 were done):

1. 94°C, 1 min → denaturation;
2. 94°C, 45 s → denaturation;
3. 60°C, 1 min → annealing;
4. 72°C, 2 min → elongation;
5. 72°C, 10 min → elongation (to end all chains).

PCR reaction products were purified using QIAquick PCR Purification Kit as described by manufacturer using water for elution. The purified products were tested on 1% agarose (w/v) gel prepared in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) with 2 µL of 10X diluted ethidium bromide. 10 µL of each sample was mixed with 2 µL of 6X Orange DNA Loading Dye. The DNA fragment size marker "Mass Ruler DNA Ladder Mix ready-to-use" was used. 10 µL of each sample loaded to gel and TAE was used as running buffer.

#### 6.1.3.2. *Insertion of LSECtin and MGL S-ECD or S-CRD coding cDNA sequences to pET30b vector*

The purified amplified DNA sequences of coding LSECtin and MGL S-ECD or S-CRD, and pET30b were consecutively digested with restrictases NdeI and then XhoI. Reaction mixtures were prepared as shown in table 6.2.

**Table 6.2.:** Preparation of LSECtin and MGL S-ECD or S-CRD cDNA sequences and pET30b digestion mixtures.

| For PCR products | For vector |
|---|---|
| 30 µL of corresponding PCR product | 30 µL plasmid |
| 3 µL restrictase | 4 µL restrictase |
| 6 µL buffer | 6 µL buffer |
| 21 µL water | 20 µL water |

The first digestion was done with NdeI using buffer *O*. Each reaction was conducted for 2 hours at 37°C. Reaction products were purified using QIAquick PCR Purification Kit as described by manufacturer.

Purified samples were digested with XhoI using buffer *R*, and after digestion again purified by the same protocol. Digested samples were analyzed in 1% agarose gel as described above.

The digested PCR products were inserted into a digested pEt30b using Rapid DNA Ligation Kit from Roche. The reaction mixture was prepared as follows: 1 µL digested vector, 7 µL digested PCR product, 2 µL dilution buffer, mixing of the three components, then addition of 10 µL of ligation buffer, mixing, and 1 µL of ligase. The control mixture was prepared in the same way except 7 µL of water was added instead of PCR product. Ligation reaction was carried out at room temperature for 1 hour. Subsequently, the TOP10 E.coli cells were transformed with ligation reaction product as described in subsection 6.1.1, except 5 µL of DNA were added to 30 µL of cells. The cells were plated on Petri dish with LB-Agar in presence of 50 µg/mL kanamycin and grown at 37°C overnight.

From selected positive clones, pET30b/lectin construct plasmids were propagated and purified as described in section 6.1.1 and were tested by NdeI/XhoI digestion (reaction preparation in table 6.3).

**Table 6.3.:** Preparation of NdeI/XhoI digestion mixtures for the test of positive cDNA insertion.

| Enzyme mixture for 20 samples | Reaction mixtures |
|---|---|
| 20 µL buffer *O* | 5 µL enzyme mixture |
| 3.75 µL NdeI | 5 µL plasmid with cDNA |
| 7.5 µL XhoI | (for control: 3 µL of empty |
| 68.75 µL $H_2O$ | pEt30b + 2 µl $H_2O$) |

Reaction proceeded for 1 hour at 37°C. The reaction products were visualized in 1% agarose gel as described above. The positive samples were sent for sequencing to COGENICS.

#### 6.1.3.3. *Insertion of LSECtin and MGL S-ECD or S-CRD coding cDNA sequences to pET20b vector*

To extract the cDNA inserts, the generated pET30b plasmids containing corresponding lectin coding cDNA sequences were digested consecutively with NdeI and XhoI. The first step reaction mixtures contained 15 µL of pET30b/lectin cDNA, 2 µL NdeI, 3 µL buffer *O* and 10 µL water. Digestion was carried out for 2 h at 37°C, and reaction products were purified using QIAquick PCR Purification Kit following manufacturer's instructions and using elution with 30 µl water.

The second step digestion reaction mixtures contained 30 µL of purified first step reaction products, 2 µL XhoI, 4 µL buffer *R* and 4 µL water. Reaction of 2 h at 37°C was followed by purification with QIAquick PCR Purification Kit using elution with 30 µl water.

Further, the purified digestion products were inserted into pEt20b, which was prior digested with NdeI and XhoI in the same manner as described for pEt30b above. Ligation and subsequent TOP10 cell transformation were performed as described above for pEt30b construct generation. The cells were plated on Petri dish with LB-Agar in presence of 0.1 mg/mL ampicillin and grown at 37°C overnight.

After amplification and purification of the plasmids (described in subsection 6.1.1), they were tested for positive cDNA sequence insertion as described above for pEt30b constructs. The plasmids containing positive insertions were sent for sequencing to COGENICS.

### 6.1.4. *Cloning of LSECtin S-CRD construct for targeting to periplasm*

#### 6.1.4.1. *Construction of pEt30b vector with encoded OmpA signal sequence*

OmpA coding sequence was excised from pASK-IBA6 vector by consecutive digestion with XbaI and HindIII restrictases (fig. 6.1).

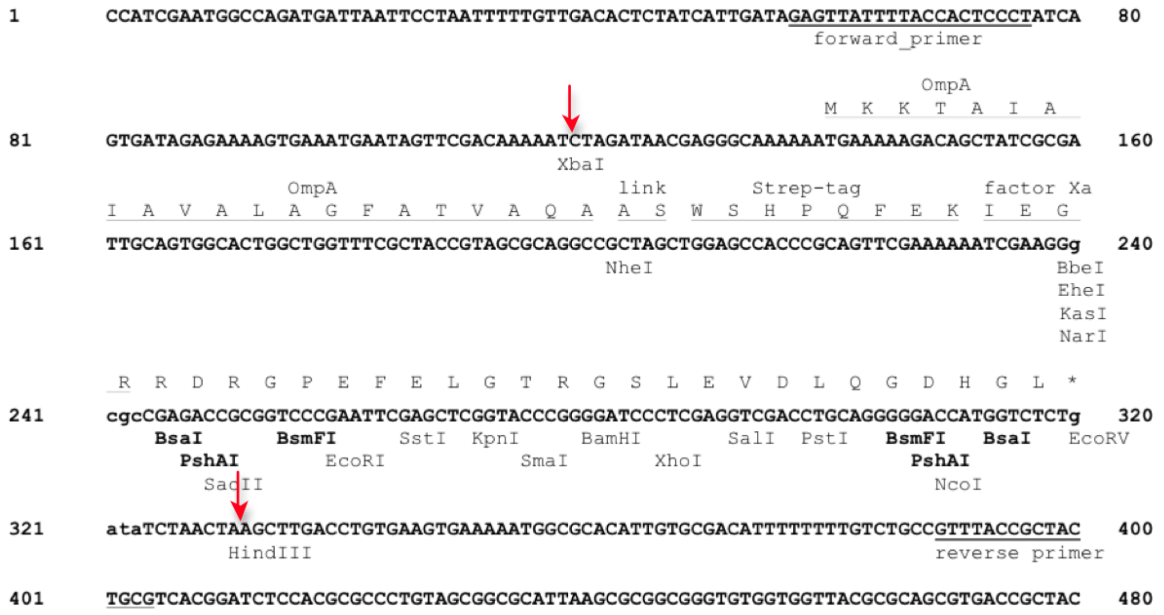

1 CCATCGAATGGCCAGATGATTAATTCCTAATTTTTGTTGACACTCTATCATTGATAGAGTTATTTTACCACTCCCTATCA 80
forward_primer
OmpA
M K K T A I A
81 GTGATAGAGAAAAGTGAAATGAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGAAAAAGACAGCTATCGCGA 160
XbaI
OmpA
link
Strep-tag
factor Xa
I A V A L A G F A T V A Q A A S W S H P Q F E K I E G
161 TTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCGCTAGCTGGAGCCACCCGCAGTTCGAAAAAATCGAAGGg 240
NheI
BbeI
EheI
KasI
NarI
R R D R G P E F E L G T R G S L E V D L Q G D H G L *
241 cgcCGAGACCGCGGTCCCGAATTCGAGCTCGGTACCCGGGGATCCCTCGAGGTCGACCTGCAGGGGGACCATGGTCTCTg 320
BsaI
BsmFI
SstI
KpnI
BamHI
SalI
PstI
BsmFI
BsaI
EcoRV
PshAI
EcoRI
SmaI
XhoI
PshAI
SacII
NcoI
321 ataTCTAACTAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTTGTCTGCCGTTTACCGCTAC 400
HindIII
reverse primer
401 TGCGTCACGGATCTCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTAC 480

**Figure 6.1.:** Multiple cloning site of pASK-IBA6.

First step digestion reaction mixture contained 100 µL plasmid, 3 µL XbaI, 12 µL *Tango* buffer and 5 µL water. In parallel, pEt30b vector was also digested with XbaI using the same reaction mixture composition. For the second step, the purified XbaI digests were further digested with HindIII (reaction mixture: 45 µL purified XbaI digests, 4 µL HindIII, 7 µL buffer *R* and 14 µL water) and subsequently purified. All digestions were performed for 2 h at 37°C, and reaction products were purified using QIAquick PCR Purification Kit following manufacturer's instructions and using elution with 50 µL of water.

Purified digests were loaded to 1.3% UltraPure agarose gel prepared in TAE buffer without ethidium bromide, other conditions as previously described. After staining the test gel in ethidium bromide bath, the digestion products, i.e. the digested pEt30b vector or ompA sequence from pASK-IBA6, were visualized, then excised from corresponding part of non-stained gel and purified using QIAquick gel purification kit according to manufacturer's instructions.

Ligation of purified ompA sequence to XbaI/HindIII digested pEt30b was performed using ligation kit from Fermentas and the following reaction mixture: 1 µL pEt30b, 13 µL ompA, 4 µL 5X buffer, 1 µL ligase. Ligation continued for 1 h at room temperature and subsequently the TOP10 cells were transformed with the reaction product by the protocol described in subsection 6.1.1. The cells were plated on Petri dish with LB-Agar in presence of 50 µg/mL kanamycin and grown at 37°C overnight. After propagating several clones and purifying plasmid as described in subsection 6.1.1, ompA insertion was checked by 4-hours long digestion by BsaI at 37°C, and positive samples were sent for sequencing to COGENICS.

### 6.1.4.2. *Generation of pEt30b-ompA/LSECtin S-CRD construct*

**Preparation of LSECtin S-CRD cDNA sequences.** At the first step, LSECtin S-CRD coding cDNA was amplified by PCR using previously constructed pEt30b/LSECtin S-CRD vector (subsubsection 6.1.3.2) as a matrix and the following primers (ordered from COGENICS):

✧ Forward *lsec-c-ompA-f.* 5´-CG CAG TTC GAA AAA ATC GAA GGG CGC AAC AAC TCC TGC GAG CCG TGC-3´;

✧ Reverese *lsec-c-ompA-r.* 5´- CAG AGA CCA TGG TCC CCC TGC TCA GCA GTT GTG CCT TTT CTC-3´,

where sequence part in blue anneals within ompA part in pEt30b-ompA, and the red sequence anneals to the beginning of CRD coding cDNA in pEt30b/LSECtin S-CRD. The preparation of primers shown in table 6.4.

**Table 6.4.:** Preparation of primers for LSECtin S-CRD cDNA sequence amplification.

| Primer | Quantity, nmol | Dissolved in TE, µL | Final conc., µM | Working conc. (dilution with water), µM |
|---|---|---|---|---|
| *lsec-c-ompA-f.* | 17.6 | 196 | 89.8 | 20 |
| *lsec-c-ompA-r.* | 37.8 | 420 | 90 | 20 |

The mixture for PCR was prepared using reagents from QuikChange Lightning: Site Directed Mutagengesis Kit as follows: 5 µL 10x buffer, 2 µL dNTP, 3 µL QuikSolution, 2 µL pEt30b/LSECtin S-CRD vector, 2.5 µL of each prime (20 µM); 1 µL Pfu DNA polymerase (L QC Enzyme), and 32 µL water.

The reaction proceeded as follows (35 cycles of steps 2-4 were done):

1. 95°C, 2 min;
2. 95°C, 1 min;
3. 55°C, 1 min;
4. 72°C, 3 min;
5. 72°C, 6 min.

The PCR was loaded to 1% UltraPure$^{\mathrm{TM}}$ agarose gel prepared in TAE buffer without ethidium bromide and other conditions were as described above. After staining in ethidium bromide bath, the band containing PCR product was excised and purified using QIAquick gel purification kit according to manufacturer's instructions and rechecked on 1% agarose gel.

**Insertion of LSECtin S-CRD coding cDNA sequence to pEt30b-ompA vector.** This step was accomplished by PCR using the generated LSECtin S-CRD coding cDNA as a primer and the generated pEt30b-ompA vector as a template.

The mixture for PCR was prepared using reagents from QuikChange Lightning: Site Directed Mutagengesis Kit and contained 5 µL 10X buffer, 1 µL dNTP, 3 µL QuikSolution, 5 µL pET30b-ompA vector, 8 µL generated LSECtin S-CRD coding cDNA, 1 µL Pfu DNA polymerase (L QC Enzyme) and 27 µL water.

The reaction proceeded as follows (22 cycles of steps 2-4 were done):

1. 95°C, 2 min;
2. 95°C, 1 min;
3. 55°C, 1 min;
4. 68°C, 3 min;
5. 68°C, 6 min.

In order to get rid of the methylated template DNA, the PCR product was digested with DpnI by adding 1 µL of the enzyme to the PCR product mix and performing reaction at 37°C for 3 h. TOP10 cells were transformed with the reaction product as described in subsection 6.1.1 and the cells were plated on Petri dish with LB-Agar in presence of 50 µg/mL kanamycin and grown at 37°C overnight.

After propagation of several clones and plasmid purification as described in subsection 6.1.1, the clones were checked by digestion with XbaI and HindIII (reaction mixture of 1 µL XbaI, 2 µL HindIII, 13 µL 10X *Tango* buffer and 14 µL water; reaction for 3 h at 37°C). The positive close was sent for sequencing to COGENICS.

## 6.1.5. *Evaluation of LSECtin and MGL construct functionality*

### 6.1.5.1. *The functionality of pEt20b and pEt30b constructs*

**Over-expression of proteins.** The BL21(DE3) were transformed (subsection 6.1.1) with corresponding tested pEt20b or pEt30b vectors containing LSECtin S-ECD, MGL S-ECD and S-CRD coding cDNA sequences. Overnight precultures were started in LB containing either 0.1 mg/mL ampicillin (in case of pEt20b) or 50 µg/mL kanamycin (in case of pEt30b). 1 mL of each overnight precultures was inoculated to 25 mL of LB with the same concentrations of either ampicillin or kanamycin. After 3 h of cultivation at 37°C (when optical densities at 595 nm ($OD_{595}$) reached approximately 0.8) and after taking 5 mL samples for test, the over-expression of proteins was induced by addition of IPTG to final concentration of 1 mM and the cultures were continued for 3.5 h at 37°C. Two samples of 5 mL at the end of each culture were taken to test by SDS-PAGE.

**Preparation of bacterial culture samples for SDS-PAGE.** For each construct, a 5 mL sample of the culture before the induction and two samples of culture after cultivation were centrifuged at 4000xg for 20 min, 4°C. The supernatants were discarded. To the pellets of the sample before induction and one of the two samples after cultivation, 0.5 mL of 6X SDS-PAGE sample loading buffer (with β-mercaptoethanol) was added. After resuspending, the cells in the two samples were disrupted by sonication for 30 s at 10% amplitude.

To the second sample with the pellet after cultivation, 0.5 mL of buffer consisting of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$. The cells were resuspended and kept on ice while sonicating for the total of 30 s at 40% amplitude with 10 s long breaks between 2 s long sonications. Subsequently, the samples were centrifuged for 10 min at 20800xg, 4°C. Both the pellet and the supernatant were prepared for the analysis by SDS-PAGE: 0.5 mL of SDS-PAGE sample loading buffer was added on the pellet, which was then resuspended and heated for 5 min, and equal volumes of supernatant and SDS-PAGE sample buffer were mixed and heated for 5 min. Typically 5 µL of samples were loaded to gel, except the supernatant sample (10 µL).

SDS-PAGE was performed according to Laemmli protocol [159] using 15% separating and 5% stacking layers of the gel. The PageRuler Unstained Protein Ladder was used as MW marker. The electrophoresis was run at 220 V and 130 mA, typically for 45-50 min.

### 6.1.5.2. *The functionality of pEt30b-ompA/LSECtin S-CRD construct*

The BL21(DE3) cells were transformed with tested pEt30b-ompA/ LSECtin S-CRD vector. Overnight preculture was started in LB containing 50 µg/mL kanamycin. 1 mL of overnight preculture was inoculated to 20 mL of LB with 50 µg/mL kanamycin and the culture carried out at 25°C. Protein expression was induced, after taking a 5 mL sample for test, by addition of IPTG to final concentration of 1 mM when $OD_{600}$ was 0.6, and $CaCl_2$ was also added to final concentration of 100 mM. The culture was continued at 25°C overnight. Two samples of 5 mL of culture were taken for SDS-PAGE and western blot (WB) analysis.

The samples for SDS-PAGE and WB were prepared and SDS-PAGE was performed as described above. The WB procedure was carried out as follows:

1. Transfer of the proteins from non-stained SDS-PAGE gel to PVDF membrane at 300 mA, 120 V, 150 W for 50 min using Tris-glycine buffer (25 mM Tris, 0.19 M glycine).
2. Membrane blocking: 1 h incubation in 5% skim milk solution prepared in PBS-Tween buffer (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4$, 1.4 mM $KH_2PO_4$, 0.05% Tween 20).
3. Incubation with mouse anti-StrepTagII antibody for 1 h: blocking solution replaced by 10 mL PBS-Tween buffer with 5 µL of antibody stock solution (0.2 mg/mL).
4. Rinsing: 3x (for 10 min each time) incubation with PBS-Tween buffer, each time buffer changed.

5. Incubation with anti-mouse-HRP conjugate (2 µL of 0.1 mg/mL stock in 10 mL of PBS-Tween buffer) for 1 h.
6. Rinsing, same as step 4.
7. Staining: a tablet of the stainer (3,3'-Diaminobenzidine) and of urea were dissolved in 1 mL of water, then the membrane repeatedly covered with the prepared stainer.

The PageRuler Prestained Protein Ladder was used as the size marker.

### 6.1.6. *Over-expression and purification of MGL S-ECD*

**Over-expression.** The *E.coli* strain BL21(DE3) was used for over-expression of DC-SIGN S-ECD. The cells were transformed with pEt30b/MGL S-ECD plasmid. The culture for over-expression of MGL S-ECD was started by inoculation of 25 mL of overnight preculture to 0.5 L of LB broth with 50 µg/mL kanamycin. The cells were grown at 37°C and the protein over-expression was induced by addition of IPTG to final concentration of 1 mM when $OD_{600}$ was about 2.0. The culture was continued at 37°C overnight. The cells were harvested by centrifugation at 5000xg for 30 min at 4°C, rinsed with 25 mM Tris-HCl pH 8 buffer by resuspension followed by centrifugation.

**Inclusion body extraction.** The pellet of the cells of 0.5 L culture was resuspended in 30 mL of the buffer consisting of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ supplemented with 1 tablet of “complete EDTA-free” protease inhibitor cocktail, and all other procedures were the same as described for DC-SIGN ECD in sub-subsection 6.1.2.1.

**Refolding.** Refolding of solubilized MGL S-ECD (1.3 mg/mL, 150 mL in Gdn-HCl) proceeded in the same manner as described (sub-subsection 6.1.2.1).

**Purification.** The purification of refolded MGL S-ECD was performed by two steps of affinity chromatography. All purifications were carried out at 4°C.

Firstly, the refolded MGL S-ECD solution (860 mL, 0.106 mg/mL) was loaded to 5 mL Strep-Tactin column prior equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer. After washing the column with the same buffer (100 mL), a step elution of the bound protein was done using the same buffer supplemented with 2.5 mM desthiobiotin (70 mL). The flow rate of 2.5 mL/min was maintained during all procedure.

The isolation of functional protein was carried out using agarose-based medium containing Superose12 column. Because agarose is galactose-based polymer and MGL binds galactose, this medium could be used for affinity purification. The part of the protein eluted from StrepTactin column was concentrated and injected (1 mL, 3 mg/mL) to 125 mL Superose 12 column prior equilibrated in 25 mM MES pH 7, 150 mM NaCl, 4 mM $CaCl_2$ buffer, and after 170 mL flow of the buffer a step elution with 25 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA was done. The flow rate of 1 mL/min was maintained throughout the procedure. All eluted protein species were concentrated.

Finally, to assess the functionality of purified MGL S-ECD, concentrated sample after Strep-Tactin was injected to 5 mL “immobilized D-galactose gel” column prior equilibrated in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer. After 25 mL of buffer run the step elution was performed with 25 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA buffer. The flow rate of 2.5 mL/min was maintained during all procedure.

## 6.2. Biochemical/biophysical protein characterization and glycomimetic compound analysis assays

### 6.2.1. *SDS-PAGE of protein samples*

Protein samples for SDS-PAGE analysis were prepared by mixing equal volumes of protein solution with SDS-PAGE sample loading buffer with (reducing) or without (non-reducing) β-mercaptoethanol, and the samples were heated for 3-5 min. Typically 10 µL of the samples were loaded to the gel. The rest of the procedure was performed as already described (sub-subsection 6.1.5.1).

### 6.2.2. *Protein concentration determination*

The concentrations of protein solutions were determined measuring absorbance at 280 nm and using extinction coefficients calculated from amino acid sequence in ExPASy server (ProtParam tool). NanoDrop 2000c spectrophotometer (from ThermoScientific) was used for protein solution absorbance measurements.

### 6.2.3. *Carbohydrate and glycomimetic compound preparation for SPR assays*

Monovalent glycomimetic compounds were dissolved in SPR running buffer to final concentration of 5 mM or 20 mM. If the compounds had lower water solubility, they were attempted to be dissolved by adding DMSO up to 4% of final concentration (in these cases, the DMSO concentration in the running buffer was adjusted accordingly). All of the prepared compound stock solutions were centrifuged at 13 krpm for 10 min in order to separate the insoluble particles and prevent injecting them to Biacore instrument.

The multivalent compounds were dissolved in running buffer to final concentration of 1 mM or 5 mM, and treated in same way as monovalent compounds in case of low water solubility. Typically, 100 mM stock solutions in water were prepared of natural carbohydrates.

### 6.2.4. *The SPR competition assays*

#### 6.2.4.1. *Single point inhibition assay*

For fast screening of compound activity and selectivity the single point inhibition assay was used. The sensor chip CM4 surface was functionalized with α1-3(α1-6)-mannotriose BSA (ManBSA) by amine coupling procedure. Carboxymethyl (CM) dextran matrix was activated by 50 µL injection of 0.2 M EDC / 0.05 M NHS mixture followed by the injection of typically 25 µL of 60 µg/mL ManBSA in 10 mM sodium acetate buffer pH 4. The remaining unbound carboxyl groups were blocked by the injection of 30 µL of 1 M ethanolamine pH 8. Finally, the surface was rinsed with 10 mM HCl injection (10 µL) and conditioned with the injection of the regeneration solution (20 µL of 50 mM EDTA pH 8). The flow cell with EDC/NHS activated and ethanolamine deactivated CM-dextran was used as the reference surface. The HBS-P buffer was used as the running buffer and the flow rate of 5 µL/min was maintained.

The mixtures of the lectin, DC-SIGN ECD or langerin ECD, at 20 µM and mannose-based (150 µM) or fucose-based (300 µM) compounds were prepared in the running buffer composed of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and 0.005% surfactant P20. The incubated mixtures (20 µL) were injected over ManBSA surface, which was subsequently regenerated with 1 min injection of 50 mM EDTA pH 8. The flow rate of 5 µL/min was maintained. After subtraction of the reference surface response, the lectin steady state binding responses were extracted from the sensorgrams using BiaEval version 4.1 and compared with the binding responses of compound-free lectin injections and converted to inhibition percent values.

#### 6.2.4.2. *Competition assay for $IC_{50}$ determination*

A range of glycomimetic compound concentrations were prepared in the presence of DC-SIGN ECD (20 µM) or langerin ECD (15 µM) and 20 µL or 13 µL of the incubated samples were injected. Compound-free lectin injections at the same concentration were used as full activity reference and also to monitor ManBSA surface activity. The same running buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and 0.005% P20) was used, and in case of compound solubility problems it was supplemented with 3-4% DMSO. The flow rate of 5 µL/min was maintained and all other conditions were the same as in the single point inhibition assay.

The resulting binding responses were converted to residual lectin activity values, which were plotted against concentration values of the compounds. The $IC_{50}$ values for each compound were determined by fitting four-parameter logistic model (eq. 6.1), and calculated using derivation of equation 6.1 (eq. 6.2):

$$Res.act. = R_{hi} - \frac{R_{hi} - R_{lo}}{1 + \left(\frac{conc.}{A_1}\right)^{A_2}} \tag{6.1}$$

$$IC_{50} = A_1 \left(\frac{R_{hi} - R_{lo}}{R_{hi} - 50} - 1\right)^{\frac{1}{A_2}} \tag{6.2}$$

where *Res.act.* is the residual lectin activity, *conc.* is the concentration of the compounds, $R_{hi}$ and $R_{lo}$ are the lowest and the highest values of percent activity, $A_1$ and $A_2$ are the inflection point and the slope of the linear region of the curve, respectively. In equation 6.2, *50* stands for 50% activity of the lectin.

### 6.2.5. *The SPR direct interaction assay*

#### 6.2.5.1. *Preparation of oriented DC-SIGN surface*

Sensor chips CM5 or CM3 were used. Firstly, CM-dextran was functionalized with StrepTactin by amino coupling. Surface was activated with 50 µL injection of 0.2 M EDC / 0.05 M NHS mixture, then StrepTactin (150 µL injection, typically at 150 µg/mL) was covalently coupled to the activated surface and remaining carboxyl groups were blocked by 50 µL 1 M ethanolamine pH 8 injection. HBS-P was used as the running buffer at 5 µL/min flow rate. The prepared surface was rinsed with 10 mM HCl (100 µL at 100 µL/min) and 50 mM NaOH/1 M NaCl (100 µL at 100 µL/min) solutions.

The second step involved the biospecific capture of DC-SIGN S-ECD. The lectin was injected (typically 150 µL) either with or without prior EDC/NHS reactivation of dextran-StrepTactin surface. in case DC-SIGN S-ECD was injected over non-reactivated dextran/StrepTactin surface, it was prepared in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer, otherwise it was prepared in HBS-P (Biacore) buffer. The prepared surface was rinsed with 5 µL of 50 mM EDTA pH 8 injection. 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ or HBS-P were used as running buffers at 5 µL/min flow rate.

The CM-dextran with covalently immobilized StrepTactin (same preparation protocol as described above) was used as a reference surface.

### 6.2.5.2. *Establishment of regeneration conditions*

Injections of 5 µL of 0.5 µM ManBSA prepared in running buffer were used to monitor lectin surface regeneration and activity. The following solutions were screened for surface regeneration capacity:

1. 50 mM EDTA pH 8
2. 5 mM NaOH/0.1 M NaCl
3. 50 mM NaOH/1 M NaCl
4. 20% v/v ethylen glycol in 40 mM EDTA pH 8
5. 50% v/v ethylen glycol in 25 mM EDTA pH 8
6. 0.1 M Gly-NaOH pH 11.9/0.3% TritonX100
7. 10 mM Gly-NaOH pH 11.9+0.03% TritonX100
8. 50 mM Gly-NaOH pH 11.9 / 0.15% TritonX100 / 25 mM EDTA pH 8

The variable volumes of regeneration candidates were injected at 100 µL/min flow rate, except for EDTA (5 µL/min). The flow rate of 5 µL/min was maintained of the running buffer composed of 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and 0.005% P20.

### 6.2.5.3. *Compound interaction with oriented DC-SIGN surface*

The concentration ranges (typically from submicromolar to 1 mM or 2 mM) of polyvalent compounds were prepared in the running buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ and 0.005% P20) supplemented with 3-4% DMSO in case of compound solubility problems. 15-50 µL of prepared compound solutions were injected over the sensor chip surfaces at 5 µL/min flow rate. DC-SIGN surface activity was monitored using ManBSA injections in the beginning and the end of each run.

The sensorgrams were reference surface (CM-dextran with immobilized StrepTactin) corrected and analyzed in BiaEval 4.1.

## 6.2.6. *Other biophysical assays*

The experimental part of AUC, ITC and DLS experiments is described in the further presented paper manuscript **n°5** (see table 7.1 for the list of articles). The NMR experimental procedures are described in article **n°6**.

# Part III.

# Results

# 7. The explanation of the organization of the results in this thesis.

The work of this thesis consists of two different parts presented in two separate chapters as shown in figure below:

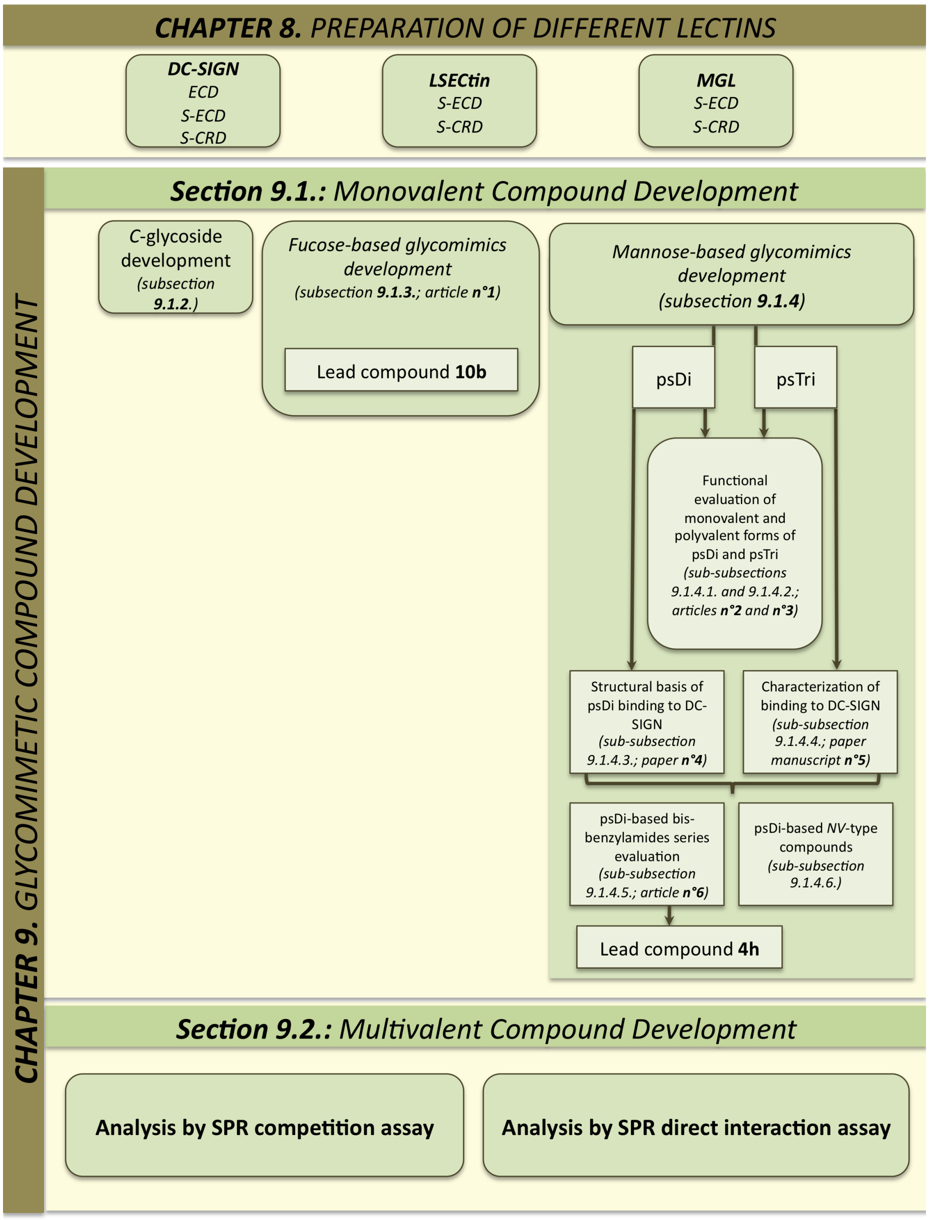

CHAPTER 8. PREPARATION OF DIFFERENT LECTINS
DC-SIGN
ECD
S-ECD
S-CRD
LSECtin
S-ECD
S-CRD
MGL
S-ECD
S-CRD
CHAPTER 9. GLYCOMIMETIC COMPOUND DEVELOPMENT
Section 9.1.: Monovalent Compound Development
C-glycoside development (subsection 9.1.2.)
Fucose-based glycomimics development (subsection 9.1.3.; article n°1)
Lead compound 10b
Mannose-based glycomimics development (subsection 9.1.4)
psDi
psTri
Functional evaluation of monovalent and polyvalent forms of psDi and psTri (sub-subsections 9.1.4.1. and 9.1.4.2.; articles n°2 and n°3)
Structural basis of psDi binding to DC-SIGN (sub-subsection 9.1.4.3.; paper n°4)
Characterization of binding to DC-SIGN (sub-subsection 9.1.4.4.; paper manuscript n°5)
psDi-based bis-benzylamides series evaluation (sub-subsection 9.1.4.5.; article n°6)
psDi-based NV-type compounds (sub-subsection 9.1.4.6.)
Lead compound 4h
Section 9.2.: Multivalent Compound Development
Analysis by SPR competition assay
Analysis by SPR direct interaction assay

**Figure 7.1.:** A schematic representation of the organization of the results in this thesis.

The first part includes the results of the **preparation of different lectins**. To begin with, the production of different ***DC-SIGN*** constructs, which **were** used for glycomimetic compound structural and functional evaluation, will be presented in section 8.1 on page 123. Apart from DC-SIGN, different constructs of two other C-type lectins, namely ***LSECtin*** and ***MGL***, were designed in order to expand the compound selectivity screening tools. The results of cloning and construct expression evaluation will be presented in section 8.2 on page 133.

The second chapter describes the **development of glycomimetic compounds**, and is also divided to two major subparts, one of which describes development of ***monovalent glycomimics***, and in the second one the results of the development of ***multivalent compounds*** are presented.

The monovalent compound development was conducted in three parallel streams as shown in figure 7.1. Pr. Jitka Moravcová team was developing the *C-glycosides*, and the results of their evaluation are presented in subsection 9.1.2 on page 153. Pr. Anna Bernardi's team worked on the two concepts, namely *fucose-based* and *mannose-based* glycomimetics, and most of the results of their development are presented as papers (table 7.1).

**Table 7.1.:** The list of papers presented in this thesis.

| ***N<sup>O</sup>*** | ***Status*** | ***Title*** | ***Location*** |
|---|---|---|---|
| **n°1** | Published article | Andreini et al. Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity versus Langerin. *Org. Biomol. Chem.* **9** (2011) 5778-5786 | Subsection 9.1.3 on page 157 and appendix on page 161 |
| **n°2** | Published article | Berzi et al. A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models. *AIDS* **26** (2012) 127-137 | Sub-subsection 9.1.4.1 on page 174 and appendix on page 177 |
| **n°3** | Published article | Luczkowiak et al. Pseudosaccharide Functionalized Dendrimers as Potent Inhibitors of DC-SIGN Dependent Ebola Pseudotyped Viral Infection. *Bioconjugate chemistry* **22** (2011) 1354-1365 | Sub-subsection 9.1.4.2 on page 191 and appendix on page 195 |
| **n°4** | Submitteed paper | Thépaut et al. Structure of a glycomimetic ligand in the Carbohydrate Recognition Domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design. Submitted to *JACS* | Sub-subsection 9.1.4.3 on page 209 and appendix on page 213 |
| **n°5** | Paper manuscript | Sutkevičiūtė et al. Lectin clustering by a glycomimetic ligand without any multivalent presentation a case study in DC-SIGN antagonist development. | Sub-subsection 9.1.4.4 on page 241 and appendix on page 245 |
| **n°6** | Article in press | Varga et al. Selective targeting of DC-SIGN with mannose-based glycomimetics. Synthesis and interaction studies of bis-benzylamide derivatives of a pseudomannobioside. *Chem. Eur. J.* | Sub-subsection 9.1.4.5 on page 265 and appendix on page 271 |

While the results of the development of fucose-based glycomimetics are presented in article **n°1**, the third stream of monovalent compound development, i.e. mannose-based glycomimics, was more elaborated. It begins with functional characterization of two mannose-based glycomimetic compounds, pseudomannobioside (psDi) and pseudomannotrioside (psTri) [49]. From these two compounds, psTri turned out to have much better inhibitory efficiency than psDi, hence it was used to prepare a tetravalent dendron and tested for capacity to inhibit HIV *trans* infection of T lymphocytes [49]. Later on, the potency of this compound to inhibit HIV transmission in cervical explants was also studied and the results are presented in the article **n°2**.

The 32-valent forms of psDi and psTri were also generated and tested comparing the activities with the mono- and tetravalent psDi and psTri compounds in SPR competition and Ebola virus infection inhibition assays. The results are presented in article **n°3**. The latter studies of the compounds indicated discrepancies regarding psTri structural and functional characteristics, i.e. monovalent forms of psDi and psTri had an order of magnitude difference in their activities in favor for psTri, but this difference was lost comparing multivalent compounds. Therefore, the interaction of both monovalent compounds was characterized in detail: the X-ray structures of both compounds in complex with DC-SIGN CRD were solved, structural properties of the interaction in solution was studied by NMR, and the complexes were characterized by other biophysical techniques (ITC, AUC, DLS). The results of these studies are presented in two papers, i.e. submitted article **n°4** and paper manuscript **n°5**.

Finally, these studies disapproved psTri as the best lead compound, but revealed psDi to possess all the desirable qualities of a compound for further development. Thus the development of mannose-based glycomimics continued with a focus on psDi optimization. Two different types of psDi-based compound series were synthesized and evaluated for inhibition and selectivity to DC-SIGN versus langerin. The results of the evaluation of one of the two series (bis-benzylamides) are presented in paper **n°6**, and the analysis of the other series ($NV$-type) is described in subsubsection 9.1.4.6 on page 290.

Simultaneously to the development of monovalent glycomimics, various polyvalent scaffolds were designed and synthesized by the team of Dr. Javier Rojo. Then the chosen monovalent glycomimics were tethered to these polyvalent scaffolds and tested as DC-SIGN antagonists. The results of corresponding polyvalent compound evaluation by SPR competition and direct interaction assays are presented in subsections 9.2.2 and 9.2.3, respectively.

# 8. Results of preparation of recombinant lectins

## 8.1. Preparation of different DC-SIGN constructs

Three different constructs of DC-SIGN were used in this work (fig. 8.1):

- DC-SIGN ECD (a.a. 66-404);
- DC-SIGN S-CRD (a.a. 250-404 and includes N-terminal StrepTag: MASWSHPQFEKIEGR, where green, red and blue parts correspond to linker, StrepTag II and Factor Xa cleavage site sequences, respectively);
- DC-SIGN S-ECD (a.a. 66-404 and includes N-terminal StrepTag).

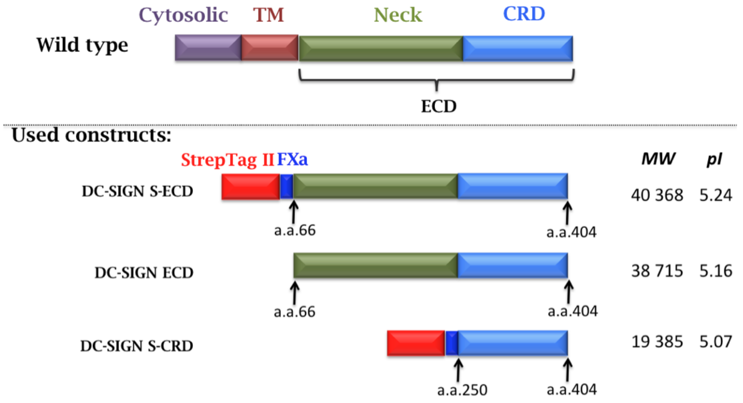


**Figure 8.1.:** The DC-SIGN constructs used in this thesis.
TM stands for transmembrane domain, a.a. – amino acids, MW – molecular weight (in Da), pI – isoelectric point. The shown MW and pI values are theoretical, calculated from corresponding amino acid sequences with ProtParam tool in ExPASy server.

DC-SIGN ECD in this thesis is used mainly in SPR competition assay, and also was used in ITC, AUC and DLS experiments. The CRD construct of DC-SIGN is used only for co-crystallization with selected compounds. Because monomeric CRD doesn't have high affinity to sugars, the StrepTag was included in order to facilitate the purification of the protein by the means of the affinity chromatography using the StrepTactin column. There were different reasons to add StrepTag to the ECD construct of DC-SIGN: this tag served for the preparation of oriented DC-SIGN surfaces for SPR direct interaction assay, where the lectin construct is captured by covalently immobilized StrepTactin; StrepTag also aided the purification of this construct.

All of the three constructs were designed and their production was developed before I started my PhD.

### 8.1.1. *DC-SIGN ECD over-expression and purification results*

DC-SIGN ECD (38.7 kDa) pEt30b construct was expressed in BL21(DE3) *E.coli* strain (fig. 8.2) and purified from 1 L of culture.

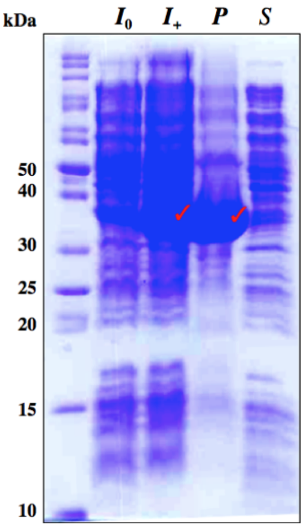


**Figure 8.2.:** The SDS-PAGE analysis of DC-SIGN ECD expression.
Samples before ($I_0$) and after induction containing whole disrupted cells ($I_+$), supernatant ($S$) and pellet ($P$) were loaded. Red ticks mark the target protein bands.

The inclusion bodies were isolated by several steps of selective centrifugations, then solubilized in 6 M Gdn-HCl buffer and refolded by flash dilution and subsequent dialyses. The functional protein was recovered by an affinity chromatography step using Mannan-Agarose column followed by gel-filtration on Superose12 column to eliminate EDTA and recharge the lectin with $Ca^{2+}$ ions. The examples of the results of DC-SIGN ECD purification on Mannan-Agarose and Superose12 columns are shown in figure 8.3.

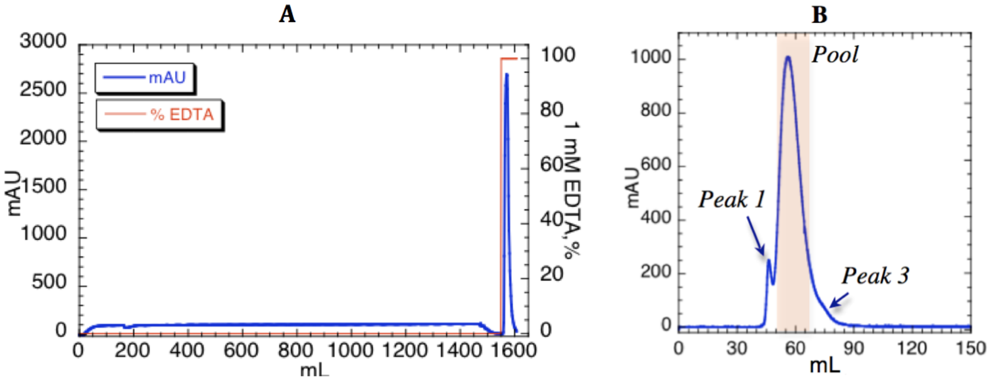


**Figure 8.3.:** The chromatograms of DC-SIGN ECD purification.
**A**, Loading and elution from Mannan-Agarose column; **B**, Gel-filtration on Superose12 column with different forms of protein marked: 'peak 1' - possibly higher size oligomers, 'pool' highlighted in orange – the collected well-folded protein, and the 'peak 3' – possibly lower size protein form.

The typical yield of the protein from 1 L of bacterial cell culture evolved as follows: 237 mg of protein was loaded to Mannan-Agarose column and 78.4 mg were eluted from the column; after all injections to Superose12 column and concentration of pooled fractions the total amount of protein was 52.4 mg. All steps of the purification were monitored by SDS-PAGE, and the results of reduced and non-reduced DC-SIGN ECD samples from different purification stages are shown in figure 8.4.

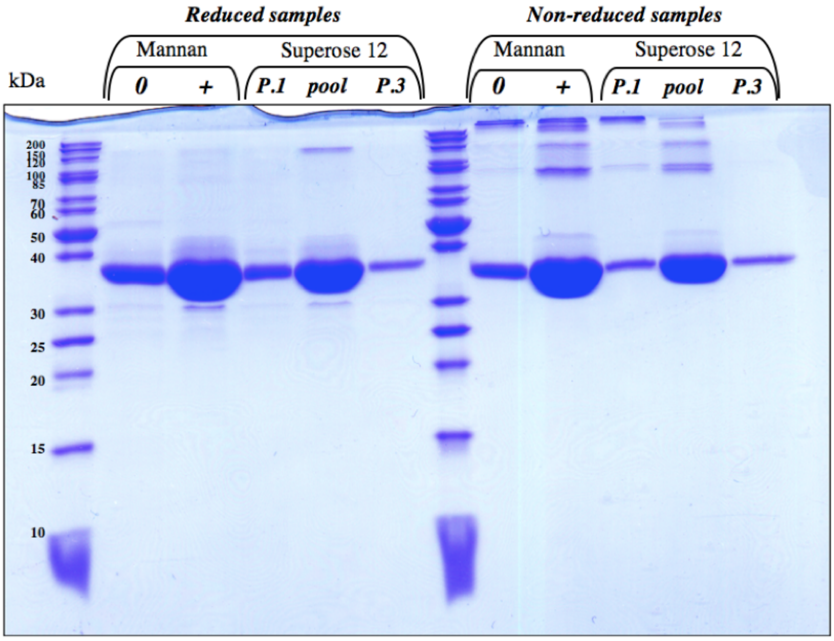


**Figure 8.4.:** The SDS-PAGE analysis of DC-SIGN ECD purification.
Reduced or non-reduced samples after refolding (i.e. before Mannan-Agarose, marked Mannan '0'), concentrated samples of protein eluted from Mannan-Agarose column (marked Mannan '+') and different samples from Superose12 as shown in fig. 8.3B (peak 1 marked 'P.1', peak 3 marked 'P.3') were loaded to the gel.

DC-SIGN ECD forms disulphide-linked oligomers, which comprise nearly a half of all DC-SIGN ECD sample after refolding, and the following steps of purification gradually remove the major part of those high oligomers. It is possible that the size-exclusion column is not capable to separate these disulphide-linked oligomers because well-folded protein is a tetramer (155 kDa), thus both disulphide-linked oligomers and well-folded DC-SIGN ECD are eluted at the same volume. The 'peak 1', which is excluded from the major pool of DC-SIGN ECD, contains disulphide-linked oligomers of high size, while 'peak 3' doesn't have disulphide-linked oligomers at all, and the lower size of this DC-SIGN ECD form could result from folding into dimers or trimers but not tetramers. It is also noticeable that the sample also contains smaller fragments, which are linked to correct size or higher oligomers of DC-SIGN ECD by disulphide bridges, since the lower MW bands in SDS-PAGE that are visible in reduced sample disappear at non-reducing conditions. Nevertheless, the vast majority of the functional protein has the correct size as seen in SDS-PAGE at both reducing and non-reducing conditions.

### 8.1.2. *DC-SIGN S-ECD over-expression and purification results*

DC-SIGN S-ECD (40.4 kDa) pEt20b construct was expressed in Rosetta(DE3) *E.coli* strain (fig. 8.5), which was previously chosen among several strains as the most optimal in terms of expression capacity, and purified from 2 L of culture due to relatively low expression level.

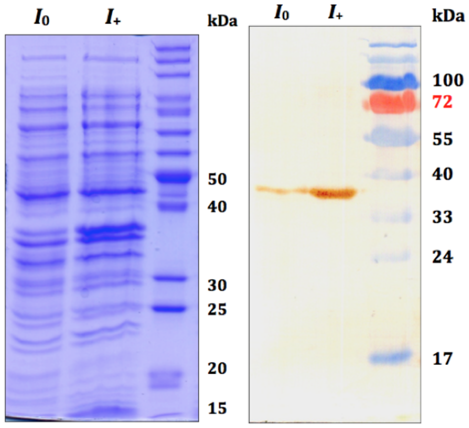


**Figure 8.5.:** The expression of DC-SIGN S-ECD.
The left and right panels show SDS-PAGE and WB analyses, respectively, of DC-SIGN S-ECD expression. Lanes marked $I_0$ and $I_+$ corresponded to loaded samples before and after induction. The WB was revealed using anti-StrepTag II antibody.

DC-SIGN S-ECD was "leaking", i.e. a band corresponding to target protein is seen before induction, due to not as tight expression control as in case of pEt30b.

The inclusion bodies were extracted and refolded in the same manner as for DC-SIGN ECD construct. The functional protein was recovered by two-step affinity purification followed by gel-filtration: at first the protein was captured and concentrated on StrepTactin column, then the functional protein was isolated using Mannan-Agarose column, and finally the EDTA from elution buffer of previous purification step was eliminated and the lectin was recharged with $Ca^{2+}$ ions using Superose12 column. The results of DC-SIGN S-ECD purification on StrepTactin, Mannan-Agarose and Superose12 columns are shown in figure 8.6.

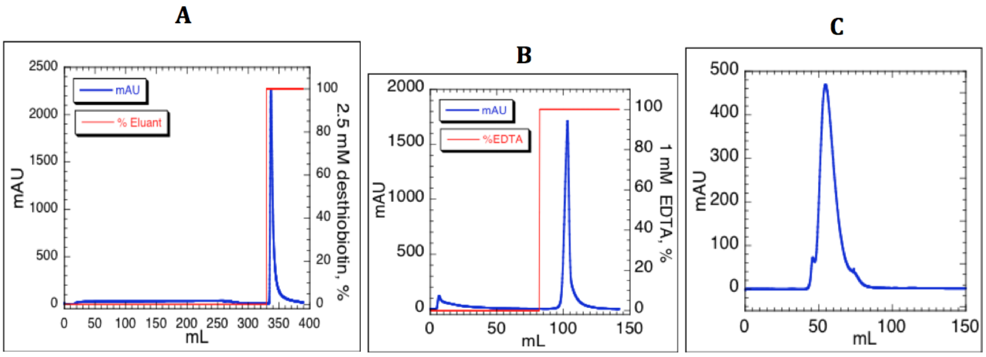


**Figure 8.6.:** The chromatograms of DC-SIGN S-ECD purification.
The chromatograms of DC-SIGN S-ECD purification. **A**, Loading and elution from StrepTactin column. **B**, Injection and elution from Mannan-Agarose column. **C**, Gel-filtration on Superose12 column.

The yield of the protein evolved as follows: 25.6 mg of protein was loaded to StrepTactin and 13.4 mg were eluted from the column → 11.4 mg were injected to Mannan-Agarose column and

7.6 mg were eluted → 7.4 mg were injected to Superose12 and 5.4 mg were eluted. Thus the general yield of DC-SIGN S-ECD is 5.4 mg from 2 L of bacterial culture.

The purification was monitored by SDS-PAGE of reduced (*R*) or non-reduced (*NR*) DC-SIGN S-ECD samples from different purification stages (fig. 8.7).

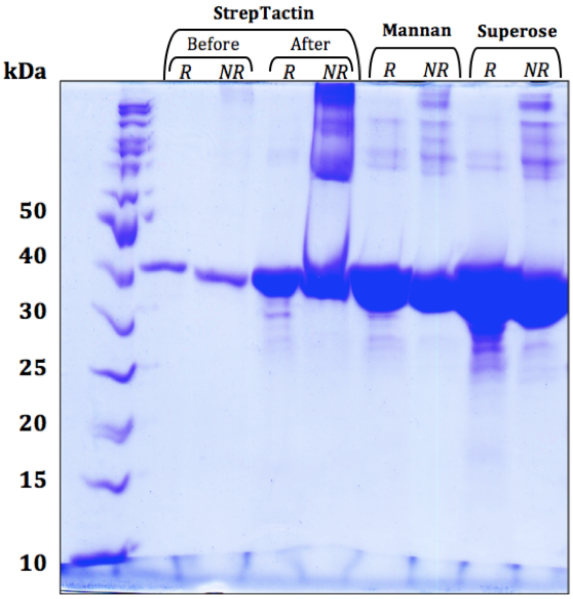


**Figure 8.7.:** The SDS-PAGE analysis of DC-SIGN S-ECD purification.
Reduced (*R*) or non-reduced (*NR*) samples after refolding (i.e. before StrepTactin), concentrated samples of protein eluted from StrepTactin column (i.e. after StrepTactin), from Mannan-Agarose (i.e. Mannan) and from Superose12 (i.e. Superose) were loaded to the gel.

As DC-SIGN ECD, the S-ECD construct of DC-SIGN also forms disulphide-linked oligomers, and the distribution of these oligomers has the same pattern as in case of DC-SIGN ECD, with the vast majority of the functional protein of a correct size.

### 8.1.3. *DC-SIGN S-CRD refolding optimization results*

The S-CRD construct of DC-SIGN is required for co-crystallization with selected glycomimetic compounds. Moreover, this construct could be a valuable tool in SPR assays for the following reasons. Because of their low MW and low affinity (thus high bulk responses when injecting high compound concentrations), the monovalent glycomimetic compounds cannot be assessed for binding affinity to the surface functionalized with DC-SIGN. The soluble tetrameric DC-SIGN ECD, which is used in SPR competition assay, does not allow to determine solution affinity of the compounds because of the presence of four binding sites within one protein entity that binds to the surface. But CRD has only one binding site meaning that this construct could be successfully used in SPR solution affinity assay.

The over-expression of DC-SIGN S-CRD is very efficient (fig. 8.9) giving a promising potential to obtain high yields of purified protein. However, all of the protein is expressed as inclusion bodies and must be refolded. Although the protocol of DC-SIGN S-CRD purification was set-up and used before my PhD, it gives low yields of protein mainly because of extensive precipitation during the refolding step. Additionally, a high proportion of the protein was found to exists as functional disulfide-linked dimers (peak 4 (*pk.4*) in fig. 8.8), therefore a double mutant construct where the first and the last cysteine residues (responsible for C0-C0´ disulfide bridge formation at the base of CRD, fig. 1.6 on page 51) were changed to serines (see figure 2.7 on page 66). Unfortunately, this construct did not yield any improvements in DC-SIGN S-CRD purification.

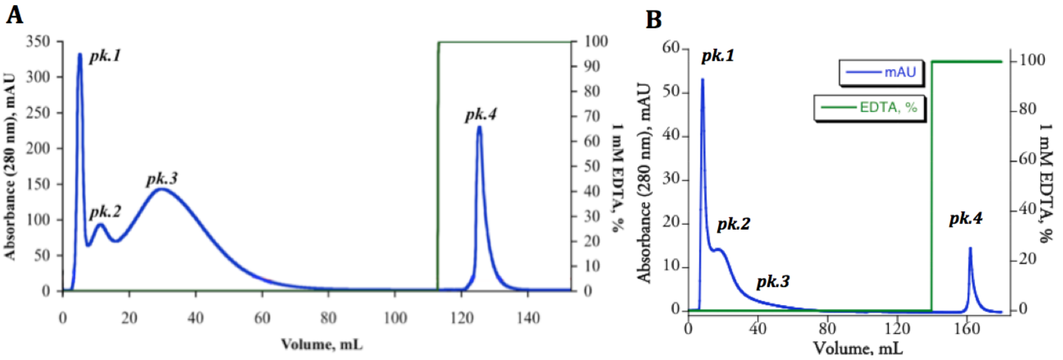


**Figure 8.8.:** The chromatograms showing elution profiles of DC-SIGN S-CRD purification on Mannan-Agarose column.

**A**, The best-case scenario of DC-SIGN S-CRD sample preparation (the result from previous experiments before my PhD): the first peak (*pk.1*) was assumed to be desthiobiotin coming from previous purification on StrepTactin column mixed with non-functional/contaminant proteins; *pk.2* and *pk.3* (a delayed peak) were assumed to be monomeric misfolded and functional DC-SIGN S-CRD, respectively; and *pk.4* was found to be a disulfide-linked DC-SIGN S-CRD dimers that were functional since EDTA was required to detach mannan-bound CRDs. **B**, A typical elution profile on Mannan-Agarose column of DC-SIGN S-CRD (injection of non-concentrated sample prior purified using StrepTactin column).

Hence, the refolding step was targeted for optimization in order to improve the yield of the functional monomeric DC-SIGN S-CRD.

This lectin (19.4 kDa) was expressed in BL21(DE3) *E.coli* strain (fig. 8.9), and inclusion bodies were extracted from 0.5 L of culture in the same manner as for DC-SIGN ECD and S-ECD. After solubilization in Gdn-HCl buffer, the estimated amount of the protein was 95 mg.

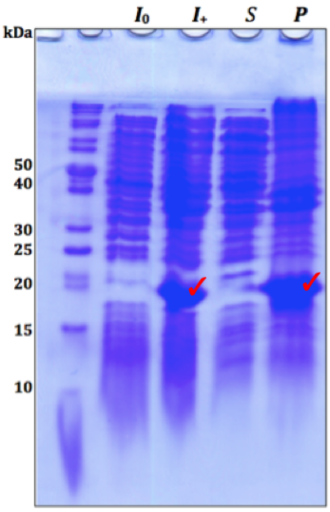


**Figure 8.9.:** The SDS-PAGE analysis of DC-SIGN S-CRD production.
Samples before ($I_0$ ) and after induction containing whole disrupted cells ($I_+$), supernatant ($S$) and pellet ($P$) were loaded. Red ticks mark the target protein bands.

Then the refolding of solubilized DC-SIGN S-CRD inclusion bodies was tested in 4 different buffers:

1. "Usual buffer" consisting of 1.25 M NaCl, 25 mM Tris-HCl pH 8, 25 mM $CaCl_2$ (*sample 1*);
2. Usual buffer supplemented with glutathiones (1 mM oxidized / 10 mM reduced) (*sample 2*);
3. Usual buffer supplemented with 0.16 M L-arginine (*sample 3*);
4. Usual buffer supplemented with glutathiones (as 2.) and 0.16 M L-arginine (*sample 4*).

The absorbance spectra of each sample and a control (same concentration of protein in inclusion body solubilization buffer) were recorded and compared (fig. 8.10).

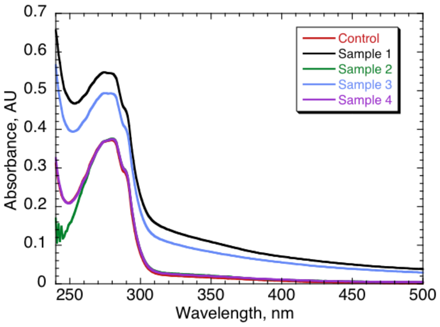


**Figure 8.10.:** The absorbance spectra of DC-SIGN S-CRD refolding samples.

As can be seen from figure 8.10, samples 2 and 4 superimposed with the control meaning no precipitation. Both of the buffers of these two samples contained glutathiones, which suggests that precipitation in other samples might have occurred due to the formation of disulphide-linked aggregates. Buffer 4, which contains both glutathiones and arginine that helps to minimize aggregate formation through hydrophobic patches, was selected for refolding of the whole batch of solubilized inclusion bodies. No precipitation was observed even after the dialyses, however, after centrifuging a gel-like pellet formed. SDS-PAGE analysis (fig. 8.11, *gel 1*) of non-reduced sample of the supernatant showed the presence of big disulphide-linked aggregates that couldn't enter the gel (a band at the bottom of the well). The analysis of this pellet (fig. 8.11, *gel 2*) showed that it contained a considerable amount of DC-SIGN S-CRD.

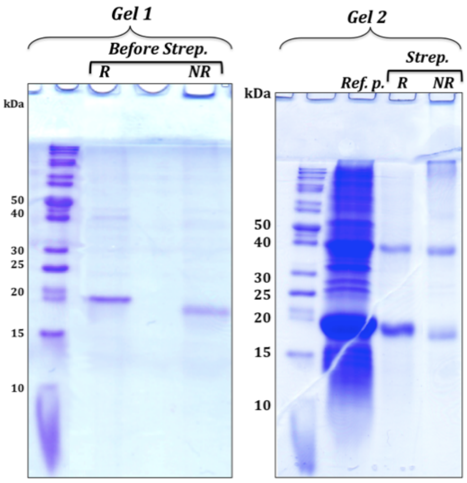


**Figure 8.11.:** The SDS-PAGE analysis of different purification stages of DC-SIGN S-CRD. *Gel 1* shows reduced (*R*) and non-reduced (*NR*) DC-SIGN S-CRD samples after refolding (the supernatant before loading to StrepTactin, *Before Strep.*). *Gel 2* shows reduced sample of gel-like pellet after DC-SIGN S-CRD refolding (*Ref.p.*) and samples of elution peak from StrepTactin (*Strep.*).

The refolded DC-SIGN S-CRD was loaded to StrepTactin column in the continued presence of glutathiones (fig. 8.12A), which were added in order to decrease the formation of disulfide-linked oligomers. The flow-through was reloaded to StrepTactin two more times (results not shown) to increase the yield of the purified protein. In total, 15.2 mg of DC-SIGN S-CRD was recovered in elution peaks from StrepTactin column.

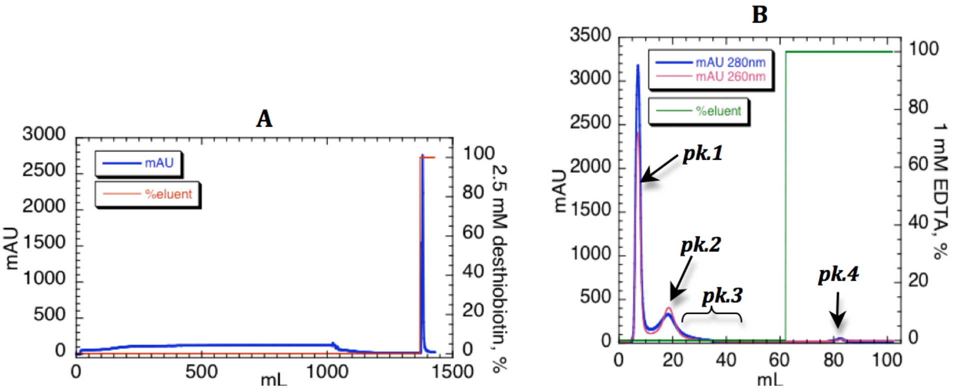


**Figure 8.12.:** The chromatograms of DC-SIGN S-CRD purification. **A**, An example of loading and elution from StrepTactin column. **B**, An example of elution profile of DC-SIGN S-CRD (injection of 1.8 mL at 1.8 mg/mL) on Mannan-Agarose column with absorbance changes monitored at 280 nm (blue line) and 260 nm (red line).

The SDS-PAGE analysis of StrepTactin elution peak showed that the sample was composed of nearly equal amounts of monomeric and disulfide-linked dimeric DC-SIGN S-CRD, but also contained high oligomers as a band was visible at the interface of stacking and running layers of the gel (fig. 8.11, *gel 2*). Comparing DC-SIGN S-CRD samples before loading to StrepTactin, when the majority of the protein was monomeric (fig. 8.11, *gel 1*), and the elution peak, it becomes clear that the concentrating of the protein influenced by binding to the column catalyzed the formation of aberrant disulfides even in the presence of glutathiones. Nevertheless, the addition of glutathiones had a slight improvement of the yield of monomeric protein as can be seen in SDS-PAGE analysis of previous DC-SIGN S-CRD purification on StrepTactin column in the absence of glutathiones, where the dimeric DC-SIGN S-CRD and higher oligomers were dominant (fig. 8.13).

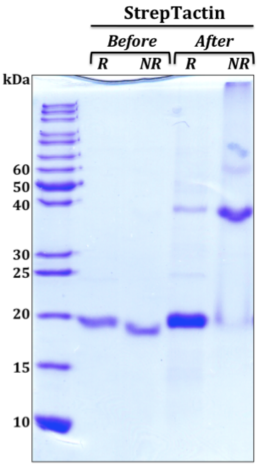

StrepTactin
Before
After
R
NR
R
NR
kDa
60
50
40
30
25
20
15
10

**Figure 8.13.:** The SDS-PAGE analysis of DC-SIGN S-CRD purification on StrepTactin column in the absence of glutathiones.

The lectin was refolded and purified on StrepTactin column under "conventional" conditions, i.e. in the same way as described for DC-SIGN ECD and S-ECD. The reduced (*R*) and non-reduced (*NR*) samples, as indicated, were analyzed.

The purification of DC-SIGN S-CRD (refolded in optimized conditions) continued, and the concentrated elution peak from StrepTactin was injected to Mannan-Agarose column (fig. 8.12B). All the peaks present in the elution profile (fig. 8.12B) were analyzed by SDS-PAGE (fig. 8.14).

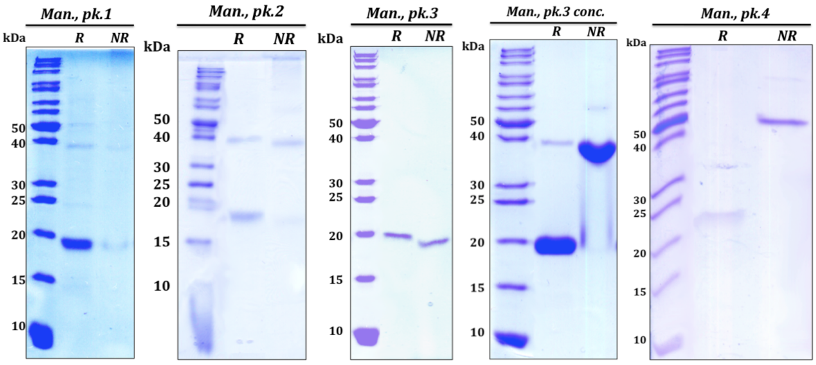

Man., pk.1
kDa
R
NR
50
40
30
25
20
15
10
Man., pk.2
kDa
R
NR
50
40
30
25
20
15
10
Man., pk.3
kDa
R
NR
50
40
30
25
20
15
10
Man., pk.3 conc.
kDa
R
NR
50
40
30
25
20
15
10
Man., pk.4
kDa
R
NR
50
40
30
25
20
15
10

**Figure 8.14.:** The SDS-PAGE analysis of DC-SIGN S-CRD purification on Mannan-Agarose column.

The reduced (*R*) and non-reduced (*NR*) samples of the corresponding peaks from Mannan-Agarose column (*Man., pk.*), as indicated in fig. 8.12B, were analyzed. "*Man., pk.3 conc.*" is a concentrated sample of peak 3 from Mannan-Agarose.

It appeared that the first peak after Mannan-Agarose (*pk.1*) was composed mainly of high oligomers of DC-SIGN S-CRD formed due to aberrant intermolecular disulfide formation: in reducing conditions the major part of the protein is in the band corresponding to the correct size of DC-SIGN S-CRD, but the major part of non-reduced sample stays in the interface of stacking and running layers of the gel (fig. 8.14, *Man., pk1.*).

Initially, the second peak (*pk.2*) and the third broad peak (*pk.3*) were considered to be monomeric misfolded and correctly folded DC-SIGN S-CRD species, respectively. The SDS-PAGE analysis showed that although *pk.2* contained a small amount of monomeric protein, it is majorly composed of dimers and high oligomers as indicated by the band present in the interface of stacking and running layers of the gel (fig. 8.14, *Man., pk.2*). Moreover, in addition to standard 280 nm, the absorbance during the purification on Mannan-Agarose was monitored at 260 nm and appeared to be stronger than at 280 nm, which means that this sample contained residual amounts of nucleic acids (absorbance maxima for nucleic acids is at 260 nm). Thus *pk.2* and *pk.3* are probably contributed by the same species of the protein (later referred as *pk.2/pk.3*), but the presence of nucleic acids influences the emergence of *pk.2* as a separate peak.

The last EDTA elution peak (*pk.4*) contains active protein as deduced from requirement of EDTA to detach the protein from the column. SDS-PAGE analysis of *pk.4* (fig. 8.14, *Man., pk.4*) showed that this sample was composed of disulfide-linked dimers of DC-SIGN S-CRD, which are functional, explaining the increased affinity of the lectin to mannan.

Considering a very broad shape of *pk.2/pk.3*, it is possible that the protein species in *pk.2/pk.3* were slowly “leaking” and did not completely leave the column prior to the switch to elution buffer, which means that the *pk.4* is composed of the same protein species as *pk.2/pk.3.*

The fractions making up *pk.3* were pooled and also analyzed by SDS-PAGE (fig. 8.14, *Man., pk.3*). The analysis revealed that this sample is homogeneous and contains monomeric DC-SIGN S-CRD. The pooled protein was concentrated and again analyzed in reducing and non-reducing conditions (fig. 8.14, *Man., pk.3 conc.*). It appeared that concentrating forced the monomeric DC-SIGN S-CRD to form disulfide-linked dimers leaving only a negligible amount of monomeric protein, a very disappointing outcome of these experiments.

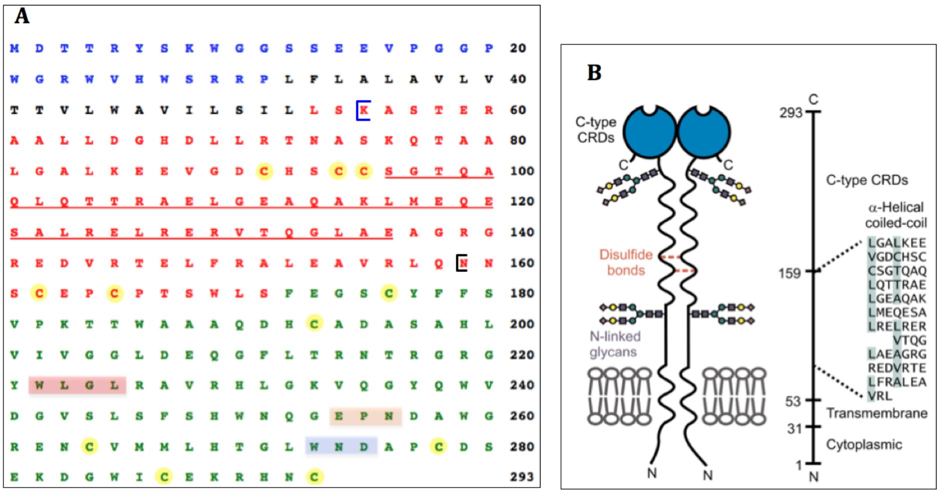
In summary, these studies showed that the presence of multiple cysteine residues in DC-SIGN S-CRD is the limiting factor for the preparation of purified samples of this lectin construct. Even though the refolding can be achieved keeping nearly all protein soluble, the subsequent purification includes concentrating steps and this forced the formation of disulfide-linked DC-SIGN S-CRD oligomers dramatically reducing the quantity of functional monomeric protein.

## 8.2. Generation of recombinant LSECtin and MGL S-ECD and S-CRD constructs

### 8.2.1. *General information about LSECtin and MGL*

In order to expand the tools for glycomimetic compound selectivity screening, the extracellular and carbohydrate recognition domains of LSECtin and MGL were constructed. Both of these type II membrane proteins are also type II C-type lectins as DC-SIGN and langerin.

**LSECtin** LSECtin is encoded in the same gene cluster as DC-SIGN and DC-SIGNR (Liu:2004p28). It also contains an EPN motif in its CRD and it was found to have an overlapping specificity for mannose and GlcNAc containing oligosaccharides [25] with the most effective glycan ligand, a disaccharide GlcNAcβ1–2Man [160]. This lectin presumably exists as a dimer stabilized by two disulphide bonds in its neck region (fig. 8.15) [160]. LSECtin is predominantly expressed on sinusoidal endothelial cells of human liver and lymph node (thereof its name), but it has been also found to be present on macrophages, DCs and Kupffer cells [161, 162], where it functions as an endocytic receptor, although the physiological role of this lectin is not yet fully clear. LSECtin was reported to serve as an attachment factor for enveloped viruses such as Ebola, Marburg and SARS coronavirus [161, 163, 164, 160], and in case of Ebola virus LSECtin was reported to mediate the infection [163]. In contrast to DC-SIGN, LSECtin does not bind lentiviral particles such as HIV or HCV viruses [161].



A

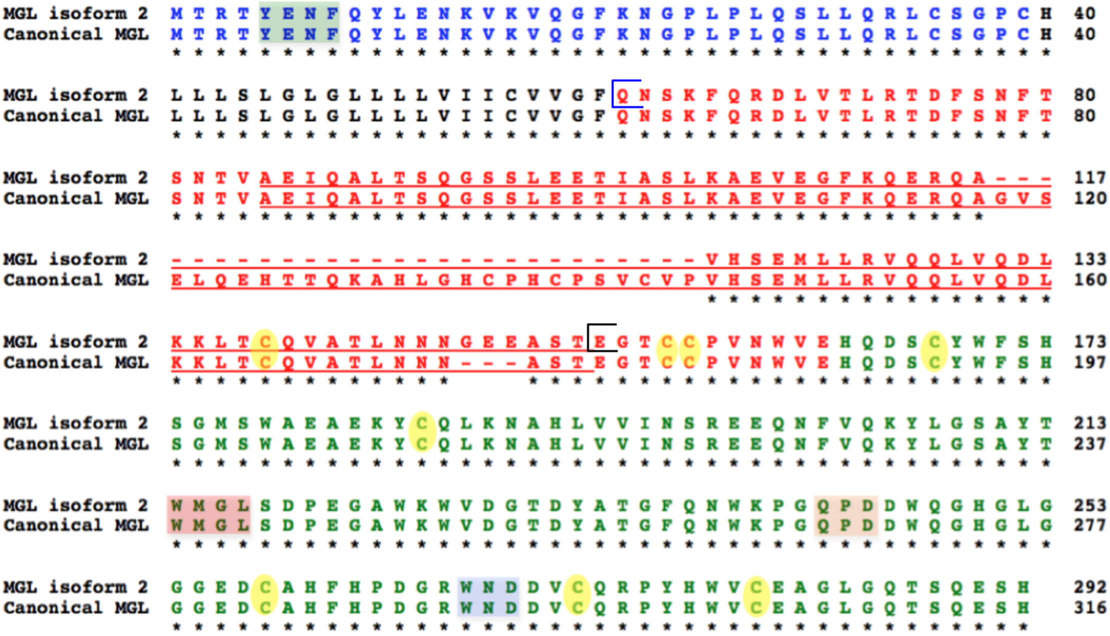
M D T T R Y S K W G G S S E E V P G G P 20
W G R W V H W S R R P L F L A L A V L V 40
T T V L W A V I L S I L L S K A S T E R 60
A A L L D G H D L L R T N A S K Q T A A 80
L G A L K E E V G D C H S C C S G T Q A 100
Q L Q T T R A E L G E A Q A K L M E Q E 120
S A L R E L R E R V T Q G L A E A G R G 140
R E D V R T E L F R A L E A V R L Q N N 160
S C E P C P T S W L S F E G S C Y F F S 180
V P K T T W A A A Q D H C A D A S A H L 200
V I V G G L D E Q G F L T R N T R G R G 220
Y W L G L R A V R H L G K V Q G Y Q W V 240
D G V S L S F S H W N Q G E P N D A W G 260
R E N C V M M L H T G L W N D A P C D S 280
E K D G W I C E K R H N C 293

**Figure 8.15.:** The structure of LSECtin.

**A**, The amino acid sequence of full length LSECtin with cytosolic, transmembrane, neck region, and CRD parts in blue, black, red and green letters (domain organization based on information from UniProtKB); the potential coiled-coil region in the neck domain is underlined (based on information from UniProtKB); the conserved motifs defining sugar specificity and $Ca^{2+}$-coordination are highlighted in orange and blue, and the third conserved motif, WLGL instead of WIGL in case of LSECtin, is highlighted in red; all cysteine residues highlighted in yellow. **B**, The presumed organization of full length LSECtin (from [160]). The beginning of ECD and CRD sequences chosen for cloning are marked by blue and black sticks.

For cloning, amino acids 55-293 were chosen to produce LSECtin ECD construct, and CRD construct comprised amino acids 159-293.

**MGL** Although MGL, a macrophage galactose-type C-type lectin (also called DC-asialoglycoprotein receptor (DC-ASGP-R) or human macrophage lectin (HML)), has a different sugar specificity than DC-SIGN since it contains galactose specificity determining QPD motif in its CRD, it is a general interest to have this lectin as an extended means for compound selectivity screening. Like DC-SIGN, MGL is a type II membrane protein that contains a single C-type CRD, and it is expressed on macrophages and immature myDCs [165], and it is likely a sole galactose-type CLR on DCs [166]. Presumably, MGL exists as trimers although dimeric forms could be identified too [167]. Although initially MGL was reported to recognize galactose and GalNAc sugars [165], recombinantly in bacteria expressed MGL bound only GalNAc [167], and later studies found MGL strongly specific to α- or β-linked GalNAc, including Tn (GalNAcα1-O-Ser/Thr, carcinoma-associated antigen), LDN (GalNAcβ1-4GlcNAc-R) and LDNF (GalNAcβ1-4(Fucα1-3)GlcNAc-R) antigens (although present in humans, LDN and LDNF are abundantly expressed by human helminth parasite Schistosoma mansoni), and not to galactose-containing sugar analogues [168]. Such carbohydrate specificity suggests that MGL may have a role in tolerance to self-gangliosides, recognition of tumor antigens and the eggs of human helminth parasite S. mansoni [168]. MGL has been also found to promote cellular entry of filoviruses like Ebola and Marburg virus [169], and together with macrophage mannose receptor was shown to be implicated in influenza infection of macrophages [170].



```
MGL isoform 2  M T R T Y E N F Q Y L E N K V K V Q G F K N G P L P L Q S L L Q R L C S G P C H  40
Canonical MGL  M T R T Y E N F Q Y L E N K V K V Q G F K N G P L P L Q S L L Q R L C S G P C H  40
               * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

MGL isoform 2  L L L S L G L G L L L L V I I C V V G F Q N S K F Q R D L V T L R T D F S N F T  80
Canonical MGL  L L L S L G L G L L L L V I I C V V G F Q N S K F Q R D L V T L R T D F S N F T  80
               * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

MGL isoform 2  S N T V A E I Q A L T S Q G S S L E E T I A S L K A E V E G F K Q E R Q A - - -  117
Canonical MGL  S N T V A E I Q A L T S Q G S S L E E T I A S L K A E V E G F K Q E R Q A G V S  120
               * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

MGL isoform 2  - - - - - - - - - - - - - - - - - - - - - - - - V H S E M L L R V Q Q L V Q D L  133
Canonical MGL  E L Q E H T T Q K A H L G H C P H C P S V C V P V H S E M L L R V Q Q L V Q D L  160
                                                               * * * * * * * * * * * * * * * *

MGL isoform 2  K K L T C Q V A T L N N N G E E A S T E G T C C P V N W V E H Q D S C Y W F S H  173
Canonical MGL  K K L T C Q V A T L N N N - - - A S T E G T C C P V N W V E H Q D S C Y W F S H  197
               * * * * * * * * * * * * *       * * * * * * * * * * * * * * * * * * * * * * * *

MGL isoform 2  S G M S W A E A E K Y C Q L K N A H L V V I N S R E E Q N F V Q K Y L G S A Y T  213
Canonical MGL  S G M S W A E A E K Y C Q L K N A H L V V I N S R E E Q N F V Q K Y L G S A Y T  237
               * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

MGL isoform 2  W M G L S D P E G A W K W V D G T D Y A T G F Q N W K P G Q P D D W Q G H G L G  253
Canonical MGL  W M G L S D P E G A W K W V D G T D Y A T G F Q N W K P G Q P D D W Q G H G L G  277
               * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

MGL isoform 2  G G E D C A H F H P D G R W N D D V C Q R P Y H W V C E A G L G Q T S Q E S H  292
Canonical MGL  G G E D C A H F H P D G R W N D D V C Q R P Y H W V C E A G L G Q T S Q E S H  316
               * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

**Figure 8.16.:** ClustalW sequence alignment of canonical MGL and isoform 2. Cytosolic, transmembrane, neck region, and CRD parts are in blue, black, red and green letters (domain organization based on information from UniProtKB); the putative coiled-coil region in the neck domain is underlined (based on information from UniProtKB); the endocytosis signal motif is highlighted in green; the conserved motifs defining sugar specificity and $Ca^{2+}$-coordination are highlighted in orange and blue, and the third conserved motif, WMGL instead of WIGL in case of MGL, is highlighted in red; all cysteine residues in extracellular part of isoform 2 are highlighted in yellow. The beginning of ECD and CRD sequences chosen for cloning are marked by blue and black sticks.

Three alternative splicing variants of MGL are known: isoform 1, also called "canonical" is the longest one; isoform 2 has a deletion of amino acids 118-144 and an insertion of GEE after N173 with respect to canonical MGL; and isoform 3 has two deletions (118-144, and 284-316 amino acids with respect to canonical MGL), the same insertion as isoform 2 and the sequence replacement at position 226-283 (NFVQKYLGSA ... HGLGGGEDCA → VRASGTQFLR ... IQRNISKLLS). Isoform 2 of MGL was chosen for cloning (fig. 8.16).

To clone ECD and CRD parts of MGL isoform 2, the sequences comprising amino acids 61-292 and 153-292, respectively, were chosen.

To all of the cloned constructs, LSECtin or MGL, N-terminal tag was included:

MASWSHPQFEKIEGR (5´-ATG GCT AGC TGG AGC CAC CCG CAG TTC GAA AAA ATC GAA GGG CGC-3´), where green, red and blue parts correspond to linker, StrepTag II and Factor Xa cleavage site sequences, respectively.

### 8.2.2. *Results of LSECtin and MGL S-ECD and S-CRD construct cloning into pEt20b and pEt30b vectors*

#### 8.2.2.1. *Generation of pEt20b and pEt30b constructs of LSECtin and MGL S-ECD and S-CRD*

LSECtin and MGL S-ECD and S-CRD pEt20b and pEt30b constructs were generated by a general strategy represented in figure 8.17.

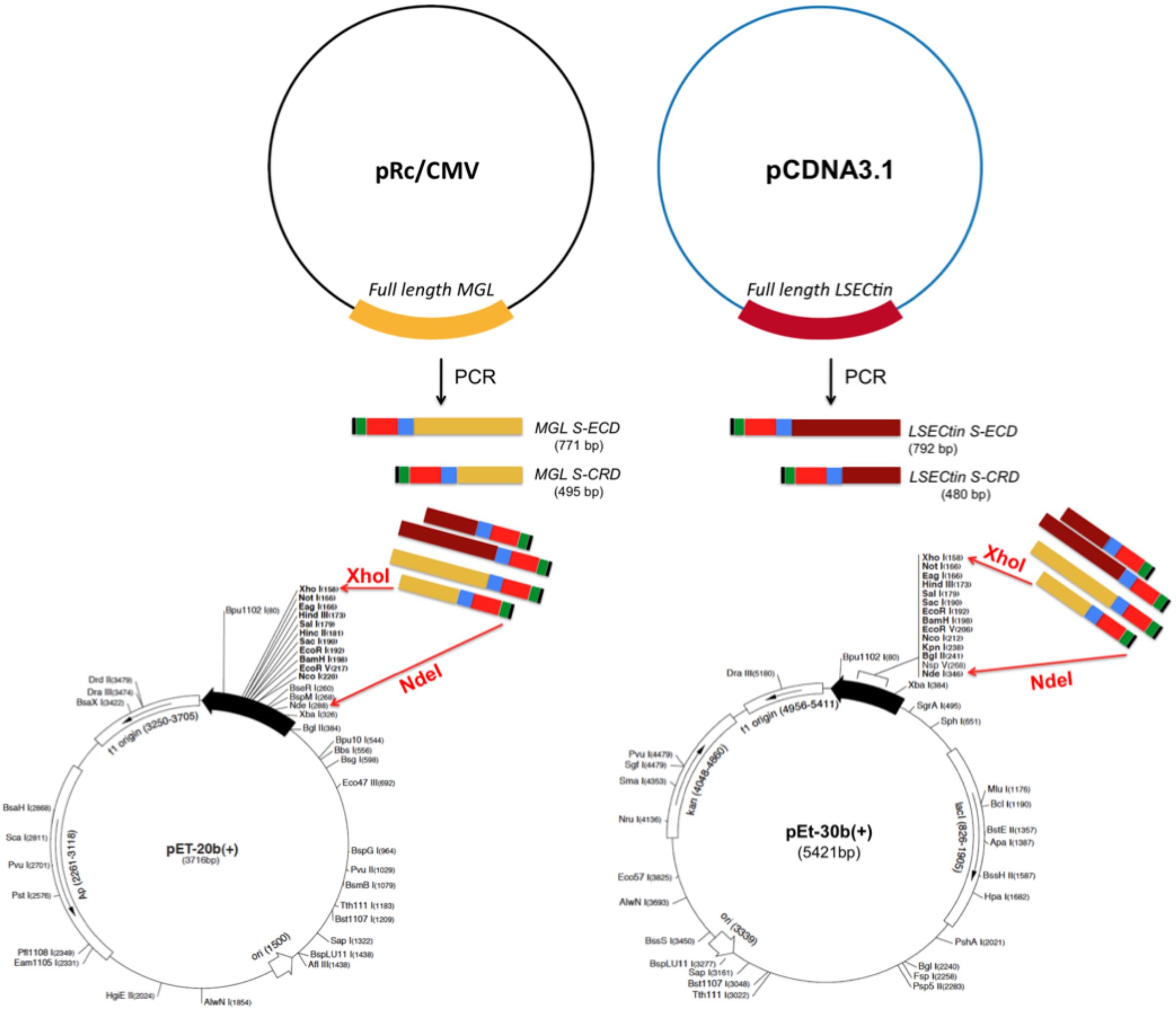

pRc/CMV
Full length MGL
pCDNA3.1
Full length LSECtin
PCR
PCR
MGL S-ECD
(771 bp)
MGL S-CRD
(495 bp)
LSECtin S-ECD
(792 bp)
LSECtin S-CRD
(480 bp)
XhoI
NdeI
XhoI
NdeI
pET-20b(+)
(3716bp)
f1 origin (3250-3705)
Ap (2261-3118)
ori (1500)
Bpu1102 I(80)
Xho I(158)
Not I(166)
Eag I(166)
Hind III(173)
Sal I(179)
Hinc II(181)
Sac I(190)
EcoR I(192)
BamH I(198)
EcoR V(217)
Nco I(220)
BseR I(260)
BspM I(268)
Nde I(288)
Xba I(326)
Bgl II(384)
Bpu10 I(544)
Bbs I(556)
Bsg I(598)
Eco47 III(692)
BspG I(964)
Pvu II(1029)
BsmB I(1079)
Tth111 I(1183)
Bst1107 I(1209)
Sap I(1322)
BspLU11 I(1438)
Afl III(1438)
AlwN I(1854)
HgiE II(2024)
Eam1105 I(2331)
Pfl1108 I(2349)
Pst I(2576)
Pvu I(2701)
Sca I(2811)
BsaH I(2868)
BsaX I(3422)
Dra III(3474)
Drd II(3479)
pEt-30b(+)
(5421bp)
f1 origin (4956-5411)
kan (4048-4860)
lacI (826-1905)
ori (3339)
Bpu1102 I(80)
Xho I(158)
Not I(166)
Eag I(166)
Hind III(173)
Sal I(179)
Sac I(190)
EcoR I(192)
BamH I(198)
EcoR V(206)
Nco I(212)
Kpn I(238)
Bgl II(241)
Nsp V(268)
Nde I(346)
Xba I(384)
SgrA I(495)
Sph I(651)
Mlu I(1176)
Bcl I(1190)
BstE II(1357)
Apa I(1387)
BssH II(1587)
Hpa I(1682)
PshA I(2021)
Bgl I(2240)
Fsp I(2258)
Psp5 II(2283)
Tth111 I(3022)
Bst1107 I(3048)
Sap I(3161)
BspLU11 I(3277)
BssS I(3450)
AlwN I(3693)
Eco57 I(3825)
Nru I(4136)
Sma I(4353)
Sgf I(4479)
Pvu I(4479)
Dra III(5180)

**Figure 8.17.:** The strategy of LSECtin and MGL S-ECD and S-CRD cloning.
Upper panel shows the generation of corresponding lectin constructs coding cDNA sequences; lower panel illustrates generated cDNA insertion into selected vectors. Black, green red and blue tag colors correspond to methionine, linker region, StrepTag II and Factor Xa cleavage site, respectively.

The PCR products (fig. 8.18) were inserted at first into pEt30b then pEt20b vectors between XhoI and NdeI restriction sites, and the results of insertion to pEt30b vector are shown in figure 8.19.

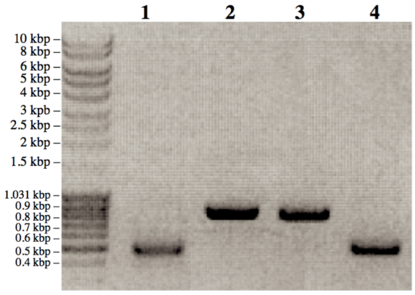

1
2
3
4
10 kbp
8 kbp
6 kbp
5 kbp
4 kbp
3 kpb
2.5 kbp
2 kbp
1.5 kbp
1.031 kbp
0.9 kbp
0.8 kbp
0.7 kbp
0.6 kbp
0.5 kbp
0.4 kbp

**Figure 8.18.:** The results of generation of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences by PCR.

The agarose gel of amplified corresponding cDNA: lane 1 – LSECtin S-CRD, lane 2 – LSECtin S-ECD, lane 3 – MGL S-ECD, lane 4 – MGL S-CRD.

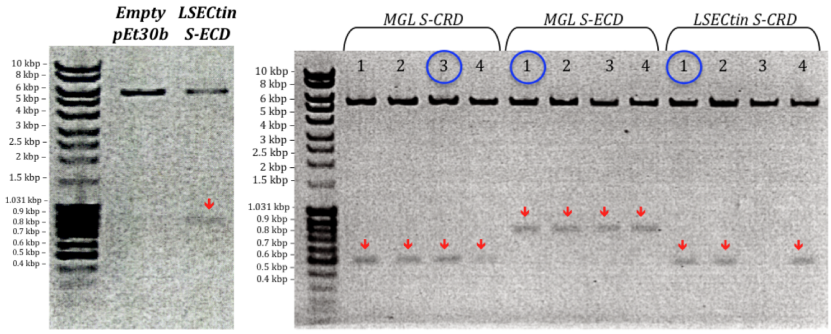

Empty pEt30b
LSECtin S-ECD
MGL S-CRD
MGL S-ECD
LSECtin S-CRD
1
2
3
4
1
2
3
4
1
2
3
4
10 kbp
8 kbp
6 kbp
5 kbp
4 kbp
3 kbp
2.5 kbp
2 kbp
1.5 kbp
1.031 kbp
0.9 kbp
0.8 kbp
0.7 kbp
0.6 kbp
0.5 kbp
0.4 kbp

**Figure 8.19.:** The results of insertion of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences to pEt30b vector.

The agarose gels of XhoI/NdeI digested corresponding plasmids: left panel shows empty pEt30b vector and pEt30b/LSECtin S-ECD; right panel shows digested pEt30b vector containing corresponding lectin coding cDNA (as marked in the figure) extracted from 4 different clones in each case. The target inserts are marked by red arrows, and the clones marked by blue circles were sequence checked.

Most of the clones contained the target cDNA inserts. The shown clone of LSECtin S-ECD, and LSECtin S-CRD clone 1, MGL S-CRD clone 3 and MGL S-ECD clone 1 were sequence checked and proved to be correct, thus chosen for further investigation.

The results of generated lectin cDNA insertion to pEt20b vector are shown in figure 8.20.

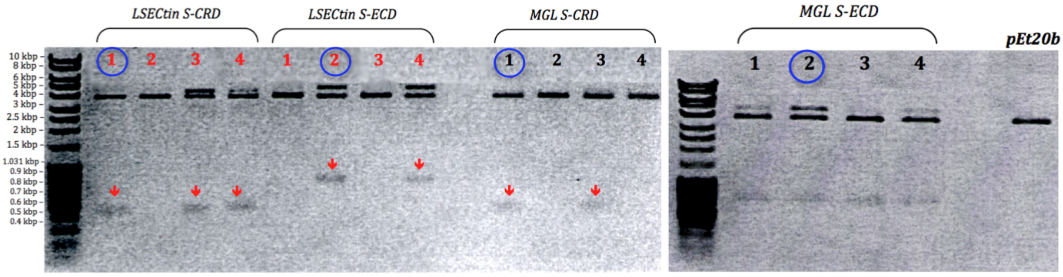

LSECtin S-CRD
LSECtin S-ECD
MGL S-CRD
MGL S-ECD
pEt20b
1
2
3
4
1
2
3
4
1
2
3
4
1
2
3
4
10 kbp
8 kbp
6 kbp
5 kbp
4 kbp
3 kbp
2.5 kbp
2 kbp
1.5 kbp
1.031 kbp
0.9 kbp
0.8 kbp
0.7 kbp
0.6 kbp
0.5 kbp
0.4 kbp

**Figure 8.20.:** The results of insertion of LSECtin and MGL S-ECD and S-CRD coding cDNA sequences to pEt20b vector.

The agarose gels of XhoI/NdeI digested prepared plasmids: both panels show digested pEt20b vector containing corresponding lectin cDNA (as marked in the figure) extracted from 4 different clones in each case, and an empty digested pEt20b vector. The target inserts are marked by red arrows, and the clones marked by blue circles were sequence checked.

Most of the clones contained the target cDNA inserts. LSECtin S-ECD clone 2, and LSECtin S-CRD clone 1, MGL S-CRD clone 1 and MGL S-ECD clone 2 were sequence checked and proved to be correct, thus chosen for further investigation.

#### 8.2.2.2. *Results of LSECtin and MGL S-ECD and S-CRD pEt20b and pEt30b construct functionality evaluation*

The capability of generated lectin constructs to express corresponding proteins was tested in *E.coli* BL21(DE3) strain. The protein expression location was also checked by preparing samples of total expression (whole cells), inclusion bodies (the pellet after cell disruption) and cytoplasm/periplasm (supernatant after cell disruption). The results of expression test of pEt30b constructs are shown in figure 8.21 (the expected molecular weights are 28 kDa and 17.8 kDa for MGL S-ECD and S-CRD constructs, and 28.3 kDa and 17 kDa for LSECtin S-ECD and S-CRD constructs, respectively).

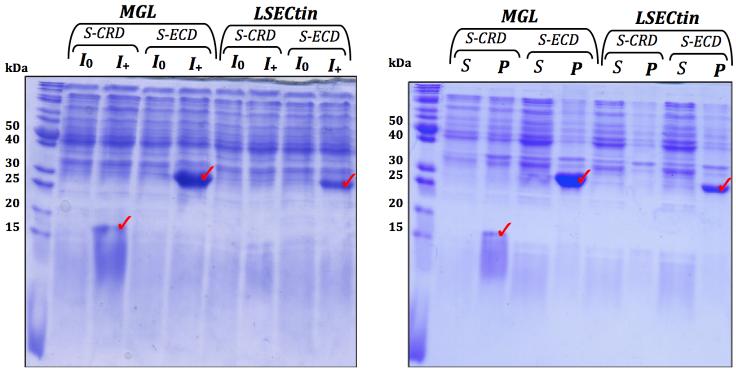

MGL
LSECtin
S-CRD
S-ECD
S-CRD
S-ECD
$I_0$
$I_+$
$I_0$
$I_+$
$I_0$
$I_+$
$I_0$
$I_+$
kDa
50
40
30
25
20
15
MGL
LSECtin
S-CRD
S-ECD
S-CRD
S-ECD
S
P
S
P
S
P
S
P
kDa
50
40
30
25
20
15

**Figure 8.21.:** The results of LSECtin and MGL S-ECD and S-CRD production by pEt30b constructs.

The left panel shows total expression results with loaded samples as marked in the figure ($I_0$ and $I_+$ notations correspond to samples before and after induction, respectively), and the right panel shows protein expression localization ($S$ and $P$ notations correspond to supernatant and pellet of disrupted cells, respectively). Red ticks mark the target protein bands.

All of the pEt30b constructs, except for LSECtin S-CRD, appeared to express corresponding lectins at high quantities in inclusion bodies. No expression at all was observed for LSECtin S-CRD construct. Similar results were obtained for pEt20b constructs (fig. 8.22).

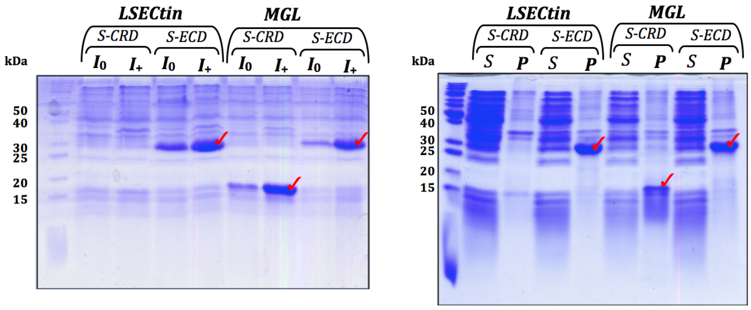

LSECtin
MGL
S-CRD
S-ECD
S-CRD
S-ECD
$I_0$
$I_+$
$I_0$
$I_+$
$I_0$
$I_+$
$I_0$
$I_+$
kDa
50
40
30
25
20
15
LSECtin
MGL
S-CRD
S-ECD
S-CRD
S-ECD
S
P
S
P
S
P
S
P
kDa
50
40
30
25
20
15

**Figure 8.22.:** The results of LSECtin and MGL S-ECD and S-CRD production by pEt20b constructs.

The left panel shows total expression results with loaded samples as marked in the figure, and the right panel shows protein expression localization. Notations are same as in figure 8.21. Red ticks mark the target protein bands.

As expected, pEt20b constructs had a “leaking” expression. Similarly as pEt30b constructs, the lectins were expressed at relatively high quantities in inclusion bodies, except for LSECtin S-CRD, which showed apparently no expression. Therefore, a different LSECtin S-CRD construct had to be designed.

### 8.2.3. *Results of generation of LSECtin S-CRD construct to direct expression to periplasm*

#### 8.2.3.1. *The results of LSECtin S-CRD construct cloning*

Although it was not clear what were the underlying reasons for the absence of LSECtin S-CRD expression in pEt20b and pEt30b constructs, it was chosen to design a new construct, which would be directed to the periplasm of E.coli. This strategy was based on the use of OmpA protein signal sequence in pEt30b vector. Because OmpA signal peptide coding sequence (ompA, 213 bp) is not included in commercially available pEt30b vector, it was cleaved out from pASK-IBA6 vector using HindIII and XbaI restrictases as shown in figure 6.1 on page 108 (results of excision in fig. 8.23), and then inserted to pEt30b vector between HindIII and XbaI cleavage sites.

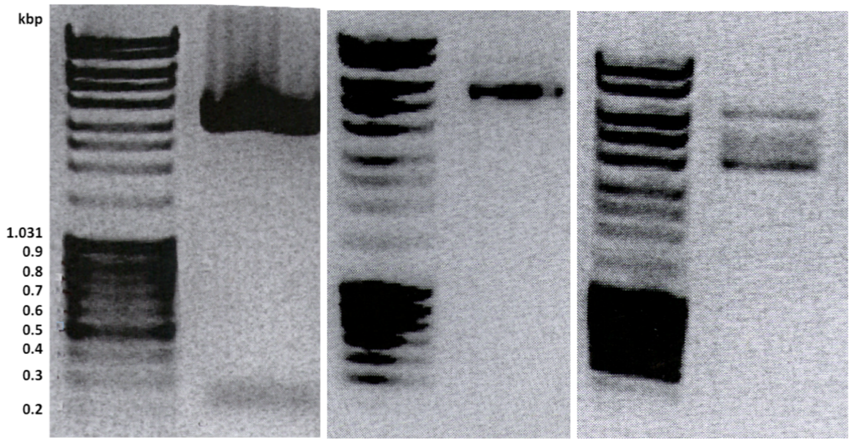


**Figure 8.23.:** The results of design of pEt30b vector containing OmpA signal peptide coding sequence.
The left panel: an agarose gel showing ompA (213 bp) cleaved out from pASK-IBA6 vector by HindIII/XbaI restrictases. The middle and right panels correspond to BsaI digested and non-digested pEt30b-ompA vector, respectively.

The success of the excised ompA sequence insertion into pEt30b vector was checked by digestion with BsaI restrictase: since BsaI site is present only in ompA sequence, the cleavage would relax plasmid DNA as it would become linear and thus result in a single band in agarose gel compared to multiple bands of supercoiled non-cleaved plasmid. Indeed, the digested pEt30b-ompA vector gave a single band (fig. 8.23) and after sequence checking the insertion approved to be correct.

The next step was the insertion of LSECtin S-CRD coding cDNA into the constructed pEt30b-ompA vector. This was accomplished by site-directed mutagenesis. At first, LSECtin S-CRD coding cDNA (455 bp) was amplified by PCR from previously constructed pEt30b/LSECtin S-CRD construct using primers that included a part annealing to ompA sequence (at StrepTag site) and another part annealing to the beginning of CRD coding sequence. PCR results are shown in figure 8.24. Then this new LSECtin S-CRD coding cDNA sequence was used as a primer in the following

PCR to insert it into pEt30b-ompA vector (used as a template DNA in the PCR).

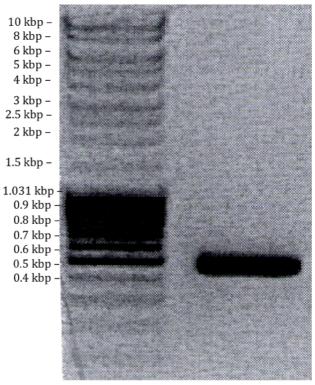


**Figure 8.24.:** The PCR results of LSECtin S-CRD coding cDNA that included ompA-annealing sequences (455 bp) amplification.

The resulting construct was tested by XbaI/HindIII digestion, and after sequence checking proved to be correct (the results of digestion are not shown because the intensity of the corresponding band was too low to be visible in the scan of the gel photo).

### 8.2.3.2. *The results of pEt30b-ompA/LSECtin S-CRD construct functionality evaluation*

The capability of constructed pEt30b-ompA/LSECtin S-CRD vector to express LSECtin S-CRD together with expression location was tested in BL21(DE3) *E.coli* strains. The samples before induction ($I_0$), after induction whole disrupted cells ($I_+$), insoluble ($P$) and soluble parts ($S$) of disrupted cells were analyzed by SDS-PAGE. Because the presence of protein expression was not evident on the gel (fig. 8.25 left panel), it was blotted on PVDF membrane and protein was visualized using anti-StrepTag II antibodies (fig. 8.25 right panel).

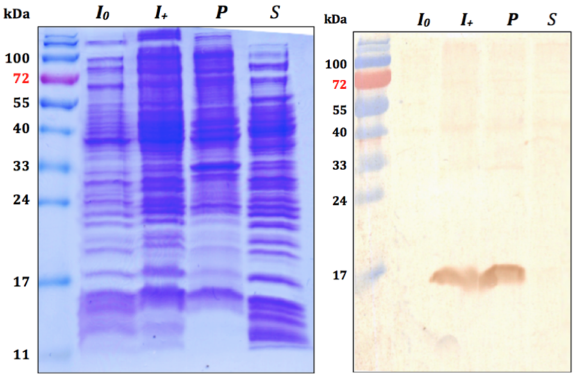


**Figure 8.25.:** The functionality of pEt30b-ompA/LSECtin S-CRD construct to express LSECtin S-CRD.

The left and right panels show SDS-PAGE and WB results, respectively, with the corresponding samples as marked on the figure. The WB was revealed using anti-StrepTag II antibody.

It appeared that LSECtin S-CRD was expressed in inclusion bodies but not as a soluble protein. Because the presence of protein expression is poorly seen on the gel, it is possible that earlier pEt20b and pEt30b constructs also express lectin but at low quantities (protein expression presence wasn't analyzed by WB).

### 8.2.4. *MGL S-ECD over-expression and purification results*

The expression level of MGL S-ECD (28 kDa) was high (Figure 8.26). The lectin was purified from two batches of 0.5 L cell culture.

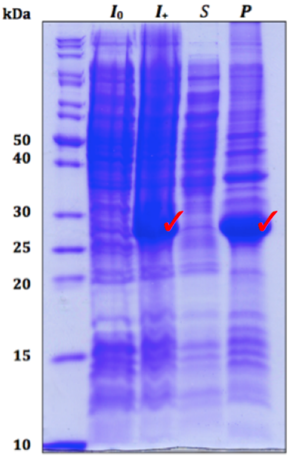


**Figure 8.26.:** The SDS-PAGE analysis of MGL S-ECD expression.
Samples before ($I_0$) and after induction containing whole disrupted cells ($I_+$), supernatant ($S$) and pellet ($P$) were loaded.

Inclusion bodies were extracted, solubilized in Gdn-HCl and refolded in the same manner as DC-SIGN ECD and S-ECD. During refolding of solubilized inclusion bodies, extensive precipitation was observed. The refolded protein was purified in two steps. At first the protein solution (860 mL, 0.106 mg/mL) was loaded to StrepTactin column and eluted with 2.5 mM desthiobiotin (fig. 8.27).

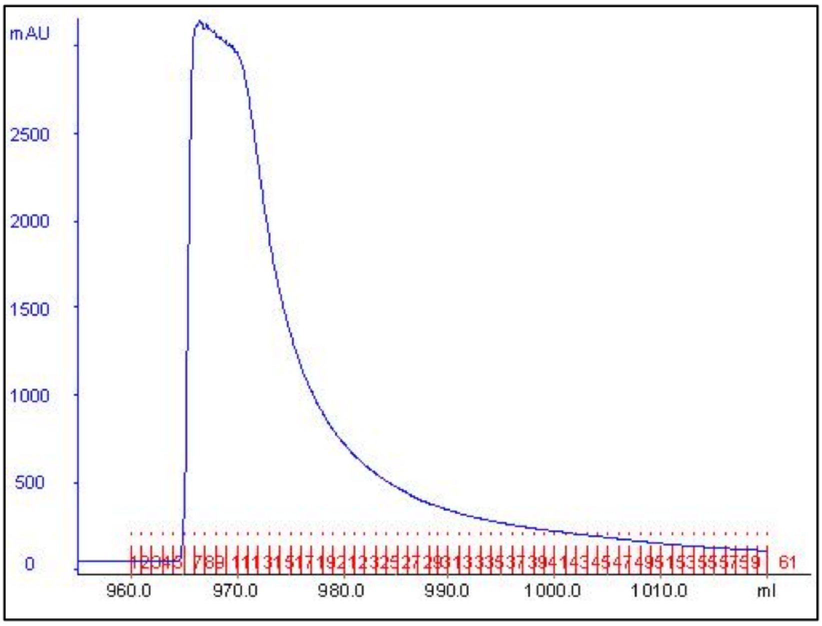


**Figure 8.27.:** The chromatogram of MGL S-ECD elution by desthiobiotin from StrepTactin column.

The next step of purification was performed to isolate functional protein. Because of MGL specificity to galactose, Superose12 column, which contains galactose-based polymer agarose, could be used as a combination of an affinity and size-exclusion column. The sample of MGL S-ECD eluted from StrepTactin, was concentrated to 3 mg/mL and injected (1 mL) to 125 mL Superose12 column (fig. 8.28).

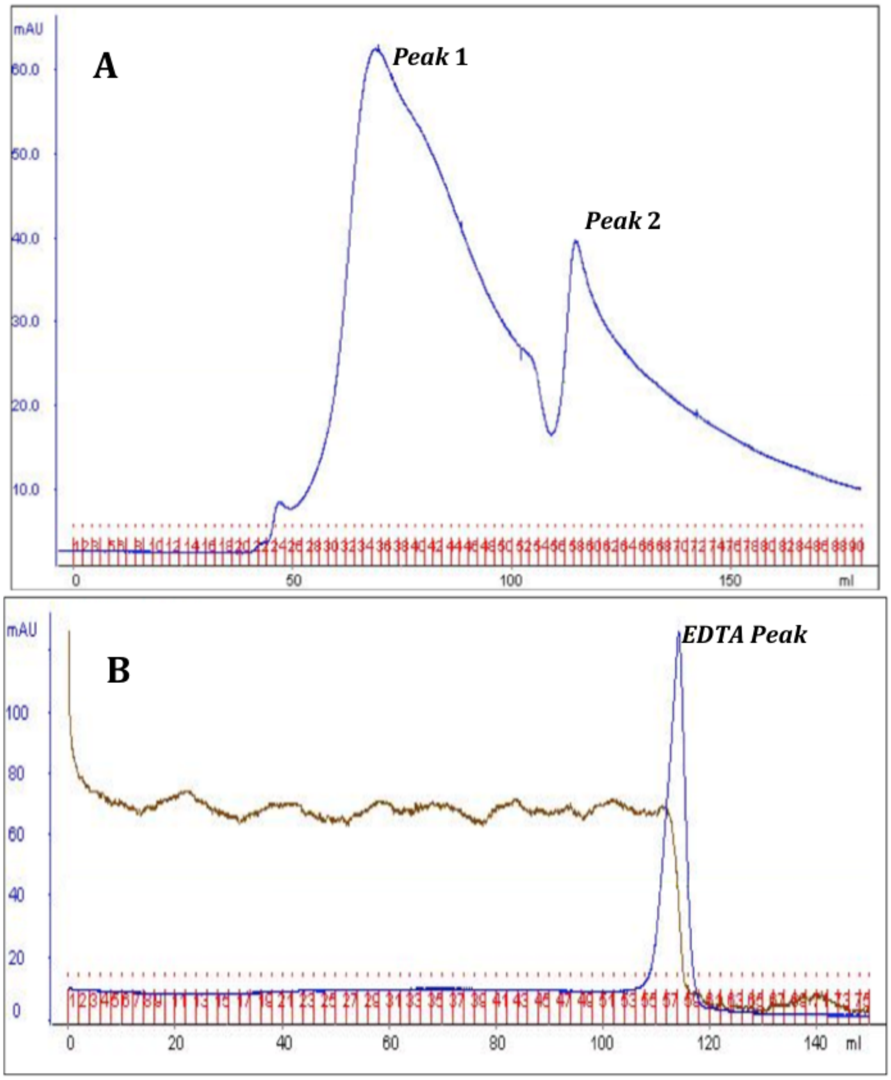


**Figure 8.28.:** The chromatograms of MGL S-ECD injection to 125 mL Superose12 column. A, The elution profile in running buffer. B, The retained protein elution with buffer containing 1 mM EDTA (brown line represents conductivity changes).

Two major species of MGL S-ECD migrating through the column could be observed (fig. 8.28A), and deduced from elution time they both contained most likely non-functional incorrectly folded protein of various sizes. However, the activity of part of MGL S-ECD preparation was suggested by the observation that using buffer containing EDTA, the bound protein, i.e. functional, was washed out (fig. 8.28B).

To confirm the functionality of purified MGL S-ECD, a column with immobilized galactose (later refered “galactose column”) was used. The sample of MGL S-ECD eluted from StrepTactin was concentrated and injected (1 mL) to the galactose column (fig. 8.29).

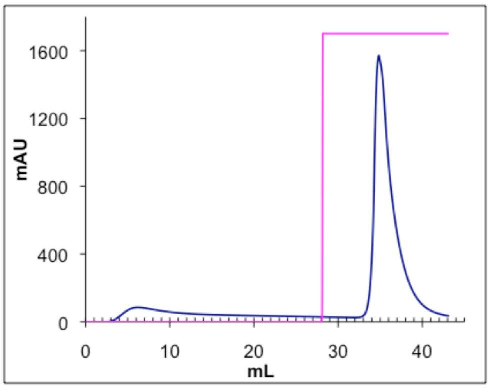


**Figure 8.29.:** The chromatogram of MGL S-ECD elution profile on immobilized galactose column. Blue and pink lines represent absorbance changes and the switch from running to elution buffer (from 0% to 100% of buffer containing 1 mM EDTA), respectively.

It turned out that the vast majority of the MGL S-ECD sample eluted from StrepTactin column was functional as almost of the all injected protein left the column in EDTA elution peak. Thus the sample heterogeneity observed on Superose12 column (fig. 8.28) is likely caused by different oligomeric forms of probably active MGL S-ECD.

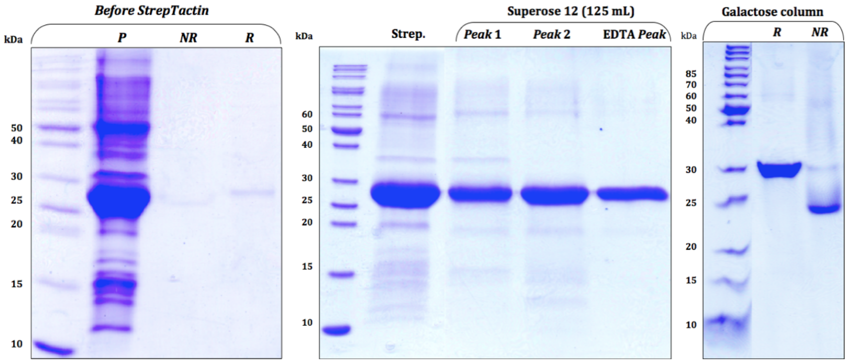


**Figure 8.30.:** The SDS-PAGE analysis of MGL S-ECD purification steps.
Reduced (*R*) or non-reduced (*NR*) samples were loaded. Markings: the gel "Before StrepTactin" shows samples after refolding, i.e. the reduced sample of precipitate after refolding (*P*) and supernatant (*R* and *NR*); the middle gel shows concentrated reduced sample of protein eluted from StrepTactin column (i.e. "*Strep.*") and concentrated reduced samples from Superose12 (fig. 8.28); the gel "Galactose column" shows samples eluted from galactose column.

All the purification steps were monitored by SDS-PAGE loading reduced (*R*) or non-reduced (*NR*) MGL S-ECD samples (fig. 8.30). The major part of the precipitate after refolding contained MGL S-ECD. The Superose12 purification with EDTA elution seems to have worked well since the resulting sample (fig. 8.28, "*EDTA Peak*") is rather homogeneous; the "*Peak 1*" and "*Peak 2*" samples contain various size contaminants, partially explaining the elution profile heterogeneity, though the major part of the samples is comprised by MGL S-ECD. The faster migration of the non-reduced sample after galactose column might indicate that intramolecular disulphide bonds render the molecule more compact compared to disulphide-free molecule allowing the faster migration. It is also noticeable that the sample eluted from galactose column contains disulphide-linked oligomers of various sizes (the blue trail above the major band).

The yield of the protein evolved as follows: 91 mg of total protein after refolding were loaded to StrepTactin and 29 mg were recovered from the column in a desthiobiotin elution peak. From 3 mg of protein injected to Superose12 column, 0.2 mg were eluted (EDTA peak) as homogeneous active protein. From these results it can be extrapolated that the injection of 29 mg could be expected to yield 1.9 mg of pure protein and this quantity could be considered as the yield from 1 L culture.

Generally, the MGL S-ECD production process should be improved most importantly at the refolding stage, and the buffer formulation to stabilize protein should be found (the purified samples were precipitating). The order of purification procedure should be also changed: first purification on StrepTactin column followed by affinity purification on galactose column and finally gel-filtration using the column with the media, for which MGL doesn't have affinity. Also the loading to StrepTactin seemed to over-saturate the column, as judged from absorbance and peak profile, thus loading in smaller portions or loading the flow-through could probably help to get a higher yield of purified protein. It is also possible that StrepTactin column is not necessary at all, since the ECD construct presents three CRDs and may have sufficient affinity for galactose column at the first step of purification.

## 8.3. Concluding remarks on different lectin production

The production of DC-SIGN ECD construct is efficient without any major bottleneck, and considerable yields of purified functional protein are obtained (typically 50-80 mg from 1 L of cell culture). Surprisingly, the addition of StrepTag II at N-terminus of DC-SIGN ECD markedly decreases the yield of purified protein mainly due to strong precipitation during refolding. Nevertheless, only low quantities of this construct are required for the preparation of oriented surface for SPR direct interaction assay.

Unfortunately, the efforts to optimize DC-SIGN S-CRD production yield did not give any improvements. However, the reason for low yields was revealed: while being concentrated, the protein forms nonfunctional disulfide-linked oligomers.

Four new C-type lectin constructs were developed for the use in the laboratory in the future: LSECtin S-CRD and S-ECDs as well as MGL S-CRD and S-ECDs. Only with the exception for LSECtin S-CRD, all other constructs were capable to efficiently over-express corresponding proteins. Nonetheless, LSECtin S-CRD construct is also produced but expression level is lower.

The production of MGL S-ECD was investigated a step further, and the first trials of purification of this construct were performed. Although the promising results were obtained, the procedure should be optimized.

The development of these four new constructs discontinued when the first series of the glycomimetic compounds were synthesized. From then on, the SPR analysis of the glycomimetic compounds, which were continuously synthesized and sent for analysis by different chemists groups, became the priority of the work. From that point, only DC-SIGN and langerin ECD constructs were used for the studies (I have been preparing DC-SIGN, and langerin ECD was prepared by another PhD student in the group, Eric Chabrol).

# 9. The results of the development of glycomimetic DC-SIGN antagonists

## 9.1. Development of monovalent glycomimics

### 9.1.1. *The SPR competition assay characteristics*

The monovalent glycomimetic compounds were evaluated as DC-SIGN binders by SPR competition assay, which allows to assess and compare the capacities of the compounds to block DC-SIGN ECD interaction with ManBSA, used as an artificial ligand of DC-SIGN due to its lower cost as compared to gp120. All of the compounds were prepared and analyzed in the running buffer without DMSO (see subsection 6.2.3 on page 113), and if DMSO was required, the concentration used will be noted.

#### 9.1.1.1. *The experimental design of SPR competition assay*

Figure 9.1 illustrates the principal design of the SPR competition assay. ManBSA, bearing 12 branched trisaccharides ($\alpha$1-3, $\alpha$1-6 mannotriose) per BSA molecule in average, was immobilized on the CM-dextran surface of the sensor chip via amine coupling chemistry, and the quantity of immobilized ManBSA varied from 1200 RU to 5000 RU. Then the samples containing a constant amount of lectin and varying concentrations of the compound were injected over the reference (activated/deactivated CM-dextran) and ManBSA surfaces.

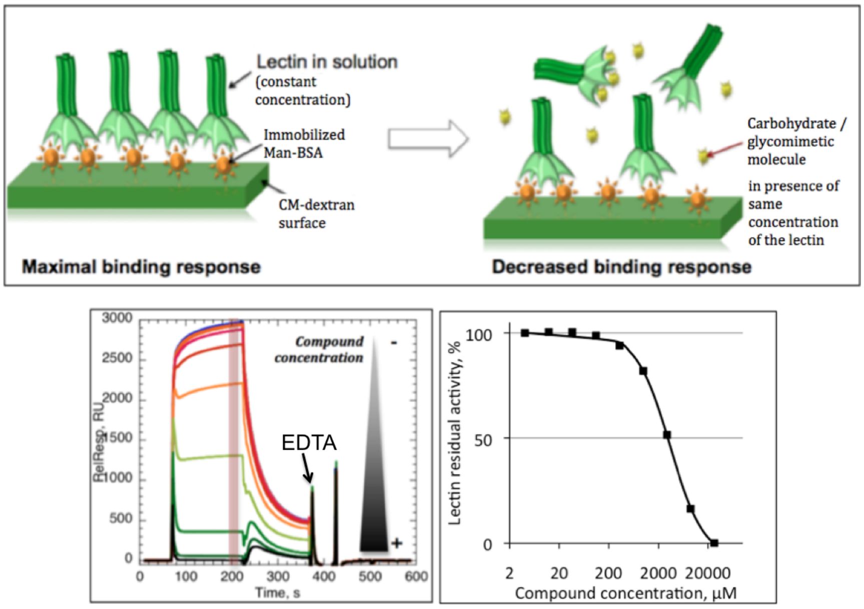


**Figure 9.1.:** The SPR competition assay.

The top part illustrates the experimental design of the assay, and the bottom part demonstrates the overlaid reference surface corrected sensorgrams showing the binding responses of DC-SIGN ECD (20 µM) in the presence of a compound at different concentrations (the red bar on the sensorgrams indicates an approximate interval, over which the binding responses were extracted, and the arrow shows the point of EDTA injection to regenerate the surface), and the corresponding inhibition curve.

The residual lectin activity values were calculated by normalizing the binding responses at equilibrium (extracted from reference surface corrected sensorgrams) with respect to the response of lectin alone sample, which was assigned a 100% activity value. Subsequently, the $IC_{50}$ for each compound were calculated by fitting eq. 6.2 (see sub-subsection 6.2.4.2 on page 114) to the plots of residual activity versus compound concentration.

In a few cases a simpler set-up of competition assay, so-called single-point assay, was used: only a single concentration of monovalent compounds (150 µM for mannose-based or 300 µM for fucose-based compounds) in a mixture with the lectin was used, and lectin binding responses in the presence of individual compounds were compared to lectin alone binding.

All of the sensorgrams and inhibition curves are presented in appendix.

#### 9.1.1.2. *The surface activity control during the experiment*

The activity of ManBSA surface in the course of the experiment was monitored in two ways. Firstly, the apparent affinity of DC-SIGN ECD to ManBSA surface was determined before and after the competition assay by fitting the steady state affinity model (eq. 9.1) to the plots of DC-SIGN binding responses versus concentration (fig. 9.2), and the apparent $K_\mathrm{D}$ values were compared. For all experiments very close affinities ($K_\mathrm{D}$ 0.8-1.5 µM) were obtained.

$$R_{eq} = \frac{K_A{\cdot}C{\cdot}R_{max}}{1 + K_A} \tag{9.1}$$

where $R_{eq}$ is equilibrium binding response, $K_\mathrm{A}$ equilibrium association constant ($K_\mathrm{D}{=}1/K_\mathrm{A}$), $C$ – concentration of the injected analyte, and $R_{max}$ – surface binding capacity (the maximum binding response that can be achieved for a given analyte with a given surface).

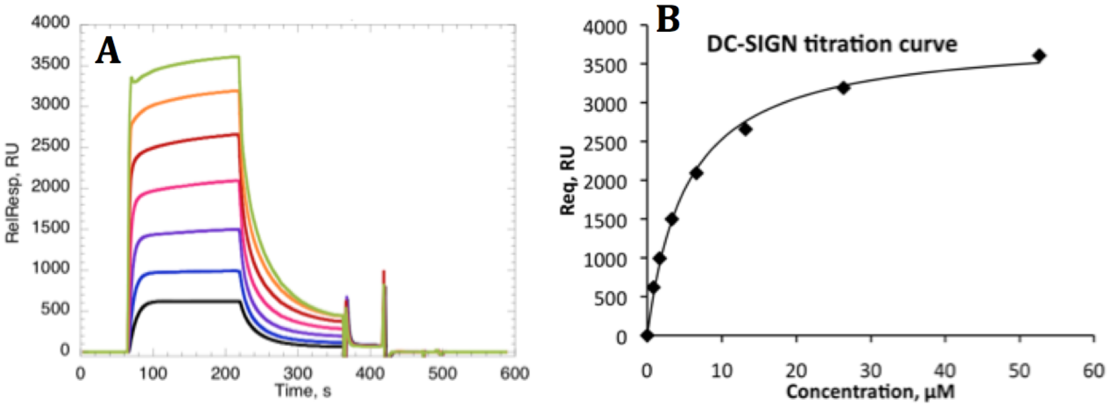


**Figure 9.2.:** The ManBSA surface titration with DC-SIGN ECD.
**A**, Reference surface corrected sensorgrams showing the binding of DC-SIGN ECD at concentrations ranging from 0.82 µM to 52.7 µM (with respect to binding sites). **B**, The plot of the corresponding DC-SIGN ECD binding responses as a function of DC-SIGN monomer concentration. The experimental points (◆) and the fit of eq. 9.1 (solid line) are shown. ManBSA was immobilized to 1700 RU, and the calculated apparent $K_D$ 1.3 µM for the whole tetrameric ECD.

In parallel, a second approach was used to monitor surface activity. In this case, the same sample of the lectin alone was injected in the beginning, in the middle and in the end of the experiment, and the binding responses were compared. Typically, 50-90 cycles of sample injections followed by surface regenerations were performed per experiment, and in the end of a such run the surface retained at least 95% of its primary activity.

Additional control of the experiments was done by including a reference compound (e.g. D-mannose, L-fucose) to the run and comparing the obtained $IC_{50}$ values between different experiments.

#### 9.1.1.3. *The evaluation of the variation of $IC_{50}$ values*

The variation of $IC_{50}$ values was evaluated comparing the values obtained for the same compound in several different experiments (table 9.1) where the ManBSA surface densities and DC-SIGN ECD preparations were different, and some of the experiments were performed by different operators.

**Table 9.1.:** Evaluation of variation of $IC_{50}$ values.

| Compound | Mean $IC_{50}$± SD, mM | Number of measurements |
|---|---|---|
| **D-mannose** | 3.29 ± 0.34 | 8 |
| **Man α1-2 Man** | 0.95 ± 0.11 | 7 |
| **psDi** (fig. 9.14 on page 173) | 1.02 ± 0.11 | 14 |
| **psTri** (fig. 9.14 on page 173) | 0.145 ± 0.083 | 9 |
| **4h** (fig. 9.17 on page 265) | 0.31 ± 0.04 | 24 |
| **L-fucose** | 2.06 ± 0.18 | 6 |
| **Lewis$^X$** | 0.77 ± 0.10 | 6 |
| **10b** (fig. 9.13 on page 172) | 0.395 ± 0.043 | 7 |
| **10b-azide1** (fig. 9.13 on page 172) | 0.328 ± 0.043 | 9 |

Under these conditions, the calculated standard deviations (SD), when normalized to the $IC_{50}$ mean value, varied at 10-13% range over a total of up to 24 measurements.

Interestingly, very high variation of $IC_{50}$ values for psTri was observed. This might be related to the ability of this compound to bridge DC-SIGN tetramers (see manuscript **n°5**), which could lead to the sensitivity of the DC-SIGN ECD preparation, i.e. number of active CRDs per tetramer.

#### 9.1.1.4. *The set-up of the evaluation of compound selectivity to DC-SIGN versus langerin*

The same principle of competition assay was used to determine the compounds' capacities to inhibit langerin ECD (15 µM, the concentration was lower than DC-SIGN ECD due to the limited availability of this lectin) binding to Man-BSA surface. However, langerin strongly interacted with the CM-dextran surface (fig. 9.3).

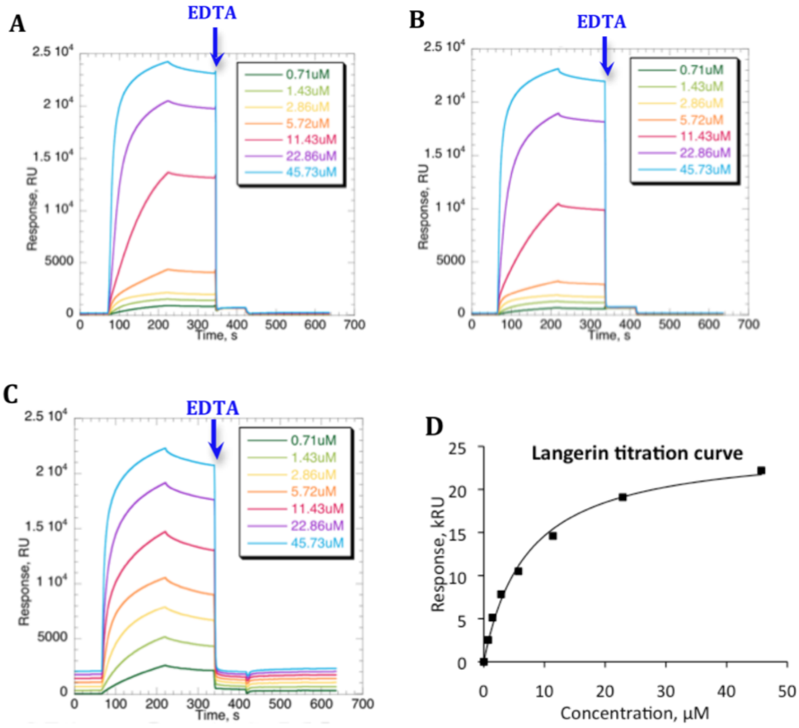


**Figure 9.3.:** The raw sensorgams representing titration of difference surfaces with langerin ECD. A range of langerin ECD concentrations was injected over the following surfaces: **A**, unmodified CM-dextran surface; **B**, activated/deactivated CM-dextran surface; and **C**, immobilized ManBSA (5000 RU) surface (the beginning of sensorgrams was not normalized to 0 RU to take into account the accumulation of protein not fully removed by regeneration after each cycle). **D**, a surface saturation curve by langerin ECD titration: binding responses corresponding to sensorgrams in panel C were plotted against langerin monomer concentration; the solid line represents the fit of eq. 9.1 to the experimental points and yielded an apparent $K_D$ of 2.4 µM for the whole trimeric ECD. The blue arrows mark the injection of surface regeneration solution, 50 mM EDTA pH 8.

Efforts were made to eliminate or significantly reduce this binding by preparing the reference surface with high amount of immobilized BSA in order to shield the dextran (fig. 9.4B), or by performing several cycles of CM-dextran activation/deactivation to decrease the carboxyl groups in order to reduce possible electrostatic interactions (fig. 9.4C).

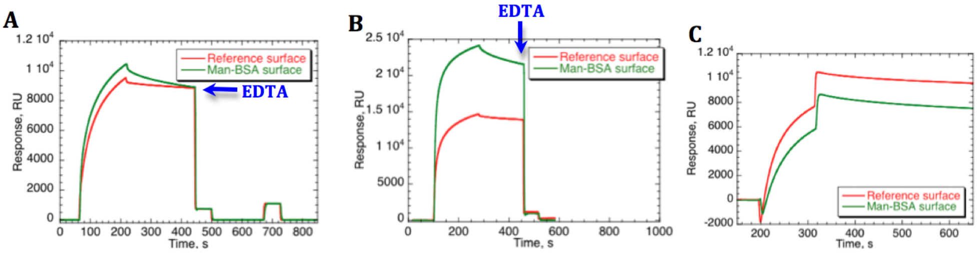


**Figure 9.4.:** The sensorgrams comparing langerin interaction with ManBSA and reference surfaces. Langerin ECD (15 µM, 50 µM and 16 µM for A, B and C, respectively) injections over corresponding reference (red lines) and ManBSA surfaces (green lines). **A**, the reference surface was EDC/NHS activated and ethanolamine deactivated CM-dextran, and Man-BSA was immobilized 1900 RU. **B**, the reference surface was coated with BSA (3300 RU) and Man-BSA was immobilized to 2000 RU. **C**, the reference and Man-BSA surfaces were pretreated with four alternating injections (50 µL) of EDC/NHS followed by ethanolamine, then 300 RU of Man-BSA was immobilized on the active surface; the high bulk response of langerin injection arose due to the difference in running (4% DMSO) and sample (no DMSO) buffers. The blue arrows mark the injection of surface regeneration solution, 50 mM EDTA pH 8.

All of these efforts did not give any plausible results. However, as can be seen from figures 9.3 and 9.4, the injection of EDTA abolished the interaction of langerin with the surface suggesting that this binding was $Ca^{2+}$-dependent and sugar determined, i.e. the interacting entity was the CRD of langerin. Indeed, langerin recognizes maltose [60], hence its specificity to dextran matrix is not surprising.

Finally, ManBSA/dextran surface was considered as a combined ligand of langerin. No reference surface subtraction was performed prior to extraction of binding responses, which were calculated using similar intervals as in case of DC-SIGN (fig. 9.1) just after the end of the injection in order to avoid bulk responses. Moreover, it appeared that the inhibition of langerin binding is highly dependent on ManBSA surface density, which was in the contrary to DC-SIGN, since it didn't show such dependency (an example from studies in article **n°4** are shown figure 9.5).

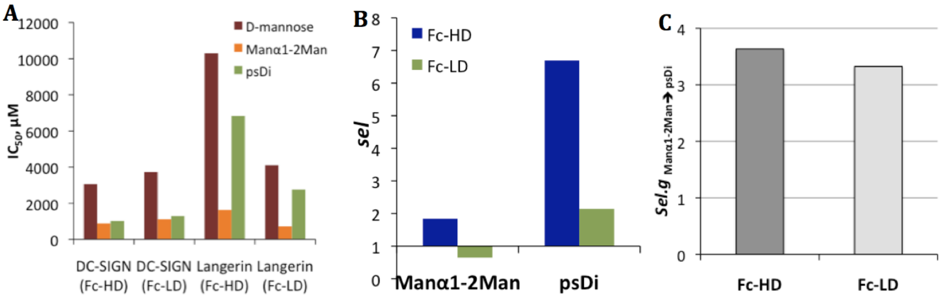


**Figure 9.5.:** The langerin inhibition dependence on ManBSA surface density. **A**, $IC_{50}$ values of langerin and DC-SIGN inhibition by indicated compounds that were obtained on high density (Fc-HD, 5000 RU of immobilized ManBSA) and low density (Fc-LD, 1100 RU of immobilized ManBSA) surfaces. **B**, the corresponding selectivities of Man α1-2 Man and psDi calculated according to eq. 9.2. **C**, the corresponding selectivity gains achieved by changing Man α1-2 Man to psDi, calculated according to eq. 9.3.

Despite high differences in absolute $IC_{50}$ values for different density surfaces, the general pattern of $IC_{50}$ value distribution among the compounds remained the same.

In order to get a more tangible evaluation of the selectivity of the compounds, the following approach was envisioned. The absolute selectivity of a particular compound ($sel_{comp.}$) for DC-SIGN *vs* langerin can be defined as a following ratio:

$$sel_{comp.} = \frac{IC_{50\,Langerin}}{IC_{50\,DC-SIGN}} \tag{9.2}$$

This means that $sel_{comp.} > 1$ indicates a selectivity of the compound in favor for DC-SIGN. However, selcomp. term is still entirely surface density-dependent (fig. 9.5B). To exclude this effect, a selectivity gain (*sel.g*) achieved by switching from one compound (reference, *comp.1*) to another (*comp.2*) can be considered:

$$sel.g_{comp.1 \rightarrow comp.2} = \frac{sel_{comp.2}}{sel_{comp.1}} \tag{9.3}$$

This term also provides a relative comparison of the selectivity improvement between the two lectins when switching from one compound to another, i.e. it answers the question "how much the selectivity of compound 2 is better than compound 1", and $sel.g_{comp.1 \rightarrow comp.2} > 1$ means a gain of selectivity to DC-SIGN *vs* langerin for compound 2 compared to compound 1, while values <1 mean a loss of selectivity for DC-SIGN. Indeed, the *sel.g* values are no longer dependent on surface density as can be seen in the example in figure 9.5C, which tells that chemically modifying a natural Manα1-2Man disaccharide, improved a selectivity to DC-SIGN of the resulting compound psDi by a factor of 3.5.

### 9.1.2. *C-glycoside development*

In order to generate compounds with higher resistance to the cleavage by glycosidases, the C-glycosides were constructed and assayed for their binding to DC-SIGN (C-glycosides are sugars where the anomeric oxygen is replaced by a carbon). The design and synthesis of these C-glycosides was conducted in the group of Pr. J. Moravcová by PhD student Benedetta Bertolotti.

The first trials were made with D-mannose (BB112) and L-fucose (BB133) (fig. 9.6 top panel) with the short linkers attached by C-glycosidic bond were synthesized. Then various disaccharides (fig. 9.6 middle and bottom panels) were synthesized and tested. All of the disaccharides, except ZL1(D), ZL2(L), were analyzed in the presence of 4% DMSO; ZL1(D) and ZL2(L) were analyzed in the absence of DMSO.

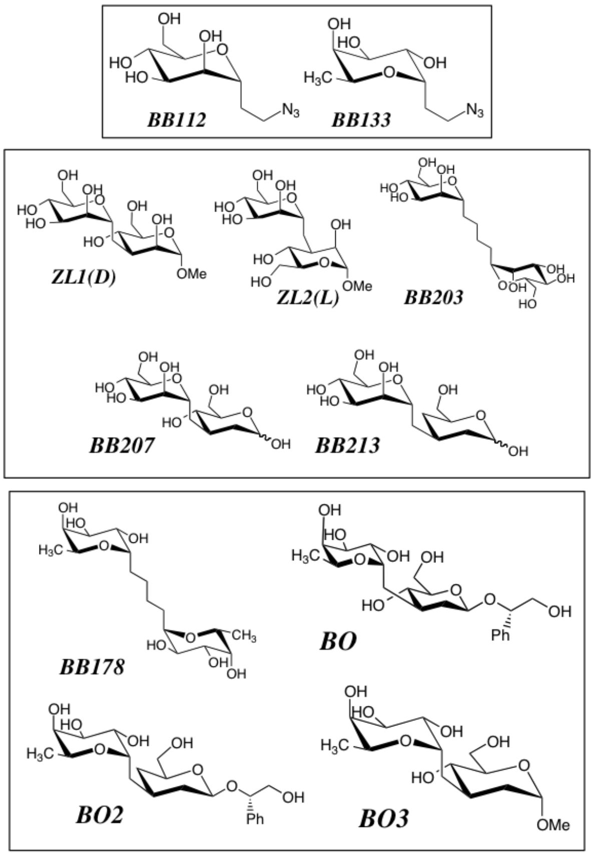


**Figure 9.6.:** The structures of C-glycosides.
Top panel shows the structure of D-mannose (BB112) and L-fucose (BB133) with the short linkers attached by C-glycosidic bond. Middle and bottom panels show the structures of mannose- and fucose-based pseudo-disaccharides, respectively.

All of these compounds were tested by SPR competition assay to assess their capacity of inhibiting DC-SIGN ECD binding to ManBSA surface and compared with the available natural counterparts, i.e. D-mannose, L-fucose, Manα1-2Man and Manα1-3Man.

Surprisingly, the L-fucose C-glycoside BB133 appeared to have 2-fold better activity of DC-SIGN inhibition than the natural L-fucose (fig. 9.7), although the underlying structural basis for such results is not clear. The mannose C-glycoside had exactly the same activity as D-mannose.

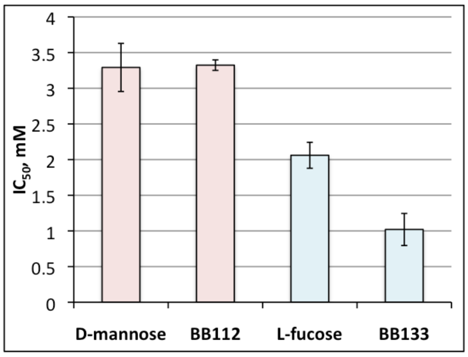


**Figure 9.7.:** DC-SIGN inhibition activities of D-mannose and L-fucose C-glycosidic monosaccharides compared to their natural counterparts.

D-mannose and L-fucose $IC_{50}$ values are the means as shown in table 9.1; the error bars are standard deviations (over two measurements of BB112 and BB133).

Among the pseudo-disaccharides, there were compounds that had very low solubility (i.e. BB178 and BO). They were solubilized in the running buffer containing 4% DMSO, and thus the assay was performed in DMSO (4%) presence.

The results of mannose-based pseudo-disaccharides showed that these compounds have similar activity as D-mannose (fig. 9.8).

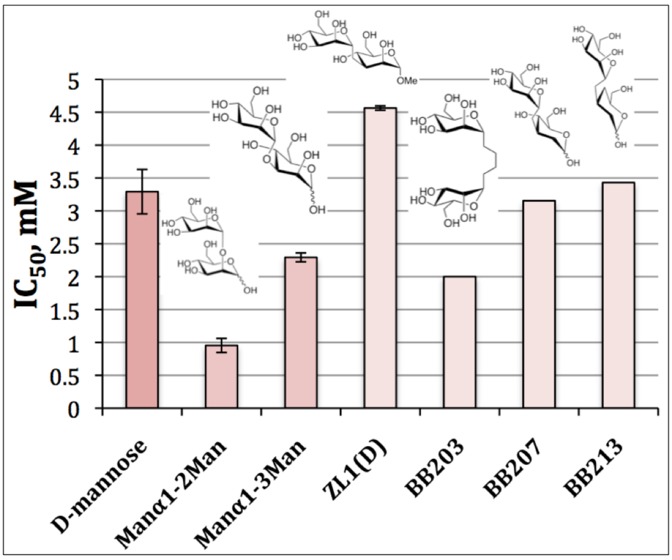


**Figure 9.8.:** DC-SIGN inhibition activities of D-mannose C-glycosidic pseudo-disaccharides compared to D-mannose, Manα1-2Man and Manα1-3Man.

D-mannose and Manα1-2Man $IC_{50}$ values are the means as shown in table 9.1; the error bars are standard deviations (over two measurements of Manα1-3Man and ZL1(D), other compounds run only once).

Only BB203, which has an elongated linker between two mannose moieties connected via anomeric carbons, showed a better activity than mannose. All other compounds had the same (BB207, BB213) or lower (ZL1(D)) activity as D-mannose suggesting that only non-reducing mannose moiety was interacting with the lectin and the second moiety had either no effect on binding or probably imposed some kind of hindrance for the binding (in case of ZL1(D)). Manα1-3Man has more than 2-fold lower activity than Manα1-2Man, and none of Manα1-3Man mimics had higher activity than the natural disaccharide.

Interestingly, all of the fucose-based pseudo-disaccharides had a better apparent affinity to DC-SIGN than L-fucose (fig. 9.9).

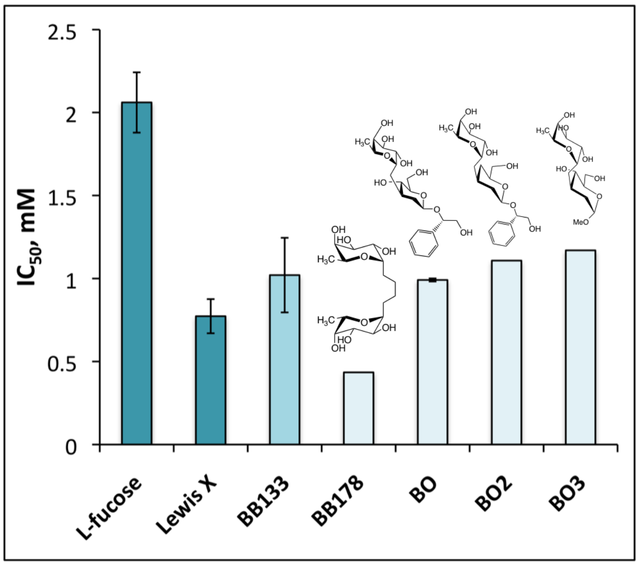


**Figure 9.9.:** DC-SIGN inhibition activities of L-fucose C-glycosidic pseudo-disaccharides compared to L-fucose and Lewis$^{X}$.

L-fucose and Lewis$^{X}$ IC$_{50}$ values are the means as shown in table 9.1; the error bars are standard deviations (over two measurements for BO, other compounds run only once).

Similarly to mannose-based pseudo-disaccharides, the best activity in these fucose-based compound series belonged to the compound with an elongated alkyl linker between two L-fucose moieties attached via anomeric carbons, i.e. compound BB178. The activity of this compound was even higher than of Lewis$^{X}$ trisaccharide. It is noteworthy that BB178 was the least soluble compound (part of the compound remained non-dissolved when final concentration of 20 mM (7 mg/mL) was aimed) meaning that its activity is underestimated. However, low water solubility is a disadvantage in the development of drug candidates.

The other three fucose-based pseudo-disaccharides had about 2-fold better activity than L-fucose suggesting that the addition of corresponding moieties to L-fucose improved their affinity to DC-SIGN.

It is interesting to notice that among mannose-based pseudo-disaccharides, there was a compound, namely ZL2(L), with a bizarre effect observed (fig. 9.10). The presence of higher concentrations ($\geq$2 mM) of ZL2(L) seemed to trigger DC-SIGN binding to dextran surface (fig. 9.10B), since 10 mM ZL2(L) alone did not show interaction with the reference surface, but slightly interacted with ManBSA surface (100 RU bound) bound (fig. 9.10C). The same binding to ManBSA surface was also observed for compound ZL1(D) (fig. 9.10D), which did not have such effect as ZL2(L).

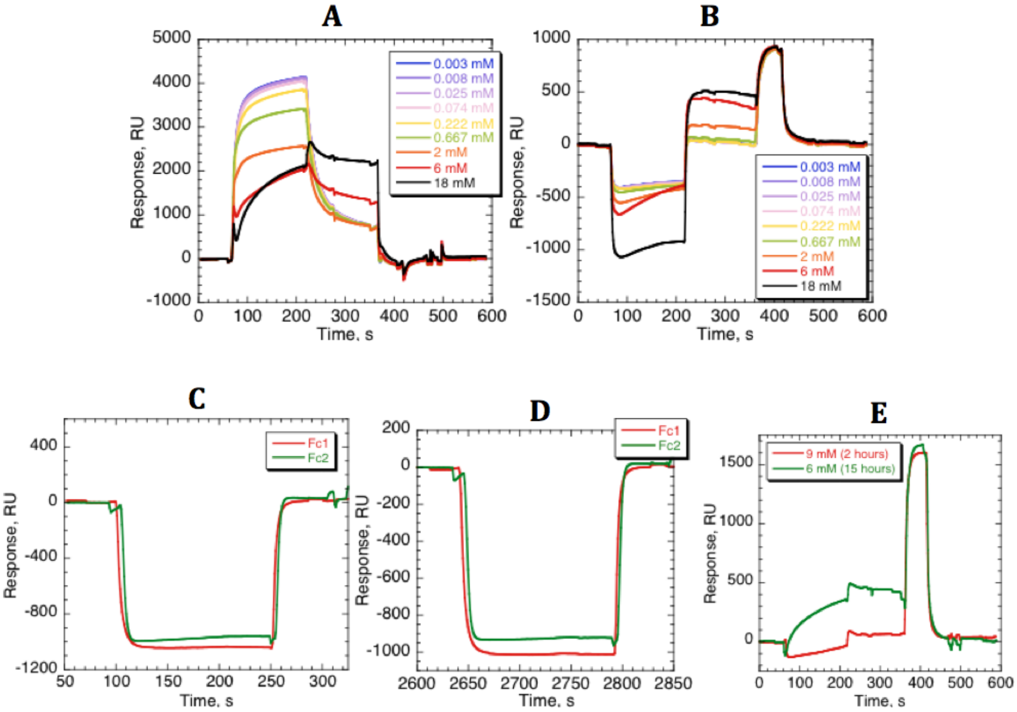


**Figure 9.10.:** Sensorgrams illustrating the effect of compound ZL2(L) for DC-SIGN ECD binding to CM-dextran and ManBSA surfaces.

**A**, Reference surface corrected sensorgrams of injections of ZL2(L) (at concentrations as shown in figure legend) in presence of 20 µM DC-SIGN ECD, over ManBSA surface (5300 RU immobilized). **B**, Raw sensorgrams of same injections as in panel A, over a reference surface (activated/deactivated CM-dextran). **C**, Raw sensorgrams showing injection of 10 mM ZL2(L) alone over reference (Fc1) and ManBSA (Fc2) surfaces. **D**, Raw sensorgrams showing injection of 5 mM ZL1(D) alone over reference (Fc1) and ManBSA (Fc2) surfaces. **E**, Raw sensorgrams showing 9 mM (red) and 6 mM (green) ZL1(D) with 20 µM DC-SIGN ECD injected over a reference surface; samples with 9 mM and 6 mM ZL1(D) were incubated at room temperature for 2 and 15 hours, respectively.

Moreover, the sensorgrams deviated from a usual shape, which has a characteristic rapid association phase of DC-SIGN ECD binding to ManBSA surface with binding plateau reached immediately after the injection is started, and a rapid dissociation phase (fig. 9.2 on page 149A). In the presence of higher concentrations of ZL2(L) this behavior was distorted: the association does not reach plateau and binding continues, while dissociation phase is markedly stabilized (fig. 9.10A). However, the interaction was Ca$^{2+}$-dependant (EDTA was capable to remove the bound material). In terms of binding responses, about 850 RU bound to the reference surface after 18 mM ZL2(L)/20 µM DC-SIGN ECD mixture was injected. This high response is unlikely just for a disaccharide molecule. Furthermore, this effect was dependent on incubation time of ZL2(L)/DC-SIGN ECD mixture: the longer the incubation, the more visible the effect (fig. 9.10E). The reason of such behavior of compound ZL2(L) is not clear.

### 9.1.3. *Development of fucose-based glycomimetic compounds*

The research work on fucose-based glycomimetic DC-SIGN antagonists, conducted by our team in collaboration with Pr. A. Bernardi group, has been started few years before the beginning of my PhD. Based on the structural knowledge of Lewis$^X$ binding to DC-SIGN CRD (described in subsection 2.3.1 on page 67, fig. 2.9 on page 68D), two α-fucosylamides, ***2a*** and ***17*** (fig. 9.11), were designed, which conserved the L-fucose anchor for directing them to CRD of DC-SIGN, but GlcNAc and D-galactose moieties of Lewis$^X$ were changed in order to improve binding affinity to DC-SIGN. These compounds were evaluated as DC-SIGN inhibitors, and indeed showed slightly better affinity to DC-SIGN than Lewis$^X$ trisaccharide [145]. However, the results suggested that this type of compounds could be further improved.

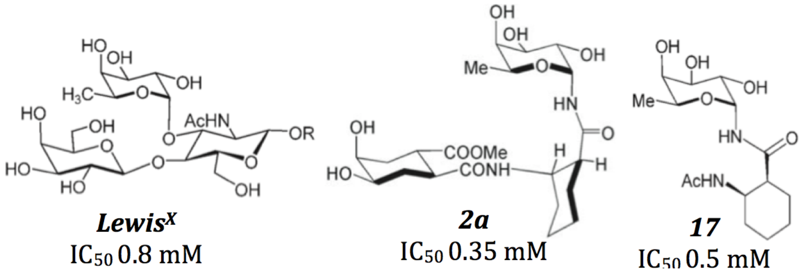


**Figure 9.11.:** The structures of previously designed fucose-based compounds ***2a*** and ***17***.

Keeping the same concept, several series of α-fucosylamides (fig. 9.12), all of which possess a β-amino acid tether and a fucosylamide anchor (general formula shown in figure 9.12A), were generated by a PhD student Daniela Doknic and tested in SPR competition assay, and the results are outlined in the following article.

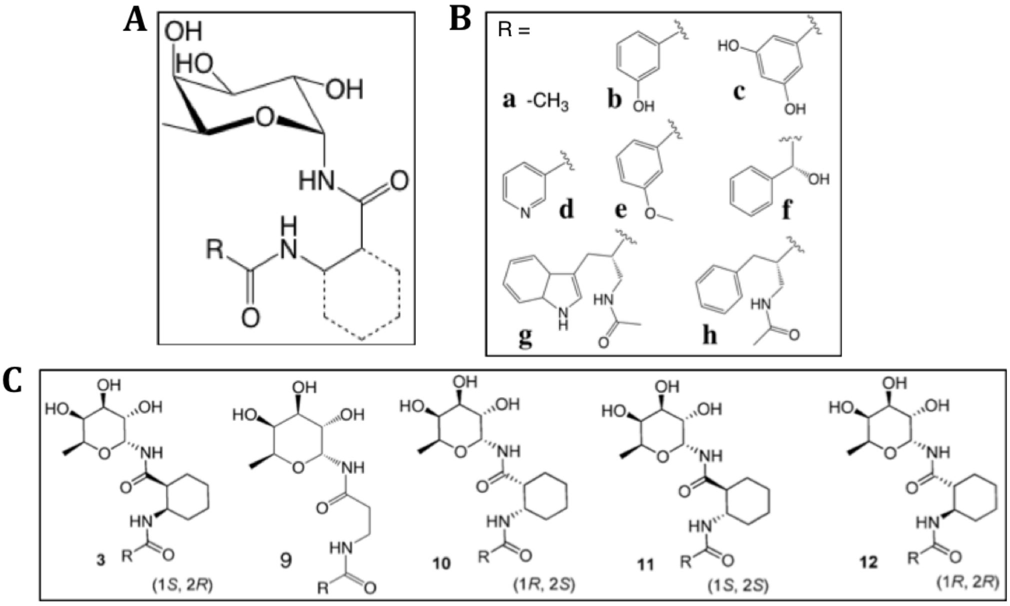


**Figure 9.12.:** The structures of newly synthesized fucose-based compounds.
**A**, The general formula of all compounds. **B**, The substituents, *R*, of the compounds. **C**, The general formulas of different series. All the structures and their names are the same as in article **n°1** (the individual compound names are derived writing series number followed by the letter of a relevant substituent).

Additionally to the evaluation of the potency of these compounds to inhibit DC-SIGN binding to immobilized ManBSA, they were also tested against langerin, a C-type lectin expressed by Langerhans cells and demonstrated to have protective role against HIV infection (see sub-subsection 2.2.1.2 on page 64). All of the compounds appeared to have low affinity to langerin, a desirable result in the development of selective DC-SIGN antagonists.

The compound **10b** was identified as the most active and selective ligand in these series.

The main outcome of these studies:

- Most of the Lewis$^{X}$ trisaccharide mimicking compounds have a similar activity towards DC-SIGN as the latter sugar.
- For the first time the selectivity to DC-SIGN *vs* langerin was addressed; as could be expected, fucose-based compounds had low affinity to langerin, thus good selectivity to DC-SIGN.

Contributions:

All of the compounds were designed and synthesized by the group of Pr. A. Bernardi. The described modeling and docking studies were conducted by Pr. A. Bernardi group in collaboration with Dr. J. Weiser team. Langerin used in SPR assays were produced in our team.

My contribution to this study:

I have prepared DC-SIGN required for SPR experiments, set up the competition assay with langerin and performed all of the described SPR assays, analyzed the data, and participated in the preparation of the paper manuscript.

# 1

# 1

***Paper n$^{o}$1:*** **Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity *versus* Langerin**

Organic &
Biomolecular
Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 5778

www.rsc.org/obc

PAPER

# Second generation of fucose-based DC-SIGN ligands : affinity improvement and specificity *versus* Langerin†

**Manuel Andreini,‡[a] Daniela Doknic,‡[a] Ieva Sutkeviciute,‡[b,e] José J. Reina,[a] Janxin Duan,[c] Eric Chabrol,[b,e] Michel Thepaut,[b,e,f] Elisabetta Moroni,[a] Fabio Doro,[a] Laura Belvisi,[a] Joerg Weiser,[c] Javier Rojo,[d] Franck Fieschi*[b,e,g] and Anna Bernardi*[a]**

***Received 11th April 2011, Accepted 20th May 2011***
**DOI: 10.1039/c1ob05573a**

DC-SIGN and Langerin are two C-type lectins involved in the initial steps of HIV infections: the former acts as a viral attachment factor and facilitates viral invasion of the immune system, the latter has a protective effect. Potential antiviral compounds targeted against DC-SIGN were synthesized using a common fucosylamide anchor. Their DC-SIGN affinity was tested by SPR and found to be similar to that of the natural ligand Lewis-X ($Le^X$). The compounds were also found to be selective for DC-SIGN and to interact only weakly with Langerin. These molecules are potentially useful therapeutic tools against sexually transmitted HIV infection.

## Introduction

Dendritic Cells (DCs) are instrumental in the development of pathogen-specific immune responses.[1] DCs are professional antigen-presenting cells that capture microbes entering skin or mucosal tissues and process them to form MHC-peptide complexes. After antigen uptake, immature DCs acquire the capacity to migrate to lymph nodes where they present processed antigens to T-cells, initiating adaptive immune responses. DCs express a repertoire of pathogen-recognition receptors (PRRs), including Toll Like Receptors (TLRs) and C-type lectins that mediate both signaling by self antigens and, in some cases, pathogen recognition.[2] C-type lectins represent a large family of $Ca^{2+}$ dependent lectins and recognize pathogen-derived carbohydrate structures. Many different C-type lectins expressed by DCs have been described,[3] including DC-SIGN.

DC-SIGN (DC-Specific ICAM-3 Grabbing Nonintegrin; CD 209) was originally defined as an intercellular adhesion molecule-3 (ICAM-3) receptor that plays an important role in establishing the first contact between DC and resting T cells.[4] It is a type II transmembrane C-type lectin with a single C-terminal Carbohydrate Recognition Domain (CRD) within its sequence. In the cellular membrane, DC-SIGN is assembled as a tetramer, thanks to an extended coiled-coil region that allows simultaneous presentation of four CRDs.[5] This oligomerization influences the lectin avidity in binding events. DC-SIGN appears to promote dissemination of a number of viruses (*e.g.*, HIV, hepatitis C virus, Ebola virus)[6] and to participate in suppressing immune responses to some pathogens, (*e.g.*, *Mycobacterium tuberculosis* and *Helicobacter pylori*).[7]

The various roles attributed to DC-SIGN have generated much interest towards the identification of ligands that can be used to explore its different functions and/or to inhibit pathogen binding.[8] However, generation of specific ligands for DC-SIGN is a challenging task, since many other C-type lectins exist and share important structural features with DC-SIGN binding site. Among the list of C-type lectin receptors closely related to DC-SIGN, Langerin, which is also expressed at the cell surface of antigen presenting cells, and L-SIGN, expressed on endothelial liver cell, placenta and lymph nodes, are particularly likely to interfere with DC-SIGN recognition.[9,10] These three lectins are all calcium-dependent carbohydrate-binding proteins and share the ability to bind high-mannose oligosaccharides. On the other hand, the three lectins show different specificity towards fucosylated oligosaccharides, a fact which may be used to design DC-SIGN specific ligands. Indeed, contrary to L-SIGN, DC-SIGN is known to bind the Lewis X ($Le^X$) epitope (Galβ4[Fucα3]GlcNAc, **1** in Fig. 1), as illustrated in the recognition mode of *Schistosoma mansoni* egg by these lectins.[5d,7,11] Moreover, both DC-SIGN and Langerin appear to recognize blood group B antigen through its fucose residue[12,13] in the primary $Ca^{2+}$ binding site, but again the $Le^X$ antigen is specific for DC-SIGN relative to Langerin. As

---

[a] *Universita' degli Studi di Milano, Dipartimento di Chimica Organica e Industriale and CISI, via Venezian 21, 20133 Milano, Italy. E-mail: anna.bernardi@unimi.it*
[b] *Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, 38027 Grenoble, France. E-mail: fieschi@ibs.fr*
[c] *Anterio consult&research, Augustaanlage 23, D-68165 Mannheim*
[d] *Glycosystems Laboratory, Instituto de Investigaciones Químicas, CSIC–Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla, Spain*
[e] *CNRS, UMR 5075, Grenoble, France*
[f] *CEA, Grenoble, France*
[g] *Institut Universitaire de France, 103 boulevard Saint-Michel, 75005 Paris, France*

† Electronic supplementary information (ESI) available: . See DOI: 10.1039/c1ob05573a

‡ These authors, listed in alphabetical order, contributed equally to this work

This journal is © The Royal Society of Chemistry 2011

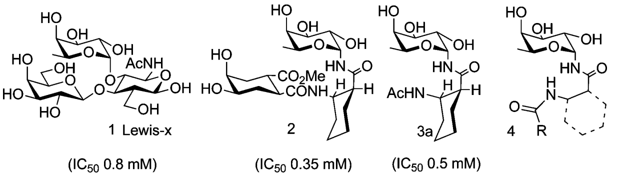


**Fig. 1** The known fucose-based DC-SIGN ligands **1–3a** and the general structure **4** of the library described in this paper. (IC$_{50}$ from ref. 14).

shown by X-ray structures,[12,13] in addition to binding the fucose residue on the Ca$^{2+}$ site, DC-SIGN is uniquely able to stabilize Le$^{X}$ galactose residue in a second binding area due to key residues that are absent in Langerin.

The existence of a secondary binding site is also suggested by a glycan array study of over 100 glycan structures.[12] This study demonstrated that the presence of a terminal fucose residue is not sufficient for DC-SIGN binding, but 14 fucose-bearing glycans with the structure of Lewis epitopes were found to bind selectively to DC-SIGN relative to L-SIGN.

We have recently described the first fucose-based artificial ligand of DC-SIGN (compound **2**, Fig. 1), designed to mimic this trisaccharide.[14] The ligand was built by using an α-fucosylamide anchor to drive the molecule to the DC-SIGN primary binding site and connecting it to a galactose mimic using a cyclic *cis*-β-amino acid ((1*S*,2*R*)-2-amino-cyclohexanecarboxylic acid, Fig. 1). The layout of these residues and more specifically the linker β-amino acid allowed the molecule to adopt a three-dimensional shape similar to the Le$^{X}$ trisaccharide (Fig. 1).[15] Amide bonds were chosen to connect the three elements of ligand **2** to achieve synthetic simplicity as well as chemical and metabolic stability of the target molecule. DC-SIGN binding studies performed by SPR showed that ligand **2** and surprisingly its simplified version **3a**, which does not contain the galactose-mimic moiety, inhibit DC-SIGN better than the natural ligand **1** (IC$_{50}$ = 0.35 mM, 0.5 mM and 0.8 mM, respectively).[14] The weak difference of affinity between **2** and **3a**, however, suggested that the galactose-like fragment in **2** gives a limited contribution to the binding interaction. Building on this knowledge, the goals of the work we report in this paper were: 1) to establish a minimal structure easily accessible in large scale and able to engage the receptor with an affinity similar to that of the natural ligand Le$^{X}$; 2) to improve the binding affinity of the fucosylamides by optimizing the interactions in the secondary binding site. To achieve these goals, a library of *ca.* 40 derivatives of general formula **4** was designed, synthesized and assayed by SPR to determine the ability of the compounds to inhibit DC-SIGN binding to immobilized mannosylated Bovine Serum Albumin (Man-BSA). Moreover, a preliminary selectivity screening was introduced to test some library members for inhibition of Langerin, using SPR. Selectivity for DC-SIGN *versus* Langerin is specially important to develop inhibitors of sexually transmitted HIV infections. As discussed above, interaction with DC-SIGN on mucosal DC is used by the virus to invade the host immune system. On the contrary, Langerin is suggested to have protective effects against HIV infection.[16] Indeed, some of the fucosylamides examined displayed an interesting DC-SIGN selectivity and have the potential of being developed as antiviral agents.

## Results and discussion

In order to select reasonable ligand candidates, the properties of the protein surface in the vicinity of the Ca$^{2+}$ site in the Lewis$^{X}$-DC-SIGN complex (1SL5)[12] were examined using GRID.[17] Both the DRY probe and the WATER probe were used to identify hydrophilic and lipophilic regions of the binding site. The molecular representations shown in Fig. 2 were obtained with the Maestro graphical interface. Various minima for the WATER probe were identified in the vicinity of the Ca$^{2+}$-binding region: in the crystal structure, they are occupied by crystallographic water molecules W13, W34 and W36 (Fig. 2a). W13 and W36 are located in two well-defined low interaction energy sites, both below −11 kcal mol$^{-1}$. W13 is in the vicinity of the fucose residue and mediates the interaction of Fuc-O2 with Glu354 and Lys368, and of Gal-O6 with Asp367 (Fig. 2a). W36 mediates the interaction of Gal-O4 with Glu358. The W36 site is occupied by a crystallographic water molecule also in 3 out of 4 known X-ray structures of DC-SIGN in complex with oligomannosides (1SL4,[12] 1K9I,[18] 2IT5[19]) and it is replaced by one sugar hydroxyl group in the fourth one (2IT6[19]). The W34 site belongs to a larger isoenergetic area with a less favorable GRID interaction energy (*ca.* −7 kcal mol$^{-1}$), occupied by water molecules in 2 out of 4 oligomannoside-DC-SIGN complexes (1SL4, 1K9I) and loosely

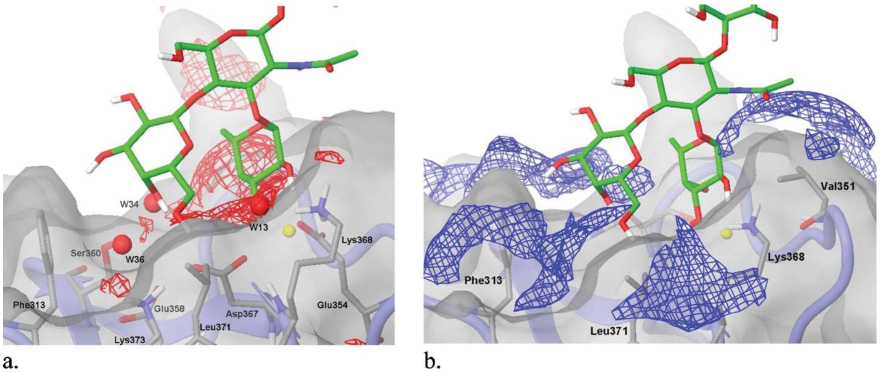
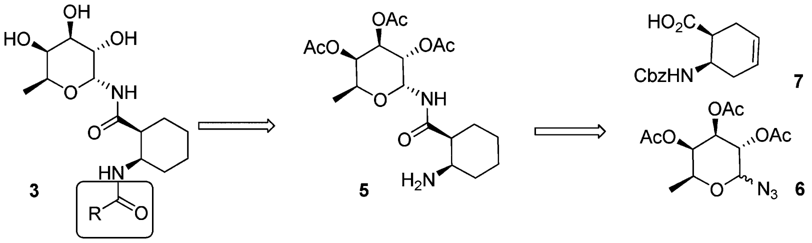


**Fig. 2** Energetic maps of GRID interactions (the Ca$^{2+}$ ion is shown in yellow) in the structure of the DC-SIGN- Le$^{X}$ complex (pdb code : 1SL5, from ref. 11). a) WATER probe, −7.0 kcal mol$^{-1}$ isosurface, showing the binding sites for crystallographic water molecules W13, W34 and W36 in 1SL5. b) DRY probe, −0.5 kcal mol$^{-1}$ isosurface, showing hydrophobic areas near the binding region and the groove formed by Phe313 and Leu 371.



**3** **5** **7** **6**

**Scheme 1** Retrosynthetic analysis of fucosyl derivative **3**.

replaced by mannose hydroxyl groups in the other 2 (2IT5 and 2IT6). In the 1SL5 structure, W34 mediates binding of Gal-O4 to Ser360 and, together with W36, contributes to the creation of a secondary binding site involving Leu371, Asp367, Lys373, Glu358 and flanked by Phe313 (Fig. 2a).

GRID analysis with the DRY probe allowed identification of the hydrophobic areas near the binding region, which are shown in Fig. 2b. Two of them, formed by Val351 and Asn362/Asn344, are in the immediate vicinity of the $Ca^{2+}$-binding site and establish Van der Waals contact with the ligand. Phe313 and Leu371/Lys368 side chains form a major hydrophobic groove which flank the W36 crystallographic site.

Docking of mimic **2** in the 1SL5 structure was obtained using Glide.[20] The complex, which included protein, ligand and the two water molecules W13 and W36, was prepared with the standard Preparation Wizard routine of Glide, but the final minimization was performed in implicit (GB/SA[21]) water with the AMBER* force field. It was found that this procedure allowed achievement of a better orientation of the water molecules and optimization of their hydrogen bonding pattern, which in turn avoided steric clashes in the following re-docking step. Docking obtained with this model suggested that optimal interaction is reached with the ligand in an extended conformation, which would allow the galactose mimic to place two hydroxyl groups in a hydrophilic patch of the protein near the side chain of Phe313 while nesting the cyclohexane ring in the groove formed by Phe313 and Leu371 (Fig. 3).

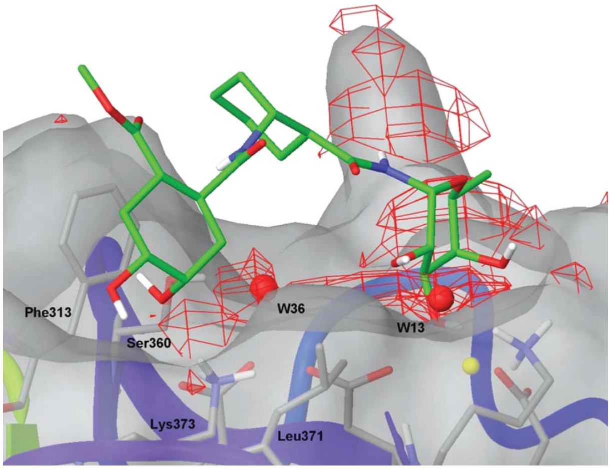
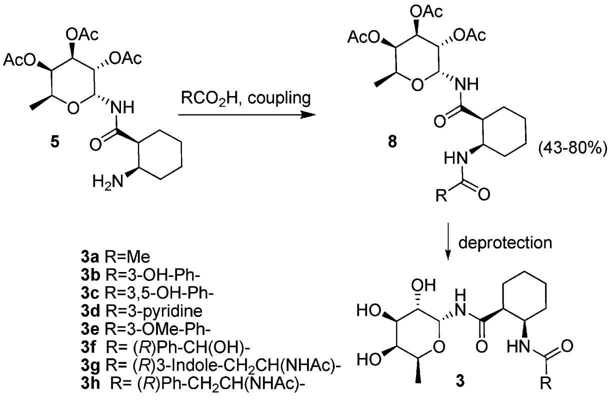


**Fig. 3** Docking of mimic **2** in 1SL5. (the $Ca^{2+}$ ion is shown in yellow).

Based on this model, interesting candidates to replace the galactose moiety should be molecules able to take advantage of lipophilic interactions and to interact specifically with the secondary hydrophilic regions. Further docking experiments suggested that favorable interactions could also be established by positively charged groups in the ligands and the negatively charged regions of the protein created by Asp 367 and Glu 358 side chains. Following this analysis, candidates for the R group in **4** were selected among commercially available carboxylic acids featuring aromatic groups and/or hydroxyl groups, amino groups or acetamides.

The initial set of compounds were synthesized starting from amine **5**,[14] which in turn was obtained from tri-*O*-acetyl-L-fucosylazide **6** and the protected (1*S*,2*R*)-β-amino acid **7**, as we have previously described (Scheme 1).[14]

The coupling reactions between amine **5** and the acid partners ($RCO_2H$, Scheme 2) were performed using either HBTU, acid chloride or EDC/HOBt activation, as described in the Supplementary Information. These conditions afforded the protected ligands **8** in variable yields (between 43 and 80%) after isolation by solid phase extraction and chromatographic purification. Removal of the protecting groups under standard Zemplen's conditions gave the required compounds **3**. If the acid partner carried a Boc protection, this was first removed using a mixture of $TFA/CH_2Cl_2$ (1/5). In this way *ca.* 30 different compounds were prepared (see Supplementary Information for the structure and characterization of the entire library). Scheme 2 shows the structure of those that will be used in the following discussion.



**Scheme 2** Coupling reaction of amine **5** with acid partners.

DC-SIGN affinity for the entire set of 30 compounds of general formula **3** was estimated using a surface plasmon resonance (SPR) biosensor in a competition assay which we have previously described.[14] The assay allows an affinity evaluation of all ligands relative to one another on the basis of their percentage inhibition

This journal is © The Royal Society of Chemistry 2011

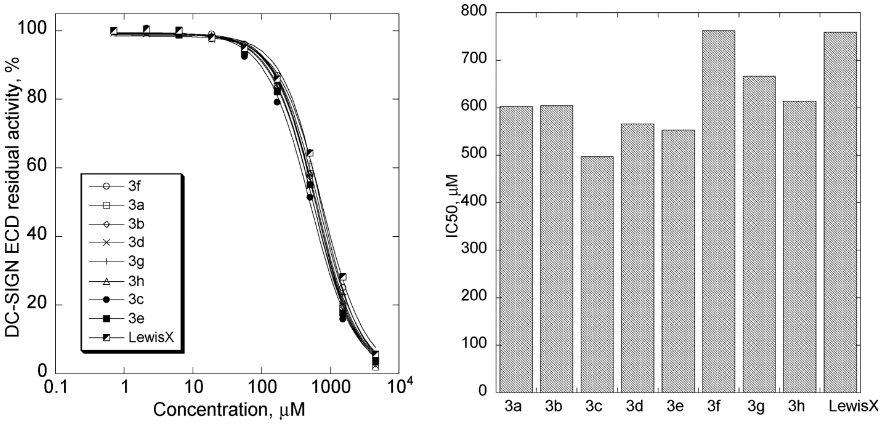


**Fig. 4** The dependency of DC-SIGN ECD percent activity on concentrations of corresponding compounds. (see original sensorgrams in SI-Fig.4).

of DC-SIGN binding to immobilized mannosylated bovine serum albumin (Man-BSA). The commercially available Man-BSA used in these assays contain an average of 12 glycosylation sites displaying the Manα1-3[Manα1-6]Man branched trisaccharide. Man-BSA was covalently attached to a carboxymethyl dextran-functionalized gold SPR sensor chip CM4. Inhibition studies were then performed using extracellular domain (ECD) of DC-SIGN (20 μM) injected alone or in the presence of a constant concentration (300 μM) of the ligands. At this concentration, for the particular chip used in the assay, $Le^X$ exhibited 25% of inhibition. All the molecules **3** synthesized showed a similar efficiency, independent of the nature of the R group, and none improved significantly over the activity of **3a**. To confirm these data, complete inhibition curves were obtained and $IC_{50}$ values were estimated for the selected group of compounds **3a–h** shown in Scheme 2. The results were totally consistent with the previous observation (Fig. 4).

Interestingly, when an analogous group of compounds **9** (Fig. 5), obtained by reaction of fucosylazide **6** with β-alanine rather than with **7**, was examined in SPR single point assays at 300 μM concentration, similar inhibition values (25–30%) were obtained. Thus, these simple α-fucosyl-β-alanyl amides showed a similar affinity for DC-SIGN as $Le^X$ and all the compounds **3** synthesized.

**9a** R=Me
**b** R=3-OH-Ph-
**c** R=3,5-OH-Ph-
**g** R=(*R*)-3-Indole-CH$_2$CH(NHAc)-
**h** R=(*R*)-Ph-CH$_2$CH(NHAc)-

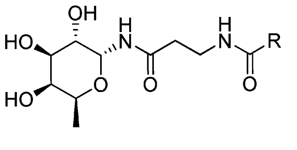


**Fig. 5** β-Alanine derivatives **9**.

These observations, that are in striking contrast with the expectations derived from docking studies, strongly suggest that the R substituent in **3** is not reaching the secondary binding site identified by the docking algorithm and may not be interacting at all with the protein. The unexpected results obtained with α-fucosyl-β-alanyl amides **9**, whilst providing us with very simple ligands of high efficiency,[22] confirm that the (1*S*,2*R*)-2-amino-cyclohexanecarboxylic acid scaffold selected for the synthesis of **3** does not enforce optimal interaction of the secondary residue with the protein. NMR studies are currently in progress to assess the structural details of the interaction of DC-SIGN with compounds **3**.[23] To further explore the role of the β-amino acid structure in defining ligand–protein interaction, the configuration of the scaffold was varied systematically and a third set of compounds was synthesized, where the R fragment was kept unchanged and the β-configuration was systematically permutated.

The new set of compounds **10–12** (Fig. 6) were synthesized starting from β-amino acids **13–15** (Fig. 6) using the synthetic sequence employed for **3** and shown in Schemes 1 and 2. The syntheses of the enantiomerically pure isomeric amino acids **13–15** and of the corresponding fucosylamide derivatives **10–12** are described in the Supplementary Information. The $IC_{50}$ obtained by SPR analysis of **10–12** are collected in Fig. 7 and compared to selected data obtained for **3**, for $Le^X$ and for L-fucose (see SI-Fig 5 for original sensorgrams).

The $IC_{50}$ values obtained for **3a** and **3b** are consistent with previous measurements (Fig. 4). The data confirm that the activity of most fucosylamides is close to that of $Le^X$. The series of compounds **10a–d** show a larger increase of the affinity (a factor of 3) on passing from the acetamide **10a** (R = Me) to the aromatic amides **10b–d**, suggesting a possible role of the aromatic group in the interaction with the protein. This series, which is built on the (1*R*,2*S*)-2-amino-cyclohexanecarboxylic acid scaffold **13**, also contains the strongest ligands so far, the hydroxybenzoic acid derivatives **10b** and **10c** ($IC_{50}$ 470 μM) and is therefore the best current candidate for further optimization.

To analyze the selectivity properties of the compounds screened, we developed an additional SPR analysis for Langerin binding properties. As for DC-SIGN, the ability of Langerin to bind to a surface functionalized with Man-BSA was tested. In the case of Langerin, binding to Man-BSA as well as to the dextran matrix (SI-Fig 1) was observed. Therefore, the dextran/Man-BSA surface was considered as a combined ligand of Langerin ECD. Upon titration of the surface with Langerin, a similar saturation curve than for DC-SIGN was obtained (SI-Fig. 2 and 3). Langerin displayed an apparent $K_d$ for this surface of 10.3 μM. Indeed, the two lectins exhibited comparable affinity for this surface, thus the same fixed concentration of Langerin was used in the SPR-based competition assay, which allowed a direct comparison of the binding inhibition properties of the compounds for Langerin *versus* DC-SIGN.

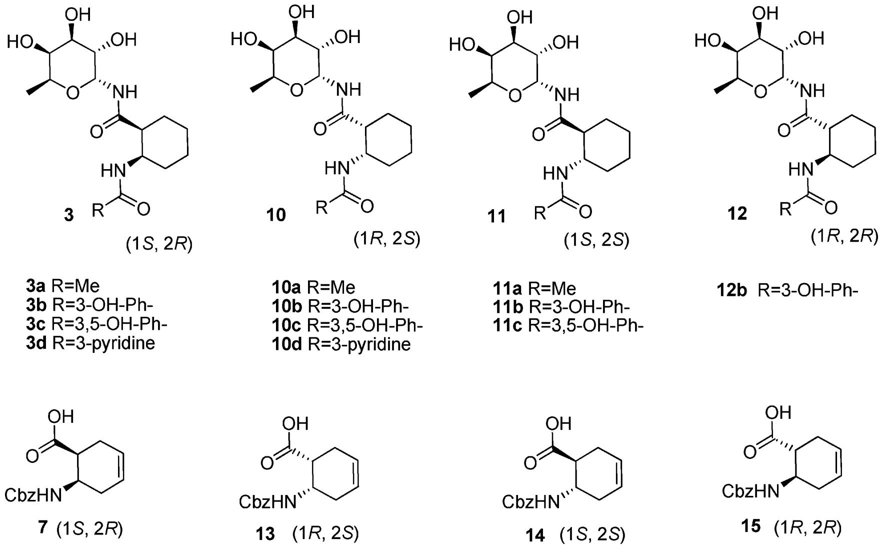


**Fig. 6** Library of compounds with different stereochemistry in the cyclohexane scaffold.

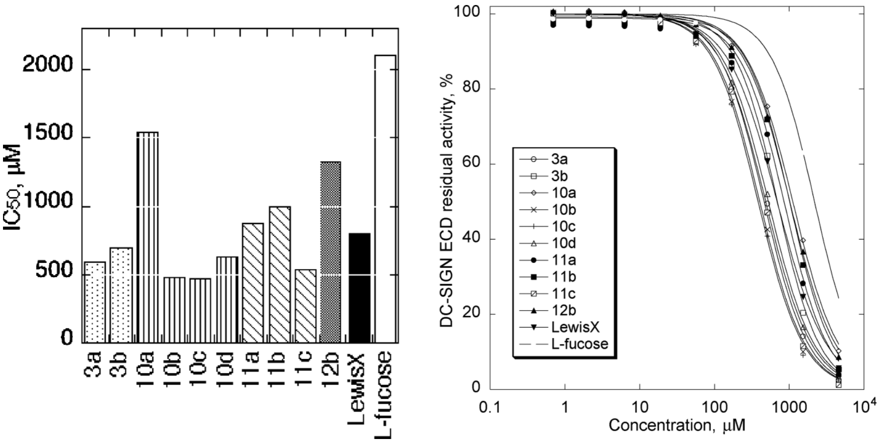


**Fig. 7** The $IC_{50}$ values obtained for compounds **10–12** by SPR inhibition assay. $IC_{50}$ of L-fucose, $Le^X$, **3a** and **3b** measured in the same conditions are shown for comparison.

Compounds initially tested for their DC-SIGN inhibitory potency (Fig. 7) were evaluated with Langerin ECD. The results are shown in Fig. 8A (see also SI-Fig.6). The inhibitory potency of the fucosylated mimics is so low against Langerin that it was not possible to determine an $IC_{50}$. A crude comparison of the inhibitors' properties towards DC-SIGN and Langerin was obtained by comparing the residual activity of both lectins at the highest concentration tested for each compound (Fig. 8B).

The data show that many of the fucosylamides tested display a larger DC-SIGN selectivity than $Le^X$ and confirm **10b** as one of the most interesting elements of this group, both for its DC-SIGN affinity and for its specificity.

## Conclusions

In this paper, we have presented a new library of fucose-based ligands of DC-SIGN, all characterized by the presence of a β-amino acid tether and of a fucosylamide anchor, able to direct the molecules to the DC-SIGN CRD binding site. We were able to identify many compounds that, compared to the natural ligand $Le^X$ and its previously reported mimic **2**, display a similar DC-SIGN inhibition efficiency at a fraction of the synthetic cost. In particular, α-fucosylamides **9**, derived from β-alanine, are interesting candidates for polyvalent presentations[8d–k] due to their high synthetic accessibility and good ligand efficiency.[22] A second group of compounds, α-fucosylamides **10** derived from (1*R*,2*S*)-2-amino-cyclohexanecarboxylic acid **13**, were also of interest because they yielded the most active and selective ligand of this group (**10b**). Indeed, it may be important that molecules directed to block the action of DC-SIGN do not interfere with the action of other lectins. We have recently shown that α-*N*-fucosylamides of general formula **3** and **9** interact strongly with the L-fucose binding lectin PA-IIL of *Pseudomonas aeruginosa*.[24] In particular, to develop inhibitors of sexually transmitted HIV infections it may be necessary to select DC-SIGN antagonists that do not interfere with the action of Langerin. Like DC-SIGN, Langerin is a

This journal is © The Royal Society of Chemistry 2011

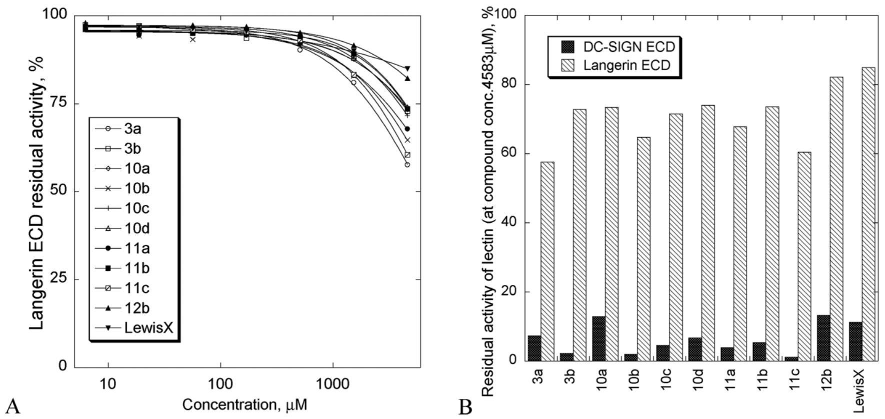


**Fig. 8** A) Inhibition of Langerin ECD binding to Man-BSA immobilized on the dextran surface. B) Residual lectin activity at the highest tested concentration of the ligands (4.6 mM). Langerin, white bars; DC-SIGN, black bars.

membrane C-type lectin known to bind to HIV-1. However, whilst interaction with DC-SIGN is used by the virus to invade the host immune system, Langerin is suggested to have protective effects against HIV infection.[16,25] Indeed, Langerhans cells, which are the first dendritic cells to encounter HIV *via* genital mucosa, have been described as a natural barrier for HIV-1 transmission, which is dependent on Langerin expression.[15] The importance of Langerin in HIV protection has been again emphasized in the context of HIV-1/Herpes Simplex Virus type II (HSV-2) co-infection.[26] In this last case, it has been demonstrated that HIV susceptibility of Langerhans cells, and the subsequent virus transmission, could be promoted by HSV-2-dependent abrogation of Langerin's functions. Conversely, DC-SIGN has a well-established role in dendritic cells-mediated HIV-1 transmission.[6] Thus, the selectivity of the ligands for DC-SIGN relative to Langerin was also tested. Indeed, we described here for the first time, simultaneous screening of artificial compounds towards both lectins. In agreement with literature data on the natural ligands of these lectins, a low inhibitory potency of Le$^X$ towards Langerin has been observed in contrast to DC-SIGN. The capacity of Le$^X$ and of the other fucosylated derivatives to inhibit Langerin binding is so low that we could not perform a full inhibition curve in a reasonable range of concentration. Therefore, selectivity of the fucose-based ligands tested was assessed by comparing residual lectin activity at the highest ligand concentration tested. Most of the α-fucosylamides assayed were found to be more DC-SIGN selective than the natural ligand Le$^X$ and therefore they are more likely to be turned into therapeutically useful tools against sexually transmitted HIV infection.[8]

## Experimental

### Langerin and DC-SIGN ECD expression and purification

Langerin ECD constructs (comprising residue 68–328) have been overexpressed using a pET30b derived vector in BL21(DE3) as described previously.[27] The protein was expressed as inclusion body, refolded and purified to homogeneity in a functional form as already described.[22] DC-SIGN ECD protein (residue 66–404) has been overexpressed and purified as described previously.[5a]

### Surface plasmon resonance analysis

All experiments were performed on a Biacore 3000 using functionalized CM4 sensor chips and the corresponding reagents from Biacore. Two flow cells were activated as previously described.[28] Flow cell one was then blocked with 30 µL of 1 M ethanolamine and used as a control surface. The second one was treated with BSA-Manα1-3[Manα1-6]Man (Man-BSA, Dextra) (60 µg mL$^{-1}$) in 10 mM acetate buffer, pH 4. Remaining activated groups were blocked with 30 µL of 1 M ethanolamine. The final density immobilized on the surface of the second flow cell was 2000 RU. The Man-BSA used to functionalize CM4 chip harbours 12 glycosylation sites according to manufacturer.

Two types of SPR tests were set up for the evaluation of glycomimic compounds. The single point inhibition assay was used for fast screening of compound selectivity. Here, either DC-SIGN or Langerin at concentration of 20 µM were incubated with corresponding compounds (300 µM final concentration) and 20 µl of the samples were co-injected over Man-BSA surface. The lectin steady state binding responses were extracted from the sensorgrams and compared with the responses of compound-free lectin injections and converted to inhibition percent values. In the case of Langerin inhibition assay, even with parallel functionalization of the reference surface with non-glycosylated BSA, some interaction of Langerin with the dextran matrix still remains. Indeed, the dextran/Man-BSA surface has been considered as a combined ligand of Langerin ECD (see Supplementary Information for more details).

The second type of SPR test was used to estimate the relative compound affinity to the lectins on the basis of their IC50 values. This was accomplished in the same manner as in the single point inhibition assay, except that both lectins were incubated with increasing concentrations (from 0 to 5000 µM) of the corresponding compounds, the injected sample volumes were 13 µl, and the resulting binding responses were converted to residual lectin activity values, which were plotted against concentration values of the compounds. The relative IC50 values for each compound were determined by fitting four parameter logistic model (eqn (1)) to the experimental data.

$$y = \mathrm{bot} + \frac{\mathrm{top} - \mathrm{bot}}{1 + \left(\frac{x}{\mathrm{IC}_{50}}\right)^{\mathrm{slope}}} \tag{1}$$

where $y$ is the percent activity, $x$ is the corresponding concentration, bot and top are the lowest and the highest values of percent activity, respectively.

In both type of experiments 5 μl flow rate was used and the running buffer was 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mM CaCl2, and 0.005% of P20 surfactant. All the samples were prepared in the running buffer.

The stability of the surface during a campaign was evaluated by DC-SIGN ECD binding capacity as a function of the number of cycles. The chip surfaces demonstrated a strong stability with negligible decrease of the binding capacity of only 0.06% per cycle (see data in the Supplementary Information).

## Modeling

**Grid analysis.** The properties of the protein surface in the vicinity of the $Ca^{2+}$ site were determined using the protein crystal structure derived from the complex DC-SIGN-$Le^X$ (pdb code 1SL5[12]) and the program GRID[17] (version 22). In particular the DRY probe and the WATER probe were used to identify hydrophilic and lipophilic regions of the binding site. The accompanying program GREAT was used to check the crystal structure file and to prepare the file of coordinates in standard PDB format, which is used as an input for the program GRIN. This program, which prepares the input for the main program GRID, was used to automatically assign atom types and charges for every atom of the protein, using provided standard parameters. Calculations of the interaction energy between the probe and each atom of the protein were performed on a box (36.6 Å × 23.0 Å × 16.6 Å per side) centered on the protein, with a grid spacing of 0.2 Å (NPLA = 5) and its value was evaluated at each grid point. A dielectric constant of 80 was used to simulate a bulk aqueous phase, while a dielectric constant of 4 was assigned to the interior of the protein.

The output, which consists of an array of interaction energies, can be visualized as contour surfaces at appropriate energy levels together with the protein structure. Contours at negative energy levels delineate regions of attraction between probe and protein, whereas positive energy levels define the surface of the protein. Visual inspection of the contour surfaces superimposed on the active site of DC-SIGN enabled the identification of the most favored hydrophilic and lipophilic regions, facilitating the interpretation of protein–ligand interaction. Moreover interaction energy values between the WATER probe and the protein were used to identify the important structural water molecules out of all the crystallized water molecules found in the X-ray structure. An arbitrary cut-off of $-10.42$ kcal $mol^{-1}$ (the most negative energy value was $-14.97$ kcal $mol^{-1}$) was chosen to detect the most favorable hot spots for a water molecule. This cut off allowed us to identify two important structural water molecules, corresponding to W13 and W36 in the crystal structure of DC-SIGN-$Le^X$ complex (1SL5). The DC-SIGN crystal structure including these two water molecules was subsequently used for additional docking runs.

**Docking calculations.** Docking calculations were performed using the program GLIDE 4.5.[20] The initial setup for the receptor preparation before docking runs was performed using Schrödinger's 'Protein Preparation Wizard', starting from the X-ray crystal structure of the DC-SIGN-$Le^X$ complex. All crystallographic water molecules, except for W13 and W36, were deleted, bond orders assigned and hydrogen atoms added. The assignments of protonation states for basic and acidic residues were based on the optimization of hydrogen bonding patterns. The final minimization of the complex was not performed with the Preparation Wizard default, but the complex was minimized (500 steps, Truncated Newton Conjugate Gradient method) in implicit water (GB/SA[21] model) using MacroModel[29] with the AMBER* force field (dielectric constant 1, cut off extended, convergence on gradient with threshold of 0.05). The oxygen atoms of W13 and W36 as well as the $Ca^{2+}$ ion were anchored to the original position through a harmonic potential during minimization. At the end of the minimization, the root-mean-square deviation (RMSD) of all heavy atoms was within 0.34 Å of the crystallographic positions. Docking calculations were performed in Standard Precision mode with standard OPLS-AA(2001)[30] force field; non-planar conformations of amide bonds were penalized, Van der Waals radii were scaled by 0.80 and the partial charge cut off was fixed to 0.15. The shape and properties of the binding site were mapped onto grids with dimensions of 36 Å (enclosing box) and 14 Å (ligand diameter midpoint box), centered on the ligand in the X-ray structure of the DC-SIGN-$Le^X$ complex. Docking was constrained in the $Ca^{2+}$ binding site by specifying a reference core corresponding to the C1–C6 carbon and O3–O4–O5 oxygen atoms of the fucose residue of the reference ligand in the X-ray structure of the DC-SIGN-$Le^X$ complex: ligands that feature the same core moiety as the reference ligand are subject to the constraint. The RMSD tolerance for the position of the core was set to 3.5 Å. This parameter enforces the fucose moiety to be located within 3.5 Å of the fucose residue in the reference ligand: ligand poses that do not match this constraint are screened out. These constraints allow coordination of the $Ca^{2+}$ ion by the hydroxy groups OH-2 and OH-3 of fucose, as observed in several crystal structure of C-type lectins complexed with fucose-containing ligands. The RMSD tolerance of 3.5 Å was selected to allow docking poses to explore both possible binding modes of the vicinal diol. Indeed, crystallographic and NMR data on DC-SIGN show that the monosaccharide moiety of ligands can coordinate the $Ca^{2+}$ ion with the vicinal hydroxyl groups in two possible orientations, differing by a 180° rotation. The quality of this docking protocol was validated by re-docking the $Le^X$ ligand in the DC-SIGN-$Le^X$ complex, which yielded a ligand pose that could be superimposed with crystalline $Le^X$ with RMSD of 1.71 Å (0.31 Å for the fucose residue).

## Synthesis

Solvents were dried by standard procedures: dichloromethane, methanol, *N*,*N*-diisopropylethylamine and triethylamine were dried over calcium hydride, chloroform and pyridine were dried over activated molecular sieves. Reactions requiring anhydrous conditions were performed under nitrogen. $^{1}H$, $^{13}C$ and $^{31}P$-NMR spectra were recorded at 400 MHz on a Bruker AVANCE-400 instrument. Chemical shifts ($\delta$) for $^{1}H$ and $^{13}C$ spectra are

This journal is © The Royal Society of Chemistry 2011

expressed in ppm relative to internal $Me_4Si$ as standard. Signals are abbreviated as s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet. Mass spectra were obtained with a Bruker ion-trap Esquire 3000 apparatus (ESI ionization) or an Autospec Fission Instrument (FAB ionization). HRMS (FT-ICR, ESI) were obtained with an Apex II instrument. Thin layer chromatography (TLC) was carried out with pre-coated Merck $F_{254}$ silica gel plates. Flash chromatography (FC) was carried out with Macherey-Nagel silica gel 60 (230–400 mesh). The libraries were synthesized through the common approach shown in Scheme 1.[14] All synthetic schemes and procedures, the synthesis and characterization of new intermediates, the synthesis and characterization of the full library of ligands (**3**, **9**, **11** and **12**) as well as their $^1$H- and $^{13}$C-NMR spectra are collected in the Supplementary Information. Below we report the characterization of compounds **10a–d** (from (1*R*,2*S*)-2-aminocyclohexanecarboxylic acid). Compounds **3a** and **5** were previously described.[14]

### *N*-((1*R*,2*S*)-2-Acetamido-cyclohexanecarboxyl)-α-L-fucopyranosylamine (10a)

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was used in the general acetylation method (see Supplementary Information). The product was purified by flash chromatography on silica gel (AcOEt, $R_f$ 0.29). Yield: 27 mg (82%). Zemplen deprotection and flash chromatography (85 : 15 $CHCl_3$ : MeOH, $R_f$ 0.18) afforded **10a**. Yield: 13 mg (65%). $^1$H-NMR (400 MHz, $CD_3OD$): $\delta$ (ppm) = 1.18 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.32–1.50 (m, 2H, $H_{Cy4ax}$ and $H_{Cy5ax}$), 1.55 –1.68 (m, 3H, $H_{Cy6ax}$, $H_{C3ax}$ and $H_{C4eq}$ or $H_{C5eq}$), 1.73–1.78 (m, 1H, $H_{Cy4eq}$ or $H_{Cy5eq}$), 1.83–1.92 (m, 1H, $H_{Cy3eq}$ and $H_{C6eq}$), 1.94 (s, 3H, Ac-Me), 2.71–2.75 (m, 1H, $H_{Cy1}$), 3.63–3.66 (m, 1H, $H_{F4}$), 3.74 (dd, $J_{3\text{-}4}$ = 3.3 Hz, $J_{3\text{-}2}$ = 10.3 Hz, 1H, $H_{F3}$), 3.79 (q, $J_{5\text{-}6}$ = 6.5 Hz, 1H, $H_{F5}$), 3.94 (dd, $J_{1\text{-}2}$ = 5.6 Hz, $J_{2\text{-}3}$ = 10.2 Hz, 1H, $H_{F2}$), 4.18–4.25 (m, 1H, $H_{Cy2}$), 5.51 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$). $^{13}$C-NMR (100 MHz, $CD_3OD$): $\delta$ (ppm) = 17.0 ($C_{F6}$), 23.0 (Ac-Me), 23.6 ($C_{Cy4}$ or $C_{Cy5}$), 24.2 ($C_{Cy4}$ or $C_{Cy5}$), 26.8 ($C_{Cy6}$), 30.7 ($C_{Cy3}$), 40.3 ($C_{Cy1}$), 46.4 ($C_{Cy2}$), 68.3 ($C_{F2}$), 68.8 ($C_{F5}$), 71.8 ($C_{F3}$), 73.3 ($C_{F4}$), 78.3 ($C_{F1}$), 173.1 (NHCO), 177.5 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{15}H_{26}N_2O_6$: 353.16831 [M + Na]$^+$; found: 353.16838. $[\alpha]_D$ −77.5 (*c* 0.35, EtOH)

### *N*-[(1*R*,2*S*)-2-(3-Hydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (10b)

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was coupled with 3-hydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (6 : 4 AcOEt : petroleum ether, $R_f$ 0.38). Yield: 13 mg (31%). Zemplen deprotection and flash chromatography (85 : 15 chloroform : methanol, $R_f$ 0.17) afforded **10b**. Yield: 8 mg (89%). $^1$H-NMR (400 MHz, $CD_3OD$): $\delta$ (ppm) = 1.18 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.42–1.55 (m, 2H, $H_{Cy4ax}$ and $H_{Cy5ax}$), 1.61–1.76 (m, 4H, $H_{Cy6ax}$, $H_{Cy3ax}$, $H_{Cy4eq}$ and $H_{Cy5eq}$), 1.94–2.02 (m, 1H, $H_{Cy6eq}$), 2.16–2.24 (m, 1H, $H_{Cy3eq}$), 2.86–2.91 (m, 1H, $H_{Cy1}$), 3.63–3.66 (m, 1H, $H_{F4}$), 3.74–3.80 (m, 2H, $H_{F5}$ and $H_{F3}$), 3.91–3.95 (m, 1H, $H_{F2}$), 4.28–4.33 (m, 1H, $H_{Cy2}$), 5.52–5.54 (m, 1H, $H_{F1}$) 6.90–6.95 (m, 1H, Ar), 7.18–7.28 (m, 3H, Ar). $^{13}$C-NMR (100 MHz, $CD_3OD$): $\delta$ (ppm) = 17.0 ($C_{F6}$), 23.8 ($C_{Cy4}$ or $C_{Cy5}$), 24.2 ($C_{Cy4}$ or $C_{Cy5}$), 27.8 ($C_{Cy6}$), 30.4 ($C_{Cy3}$), 46.2 ($C_{Cy1}$), 50.2 ($C_{Cy2}$), 68.2 ($C_{F2}$), 68.8 ($C_{F5}$), 71.7 ($C_{F3}$), 73.3 ($C_{F4}$), 78.4 ($C_{F1}$), 115.3 (Ar), 119.4 (Ar), 119.7 (Ar), 130.8 (Ar), 137.5 ($C_{quart.Ar.}$), 158.9 ($C_{quart.Ar.}$), 170.0 (NHCO), 178.0 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{20}H_{28}N_2O_7$: 431.17887 [M + Na]$^+$; found: 431.17948. $[\alpha]_D$ −53.1 (*c* 0.25, MeOH)

### *N*-[(1*R*,2*S*)-2-(3,5-Dihydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (10c)

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was coupled with 3,5-dihydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (7 : 3 ethyl acetate : n-hexane, $R_f$ 0.43). Yield: 17 mg (37%). Zemplen deprotection and flash chromatography (85 : 15 chloroform : methanol, $R_f$ 0.14) afforded **10c**. Yield: 10 mg (91%). $^1$H-NMR (400 MHz, $CD_3OD$): $\delta$ (ppm) = 1.19 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.42–1.55 (m, 2H, $H_{Cy4ax}$ and $H_{Cy5ax}$), 1.61–1.73 (m, 4H, $H_{C6ax}$, $H_{Cy3ax}$, $H_{Cy4eq}$ and $H_{Cyeq}$), 1.92–2.02 (m, 1H, $H_{Cy6eq}$), 2.16–2.22 (m, 1H, $H_{Cy3eq}$), 2.85–2.90 (m, 1H, $H_{Cy1}$), 3.64–3.66 (m, 1H, $H_{F4}$), 3.77 (dd, $J_{3\text{-}4}$ = 3.3 Hz, $J_{2\text{-}3}$ = 10.3 Hz, 1H, $H_{F3}$), 3.79 (q, $J_{5\text{-}6}$ = 6.5 Hz, 1H, $H_{F5}$), 3.95 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $J_{2\text{-}3}$ = 10.3 Hz, $H_{F2}$), 4.25–4.30 (m, 1H, $H_{Cy2}$), 5.55 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$) 6.39 (bs, 1H, Ar), 6.68 (bs, 2H, Ar). $^{13}$C-NMR (100 MHz, $CD_3OD$): $\delta$ (ppm) = 17.0 ($C_{F6}$), 23.8 ($C_{Cy4}$ or $C_{Cy5}$), 24.1 ($C_{Cy4}$ or $C_{Cy5}$), 27.9 ($C_{Cy6}$), 30.4 ($C_{Cy3}$), 46.2 ($C_{Cy1}$), 50.2 ($C_{Cy2}$), 68.2 ($C_{F2}$), 68.8 ($C_{F3}$ or $C_{F5}$), 71.7 ($C_{F3}$ or $C_{F5}$), 73.3 ($C_{F4}$), 78.4 ($C_{F1}$), 106.8 (Ar), 106.8 (Ar), 138.1 (Cquart., Ar.), 160.0 (Cquart., Ar.), 170.0 (NHCO), 178.0 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{20}H_{28}N_2O_8$: 447.17379 [M + Na]$^+$; found: 447.17407. $[\alpha]_D$ −65.7 (*c* 0.20, MeOH)

### *N*-[(1*R*,2*S*)-2-(3-Pyridinecarboxamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (10d)

The crude hydrogenation product of **28** (see Supplementary Information–SI-Scheme 3) was coupled with nicotinic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (97 : 3 ethyl acetate : triethyl amine, $R_f$ 0.17). Yield: 19 mg (50%) Zemplen deprotection and flash chromatography (85 : 15 chloroform : methanol, $R_f$ 0.17) afforded **10d**. Yield: 6 mg (43%). $^1$H-NMR (400 MHz, $CD_3OD$): $\delta$ (ppm) = 1.19 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.43–1.55 (m, 2H, $H_{Cy4ax}$ and $H_{Cy5ax}$), 1.67–1.78 (m, 4H, $H_{Cy6ax}$, $H_{Cy3ax}$, $H_{Cy4eq}$ and $H_{C5eq}$), 1.94–2.04 (m, 1H, $H_{Cy3eq}$), 2.14–2.21 (m, 1H, $H_{Cy6eq}$), 2.88–2.92 (m, 1H, $H_{Cy1}$), 3.63–3.66 (m, 1H, $H_{F4}$), 3.74 (dd, $J_{3\text{-}4}$ = 3.4 Hz, $J_{2\text{-}3}$ = 10.3 Hz, 1H, $H_{F3}$), 3.78 (q, $J_{5\text{-}6}$ = 6.5 Hz, 1H, $H_{F5}$), 3.93 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $J_{2\text{-}3}$ = 10.2 Hz, $H_{F2}$), 4.37–4.42 (m, 1H, $H_{Cy2}$), 5.53 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 7.50–7.53 (m, 1H, $H_{Ar5}$), 8.20 (d, 1H, $J_{Ar5}$-$J_{Ar6}$ = 8.0 Hz, $H_{Ar6}$), 8.65 (d, 1H, $J_{Ar4}$-$J_{Ar5}$ = 4.5 Hz, $H_{Ar4}$), 8.92 (s, 1H, $H_{Ar2}$). $^{13}$C-NMR (100 MHz, $CD_3OD$): $\delta$ (ppm) = 17.0 ($C_{F6}$), 23.7 ($C_{Cy4}$ and $C_{Cy5}$), 24.1 ($C_{Cy3}$), 27.4 ($C_{Cy6}$), 30.4 ($C_{Cy1}$), 50.4 ($C_{Cy2}$), 68.2 ($C_{F2}$), 68.8 ($C_{F5}$), 71.7 ($C_{F3}$), 73.3 ($C_{F4}$), 78.3 ($C_{F1}$), 125.1 ($C_{Ar5}$), 132.7 ($C_{quart.Ar}$), 137.4 ($C_{Ar6}$), 149.5 ($C_{Ar2}$), 152.5 ($C_{Ar4}$), 167.95 (NHCO), 177.76 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{19}H_{27}N_3O_6$: 416.17921 [M + Na]$^+$; found: 416.17934. $[\alpha]_D$ −81.9 (*c* 0.15, MeOH)

## Acknowledgements

This work was supported by the FIRB program CHEM-PROFARMANET (RBPR05NWWC), the Marie Curie ITN FP7 project CARMUSYS (PITN-GA-2008-213592) and Comune di Milano (Convenzione 55/2008). Exact mass were obtained from CIGA (Centro interdipartimentale grandi apparecchiature, Università degli Studi di Milano).

## References

1 (*a*) J. Banchereau and R. M. Steinman, *Nature*, 1998, **392**, 245–252; (*b*) D. N. J. Hart, *Blood*, 1997, **90**, 3245–3287; (*c*) I. Mellman and R. M. Steinman, *Cell*, 2001, **106**, 255–258.

2 (*a*) A. Cambi and C. G. Figdor, *Curr. Opin. Immunol.*, 2005, **17**, 345–351; (*b*) E. McGreal, J. Miller and S. Gordon, *Curr. Opin. Immunol.*, 2005, **17**, 18–24; (*c*) C. Théry and S. Amigorena, *Curr. Opin. Immunol.*, 2001, **13**, 45–51.

3 C. G. Figdor, Y. van Kooyk and G. J. Adema, *Nat. Rev. Immunol.*, 2002, 77–84.

4 T. B. H. Geijtenbeek, R. Torensma, S. J. Van Vliet, G. C. F. van Duijnhoven, G. J. Adema, Y. van Kooyk and C. G. Figdor, *Cell*, 2000, **100**, 575–585.

5 (*a*) G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand and F. Fieschi, *J. Biol. Chem.*, 2009, **284**, 21229–21240; (*b*) D. Serrano-Gomez, E. Sierra-Filardi, R. T. Martinez-Nunez, E. Caparros, R. Delgado, M. A. Munoz-Fernandez, M. A. Abad, J. Jiménez-Barbero, M. Leal and A. L. Corbi, *J. Biol. Chem.*, 2008, **283**, 3889–3903; (*c*) H. Feinberg, Y. Guo, D. A. Mitchell, K. Drickamer and W. I. Weis, *J. Biol. Chem.*, 2005, **280**, 1327–1335; (*d*) D. A. Mitchell, A. J. Fadden and K. Drickamer, *J. Biol. Chem.*, 2001, **276**, 28939–28945.

6 (*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. Kewal Ramani, D. R. Littman, C. G. Figdor and Y. Van Kooyk, *Cell*, 2000, **100**, 587–597; (*b*) P.-Y. Lozach, H. Lortat-Jacob, A. De Lacroix De Lavalette, I. Staropoli, S. Foung, A. Amara, C. Houles, F. Fieschi, O. Schwartz, J.-L. Virelizier, F. Arenzana-Seisdedos and R. Altmeyer, *J. Biol. Chem.*, 2003, **278**, 20358–20366; C. P. Alvarez, F. Lasala, J. Carrillo, O. Muniz, A. L. Corbi and R. Delgado, *J. Virol.*, 2002, **76**, 6841–6844.

7 (*a*) T. B. H. Geijtenbeek, S. J. van Vliet, E. A. Koppel, M. Sanchez-Hernandez, C. Vandenbroucke-Grauls, B. Appelmelk and Y. van Kooyk, *J. Exp. Med.*, 2003, **197**, 7–17; (*b*) L. Tailleux, O. Schwartz, J. L. Herrmann, E. Pivert, M. Jackson, A. Amara, L. Legres, D. Dreher, L. P. Nicod, J. C. Gluckman, P. H. Lagrange, B. Gicquel and O. Neyrolles, *J. Exp. Med.*, 2003, **197**, 121–127.

8 (*a*) J. J. Reina, S. Sattin, D. Invernizzi, S. Mari, L. Martinez-Prats, G. Tabarani, F. Fieschi, R. Delgado, P. M. Nieto, J. Rojo and A. Bernardi, *ChemMedChem*, 2007, **2**, 1030–1036; (*b*) D. A. Mitchell, N. A. Jones, S. J. Hunter, J. M. D. Cook, S. F. Jenkinson, M. R. Wormald, R. A. Dwek and G. W. J. Fleet, *Tetrahedron: Asymmetry*, 2007, **18**, 1502–1510; (*c*) M. J. Borrok and L. L. Kiessling, *J. Am. Chem. Soc.*, 2007, **129**, 12780–12785; (*d*) F. Lasala, E. Arce, J. R. Otero, J. Rojo and R. Delgado, *Antimicrob. Agents Chemother.*, 2003, **47**, 3970–3972; (*e*) J. Rojo and R. Delgado, *J. Antimicrob. Chemother.*, 2004, **54**, 579–581; (*f*) G. Tabarani, J. J. Reina, C. Ebel, C. Vives, H. Lortat-Jacob, J. Rojo and F. Fieschi, *FEBS Lett.*, 2006, **580**, 2402–2408; (*g*) S.-K. Wang, P.-H. Liang, R. D. Astronomo, T.-L. Hsu, S.-L. Hsieh, D. R. Burton and C.-H. Wong, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 3690–3695; (*h*) O. Martinez-Avila, L. M. Bedoya, M. Marradi, C. Clavel, J. Alcami and S. Penades, *ChemBioChem*, 2009, **10**, 1806–1809; (*i*) S. Sattin, A. Daghetti, Anna, M. Thepaut, A. Berzi, M. Sanchez-Navarro, G. Tabarani, J. Rojo, F. Fieschi, M. Clerici and A. Bernardi, *ACS Chem. Biol.*, 2010, **5**, 301–312; (*j*) O. Martinez-Avila, K. Hijazi, M. Marradi, C. Clavel, C. Campion, C. Kelly and S. Penades, *Chem.–Eur. J.*, 2009, **15**, 9874–9888; (*k*) K. C. A. Garber, K. Wangkanont, E. E. Carlson and L. L. Kiessling, *Chem. Commun.*, 2010, **46**, 6747–6749.

9 E. J. Soilleux, *Clin. Sci.*, 2003, **104**, 437–446.

10 S. Pöhlmann, F. Baribaud and R. W. Doms, *Trends Immunol.*, 2001, **22**, 643–646.

11 (*a*) S. Pöhlmann, E. J. Soilleux, F. Baribaud, G. J. Leslie, L. S. Morris, J. Trowsdale, B. Lee, N. Cleman and R. W. Doms, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 2670–2675; (*b*) E. Van Liempt, C. M. C. Bank, P. Mehta, J. J. García-Vallejo, Z. S. Kawar, R. Geyer, R. A. Alvarez, R. D. Cummings, Y. van Kooyk and I. van Die, *FEBS Lett.*, 2006, **580**, 6123–6131; (*c*) E. Van Liempt, A. Imberty, C. M. C. Bank, S. J. Van Vliet, Y. Van Kooyk, T. B. H. Geijtenbeek and I. Van Die, *J. Biol. Chem.*, 2004, **279**, 33161–33167.

12 Y. Guo, H. Feinberg, E. Conroy, D. A. Mitchell, R. Alvarez, O. Blixt, M. E. Taylor, W. I. Weis and K. Drickamer, *Nat. Struct. Mol. Biol.*, 2004, **11**, 591–598.

13 H. Feinberg, M. E. Taylor, N. Razi, R. McBride, Y. A. Knirel, S. A. Graham, K. Drickamer and W. I. Weis, *J. Mol. Biol.*, 2011, **405**, 1027–1039.

14 G. Timpano, G. Tabarani, M. Anderluh, D. Invernizzi, F. Vasile, D. Potenza, P. M. Nieto, J. Rojo, F. Fieschi and A. Bernardi, *ChemBioChem*, 2008, **9**, 1921–1930.

15 For a review on structural studies on the conformation of Lewis oligosaccharides see: E. Yuriev, W. Farrugia, A. M. Scott and P. A. Ramsland, *Immunol. Cell Biol.*, 2005, **83**, 709–717.

16 L. de Witte, A. Nabatov, M. Pion, D. Fluitsma, M. A. de Jong, T. de Gruijl, V. Piguet, Y. van Kooyk and T. B. H. Geijtenbeek, *Nat. Med.*, 2007, **13**, 367–371.

17 P. J. Goodford, *J. Med. Chem.*, 1985, **28**, 849–857.

18 H. Feinberg, D. A. Mitchell, K. Drickamer and W. I. Weis, *Science*, 2001, **294**, 2163–2166.

19 H. Feinberg, R. Castelli, K. Drickamer, P. H. Seeberger and W. I. Weis, *J. Biol. Chem.*, 2007, **282**, 4202–4209.

20 Glide version 4.5, Schrödinger, LLC NY 2007.

21 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.

22 A. L. Hopkins, C. R. Groom and A. Alex, *Drug Discovery Today*, 2004, **9**, 430–431.

23 C. Guzzi, J. Angulo, F. Doro, J. J. Reina, M. Thépaut, F. Fieschi, A. Bernardi, J. Rojo and P. M. Nieto, submitted.

24 M. Andreini, M. Anderluh, A. Audfray, A. Bernardi and A. Imberty, *Carbohydr. Res.*, 2010, **345**, 1400–1407.

25 L. de Witte, A. Nabatov and T. B. H. Geijtenbeek, *Trends Mol. Med.*, 2008, **14**, 12–19.

26 M. A. W. P. de Jong, L. de Witte, M. Taylor and T. B. H. Geijtenbeek, *J. Immunol.*, 2010, **185**, 1633–1641.

27 M. Thépaut, J. Valladeau, A. Nurisso, R. Kahn, B. Arnou, C. Vives, S. Saeland, C. Ebel, C. Monnier, C. Dezutter-Dambuyant, A. Imberty and F. Fieschi, *Biochemistry*, 2009, **48**, 2684–2698.

28 F. Halary, A. Amara, H. Lortat-Jacob, M. Messerie, T. Delaunay, C. Houlès, F. Fieschi, F. Arenzana-Seisdedos, J.-F Moreaus and J. Déchanet-Merville, *Immunity*, 2002, **17**, 653–664.

29 MacroModel, version 9.5, Schrödinger, LLC, New York, NY, 2007.

30 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11235.

**Additional studies.** After the latter studies were published, a couple of modifications were introduced to compound **10b** yielding compounds **10b-azide1** and **10b-azide2**, shown in figure 9.13A (an ethylazide linker was introduced in order to make it possible to tether these ligands on polyvalent scaffolds). They were also tested in SPR competition assay for DC-SIGN inhibition and the results are shown in figure 9.13B.

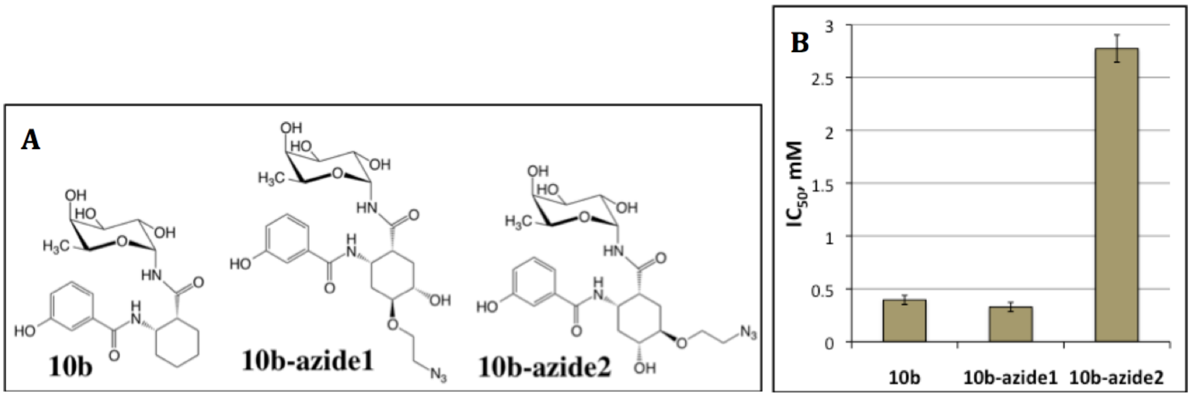


**Figure 9.13.:** The structures and activities of compound **10b** and its derivatives.
**A**, The structures of the three compounds. **B**, The $IC_{50}$ values of DC-SIGN inhibition by the corresponding compounds; the error bars show standard deviations (for **10b** and **10b-azide1** values are as in table 9.1, for **10b-azide2** values are from 3 measurements).

The addition of a *meta* hydroxyl and the azide linker to yield compound **10b-azide1** apparently had no significant effect on the activity of the compound as compared to **10b**. However, the swapping positions of hydroxyl and the azide linker groups in **10b-azide1** leading to compound **10b-azide2** resulted in a dramatic decrease of the activity of this compound, probably suggesting the steric hindrances imposed to the binding by the following swap.

Compound **10b-azide1** was selected for further development to tether it on a multivalent scaffold.

### 9.1.4. *Development of mannose-based glycomimetic compounds*

Two mannose-based glycomimetic compounds, psDi [143] and psTri [49] (fig. 9.14), were designed and evaluated as DC-SIGN inhibitors before my PhD started, and psTri appeared to have an order of magnitude better inhibitory potency than psDi [49]. Hence it was chosen to generate the first multivalent (tetravalent dendron) glycomimetic compound and tested in cell-based HIV infection inhibition assay [49].

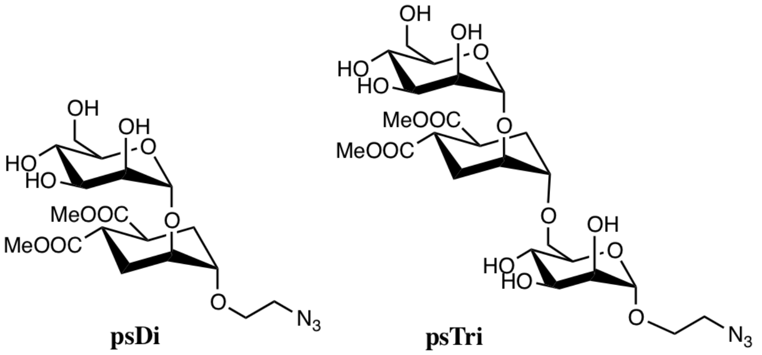


**Figure 9.14.:** The structures of previously designed mannose-based glycomimics psDi and psTri.

The results of these cellular studies were promising, thus psTri dendron later was also assessed for the potency to inhibit HIV infection in cervical explants (article **n°2**), a real-life closer experimental set-up.

When my PhD started, in addition to psTri, polyvalent form of psDi were also generated, and one of my very first tasks was the evaluation of these multivalent compounds as DC-SIGN inhibitors in SPR competition assay. Simultaneously, the same compounds were evaluated by our collaborators in Spain, Dr. Rafael Delgado team, in a cell-based assay as Ebola virus infection inhibitors (article **n°3**).

Although in all of these studies psTri was observed as markedly more potent DC-SIGN inhibitor than psDi, there was also a discrepancy in the results observed as psTri was the better inhibitor only in monovalent form, and the potencies of both psDi and psTri were very similar once they were tethered to multivalent scaffolds. Therefore the efforts were invested to study the interaction with DC-SIGN of these both compounds in more detail. After solving the structures of DC-SIGN in complex with psDi and psTri (papers **n°4** and **n°5**), even more discrepancy arose: despite rather different activities of these two compounds, they had virtually the same binding modes within DC-SIGN CRD. In addition, the NMR studies of psDi interaction with DC-SIGN indicated a unique binding mode for this compound, which was in a good agreement with X-ray data, while NMR data for psTri showed complex interaction properties (results not shown). Thermodynamic and hydrodynamic characterization of the interaction of both compounds with the whole extracellular domain of DC-SIGN was also performed. These studies revealed that the higher potency of psTri was an artifact resulting from a competition assay set-up: soluble DC-SIGN tetramers are used in this assay and psTri molecule showed a unique capability to bridge the protein tetramers.

These studies disapproved psTri as the best lead compound. On the other hand, they revealed that synthetically simpler psDi had all the desirable qualities of a compound for further development as a drug candidate: a well-defined unique binding mode to DC-SIGN, favorable selectivity to

DC-SIGN combined with a good synthetic availability. Hence, further design of mannose-based glycomimetics went on using psDi as a starting scaffold and the structural knowledge of its binding to DC-SIGN. Two different types of psDi-based compound series were synthesized and evaluated for inhibition and selectivity to DC-SIGN versus langerin. The first series included the aromatic bis-benzylamide substituents on the cyclohexane ring of psDi in place of methyls (paper **n°6**). From this series, one compound, namely **4h**, was identified as a new lead compound. The second psDi-based series ($NV$-type, sub-subsection 9.1.4.6 on page 290) had different substituents at position 6 of non-reducing mannose of psDi.

#### 9.1.4.1. *Evaluation of tetravalent psTri to inhibit HIV infection in human cervicovaginal explants, and estimation of monovalent psTri selectivity to DC-SIGN versus langerin*

The previous encouraging results of tetravalent psTri dendron (fig. 9.15 on page 191) tested in cell-based HIV *trans* infection assay [49] prompted to evaluate this compound in a more biologically relevant assay. The following article by Berzi et al. presents the characterization of this compound as the inhibitor of HIV infection of cervicovaginal explants. The dendron proved to be non-toxic at used concentrations, it was capable to significantly reduce R5-tropic HIV infection and had a moderate activity against X4-tropic virus. It also induced the production of anti-HIV activity possessing β-chemokines.

Because Langerhans cells are present in the endothelial layer of the cervix, it was interesting to assess the selectivity of psTri to DC-SIGN versus langerin. Therefore SPR competition assay with soluble DC-SIGN and langerin was performed and showed that monovalent psTri, as compared to D-mannose, had a desirable increased selectivity to DC-SIGN.

The main outcome of these studies:

- ✧ The strategy of HIV infection inhibition using glycomimetics works not only in a cell-based assay but also in a more relevant *ex vivo* assay.
- ✧ psTri had better selectivity to DC-SIGN than langerin ($sel.g._{mannose \rightarrow psTri} = 5.4 \pm 0.9$, not presented in the article).

Contributions:

In the article presented dendrons and dendrimers were synthesized by the group of Dr. J. Rojo. The assays of HIV infection of cervicovaginal explants were carried out by the group of Pr. Mario Clerici.

My contribution to this study:

I have conducted the SPR assay and analyzed the data.

# 2

# 2

**Paper n$^{o}$2: A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models**

# A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models

Angela Berzi[a], José J. Reina[d], Roberta Ottria[d], Ieva Sutkeviciute[e,f,g], Patrizio Antonazzo[i], Macarena Sanchez-Navarro[j], Eric Chabrol[e,f,g], Mara Biasin[a], Daria Trabattoni[a], Irene Cetin[i], Javier Rojo[j], Franck Fieschi[e,f,g,h], Anna Bernardi[d,k] and Mario Clerici[b,c]

**Objective:** Dendritic cell-specific intercellular adhesion molecule (ICAM)-3 grabbing nonintegrin (DC-SIGN) participates in the initial stages of sexually transmitted HIV-1 infection by recognizing highly mannosylated structures presented in multiple copies on HIV-1 gp120 and promoting virus dissemination. Inhibition of HIV interaction with DC-SIGN thus represents a potential therapeutic approach for viral entry inhibition at the mucosal level.

**Design:** Herein we evaluate the efficacy in inhibiting HIV-1 infection and the potential toxicity of a multimeric glycomimetic DC-SIGN ligand (Dendron **12**).

**Methods:** The ability of Dendron **12** to block HIV-1 infection was assessed in cellular and human cervical explant models. Selectivity of Dendron **12** towards DC-SIGN and langerin was evaluated by surface plasmon resonance studies. β chemokine production following stimulation with Dendron **12** was also analyzed. Toxicity of the compound was evaluated in cellular and tissue models.

**Results:** Dendron **12** averted HIV-1 *trans* infection of $CD4^+$ T lymphocytes in presence of elevated viral loads and prevented HIV-1 infection of human cervical tissues, under conditions mimicking compromised epithelial integrity, by multiple clades of R5 and X4 tropic viruses. Treatment with Dendron **12** did not interfere with the activity of langerin and also significantly elicited the production of the β chemokines MIP-1α, MIP-1β and RANTES.

**Conclusion:** Dendron **12** thus inhibits HIV-1 infection by competition with binding of HIV to DC-SIGN and stimulation of β-chemokine production. Dendron **12** represents a promising lead compound for the development of anti-HIV topical microbicides.

© 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins

*AIDS* 2012, **26**:127–137

**Keywords: cervical explants, DC-SIGN, glycomimetic drugs, HIV, topical microbicides**

---

[a]Chair of Immunology, Department of Clinical Sciences L. Sacco, [b]Chair of Immunology, Department of Biomedical Sciences and Technologies, University of Milan, Segrate, [c]Don C. Gnocchi Foundation IRCSS, [d]Department of Organic and Industrial Chemistry, University of Milan, Milan Italy, [e]Institut de Biologie Structurale, Université Joseph Fourier, [f]CEA, DSV, [g]CNRS, UMR 5075, Grenoble, [h]Institut Universitaire de France, Paris, France, [i]Unit of Obstetrics and Gynaecology, Department of Clinical Sciences L. Sacco, University of Milan, Milan, Italy, [j]Instituto de Investigaciones Químicas, CSIC - University of Sevilla, Sevilla, Spain, and [k]CNR-ISTM, Institute of Molecular Science and Technologies, Milan, Italy.

Correspondence to Mario Clerici, Chair of Immunology, Department of Biomedical Sciences and Technologies, University of Milan, Via Fratelli Cervi 93, 20090 Milan, Italy.

Tel: +39 025030412; fax: +39 025030414; e-mail: mario.clerici@unimi.it

Received: 7 September 2011; revised: 5 October 2011; accepted: 12 October 2011.

DOI:10.1097/QAD.0b013e32834e1567

ISSN 0269-9370 © 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

## Introduction

HIV remains one of the leading causes of mortality and morbidity [1]. As the vast majority of HIV-1 infections occur via sexual transmission through mucosal surfaces, the development of vaginal and rectal topical microbicides represents a promising approach to prevent sexually transmitted HIV-1 infection.

Dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin (DC-SIGN) is involved in the initial step of HIV-1 sexually transmitted infection and it may be considered a promising therapeutic target [2,3].

Myeloid immature dendritic cells located in mucosal tissue of vagina, cervix and rectum express DC-SIGN and are among the first cells to encounter sexually transmitted HIV [4–6]. Upon HIV-1 binding to DC-SIGN, the virus escapes (at least partially) degradation into lysosomes. Rather, it is internalized into endosomes and multivescicolar bodies, in which it is protected from degradation and retained in a high infective state [7–9]. Dendritic cells transmit the virus in *trans* to adjacent $CD4^+$ T lymphocytes in genital mucosae or, after migration to lymphoid tissue, to $CD4^+$ T lymphocytes resident in lymphoid tissue, promoting HIV-1 dissemination [10,11]. DC-SIGN, facilitating HIV-1 interaction with CD4 and co-receptors, also enhances dendritic cell infection in *cis*, that results in long-term transmission of HIV [12,13].

In addition, binding of HIV-1 to DC-SIGN activates signaling pathways that modulate Toll-like receptor signaling, inducing immunosuppressive responses and triggering HIV replication and transmission [14–17].

DC-SIGN specifically recognizes the high mannose glycan ($Man_9$), presented in multiple copies on HIV-1 gp120 [8,18]. Fragments of $Man_9$ terminated by a dimannoside or a tri-mannoside bind to DC-SIGN almost as efficiently as the entire $Man_9$ [19]. Chemically synthesized analogs of $(Man)_9$ terminal di-saccharides and tri-saccharides are more resistant to hydrolysis by glycosydases than the corresponding natural oligosaccharides and interact efficiently with DC-SIGN [20]. These analogs can be linked to tetravalent (dendrons) scaffolds to obtain compounds endowed with stronger binding affinities to DC-SIGN [21,22]. Such compounds can compete with binding of HIV gp120 to DC-SIGN and are suitable for the development of new anti-HIV microbicides.

We recently demonstrated that a tetravalent dendron containing four copies of a linear pseudo-mannotrioside (Dendron **12**) was able to inhibit HIV-1 *trans* infection of $CD4^+$ T lymphocytes [23]. Experimental models based on infection with HIV of explants taken from human uterine cervix, albeit with some limitations, allow a better approximation of the conditions *in vivo* compared to cellular models [24–27]. Thus, in this study a cervical explant model was exploited to assess the efficacy in inhibiting HIV-1 infection and the toxicity of Dendron **12**, with the purpose of evaluating if the compound is a suitable candidate for the development as topical microbicide.

## Material and methods

### Cell culture

B-THP1 and B-THP1/DC-SIGN cells (contributed by Drs Li Wu and Vinet N. KevalRamani) were cultured as described [23].

### Virus

The following HIV-1 strains were used: BaL (contributed by Drs. S. Gartner, M. Popovic and R. Gallo; DU174 (Source: Dr L Morris); the R5 and X4 tropic primary isolates 8 g and DPMVF (provided by Professor Stefano Aquaro).

### Inhibition of HIV infection in *trans*

Human $CD4^+$ T lymphocytes were purified from peripheral blood of volunteer healthy donors following written consent, and activated as described [23]. Dendron **12**, synthesized as described [23], was diluted to desired concentration into culture medium [Roswell Park Memorial Institute (RPMI) 160 with 20% fetal bovine serum (FBS), penicillin and streptomycin and L-glutamine, all from Euroclone, Siziano, Italy]. B-THP1/DC-SIGN or B-THP1 cells ($10^6$ cells/ml) were preincubated with Dendron **12** (250 μmol/l), or culture medium alone for 30 min prior to exposure to BaL (virus titer ranging from 5 $TCID_{50}$ to 80 $TCID_{50}$), in the continued presence of the inhibitor 3 h at 37°C. After extensive washing, B-THP1/DC-SIGN cells were co-cultured with activated human $CD4^+$ T cells as previously described [23].

### Human cervical explants infection

Cervical tissue was obtained, following written informed consent, from premenopausal women, HIV, hepatitis B virus (HBV) and hepatitis C virus (HCV) seronegative and without current genital infection, undergoing therapeutic hysterectomy at Unit of Obstetrics and Gynaecology of Sacco Hospital (Milan, Italy). The study was approved by the local ethic committee and conducted in compliance with international guidelines and local laws. Endocervical 3 mm × 3 mm explant biopsies comprised both epithelium and stromal tissue.

Within 1 h after obtaining tissues the explants were extensively washed. Then explants were either

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

immediately treated with the Dendron **12** and infected with R5 tropic strains or were prestimulated for 2 days in presence of IL-2 (R&D systems, Minneapolis, Minnesota, USA) and phytohaemagglutinin (PHA) (Sigma–Aldrich, Saint Louis, Missouri, USA) prior to X4 tropic DPMVF exposure.

Explants were pretreated 30 min at 37°C with different concentrations of Dendron **12**. Afterwards explants were exposed to HIV-1 BaL, DU174 (both $2.6 \times 10^4$ $TCID_{50}$), 8 g or DPMVF (both $10^4$ $TCID_{50}$) in the continued presence of the compound 3 h at 37°C. Subsequently explants were washed with RPMI. Unstimulated explants were cultured in RPMI medium supplemented with 20% FBS, penicillin and streptomycin, L-glutamine (Euroclone) and gentamycin (Sigma–Aldrich) at 37°C and 5% $CO_2$. Stimulated explants were cultured under the same conditions in presence of IL-2. Supernatants were collected 3 and 7 days post infection.

## p24 ELISA

p24 concentration in the supernatants was assayed by Alliance HIV-1 p24 Antigen kit (Perkin Elmer, Waltham, Massachusetts, USA). Plates were read using the IMark microplate reader equipped with Microplate Manager 6 software (Biorad, Segrate, Italy).

## Toxicity on peripheral blood mononuclear cells

Peripheral blood mononuclear cells (PBMC), purified from peripheral blood of healthy donors as described [23], were incubated with different concentrations of Dendron **12** for 3 or 7 days. The apoptosis was monitored evaluating the percentage of dead cells by staining with 7-AAD (Beckman Coulter, Fullerton, California, USA). Flow cytometric analyses were performed using a CYTOMICS FC-500 flow cytometer interfaced with CXP21 software (Beckman Coulter).

## 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay

Toxicity of Dendron **12** was determined by a 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide-based assay (Sigma–Aldrich). Viable explants reduce MTT to formazan crystals. Explants were cultured with increasing concentration of Dendron **12** diluted in medium culture for 3 and 7 days. The 50% of medium culture (containing the different concentration of Dendron **12**) was changed at day 3. After culturing, explants were washed and incubated in medium RPMI without phenol red and 10% FBS and MTT. Formazan was dissolved by MTT solubilization solution and formazan absorbance was measured at 595 nm.

## Differentiation and treatment of monocyte-derived dendritic cells

$CD14^+$ monocytes were separated from PBMCs using the $CD14^+$ microbeads (Miltenyi Biotech, Bergisch Gladbach, Germany) following manufacturer's instruction. Monocytes were differentiated into iDCs by culturing them in presence of IL-4 (20 ng/ml) and granulocyte-macrophage colony-stimulating factor (20 ng/ml) (R&D Systems) for 6 days. DC-SIGN expression was checked by staining with anti human DC-SIGN-PE monoclonal antibody (clone AZND1, Beckman Coulter) and flow cytometric analysis. Monocyte-derived dendritic cells (MDDCs) were treated with Dendron **12**. Supernatants were harvested and RNA extracted from cells after 3, 24 and 72 h. β chemokine concentration in culture supernatants was evaluated using DuoSet kits (R&D Systems).

## RNA extraction and real-time PCR

RNA was extracted using the acid guanidium thiocyanate–phenol–chloroform method and purified from genomic DNA with RNase-free DNase (RQ1 DNase, Promega, Madison, Wisconsin, USA). RNA was reverse-transcribed using random examer primers and M-MLV reverse transcriptase (Clontech, Palo Alto, California, USA). cDNA quantification for macrophage inflammatory protein (MIP)-1α, MIP-1β, RANTES (regulated upon activation, normal T-cell expressed, and secreted) and glycerAldehyde 3-phosphate dehydrogenase (GAPDH) was performed by real-time PCR (DNA Engine Opticon 2; MJ Research, Ramsey, Minnesota, USA). Reactions were performed using a SYBR Green PCR mix (Finnzymes, Espoo, Finland). Results were expressed as ΔΔCt and presented as ratios between the target gene and the GAPDH housekeeping mRNA.

## Surface plasmon resonance analysis

Extracellular domain (ECD) of langerin (residue 68–328) and DC-SIGN (residue 66–404) was overexpressed and purified as described [28,29]. Surface plasmon resonance (SPR) experiments were performed on a Biacore 3000 using functionalized CM4 sensor chips and the corresponding reagents from Biacore. Two flow cells were activated as described [30]. Flow cell one was blocked with ethanolamine and used as a control surface. The second one was treated with BSA-Man**α**1–3[Man**α**1–6] Man (Man-BSA, Dextra) (60 μg/ml) in 10 mmol/l acetate buffer, pH 4. Remaining activated groups were blocked with ethanolamine. The final density immobilized on the surface of the second flow cell was 5200 RU. The Man-BSA used to functionalize CM4 chip harbors 12 glycosylation sites. The affinities for DC-SIGN ECD and langerin ECD of pseudo-mannotrioside and D-mannose were estimated by an inhibition assay, as described [22,31]. Each lectin was injected onto the Man-BSA surface, at 20 μmol/l alone or in presence of an increasing concentration of compounds. Injections were performed at 5 μl/min using 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mmol/l $CaCl_2$, and 0.005% of P20 surfactant as running buffer. The surface was regenerated by 1' injection of 50 mmol/l EDTA, pH 8. The $IC_{50}$ values were determined as described [22,31].

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

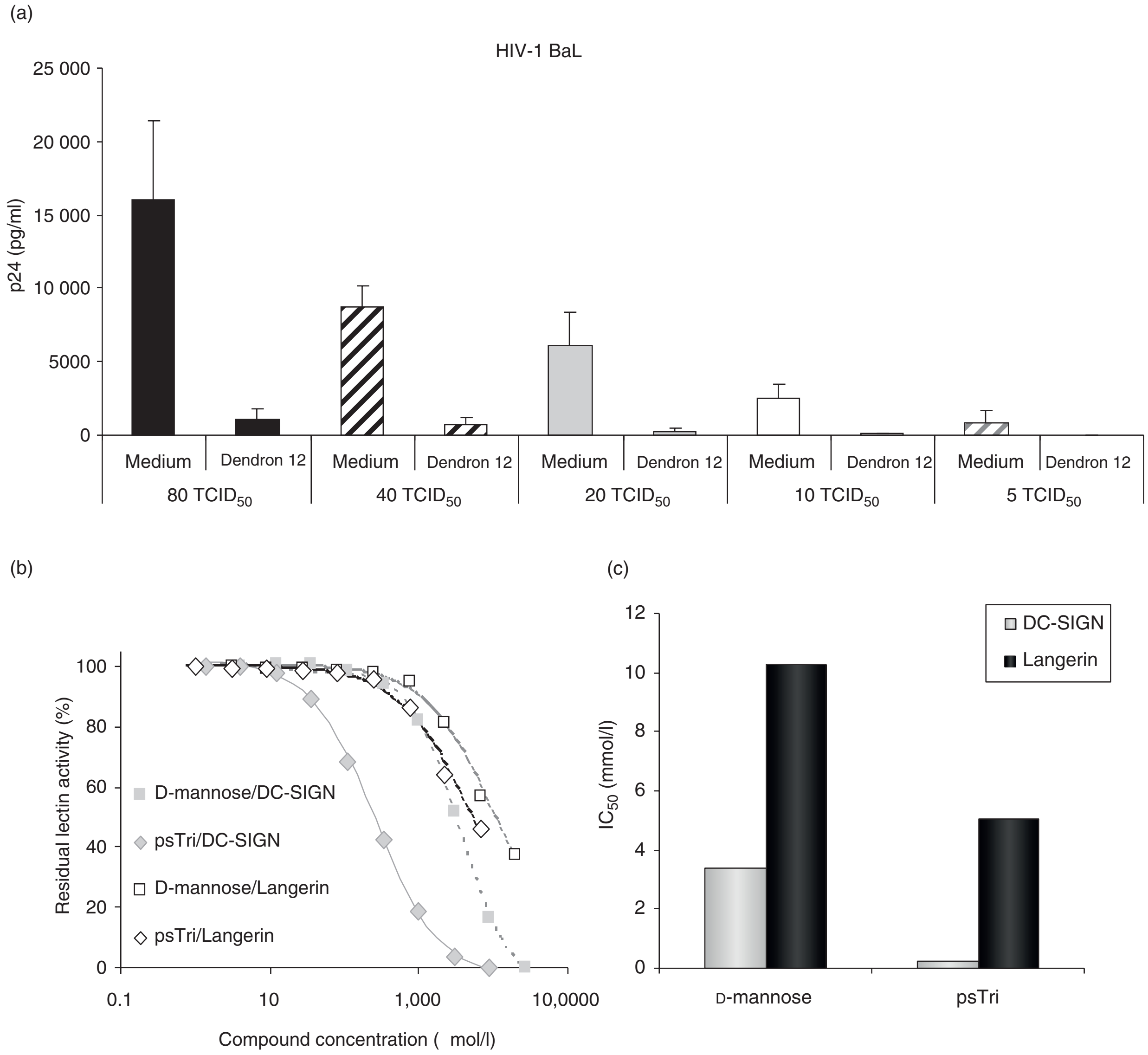


**Fig. 1. Evaluation of Dendron *12* ability to inhibit HIV-1 Bal *trans* infection in presence of increasing amounts of the virus (a) and selectivity to DC-SIGN (b,c).** (a) After pretreatment with Dendron **12** (250 μm) or medium culture, B-THP-1/DC-SIGN cells were pulsed with BaL for 3 h. After washing cells were co-cultured 3 days with CD4$^+$ T lymphocytes from healthy donors. Levels of infection were quantified by measuring p24 in the supernatants of co-cultures by ELISA. Data were obtained from three different healthy donors. Each donor was tested in duplicate. Values are mean ± SD. (b, c) SPR experiment results of DC-SIGN ECD and langerin ECD binding to Man-BSA/dextran surface inhibition by pseudo-mannotrioside (psTri) and D-mannose. (b) Inhibition curves, and (c) lectin selectivity histograms.

## Results

### Dendron 12 inhibits *trans* infection in presence of HIV-1 elevated viral load

We firstly examined whether the Dendron **12** was able to inhibit the *trans* infection in the presence of elevated viral load. B-THP1/DC-SIGN cells were used as model to mimic HIV transmission to CD4$^+$ T cells, as previously described [6,23,32]. B-THP1/DC-SIGN cells were preincubated 30 min in the presence or in absence of Dendron **12** and then exposed to different inoculums of HIV-1 Bal. Then, B-THP1/DC-SIGN cells were washed and co-cultured with activated CD4$^+$ T cells. At lower viral concentration inhibition of BaL transmission to CD4$^+$ T cells was almost complete. Even at higher viral load (40 and 80 TCID$_{50}$) the Dendron **12** was able to counteract the transmission of the virus to CD4$^+$ T lymphocytes (more than 92% of inhibition) (Fig. 1a).

### Selectivity towards DC-SIGN

Both langerin and DC-SIGN recognize Man$_9$ on gp120, but have distinct specificities towards complex

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

oligosaccharides [33–36]. Specificity of pseudo-mannotrioside ligand for langerin and DC-SIGN was tested by competition experiments using SPR, as previously described [37]. A CM4 sensor chip was functionalized with BSA-mannotriose and a fixed amount of the extracellular domain of DC-SIGN and langerin was injected over the surface in the presence or absence of pseudo-mannotrioside or mannose (control). From the inhibition curves (Fig. 1b), an $IC_{50}$ of the two ligands towards each lectins was evaluated (Fig. 1c). A limited difference in favor of DC-SIGN was observed for mannose. On the contrary, pseudo-mannotrioside is 20 times more potent toward DC-SIGN than against langerin. Moreover, pseudo-mannotrioside is 14 times more potent than D-mannose towards DC-SIGN. Indeed, the use of pseudo-mannotrioside allows real improvement in affinity and in selectivity.

## Inhibition of HIV-1 infection of human cervical tissue by Dendron 12

Endocervical tissue was obtained from premenopausal women, HIV, HBV and HCV seronegative, undergoing planned therapeutic hysterectomy. Explants were exposed to HIV-1 in a nonpolarized manner, analogous to condition of compromised epithelium *in vivo*. As the laboratory adapted R5 strain HIV-1 BaL is able to infect resting tissue [25], explants were not activated to mimic physiological conditions.

Explants were pretreated 30 min in absence or in presence of increasing concentration of Dendron **12** and then were exposed to BaL in the continued presence of compound. After washing to remove unbound Dendron **12** and virus, explants were maintained in culture up to 7 days. Data represent p24 levels and are presented as the mean of five independent experiments, using explants from separate donors. Dendron **12** inhibited cervical explants BaL infection in dose-dependent manner (Fig. 2). At the higher concentration tested, Dendron **12** reduced the infection by about 80%, at 3 and 7 days post infection.

The ability of Dendron **12** to avert explant infection mediated by primary HIV-1 isolates was also verified. R5 tropic isolate 8 g was able to infect unstimulated tissue, but X4 tropic isolate DPMVF needed preactivation to induce infection (not shown). Explants unstimulated or preactivated 2 days with IL-2 and PHA were pretreated with increasing concentration of Dendron **12** and infected, respectively, with 8 g or DPMVF, as described before. Infection inhibition was dose-dependent. At the concentration of 1 mmol/l, infection by both isolates was reduced by more than 85%. At 0.05 mmol/l Dendron **12** decreases by 56% (day 3) and by 40% (day 7) the infection mediated by 8 g (Fig. 3a and b), but the inhibitory effect against DPMVF was largely lost (Fig. 3c and d).

Furthermore we evaluated the capability of Dendron **12** to block explant infection by Clade C R5 tropic strain DU174. Unstimulated explants were pretreated with the compound, exposed to DU174 and cultured as described. Dendron **12** reduced DU174 infection in dose-dependent manner and by about 90% at 1 mM (Fig. 3e and f).

## Induction of β1 chemokines production by Dendron 12

We wondered if, in addition to competitive inhibition of DC-SIGN, other mechanisms account for the antiviral effect of the compound **12**. So we investigated if Dendron **12** stimulates the production of factors interfering with

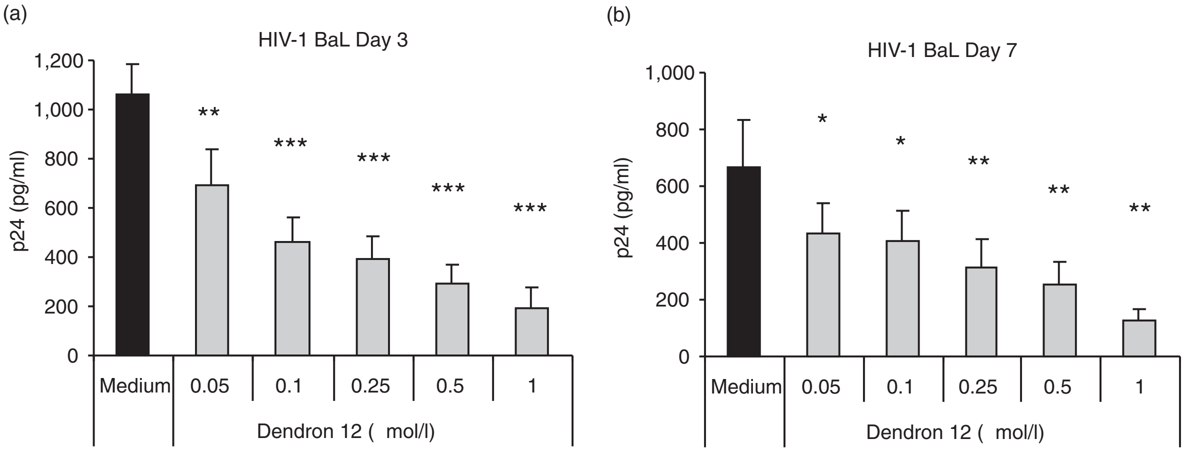


**Fig. 2. Inhibition of endocervical tissue infection mediated by HIV-1 Bal.** After a 30 min pretreatment with the Dendron **12** or medium culture, endocervical explants were incubated 3 h with BaL in the continued presence of indicated concentrations of Dendron **12**. Then explants were washed and cultured for 7 days. Infection was monitored by ELISA measurement of p24 in explant culture supernatants at day 3 (a) and 7 (b) post infection. Values represent the mean ± SD of five independent experiments. $^{*}P < 0.05$ (Student's *t*-test), $^{**}P < 0.01$, $^{***}P < 0.001$.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

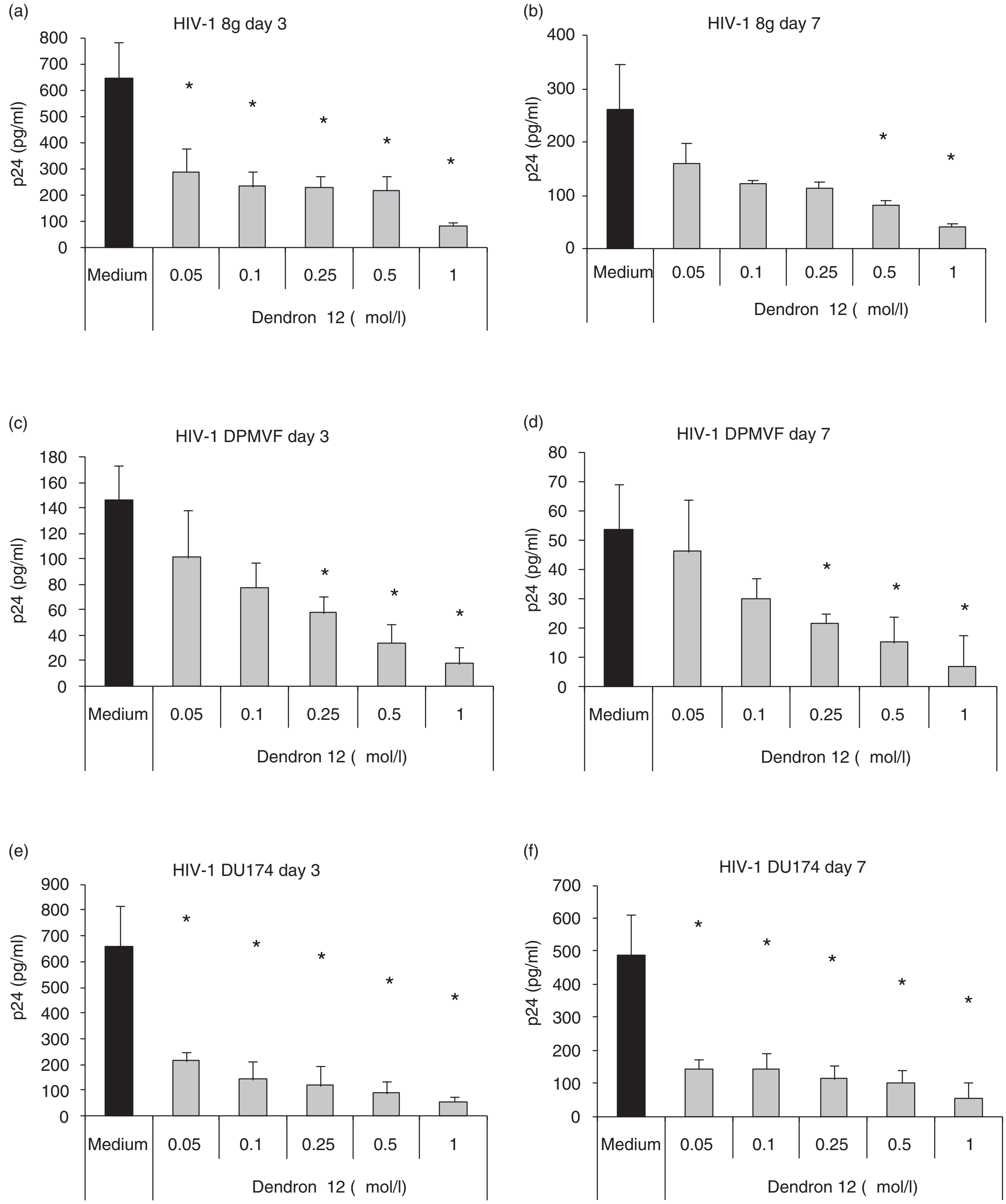


**Fig. 3. Inhibition of endocervical infection induced by HIV-1 primary isolates (8 g and DPMVF) and HIV-1clade C DU174.** (a, b, e, f) Unstimulated explants were pretreated 30′ with Dendron **12** or medium culture and challenged with 8 g or DU174. (c, d) Explants immuno-stimulated were pretreated 30′ with Dendron **12** or medium culture and infected with DPMVF. After washing explants were put in culture. Infection was monitored by ELISA measurement of p24 in explant culture supernatants at day 3 and 7 post infection. Values represent the mean ± SD of three independent experiments. $^{*}P < 0.05$ (Student's *t*-test).

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

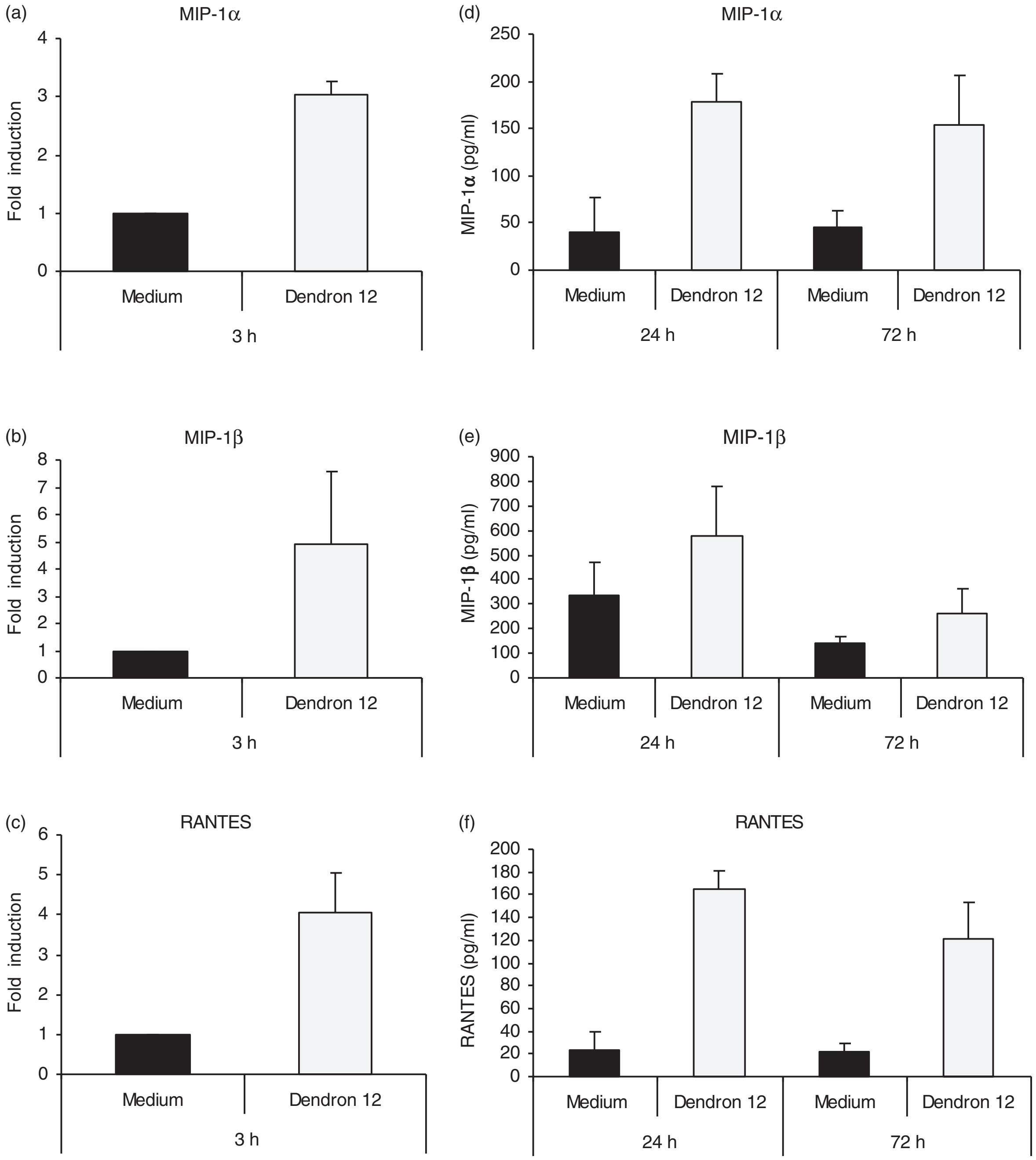


**Fig. 4. β chemokines induction upon Dendron *12* (250 μmol/l) stimulation of human iMDDCs from healthy donors.** (a–c) β chemokines expression following 3 h Dendron **12** (250 μmol/l) treatment. mRNA levels of MIP-1α (a), MIP-1β (b) and RANTES (c) were assessed by quantitative real-time PCR. Expression is normalized to GAPDH and shown as fold changes expression from the unstimulated sample, set as 1. (d–f) β chemokines protein production after stimulation with Dendron **12** (250 μmol/l). The concentration of MIP-1α (d), MIP-1β (e) and RANTES (f) in the culture supernatants was assayed by ELISA at 24 and 72 h. (a–f) Values represent the mean ± SD.

HIV infection. Due to difficulty of isolating sufficient amounts of primary mucosal dendritic cells, immature monocyte-derived dendritic cells (iMDDCs) that share with mucosal dendritic cells similar features and DC-SIGN expression, were used as a model [38]. iMDDCs were treated with Dendron **12** for 3, 24 and 72 h. Expression and production of β chemokines MIP-1α, MIP-1β, and RANTES following stimulation were analyzed by quantitative real-time PCR and ELISA. The treatment increased expression level of mRNA specific

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

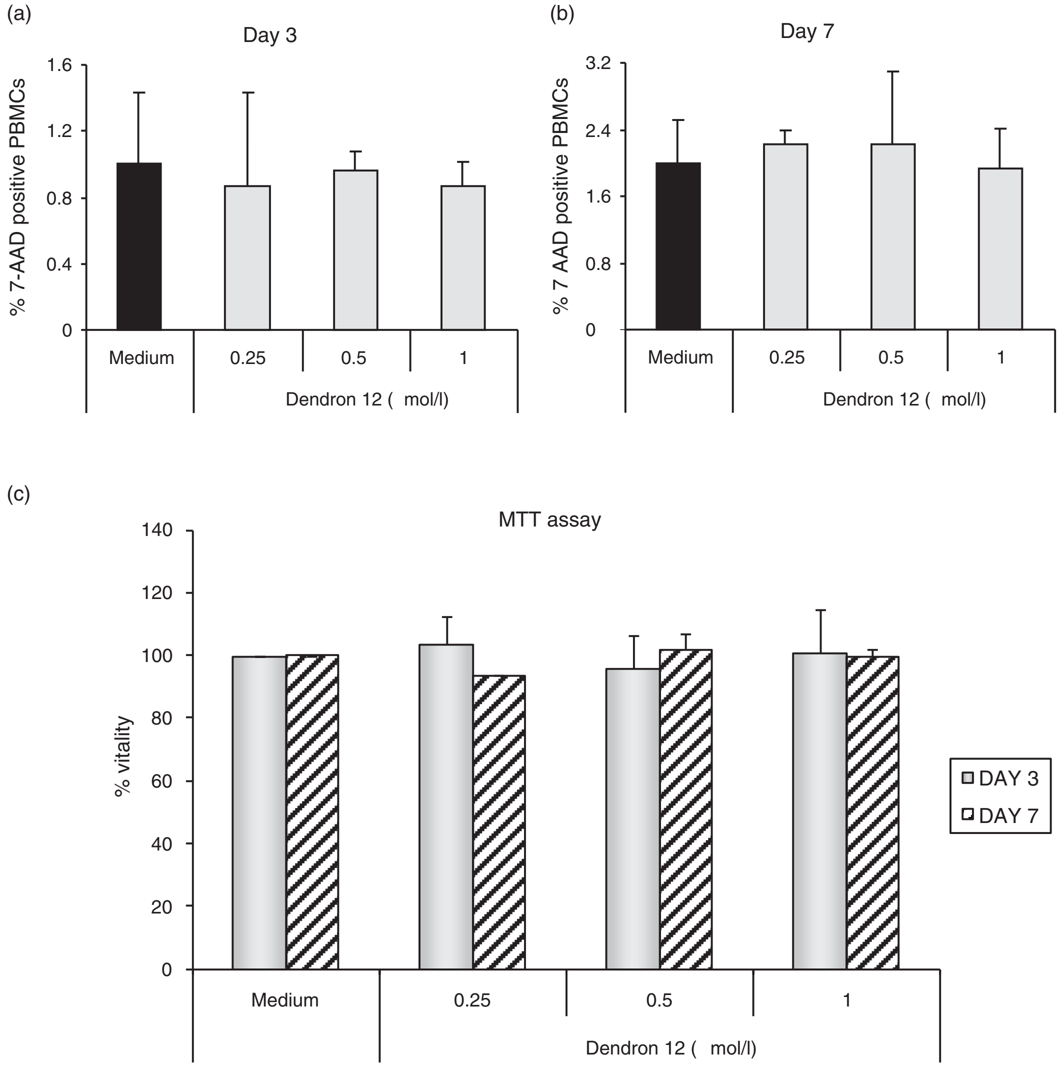


**Fig. 5. Toxicity of compound *12*.** (a, b) Percentage of 7-AAD-positive (nonviable) PBMCs after 3 or 7 days of incubation with different concentrations of Dendron **12** (0.25, 0.5 and 0.1 mmol/l), or in the absence of the inhibitor (medium). Percentage of 7-AAD was determined by flow cytometry. Experiments were performed on PBMCs from three healthy donors. Values are mean ± SD. Effect of Dendron **12** on cervical explants viability (c). Endocervical explants, derived from the same donors, were exposed in nonpolarized manner to Dendron **12** or culture medium (control) for 3 and 7 days. Effect of Dendron **12** on tissue viability was determined by the MTT assay. Explants were weighted to normalize optical density of formazan yielded. Viability was expressed as percentage. Experiments were performed on explants from three donors. Data are reported as the average percentage viability (±SD).

for MIP-1α, MIP-1β and RANTES (Fig. 4a–c) after 3 h. Also MIP-1α, MIP-1β and RANTES production increased after 24 and 72 h of stimulation (Fig. 4d–f).

## Evaluation of Dendron 12 toxicity

Cytotoxicity against PBMCs was evaluated by labeling with 7-amino-actinomycin D (7-AAD) that identifies nonviable cells after Dendron **12** treatment. Compound exposure for 3 and 7 days did not alter significantly the viability of PBMCs (Fig. 5a and b).

To assess toxicity of Dendron **12** towards cervical tissue, effect of nonpolarized exposure of the compound on explants viability was monitored (Fig. 5c). After 3 and 7 days treatment in absence or in presence of different concentration of Dendron **12**, viability was evaluated by a MTT-based assay. Viability of the compound treated explants was compared to viability of untreated control. No significant difference between control and treated explants was observed up to a concentration of 1 mM (the higher concentration tested in assessment of compound efficacy against HIV-1 infection).

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

## Discussion

Three decades after HIV discovery HIV-AIDS pandemic continues and millions of people are infected every year. Thus, the development of effective, nontoxic and low-cost topical microbicides represents a valid alternative approach to prevent the sexual transmission of HIV [3]. However, so far almost all compounds failed to prevent HIV transmission in efficacy trials. A recent exception was a vaginal gel formulation of tenofovir that reduced HIV infection by 50% [39].

Tenofovir and other topical microbicides can prevent localized infection of target cells in genital mucosae. However, dendritic cells DC-SIGN$^+$ internalize and transport HIV to secondary lymphoid organs, rendering the virus inaccessible to inhibitory effect of the microbicides. Therefore inhibition of DC-SIGN is essential to block HIV-1 uptake and dissemination from migratory dendritic cells.

We have previously reported that the tetravalent Dendron **12** was able to block almost completely the HIV-1 *trans* infection of CD4$^+$ T cells at micromolar range [23]. This compound exerts its activity by competitive inhibition of HIV-1 gp120 binding to DC-SIGN. In our initial experiments we demonstrated that Dendron **12** (at the same concentration previously assayed) even in presence of higher viral loads retains its ability to inhibit HIV-1 *trans* infection. This potent inhibitory activity is due both to elevate affinity for DC-SIGN of Dendron **12** pseudo-trisaccharide units and to high avidity of binding, guaranteed by the tetravalent presentation on the compound scaffold.

Different DC-SIGN inhibitors have been described so far. Dendrons displaying complex oligomannoses in high density inhibited binding of gp120 to DC-SIGN with $IC_{50}$ in nanomolar range [40]. Nevertheless the complexity of the oligosaccharides used limits the possibility of a therapeutic application. Gold nanoparticles displaying mannosyl oligosaccharides are potent inhibitors of DC-SIGN-mediated HIV-1 *trans* infection of human PBMCs [41], but may have toxic effects because of gold accumulation.

Unlike DC-SIGN, the C-type lectin langerin, expressed on Langerhans cells, appears to play a protective role against HIV infection. Langerhans cells are located in the stratified mucosal epithelia of female and male genital tissue. Both DC-SIGN and langerin recognize high mannose glycans on gp120 through their carbohydrate recognition domain (CRD). Recent data demonstrate that langerin prevents HIV transmission by Langerhans cells, at least in the presence of low concentration of the virus, promoting rapid degradation and clearance of HIV-1 [32,33]. Using a biosensor with a SPR detection method we showed that the Dendron **12** is much more selective for DC-SIGN than for langerin. Considering the similarity between the CRD of DC-SIGN and langerin, the fact that the Dendron **12** does not interfere with langerin function is a remarkable advantage in view of developing new microbicides.

The efficacy and the safety of the Dendron **12** were evaluated in a human cervical explant model. Endocervix is more susceptible to HIV-1, being lined by a single layer of columnar epithelium, and contains dendritic cell DC-SIGN$^+$ in the subepithelial region [4,5,24]. Explants were exposed to HIV-1 to mimic a condition of compromised epithelium *in vivo*; condition that highly increases the risk of HIV infection.

R5-tropic virus strains predominate during HIV-1 transmission *in vivo* and cause the majority of new infections [42]. The results obtained showed that Dendron **12** strongly reduced the infection of cervical explants by different HIV-1 R5 tropic strains, such as BaL, the primary isolate 8 g and the Clade C DU174. This may have a great impact, considering that Clade C is the most abundant subtype in all countries of Southern Africa and in some countries of eastern Africa and Asia, areas where the majority of HIV-1-infected people resides [43,44].

Rare cases of HIV-1 infection by X4-tropic strains were observed in CCR5Δ32 homozygotes and X4-tropic strain can infect immune activate human cervical tissue [25,42]. Dendron **12** was able to inhibit DPMVF primary X4-tropic strain infection at higher concentration assayed, but this effect was in part lost at lower concentrations of the compound.

Data obtained suggest that Dendron **12** might have additional mechanisms of action apart from competitive inhibition of HIV-1 binding to DC-SIGN. Flow cytometric studies indicate that treatment with Dendron **12** reduces DC-SIGN expression on B-THP1/DC-SIGN$^+$ cells [23]. This effect may be due to increased internalization of the receptor after binding of the compound to DC-SIGN.

Furthermore, Dendron **12** induced an increase of the production of $\beta$ chemokines, such as MIP-1$\alpha$, MIP-1$\beta$ and RANTES, by iMDDCs. These $\beta$ chemokines, natural ligands of CCR5, suppress HIV-1 R5 tropic strain replication, competing with the binding of the virus to CCR5 or inducing receptor internalization [45–47]. The Dendron **12**, enhancing $\beta$ chemokines production, could interfere, at least partially, with the direct infection of CCR5$^+$ CD4$^+$ T lymphocytes and macrophages located in genital mucosae.

The Dendron **12** does not reduce the vitality of PBMCs and of explants. However, additional experiments, such as rabbit vaginal irritation assay [48], would be needed for a more accurate evaluation of potential toxic effects.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Vaginal epithelium has limited permeability to particles greater than 30 nm [2]. However, the diameter of Dendron **12** is certainly below that threshold, so the compound could enter and diffuse into intact mucosal tissue. Furthermore the Dendron **12** scaffold can be easily modified to improve absorption of the compound, without decreasing affinity to DC-SIGN.

The tetravalent Dendron **12** prevents HIV *trans* infection of $CD4^+$ T lymphocytes at micromolar range, even in presence of elevated viral load, and displays high solubility in physiological media, a neglectable toxicity and a long-lasting effect. Moreover it inhibits in dose-dependent manner HIV-1 infection of human cervical explants. These features make the Dendron **12** a good candidate as a lead compound to develop new microbicide drugs.

However, the Dendron **12** inhibits 80–90% of cervical explants HIV-1 infection, but cannot block it completely. To overcome these limitations, the structure of this compound can be improved in both the scaffold and the active pseudo-saccharide ligand to develop new ligands of DC-SIGN more effective and easier to synthesize. Furthermore, Dendron **12** (or its derivatives) can be used in combination with other molecules directed against different HIV targets. In particular multivalent structures, presenting on the same scaffold multiple copies of DC-SIGN and HIV co-receptor inhibitors, simultaneously blocking different HIV targets, could protect against different routes of HIV transmission.

## Acknowledgements

M.C., A. Bernardi, D.T. conceived the study; A. Berzi, M.C. wrote the paper, A. Berzi, M.B. performed the experiments and analyzed the data; Jo.Re., R.O. synthesized the pseudo-mannotrioside; Ja.Ro., M.S.N. synthesized the Dendron scaffold; F.F., I.S. performed SPR experiments and analyzed the data; E.C. expressed and purified DC-SIGN and langerin; I.C., P.A. obtained cervical explants.

### Conflicts of interest

Conflict of interest and source of funding: the authors declare no conflict of interest.

B-THP-1 and B-THP-1/DC-SIGN cells, HIV-1 BaL and DU174 were provided through the EU programme EVA centre for AIDS Reagents NIBSC, UK. Buffy coats from healthy donors were provided by the Transfusional Unit of Vimercate Hospital (Italy). This work was supported by: the FIRB program CHEM-PROFARMANET (RBPR05NWWC), Istituto Superiore di Sanità 'Programma Nazionale di Ricerca sull'AIDS', EMPRO and AVIP EC WP6 Projects, the nGIN EC WP7 Project, 2008 Ricerca Finalizzata and Ricerca Corrente (Italian Ministry of Health), the MICINN of Spain CTQ2008-01694/BQU European FP7 project: EU ITN CARMUSYS (PITN-GA-2008-213592), Sidaction – ensemble contre le sida. A. Berzi was supported by a fellowship of Doctorate School of Molecular Medicine, University of Milan; I.S. by ITN Carmusys, E.C. by a fellowship from 'ministère de la recherché et de l'enseignement supérieur'; M.S.N. by FPU fellowship.

## References

1. UNAIDS Report on the global AIDS epidemic, 2010.
2. Cutler B, Justman J. **Vaginal microbicides and the prevention of HIV infection.** *Lancet Infect Dis* 2008; **8**:685–697.
3. Reina JJ, Bernardi A, Clerici M, Rojo J. **HIV microbicides: state-of-the-art and new perspectives on the development of entry inhibitors.** *Future Med Chem* 2010; **2**:1141–1159.
4. Geijtenbeek TBH, Torensma R, van Vliet SJ, van Duijnhoven GCF, Adema GJ, van Kooyk Y, *et al.* **Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses.** *Cell* 2000; **100**:575–585.
5. Pope M, Haase AT. **Transmission, cute HIV infection and the quest for strategies to prevent infection.** *Nat Med* 2003; **9**:847–852.
6. Wu L, Wu L, KevalRamani VL. **Dendritic-cell interactions with HIV: infection and viral dissemination.** *Nat Rev Immunol* 2006; **6**:859–868.
7. Geijtenbeek TBH, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GCF, Middel J, *et al.* **DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T Cells.** *Cell* 2000; **100**:587–597.
8. Van Kooyk Y, Geijtenbeeck TBH. **DC-SIGN: escape mechanism for pathogens.** *Nat Rev Immunol* 2003; **3**:697–709.
9. Kwon DS, Gregorio G, Bitton N, Hendrikson WA, Littman DR. **DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection.** *Immunity* 2002; **16**:135–144.
10. McDonald D, Wu L, Bohks SM, KewalRamani VN, Unutmaz D, Hope TJ. **Recruitment of HIV and its receptors to dendritic cell–T cell junctions.** *Science* 2003; **300**:1295–1297.
11. Wiley RD, Gummuluru S. **Immature dendritic cells derived exosomes can mediate HIV-1 trans infection.** *Proc Natl Acad Sci USA* 2006; **103**:738–743.
12. Lee B, Leslie G, Soilleux E, O'Doherty U, Baik S, Levroney E, *et al.* **Cis expression of DC-SIGN allows for more efficient entry of human and simian immunodeficiency viruses via CD4 and a co-receptor.** *J Virol* 2001; **75**:12028–12038.
13. Turville SG, Santos JJ, Frank I, Cameron PU, Wilkinson J, Miranda-Saksena M, *et al.* **Immunodeficiency virus uptake, turnover, and 2-phase transfer in human dendritic cells.** *Blood* 2003; **103**:2170–2179.
14. Gringhuis SI, den Dunnen J, Litjens M, van Het Hof B, van Kooyk Y, Geijtenbeek TBH. **C-type lectin DC-SIGN modulates Toll-like receptor signaling via Raf-1 kinase-dependent acetylation of transcription factor NF-κB.** *Immunity* 2007; **26**: 605–616.
15. Hodges A, Sharrocks K, Edelmann M, Baban D, Moris A, Schwartz O, *et al.* **Activation of the lectin DC-SIGN induces an immature dendritic cell phenotype triggering Rho-GTPase activity required for HIV-1 replication.** *Nat Immunol* 2007; **8**:569–577.
16. Gringhuis SI, van der Vlist M, van den Berg LM, den Dunnen J, Litjens M, Geijtenbeek TBH. **HIV-1 exploits innate signaling by TLR8 and DC-SIGN for productive infection of dendritic cells.** *Nat Immunol* 2010; **11**:419–426.
17. Gringhuis SI, den Dunnen J, Litjens M, van der Vlist M, Geijtenbeek TBH. **Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and *Helicobacter pylori*.** *Nat Immunol* 2010; **10**:1081–1089.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

18. Feinberg H, Mitchell DA, Drickamer K, Weis WI. **Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR.** *Science* 2001; **294**:2163–2166.
19. Adams EW, Ratner DM, Bokesch HR, McMahon JB, O'Keefe BR, Seeberger PH. **Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology; glycan-dependent gp120/protein interactions.** *Chem Biol* 2004; **11**:875–881.
20. Reina JJ, Sattin S, Invernizzi D, Mari S, Martínez-Prats L, Tabarani G, *et al.* **1,2 Mannobioside mimic: synthesis, DC-SIGN interaction by NMR and docking, and antiviral activity.** *Chem MedChem* 2007; **2**:1030–1036.
21. Rojo J, Delgado R. **Glycodendritic structures: promising new antiviral drugs.** *J Antimicrob Chemother* 2004; **54**:579–581.
22. Luczkowiak J, Sattin S, Sutkevičiūte I, Reina JJ, Sánchez-Navarro M, Thépaut M, *et al.* **Pseudosaccharide functionalized dendrimers as potent inhibitors of DC-SIGN dependent Ebola pseudotyped viral infection.** *Bioconj Chem* 2011; **22**:1354–1365.
23. Sattin S, Daghetti A, Thépaut M, Berzi A, Sanchez-Navarro M, Rojo J, *et al.* **Inhibition of DC-SIGN-mediated infecion by a linear trimannoside mimic in a tetravalent presentation.** *ACS Chem Biol* 2010; **5**:301–312.
24. Anderson DJ, Pudney J, Shust DJ. **Caveat associated with the use of human cervical tissue for HIV and microbicide research.** *AIDS* 2010; **24**:1–4.
25. Greenhead P, Hayes P, Watts P, Laing K, Griffin G, Shattock R. **Parameters of Human Immunodeficiency Virus Infection of human cervical tissue and inhibition by vaginal virucides.** *J Virol* 2000; **74**:5577–5586.
26. Wallace GS, Cheng-Mayer C, Schito ML, Fletcher P, Miller Jenkins LM, Hayashi R, *et al.* **Human immunodeficiency virus type 1 nucleocapsid inhibitors impede trans infection in cellular and explant models and protect nonhuman primates from infection.** *J Virol* 2009; **83**:9175–9182.
27. Cummins JE, Guarner J, Flowers L, Guenthner PC, Bartlett J, Morken T, *et al.* **Preclinical testing of candidate topical microbicides for antihuman immunodeficiency virus type 1 activity and tissue toxicity in a human cervical explant culture.** *Antimicrob Agents Chemother* 2007; **51**:1770–1779.
28. Thépaut M, Valladeau J, Nurisso A, Kahn R, Arnou B, Vivès C, *et al.* **Structural studies of Langerin and Birbeck granule: a macromolecular organization model.** *Biochemistry* 2009; **48**:2684–2698.
29. Tabarani G, Thépaut M, Stroebel D, Ebel C, Vivès C, Vachette P, *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.
30. Halary F, Amara A, Lortat-Jacob H, Messerle M, Delaunay T, Houlès C, *et al.* **Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection.** *Immunity* 2002; **17**:653–664.
31. Andreini M, Doknic D, Sutkeviciute I, Reina JJ, Duan J, Chabrol E, *et al.* **Second generation of fucose-based DC-SIGN ligands: affinity improvement and specificity versus Langerin.** *Org Biomol Chem* 2011; **9**:5778–5786.
32. Wu L, Martin TD, Carrington M, KewalRamani VN. **Raji B cells, misidentified as THP-1 cells, stimulate DC-SIGN- mediated HIV transmission.** *Virology* 2004; **318**:17–23.
33. de Witte L, Nabatov A, Pion M, Fluitsma D, Marein AW, de Jong P, *et al.* **Langerin is natural barrier to HIV-1 transmission by langerhans cells.** *Nat Med* 2007; **13**:367–371.
34. de Witte L, Nabatov A, Geijtenbeek TBH. **Distinct roles for DC-SING-dentritic cells and langherans cells in HIV-1 transmission.** *Trends Mol Med* 2007; **14**:12–19.
35. Guo Y, Feinberg H, Conroy E, Mitchell DA, Alvarez R, Blixt O, *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**:591–598.
36. Galustian C, Park CG, Chai W, Kiso M, Bruening SA, Kang YS, *et al.* **High and low affinity carbohydrate ligands revealed for murine SIGN-R1 by carbohydrate array and cell binding approaches, and differing specificities for SIGN-R3 and Langerin.** *Int Immunol* 2004; **16**:853–866.
37. Timpano G, Tabarani G, Anderluh M, Invernizzi D, Vasile F, Potenza D, *et al.* **Synthesis of novel DC-SIGN ligands with an alpha-fucosylamide anchor.** *ChemBioChem* 2008; **9**:1921–1930.
38. Sallusto F, Lanzavecchia A. **Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha.** *J Exp Med* 1994; **179**:1109–1118.
39. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, *et al.* **Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women.** *Science* 2010; **329**:1168–1174.
40. Wang SK, Liang PH, Astronomo RD, Hsu TL, Hsieh SL, Burton DR, *et al.* **Targeting the carbohydrates on HIV-1: interaction of oligomannose dendrons with human monoclonal antibody 2G12 and DC-SIGN.** *Proc Natl Acad Sci USA* 2008; **105**:3690–3695.
41. Martinez-Avila O, Bedoya LM, Marradi M, Clavel C, Alcami J, Penades S. **Multivalent manno-glyconanoparticles inhibit DC-SIGN-mediated HIV-1 trans-infection of human T cells.** *Chem Biochem* 2009; **10**:1806–1809.
42. Nazari R, Joshi S. **CCR5 as target for HIV-1 gene therapy.** *Curr Gene Ther* 2008; **8**:1–9.
43. McCutchan FE. **Understanding the genetic diversity of HIV-1.** *AIDS* 2000; **14 (Suppl 3)**:S31–S44.
44. Spira S, Wainberg MA, Loemba H, Turner D, Brenner BG. **Impact of clade diversity on HIV-1 virulence, antiretroviral drug sensitivity and drug resistance.** *J Antimicrob Chemother* 2003; **51**:229–240.
45. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, *et al.* **Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene.** *Nature* 1996; **26**:722–725.
46. Wu L, LaRosa G, Kassam N, Gordon CJ, Heath H, Ruffing N, *et al.* **Interaction of chemokine receptor CCR5 with its ligands: multiple domains for HIV-1 gp120 binding and a single domain for chemokine binding.** *J Exp Med* 1997; **186**:1373–1381.
47. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC, Lusso P. **Identification of RANTES, MIP-1(, and MIP-1( as the major HIV-1-suppressive factors produced by CD8 T cells.** *Science* 1995; **270**:1811–1815.
48. Eckstein P, Jackson MC, Millman N, Sobrero AJ. **Comparison of vaginal tolerance tests of spermicidal preparations in rabbits and monkeys.** *J Reprod Fertil* 1969; **20**:85–93.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

#### 9.1.4.2. *Evaluation of psDi and psTri as DC-SIGN antagonists presented in monovalent and multivalent manner using Ebola infection inhibition and SPR competition assays*

The glycodendritic structures based on hyperbranched Boltorn type dendrimers (fig. 9.15) functionalized with D-mannose (BH30sucMan) have been previously explored as DC-SIGN inhibitors by our group with our collaborators Dr. J. Rojo and Dr. R. Delgado [157, 158]. The following article by Luczkowiak et al. for the first time addresses these dendritic structures functionalized with glycomimetic ligands, i.e. psDi and psTri (fig. 9.14) as Ebola infection inhibitors.

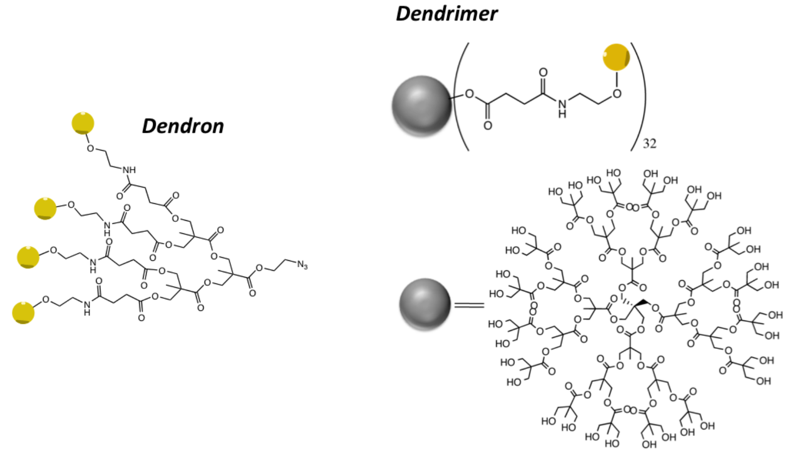


**Figure 9.15.:** The structures of the scaffolds of tetravalent dendron and multivalent Boltorn type dendrimer.

The yellow spheres mark attachment points of psDi or psTri (fig. 9.14 on page 173). The grey sphere represents Boltorn type $3^{\text{rd}}$ generation dendrimer with the shown structure. The dendrimers with psDi and psTri are further called G3(pseudodi)$_{32}$ and G3(pseudotri)$_{32}$, respectively.

Indeed, such glycodendrimers have an order of magnitude higher potency (IC$_{50}$ values range 0.02 µM – 0.06 µM) to inhibit DC-SIGN-mediated Ebola infection than the analogous BH30sucMan dendrimer (IC$_{50}$ 0.3 µM [157]). The tetravalent dendrons with psDi and psTri tethers were also tested, and as expected, had lower potencies to inhibit DC-SIGN-mediated Ebola infection.

The same compounds were analyzed in SPR competition assay. The absolute IC$_{50}$ values from SPR competition assay were 2 orders of magnitude higher than values in Ebola infection inhibition assay, however, the general pattern of IC$_{50}$ values for the compounds corresponded in both types of experiments.

The main outcome of these studies:

- The multivalent glycomimetic compounds were tested as the inhibitors of another type pathogen, the Ebola virus, and proved to be efficient in inhibiting the infection.
- psTri showed a discrepancy when comparing results for monovalent psDi and psTri compounds and the multivalent forms.

Contributions:

In the article presented dendrons and dendrimers were synthesized by the group of Dr. J. Rojo. The assays of DC-SIGN-mediated Ebola infection in *cis* and in *trans* were performed in the group of Dr. R. Delgado by a PhD student Joanna Luczkowiak. DC-SIGN used in SPR assay was produced in our team by Dr. Michel Thépaut.

My contribution to this study:

I have performed all of the described SPR assays, analyzed the data, and participated in the preparation of the paper manuscript.

3

# 3

*Paper n$^o$3:* **Pseudosaccharide Functionalized Dendrimers as Potent Inhibitors of DC-SIGN Dependent Ebola Pseudotyped Viral Infection**

Bioconjugate Chemistry
ARTICLE
pubs.acs.org/bc

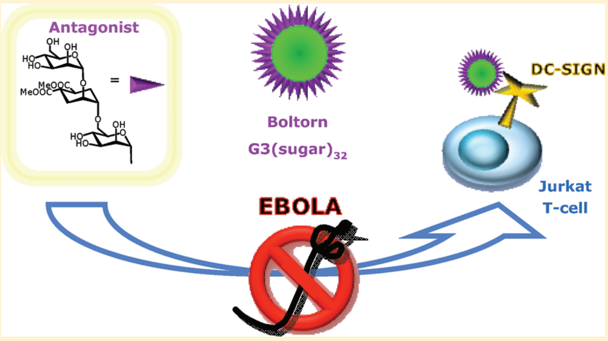
# Pseudosaccharide Functionalized Dendrimers as Potent Inhibitors of DC-SIGN Dependent Ebola Pseudotyped Viral Infection

Joanna Luczkowiak,[§] Sara Sattin,[‡] Ieva Sutkevičiūtė,[∥,#] José Juan Reina,[‡] Macarena Sánchez-Navarro,[†] Michel Thépaut,[∥,⊥] Lorena Martínez-Prats,[§] Anna Daghetti,[‡] Franck Fieschi,[*,∥,#] Rafael Delgado,[*,§] Anna Bernardi,[*,‡,¶] and Javier Rojo[*,†]

[†]Glycosystems Laboratory, Instituto de Investigaciones Químicas, CSIC — Universidad de Sevilla, Américo Vespucio 49, 41092 Seville, Spain
[‡]Università degli Studi di Milano, Dipartimento di Chimica Organica e Industriale, and CISI, Milano, Italy
[§]Laboratorio de Microbiología Molecular, Instituto de Investigación Hospital 12 de Octubre (imas12), 28041 Madrid, Spain
[∥]Institut de Biologie Structurale, CNRS, UMR 5075, 41 rue Jules Horowitz, 38027 Grenoble France
[⊥]CEA, DSV, 38027 Grenoble France
[#]Université Joseph Fourier, Institut Universitaire de France, 38000 Grenoble, France
[¶]CNR-ISTM, Institute of Molecular Sciences and Technologies, Milano, Italy

S *Supporting Information*

**ABSTRACT:** The development of compounds with strong affinity for the receptor DC-SIGN is a topic of remarkable interest due to the role that this lectin plays in several pathogen infection processes and in the modulation of the immune response. DC-SIGN recognizes mannosylated and fucosylated oligosaccharides in a multivalent manner. Therefore, multivalent carbohydrate systems are required to interact in an efficient manner with this receptor and compete with the natural ligands. We have previously demonstrated that linear pseudodi- and pseudotrisaccharides are adequate ligands for DC-SIGN. In this work, we show that multivalent presentations of these glycomimetics based on polyester dendrons and dendrimers lead to very potent inhibitors (in the nanomolar range) of cell infection by Ebola pseudotyped viral particles by blocking DC-SIGN receptor. Furthermore, SPR model experiments confirm that the described multivalent glycomimetic compounds compete in a very efficient manner with polymannosylated ligands for binding to DC-SIGN.



## INTRODUCTION

Carbohydrates are involved in several biological processes including cell-differentiation, migration, tumor progression and metastasis, inflammation, pathogen infection, and so forth.[1] Carbohydrates participate in these events through complex and selective recognition processes triggered by interaction with receptors.

One of the processes that initialize immune response to pathogen invasion is the interaction of the pathogen surface glycans with C-type lectin receptors expressed on the antigen-presenting cells (APCs), such as dendritic cells (DCs). Although usually individual carbohydrate unit binding to lectin is weak ($K_D$ in millimolar range), the clustered organization of these receptors as well as high glycosylation of their ligands create the conditions for more specific multivalent interactions, and thus overcome low affinity problems.[2−5] The use of scaffolds to prepare carbohydrate multivalent systems is nowadays a very useful and common approach to develop tools to understand and intervene in these biological processes where carbohydrates play a key role.[6,7]

The lack of design in these multivalent systems is partly a result of the often scarce information about the structural details of the multivalent presentation of the corresponding receptors at the cell surface. However, even if these details were known, topological design of large, multivalent molecules would still not be straightforward and the valency and 3D structures required to achieve a strong cluster effect on a given receptor cannot be estimated a priori. These facts explain the huge number of examples illustrating different strategies to target cellular receptors in a multivalent manner. In this context, our research is focused on the development of glycodendrimers to block a C-type lectin, DC-SIGN, which recognizes glycoconjugates present on the surfaces of several pathogens including viruses (HIV, Ebola, Cytomegalovirus, Dengue, SARS), bacteria (*M. tuberculosis*, *S. pneumoniae*) fungi (*C. albicans*, *A. fumigatus*), and

**Received:** January 21, 2011
**Revised:** March 31, 2011
**Published:** June 08, 2011

ACS Publications © 2011 American Chemical Society

parasites (*Leishmania, S. mansoni*).[8−14] It has been proven that this lectin plays a key role in the early stages of the infection processes caused by some of these pathogens. Therefore, DC-SIGN can be selected as a new therapeutic target for the design of antiviral drugs.[15−18] Previous results from our groups indicate that glycodendrimers based on Boltorn-type structures bearing mannoses exhibit a strong antiviral activity in a pseudotype Ebola viral particles infection model both for in *cis* and in *trans* infection.[19−21] Additionally, we have demonstrated that carbohydrate mimics such as a pseudomannobioside present an activity around 1 order of magnitude higher than mannose in such infection studies.[22,23] These new glycomimetics are more stable against enzymatic degradation than the corresponding natural counterparts, and therefore, they are more adequate compounds to be used in clinical applications.

In this study, we combined the improved binding properties of these new pseudomannoside ligands with the potency of multivalent presentation on Boltorn scaffolds. Indeed, nanomolar inhibition levels have been reached in a pseudotyped Ebola viral particles infection model. Binding properties between these compounds and their receptor, DC-SIGN, have been analyzed at the molecular level using biosensors with surface plasmon resonance (SPR) detection method. Altogether, this work opens road to new compounds in antiviral strategy.

## ■ MATERIALS AND METHODS

Reagents were purchased from Sigma-Aldrich, Senn Chemicals, Flucka and were used without purification. Solvents were dried by standard procedures. Reactions requiring anhydrous conditions were performed under nitrogen. Synthetic compounds were purified by flash chromatography using medium or fine silica gel or by Sephadex (LH20, G25). Thin layer chromatography (TLC) was carried out with precoated Merck $F_{254}$ silica gel plates. Flash chromatography (FC) was carried out with Macherey-Nagel silica gel 60 (230−400 mesh). Reaction completion was observed by TLC using as development reagents phosphomolibdic acid, 10% sulfuric acid in methanol or anisaldehyde. $^1$H and $^{13}$C spectra were recorded at 400 MHz on Bruker Avance DPX 300, DRX 400, and DRX 500 MHz spectrometers. Chemical shifts ($\delta$) for $^1$H and $^{13}$C spectra are expressed in ppm relative to internal TMS (tetramethylsilane) using manufacturer indirect referencing method. Signals were abbreviated as s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; and m, multiplet. Mass spectra were obtained with a Bruker ion-trap Esquire 3000 or 6000 apparatus (ESI ionization) and Microflex apparatus (MALDI ionization) from Bruker. HRMS (FT-ICR, ESI) were obtained with an Apex II instrument. Infrared spectra were recorded with a Bruker Vector 22 FT-IR spectrometer. Elemental analyses were obtained with a Leco CNHS instrument. All hydrogenation reactions were carried out under $H_2$ atmosphere at atmospheric pressure.

**Synthesis of Compounds.** Compounds **G1(Bn)$_4$**,[24] **G1(OH)$_8$**,[25] **1**,[26] **2**,[27] **3**,[28] **6**,[29] **11**,[23] **13**,[22] and **14**−**15**[23] were prepared as described previously in the literature.

**Synthesis of G2(Bn)$_8$.** **G1(OH)$_8$** (100 mg, 0.17 mmol) and DMAP (70 mg) were dissolved in dry pyridine (1 mL) and then diluted with $CH_2Cl_2$ (2 mL). Anhydride **1** (812 mg, 2.04 mmol) was added and the reaction mixture was stirred at room temperature for 15 h. Then, the excess of anhydride **1** was quenched by stirring the reaction mixture with 1 mL of pyridine/$H_2O$ (1:1) solution overnight. The organic phase was diluted with $CH_2Cl_2$ (100 mL), and washed with $NaHSO_4$ 1 M (2 × 40 mL), $Na_2CO_3$ (10%) (2 × 40 mL), and NaCl sat. (40 mL). The organic phase was dried using $MgSO_4$ anh., and the solvent was evaporated under vacuum. The residue was recrystallized from $CH_2Cl_2$/hexane to give **G2(Bn)$_8$** (370 mg, 97%) as white crystals. $^1$H NMR (300 MHz, $CDCl_3$): $\delta$ (ppm) 7.40−7.38 (m, 16 H, $H_{Ar}$), 7.32−7.26 (m, 30H, $H_{Ar}$), 5.36 (s, 8H, 8 × $CH_{Bn}$), 4.53 (d, 16H, $J$ = 11.7 Hz, $CH_{2G2}$), 4.36 (d, 8H, $J$ = 11.4 Hz, $CH_{2G1}$), 4.29 (d, 8H, $J$ = 11.4 Hz, $CH_{2G1}$), 3.89 (s, 8H, $CH_{2Pentaerythritol}$), 3.52 (d, 16H, $J$ = 11.7 Hz, $CH_{2G2}$), 1.15 (s, 12H, $CH_{3G1}$), 0.88 (s, 24H, $CH_{3G1}$); $^{13}$C NMR (100 MHz, $CDCl_3$) $\delta$ (ppm) 175.9 ($\mathrm{COO_{G2}}$), 174.1 ($\mathrm{COO_{G1}}$), 140.2 ($\mathrm{C_{Ph}}$), 131.2 ($\mathrm{C_{Ph}}$), 130.4 ($\mathrm{CH_{Ph}}$), 128.5 ($\mathrm{CH_{Ph}}$), 104.0 ($\mathrm{CH_{Bn}}$), 75.7, 75.7, 67.2, 63.6, 49.2, 44.8, 36.2, 20.0 ($\mathrm{CH_{3G2}}$), 19.9 ($\mathrm{CH_3\ _{G1}}$); ESI-MS for $C_{121}H_{140}O_{40}$; calcd 2232.9 $[M]^+$; found 2256.7 $[M+Na]^+$ and 1139.8 $[M+2Na]^{2+}$; Elemental analysis calcd (%) for $C_{121}H_{140}O_{40}$: C, 65.04%; H, 6.32%; found: C, 65.32%; H, 6.31%.

**Synthesis of G2(OH)$_{16}$.** **G2(Bn)$_8$** (3.8 g, 1.70 mmol) was dissolved in $CH_2Cl_2$ (50 mL) and diluted with MeOH (50 mL). A catalytic amount of Pd(C) was added to the solution and the reaction mixture was hydrogenated under vigorous stirring for 12 h. Then, the reaction mixture was filtered over a pad of Celite. The solvent was evaporated to give **G2(OH)$_{16}$** (860 mg, 95%) as a white solid. $^1$H NMR (300 MHz, $CD_3OD$): $\delta$ (ppm) 4.39 (d, 8H, $J$ = 11.1 Hz, $CH_{2G1}$), 4.30 (d, 8H, $J$ = 11.1 Hz, $CH_{2G1}$), 4.29 (s, 8H, $CH_{2Pentaerythritol}$), 3.70 (d, 16H, $J$ = 10.8 Hz, $CH_{2G2}$), 3.61 (d, 16H, $J$ = 10.8 Hz, $CH_{2G2}$), 1.35 (s, 12H, $CH_{3G1}$), 1.17 (s, 24H, $CH_{3G2}$); $^{13}$C NMR (75 MHz, $CD_3OD$): $\delta$ (ppm) 176.0 ($\mathrm{COO_{G2}}$), 173.8 ($\mathrm{COO_{G1}}$), 66.1 ($\mathrm{CH_{2Pentaerythritol}}$), 65.9 ($\mathrm{CH_{2G2}}$), 63.6 ($\mathrm{C_{G2}}$), 51.8 ($\mathrm{CH_{2G1}}$), 44.3 ($\mathrm{C_{G1}}$), 18.3 ($\mathrm{CH_{3G1}}$), 17.4 ($\mathrm{CH_{3G2}}$); ESI-MS for $C_{65}H_{108}O_{40}$; calcd: 1528.6 $[M]^+$; found: 1551.4 $[M+Na]^+$.

**Synthesis of G3(Bn)$_{16}$.** **G2)OH)$_{16}$** (1.00 g, 0.654 mmol) and DMAP (445 mg) were dissolved in dry pyridine (20 mL) and then diluted with $CH_2Cl_2$ (40 mL). Anhydride **1** (6.02 g, 14.13 mmol) was added, and the reaction mixture was stirred at room temperature for 15 h. Then, the excess of anhydride **1** was quenched by stirring the reaction mixture with 20 mL of pyridine/$H_2O$ (1:1) solution overnight. The organic phase was diluted with $CH_2Cl_2$ (300 mL), and washed with $NaHSO_4$ 1 M (2 × 150 mL), $Na_2CO_3$ (10%) (2 × 150 mL), and NaCl sat. (150 mL). The organic phase was dried using $MgSO_4$ anh., and the solvent was evaporated under vacuum. The residue was recrystallized from $CH_2Cl_2$/MeOH to give **G3(Bn)$_{16}$** (2.7 g, 97%) as white crystals. $^1$H NMR (500 MHz, $CDCl_3$): $\delta$ (ppm) 7.37−7.34 (m, 32H, $CH_{Ph}$), 7.27−7.21 (m, 48H, $CH_{Ph}$), 5.32 (s, 16H, $CH_{Bn}$), 4.49 (d, 32H, $CH_{2G3}$), 4.34−4.25 (m, 32H, 2 × 16 $CH_{2G2}$), 4.12 (d, 8H, $\mathrm{CH2_{G1}}$), 4.11 (s, 8H, $CH_{2Pentaerythritol}$), 4.02 (d, 8H, $CH_{2G1}$), 3.55−3.51 (m, 32H, $CH_{2G3}$), 1.18 (s, 24H, $CH_{3G2}$), 1.08 (s, 12H, $CH_{3G1}$), 0.88 (s, 48H, $CH_{3G3}$); $^{13}$C NMR (125 MHz, $CDCl_3$): $\delta$ (ppm) 173.2 ($\mathrm{COO_{G3}}$), 171.9 ($\mathrm{COO_{G2}}$), 171.5 ($\mathrm{COO_{G1}}$), 138.0 ($\mathrm{C_{Ph}}$), 128.8 ($\mathrm{CH_{Ph}}$), 128.1 ($\mathrm{CH_{Ph}}$), 126.2 ($\mathrm{CH_{Ph}}$), 101.6 ($\mathrm{CH_{Bn}}$), 73.4, 73.3, 64.9, 46.9, 46.6, 42.5 ($\mathrm{CH_{2G3}}$), 34.0, 25.0, 17.7 ($\mathrm{CH_{3G2}}$), 17.6 ($\mathrm{CH_{3G3}}$), 17.2 ($\mathrm{CH_{3G1}}$); ESI-MS for $C_{257}H_{300}O_{88}$; calcd: 4793.9 $[M]^+$; found: 2420.5 $[M+2Na]^{2+}$ and 1621.8 $[M+3Na]^{3+}$; Elemental analysis calcd (%) for $C_{257}H_{300}O_{88}$: C, 64.34%; H, 6.30%; found: C, 64.07%; H, 6.46%.

**Synthesis of G3(OH)$_{32}$.** **G3(Bn)$_{16}$** (1.7 g, 0.354 mmol) was dissolved in $CH_2Cl_2$ (30 mL) and diluted with MeOH (30 mL). A catalytic amount of Pd(C) was added to the solution and the reaction mixture was hydrogenated under vigorous stirring for 12 h. Then, the reaction mixture was filtered over a pad of Celite. The solvent was evaporated to give **G3(OH)$_{32}$** (1.14 g, 95%) as a white solid. $^1$H NMR (500 MHz, $CD_3OD$): $\delta$ (ppm) 4.36−4.22 (m, 56H, $CH_{2Pentaerythritol} + CH_{2G1} + CH_{2G2}$), 3.68 (m, 32H, $CH_{2G3}$), 3.58 (d, 32H, $J$ = 10.0 Hz, $CH_{2G3}$), 3.39 (s, 32H, $CH_2OH$), 1.34 (s, 12H,

$CH_{3G1}$), 1.30 (s, 24H, $CH_{3G2}$), 1.12 (s, 48H, $CH_{3G3}$); $^{13}$C NMR (75 MHz, $CD_3OD$): δ (ppm) 176.0 ($COO_{G2}$), 173.8 ($COO_{G1}$), 66.1 ($CH_{2Pentaerythritol}$), 65.9 ($CH_{2G2}$), 63.6 ($C_{G2}$), 51.8 ($CH_{2G1}$), 44.3 ($C_{G1}$), 18.3 ($CH_{3G1}$), 17.8 ($CH_{3G3}$), 17.4 ($CH_{3G2}$); ESI-MS for $C_{65}H_{108}O_{40}$; calcd: 3385.4 $[M]^+$; found: 1716.0 $[M+2Na]^{2+}$ and 1151.0 $[M+3Na]^{3+}$.

**Synthesis of G3(sucBn)$_{32}$.** (460 mg, 1.36 mmol), succinic acid derivate **2** (1.086 g, 5.22 mmol) and DPTS (512 mg, 1.74 mmol) were disolved in $CH_2Cl_2$ (15 mL) under argon atmosphere. DCC (1.433 g, 6.96 mmol) was added and the reaction mixture was stirred for 24 h at room temperature. Then, the reaction mixture was diluted with (6 mL), filtered over a pad of Celite, and the solvent was evaporated. The residue was purified by flash chromatography on silica gel (Hex-AcOEt, 1:1.2), to afford **G3(sucBn)$_{32}$** (1.10 g, 81%) as an oil. $^1$H NMR (500 MHz, $CDCl_3$): δ (ppm) 7.41−7.33 (m, 160H, $CH_{Ar}$), 5.13 (bs, 64H, $CH2_{Bn}$), 4.50−4.21 (m, 120H, $CH2_{dendrimer}$), 2.63 (bs, 128H, $CH_{2suc}$), 1.39 (bs, 12H, $CH_{3G1}$), 1.31 (bs, 24H, $CH_{3G2}$), 1.23 (bs, 48H, $CH_{3G3}$); $^{13}$C NMR (125 MHz, $CDCl_3$) δ (ppm) 171.9 ($COO_{G3}$), 171.7 ($COO_{G2}$), 171.6 ($COO_{G1}$), 135.9 ($C_{ArBn}$), 128.5 ($CH_{ArBn}$), 128.4 ($CH_{ArBn}$), 128.2 ($CH_{ArBn}$), 66.4 ($CH_{2Bn}$), 65.1, 64.9 ($CH_{2dendrimer}$), 46.7 ($C_{G1}$), 46.6 ($C_{G2}$), 46.3 ($C_{G3}$), 28.9 ($CH_{2suc}$), 28.8 ($CH_{2suc}$), 17.7 ($CH_{3G3}$), 17.6 ($CH_{3G2}$), 17.4 ($CH_{3G1}$); ESI-MS for $C_{497}H_{556}O_{184}$; calcd: 9467.4 $[M]^+$; found: 4758.5 $[M+2Na]^{2+}$ and 3185.3 $[M+3Na]^{3+}$; Elemental analysis calcd (%) for $C_{497}H_{556}O_{184}$: C, 63.01%; H, 5.92%; found: C, 62.54%; H, 6.00%.

**Synthesis of G3(suc)$_{32}$.** **G3(sucBn)$_{32}$** (87 mg, $9.19 \times 10^{-3}$ mmol) was dissolved in $CH_2Cl_2$ (3 mL) and diluted with MeOH (6 mL). A catalytic amount of Pd(C) was added to the solution and the reaction mixture was hydrogenated under vigorous stirring for 12 h. Then, the reaction mixture was filtered over a pad of Celite. The solvent was evaporated to give **G3(suc)$_{32}$** (79 mg, quant.) as a white solid. $^1$H NMR (300 MHz, $D_2O$): δ (ppm) 4.50−4.20 (m, 120, 60 × $CH_2OCO$), 2.70−257 (m, 128H, 64 × $CH_{2suc}$), 1.42 (s, 12H, 4 × $CH_{3G1}$), 1.35 (s, 32H, 8 × $CH_{3G2}$), 1.28 (s, 52H, 16 × $CH_{3G3}$); $^{13}$C NMR (125 MHz, $CD_3CN$): δ (ppm) 174.2 (COO), 173.0 (COO), 172.9 (COO), 66.3 ($CH_2O$), 47.4 (C), 29.1 ($CH_{2suc}$), 17.9 ($CH_3$).

**G3(pseudodi)$_{32}$.** To a solution of **G3(suc)$_{32}$** (3.5 mg, $5.31 \times 10^{-4}$ mmol, 1 equiv) in dry DMA (100 μL) under nitrogen atmosphere, HATU (13 mg, 0.034 mmol, 64 equiv) and DIPEA (12 μL, 0.068 mmol, 128 equiv) were added. After 15 min a solution of **13** (16.0 mg, 0.037 mmol, 69 equiv) in dry DMA (170 μL) was added. The reaction was stirred at room temperature for 1 day. MALDI mass analysis of a sample of the reaction mixture showed the completion of the reaction. The reaction mixture was diluted in methanol and charged directly onto a Sephadex LH-20 column in methanol. Slow elution led to the purification of the product **G3(pseudodi)$_{32}$** that was isolated in good yield (9.2 mg, 88%). $^1$H NMR (400 MHz, $D_2O$): δ (ppm) 5.00 (s, $H_{1M}$), 4.43−4.14 (m, $CH_2O_{G1} + CH_2O_{G2} + CH_2O_{G3}$), 4.02−3.95 (m, $H_{D2} + H_{2M}$), 3.92 (s, $H_{2M}$), 3.86 (d, $J$ = 11.2 Hz, $H_{6MB}$), 3.81 (dd, $J$ = 3.2 and 8.8 Hz, $H_{3M}$), 3.78−3.56 (m, $H_{6MA} + H_{D1} + H_{4M} + H_{5M} + H_7$), 3.69 (s, $COOCH_3$), 3.69 (s, $COOCH_3$), 3.42−3.32 (m, $H_8$), 2.97−2.81 (m, $H_{D4} + H_{D5}$), 2.72−2.61 (m, $CH_2COO$), 2.61−2.49 (m, $CH_2CONH$), 2.16−2.03 (m, $H_{D6eq} + H_{D3eq}$), 1.89−1.71 (m, $H_{D6ax} + H_{D3ax}$), 1.39−1.18 (m, $CH_{3G1} + CH_{3G2} + CH_{3G3}$); $^{13}$C NMR (100 MHz, $D_2O$): 177.1, 176.8, 173.4 ($COOCH_3$, $COOCH_2$, CONH), 98.7 ($C_{1M}$), 73.9 ($C_{D1}$), 73.6 ($C_{4M}$), 71.1 ($C_{D2}$), 70.6 ($C_{2M} + C_{3M}$), 67.1 ($C_7$), 66.9 ($C_{5M}$), 65.5 ($CH_2O_{G1} + CH_2O_{G2} + CH_2O_{G3}$), 61.1 ($C_{6M}$), 52.5 ($CH_3O$), 46.5 ($C_{quatG1} + C_{quatG2} + C_{quatG3}$), 39.5 ($C_8$), 39.1 ($C_{D4}$, $C_{D5}$), 30.2 ($CH_2CONH$), 29.3 ($CH_2COO$), 27.3, 26.9 ($C_{D6}$, $C_{D3}$), 17.3 ($CH_{3G1} + CH_{3G2} + CH_{3G3}$); MALDI-ToF MS (matrix, SA): distribution centered on 31 sugar loaded, for 31 sugars $[M+Na]^+$ calculated: 19614.8, found: 19689.5.

**G3(pseudotri)$_{32}$.** To a solution of **G3(suc)$_{32}$** (7.6 mg, $1.15 \times 10^{-3}$ mmol, 1 equiv) in 100 μL of dry DMA under nitrogen atmosphere, HATU (28.1 mg, 0.074 mmol, 64 equiv) and DIPEA (26 μL, 0.148 mmol, 128 equiv) were added. After 15 min, a solution of **15** (35.7 mg, 0.06 mmol, 52 equiv) in dry DMA (230 μL) was added. The reaction was stirred at room temperature for 2 days. MALDI mass analysis of a sample of the reaction mixture showed the completion of the reaction. The reaction mixture was diluted in methanol and charged directly onto a Sephadex LH-20 column in methanol. Slow elution led to the purification of the product **G3(pseudotri)$_{32}$** that was isolated in good yield (25.5 mg, 88% yield). $^1$H NMR (400 MHz, $D_2O$): δ (ppm) 5.00 (s, $H_{1M'}$), 4.84 (s, $H_{1M}$), 4.40−4.10 (m, $CH_2O_{G1} + CH_2O_{G2} + CH_2O_{G3}$), 4.04 (s, $H_{D2}$), 3.97 (bs, $H_{2M'}$), 3.92 (s, $H_{2M}$), 3.89−3.53 (m, $H_{3M'} + H_{6M'B} + H_{6MB} + H_{D1} + H_{6M'A} + H_{6MA} + H_{3M} + H_{4M} + H_{4M'} + H_7 + H_{5M} + H_{5M'}$), 3.69 (s, $COOCH_3$), 3.68 (s, $COOCH_3$), 3.49−3.34 (m, $H_8$), 3.00−2.85 (m, $H_{D4} + H_{D5}$), 2.73−2.59 (m, $CH_2COO$), 2.59−2.46 (m, $CH_2CONH$), 2.27−2.03 (m, $H_{D6eq} + H_{D3eq}$), 1.89−1.71 (m, $H_{D6ax} + H_{D3ax}$), 1.40−1.15 (m, $CH_{3G1} + CH_{3G2} + CH_{3G3}$); $^{13}$C NMR (100 MHz, $D_2O$): δ (ppm) 177.3, 177.0, 173.5, 173.0 ($COOCH_3$, $COOCH_2$, CONH), 99.9 ($C_{1M}$), 98.7 ($C_{1M'}$), 74.7 ($C_{D1}$), 73.5, 72.2 ($C_{4M} + C_{4M'}$), 71.0 ($C_{D2}$), 70.9, 70.7, 70.6 ($C_{2M'} + C_{3M} + C_{3M'}$), 70.2 ($C_{2M}$), 68.2 ($C_{6M}$), 67.1, 66.9 ($C_{5M} + C_{5M'}$), 66.2 ($C_7$), 65.4 ($CH_2O_{G1} + CH_2O_{G2} + CH_2O_{G3}$); 61.1 ($C_{6M'}$), 52.6 ($CH_3O$), 46.4 ($C_{quatG1} + C_{quatG2} + C_{quatG3}$), 39.1, ($C_{D4}$, $C_{D5}$), 39.0 ($C_8$), 30.1 ($CH_2CONH$), 29.2 ($CH_2COO$), 27.2, 26.9 ($C_{D6}$, $C_{D3}$), 17.3 ($CH_{3G1} + CH_{3G2} + CH_{3G3}$). MALDI-ToF MS (matrix: SA): distribution centered on 31 sugar loaded, for 31 sugar $[C_{1041}H1_{612}N_{32}O_{664}Na]^+$ calcd: 25222.75, found: 25227.9.

**2-Azidoethyl Isopropylidene-2,2-bis(oxymethyl)propionate (4).** 2-Bromoethanol (1.2 mL, 17.0 mmol) and DMAP (0.325 g, 2.6 mmol) were dissolved in pyridine (25 mL), followed by the addition of $CH_2Cl_2$ (70 mL). The anhydride of isopropylidene-2,2-bis(oxymethyl)propionic acid **6** (7.4 g, 22.4 mmol) was added slowly. The solution was stirred at room temperature overnight and then was quenched with water (1 mL) under vigorous stirring, followed by dilution with of $CH_2Cl_2$ (200 mL), and the solution was washed with 10% of $NaHSO_4$ (3 × 40 mL), 10% of $Na_2CO_3$ (3 × 40 mL), and brine (20 mL). The organic phase was dried with $MgSO_4$, filtered, and concentrated. The crude product was purified by flash chromatography on silica gel, eluting with hexane (100 mL) and gradually increasing the polarity to EtOAc/hexane (9:1; 6:1) to give 2-bromoethyl isopropylidene-2,2-bis(oxymethyl)propionate as colorless oil (4.9 g, 86%). This bromo derivative (4.9 g, 14.6 mmol) and $NaN_3$ (2.6 g, 40.0 mmol) were dissolved in dry DMF (50 mL) and the reaction mixture was stirred at 50 °C for 15 h. The solvent was removed under high vacuum and AcOEt was added to the residue. The $NaN_3$ on excess was taken off by filtration, and the solvent was removed to obtain **4** as colorless oil (3.53 g, 99%). $^1$H NMR (300 MHz, $CDCl_3$) δ: 4.30 (t, 2H, $J$ = 5.1 Hz, $CH_2CH_2N_3$), 4.18 (d, 2H, $J$ = 11.9 Hz, $CH_2OCO$), 3.64 (d, 2H, $J$ = 11.9 Hz, $CH_2OCO$), 3.67 (t, 2H, $J$ = 5.1 Hz, $CH_2CH_2N_3$), 1.41 (s, 3H, $CH_3$), 1.36 (s, 3H, $CH_3$), 1.19 (s, 6H, 2 × $CH_3$); $^{13}$C NMR (100 MHz, $CDCl_3$) δ: 173.8 (COO), 98.0 ($C_{ipr}$), 65.8 ($CH_2OCO$), 63.5 ($CH_2CH_2N_3$), 49.7 ($CH_2CH_2N_3$), 41.9 (C), 24.8, 22.2 ($CH_3$), 17.7 ($CH_3$); ESI-MS for

$C_9H_{15}N_3O_4$; calcd: 243.1, $[M]^+$; found 266.1 $[M+Na]^+$. IR ($\upsilon$, $cm^{-1}$): 3448 (bs), 2992, 2940, 2105, 1734, 1638.

**2-Azidoethyl 2,2-Bis(hydroxymethyl)propionate (5).** Azido derivative **4** (11.0 g, 45.2 mmol) was dissolved in MeOH (50 mL). One teaspoon of a Dowex, H+ resin was added, and the reaction mixture was stirred at 40 °C until complete disappearance of starting material. When the reaction was complete, the Dowex H+ resin was filtered off in a glass filter and carefully washed with methanol. The methanol was evaporated to give **5** as colorless oil (9.0 g, 98%). $^1H$ NMR ($CDCl_3$, 300 MHz) $\delta$: 4.34 (t, 2H, $J$ = 5.1 Hz, $CH_2CH_2N_3$), 3.92 (d, 2H, $J$ = 11.2 Hz, $CH_2OCO$), 3.74 (d, 2H, $J$ = 11.2 Hz, $CH_2OCO$), 3.52 (t, 2H, $J$ = 5.1 Hz, $CH_2CH_2N_3$), 3.02 (bs, 2H, $OH$), 1.12 (s, 3H, $CH_3$); $^{13}C$ NMR ($CDCl_3$, 100 MHz) $\delta$: 175.3 (COO), 67.8 ($CH_2OCO$), 63.5 ($CH_2CH_2N_3$), 49.8 ($CH_2CH_2N_3$), 49.4 (C), 17.0 ($CH_3$); ESI-MS for $C_7H_{13}N_3O_4$; calcd: 203.1 $[M]^+$; found: 226.0 $[M+Na]^+$. IR ($\upsilon$, $cm^{-1}$): 3405 (bs), 2945, 2885, 2105, 1729, 1660, 1459. Elemental analysis calcd. (%) for $C_7H_{13}N_3O_4$: C, 41.38%, H, 6.45%; N, 20.68%; found: C, 41.83%; H, 6.79%; N, 20.50%.

**Compound 7.** Azido derivative **5** (2.4 g, 11.8 mmol) and DMAP (0.43 g, 3.54 mmol) were dissolved in pyridine (6 mL), followed by the addition of $CH_2Cl_2$ (12 mL). The anhydride of isopropylidene-2,2-bis(hydroxymethyl)propionic acid **6** (9.3 g, 30.7 mmol) was added slowly. The solution was stirred at room temperature until completion. The reaction was quenched with water (1 mL) under vigorous stirring, followed by dilution with $CH_2Cl_2$ (50 mL), and the solution was washed with 10% of $NaHSO_4$ (3 × 20 mL), 10% of $Na_2CO_3$ (3 × 20 mL), and brine (10 mL). The organic phase was dried with $MgSO_4$, filtered, and concentrated. The crude product was purified by flash chromatography on silica, eluting with hexane (100 mL) and gradually increasing the polarity to EtOAc/hexane (1:3) to give **7** as colorless oil (4.4 g, 73%). $^1H$ NMR ($CDCl_3$, 300 MHz) $\delta$: 1.17 (s, 6H, $CH_3$), 1.34 (s, 3H, $CH_3$), 1.38 (s, 6H, $CH_3$), 1.43 (s, 6H, $CH_3$), 3.51 (t, $J$ = 5.1 Hz, 2H, $CH_2CH_2N_3$), 3.64 (d, $J$ = 11.9 Hz, 4H, $CH_2OCO$), 4.17 (d, $J$ = 11.8 Hz, 4H, $CH_2OCO$), 4.34−4.28 (t, $J$ = 5.1 Hz, 2H, $CH_2CH_2N_3$), 4.36 (s, 4H, $CH_2OCO$); $^{13}C$ NMR ($CDCl_3$, 75.5 MHz) $\delta$: 173.5, 172.3 (COO), 98.1 ($C^ipr$), 66.0, 66.0, 65.3 ($CH_2OCO$), 63.9 ($CH_2CH_2N_3$), 49.6 ($CH_2CH_2N_3$), 46.9, 42.1 (C), 25.3, 22.0 ($CH_3{}^ipr$), 18.0, 17.0 ($CH_3$); ESI-EM calc for $C_{23}H_{37}N_3O_{10}$($m/z$): 515.2; found 538.2 $[M+Na]^+$. IR ($\upsilon$, $cm^{-1}$): 3453 (bs), 2991, 2940, 2876, 2106, 1738, 1643, 1455.

**Compound 8.** Azido derivative **7** (4.3 g, 8.4 mmol) was dissolved in MeOH (20 mL). One teaspoon of a Dowex H+ resin was added, and the reaction mixture was stirred at 40 °C until complete disappearance of starting material by TLC. When the reaction was complete, the Dowex H+ resin was filtered off in a glass filter and carefully washed with methanol. The methanol was evaporated to give **8** as colorless oil (3.5 g, 97%). $^1H$ NMR ($CDCl_3$, 300 MHz) $\delta$: 4.46 (d, $J$ = 11.1 Hz, 2H, $CH_2OCO$), 4.36−4.26 (m, 4H, $CH_2OCO$, $CH_2CH_2N_3$), 3.90−3.78 (m, 4H, $CH_2OCO$), 3.76−3.67 (m, 4H, $CH_2OCO$), 3.51 (t, $J$ = 5.1 Hz, 2H, $CH_2CH_2N_3$), 3.19−2.87 (bs, 4H), 1.36 (s, 3H, $CH_3$), 1.08 (s, 6H, $CH_3$); $^{13}C$ NMR ($CDCl_3$, 75.5 MHz) $\delta$: 175.0, 175.0, 172.8 (COO), 67.3, 67.2, 67.2, 64.8 ($CH_2OCO$), 64.0 ($CH_2CH_2N_3$), 49.8 ($CH_2CH_2N_3$), 49.8, 49.6 (C), 18.0, 17.0 ($CH_3$); ESI-EM calcd for $C_{17}H_{29}N_3O_{10}$ ($m/z$): 435.2; found 458.2 $[M+Na]^+$. IR ($\upsilon$, $cm^{-1}$): 3405 (bs), 2943, 2886, 2108, 1731, 1469.

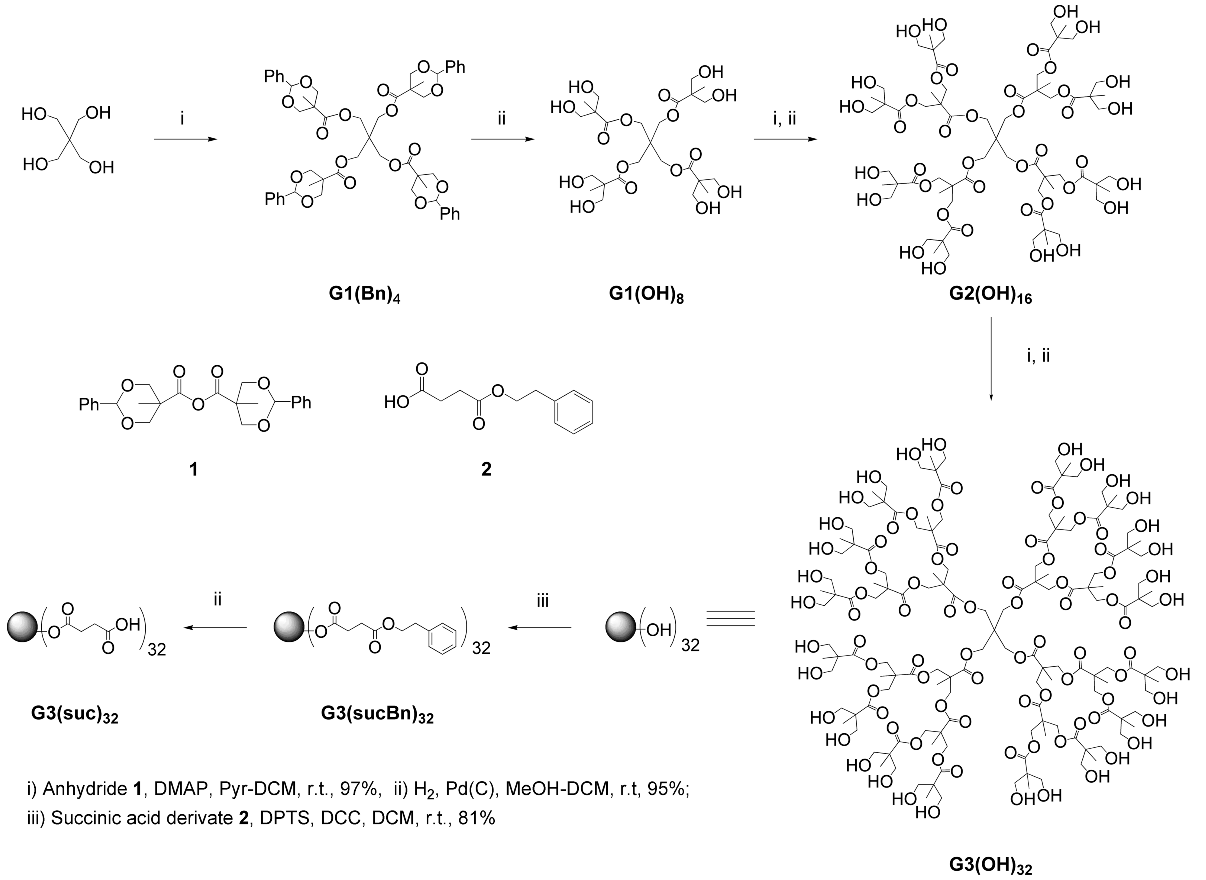
**Compound 10.** 4-Oxo-4-(2-(trimethylsilyl)ethoxy)butanoic acid (**9**) (1.43 g, 6.57 mmol), tetraol **8** (0.50 g, 1.09 mmol), and DPTS (0.65 g, 2.18 mmol) were mixed in $CH_2Cl_2$ anhydrous (10 mL). Then, the reaction flask was flushed with argon and DCC (1.35 g, 6.57 mmol) was added. Stirring at 40 °C was continued for 15 h under argon atmosphere. Once the reaction was complete, the formed urea was filtered off on a glass filter and washed with a small volume of $CH_2Cl_2$. The crude product was purified by liquid chromatography Sephadex LH20 ($CH_2Cl_2$/MeOH; 1:1) to give **10** as colorless viscous oil (1.11 g, 81%). $^1H$ NMR ($CDCl_3$, 300 MHz) $\delta$: 4.43−3.97 (m, 14H, 4 $CH_{2G1}$, 8 $CH_{2G2}$, 2 $CH_2CH_2N_3$, 8 $CH_2CH_2COCH_2$), 3.53−3.43 (t, $J$ = 5.1 Hz, 2H, $CH_2CH_2N_3$), 2.65−2.47 (m, 16H, $CH_2$suc), 1.26 (s, 3H, $CH_{3\ G\acute{1}}$), 1.20 (s, 6H, $CH_{3\ G\acute{2}}$), 1.00−0.88 (t, $J$ = 8.5 Hz, 8H, $CH_2$-TMS), 0.00 (s, 36 H, $CH_3$TMS); $^{13}C$ NMR ($CDCl_3$, 75.5 MHz) $\delta$: 172.1, 171.9, 171.9, 171.7 (COO), 65.6, 65.3, 64.0, 62.9 ($CH_2OCO$, $CH_2CH_2N_3$, $CH_2CH_2COCH_2$), 49.6 ($CH_2CH_2N_3$), 46.7, 46.4 ($CCH_2O$), 29.1, 28.8 ($CH_2$suc), 17.7, 17.4 ($CH_2$-TMS, $CH_3$), 17.2, −1.5, −1.9 ($CH_3$TMS); ESI-HRMS calcd for $C_{53}H_{93}N_3O_{22}Si_4Na$ ($m/z$): 1258.5291; found: 1258.5226. IR ($\upsilon$, $cm^{-1}$): 2955, 2898, 2104, 1730, 1471.

**Compound 11.** TFA (5 mL) was added to a solution of **10** (142 mg, 0.115 mmol) in $CH_2Cl_2$ (5 mL). The mixture was stirred at room temperature for 2 h. After complete conversion, the solvent was evaporated and dried in high vacuum to afford **11** as pale yellow syrup (96 mg, 100%). The compound was used in the next step without purification. $^1H$ NMR (MeOH-$d_4$, 400 MHz) $\delta$: 4.30 (m, 28H, 4$CH_{2\ G1}$ 8$CH_{2\ G2}$, 2$CH_2CH_2N_3$), 3.57 (t, $J$ = 5.1 Hz, 2H, $CH_2CH_2N_3$), 2.62 (m, 16H, $CH_2CONH$, $CH_2COO$), 1.34 (s, 3H, $CH_{3G1'}$), 1.27 (s, 6H, $CH_{3G2'}$); $^{13}C$ NMR ($CDCl_3$, 100 MHz) $\delta$: 174.4, 172.2, 172.2, 172.0 (COO), 65.5, 65.3, 65.2, 64.0 ($CH_2OCO$, $CH_2CH_2N_3$, $CH_2CH_2COCH_2$), 49.4 ($CH_2CH_2N_3$), 46.9, 46.5, 46.3 ($CCH_2O$), 28.6, 28.5, 28.3, 28.2 ($CH_2CONH$, $CH_2COO$), 16.7, 16.5 ($CH_3$); ESI-HRMS calcd for [M+Na] $C_{33}H_{45}N_3O_{22}Na$ ($m/z$): 858.2392; found: 858.2435.

**Compound 12.** To a solution of the scaffold dendron **11** (6.0 mg, 0.00718 mmol, 1 equiv) in 150 $\mu$L of dry DMA under nitrogen atmosphere, HATU (22 mg, 0.0575 mmol, 8 equiv) and DIPEA (20 $\mu$L, 0.1157 mmol, 16 equiv) were added. After 15 min, a solution of **13** (23.4 mg, 0.05349 mmol, 8 equiv) in dry DMA (210 $\mu$L) was added. The reaction was stirred at room temperature for 3 days. MALDI mass analysis of a sample of the reaction mixture showed the completion of the reaction. The reaction mixture was diluted in methanol and charged directly onto a Sephadex LH-20 column (methanol). Slow elution led to the purification of the product **12** that was isolated in good yield (16.9 mg, 94%). $^1H$ NMR (400 MHz, $D_2O$): $\delta$ (ppm) 4.93 (s, 4H, $H_1$), 4.27 (bs, 6H, $OCH_2CH_2N_3$, $CH_2O_{G1}$), 4.18 (bs, 8H, $CH_2O_{G2}$), 3.91 (bs, 8H, $D_2$, $H_2$), 3.78 (d, 4H, $H_6B$, $J$ = 12 Hz), 3.76−3.71 (m, 4H, $H_3$), 3.69−3.59 (m, 8H, $H_{6A}$, $H_{D1}$), 3.63 (s, 24H, $OCH_3$), 3.59−3.49 (m, 18H, $H_4$, $H_5$, $H_7$, $CH_2N_3$), 3.30 (bt, 8H, $H_8$), 2.91−2.72 (m, 8H, $H_{D4}$, $H_{D5}$), 2.65−2.56 (m, 8H, $CH_2COO$), 2.54−2.45 (m, 8H, $CH_2CONH$), 2.10−1.99 (m, 8H, $H_{D6eq}$, $H_{D3eq}$), 1.81−1.64 (m, 8H, $H_{D6ax}$, $H_{D3ax}$), 1.26 (s, 3H, $CH_{3G1'}$), 1.19 (s, 6H, $CH_{3G2'}$); $^{13}C$ NMR (100 MHz, $D_2O$): 177.4, 177.2, 174.1, 174.0, 173.8, 173.7 ($COOCH_3$, $COOCH_2$, CONH), 98.6 ($C_{1M}$), 73.8 ($C_{D1}$), 73.5 ($C_{4M}$), 71.0, 70.5 ($C_{D2}$, $C_{2M}$, $C_{3M}$), 66.9 ($C_7$), 66.8 ($C_{5M}$), 65.9 ($OCH_{2G1}$, $OCH_{2G2}$), 64.8 ($OCH_2CH_2N_3$), 61.0 ($C_{6M}$); 52.6 ($CH_3O$); 49.4 ($CH_2N_3$); 46.6, 46.4 ($C_{quat\ G1,\ G2}$); 39.4 ($C_8$); 39.0 ($C_{D4}$, $C_{D5}$); 30.1 ($CH_2CONH$); 29.3 ($CH_2COO$); 27.0, 26.7 ($C_{D6}$, $C_{D3}$); 16.9 ($CH_{3G1,G2}$); ESI-MS calcd for $[C_{105}H_{161}N_7O_{62}Na_2]^{+2}$ 1279.7, found: 1279.2; MALDI-ToF MS (matrix DHB) $[C_{105}H_{161}N_7O_{62}Na]^+$ calcd: 2536.4, found: 2537.2.

**Production of Recombinant Viruses.** Recombinant viruses were produced in 293 T cells. The viral construction was pseudotyped with Ebola virus envelope glycoprotein (EboGP) and

Scheme 1. Synthesis of Compound G3(suc)$_{32}$ from Pentaerythritol



i) Anhydride **1**, DMAP, Pyr-DCM, r.t., 97%, ii) $H_2$, Pd(C), MeOH-DCM, r.t, 95%;
iii) Succinic acid derivate **2**, DPTS, DCC, DCM, r.t., 81%

expressed luciferase as a reporter of the infection. The infection of Jurkat cells (a CD+ T-lymphocyte cell line) by EboGP is absolutely dependent on virus envelope interaction with DC-SIGN. As a control of DC-SIGN independent infection, we produced recombinant viruses pseudotyped with vesicular stomatitis virus envelope glycoprotein (VSV-G).

One day (18−24 h) before transfection, 6 × $10^6$ 293 T cells were seeded onto 10 cm plates. Cells were cultured in DMEM medium supplemented with 10% heat-inactivated FBS, 25 mg Gentamicin, and 2 mM L-glutamine. Few minutes before transfection, the medium on transfection plates was changed to 9 mL DMEM and 25 $\mu$M chloroquine. Transfection reaction with all reagents at room temperature (r.t.) was prepared in 15 mL tubes: 183 $\mu$L of 2 M $CaCl_2$, 450 ng of Ebola virus envelope, 21 $\mu$g of pNL4−3 luc, and 1200 $\mu$L of $H_2O$. Next, 1.5 mL of 2×HBS (Hepes Buffer Saline) pH 7.00 was added quickly to the tubes and bubbled for 30 s. HBS/DNA solution was dropwise added onto medium. After 8 h of incubation at 37 °C with 5% $CO_2$, medium on transfection plates was changed to 10 mL DMEM and once again one day after transfection to 7 mL DMEM. Transfection supernatants were harvested after 48 h, centrifuged at 1200 rpm for 10 min at r.t. to remove cell debris, and stored frozen at −80 °C.

**Infection Assay.** *Infection in cis.* Infection was performed on Jurkat cells expressing DC-SIGN. Since Ebola virus does not infect T-lymphocytes, its entry into Jurkat cells is absolutely dependent on the interaction with DC-SIGN.[30]

DC-SIGN$^+$ Jurkat cells (2.5 × $10^5$) were incubated with carbohydrate multivalent compounds on 24-well suspension plates for 30 min at r.t. and then challenged with 10 000 tissue culture infective dose (TCID) of recombinant viruses. After 48 h of incubation, cells were washed twice with PBS and lysed with 100 $\mu$L of 1× Lysis Buffer for luciferase assay (Promega).

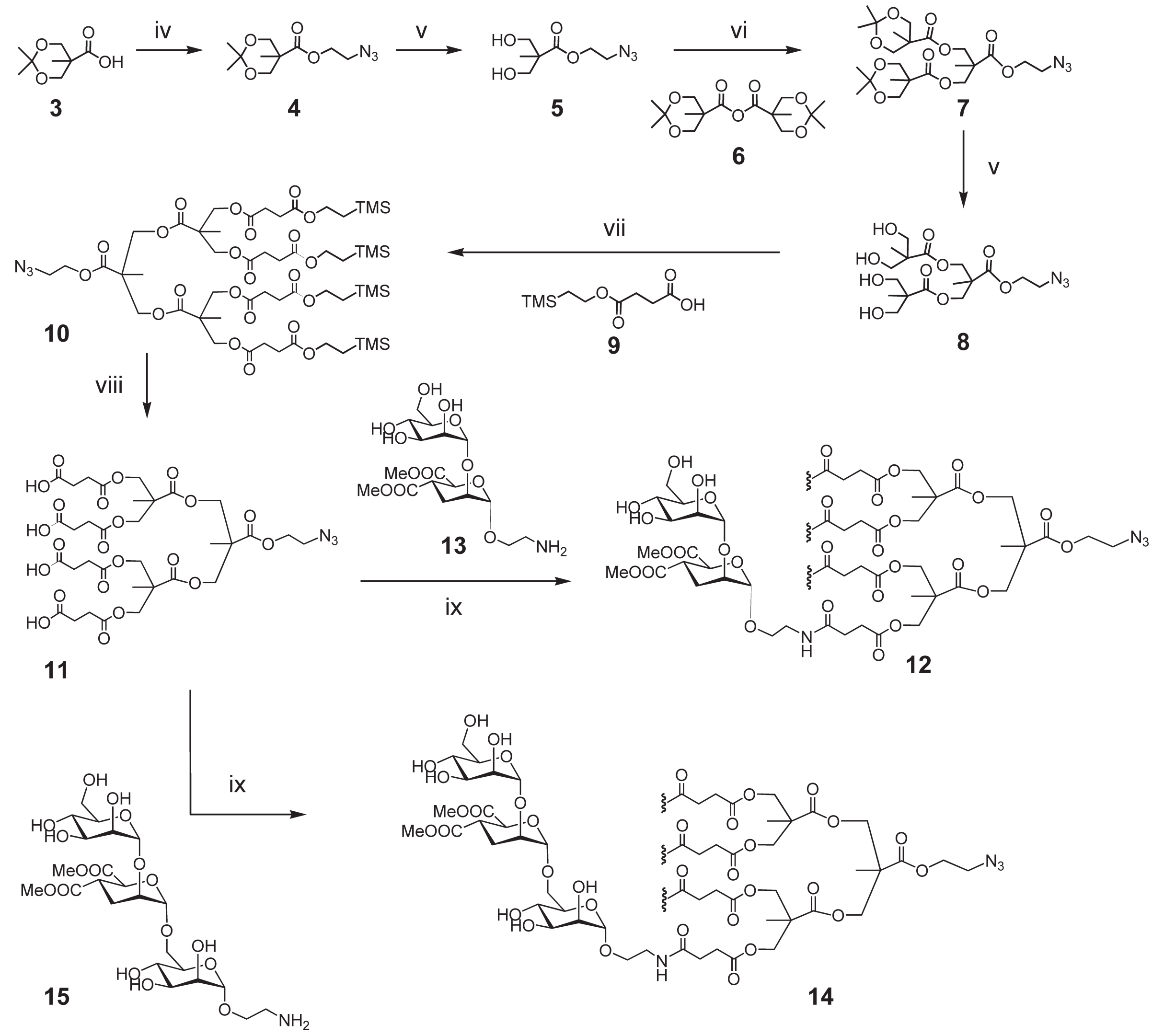
Infection control experiment was performed with VSV-G pseudoviruses under the same conditions.

*Infection in trans.* DC-SIGN$^+$ Jurkat cells (2.5 × $10^5$) were preincubated for 20 min at r.t. with carbohydrate multivalent compounds. Then, they were challenged with 10 000 TCID of recombinant viruses and incubated for 2 h at r.t. with rotation. After 2 h, cells were centrifuged at 1000 rpm for 5 min and washed twice with 1 mL of PBS supplemented with 0.5% bovine serum albumin (BSA) and 1 mM $CaCl_2$. DC-SIGN$^+$ Jurkat cells were resuspended in 500 $\mu$L of RPMI medium and co-cultivated with adherent HeLa cells ($10^5$ cells/well) on 24-well plate. After 48 h, supernatant was removed and monolayer of HeLa was washed twice with 1 mL of PBS and lysed with 100 $\mu$L of 1× Lysis Buffer for luciferase assay.

As control experiment, *trans*-infection test with VSV-G pseudoviruses was performed.

**Surface Plasmon Resonance.** Extracellular domain (ECD) of DC-SIGN protein (residue 66−404) has been overexpressed and purified as described previously.[31] All experiments were performed on a Biacore 3000 using functionalized CM4 sensor chips and the corresponding reagents from BIAcore. Two flow

**Scheme 2. Synthesis of Tetravalent Compounds 12 and 14**



iv) a) 2-bromoethanol, DMAP, DCM, r.t.; b) $NaN_3$, DMF, 86%; v) Dowex, H+ resin, MeOH, r.t., 98%; vi) DMAP, Anhydride **6**, DCM, r.t., 73%, vii) Acid **9,** DPTS, DCC, DCM, 40ºC, 81%; viii) TFA, DCM, r.t., quant.; ix) pseudomannobioside **13** or pseudomannotrioside **15**, HATU, DIPEA, DMA, r.t., 94%

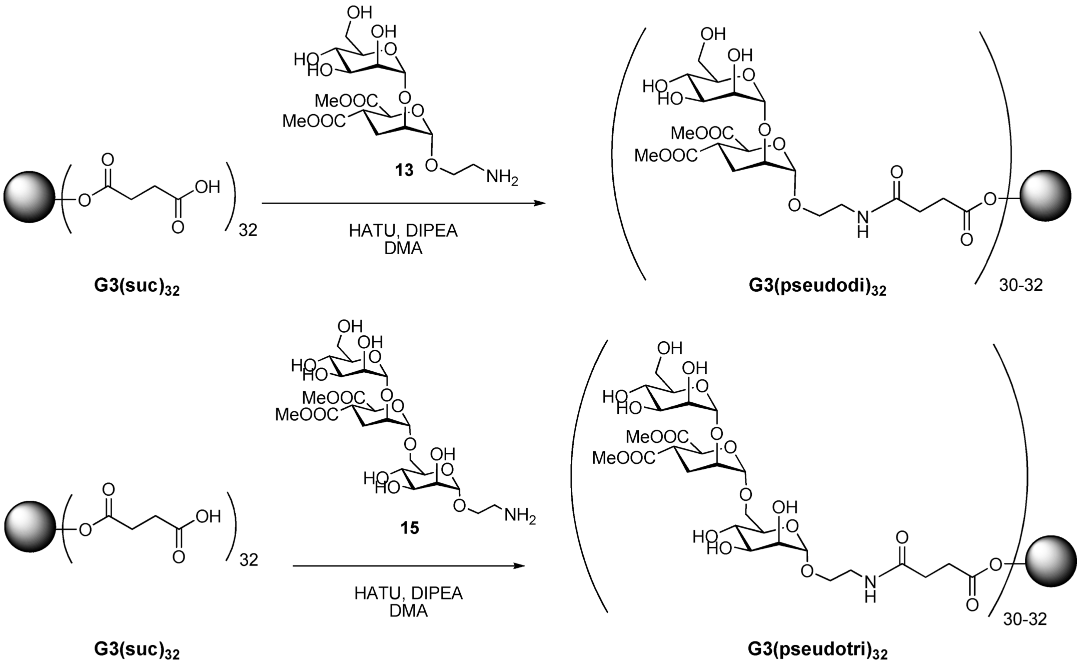
cells were activated as previously described.[32] Flow cell one was then blocked with 30 μL of 1 M ethanolamine and used as a control surface. The second one was treated with BSA-Manα1−3[Manα1−6]Man (BSA-Mannotriose, Dextra) (60 μg/mL) in 10 mM acetate buffer, pH 4. Remaining activated groups were blocked with 30 μL of 1 M ethanolamine. The final density immobilized on the surface of the second flow cell was 1200 RU. The BSA-Mannotriose used to functionalize CM4 chip harbors 12 glycosylation sites according to manufacturer. The affinity of the various sugars and mimics was then estimated through a DC-SIGN ECD binding inhibition assay. The ECD of DC-SIGN was injected onto the BSA-Mannotriose surface, at 20 μM alone or in presence of an increasing concentration of the sugar derivatives. Injections were performed at 20 μL/min using 25 mM Tris-HCl, pH 8, 150 mM NaCl, 4 mM $CaCl_2$, and 0.005% of P20 surfactant as running buffer. The original sensorgrams are reported as Supporting Information.

To determine $IC_{50}$ values for sugar derivatives, the steady state binding responses of DC-SIGN ECD to BSA-Mannotriose surface were obtained from sensorgrams and converted to relative residual activity values. Relative $IC_{50}$ values were determined from the plots of sugar derivative concentration vs relative residual DC-SIGN ECD activity by fitting four-parameter logistic model (see eq 1 in SI) to the experimental data.

## ■ RESULTS

In previous works, we have demonstrated that the commercially available and very cheap third generation of a Boltorn

Scheme 3. Coupling of Glycomimetics 13 and 15 on G3(suc)$_{32}$



dendritic polymer was a very convenient platform for a multivalent presentation of carbohydrates.[19] This polydisperse dendritic polymer, bearing 32 copies of mannose, exhibited a high antiviral activity ($IC_{50} = 0.3\ \mu M$) in an Ebola infection model.[20] The polydispersity of the commercial Boltorn scaffold is a major drawback of this strategy. On the basis of these preliminary results, we decided to prepare a Boltorn-type polyester dendrimer in a completely monodisperse manner. To approach this synthesis, we used a divergent strategy based on the anhydride **1** as a building block monomer and pentaerythritol as a central core, a strategy previously used to prepare polyester dendrons.[26] A sequential set of two reactions was employed as described in Scheme 1. A coupling reaction with anhydride **1** to grow to the next generation, followed by hydrogenolysis using palladium on carbon as catalyst to remove the benzylidene protecting groups, yielded the first ($G1(OH)_8$), second ($G2(OH)_{16}$), and third ($G3(OH)_{32}$) generations of the polyester dendrimers in a very efficient way and with complete monodispersity, as shown by mass spectrometry. These dendrimers were prepared in multigram-scale and purified by recrystallization in a very straightforward way.

Carboxylic acids were introduced as functional groups at the surface of the third-generation dendrimer $G3(OH)_{32}$ to allow coupling with conveniently functionalized ligands by the formation of amide bonds. In contrast to our previous experience, where the carboxylic acids were introduced by reaction of $G3(OH)_{32}$ with a large excess of succinic anhydride, succinic acid monobenzyl ester and DCC were used for the coupling reaction (Scheme 1). Using this strategy, the resulting compound $G3(sucBn)_{32}$ was soluble in organic solvents and very easy to purify by flash chromatography on silica gel. Again, a complete analysis of the compound, especially by MS, confirmed the monodispersity of the molecule. After a simple step of benzyl ester deprotection by hydrogenolysis at atmospheric pressure using Pd(C) (10%) at room temperature, the dendrimer $G3(suc)_{32}$ with 32 carboxylic acids was obtained almost quantitatively.

To evaluate the importance of the valency, a tetravalent system also based on 2,2′-bis(hydroxymethyl) propionic acid (bis-MPA) (Scheme 2) was considered as a scaffold for presentation of four copies of selected ligands.

Starting from the isopropylidene-2,2-bis(methoxy)propionic acid (**3**) prepared as reported in the literature[28] and following the experimental procedure described in Scheme 2, we obtained **4** in good yields with adequate functionalization. As a protecting group of the carboxylic acid, the azidoethanol was used. This terminal azide has been used in our laboratory in a convergent strategy to create higher valency compounds applying a click chemistry Cu catalyzed cycloaddition (unpublished results). Compound **4** was prepared in good yields by reaction of **3** with 2-bromoethanol in the presence of DCC and 4-(dimethylamino)pyridinum 4-toluenesulfonate (DPTS) followed by treatment with sodium azide in DMF at 40 °C. After deprotection of the hydroxyl groups using an acidic resin in methanol, the tetraol **8** was prepared using a two-step synthetic sequence reported in the literature to prepare similar polyester dendrons.[28,29,33] This synthetic sequence begun coupling **5** with 2 equiv of anhydride **6** which, in turn, was prepared in one step from acid **3** using DCC,[29] in pyridine and dimethylaminopyridine as catalyst. The second step was the deprotection of the isopropylidene acetals of **7** using Dowex H+. Tetraol **8** was obtained in good yield after chromatographic purification. The introduction of the carboxylic acid linkers on **8** was performed using the trimethylsilylethanol monosuccinate **9** prepared by ring-opening of succinic anhydride by trimethylsilylethanol as described in the literature.[34] In this way, the silyl protected compound was prepared in good yields and easily purified. After deprotection using trifluoroacetic acid at room temperature, the tetraacid **11** was obtained in very good yield and was used in the next step without any further purification.

Our previous studies have demonstrated that compounds **13** and **15**, linear glycomimetics of the disaccharide Manα1,2Man and of the trisaccharide Manα1,2Manα1,6Man, respectively, were

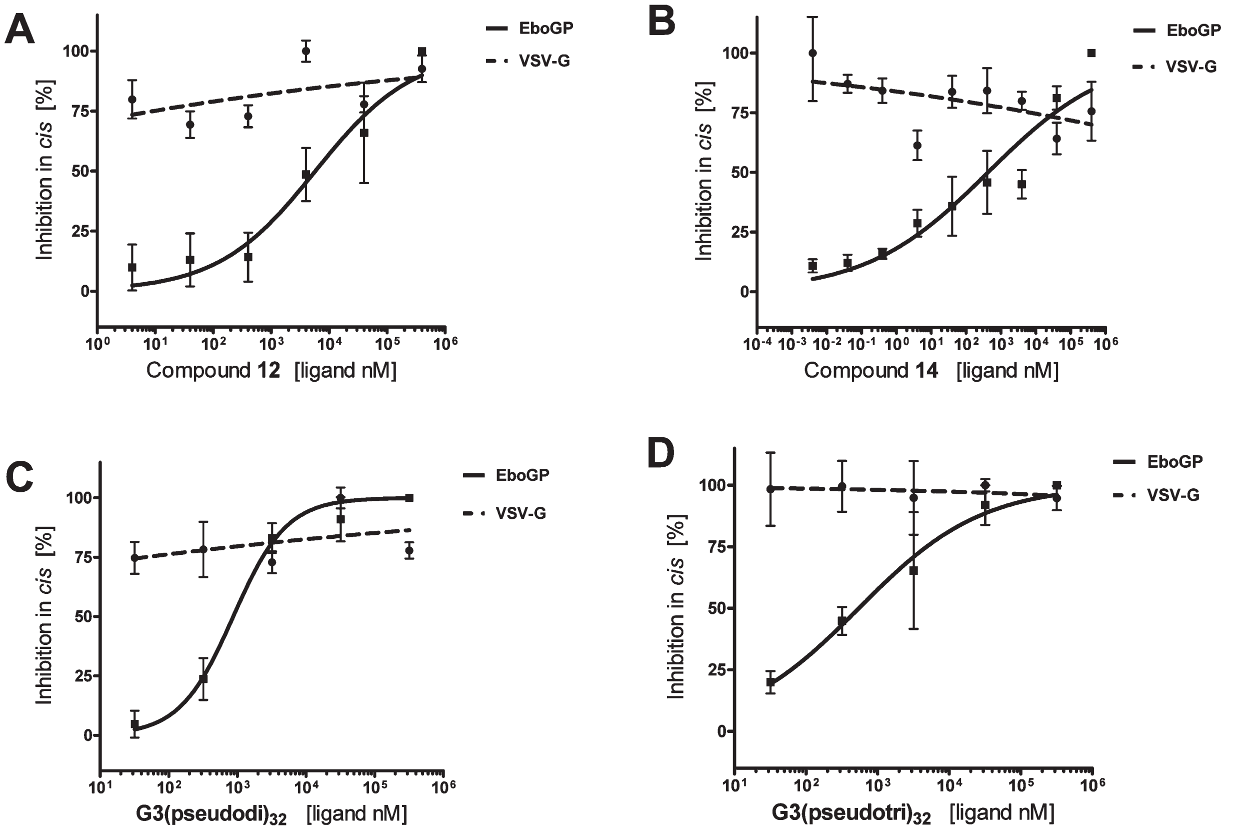


**Figure 1.** Infection results of the *cis*-experiments using **12**, **14**, G3(pseudodi)$_{32}$, and G3(pseudotri)$_{32}$ compounds as inhibitors. Concentrations are expressed based on ligand concentration.

very adequate ligands for DC-SIGN providing better stability against enzymatic degradation than the corresponding natural counterparts. The synthesis of these glycomimetics has been previously described.[22,23]

Conjugation of **13** and **15** on the tetra- and multivalent scaffolds was performed using HATU as an activating reagent in the presence of DIPEA at room temperature and afforded the corresponding multivalent systems with high yields after purification by size exclusion chromatography as described in Schemes 2 and 3.

The tetravalent compounds were fully characterized by NMR, IR, and MS analysis. MALDI-MS of the compounds G3(pseudodi)$_{32}$ and G3(pseudotri)$_{32}$ using sinapinic acid (SA) as matrix allowed to observe a complex signal centered on the molecular weight corresponding to the 31-loaded dendrimer both for pseudodi- and pseudotrimultivalent systems (see SI for MS spectra). This is consistent with the number of 30−32 ligands which could be estimated based on the integration of significant signals in the $^1$H NMR spectrum (see SI). Although the MS spectra were not conclusive, we concluded that almost all the reactive sites of the scaffold were conjugated with ligands to obtain a high load narrowly disperse system.

**Infection Studies.** The antiviral activity of these tetra- and multivalent systems was tested using pseudotyped viral particles presenting the EboGP and Jurkat cell line expressing DC-SIGN on the surface. Two kinds of experiments were set up: a *cis*-infection assay of DC-SIGN+ Jurkat cells with this artificial Ebola virus, and a *trans*-infection experiment of HeLa cells using DC-SIGN+ Jurkat cells as viral particle transporters.

We have used these infection models previously to test the activity of mannosylated Boltorn-type hyperbranched dendritic polymers.[20] The DC-SIGN blocking efficiency has been evaluated as dependence of multivalency and of the different pseudomannoside ligands used in the compounds synthesized. The infection inhibition analysis has been done as a function of compound concentration. All compounds were used in at least 3 independent experiments. The results of the infection assays are represented as a percentage of infection inhibition compared with DC-SIGN+ Jurkat cells infected by pseudoviruses without addition of any compound (blank experiment). The IC$_{50}$ values were calculated with the 95% confidence intervals. Control experiment with VSV-G pseudoviruses was performed with final concentrations of 10 nM and 10 $\mu$M (negative control).

The results obtained in the *cis* infection experiments indicate that the tetravalent systems presenting four copies of the pseudomannotrioside, **14**, show an IC$_{50}$ 1 order of magnitude lower than the system with four pseudomannobiosides, **12** (Figure 1). In a previous publication, we have shown that the tetravalent system **14** bearing four copies of linear pseudomannotrioside **15** was an effective inhibitor of T-cell *trans* infection by HIV.[23]

Multivalent systems with an average of 30−32 copies of glycomimetic ligands were tested. In this case, no significant differences were found as a function of the pseudosaccharide presented. The IC$_{50}$ values found for the multivalent systems were around 20 nM (around 0.5 $\mu$M, based on ligand concentration), which shows that their inhibitory potency is between

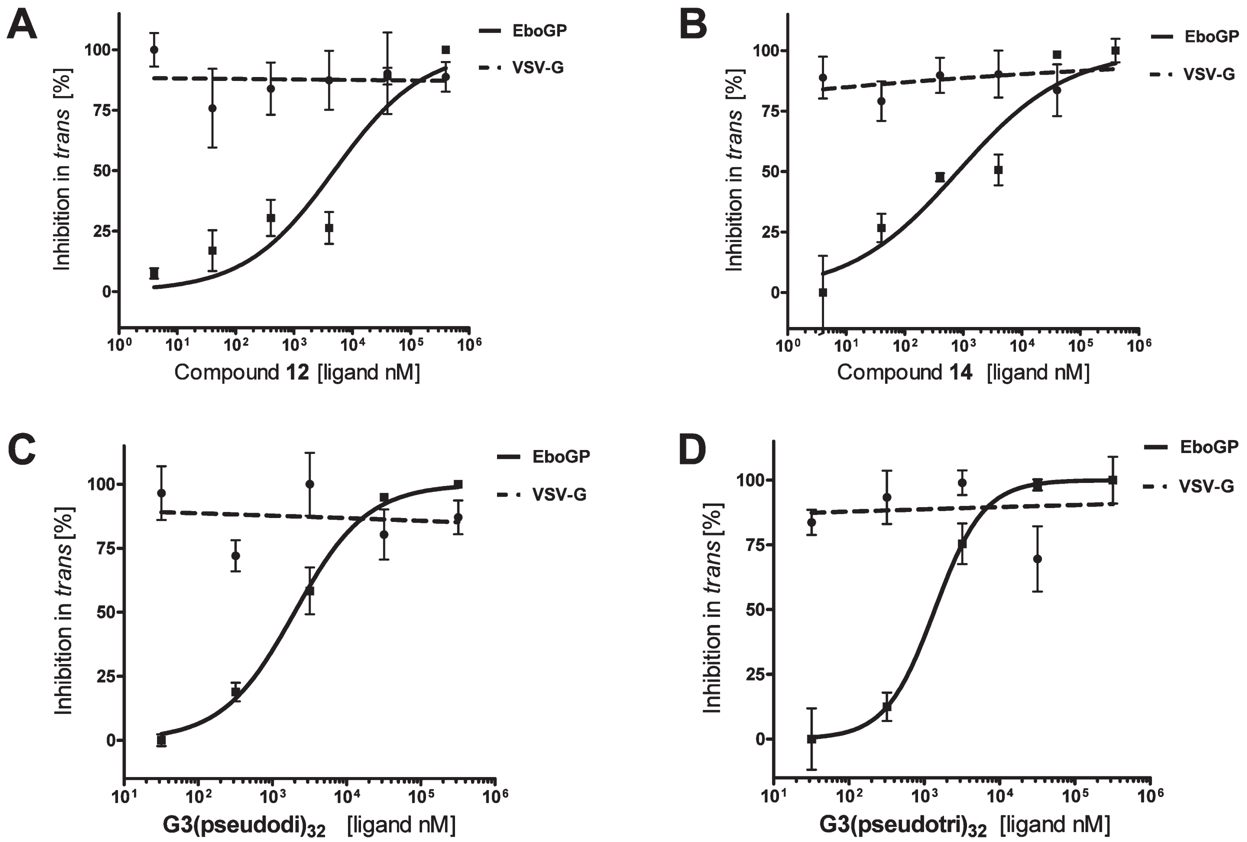


**Figure 2.** Infection results of the *trans* experiments using **12**, **14**, G3(pseudodi)$_{32}$, and G3(pseudotri)$_{32}$ compounds as inhibitors. Concentrations are expressed based on ligand concentration.

1 to 2 orders of magnitude higher than that of the corresponding tetravalent systems.

Similar results were obtained in the *trans* infection experiments (Figure 2). The IC$_{50}$ of 32-valent systems bearing pseudomannobioside **13** or pseudomannotrioside **15** were 62 nM and 31.5 nM, respectively (2 and 1 $\mu$M, based on ligand concentration, respectively). The tetravalent dendron **14** with four copies of pseudomannotrioside showed an IC$_{50}$ of 203 nM (0.8 $\mu$M, based on ligand concentration), and the tetravalent dendron **12** with four copies of pseudomannobioside was 1 order of magnitude less potent with an IC$_{50}$ of 1.22 $\mu$M (4.9 $\mu$M, based on ligand concentration) (Figure 2).

**SPR Studies.** Competition experiments using Biacore SPR were performed to analyze the interaction of these glycomimetic tetra- and multivalent systems with the receptor DC-SIGN and to characterize the interaction at the molecular level.

For this aim, Biacore CM4 sensor chip was functionalized as described in the experimental section with BSA-Mannotriose. A fixed amount of DC-SIGN ECD was injected over the sensor chip flow cells in the presence or absence of different concentrations of the glycodendritic compounds to be evaluated. Analysis of the sensorgrams (see Supporting Information) allowed estimating IC$_{50}$ for the monovalent, tetravalent, and multivalent compounds. At the monovalent level, strong differences were found (almost 1 order of magnitude) between the IC$_{50}$ of monovalent pseudodi- and pseudotrisaccharides (IC$_{50}$ 1.19 mM for the pseudodisaccharide **13** and 0.16 mM for the pseudotrisaccharide **15**). However, these differences decrease notably when the ligands are presented on the tetravalent scaffold (IC$_{50}$ 227 $\mu$M for the tetravalent pseudodi **12** and 120 $\mu$M for the tetravalent pseudotri **14**) and are very similar when a dendrimer of third generation is used as multivalent platform (IC$_{50}$ 2 $\mu$M for the multivalent pseudodi G3(pseudodi)$_{32}$ and 1.25 $\mu$M for the multivalent pseudotri G3(pseudotri)$_{32}$. (Figure 3)

These data show that tetravalent presentation of the pseudodisaccharide **13** on dendron **12** affords a very modest (1.3 per pseudosugar), if any, polyvalency effect ($\beta$), which improves significantly (18 per pseudosugar) in the full dendrimer G3-(pseudodi)$_{32}$. Strikingly, the tetravalent presentation of the pseudotrisaccharide **15** on dendron **14** does not afford any affinity increase ($\beta$ = 0.3 per pseudosugar) and only a modest improvement (4 per pseudosugar) is observed for the full dendrimer G3(pseudotri)$_{32}$.

## ■ DISCUSSION

DC-SIGN is considered a very interesting therapeutic target for the design of new antiviral drugs. In addition, the ability of DC-SIGN to trigger signaling pathways that lead to a specific immune response has attracted attention on this lectin.[35−42] The development of specific tools to interact with and to understand the role and function of this lectin is of great interest. Furthermore, specific ligands designed to block the carbohydrate recognition domain of this lectin could be good candidates to be tested as antimicrobial agents.

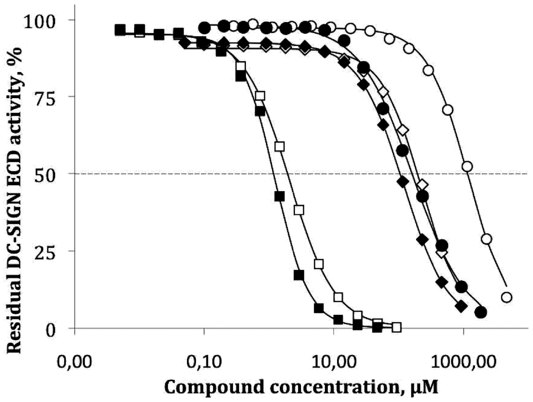


**Figure 3.** Inhibition of DC-SIGN ECD interaction with BSA-Mannotriose surface. DC-SIGN ECD (20 $\mu$M) was incubated with pseudomannobioside and pseudomannotrioside compounds at increasing concentrations: from 0 to 5000 $\mu$M and to 2000 $\mu$M for monovalent pseudomannobioside and pseudomannotrioside, respectively, from 0 to 1000 $\mu$M for compounds **12** and **14**, and from 0 to 100 $\mu$M for G3(pseudodi)$_{32}$, and G3(pseudotri)$_{32}$. The samples were coinjected over BSA-Mannotriose surface (1200 RU functionalization level). The steady state responses were extracted from the sensorgrams (see Supporting Information), converted to DC-SIGN ECD residual activity, and plotted against corresponding compound concentration. ○,● — monovalent, ◇,♦ — tetravalent, and □,■ — multivalent compounds; the empty and filled symbols represent pseudomannobioside and pseudomannotrioside compounds, respectively.

In this context, we have selected two linear pseudosaccharides **13** and **15** with good affinity for this receptor. The affinity has been remarkably enhanced *via* a multivalent presentation of these ligands. Tetra- and multivalent scaffolds based on bis-MPA as building block have been prepared with adequate functionalization on the surface for attachment of the selected ligands. The corresponding glycoconjugates with 4 and 32 copies of these ligands were prepared in good yields and were characterized using typical spectroscopic methods. MS analysis demonstrated the monodispersity of the tetravalent systems (compounds **12** and **14**). Higher valency systems obtained from monodisperse G3 Boltorn-type dendrimer G3(OH)$_{32}$ did not give good MS spectra and by analysis of the NMR spectra could only be described as systems with 30−32 ligands as an average loading; therefore, the monodispersity of these systems was not confirmed.

These tetra- and multivalent systems were tested *in vitro* as antiviral compounds in a well-established infection model based on pseudotyped viral particles with the Ebola virus envelope glycoprotein GP1. This infection system is a very efficient and “clean” model where the infection is exclusively dependent on DC-SIGN.[20] Cells that do not express the receptor and viruses able to infect the cells independently of DC-SIGN have been used as controls in these experiments. The results obtained have clearly demonstrated inhibition of the infection through blockage of the DC-SIGN receptor. The tetravalent systems **12** and **14** were very active in the low micromolar range, and the multivalent systems G3(pseudosugar)$_{32}$ showed a very strong inhibition effect with IC$_{50}$ in the nanomolar range. These systems could be considered as very potent inhibitors to be tested in further experiments. On the other hand, relatively small differences were observed between the two selected monovalent ligands the pseudodi **13** and the pseudotrisaccharide **15**.

In order to analyze this in more detail and to have additional molecular-level data on the interaction of these compounds with DC-SIGN, competition experiments using a biosensor with SPR detection were performed. A chip functionalized with mannosylated BSA was used to compete with our ligands for DC-SIGN. Analysis of the results obtained in these competition experiments demonstrates that the pseudosugars are good ligands for DC-SIGN, and that the pseudotrisaccharide **15** is 1 order of magnitude better than the pseudodisaccharide **13** when used as monovalent units. Surprisingly, this trend was not observed when the ligands are used in multivalent presentations: they both present a similar affinity when are exposed on the multivalent scaffolds. A plausible explanation for these facts could be the loss of different binding modes of the ligands when they are linked to a scaffold. In the particular case of DC-SIGN, it has been experimentally determined by X-ray[43] and also by NMR[44] that different binding modes are possible for carbohydrate ligands. These binding modes were observed for free ligands, but once they are conjugated to scaffolds, some or all of these alternative binding modes could become inaccessible; consequently, this could notably influence binding affinity. As a consequence, despite the improved affinity of the pseudomannotrioside at the monovalent level, our data suggest that the pseudomannobioside is the effective lead compound to use for further improvement of our multivalent pseudosaccharide compounds.

More structural experiments are currently being carried out in our laboratory at the monovalent level to obtain more detailed information and to investigate the factors that govern this phenomenon.

In summary, we present in this work the preparation of carbohydrate mimics on a tetra- and multivalent presentations and the analysis of their antiviral activity using a pseudotyped Ebola virus infection model. A strong inhibitory activity was found for the higher valency compounds with IC$_{50}$ in the nanomolar range. Competition experiments using SPR have also demonstrated the ability of these compounds to interact with DC-SIGN and compete for this receptor. Our results confirm pseudoglycosylated multivalent systems as very promising antiviral drugs with strong activities and encourage us to pursue more *in vitro* and *in vivo* studies concerning the potential biomedical applications.

## ■ ASSOCIATED CONTENT

**Ⓢ Supporting Information.** Additional information as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author

*E-mail: franck.fieschi@ibs.fr (F.F); rdelgado.hdoc@salud.madrid.org (R.D.); anna.bernardi@unimi.it (A.B.); javier.rojo@iiq.csic.es (J.R.).

## ■ ACKNOWLEDGMENT

We would like to acknowledge the financial support by the MICINN of Spain CTQ2008-01694, the EU RTN CARMUSYS (PITN-GA-2008-213592), Comune di Milano (Convenzione 55/2008), Sidaction — Ensemble contre le Sida, Fundación Mutua Madrileña (FMM2007-700) and the European FEDER funds. The expression plasmid for the glycoprotein of the Zaire strain of Ebola virus was kindly provided by Dr. A. Sanchez, Centers for Disease Control and Prevention, Atlanta, GA.

## REFERENCES

(1) Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology 2*, 97–130.

(2) Lundquist, J. J., and Toone, E. J. (2002) The cluster glycoside effect. *Chem. Rev. 102*, 555–578.

(3) Lee, R. T., and Lee, Y. C. (2000) Affinity enhacement by multivalent lectin-carbohydrate interaction. *Glycoconj. J. 17*, 543–551.

(4) Mammen, M., Choi, S.-K., and Whitesides, G. M. (1998) Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem., Int. Ed. 37*, 2754–2794.

(5) Lee, Y. C., and Lee, R. T. (1995) Carbohydrate-protein interactions: basis of glycobiology. *Acc. Chem. Res. 28*, 321–327.

(6) Chabre, Y. M., and Roy, R. (2010) Designs and creativity in synthesis of multivalent neoglycoconjugates. *Adv. Carb. Chem. Biochem. 63*, 168–393.

(7) Lahmann, M. (2009) Architectures of multivalent glycomimetics for probing carbohydrate-lectin interactions. *Top. Curr. Chem. 288*, 17–65.

(8) Curtis, B. M., Scharnowske, S., and Watson, A. J. (1992) Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120. *Proc. Natl. Acad. Sci. U.S.A. 89*, 8356–8360.

(9) Geijtenbeek, T. B. H., Kwon, D. S., Torensma, R., Van Vliet, S. J., Van Duijnhoven, G. C. F., Middel, J., Cornelissen, I. L. M. H. A., Nottet, H. S. L. M., 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.

(10) Geijtenbeek, T. B. H., Torensma, R., Van Vliet, S. J., Van Duijnhoven, G. C. F., Adema, G. J., Van Kooyk, Y., and Figdor, C. G. (2000) Identification of DC-SIGN, a novel dendritic cell−specific ICAM-3 receptor that supports primary immune responses. *Cell 100*, 575–585.

(11) Feinberg, H., Mitchell, D. A., Drickamer, K., and Weis, W. (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science 294*, 2163–2166.

(12) Mitchell, D. A., Fadden, A. J., and Drickamer, K. (2001) A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. *J. Biol. Chem. 276*, 28939–28945.

(13) van Kooyk, Y., Geijtenbeek, T. B. H. (2003) DC-SIGN: escape mechanism for pathogens. *Nat. Rev. Immunol. 3*, 697−709.

(14) Geijtenbeek, T. B. H., and van Kooyk, Y. (2003) Pathogens target DC-SIGN to influence their fate DC-SIGN functions as a pathogen receptor with broad specificity. *APMIS 111*, 698–714.

(15) Rojo, J., and Delgado, R. (2004) Glycodendritic structures: promising new antiviral drugs. *J. Antimicrob. Chemother. 54*, 579–581.

(16) Borrok, M. J., and Kiessling, L. L. (2007) Non-carbohydrate inhibitors of the lectin DC-SIGN. *J. Am. Chem. Soc. 129*, 12780–12785.

(17) Becer, C. R., Gibson, M. I., Geng, J., Ilyas, R., Wallis, R., Mitchell, D. A., and Haddleton, D. M. (2010) High-affinity glycopolymer binding to human DC-SIGN and disruption of DC-SIGN interactions with HIV envelope glycoprotein. *J. Am. Chem. Soc. 132*, 15130–15132.

(18) Sánchez-Navarro, M., and Rojo, J. (2010) Targeting DC-SIGN with multivalent systems. *Drug News and Perspectives 23*, 557–572.

(19) Arce, E., Nieto, P. M., Díaz, V., Castro, R. G., Bernad, A., and Rojo, J. (2003) Glycodendritic structures based on Boltorn hyperbranched polymers and their interactions with Lens Culinaris lectin. *Bioconjugate Chem. 14*, 817–823.

(20) Lasala, F., Arce, E., Otero, J. R., Rojo, J., and Delgado, R. (2003) Mannosyl glycodendritic structure inhibits DC-SIGN-mediated Ebola virus infection in cis and in trans. *Antimicrob. Agents Chemother. 47*, 3970–3972.

(21) Rojo, J., and Delgado, R. (2007) Dendrimers and dendritic polymers as anti-infective agents: new antimicrobial strategies for therapeutic drugs. *Anti-infect. Agents Med. Chem. 6*, 151–174.

(22) Reina, J. J., Sattin, S., Invernizzi, D., Mari, S., Martinez-Prats, L., Tabarani, G., Fieschi, F., Delgado, R., Nieto, P. M., Rojo, J., and Bernardi, A. (2007) 1,2-Mannobioside mimic: synthesis, DC-SIGN interaction by NMR and docking, and antiviral activity. *ChemMedChem 2*, 1030–1036.

(23) Sattin, S., Daghetti, A., Thépaut, M., Berzi, A., Sánchez-Navarro, M., Tabarani, G., Rojo, J., Fieschi, F., Clerici, M., and Bernardi, A. (2010) Inhibition of DC-SIGN-mediated HIV infection by a linear trimannoside mimic in a tetravalent presentation. *ACS Chem. Biol. 3*, 301–312.

(24) Annby, U., Malmberg, M., Pettersson, B., and Rehnberg, N. (1998) Benzylidene protected Bis-MPA A convenient dendrimer building block. *Tetrahedron Lett. 39*, 3217–3220.

(25) Goodwin, A. P., Lam, S. S., and Fréchet, J. M. J (2007) Rapid, efficient synthesis of heterobifunctional biodegradable dendrimers. *J. Am. Chem. Soc. 129*, 6994–6995.

(26) Ihre, H., Padilla de Jesús, O. L., and Fréchet, J. M. J. (2001) Fast and convenient divergent synthesis of aliphatic ester dendrimers by anhydride coupling. *J. Am. Chem. Soc. 123*, 5908–5917.

(27) Isomura, S., Wirsching, P., and Janda, K. D. (2001) An immunotherapeutic program for the treatment of nicotine addiction: hapten design and synthesis. *J. Org. Chem. 66*, 4115–4121.

(28) Ihre, H, Hult, A., Fréchet, J. M. J., and Gitsov, I. (1998) Double-stage convergent approach for the synthesis of functionalized dendritic aliphatic polyesters based on 2,2-Bis(hydroxymethyl)propionic acid. *Macromolecules 31*, 4061–4068.

(29) Malkoch, M., Malmström, E., and Hult, A. (2002) Rapid and efficient synthesis of aliphatic ester dendrons and dendrimers. *Macromolecules 35*, 8307–8314.

(30) Alvarez, C. P., Lasala, F., Carrillo, J., Muñiz, O., Corbí, A. L., and Delgado, R. (2002) C-Type lectins DC-SIGN and L-SIGN mediate cellular entry by ebola virus in *cis* and in *trans*. *J. Virol. 76*, 6841–6844.

(31) Tabarani, G., Thepaut, M., Stroebel, D., Ebel, C., Vives, 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*, 21229–21240.

(32) Halary, F., Amara, A., Lortat-Jacob, H., Messerie, M., Delaunay, T., Houlès, C., Fieschi, F., Arenzana-Seisdedos, F., Moreaus, J.-F., and Déchanet-Merville, J. (2002) Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell *trans*-infection. *Immunity 17*, 653–664.

(33) Yim, S.-H., Huh, J., Ahn, C.-H., and Park, T. G. (2007) Development of a novel synthetic method for aliphatic ester dendrimers. *Macromolecules 40*, 205–210.

(34) Anderson, G. T., Alexander, M. D., Taylor, S. D., Smithrud, D. B., Benkovic, S. J., and Weinreb, S. M. (1996) Catalytic antibodies in synthesis: design and synthesis of a hapten for application to the preparation of a scalemic pyrrolidine ring synthon for Ptilomycalin A. *J. Org. Chem. 61*, 125–132.

(35) Zhou, T., Chen, Y., Hao, L., and Zhang, Y. (2006) DC-SIGN and immunoregulation. *Cell. Mol. Immunol. 3*, 279–283.

(36) Caparrós, E., Muñoz, P., Sierra-Filardi, E., Serrano-Gómez, D., Puig-Kröger, A., Rodríguez-Fernández, J. L., Mellado, M., Sancho, J., Zubiaur, M., and Corbí, A. L. (2006) DC-SIGN ligation on dendritic cells results in ERK and PI3K activation and modulates cytokine production. *Blood 107*, 3950–3958.

(37) Shan, M., Klasse, P. J., Banerjee, K., Dey, A. K., Iver, S. P., Dionisio, R., Charles, D., Campbell-Gardener, L., Olson, W. C., Sanders, R. W., and Moore, J. P. (2007) HIV-1 gp120 mannose induce immunosupresive responses from dendritic cells. *PLoS Pathogens 3*, e169.

(38) van Kooyk, Y. (2008) C-type lectins on dendritic cells: key modulators for the induction of immune responses. *Biochem. Soc. Trans. 36*, 1478–1481.

(39) 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*, 1149–1157.

(40) Geijtenbeek, T. B. H., den Dunnen, J., and Gringhuis, S. I. (2009) Pathogen recognition by DC-SIGN shapes adaptive immunity. *Future Microbiol. 4*, 879–890.

(41) Gringhuis, S. I., den Dunnen, J., Litjens, M., van der Vlist, M., and Geijtenbeek, T. B. H. (2010) Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to Mycobacterium tuberculosis, HIV-1 and Helicobacter pylori. *Nat. Immunol. 10*, 1081–1089.

(42) den Dunnen, J., Gringhuis, S. I., and Geijtenbeek, T. B. H. (2010) Dusting the sugar fingerprint: C-type lectin signaling in adaptive immunity. *Immunol. Lett. 128*, 12–16.

(43) Feinberg, H., Castelli, R., Drickamer, K., Seeberger, P. H., and Weis, W. I. (2007) Multiple modes of binding enhance the affinity of DC-SIGN for high mannose *N*-linked glycans found on viral glycoproteins. *J. Biol. Chem. 282*, 4202–4209.

(44) Angulo, J., Reina, J. J., Tabarani, G., Fieschi, F., Rojo, J., and Nieto., P. M. (2008) Saturation transfer difference (STD) NMR spectroscopy characterization of dual binding mode of a mannose disaccharide to DC-SIGN. *ChemBioChem 9*, 2225–2227.

## ■ NOTE ADDED AFTER ASAP PUBLICATION

This paper was published on the Web on June 20, 2011 with a production error, the omission of the reagent lists in Schemes 1 and 2. The corrected version was reposted on June 23, 2011.

#### 9.1.4.3. *Structural characterization of psDi binding to DC-SIGN and evaluation of its selectivity to DC-SIGN versus langerin*

After psDi and psTri were found as relatively efficient DC-SIGN antagonists and psTri showed already described discrepancy, it became important to characterize of their interaction with DC-SIGN at a molecular level in order to find out, which of the two compounds is best to choose for further improvements. The following article by Thépaut et al. sheds light on the structural basis of psDi interaction with DC-SIGN: the crystal structure of DC-SIGN CRD in complex with psDi was solved and the binding properties in solution were evaluated by NMR techniques. This compound was found to have a sole binding mode, a desirable property for a molecule to be developed into a drug candidate, and this binding mode was different from its natural counterpart, Man$\alpha$1-2Man disaccharide. The revealed mode of psDi binding in the CRD of DC-SIGN gave the valuable structural knowledge for further rational optimization of this ligand.

Additionally, the selectivity of psDi to DC-SIGN versus langerin was assessed by SPR competition assay. When compared with Man$\alpha$1-2Man, psDi had an improved selectivity to DC-SIGN than langerin.

The main outcome of these studies: the structural properties of psDi interaction with DC-SIGN present all required qualities (unique binding mode, good synthetic availability, good selectivity to DC-SIGN) of a compound to be selected for further improvements.

Contributions:

DC-SIGN CRD was produced, the psDi/DC-SIGN co-crystals were obtained and the structure was solved in our group (Dr. Michel Thépaut) before my PhD started. The characterization of psDi interaction with DC-SIGN in solution was performed by a PhD student Cinzia Guzzi in the team of Dr. P. Nieto. The compounds were synthesized by the teams of Pr. A. Bernardi and Dr. J. Rojo.

My contribution to this study:

I have performed the SPR assays comparing DC-SIGN and langerin selectivity, analyzed the data, and participated in the preparation of the paper manuscript.

# 4

# 4

*Paper n$^{o}$4:* **Structure of a glycomimetic ligand in the Carbohydrate Recognition Domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design**

# Structure of a glycomimetic ligand in the Carbohydrate Recognition Domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design.

Michel Thépaut[1,2,3#], Cinzia Guzzi[5#], Ieva Sutkeviciute[1,2,3], Sara Sattin[4], Renato Ribeiro-Viana[5], Norbert Varga[4], Eric Chabrol[1,2,3], Javier Rojo[5], Anna Bernardi[4], Jesus Angulo[5], Pedro M. Nieto[5], and Franck Fieschi[1,2,6*]

[1]Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, Grenoble, F-38027, France; [2]CNRS, UMR 5075, Grenoble, F-38000, France; [3]CEA, DSV, Grenoble, F-38000, France; [4]Universita' di Milano, Dipartimento di Chimica Organica e Industriale, via Venezian 21, 20133 Milano - Italy; [5]Glycosystems Laboratory, Instituto de Investigaciones Químicas (IIQ), CSIC - Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla-Spain; [6]Institut Universitaire de France, 103 boulevard Saint-Michel 75005 Paris-France.

[#]*These authors contributed equally to the work.*
[*]*Corresponding author. E-mail: fieschi@ibs.fr; Tel: +33-(0)-4-38-78-91-77; Fax: +33-(0)-4-38-78-54-94.*

## Abstract

In genital mucosa, different fates are described for HIV according to the sub-type of dendritic cells (DCs) involved in its recognition. This notably depends on the C-type lectin receptor, langerin or DC-SIGN, involved in gp120 interaction. Langerin blocks HIV transmission by its internalization in specific organelles of Langerhans cells. On the contrary DC-SIGN enhances HIV trans-infection of T lymphocytes. Thus, approaches aiming to inhibit DC-SIGN, without blocking langerin, represent attractive anti-HIV strategies. We previously demonstrated that dendrons bearing multiple copies of glycomimetic compounds were able to block DC-SIGN-dependent HIV infection in cervical explant models. Optimization of such ligand requires detailed characterization of its binding mode. In the present work we determined the first high-resolution structure of a glycomimetic/DC-SIGN complex by X-ray crystallography. This glycomimetic, pseudo-1,2-mannobioside, shares shape and conformational properties with Manα1-2Man, its natural counterpart. However, it uses the binding epitope previously described for Lewis X, a ligand specific for DC-SIGN among the C-type lectin family. Thus, selectivity gain for DC-SIGN *vs* langerin is observed with pseudo-1,2-mannobioside as shown by surface plasmon resonance analysis. In parallel, ligand binding was also analyzed by TR-NOESY and STD NMR experiments, combined with the CORCEMA-ST protocol. These studies demonstrate that the complex, defined by X-ray crystallography, represents the unique binding mode of this ligand as opposed to the several binding orientations described for the natural ligand. This exclusive binding mode and its selective interaction properties position this glycomimetic as a good lead compound for rational improvement based on a structurally driven approach.

## Introduction

Dendritic cells (DCs) are key players in the initial response to pathogens as they are the first participants in the long series of events in host-pathogen interaction leading to activation of specific T-cells.[1,2] They are found in epidermal and mucosal tissues and are thus able to quickly recognize new invading pathogens through the identification of pathogen associated molecular patterns (PAMPs). The efficacy of DCs in their sentinel role is related to the wide diversity of pattern recognition receptors (PRRs) they express on their surface. Apart from the Toll Like receptors family of PRRs, they also possess C-type lectin receptors (CLRs) dedicated to the specific recognition of pathogen carbohydrate patterns.[3] Among these CLRs, DC-SIGN (Dendritic Cell-Specific ICAM-3 Grabbing Non-integrin) has attracted a great deal of attention during the past decade. Initially highlighted for its role in HIV transmission to T cells,[4] it has then been identified as a PRR hijacked by many other pathogens, for instance some viruses, bacteria, fungi and parasites, to escape immune response in their infectious processes.[5,6] More recently, DC-SIGN has also been involved in the modulation of the immune response.[7]

DC-SIGN is a type II transmembrane protein with a short cytosolic region, a transmembrane segment and an extended extracellular domain (ECD) projecting its carbohydrate binding domain up to 320 Å above the cell surface, as a bait to trap potential antigens.[8] This extracellular domain is divided into two structurally and functionally distinct regions: a neck region, involved in the tetramerization of the receptor, and a calcium-dependent carbohydrate recognition domain (CRD), which is at the heart of the molecular recognition processes mediated by DC-SIGN.

Therefore, many groups are developing strategies to block the sugar binding site within DC-SIGN CRD to prevent its use by pathogens' glycoproteins.[9-12] DC-SIGN/pathogen interactions are complex and imply multipoint attachment benefiting from the DC-SIGN tetrameric state and from its organization into clustered patches at the cell membrane.[13,14] For this reason almost all of these inhibition strategies exploit a multiple ligand presentation platform (polymers, dendrimers, or nanoparticles), on which the relevant ligands are presented for interaction with DC-SIGN CRD.[15-18] Several "proof of concept" studies have been performed using a simple mannose as grafted ligands onto these various display systems.[9,11,19] Indeed, we initially demonstrated that dendrimers loaded with multiple copies of mannose were able to inhibit DC-SIGN/gp120 interaction.[9] However, mannose, as natural ligand, is not specific enough for *in vivo* practical applications. The design of a ligand with good selectivity and basal affinity is of crucial importance. Several groups have invested efforts in DC-SIGN ligand improvements by increasing the complexity of the oligosaccharide (reconstructing natural high mannose oligosaccharide for instance[10,15,17]), by selecting non-sugar molecules with high throughput screening[20,21] or by designing new synthetic molecules mimicking natural sugar properties.[22,23,24,25]

Based on two natural ligands, Manα1-2Man disaccharide and Lewis X trisaccharide, we have developed glycomimetic compounds that have good affinity for DC-SIGN and low structural complexity. In the mannose series (Figure 1), we recently demonstrated that dendrons or dendrimers bearing multiple copies of pseudo-1,2-mannobioside (**1**) were able to block DC-SIGN-dependent HIV trans-infection of T cells,[16] HIV infection in cervical explant models,[26] and Ebola-pseudotyped viral infection.[16,18]

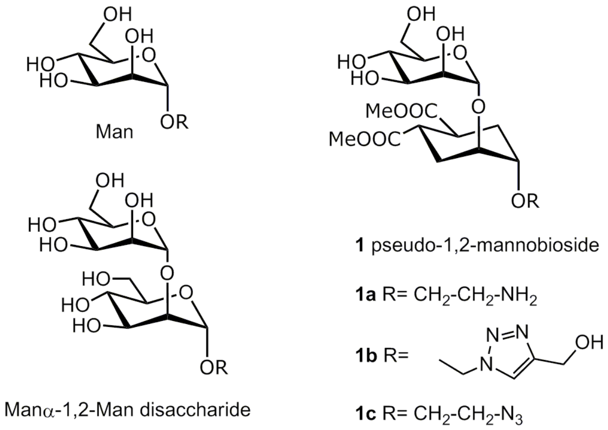


**Figure 1.** Structure of compounds used in this study.

We also developed a class of Lewis X mimics containing a fucosylamide anchor and demonstrated their potential in terms of specificity towards DC-SIGN relative to langerin, a related lectin.[25] Due to their chemical stability, resistance to enzymatic degradation by glycosidases,[27] and their rather simple and high yield synthetic route, these glycomimetics of first generation represent good candidates for optimization leading to efficient specific inhibitors for DC-SIGN.

To enable ligand improvement, a good knowledge of the particular binding mode of model compounds is of the utmost importance. To achieve this goal, glycoscientists often face a particularly hard and specific task given by the nature of glycan/protein interactions. Indeed, oligosaccharide ligands often have several binding modes within a single binding site in $Ca^{2+}$ dependent lectins. [28-30] Moreover, computational tools are often of limited use in the prediction of sugar-lectin interaction modes due to the peculiar properties of lectin binding sites. As a consequence, ligand modifications aiming to improve one binding mode may disfavour another one, leading to unpredictable global effects on the affinity for the receptor. Therefore, the optimization process becomes a challenge. For DC-SIGN, it has been shown that the natural ligand Manα1-2Man disaccharide displays at least two different binding modes within the $Ca^{2+}$ binding site[28,29] : the major mode is achieved by coordination to the $Ca^{2+}$ atom of the 3-OH and 4-OH equatorial groups of the reducing end mannose residue, the minor one occurs through the same groups of the non-reducing end mannose ring. The pseudo-1,2-

mannobioside **1** and the natural mannobioside Man$\alpha$1-2Man are known to share a similar shape and similar conformational properties.[27] However, **1** contains a single mannose unit at the non-reducing end, and could be expected to bind DC-SIGN similarly to the minor binding mode of Man$\alpha$1-2Man. Preliminary docking and NMR analysis of its binding properties to DC-SIGN suggested that several orientations of the ligand within the binding site were still possible and none could be selected based on the available data.[22]

In the present work we describe the first high-resolution structure of a glycomimetic ligand in complex with DC-SIGN. Using X-ray crystallography on crystallized DC-SIGN CRD/**1** complex, we have been able to characterize at a molecular level an unpredicted binding mode for such compound with the CRD. To obtain a dynamic picture of the interaction in solution, ligand binding was also analyzed by NMR spectroscopy. Transfer NOE (TR-NOESY) and Saturation Transfer Difference (STD) NMR experiments were used combined with the CORCEMA-ST protocol, which enables prediction of STD intensities from the Cartesian atomic coordinates of the ligand-receptor complex. These studies demonstrated that in solution **1** binds as observed in the solid state, and therefore the complex defined by X-ray represents the unique binding mode of this ligand in the $Ca^{2+}$ binding site of DC-SIGN. Due to this exclusive binding mode, this ligand represents a good lead compound for a rational ligand optimization procedure and the X-ray structure reported here represents a powerful tool for virtual screening and docking as a guide to new chemical improvement of this compound.

In addition, the structural analysis of this first high-resolution glycomimetics/DC-SIGN complex, as well as its comparison with other lectin binding sites, sets the basis for the design of efficient and specific DC-SIGN inhibitors. In particular, the interaction of **1** with langerin was also examined by SPR, leading to the satisfactory conclusion that **1** present improved selectivity for DC-SIGN. Langerin, a mannose binding C-type lectin expressed on Langerhans cells[31] is present in the same mucosal environment as interstitial DCs expressing DC-SIGN, but it is known to possess a protective action against HIV infection, which should not be antagonized by therapeutic entities targeted against DC-SIGN.[32] It was recently reported that langerin binds effectively to Man$\alpha$1-2Man containing oligosaccharides, and the X-ray structure of this protein in complex with the Man$\alpha$1-2Man disaccharide was described.[33] Thus the conclusion that **1** is a selective DC-SIGN binder is non-trivial and of great importance for further improvement *en route* to a potential therapeutic use of these Man-based glycomimetics in the prevention of sexually transmitted HIV infection.

## Results and discussion

### *X-ray crystal structure of the complex of DC-SIGN CRD with pseudo-1,2-mannobioside **1a**.*

In order to characterize the binding mode of the glycomimetics and to acquire structural data for the optimization process, co-crystallization experiments of DC-SIGN CRD with some of the produced glycomimetic compounds were assayed. Crystals of monomeric DC-SIGN CRD in the presence of pseudo-1,2-mannobioside **1a** (Figure 1) were obtained in crystallization conditions derived from those reported for the DC-SIGN CRD/Manα1-2Man.[28] The crystals contain one copy of the CRD in the asymmetric unit with a $P4_32_12$ space group.

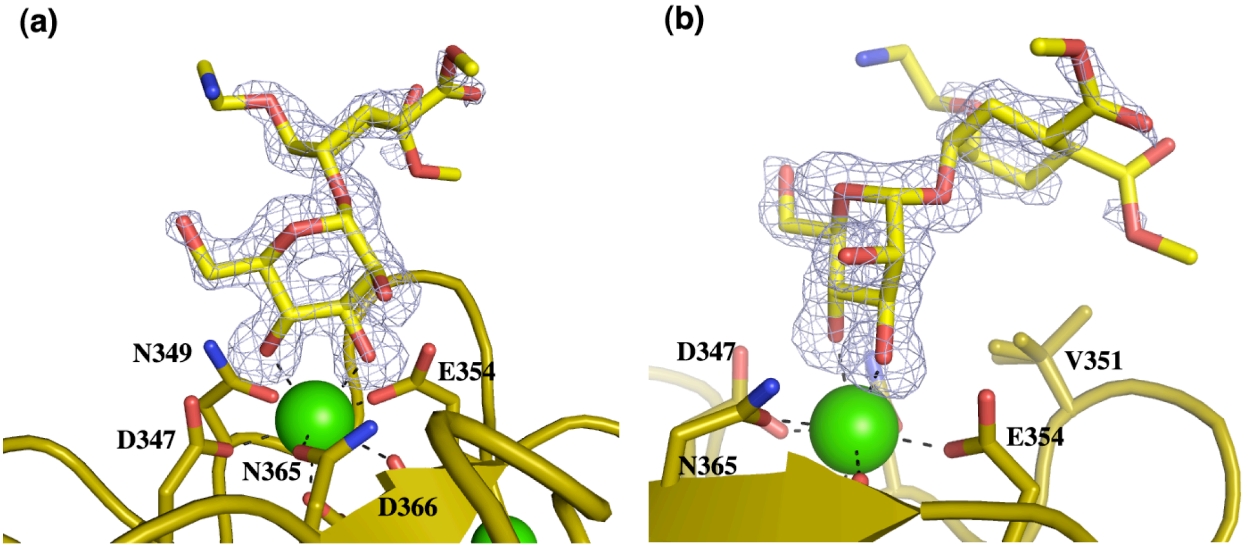


**Figure 2.** Electron density maps for bound pseudo-1,2-mannobioside. (a) and (b) correspond to two different views. The bound ligand is shown superimposed on the Fo-Fc electron density map (gray, 3 σ contour). Residues involved in $Ca^{2+}$ binding are visualized and labeled. V351 exhibits two alternative conformations with 50% occupancy in the crystal structure. Both of them are represented in b).

After structure resolution from a model without sugar, an electron density was observed on the $Ca^{2+}$ ion, confirming the presence of the pseudo-sugar in the canonical carbohydrate binding site of the CRD (Figure 2A). The structure of this complex has been solved at 1.42 Å resolution. Comparison with previously reported structures of DC-SIGN complexes, with natural ligands such as Manα1-2Man or Lewis X derivatives, shows that structure of the proteins, and more particularly residues involved in the binding site, are well conserved (see supporting information, Figure S1). As for many other mannose-based ligands, **1a** directly binds to the $Ca^{2+}$ ion through coordination bonds with equatorial 3-OH and 4-OH groups of the non-reducing mannose unit. The mannose unit, as well as the cyclohexane ring, is clearly visualized, as shown by the electron density map (Figure 2). On the contrary, the electron-density of the dimethyl ester substituents on the cyclohexane moiety (Figure 2) as well as the ethylamine appendage are poorly or not visible suggesting that these parts of the molecule remain flexible within the complex. Three water molecules are associated to the ligand and only one of them connects the molecule to the protein (see supporting information, Figure S2).

Mannose 3-OH and 4-OH groups also interact with the $Ca^{2+}$ ligand residues Glu 347, Asn 349, Glu 354 and Asn 365 through hydrogen bonds, and an additional interaction is observed between the 2-OH group of the mannose unit and Asn 365. The cyclohexane ring of the ligand contributes exclusively to the binding through van der Waals contacts with Val 351 side chain. The conformation of the ligand corresponds to the extended conformation (phi O1-C1-O2'-C2' is 66.3° ; psi'C1-O1-C2'-C1' is -141,11 °) that has been described as the most abundant in solution for the mimic.[27] The ethylamine linker on the cyclohexane is free to move and oriented towards the solvent, far from the protein surface. This suggests that functionalization of the pseudo-1,2-mannobioside through this position onto multivalent presenting scaffolds should not affect the observed binding mode in solution. Indeed, the crystal structure of the DC-SIGN CRD/pseudo-1,2-mannobioside observed here may represent a good starting point for molecular improvement.

### *NMR analysis of the interaction of pseudo-1,2-mannobioside with DC-SIGN in solution.*

Previous examples demonstrated that DC-SIGN recognizes carbohydrate ligands in a multimodal fashion, whereas X-ray diffraction studies can only show one of the possible bound poses, losing the structural information on other conformers and/or the other binding modes.[28] Despite the high resolution of DC-SIGN/ligands complex structures solved by X-ray crystallography, additional approaches have been crucial to get a deeper structural insight to carbohydrate/lectin binding modes. Particularly, in a previous work, STD NMR enabled the characterization of both binding modes of a synthetic Manα1-2Man disaccharide,[29] whilst the X-ray study was not able to fully determine the second binding mode.[28] This multimodal binding feature seems to be a common characteristic of DC-SIGN for its ligands, as NMR data on the molecular recognition between this lectin and other synthetic glycan mimics have also revealed multiple binding modes of the ligands.[30] Taking into account the above considerations, a quantitative comparison between the X-ray data obtained for the DC-SIGN/**1a** complex with data from STD NMR studies in solution was planned. The binding of **1a** has been previously studied by STD NMR techniques,[22] but the epitope was not fully characterized due to a strong spectral overlap of the NMR signals. Namely, the methylene protons in the cyclohexane unit (H3 (C) and H6 (C), both axial and equatorial) appeared at degenerated chemical shifts, precluding individual integration of their STD signals. In the course of parallel studies in the laboratory, ligand **1c** (R= $CH_2$-$CH_2$-$N_3$) was used to prepare multivalent glycodendrimers via 1,3-dipolar cycloaddition catalyzed by Cu(I), a “click chemistry” reaction. This approach produced the formation of a triazole ring which induced a significant chemical shift variation of some key signals corresponding to the cyclohexane unit, leading to a better spectral dispersion.

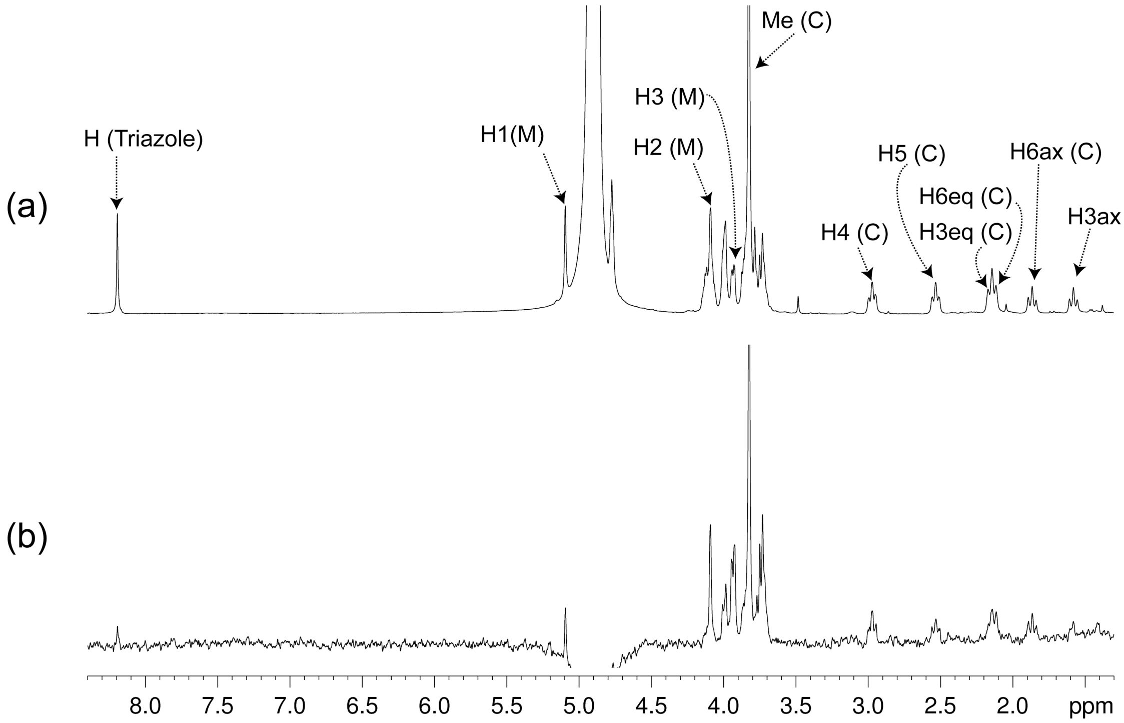
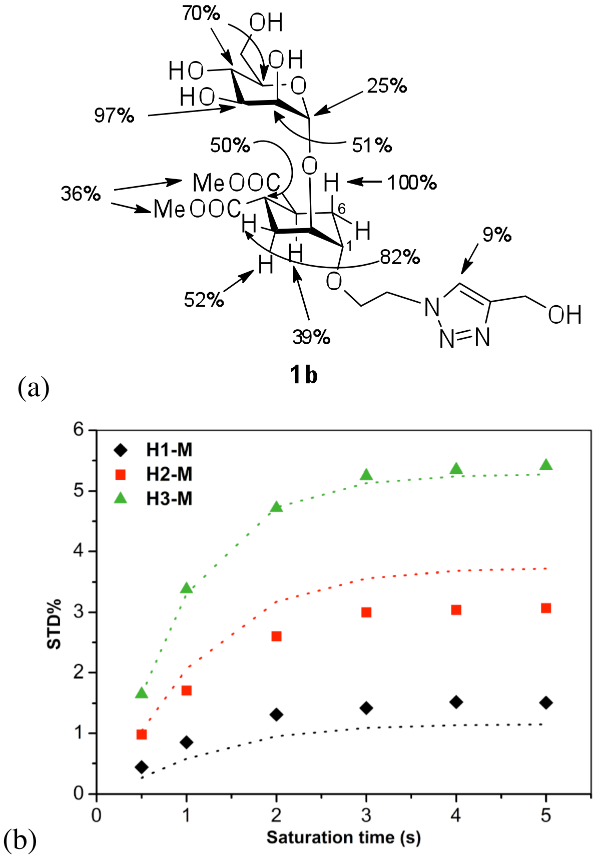
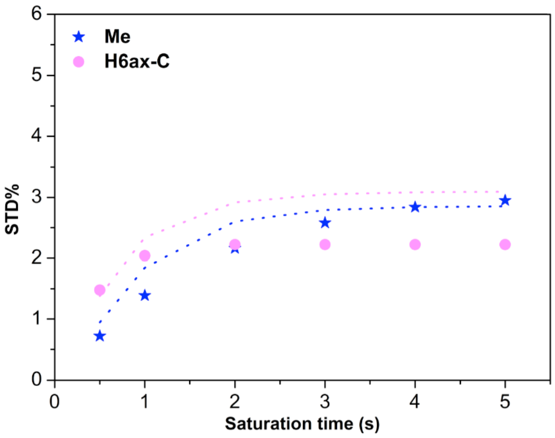


**Figure 3.** STD NMR study of the interaction of ligand 1b with DC-SIGN in solution. (a) $^1$H NMR reference spectrum (off-resonance frequency 40 ppm) and (b) STD spectrum (on-resonance frequency 0 ppm) of a sample containing 1 mM of **1b** and 19 μM of DC-SIGN ECD, at 25ºC (500 MHz). Key proton signals are labeled in (a).

Thus, we modified the structure of **1** including a triazole ring in the spacer at the pseudo-anomeric position, generating compound **1b**. This modification was expected to help signal analysis, without modifying the interaction mode of the ligand with the protein. Indeed, the NMR data (Figure 3) support this hypothesis: the triazole residue did not affect binding to DC-SIGN neither in terms of affinity (STD NMR competition experiment, see Supporting Information) nor in terms of protein-ligand contacts, (the STD signals of the triazole residue were basically null, and the pattern of intensities comparable to that of ligand **1a;** compare Figure 4 and previously described STD signals for **1a**[22]). Furthermore, as expected, this monovalent compound showed chemical shifts similar to those observed in the multivalent systems with little signal overlapping in the $^1$H-NMR spectrum and, in particular, well-resolved signals for protons H3ax (C) and H6ax (C). Therefore, compound **1b** turned out to be a very adequate model to study the binding process with DC-SIGN by STD NMR.

The ability to distinguish both axial protons allowed us to demonstrate experimentally that proton H6ax received a considerably larger amount of saturation from the protein than H3ax (Figure 3). In fact, H3ax could not be integrated accurately in the STD spectrum due to its low signal-to-noise ratio, whilst H6ax showed the strongest STD intensity among the ligand protons, along with the mannose proton H3 (M). This result indicated that H6ax (C) is in very close contact with non-exchangeable protons of some side chains of the protein binding pocket in the bound state. Interestingly, inspection of the structure of the complex obtained by X-ray diffraction (Figure 2) leads to a straightforward explanation of the strong saturation transferred from the protein to H6ax (C) of **1b**, as this proton sits on top of the methyl groups of the Val 351 side chain. In the experimental set-up, irradiation at 0 ppm leads to a very efficient saturation of these methyl protons of DC-SIGN, in agreement with the strong STD effect observed for the adjacent H6ax ligand proton.

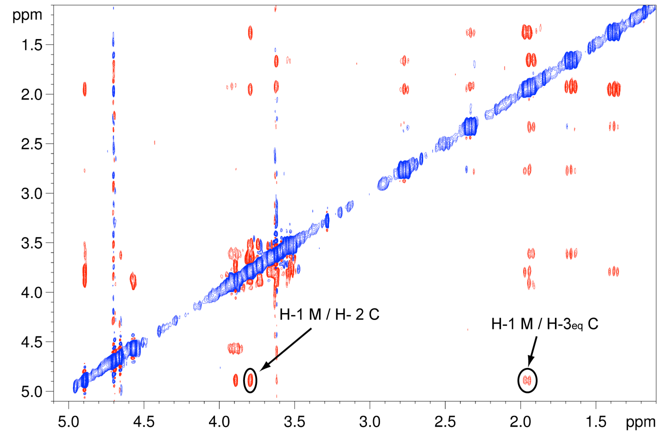




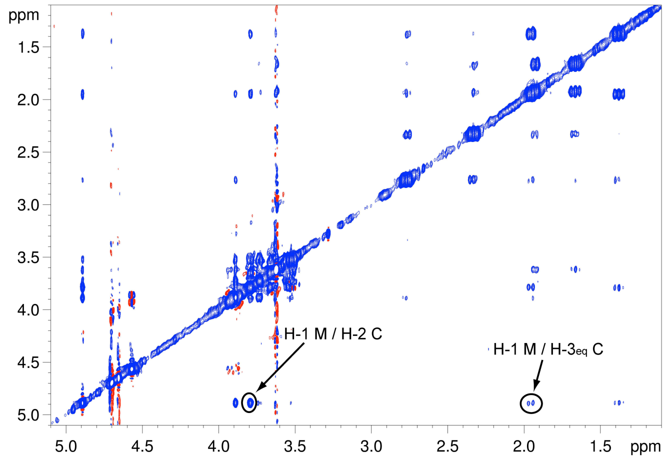
(c)

**Figure 4.** STD growth curves and CORCEMA-ST analysis. (a) Ligand epitope map of **1b** at 25° C. (b) STD build-up curves of the mannose residue of **1b**, (c) STD build-up curves of the cyclohexyl residue of **1b**. Theoretical STD intensities predicted by CORCEMA-ST using the Cartesian coordinates of the X-ray structure are shown in dashed lines; experimental data in symbols.

Having defined the binding epitope, a quantitative analysis of the agreement between the determined X-ray structure of the complex and the NMR data in solution was carried out using full matrix relaxation calculations and the Cartesian coordinates from X-ray diffraction (Figure 4). No refinement of the X-ray coordinates was carried out during CORCEMA-ST calculations, and only those protons with the most intense and well-isolated STD signals were considered in the analysis. Figure 4 compares the experimental STD build-up curves with the corresponding theoretical predictions by CORCEMA-ST calculations for the mannose (Figure 4b) and cyclohexyl (Figure 4c) residues of **1b**.[34] A good fit between theoretical and experimental curves is evident, quantitatively confirmed by the resulting low R-NOE factor of 0.2 (see Experimental Section).[35,36] This demonstrates that the binding mode of the ligand in solution, as detected by NMR, is the same as the one observed in the crystal. This result strongly supports the notion that DC-SIGN binds this ligand in a single orientation, which corresponds to that observed in the X-ray structure.



(a)



(b)

**Figure 5.** Expansions of NOESY experiments at 25ºC (500 MHz) of 1b. (a) Free state (1mM **1b**) mixing time 600 ms. (b) In the presence of 19 $\mu$M DC-SIGN ECD, mixing time 300 ms. Labels indicate some key NOE peaks.

NOESY experiments were also performed at different mixing times (see Experimental Section) to investigate the conformation of **1b** in solution, both free and bound to DC-SIGN. The resulting spectra (Figure 5) suggested that no changes occur in ligand conformation upon binding to the protein receptor, as no significant differences in the NOE fingerprints were found. Signal overlap precludes a full analysis of residual conformational motion of the ligand in the protein binding site. However, some key cross peaks clearly observed in TR-NOESY spectra support the notion that conformation observed in the X-ray structure is well represented in solution. In fact, the experimental ligand interproton distances from NMR, calculated using the Isolated Spin Pair Approximation (ISPA), displayed good agreement with distances measured on the crystal structure (see Supporting Information Table S2).

*Binding mode comparison of pseudo-1,2-mannobioside with its natural model, Man$\alpha$1-2Man.*

From the combination of the two biophysical structural approaches described above (X-ray and NMR), we demonstrated a unique binding mode for pseudo-1,2-mannobioside **1** and provided a detailed molecular view of it. Initially, **1** was designed as a structural mimic of Manα1-2Man. Notably, both molecules exhibit an equilibrium between two conformations in solution (stacked and extended) with similar φ and ψ angles around the glycosidic linkage.[27] Two alternative binding modes were previously characterized for Manα1-2Man (Figure 6A)[29] in DC-SIGN, each using a different mannose unit for $Ca^{2+}$ coordination. The absence of hydroxyl groups in the cyclohexane moiety of **1** originally suggested that interaction would occur along the Manα1-2Man minor binding mode (Figure 6A), which involves the non-reducing end mannose and would possibly generate a favorable interaction between F313 of the secondary binding site and the cyclohexane moiety of **1**. Surprisingly, **1** was found to interact through a third, distinct mode: the position of the Ca-binding diol (mannose 3-OH and 4-OH) is swapped relative to the Manα1-2Man complex, so that the 3-OH group of **1** contacts E354 (as opposed to the 4-OH group of Manα1-2Man) and that the 4-OH group of **1** contacts E347 (as opposed to the 3-OH group of Manα1-2Man). The observed conformation of **1** is close to the extended one and the cyclohexane ring participates in the interaction through van der Waals contacts with V351. Thus, despite their largely documented structural similarity, **1** and Manα1-2Man have totally different DC-SIGN binding modes (Figures 6A and 6B). On the contrary, strong similarity is observed with the binding epitope of Lewis X derivatives (lacto-N-fucopentaose III, 1SL5) (Figure 6B), which also exploits van der Waals contacts with V351.

Interestingly, whilst several DC-SIGN related C-type lectins, such as DC-SIGNR and langerin, are capable to recognize mainly mannose-based oligosaccharides, DC-SIGN is specific for Lewis-type carbohydrates as well.[37] In the case of DC-SIGNR, this difference has been previously attributed to the substitution of V351 in DC-SIGN by a serine in DC-SIGNR. Guo et al[37] showed that a simple mutation reversing this serine to a valine can convert DC-SIGNR to a Lewis X binding lectin. This indicates that Val 351 represents a key residue modulating the specificity of the primary DC-SIGN binding site. The stabilization of **1** through V351 suggests that a similar selectivity exists for our compound.

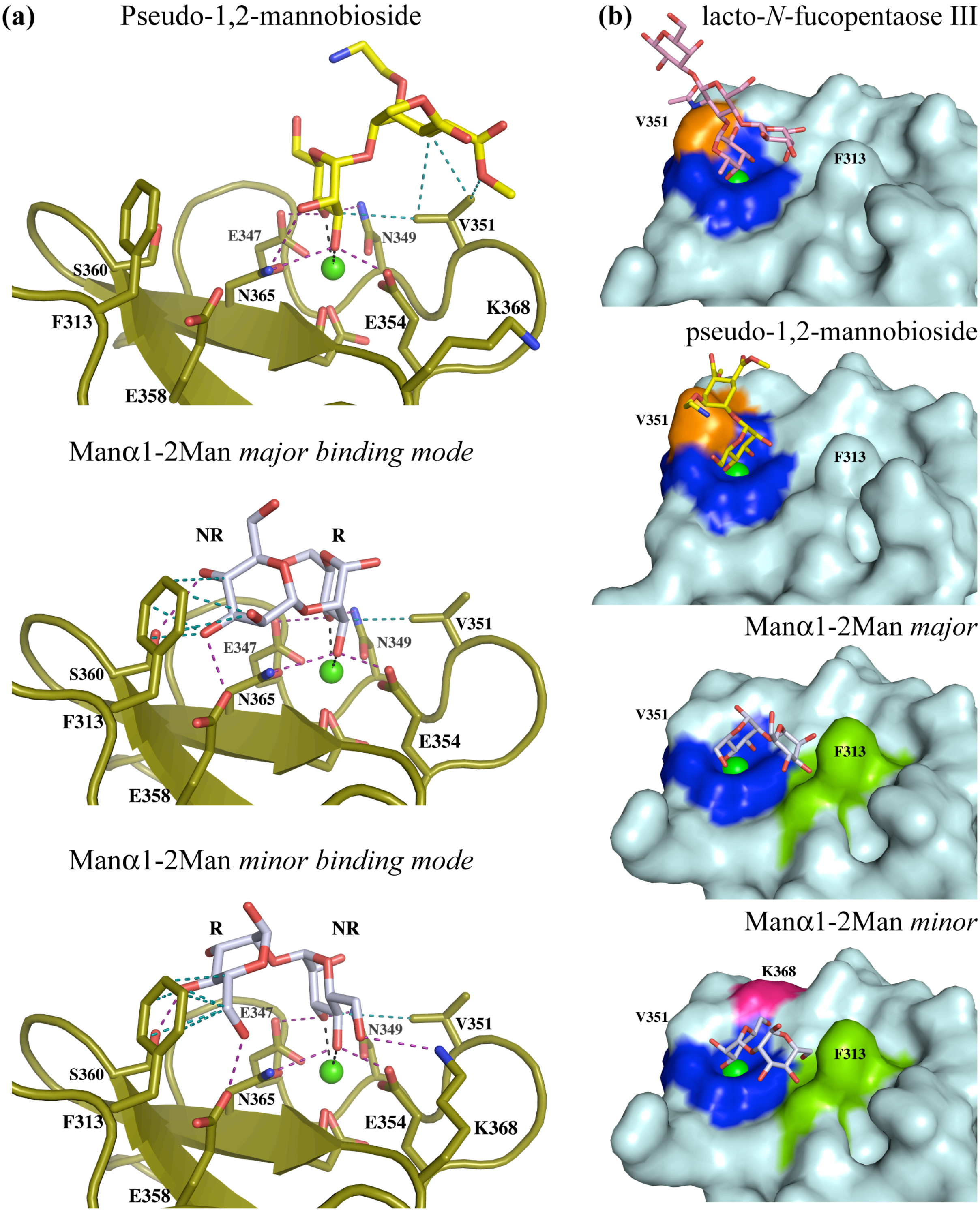

(a)
Pseudo-1,2-mannobioside
E347
N349
V351
S360
N365
F313
E354
K368
E358
Manα1-2Man *major binding mode*
NR
R
E347
N349
V351
S360
N365
F313
E354
E358
Manα1-2Man *minor binding mode*
R
NR
E347
N349
V351
S360
N365
F313
E354
K368
E358
(b)
lacto-*N*-fucopentaose III
V351
F313
pseudo-1,2-mannobioside
V351
F313
Manα1-2Man *major*
V351
F313
Manα1-2Man *minor*
K368
V351
F313

**Figure 6.** Binding of pseudo-1,2-mannobioside to DC-SIGN and comparison with the binding of other mannose and fucose-based oligosaccharides. (a) Comparison of pseudo-1,2-mannobioside **1a** and Manα1-2Man binding modes. The protein is shown in olive, the pseudo-1,2-mannobioside is in yellow, Manα1-2Man in light grey, with nitrogen, oxygen, and calcium represented as blue, red, and green spheres, respectively. Hydrogen bonds are shown as dashed purple lines, $Ca^{2+}$ coordination bonds are dashed black lines, and key van der Waals interactions are indicated by dashed blue lines. Both

Manα1-2Man binding modes are oriented highlighting the reducing (R) and non-reducing (NR) ends. (b) Comparison of carbohydrate binding surface epitope for LNFP III, pseudo-1,2-mannobioside **1a**, and both binding modes of Manα1-2Man (see Figure S4 for structure of LNFP III). Residues common to binding of all carbohydrates and involved in $Ca^{2+}$ chelation are in dark blue, residues specific for both Manα1-2Man binding modes are in green (F313, S360 and E358), K368 specific for Manα1-2Man minor binding mode is in pink, V351 involved in LNFPIII and pseudo-1,2-mannobioside **1a** binding is in orange. (a) and (b) are representations with two opposite side views, V351 and P313 are highlighted in B) for appropriate orientation. In (a) only one of the two alternatives side chain conformations of V351 is represented for clarity.

*DC-SIGN/langerin specificity.*

As mentioned above, selectivity is a key concern in the selection of lead compounds for the development of DC-SIGN antagonists. The main issue involves selectivity versus langerin, a C-type lectin known to protect the host against HIV infection, which should not be antagonized by DC-SIGN targeting compounds.[32,38]

The specificity of **1a** for DC-SIGN and langerin was evaluated using a previously described surface plasmon resonance (SPR) competition assay.[25,39] **1a** was compared to mannose and to the natural Manα1-2Man disaccharide for its ability to inhibit binding of the extracellular domains (ECD) of DC-SIGN or langerin to a mannosylated BSA (BSA-Man) surface (Figure 7A and 7B). In order to evaluate the interaction avidity effect for the lectins, we prepared two flow cells with different BSA-Man density (Fc-HD and Fc-LD in Figure 7 corresponding respectively to flow cell of high density and low density). It is evident that BSA-Man density had virtually no effect on the inhibition of DC-SIGN binding, and for this lectin an improvement by a factor of 3 in the apparent affinity was obtained for both Manα1-2Man and **1a** compared to mannose (Figure 7b, $IC_{50}$ of 730 and 880 $\mu$M, respectively, compared to 2.4 mM for mannose). Thus the natural disaccharide and the mimic display same affinity for DC-SIGN in this assay.

On the contrary, although the general pattern of the compound potencies remained the same for both surfaces (Figure 7B), it was harder to achieve inhibition of langerin binding to the Fc-HD surface (5000 RU) than to the Fc-LD (1350 RU), with a lower BSA-Man density,. For instance, at lower Man-BSA density Manα1-2Man and **1a** have an $IC_{50}$ of 1658 and 6553 $\mu$M, respectively, that become 562 and 21,124 $\mu$M at higher density. This phenomenon depends on the fact that, as previously described and detailed in the experimental section,[25] langerin displays a measurable affinity to the dextran matrix on the chip surface, and thus dextran/BSA-Man surface must be considered as a combined heterogenous ligand of langerin. The different ratios of dextran and BSA-Man, depending directly from the BSA-Man density, contribute to a complex relative affinity of the surface for langerin.

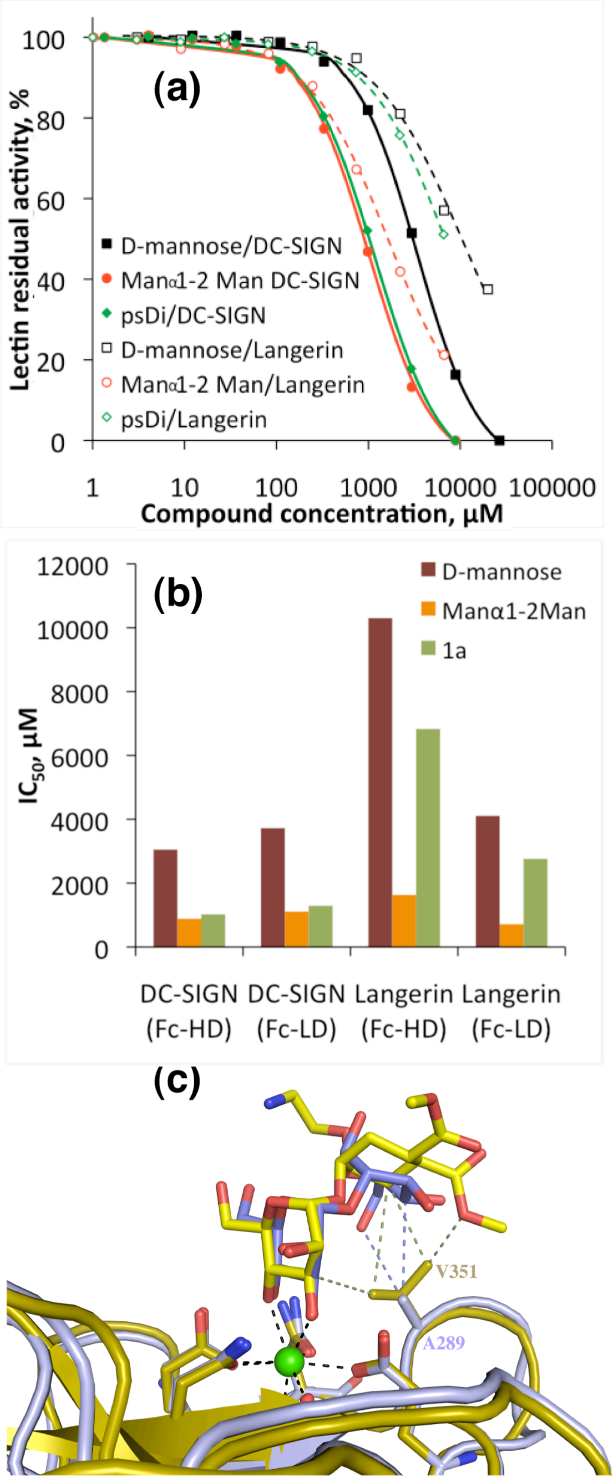


**Figure 7.** Specificity of pseudo1,2-mannobioside as a function of the C-type lectin. (a) SPR experiment results of the inhibition of DC-SIGN ECD and langerin ECD binding to BSA-Man/dextran surface by mannose, Manα1-2Man and pseudo-1,2-mannobioside. Results for higher BSA-Man density are represented. (b) Lectins selectivity : IC$_{50}$ histograms for two different flow cells (Fc-HD – 5000 RU, Fc-LD – 1350 RU of immobilized BSA-Man). (c) Superposition of Manα1-2Man/langerin and **1a**/DC-SIGN complex structures. Langerin and D-SIGN backbones are in blue and yellow respectively. Manα1-2Man and **1a** are represented in blue and yellow. Van der Waals interactions of Manα1-2Man, with A289 of

langerin, are represented as dashed blue lines while those of **1a,** with V351 of DC-SIGN, are as dashed yellow lines.

The dependence of the $IC_{50}$ values on the BSA-Man surface-density in the langerin assay, prevented a straightforward comparison of the $IC_{50}$ values obtained for the two lectins. However, comparison of the data in Figure 7b clearly shows that moving from Manα1-2Man to **1a** a loss of affinity is observed for langerin while the affinity is conserved for DC-SIGN. An attempt of quantifying the selectivity effect in relative terms is proposed in the supporting information (Figure S5). This finding further supports the hypothesis of a relation between DC-SIGN selectivity and the use of a special binding region, shared by Lewis type derivatives.

The crystal structure of a langerin CRD/Manα1-2Man complex has recently been published.[33] As for DC-SIGN, two binding modes have been observed with alternatively the reducing or the non reducing mannose involved in the primary binding through $Ca^{2+}$ chelation (see supporting information, Figure S6). The major binding mode of Manα1-2Man to langerin is close to the binding mode of **1** to DC-SIGN (Figure 7c) and the main difference is the Val 351 of DC-SIGN, which is replaced by Ala 289 in langerin (see Figure S6 and 7c). Comparing the major binding mode for the Manα1-2Man/langerin complex and the **1**/DC-SIGN structure, replacement of Ala 289 (in langerin) to Val 351 (in DC-SIGN) significantly increases van der Waals contacts within the complex. Moreover, in langerin Manα1-2Man establishes a contact with Ala 289 through the hydroxyl group of the C-6 that fills the distance to the methyl of the alanine side chain. This requires previous desolvation of the 6-OH group, an energetically costly process. In the equivalent position, **1** does not have any substituent. Formation of a **1**/langerin complex is disfavored simultaneously due to the loss of a bulky side chain in this position, valine to alanine, and to the absence of a group extending from the cyclohexane ring and capable to establish van der Waals contact with the alanine side chain. Therefore, as observed previously for the Lewis X DC-SIGN *vs* DC-SIGNR specificity, Val 351 is making the difference in the specificity of **1** for DC-SIGN**,** relative to langerin. The replacement by a serine in DC-SIGNR or an alanine in langerin, although being a moderate modification, is sufficient to lose Lewis X specificity in one case and to ensure here a preference of **1** for DC-SIGN with respect to langerin.

## Conclusions

Multiple carbohydrate binding modes have been reported as a mechanism allowing improved affinity towards DC-SIGN.[28] Indeed, it is interesting to note that in our assay **1** presents the same $IC_{50}$ as Manα1-2Man, while it has only an unique binding mode. This suggests that intrinsically **1** binds stronger than each individual binding mode of Manα1-2Man. Trying to mimic Manα1-2Man, which is also a good ligand of DC-SIGN, by using a pseudo-1,2-mannobioside, we finally discovered that this molecule mimics Lewis X binding mode to DC-SIGN, by exploiting lipophilic interactions with V351 side chain. Although this was not anticipated, it is a very interesting and convenient result considering that among a wide range of C-type lectins, DC-SIGN is the only one able to bind Lewis X.[40] This result underlines that V351 is a key residue to target for the design of DC-SIGN specific inhibitors. Indeed, this structure suggests that addition of lipophilic groups to the cyclohexane scaffold of **1** should allow to extend contacts with the protein surface, which might improve the affinity and the DC-SIGN specificity to higher level. Each affinity improvement at the monovalent level will be even more amplified upon grafting onto multivalent presenting scaffolds. Future work in our network is going along this line: the development of DC-SIGN-adapted glycomimics based on a structurally driven approach.

## Experimental Procedures.

### *Synthesis of compounds.*

**Compound 1a.** The synthesis of **1a** was previously described by Sattin et al.[16]

**Compound 1b.** Pseudo-1,2-mannobioside **1c** (10 mg, 0.02 mmol), propargyl alcohol (1.8 mg, 0.03 mmol), CuBr (2.2 mg, 0.015 mmol) and TBTA (1.6 mg, 0.03 mmol) were dissolved in 1 mL of THF/$H_2O$ (1:1). After 18 h, the solvent was evaporated and the resulting crude was purified by silica gel column chromatography ($CH_2Cl_2$/MeOH 9.5:0.5, 8.5:1.5, 8:2), affording 8.3 mg (75% yield) of compound **1b** as an oil. $^1$H NMR ($D_2O$) 300 MHz δ ppm: 8.06 (s, 1H, $H_{triazol}$), 4.96 (d, 1H, $J_{1\text{-}2}$ = 1.64 Hz, H-$1_{mann}$), 4.79 (s, 2H, $C\underline{H}_2OH$), 4.64 (t, J = 4.8 Hz, 2H, $C\underline{H}_2N$), 4.01-3.93 (m, 3H, $OC\underline{H}_2CH_2N$, H-$2_{mann}$), 3.90-3.84 (m, 2H, $C\underline{H}OCH_{mann}$, H-$6_{mann}$), 3.83-3.78 (m, 1H, H-$3_{mann}$), 3.76-3.66 (m, 8H, $C\underline{H}OCH_2CH_2N$, H-$6'_{mann}$, $CH_3$), 3.63-3.57 (m, 2H, H-$4_{mann}$, H-$5_{mann}$), 2.90-2.79 (m, 1, $C\underline{H}COOCH_3$), 2.47-2.35 (m, 1, $C\underline{H}COOCH_3$), 2.06-1.96 (m, 2H, $C\underline{H}_{2eq}C$, $C\underline{H}_{2eq}C$), 1.79-1.68 (m, 1H, $C\underline{H}_{2ax}C$), 1.51-1.39 (m, 1H, $C\underline{H}_{2ax}C$). $^{13}$C NMR ($D_2O$) 75 MHz δ ppm: 177.4, 177.2 (C=O), 146.8 ($C_{triazol}$), 124.6 ($\underline{C}H_{triazol}$), 98.5 (C-$1_{mann}$), 73.7 (C-$5_{mann}$), 73.4 ($\underline{C}HOCH_2CH_2N$), 70.8 (C-$3_{mann}$), 70.3 ($O\underline{C}H_2CH_2N$, C-$2_{mann}$), 66.7 (C-$4_{mann}$), 66.5 ($\underline{C}HOCH$), 60.9 (C-$6_{mann}$), 54.6 ($\underline{C}H_2OH$), 52.5 ($\underline{C}H_3O$), 50.3 ($OCH_2\underline{C}H_2N$), 38.8 ($\underline{C}HCOOCH_3$), 38.6 ($\underline{C}HCOOCH_3$), 26.5 ($C\underline{C}H_2C$), 26.4 ($C\underline{C}H_2C$). $[\alpha]^{25}_D$ = + 27.1° (c 0.6, MeOH). ESI-MS calc. for $C_{21}H_{33}N_3O_{12}Na$ (m/z): 542.2; found: 541.9 $[M+Na]^+$. HRMS (FAB) calc for $C_{21}H_{33}N_3O_{12}Na$ (m/z): 542.1962 found: 542.1979 $[M+Na]^+$.

### *Cloning and expression of recombinant DC-SIGN S-CRD.*

The sequence coding for carbohydrate recognition domain of DC-SIGN, comprising amino acids 254 to 404, was obtained by PCR using the forward primer 5'-gcattaggtctctgcgcatgcacccctgtccctggga-3' and the reverse primer 5'-gcagcaggtctcttatcactacgcaggagggggggtttg-3'. The PCR template used was a previous construct for DC-SIGN ECD overexpression which has been described previously.[8] The PCR product was inserted into pASK-IBA7plus vector (IBA GmbH), at the Bsa I sites, in phase with a Strep-Tag II sequence and a factor Xa cleavage site, both located at N-terminal end of the protein (the CRD with this N-terminal tag will be called hereafter S-CRD). The resulting plasmid was sequence checked and used to transform calcium competent *E. coli* BL21(DE3).

Culture was initiated from a 5% dilution of an overnight culture into LB medium with 100 mg/L ampicillin. Cells were grown for 3 h at 37°C, and DC-SIGN S-CRD expression was induced by addition of 1 mg/L anhydrotetracycline for 4 additional hours. Cells were harvested by centrifugation at 5000 g for 20 min. The protein was expressed as inclusion bodies and then a refolding step was required prior to the purification procedures.

*Protein purification.*

The pellet, containing DC-SIGN S-CRD, obtained from a 1 L culture was resuspended in 30 mL of buffer A (150 mM NaCl and 25 mM Tris pH 8). Cells were lysed by freezing at –20°C, thawing and sonication with addition of a tablet of protease inhibitors (Roche Diagnostics). Inclusion bodies were isolated by centrifugation at 100,000 g for 30 min at 4°C, and using a Potter-Elvehjem homogenizer were resuspended in 30 mL buffer A supplemented with 2 M urea and 1% Triton X-100, and recovered by a second centrifugation. Inclusion bodies were washed again with buffer A and solubilized in buffer A supplemented with 6 M guanidine hydrochloride and 0.01% 2-mercaptoethanol. Inclusion bodies solution was centrifuged at 100,000 g for 1 h at 4°C, supernatant was diluted in 120 mL buffer of 1.25 M NaCl, 25 mM Tris pH 8 and 25 mM $CaCl_2$, and the resulting solution was dialyzed overnight against 880 mL buffer of 25 mM Tris pH 8. Refolding of inclusion bodies was achieved by extensive dialysis in buffer A supplemented with 4 mM $CaCl_2$ (buffer A'), and insoluble compounds were removed by the last centrifugation step at 100,000 g for 1 h at 4°C.

Purification was then performed as previously described for langerin S-CRD[41]. Briefly, the first step of DC-SIGN S-CRD purification was performed by affinity chromatography on a 15 mL Strep-Tactin superflow column (IBA GmbH) equilibrated in buffer A' and eluted in same buffer supplemented with 2.5 mM desthiobiotin. A second purification step was performed with DC-SIGN S-CRD-containing fractions concentrated to 1.5 mL. Functional DC-SIGN S-CRD was separated from non-functional protein as it was delayed on a 15 mL mannose-agarose column equilibrated in buffer A'. Functional DC-SIGN S-CRD-containing fractions were concentrated to 11 mg/mL.

DC-SIGN ECD and Langerin ECD were produced and purified as previously described[8,42].

*DC-SIGN CRD crystallization.*

Crystallization assays were performed manually by hanging-drop vapor-diffusion method at 293 K in EasyXtal plates (Qiagen) with protein-carbohydrate stock solution made by mixing 10 $\mu$L of concentrated DC-SIGN S-CRD with 1 $\mu$L of 100 mM pseudo-1,2-mannobioside **1a**. Drops, prepared by mixing 1 $\mu$L reservoir solution with 1 $\mu$L protein-carbohydrate stock solution, were equilibrated against 1 mL reservoir solution. The best crystals appeared when reservoir solution was composed of 20% PEG 3350, 100 mM cacodylate pH 6.5 and 200 mM NaCl. A 0.06 × 0.06 × 0.65 mm crystal was cryoprotected in Paratone-N (Hampton Research) and was flash-frozen in liquid nitrogen.

*Data collection and processing.*

X-ray diffraction data were collected at ID29 beamline at ESRF Grenoble. Two datasets were collected at a wavelength of 0.9809 Å, with a crystal-to-detector distance set to 197.89 mm and an X-ray transmission of 8.1%. The first dataset was collected at maximal resolution and the second with reduced exposure time to minimize detector saturation at lower resolution. The first dataset, composed of 360 images, was collected with an oscillation range of 0.5° per image and an exposure time of 1 s. The second dataset, composed of 120 images, was collected with an oscillation range of 3° per image and an exposure time of 0.2 s.

Datasets were processed and merged using the programs XDS and XSCALE, respectively.[43] Statistics of data processing are summarized in Table 1. Matthews coefficient was calculated using the program MATTHEWS_COEF.[44]

*Phasing, model building and structure refinement*

Phasing was performed by molecular replacement with a model built from a structure of DC-SIGN CRD (pdb code 2IT6) depleted of calcium ions, carbohydrate molecules and water molecules. The best solution resulting from molecular replacement was used as the starting model for structure refinement. The structure refinement was performed by cycling between manual building using the program COOT[44] and energy minimization with the program REFMAC 5 from the CCP4 package.[44] Statistics of structure refinement are summarized in Table 1. Crystal structure of DC-SIGN CRD/**1a** has been deposited in Protein Data Bank under PDB code 2xr5.

*Surface Plasmon Resonance experiments.*

Surface plasmon resonance experiments were performed on a Biacore 3000 using a CM4 chip, functionalized at 5 $\mu$L/min. Flow cells (Fc) 1, 2 and 3 were activated with 50 µL of an 0.2 M EDC/0.05 M NHS mixture, after this step Fc2 and Fc3 were functionalized with mannosylated bovine serum albumine (BSA-Man), and finally remaining activated groups of both flow cells were blocked with 30 µL of 1 M ethanolamine. After blocking, the three flow cells were treated with 10 $\mu$L of 10 mM HCl to remove unspecific bound proteins and 20 $\mu$L of 50 mM EDTA to expose surface to regeneration protocol. After these steps 5000 RU and 1170 RU of BSA-Man were immobilized on the surfaces of Fc2 and Fc3, respectively. BSA-Man stock solution was prepared by dissolving the glycoprotein in water to final 1 mg/mL concentration, and for immobilization it was diluted to 60 μg/mL in a buffer of 10 mM sodium acetate pH 4.

To control surface activity and to determine optimal working protein concentration, 13 $\mu$L of samples of increasing concentrations of DC-SIGN and langerin ECDs (0.7–46.7 $\mu$M for DC-SIGN, and 0.7–45.7 $\mu$M for langerin) were injected onto the surfaces. The selected concentration was 20 $\mu$M for both lectins.

For inhibition studies, samples of each lectin mixed with increasing concentrations of inhibiting compounds were prepared in a running buffer (buffer A' supplemented with 0.005% P20 surfactant), and 13 μL of each sample was injected onto the surfaces at a 5 μL/min flow rate. Concentrations of inhibiting compounds ranged from 4.1 to 26,667 μM or from 3.1 to 20,000 μM for D-mannose, from 4.1 to 8,889 $\mu$M or from 1.0 to 6,667 μM for Manα1-2Man, and from 1.3 $\mu$M to 8,782 $\mu$M or from 1.0 μM to 6,587 μM for pseudo-1,2-mannobioside **1a** mixed with DC-SIGN or langerin ECDs, respectively. The resulting sensorgrams were reference surface corrected, except in the case of langerin, since this lectin displayed affinity to the dextran matrix, and thus dextran/BSA-Man surface was considered as a combined ligand of langerin (as described previously in Andreini et al).[25]

$$y = R_{hi} - \frac{R_{hi} - R_{lo}}{1 + \left(\frac{Conc}{A_1}\right)^{A_2}} \qquad (1)$$

$$IC_{50} = A_1 \cdot \left(\frac{R_{hi} - R_{lo}}{R_{hi} - 50} - 1\right)^{\frac{1}{A_2}} \qquad (2)$$

The lectin binding responses were extracted from the sensorgrams, converted to percent residual activity values ($y$) with respect to lectin alone binding, and plotted against corresponding compound concentration. The 4-parameter logistic model (eq, 1) was fitted to the plots, and the $IC_{50}$ values were calculated using the values of fitted parameters ($R_{hi}$, $R_{lo}$, $A_1$, and $A_2$) and equation 2.

*NMR spectroscopy experiments.*

NMR spectroscopy experiments were performed on a Bruker Avance DRX 500 MHz spectrometer equipped with a 5 mm inverse triple-resonance probe head. NMR samples were prepared in 550 μL of 99.9% $D_2O$ and for the experiments in presence of the receptor, in buffer $D_2O$ (150 mM NaCl, 4 mM $CaCl_2$, 25 mM d-Tris, pD 8) and with 19 μM DC-SIGN ECD. The concentration of the ligand was 1.35 mM, and the same sample was used for both, STD NMR, and TR-NOESY experiments. For the two-ligand equimolar competition experiment, a concentration of 1.7 mM was employed.

STD NMR experiments were carried out at 10, 25 and 35ºC, by using a train of Gaussian shaped pulses of 49 ms (field strength of ca. 80 Hz), an inter-pulse delay of 1 ms[45] and 15 ms spin-lock pulse (field strength of 3.7 kHz) prior to acquisition. The on-resonance frequency was set to 0 ppm and the off-resonance frequency was 40 ppm. Appropriate blank experiments were used to ensure the lack of direct saturation of the ligand protons. Saturation times of 0.5, 1, 1.5, 2, 3, 4 and 5 s were used to obtain the STD build-up curves.

The binding epitope was characterized by the analysis of initial slopes of the STD intensities at 25° C [46]: the experimental ($I_0 - I_{sat} / I_0$) curves were fitted to an exponential function described by the equation: $STD(t_{sat}) = STD_{max}(1 - e^{-ksat \cdot tsat})$, which allows to calculate STD at zero saturation time (initial slopes) by multiplying the resulting parameters $STD_{max}$ and $k_{sat}$.[47] The epitope is obtained by normalization of the whole set of initial slopes against the highest value, and expressing the result in percentage.

NOESY experiments were carried out using a phase sensitive pulse program with gradient pulses in the mixing time and with presaturation.[48,49] Mixing times of 0.2, 0.3, 0.4, 0.5, 0.6, and 0.8 s were used for NOESY spectra and 0.1, 0.2, 0.3, and 0.5 s for TR-NOESY spectra. NOE build-up curves were obtained from the normalized cross-peak volumes (ratio cross peak over diagonal peak) as function of the mixing time. For TR-NOESY experiments the growth was approximately linear up to 300 ms (see Figure S7 in the Supporting Information). Longitudinal cross-relaxation rates were obtained by averaging the ratio of the normalized volume and the mixing time, for mixing times 0.1, 0.2, and 0.3 s. From them, intramolecular ligand proton-proton distances were obtained by using the ISPA approach (isolated spin pair approximation), and taking the distance H3eq - H3ax of the cyclohexyl ring as a reference.

*CORCEMA-ST:*

The Cartesian coordinates of the crystal structure of the complex DC-SIGN CRD/**1a** were employed for the full relaxation matrix calculations. As no chemical shift assignment of the protein protons was available, they were predicted by using the program SHIFTX.[50] Although the experimental irradiation frequency for selective saturation was established at 0 ppm, all the protein protons with chemical shifts predicted to be within the [0.7, -0.7] ppm range were included, as SHIFTX does not consider the effects of line broadening under the experimental conditions. All exchangeable hydrogen atoms were excluded in the calculations, as the STD NMR experiments were performed in $D_2O$. We assumed that pdb coordinates for the bound and free protein were identical and several cycles were performed to reach the optimized parameters. For this protein-ligand system, the classical assumption of an association step limited by diffusion (on-rate $10^8$ $M^{-1}s^{-1}$) was considered. The off-rate was varied within a range that yielded a final dissociation constant within the micromolar range, typical for this kind of ligands of DC-SIGN. No large variations were observed of the R-factor during this process, and the final off-rate was 40 kHz. Assuming a spherical shape for the protein tetramer, the correlation time of bound ligand was set to 115 ns whereas 0.5 ns was used for the free ligand, and 10 ps for the internal correlation time of methyl groups. This might be considered an oversized value for the correlation time of a protein of about 155 kDa (DC-SIGN tetramer). Nevertheless this seems to be not uncommon in CORCEMA-ST calculations, [51-53] particularly when the protein shape deviates from a perfect globular shape, as it is the case with the DC-SIGN tetramer. To reduce the dimensions of the matrices, a cut off of 8 Å from the ligand was used. The STD intensities for each binding mode were calculated as percentage fractional intensity changes, $\mathbf{S}_{\mathrm{calc,k}}$, from the intensity matrix **I(t)** ($\mathbf{S}_{\mathrm{calc,k}} = (([(\mathbf{I}_{\mathbf{0k}} - \mathbf{I(t)}_{\mathbf{k}})*100]/ \mathbf{I}_{\mathbf{0k}})$, where **k** is a particular proton in the complex, and $\mathbf{I}_{\mathbf{0k}}$ its thermal equilibrium value)[34] and the calculation was carried out for the set of saturation times experimentally measured (0.5, 1, 1.5, 2, 3, 4 and 5 s). The theoretical STD values were compared to the experimental ones using the NOE R-factor[35,36] defined as:

$$\sqrt{\frac{\sum Wk\left(STD_k^{\exp} - STD_k^{calc}\right)^2}{\sum Wk\left(STD_k^{\exp}\right)^2}}$$

In this equation $STD_k^{\exp}$ and $STD_k^{calc}$ refer to experimental and calculated STD values for proton **k**. The best agreement with experimental data was achieved using the B chain of the residue V351.

## Acknowledgment

For the financial support we are grateful to: Sidaction: Ensemble contre le SIDA for Michel Thépaut post-doctoral grant and support, EU ITN Marie-Curie program (CARMUSYS - Grant number 213592) for funding Cinzia Guzzi, Ieva Sutkeviciute, Renato Ribeiro-Viana and Norbert Varga. J. A. acknowledges financial support from the MICINN through the Ramon y Cajal program.

## Reference

(1) Banchereau, J.; Steinman, R. M. *Nature* **1998**, *392*, 245–252.
(2) Bell, D.; Young, J. W.; Banchereau, J. *Adv Immunol* **1999**, *72*, 255–324.
(3) van Kooyk, Y.; Engering, A.; Lekkerkerker, A. N.; Ludwig, I. S.; Geijtenbeek, T. B. *Curr Opin Immunol* **2004**, *16*, 488–493.
(4) 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.; van Kooyk, Y. *Cell* **2000**, *100*, 587–597.
(5) van Kooyk, Y.; Geijtenbeek, T. B. *Nat Rev Immunol* **2003**, *3*, 697–709.
(6) Geijtenbeek, T. B.; van Kooyk, Y. *Apmis* **2003**, *111*, 698–714.
(7) Hodges, A.; Sharrocks, K.; Edelmann, M.; Baban, D.; Moris, A.; Schwartz, O.; Drakesmith, H.; Davies, K.; Kessler, B.; McMichael, A.; Simmons, A. *Nat Immunol* **2007**, *8*, 569–577.
(8) Tabarani, G.; Thépaut, M.; Stroebel, D.; Ebel, C.; Vivès, C.; Vachette, P.; Durand, D.; Fieschi, F. *J Biol Chem* **2009**, *284*, 21229–21240.
(9) Tabarani, G.; Reina, J. J.; Ebel, C.; Vivès, C.; Lortat-Jacob, H.; Rojo, J.; Fieschi, F. *FEBS Lett* **2006**, *580*, 2402–2408.
(10) Martínez-Avila, O.; Hijazi, K.; Marradi, M.; Clavel, C.; Campion, C.; Kelly, C.; Penadés, S. *Chemistry* **2009**, *15*, 9874–9888.
(11) Becer, C. R.; Gibson, M. I.; Geng, J.; Ilyas, R.; Wallis, R.; Mitchell, D. A.; Haddleton, D. M. *J Am Chem Soc* **2010**, *132*, 15130–15132.
(12) Ciobanu, M.; Huang, K.-T.; Daguer, J.-P.; Barluenga, S.; Chaloin, O.; Schaeffer, E.; Mueller, C. G.; Mitchell, D. A.; Winssinger, N. *Chem Commun* **2011**, *47*, 9321–9323.
(13) Cambi, A.; de Lange, F.; van Maarseveen, N. M.; Nijhuis, M.; Joosten, B.; van Dijk, E. M.; de Bakker, B. I.; Fransen, J. A.; Bovee-Geurts, P. H.; van Leeuwen, F. N.; Van Hulst, N. F.; Figdor, C. G. *J Cell Biol* **2004**, *164*, 145–155.
(14) de Bakker, B. I.; de Lange, F.; Cambi, A.; Korterik, J. P.; van Dijk, E. M.; Van Hulst, N. F.; Figdor, C. G.; Garcia-Parajo, M. F. *Chemphyschem* **2007**, *8*, 1473–1480.
(15) Martínez-Avila, O.; Bedoya, L. M.; Marradi, M.; Clavel, C.; Alcamí, J.; Penadés, S. *ChemBioChem* **2009**, *10*, 1806–1809.
(16) Sattin, S.; Daghetti, A.; Thépaut, M.; Berzi, A.; Sánchez-Navarro, M.; Tabarani, G.; Rojo, J.; Fieschi, F.; (null); Bernardi, A. *ACS Chem Biol* **2010**, *5*, 301–312.
(17) Wang, S.-K.; Liang, P.-H.; Astronomo, R. D.; Hsu, T.-L.; Hsieh, S.-L.; Burton, D. R.; Wong, C.-H. *Proc Natl Acad Sci USA* **2008**, *105*, 3690–3695.
(18) Luczkowiak, J.; Sattin, S.; Sutkeviciute, I.; Reina, J. J.; Sánchez-Navarro, M.; Thépaut, M.; Martinez-Prats, L.; Daghetti, A.; Fieschi, F.; Delgado, R.; Bernardi, A.; Rojo, J. *Bioconjug Chem* **2011**, *22*, 1354–1365.
(19) Lasala, F.; Arce, E.; Otero, J. R.; Rojo, J.; Delgado, R. *Antimicrob. Agents Chemother.* **2003**, *47*, 3970–3972.
(20) Borrok, M. J.; Kiessling, L. L. *J Am Chem Soc* **2007**, *129*, 12780–12785.
(21) Tran, T. H.; Baz, El, R.; Cuconati, A.; Arthos, J.; Jain, P.; Khan, Z. K. *J Antivir Antiretrovir* **2011**, *3*, 49–54.
(22) Reina, J. J.; Sattin, S.; Invernizzi, D.; Mari, S.; Martinez-Prats, L.; Tabarani, G.; Fieschi, F.; Delgado, R.; Nieto, P. M.; Rojo, J.; Bernardi, A. *ChemMedChem* **2007**, *2*, 1030–1036.
(23) Mitchell, D. A.; Jones, N. A.; Hunter, S. J.; Cook, J.; Jenkinson, S. F.; Wormald, M. R.; Dwek, R. A.; Fleet, G. W. J. *Tetrahedron: Asymmetry* **2007**, *18*, 1502–1510.
(24) Garber, K. C. A.; Wangkanont, K.; Carlson, E. E.; Kiessling, L. L. *Chem Commun (Camb)* **2010**, *46*, 6747–6749.
(25) Andreini, M.; Doknic, D.; Sutkeviciute, I.; Reina, J. J.; Duan, J.; Chabrol, E.; Thépaut, M.; Moroni, E.; Doro, F.; Belvisi, L.; Weiser, J.; Rojo, J.; Fieschi, F.; Bernardi, A. *Org Biomol Chem* **2011**, *9*, 5778–5786.
(26) Berzi, A.; Reina, J. J.; Ottria, R.; Sutkeviciute, I.; Antonazzo, P.; Sánchez-Navarro, M.; Chabrol, E.; (null); (null); Cetin, I.; Rojo, J.; Fieschi, F.; Bernardi, A.; Clerici, M. *AIDS* **2012**, *26*, 127–137.
(27) Mari, S.; Posteri, H.; Marcou, G.; Potenza, D.; Micheli, F.; Jiménez-Barbero, J.; Bernardi, A. *European Journal of Organic Chemistry* **2004**, *2004*, 5119–5225.
(28) Feinberg, H.; Castelli, R.; Drickamer, K.; Seeberger, P. H.; Weis, W. I. *J Biol Chem* **2007**, *282*, 4202–4209.
(29) Angulo, J.; Díaz, I.; Reina, J. J.; Tabarani, G.; Fieschi, F.; Rojo, J.; Nieto, P. M. *ChemBioChem* **2008**, *9*, 2225–2227.
(30) Guzzi, C.; Angulo, J.; Doro, F.; Reina, J. J.; Thépaut, M.; Fieschi, F.; Bernardi, A.; Rojo, J.; Nieto, P. M. *Org Biomol Chem* **2011**, *9*, 7705–7712.

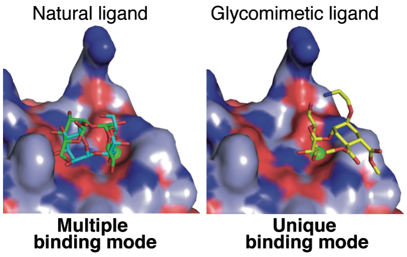
(31) Valladeau, J.; Ravel, O.; Dezutter-Dambuyant, C.; Moore, K.; Kleijmeer, M.; Liu, Y.; Duvert-Frances, V.; Vincent, C.; Schmitt, D.; Davoust, J.; Caux, C.; Lebecque, S.; Saeland, S. *Immunity* **2000**, *12*, 71–81.
(32) de Witte, L.; Nabatov, A.; Geijtenbeek, T. B. H. *Trends Mol Med* **2008**, *14*, 12–19.
(33) Feinberg, H.; Taylor, M. E.; Razi, N.; Mcbride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K.; Weis, W. I. *J Mol Biol* **2011**, *405*, 1027–1039.
(34) Jayalakshmi, V.; Rama Krishna, N. *J. Magn. Reson.* **2004**, *168*, 36–45.
(35) Xu, Y.; Sugár, I. P.; Krishna, N. R. *J Biomol NMR* **1995**, *5*, 37–48.
(36) Krishna, N. R.; Agresti, D. G.; Glickson, J. D.; Walter, R. *Biophys J* **1978**, *24*, 791–814.
(37) Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I.; Drickamer, K. *Nat Struct Mol Biol* **2004**, *11*, 591–598.
(38) de Witte, L.; Nabatov, A.; Pion, M.; Fluitsma, D.; de Jong, M. A. W. P.; de Gruijl, T.; Piguet, V.; van Kooyk, Y.; Geijtenbeek, T. B. H. *Nat Med* **2007**, *13*, 367–371.
(39) Timpano, G.; Tabarani, G.; Anderluh, M.; Invernizzi, D.; Vasile, F.; Potenza, D.; Nieto, P. M.; Rojo, J.; Fieschi, F.; Bernardi, A. *ChemBioChem* **2008**, *9*, 1921–1930.
(40) Lee, R. T.; Hsu, T.-L.; Huang, S. K.; Hsieh, S.-L.; Wong, C.-H.; Lee, Y. C. *Glycobiology* **2011**, *21*, 512–520.
(41) Thépaut, M.; Vivès, C.; Pompidor, G.; Kahn, R.; Fieschi, F. *Acta Crystallogr F Struct Biol Cryst Commun* **2008**, *64*, 115–118.
(42) Thépaut, M.; Valladeau, J.; Nurisso, A.; Kahn, R.; Arnou, B.; Vivès, C.; Saeland, S.; Ebel, C.; Monnier, C.; Dezutter-Dambuyant, C.; Imberty, A.; Fieschi, F. *Biochemistry* **2009**, *48*, 2684–2698.
(43) Kabsch, W. *J. Appl. Cryst.* **1993**, *26*, 795–800.
(44) Collaborative Computational Project, N. 4. *Acta Crystallogr. D* **1994**, *50*, 760–763.
(45) Mayer, M.; Bernd, M. *Angew. Chem. Int. Ed.* **1999**, *38*, 1784–1788.
(46) Mayer, M.; Meyer, B. *J Am Chem Soc* **2001**, *123*, 6108–6117.
(47) Mayer, M.; James, T. L. *J Am Chem Soc* **2004**, *126*, 4453–4460.
(48) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
(49) Wagner, R.; Berger, S. *Journal of Magnetic Resonance, Series A* **1996**, *123*, 119–121.
(50) Neal, S.; Nip, A. M.; Zhang, H.; Wishart, D. S. *J Biomol NMR* **2003**, *26*, 215–240.
(51) Yuan, Y.; Bleile, D. W.; Wen, X.; Sanders, D. A. R.; Itoh, K.; Liu, H.-W.; Pinto, B. M. *J Am Chem Soc* **2008**, *130*, 3157–3168.
(52) Kemper, S.; Patel, M. K.; Errey, J. C.; Davis, B. G.; Jones, J. A.; Claridge, T. D. W. *J. Magn. Reson.* **2010**, *203*, 1–10.
(53) Canales, A.; Rodríguez-Salarichs, J.; Trigili, C.; Nieto, L.; Coderch, C.; Andreu, J. M.; Paterson, I.; Jiménez-Barbero, J.; Díaz, J. F. *ACS Chem Biol* **2011**, *6*, 789–799.

Table I: DC-SIGN CRD/pseudo-1,2-mannobioside **1a** complex data collection and structure refinement statistics.

| | |
|---|---|
| Data collection statistics | |
| Wavelength (Å) | 0.9809 |
| Space group | $P4_3\,2_1\,2$ |
| Unit cell parameters (Å) | a = b = 71.45; c = 52.67 |
| Resolution (Å) | 42.41-1.42 (1.46-1.42)[a] |
| Measured reflections | 658501 (12715) |
| Unique reflections | 25925 (1761) |
| Completeness (%) | 99.4 (92.4) |
| $I/\sigma\,(I)$ | 28.67 (5.56) |
| $R_{\mathrm{merge}}$ [b] (%) | 13.0 (58.0) |
| Structure refinement statistics | |
| Resolution (Å) | 42.41-1.42 (1.46-1.42) |
| Refinement factors | |
| Used reflections/free (%) | 24628 / 5.0 |
| $R_{\mathrm{cryst}}$ [c] | 0.145 |
| $R_{\mathrm{free}}$ [c] | 0.172 |
| rmsd from ideality | |
| Bond lengths (Å) | 0.030 |
| Bond angles (deg) | 2.413 |
| Ramachandran plot (%) | |
| Most favored regions | 87.0 |
| Additional allowed regions | 11.3 |
| Generously allowed regions | 1.7 |
| Disallowed regions | 0.0 |
| Average $B$-factor (Å$^2$) | 13.30 |

[a] Values in parentheses are for the highest resolution shell. [b] $R_{\mathrm{merge}} = \Sigma_h \Sigma_m |I_m(h) - \langle I(h) \rangle| \Sigma_h \Sigma_m I_m(h)$. [c] $R_{\mathrm{cryst}} = \Sigma||F_0| - |F_c|| \Sigma |F_0|$, and $R_{\mathrm{free}} = R_{\mathrm{cryst}}$ calculated with 5% of $F_0$ sequestered before refinement.

## Table of content graphic :


Natural ligand
Glycomimetic ligand
Multiple binding mode
Unique binding mode

#### 9.1.4.4. *Structural characterization of psTri interaction with DC-SIGN; comparison of thermodynamic and hydrodynamic properties of psDi and psTri interaction with DC-SIGN*

From the two mannose-based glycomimics psDi and psTri, the latter one was previously chosen as the best lead due to its higher efficiency. However, recalling the data published by Luczkowiak et al. (article **n°3**), a surprising discrepancy can be tracked when comparing the inhibitory potencies of psDi and psTri in monovalent and multivalent presentations: the SPR data showed an order of magnitude difference in monovalent psDi and psTri $IC_{50}$ values in favor for psTri, but this difference was lost comparing corresponding multivalent compounds. Moreover, monovalent and tetravalent psTri had the same efficiency of inhibition of DC-SIGN binding to ManBSA surface. Although Ebola infection inhibition assay yielded an order of magnitude difference of the potencies of tetravalent presentations of psDi and psTri, however, it disappears for multivalent dendrimers.

The following paper manuscript aims to unravel the underlying reasons of this phenomenon. Starting with the X-ray structure of psTri/DC-SIGN CRD, the further hydrodynamic and thermodynamic characterization of psTri interaction with DC-SIGN is presented and compared with psDi. The final outcome of these studies is an unexpected finding that a relatively small glycomimic psTri is capable to bridge DC-SIGN tetramers and thus lead to an artificial overestimation of its efficacy. Therefore, the simpler and more readily to synthesize psDi ligand was approved to be as good as psTri, which is confirmed by the fact that multivalent compounds with psDi and psTri tethers have the same efficiency in cellular HIV *trans* infection inhibition assay.

The main outcome of these studies: the compound psTri has a unique intrinsic feature of clustering DC-SIGN tetramers without any multivalent presentation.

Contributions:

The compounds were synthesized by the team of Pr. A. Bernardi. The X-ray structure of psTri/DC-SIGN was solved by our group (Dr. Michel Thépaut). The initial DC-SIGN bridging model was generated in our team by Dr. Michel Thépaut. The cell-based assays of HIV inhibition were performed by a PhD student Angela Berzi in the group of Pr. Mario Clerici.

My contribution to this study:

- ✧ Prepared DC-SIGN CRD and performed its co-crystallization with psTri.
- ✧ Participated in X-ray data collection and structure solving.
- ✧ I have conducted the ITC experiments and analyzed the data with a technical support from Dr. Aymeric Audfray.
- ✧ With the help of Pr. Christine Ebel and Aline Le Roy, I have performed AUC and DLS experiments and analyzed the data.
- ✧ Prepared the initial version of the manuscript of the following paper.

5

# 5

***Paper n$^o$5:*** **Lectin clustering by a glycomimetic without any multivalent presentation: a case study in DC-SIGN antagonist development**

# Lectin clustering by a glycomimetic without any multivalent presentation: a case study in DC-SIGN antagonist development

Ieva Sutkevičiūtė[1,2,3], Michel Thépaut[1,2,3], Sara Sattin[4], Angela Berzi[5], John McGeagh[6], Aline Le Roy[1,2,3], Macarena Sanchez Navaro[7], Javier Rojo[7], Mario Clerici[8], Anna Bernardi[4], Christine Ebel[1,2,3], Franck Fieschi[1,2,3]*

*[1]Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, Grenoble, F-38027, France; [2]CNRS, UMR 5075, Grenoble, F-38000, France; [3]CEA, DSV, Grenoble, F-38000, France; [4]Universita' di Milano, Dipartimento di Chimica, via Golgi 19, 20133 Milano - Italy; [5]Chair of Immunology, Department of Clinical Sciences L. Sacco, University of Milan, Segrate; [6]Anterio Consult&Research GmbH, Augustaanlage 23, D-68165 Mannheim-Germany; [7]Glycosystems Laboratory, Instituto de Investigaciones Químicas (IIQ), CSIC - Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla-Spain; [8]Chair of Immunology, Department of Biomedical Sciences and Technologies, University of Milan, Segrate.*

---

## Abstract

DC-SIGN is a dendritic cell-specific C-type lectin receptor that recognizes highly glycosylated ligands expressed on the surfaces of various pathogens. It is known to be implicated in early stages of many viral infections, including HIV, which makes it an interesting therapeutical target. Glycomimetic compounds are good drug candidates for DC-SIGN inhibition due to their high solubility, resistance to glycosidases and non-toxicity. In this work we studied the structural properties of DC-SIGN extracellular domain (ECD), a tetrameric protein, interaction with two glycomimetic inhibitors, recently described pseudomannobioside (psDi) and pseudomannotrioside (psTri). Though the inhibitory potency of psTri is significantly higher than psDi, crystal structures of the complexes with DC-SIGN carbohydrate recognition domain show the same binding mode for both compounds. Moreover, when coupled to multivalent scaffolds, the inhibitory potencies of both compounds become uniform. Combining isothermal titration microcalorimetry, analytical ultracentrifugation and dynamic light scattering techniques to study DC-SIGN ECD interaction with these glycomimetics, revealed that psTri is capable to bridge DC-SIGN tetramers. The results of these studies allowed to select a compound for further improvements.

*Keywords:* DC-SIGN, HIV, glycomimetics, clustering

---

## Introduction

Immature dendritic cells (DCs) are the major professional antigen presenting cells of the innate immunity. They routinely survey the peripheral tissues, capture and process the invading pathogens, and finally, present the antigens to the T cells, which leads to the boost of pathogen specific adaptive immune responses Banchereau & Steinman [4]. Pattern recognition receptors (PRRs) of DCs are instrumental in capturing pathogens through the recognition of so-called pathogen associated molecular patterns (PAMPs). On their surface, DCs express a wide range of different PRRs that include toll-like receptors (TLRs) and C-type lectin receptors (CLRs), and this enables DCs to recognize the vast majority of invading pathogens Figdor *et al.* [14], Geijtenbeek *et al.* [17]. One of these PRRs is DC-SIGN (Dendritic Cell-Specific ICAM-3 Grabbing Non-integrin), a CLR that is abundantly expressed on immature DCs and that has been extensively studied during the past decade because of its intriguing roles in immunity Geijtenbeek *et al.* [15], Svajger *et al.* [28]. Apart from functioning as an adhesion molecule and a PRR, DC-SIGN has been recognized as a receptor that is usurped by various dangerous pathogens, including viruses such as HIV and Ebola, bacteria, fungi and parasites, to evade or modulate host immune responses and thus enhance their infectivity van Kooyk & Geijtenbeek [34], Geijtenbeek *et al.* [16]. These findings have highlighted DC-SIGN as an interesting therapeutical target and inspired many research groups to attempt to design its antagonists that would prevent the infections.

DC-SIGN is a type II membrane protein with a cytosolic domain followed by a transmembrane region and an extracellular domain (ECD). The latter is organized into an elongated neck region and a globular C-terminal carbohydrate recognition domain (CRD) of the C-type lectin family, which binds fucose and mannose containing carbohydrates in a calcium-dependent manner Feinberg *et al.* [13], Guo *et al.* [19]. The neck region is responsible for lectin oligomerization into tetramers and serves as a stalk that raises CRDs well above the cell membrane (320 Å) and presents them in a tetravalent manner with the distances between vicinal binding sites being ca. 40 Å Tabarani *et al.* [30]. These tetramers are further clustered to microdomains at the cell membrane Cambi *et al.* [7], de Bakker *et al.* [10].

Although monosaccharide binding affinity to DC-SIGN is very low, the described organization of DC-SIGN together with clustered presentation of the glycans on the pathogens, greatly contributes to the binding affinity enhancement through the avidity effects. Hence, the strategies of DC-SIGN inhibition involve not only the design of various monovalent fucose and mannose-based glycomimetic or small non-carbohydrate ligands with the goal to improve both affinity and selectivity, but also include extensive search for the proper scaffolds with multimeric ligand presentation that would ensure efficient competition with highly multivalent PAMPs (reviewed in Reina *et al.* [22] and Anderluh *et al.* [2]).

Our groups focus on the development of fucose- and mannose-based glycomimetic inhibitors of DC-SIGN. We recently published series of Lewis$^{X}$ trisaccharide mimics, which had slightly better affinity as compared to Lewis$^{X}$ and an improved selectivity to DC-SIGN than langerin Andreini *et al.* [3], a C-type lectin with a reported implication in the protection from HIV infection de Witte *et al.* [11]. We have also developed two mannose type glycomimetics, corresponding to a terminal branch of high mannose glycan, which are the target ligand for DC-SIGN on the gp120 HIV enveloped protein. Thus, we produced a Man$\alpha$1-2Man mimic pseudomannobioside (psDi) and a Man$\alpha$1-2Man$\alpha$1-6Man mimic pseudomannotrioside (psTri) and evaluated them as DC-SIGN inhibitors. The initial SPR competition assay indicated psTri to have an order of magnitude better affinity (in terms $IC_{50}$) than psDi Sattin *et al.* [24]. Hence, this compound was selected to invest synthetic efforts and generate a tetravalent dendron, which in turn was tested for the capacity to inhibit HIV *trans* infection of T lymphocytes mediated by B-THP-1/DC-SIGN cells Sattin *et al.* [24] as well as HIV transmission inhibition in cervical explants Berzi *et al.* [5]. Indeed, the psTri dendron displayed a very promising anti-HIV activity in these assays and was demonstrated to be non-toxic.

Moreover, the third generation (G3) Boltorn type dendrimers bearing an average of 30-32 copies of psDi (G3(pseudodi)$_{32}$) or psTri (G3(pseudotri)$_{32}$) were built and evaluated in SPR competition and DC-SIGN-mediated Ebola infection assays Luczkowiak *et al.* [21]. Both dendrimers were highly active in inhibiting Ebola infection, but surprisingly, no significant differences of their potencies were found (IC$_{50}$ ca. 20 nM). The SPR competition experiments gave the similar outcome: while the activities of the monovalent compounds have almost 1 order of magnitude difference in favor for psTri, nearly the same affinities were observed for tetravalent psDi and psTri forms (Fig. 1) as well as for their multivalent versions. Furthermore, the monovalent and tetravalent forms of psTri had basically the same activities.

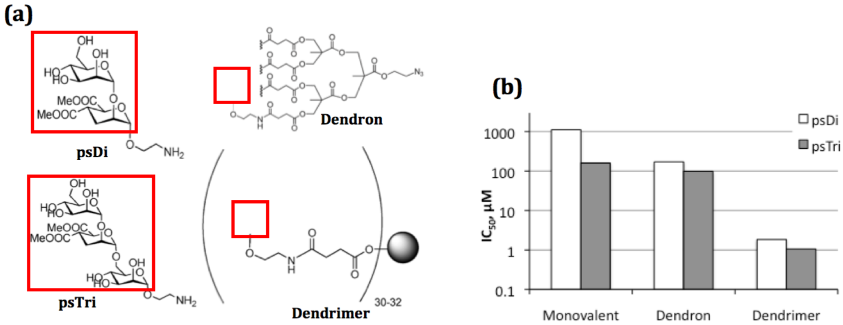


Figure 1: The comparison of psDi and psTri inhibitory efficiencies at monovalent and multivalent presentations. a) The structures of psDi and psTri, tetravalent dendron and multivalent Boltorn type dendrimer scaffolds; the red frames highlight the portions of the molecules conjugated to the scaffolds; the structure of G3 Boltorn type dendrimer, represented as a grey sphere, is shown in SI. b) The SPR competition assay results for the corresponding compounds.

These findings have raised the question why the relatively high efficiency of psTri is lost when the compound is tethered on the multivalent scaffolds and is psTri really the best lead compound. In order to make a rational choice of one of those two compounds, to which synthetic efforts should be invested for further optimization, these questions had to be addressed.

Recently, we have structurally characterized the binding of psDi to DC-SIGN CRD Thepaut *et al.* [31]. X-ray crystallography of co-crystallized DC-SIGN CRD/psDi complex revealed an unexpected binding mode of the molecule, which was confirmed by solution studies using transfer NOE (tr-NOESY) and saturation transfer difference (STD) NMR experiments. Although psDi and its natural counterpart Manα1-2Man have similar affinities to DC-SIGN as observed in surface plasmon resonance (SPR) competition test (IC$_{50}$ values 1.0 mM and 0.9 mM, respectively), our results indicated that psDi, contrary to Manα1-2Man, has only a single binding mode within DC-SIGN CRD, a prerequisite for a good lead compound for further chemical modifications to improve affinity and selectivity. Furthermore, we also found that psDi has an enhanced selectivity towards DC-SIGN compared to langerin. This is an important feature for the development of microbicides that shouldn't interfere with the protective function of langerin, especially because of its expression on Langerhans cells (LCs), which constitute the very first barrier for invading HIV virions in genital and rectal mucosa.

Herein we describe our efforts to unravel the underlying reasons for the above-described discrepancy observed for the glycomimetic compound psTri. We have co-crystallized the latter glycomimics with DC-SIGN CRD and solved the X-ray structure. Furthermore, we investigated the thermodynamic and hydrodynamic properties of DC-SIGN extracellular domain (ECD) interaction with psDi and psTri by isothermal titration microcalorimetry (ITC) and analytical ultracentrifugation (AUC). Our results indicated that monovalent psTri in fact functions as a bivalent molecule capable to bridge DC-SIGN tetramers, a very peculiar property for a rather small molecule with relatively not so high affinity. Such peculiarity of psTri resulted in an artifactual overestimation of its potency in competition assay where soluble DC-SIGN tetramers are used, and thus misled us previously to select it as the best lead compound. This work emphasizes the importance of thorough and critical investigation of the leads from the primary screenings in order to avoid the "false positives" to be selected for further, rather void, attempts to improve them.

## Results and discussion

### *The X-ray structure of DC-SIGN CRD/psTri complex and comparison with psDi binding within DC-SIGN CRD*

In order to compare psDi and psTri binding modes within DC-SIGN and thus to shed light on the underlying reasons for this discrepancy, we co-crystallized psTri with monomeric DC-SIGN CRD and solved the X-ray crystal structure.

Similarly as the recently published structure, the crystals of psTri in complex with DC-SIGN CRD had also a single copy of CRD in an asymmetric unit with $P4_32_12$ space group. The structure was solved by molecular replacement, at 1.35 Å resolution (Table 1).

When the structure was solved using the model without sugar, an electron density was observed on $Ca^{2+}$ ion in the canonical carbohydrate binding site, indicating the presence of pseusomannotrioside bound to DC-SIGN through a conventional $Ca^{2+}$-coordination by 3-OH and 4-OH groups of non-reducing mannose-moiety (Fig. 2a). The electron-density of the dimethyl ester substituents on the cyclohexane moiety as well as the reducing mannose moiety were poorly or not visible suggesting that these parts of the molecule remained flexible within the complex. In contrast to recently solved structure of pseudo-mannobioside bound to DC-SIGN Thepaut *et al.* [31], where the electron density for ethylamine linker was not visible, in the current case of psTri the density for the corresponding ethylazide linker was clearly visible, suggesting that this appendage stayed stable due to contacts in the crystal packing.

Table 1: DC-SIGN CRD/psTri complex data collection and structure refinement statistics.

| **Data collection statistics** | |
|---|---|
| Wavelength (Å) | 0.9797 |
| Space group | P $4_3$ $2_1$ 2 |
| Unit cell parameters (Å) | a = b = 71.33; c = 52.67 |
| Resolution (Å) | 50.44-1.35 (1.43-1.35)[a] |
| Measured reflections | 222778 (35175) |
| Unique reflections | 30348 (4813) |
| Completeness (%) | 99.9 (99.9) |
| $I/\sigma$ $(I)$ | 20.13 (4.82) |
| $R_{\mathrm{merge}}$ [b] (%) | 6.5 (43.3) |
| **Structure refinement statistics** | |
| Resolution (Å) | 50.44-1.35 (1.39-1.35) |
| Refinement factors | |
| Used reflections/free (%) | 28812 / 5.1 |
| $R_{\mathrm{cryst}}$ [c] | 0.143 |
| $R_{\mathrm{free}}$ [c] | 0.168 |
| rmsd from ideality | |
| Bond lengths (Å) | 0.030 |
| Bond angles (deg) | 2.631 |
| Average $B$-factor (Å$^2$) | 13.96 |

[a] Values in parantheses are for the highest resolution shell.

[b] $R_{merge} = \sum_h \sum_m |I_m(h) - \langle I(h) \rangle| \sum_h \sum_m I_m(h)$. [c] $R_{cryst} = \sum ||F_0| - |F_c|| \sum |F_0|$, and $R_{free} = R_{cryst}$ calculated with 5% of $F_0$ sequestered before refinement.

It appeared that the binding mode of psTri within DC-SIGN CRD (Fig. 2a) is the same as that of psDi, with exactly the same orientation of the psDi-corresponding portion of the molecule (Fig. 2b).

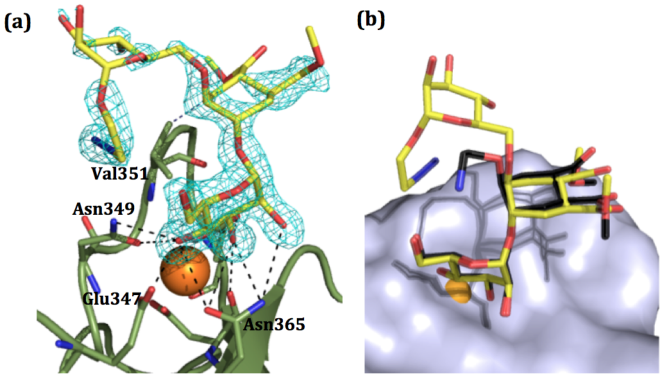


Figure 2: The binding mode of psTri within DC-SIGN CRD.
(a) The bound ligand superimposed with the Fo-Fc electron density map; the protein carbon backbone is represented in olive cartoon with amino acids contributing to binding highlighted by stick representation; psTri is shown in yellow sticks. (b) The superimposition of psDi (black sticks) and psTri (yellow sticks) crystal structures bound to DC-SIGN CRD (grey surface representation). Oxygen and nitrogen atoms are in red and blue, $Ca^{2+}$ ion is an orange sphere.

To our surprise, no additional contacts (other than in psDi structure) with the protein were observed. While non-reducing mannose moiety made coordination bonds with $Ca^{2+}$ ion by its 3-OH and 4-OH groups and the C6 methylene of the cyclohexane ring retained the van der Waals contact with Val351 side chain of the protein, the reducing mannose moiety was oriented away from the protein and apparently didn't make any contacts with the side chains. Moreover, the electron densities of this second mannose unit as well as of ethylazide linker were rather poorly defined suggesting a higher flexibility of these parts (Fig. 2a). Thus the revealed structure of psTri within DC-SIGN was more puzzling than explanatory of the described phenomenon, and prompted us to investigate the interaction further on to understand how with the same binding mode a much higher affinity might be achieved.

*Thermodynamic characterization of psDi and psTri interaction with DC-SIGN*

To inspect whether positive entropy contributions were responsible for the higher affinity of monovalent psTri with respect to psDi, we have analyzed the interaction of both psDi and psTri with tetrameric DC-SIGN ECD by isothermal titration microcalorimetry. At the first set, the same experimental conditions with identical concentrations of interaction components were used, and the pseudosaccharides (12.7 mM) were titrated into lectin solution (71 µM with respect to binding sites).

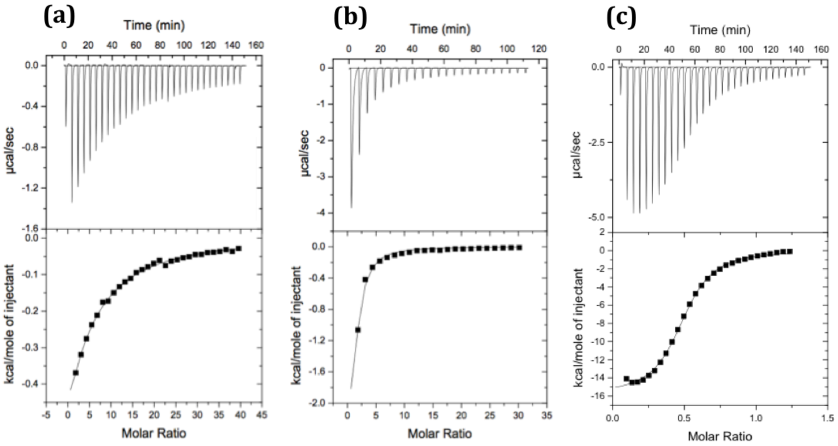


Figure 3: The ITC results of psDi and psTri titrations to DC-SIGN ECD.
a) and b) show titrations of psDi and psTri, respectively, at 12.7 mM to DC-SIGN ECD (71 µM). c), psTri (1.18 mM) titratration to 214 µM DC-SIGN ECD. Upper panels show the titration thermograms and lower panels the integration of data with fitted curves for "one binding site" model.

The ITC data (Fig. 3a and 3b) indicated the low affinity of the ligands as the titration curves did not adopt the full sigmoidal shape. Fitting one binding site model on the data with an assumed stoichiometry value n fixed to 1 yielded the $K_\mathrm{D}$ values of $990.10 \pm 19.7$ µM and $75.76 \pm 7.29$ µM for psDi and psTri, respectively, but the low affinity prevented the reliable interaction enthalpies and entropies to be obtained and compared Turnbull & Daranas [33]. Nevertheless, an intriguing outcome could be observed: the first injections of psTri to DC-SIGN solution resulted in more heat released while using the same concentrations of both ligands and the same receptor concentration. This suggested the higher affinity of psTri than psDi to DC-SIGN, which was contradictory to the X-ray data since psTri didn't seem to make any other additional contacts compared to psDi structure.

Because psTri had apparently higher affinity to DC-SIGN, we repeated the titration with adjusted concentrations (1.18 mM psTri titrated to 214 µM of DC-SIGN ECD) in order to perform titration in a more relevant concentration range (Fig. 3c). The fitting of the same model this time gave $K_\mathrm{D}$ of $5.26 \pm 0.29\,\mu\mathrm{M}$, $\Delta H = -15.80 \pm 0.15\,\mathrm{kcal\,mol^{-1}}$, $T\Delta S = -8.57\,\mathrm{kcal\,mol^{-1}}$, and $\Delta G$ of $-7.20\,\mathrm{kcal\,mol^{-1}}$. Thus a difference of 2 orders of magnitude between the $K_\mathrm{D}$ values determined for psDi and psTri molecules ( 990 µM and 5 µM, respectively), which is much higher compared to what could be initially expected from the SPR data, highlights even a stronger difference in the activities of the two molecules.

Furthermore, the fitting yielded the stoichiometry value of 0.5. Because DC-SIGN ECD was extra-purified for this experiment (see SI), the contribution of non-active binding sites to the $n<1$ value was excluded. Interaction with stoichiometry of 0.5 in this case suggests that two DC-SIGN CRDs share one psTri molecule, or in other words, two DC-SIGN ECD tetramers bind 4 psTri molecules in average. Since psTri molecule is relatively small, its binding to 2 CRDs within the same tetramer is not likely (the approximate distance between vicinal binding sites within the tetramer is 40 Å Tabarani *et al.* [30]). Hence, such a stoichiometry might mean that the same psTri molecule makes a bridge between two DC-SIGN tetramers by simultaneously binding to a CRD in each one, and would explain why under the same experimental conditions psTri titration released markedly more heat than psDi.

### *Analytical Ultracentrifugation analysis of psDi and psTri complexes with DC-SIGN ECD*

To check the hypothesis of psTri capability to bridge DC-SIGN tetramers, as suggested by ITC data, we performed the sedimentation velocity experiments on the samples retrieved from the first set of ITC measurements. The deduced final concentrations of interacting components recovered from the cell were the following: 62 µM DC-SIGN ECD (with respect to the binding sites), psDi and psTri at 1.63 mM. Thus the compounds were at the excess with respect to the lectin binding sites at a molar ratio of 26. The sedimentation velocity was also recorded for compound-free DC-SIGN ECD (62 µM) sample.

Figure 4 compares the sedimentation profiles registered at 42 000 rpm at the same time intervals (1 h) for each of the three samples. Clearly, DC-SIGN alone (Fig. 4a) and incubated with psDi (Fig. 4b) sediment similarly, while DC-SIGN incubated with psTri (Fig. 4c) moves faster, suggesting that the association of DC-SIGN tetramers is induced by psTri but not psDi.

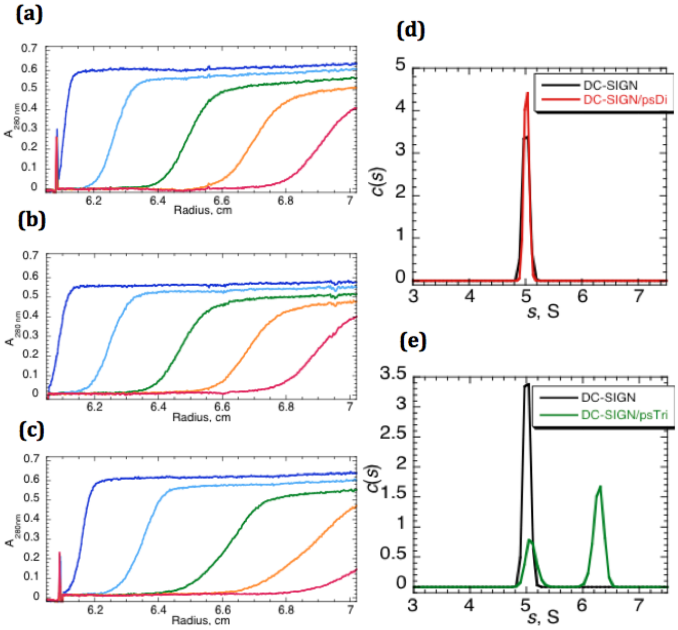


Figure 4: Sedimentation velocity experiments of 62 µM DC-SIGN ECD alone and in presence of psDi or psTri at 1.63 mM. Experimental data showing concentration profiles of DC-SIGN alone (a), in the presence of psDi (b) and psTri (c) recorded with time intervals between successive profiles of 40 min for the two first ones, then 55 min; the first profiles at panels a), b) and c) were obtained 4, 10 and 14 min, respectively, after the beginning of the centrifugation. Sedimentation velocity was registered for 4 hours at 42 000 rpm, in 0.15 cm cells, at 20°C. d) and e) shows the superimposition of the $c(s)$ curves for DC-SIGN alone and with psDi (d) and psTri (e).

The data were treated in terms of size distribution analysis. The model considers that the solution comprises a large number of species having the same frictional ratio, i.e. the same shape, characterized by their sedimentation coefficients, $s$, between a minimal and a maximal value. The result of the analysis is a plot of the “concentration”, in terms of the signal at 280 nm, as a function of sedimentation coefficient $s$. The area under the peak gives the absorbance. The frictional ratio can be fitted, and a regularization procedure can be applied, which allows smooth distributions to be obtained. Figure 4d shows a nearly perfect superposition of the $c(s)$ distributions obtained with DC-SIGN alone and in the presence of psDi. On the other hand, psTri induces the formation of larger species (fig. 4e).

DC-SIGN alone and with psDi sediments at $s_{20\mathrm{w}} = 5.2$ S, close to $s_{20\mathrm{w}} = 5.4$ S previously published for the tetramer at infinite dilution Tabarani *et al.* [29, 30]. The additional peak observed for DC-SIGN in the presence of psTri is at $s_{20\mathrm{w}} = 6.5$ S. This value may correspond to a more compact tetramer ($f/f_{\mathrm{min}} = 1.6$ instead of 1.9 for DC-SIGN alone) or to a more elongated complex of two tetramers ($f/f_{\mathrm{min}} = 2.5$). However, this experiment was done only at one ratio, and we cannot exclude that the 6.5 S peak corresponds to an intermediate value between the $s$-value of the tetramer and of the larger complex. The shape of the sedimentation boundary, thus the $c(s)$, indeed depends on the kinetics and thermodynamics of the interaction.

A second set of sedimentation velocity experiments was done with DC-SIGN ECD at 102 µM with respect to monomer incubated with psTri at molar psTri/DC-SIGN ratios of 0, 0.25, 0.5, 1, 5, 11 and 27.

Figure 5a shows the $c(s)$ plots.

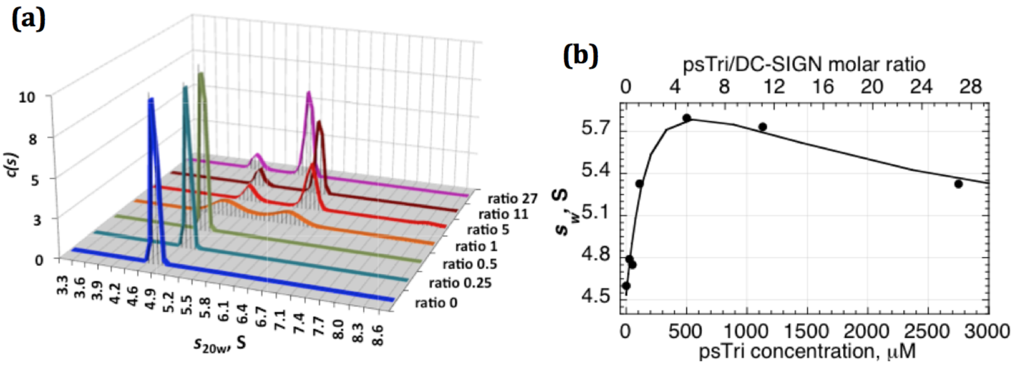


Figure 5: Sedimentation velocity experiments of 102 μM DC-SIGN ECD alone and in the presence of psTri at different psTri/DC-SIGN molar ratios.
a) The overlaid $c(s)$ curves for DC-SIGN ECD alone and with psTri at different ratios; sedimentation velocity was registered at 42 000 rpm, in 1.5 mm cells, at 20°C. b) The evolution of $s_w$-values as a function of psTri concentration; the line is shown to guide the eye.

DC-SIGN alone sediments at $s_{20\mathrm{w}} = 4.8$ S, which is a lower value than measured with DC-SIGN at 62 μM. This difference may be related to excluded volume effects, which are expected to be noticeable at this concentration ($4\,\mathrm{mg\,mL^{-1}}$) considering the elongated shape of DC-SIGN. The $s$-value of this peak is slightly increasing up to the ratio of 5, as can be seen in figure 5a. Starting from ratio 1, an additional peak emerges at approximately $s_{20\mathrm{w}} = 6.3$ S. When adding psTri, the proportion of this peak increases up to 80% of the total signal from ratios 11. The $s$-values for both peaks start to slightly decrease from the experimental ratio 5: for the largest species the observed maximum of $s_{20\mathrm{w}}$ is 6.3 S at ratio 5, and decreases to 5.7 S at ratio 27; for the smallest species at the same ratios the corresponding $s_{20\mathrm{w}}$ values are 5.2 and 4.7 S. Most likely this variation of the positions of the two peaks is related not to experimental errors, but reflects the complexity of a reaction boundary: for fast reactions indeed, association-dissociation processes affect the sedimentation boundary Schuck [25], Zhao *et al.* [35]. The integration of each $c(s)$ curve under the two peaks gives, on one hand, the total absorbance reflecting DC-SIGN concentration, which was similar at all ratios, and, on the other hand, the weight-average sedimentation coefficient value ($s_w$) that does not depend on the kinetics of the interaction. The $s_w$-isotherm was attempted to analyze in SEDPHAT but fitting hetero-association models failed most likely due to more complicated system and the non-ideality conditions. Nevertheless, the data analysis in terms of weight-average $s_w$-value (Fig. 5b) clearly shows that the addition of psTri up to 500 μM concentration to 102 μM DC-SIGN solution participates in the formation of higher macromolecular complexes as indicated by the increase of $s_w$-value, while further excess of the compound induces the disassembly of DC-SIGN complexes. Although this type of sedimentation velocity data analysis can be used to extract the interaction affinities Schuck [26], Zhao *et al.* [35], our system apparently is more complicated (due to both, the tetravalent protein and presumably bivalent small compound, and the presence of excluded volume effects), therefore, no model could be fitted to the data.

### *Characterization of DC-SIGN/psTri interaction by static and dynamic light scattering*

psTri-induced DC-SIGN association was also studied by Static (SLS) and Dynamic Light Scattering (DLS) simultaneously. DLS allows to characterize the polydispersity of the sample and determine the hydrodynamic radius ($R_{\mathrm{H}}$) of the macromolecules, while SLS gives independent values of MW in an absolute way. The examined samples of DC-SIGN alone and in the presence of psTri were the same as for AUC study. Correlation curves and $R_{\mathrm{H}}$ distributions from DLS show the presence of one contribution for most of the samples (an example is shown in the SI). Larger species ($R_{\mathrm{H}} \approx 100\,\mathrm{nm}$) were marginally detected in negligible amounts of the peak intensity and are considered as irrelevant dust.

The $R_{\mathrm{H}}$ and MW values were plotted against the psTri/DC-SIGN molar ratios (Fig. 6). SLS and DLS do not allow distinguishing monomers and dimers. The measured MW and $R_{\mathrm{H}}$ are mean values dominated by the larger species, since scattered intensity is proportional to $\sum c_{\mathrm{i}} MW_{\mathrm{i}}{}^{2}$, $c_{\mathrm{i}}$ and $MW_{\mathrm{i}}$ being the weight concentration and molar mass of species $i$. When increasing psTri/DC-SIGN ratio, MW increases from 148 kDa for DC-SIGN alone and reaches a constant value of 350 kDa from ratio 5. $R_{\mathrm{H}}$ increases from 7.5 nm without psTri to a maximum value of 10.4 nm at ratio 11, followed by a slight and perhaps irrelevant decrease to 10.2 nm at ratio 27.

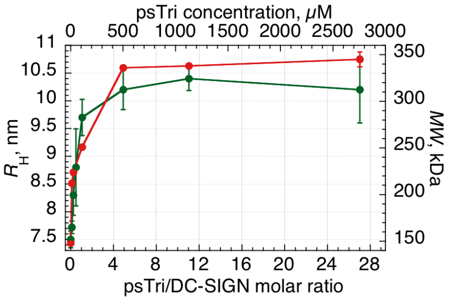


Figure 6: Results of static and dynamic light scattering of 102 μM DC-SIGN ECD alone and in presence of psTri at different psTri/DC-SIGN molar ratios.
The dependencies of registered $R_{\mathrm{H}}$ and MW on psTri/DC-SIGN ratio are plotted in green and red, respectively. The error bars represent standard deviation values.

The evolution of MW and $R_{\mathrm{H}}$ reflects the general behavior observed in AUC (Fig. 5b). We note that in the absence of psTri, MW from SLS is close to the theoretical value for a tetramer (154.827 kDa); $R_{\mathrm{H}}$ is close to the value of 7.7 nm obtained from the combination of $s_{20\mathrm{w}} = 4.8\,\mathrm{S}$ and the theoretical MW of a tetramer. The maximum MW value from SLS (350 kDa) is close to the expected value for a dimer of tetramers (309.654 kDa), while the combination of $s_{20\mathrm{w}} = 6.3\,\mathrm{S}$ with MW=309.654 kDa gives $R_{\mathrm{H}} = 11.7\,\mathrm{nm}$, close to the experimental values from DLS above psTri/DC-SIGN ratio of 5.

The DLS and SLS data at high psTri ratio reflects imperfectly the changes observed in AUC. Apart from experimental uncertainty, it may be related to the different incubation times (one hour in SLS/DLS, one day in AUC). Altogether, AUC, DLS and SLS data support that psTri promotes the association of two DC-SIGN tetramers when adding psTri up to ratio between 1 and 5, while higher excess of psTri saturates all the binding sites and consequently the situation becomes less favorable for bridging to occur.

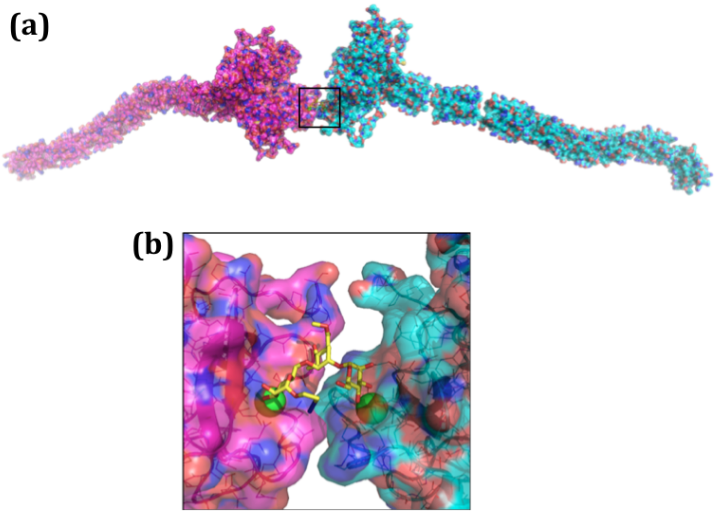


Figure 7: A possible structure of the assembly of DC-SIGN tetramers bridged by psTri molecule. a) Positioning of two DC-SIGN tetramers resulting from superposition of X-ray structures of DC-SIGN CRD in complex with mannose and psTri to each of the two envelopes of tetramers. b) The close-up of the area marked by a black frame in panel (a). the color scheme used: carbons for one of the two DC-SIGN tetrames are pink, for the other one cyan, for psTri carbons are yellow; oxygens and nitrogens are in read and blue colors; green spheres are $Ca^{2+}$ ions.

In order to get an insight to a structure of psTri-mediated DC-SIGN assembly, the molecular modeling of the supramolecular complex was performed, and the resulting model apparently showed no steric clashes (Fig. 7).

Finally, the constructed model was subjected to HydroPro calculations Torre *et al.* [32], and yielded the theoretical $s$ value for the model structure of 6.77 S, which in turn gives $R_{\mathrm{H}}$ value of 10.3 nm. This $R_{\mathrm{H}}$ value corresponds perfectly with the experimentally measured $R_{\mathrm{H}}$ maximum (Fig. 6). This suggests that the supramolecular assembly indeed adopts such an elongated conformation. However, this model is only a rough approximation of the real situation, since it assumes only one psTri molecule bound to two tetramers (while ITC data suggested an average of 4 psTri molecules bridging two tetramers). It must be kept in mind that such a system is presumably very dynamic due to relatively low affinity of single mannose residue, and thus possibly adopts different conformations.

### *Comparison of the ability of psDi and psTri-bearing multivalent systems to inhibit HIV trans infection*

Previously, due to the better activity of psTri as a monovalent compound observed in SPR competition assay, only psTri was used to generate a tetravalent dendron and tested in cellular studies of HIV *trans* infection inhibition Sattin *et al.* [24]. The promising results of these studies gave a feeling that psTri was indeed the lead for further development. From the characterization of psDi and psTri binding mode to

DC-SIGN (Thepaut *et al.* [31] and described in this work), it appeared that this initial selection might not be correct. We decided to repeat these HIV trans infection inhibition studies this time including tetravalent dendron of psDi for the comparison with previously generated psTri dendron. In addition to tetravalent dendrons, 32-valent dendrimers bearing psDi or psTri pseudosaccharides (Fig. 1a) were also included to the studies. The results showed that both dendrons were able to reduce HIV transmission to CD4$^+$ T cells by a very similar extent (Fig. 8): 5 µM concentrations displayed around 50% infection inhibition, while incubation with 50 µM dendrons reduced infection almost to 90%.

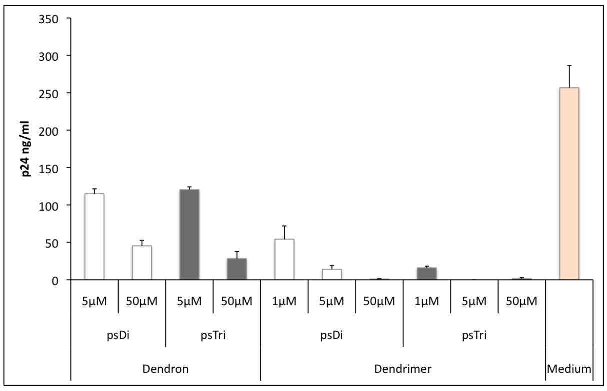


Figure 8: Comparison of HIV *trans* infection inhibition by psDi and psTri bearing multivalent systems. Data obtained in triplicate, from 3 different healthy donors. Values are mean ± SD. After 30 min of preincubation with compounds or culture medium, B-THP-1/ DC-SIGN cells were incubated for 3 h with HIV-1 BaL in the presence of the indicated concentrations of the compounds, then washed and co-cultured with CD4$^+$ T lymphocytes for 3 days. Viral infection was assessed by measuring the concentration of p24 in the co-culture supernatants.

The dendrimers with both psDi and psTri had significantly higher efficiencies than dendrons and reached nearly complete infection inhibition at 50 µM concentration.

Despite slight differences of inhibitory efficiencies of psDi and psTri bearing dendrimers at lower concentrations, these data correlate well with above described results and confirm once again that both pseudosaccharide ligands have the same activity once they are tethered to multivalent scaffolds. It definitely confirms that psDi is as effective as psTri and gives a strong reasoning to select psDi for further compound optimization.

## Conclusions

The interactions of monosaccharides with the C-type lectins are of low affinity with $K_\mathrm{D}$ values in mM range. Nature has overcome such low affinity problems by clustering the interacting partners on both sides: while the monosaccharide ligands are presented in high-density clusters such as polysaccharides that usually exist in multiple copies on proteins or lipids, the sugar binding sites are also presented in multiple copies per lectin. In order to efficiently compete such multivalent interactions, the strategies of multivalent presentation of active ligands are developed. Our groups pursue DC-SIGN antagonists for the protection from HIV and other pathogen infections by employing two strategies, i.e. the glycomimetic monovalent ligand development and the multivalent presentation of the selected monovalent leads. We designed a small glycomimetic molecule, pseudomannotrioside, as a starting monovalent compound. To our surprise, we discovered that this ligand possesses the bridging ability. Although the multivalent bridging lectin-carbohydrate interactions are common and have a biological relevance Dam & Brewer [9, 8], in biological systems the more common situation is glycoconjugate bridging by lectins allowing the formation of the carbohydrate-lectin lattices Brewer [6]. On the other side, the lectin bridging property of the small glycomimetic molecules has been engineered Gestwicki *et al.* [18], Sisu *et al.* [27], Lameignere *et al.* [20]. In this aspect, our case is unique and unexpected, especially considering the size of the molecule, the low affinity of mannose residues to DC-SIGN and the arrangement of binding sites of DC-SIGN that does not seem to allow the formation of the bridged lattices. However, such a property of psTri is unfavorable in our course of the development of DC-SIGN inhibitors where we aim to design multivalent platforms bearing multiple copies of glycomimetic leads: this ability was the underlying reason for the overestimation of this compound in the competition assay and once it is tethered to multivalent scaffold the second mannose moiety serves nothing but a part of a linker.

The competition assays are widely used in high-throughput screening of drug candidates, and this study emphasizes the high probability of the selection of "false positives" arising from an artifactual overestimation of the compound affinity in the competition assay format.

### ***Acknowledgments***

This work was supported by the Marie Curie ITN FP7 project CARMUSYS (PITN-GA-2008-213592). We are grateful to Pr. Anne Imberty and Dr. Aymeric Audfray for an access to ITC instrument and for their kind support and advices. We thank Aline Le Roy, from the IBS platform of the Partnership for Structural Biology and the Institut de Biologie Structurale in Grenoble (PSB/IBS), for the assistance and access to the instrument of AUC and DLS from PAOL.

## Experimental Procedures

### *Synthesis of compounds*

Syntheses of psDi Reina *et al.* [23], psTri Sattin *et al.* [24] and psTri–bearing dendron Sattin *et al.* [24] were described previously. The syntheses of multivalent 3$^{\text{rd}}$ generation Boltorn type dendrimers functionalized with psDi and psTri are also described previously Luczkowiak *et al.* [21].

### *DC-SIGN production and purification*

The production and purification of DC-SIGN CRD containing StrepTag II (S-CRD) and ECD constructs was conducted as described previously Thepaut *et al.* [31] and Tabarani *et al.* [30], respectively.

### *Co-crystallization of DC-SIGN/psTri and data collection*

Crystallization was performed at 20°C in EasyXtal plates (Qiagen) by hanging-drop vapor-diffusion method. The drops were prepared combining 1 µL of a purified DC-SIGN S-CRD (4.4 mg mL$^{-1}$ in 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$ buffer) and psTri (300 mM) mixture (9:1, v/v) with 1 µL of reservoir solution and equilibrated against 1 mL of reservoir solution. The best crystals were obtained with the following reservoir solution composition: 35% PEG 3350, 100 mM cacodylate pH 6.5, and 200 mM NaCl.

The X-ray diffraction data were collected from a single crystal cryprotected in Paraton-N (Hampton Research) at the European Synchrotron Radiation Facility (Grenoble, France) at FIP BM30A beamline. One dataset was collected at 0.9797 Å wavelength, with a crystal-to-detector distance of 140.92 mm. The dataset, composed of 190 images, was collected with an oscillation range of 0.5° per image and an exposure time of 30 s.

### *Structure solution and refinement*

Phasing was performed by molecular replacement with a model built from a structure of DC-SIGN CRD (pdb code 2XR5) depleted of calcium ions, carbohydrate molecules and water molecules. The best solution resulting from molecular replacement was used as the starting model for structure refinement. The structure refinement was performed by cycling between manual building using the program COOT Emsley & Cowtan [12] and energy minimization with the program REFMAC 5 from the CCP4 package 4 [1]. Statistics of structure refinement are summarized in Table 1.

### *Isothermal titration microcalorimetry*

ITC experiments were performed using Microcal VP-ITC microcalorimeter (Microcal, Northampton, MA) with a cell volume of 1.409 mL. All titration were accomplished at 25°C in Tris-HCl (25 mM, pH 8) buffer containing 150 mM NaCl and 4 mM CaCl2. The pseudosaccharides and DC-SIGN ECD were prepared in the same buffer, and the stepwise injections of 10 µL aliquots of pseudosaccharides to DC-SIGN ECD solution were done using 5 min intervals between injections. In the first set of experiments the lectin concentration was 71 µM (in terms of binding sites) and pseudosaccharides were prepared at 12.7 mM, and in the second set of measurements 1.18 mM psTri was titrated to 214 µM DC-SIGN ECD. The blank titrations of the compounds to the buffer were done in order to subtract the heat of dilution from the integrated data. Subsequently, a built-in one site-binding model was fit to the data (Microcal Origin 7) and yielded association constants ($K_\mathrm{A}$) and binding enthalpies ($\Delta H$), which allowed to calculate changes in free energy ($\Delta G$) and entropy ($\Delta S$) using equation $\Delta G = \Delta H - T\Delta S = -RT\ln K_A$, where $T$ is the absolute temperature (K), and $R = 8.314\,\mathrm{J\,mol^{-1}\,K^{-1}}$.

### *Static and Dynamic Light Scattering*

Static (SLS) and dynamic light scattering (DLS) measurements were performed using DynaPro Nanostar instrument (Wyatt Technology Corp., Santa Barbara, USA) equipped with 658 nm “red” laser at a scattering angle of 90°. Triplicate measurements of 10 scattering readings per sample in a carefully cleaned quartz cuvette were recorded at 25°C. The samples contained DC-SIGN ECD alone (102 µM with respect to binding sites) or in presence of psTri at different psTri/DC-SIGN molar ratios (0, 0.11, 0.25, 0.5, 1, 5, 11 and 27) and were prepared in Tris-HCl (25 mM, pH 8) buffer containing 150 mM NaCl and 4 mM $CaCl_2$ and centrifuged prior to analysis. The data were analyzed with Dynamics software version 7.1 (Wyatt Technology Corp.).

### *Analytical Ultracentrifugation*

Sedimentation velocity experiments were performed in a Beckman XL-I analytical ultracentrifuge using an AN-50 TI rotor (Beckman instruments), at 20°C. The experiments were carried out at 42 000 rpm, using 55 µL samples, loaded into the two-channel 0.15 cm path length centerpieces equipped with sapphire windows (Nanolytics GmbH), and the absorbance at 280 nm was monitored every 5 or 11 min for the first and second set experiments, respectively, with a radial step size of 30 µm.

The samples were prepared Tris-HCl (25 mM, pH 8) buffer containing 150 mM NaCl and 4 mM $CaCl_2$. The first set of experiments was done with the samples retrieved from first set ITC measurements and contained DC-SIGN ECD alone at 62 µM (with respect to binding sites) or in the presence of 1.63 mM psDi or psTri. Freshly prepared samples of the same concentrations were used for the control and identical results were obtained.

The second set of experiments was carried out with the samples containing 102 µM DC-SIGN ECD alone or with psTri added at different psTri/DC-SIGN binding site molar ratios (0, 0.25, 0.5, 1, 5, 11 and 27). These samples were incubated for one day as they were used for SLS/DLS measurements prior to AUC experiments.

The molar mass (MW) and partial specific volume ($\bar{v}$) of DC-SIGN ECD tetramers were estimated from the amino acid composition using the SEDNTREP software and resulted in 154827 Da and $0.733\,\mathrm{cm^3\,g^{-1}}$, respectively. The values for psDi and psTri were considered equal to value of a hexose sugar ($0.607\,\mathrm{cm^3\,g^{-1}}$). Sedimentation velocity profiles were analyzed using the size distribution analysis Schuck [25] embedded in the SEDFIT software (freely available at `http://www.analyticalultracentrifugation.com`), and for each analysis by global modeling taking typically 20 regularly spaced experimental profiles obtained over a total 4 h. The $c(s)$ analysis was performed considering 200 particles, and fitting the value of the frictional ratio $f/f_{\text{min}}$. For the regularization procedure a confidence level of 0.68 was used. All $s$-values have been corrected for solvent density and viscosity and are therefore given as $s_{20\text{w}}$ values.

### *Construction of initial bridging model*

To model the bridging of two DC-SIGN tetramers by a psTri molecule, the SAXS envelope of the DC-SIGN ECD was used Tabarani *et al.* [30]. The solved X-ray structure of CRD in complex with psTri was superimposed to the CRD part of the envelope of the tetramer, while in another tetramer the CRD in complex with mannose (pdb code 2IT5) was overlaid, and the positioning of two tetramers was attempted in a way that reducing mannose residue of psTri in one of the tetramers would superimpose with the mannose in another tetramer.

### *Calculation of theoretical hydrodynamic radius from the model*

HydroPro program Torre *et al.* [32] was used to obtain the theoretical sedimentation coefficient value for the constructed model of bridged DC-SIGN tetramers. The experimental values of solvent viscosity, density and temperature were input: $\eta = 0.01024$ poise, $\rho = 1.006\,\text{g}\,\text{cm}^{-3}$, and the temperature was 293 K. The resulting sedimentation coefficient value of 6.848 S was used to calculate theoretical $R_{\text{H}}$ according to Svedberg's equation:

$$s = \frac{MW\cdot(1 - \rho \cdot \bar{v})}{N_A 6\pi\eta R_H}$$

where $s$ is the sedimentation coefficient, MW the molecular mass of the assembly of two tetramers (309.654 kDa), and $\bar{v}$ is the partial specific volume of the dimer of tetramers ($0.733\,\text{cm}^3\,\text{g}^{-1}$); the same $\eta$ and $\rho$ values were used as shown above.

### *Infection studies*

The studies of HIV *trans* infection inhibition by psDi and psTri dendrons and dendrimers without washing prior to virus inoculation were performed as described previously Sattin *et al.* [24].

## References

[1] 4, C. C. P. N. (1994). *Acta Crystallogr D Biol Crystallogr,* **50** (Pt 5), 760–763.

[2] Anderluh, M., Jug, G., Svajger, U., & Obermajer, N. (2012). *Curr Med Chem,* **19** (7), 992–1007.

[3] Andreini, M., Doknic, D., Sutkeviciute, I., Reina, J. J., Duan, J., Chabrol, E., Thepaut, M., Moroni, E., Doro, F., Belvisi, L., Weiser, J., Rojo, J., Fieschi, F., & Bernardi, A. (2011). *Org. Biomol. Chem.* **9** (16), 5778–86.

[4] Banchereau, J. & Steinman, R. M. (1998). *Nature,* **392** (6673), 245–52.

[5] Berzi, A., Reina, J. J., Ottria, R., Sutkeviciute, I., Antonazzo, P., Sanchez-Navarro, M., Chabrol, E., Biasin, M., Trabattoni, D., Cetin, I., Rojo, J., Fieschi, F., Bernardi, A., & Clerici, M. (2012). *AIDS (London, England),* **26** (2), 127–137.

[6] Brewer, C. F. (2002). *Biochim Biophys Acta,* **1572** (2-3), 255–62.

[7] Cambi, A., de Lange, F., van Maarseveen, N. M., Nijhuis, M., Joosten, B., van Dijk, E. M. H. P., de Bakker, B. I., Fransen, J. A. M., Bovee-Geurts, P. H. M., van Leeuwen, F. N., Hulst, N. F. V., & Figdor, C. G. (2004). *J Cell Biol,* **164** (1), 145–55.

[8] Dam, T. K. & Brewer, C. F. (2008). *Biochemistry,* **47** (33), 8470–6.

[9] Dam, T. K. & Brewer, F. C. (2010). *Glycobiology,* **20** (9), 1061–4.

[10] de Bakker, B. I., de Lange, F., Cambi, A., Korterik, J. P., van Dijk, E. M. H. P., Hulst, N. F. V., Figdor, C. G., & Garcia-Parajo, M. F. (2007). *Chemphyschem,* **8** (10), 1473–80.

[11] de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. A. W. P., de Gruijl, T., Piguet, V., van Kooyk, Y., & Geijtenbeek, T. B. H. (2007). *Nat Med,* **13** (3), 367–71.

[12] Emsley, P. & Cowtan, K. (2004). *Acta Crystallogr D Biol Crystallogr,* **60** (Pt 12 Pt 1), 2126–32.

[13] Feinberg, H., Mitchell, D. A., Drickamer, K., & Weis, W. I. (2001). *Science,* **294** (5549), 2163–6.

[14] Figdor, C. G., van Kooyk, Y., & Adema, G. J. (2002). *Nat Rev Immunol,* **2** (2), 77–84.

[15] Geijtenbeek, T. B., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C., Adema, G. J., van Kooyk, Y., & Figdor, C. G. (2000). *Cell,* **100** (5), 575–85.

[16] Geijtenbeek, T. B. H., den Dunnen, J., & Gringhuis, S. I. (2009). *Future Microbiol,* **4** (7), 879–90.

[17] Geijtenbeek, T. B. H., van Vliet, S. J., Engering, A., ’t Hart, B. A., & van Kooyk, Y. (2004). *Annu Rev Immunol,* **22**, 33–54.

[18] Gestwicki, J. E., Cairo, C. W., Strong, L. E., Oetjen, K. A., & Kiessling, L. L. (2002). *Journal of the American Chemical Society,* **124** (50), 14922–33.

[19] Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. A., Alvarez, R., Blixt, O., Taylor, M. E., Weis, W. I., & Drickamer, K. (2004). *Nat Struct Mol Biol,* **11** (7), 591–8.

[20] Lameignere, E., Shiao, T. C., Roy, R., Wimmerova, M., Dubreuil, F., Varrot, A., & Imberty, A. (2010). *Glycobiology,* **20** (1), 87–98.

[21] Luczkowiak, J., Sattin, S., Sutkevičiūtė, I., Reina, J. J., Sánchez-Navarro, M., Thépaut, M., Martínez-Prats, L., Daghetti, A., Fieschi, F., Delgado, R., Bernardi, A., & Rojo, J. (2011). *Bioconjug Chem,* **22** (7), 1354–65.

[22] Reina, J. J., Bernardi, A., Clerici, M., & Rojo, J. (2010). *Future Med Chem,* **2** (7), 1141–59.

[23] Reina, J. J., Sattin, S., Invernizzi, D., Mari, S., Martínez-Prats, L., Tabarani, G., Fieschi, F., Delgado, R., Nieto, P. M., Rojo, J., & Bernardi, A. (2007). *ChemMedChem,* **2** (7), 1030–6.

[24] Sattin, S., Daghetti, A., Thépaut, M., Berzi, A., Sánchez-Navarro, M., Tabarani, G., Rojo, J., Fieschi, F., Clerici, M., & Bernardi, A. (2010). *ACS Chem Biol,* **5** (3), 301–12.

[25] Schuck, P. (2000). *Biophys J,* **78** (3), 1606–19.

[26] Schuck, P. (2003). *Anal Biochem,* **320** (1), 104–24.

[27] Sisu, C., Baron, A. J., Branderhorst, H. M., Connell, S. D., Weijers, C. A. G. M., de Vries, R., Hayes, E. D., Pukin, A. V., Gilbert, M., Pieters, R. J., Zuilhof, H., Visser, G. M., & Turnbull, W. B. (2009). *Chembiochem,* **10** (2), 329–37.

[28] Svajger, U., Anderluh, M., Jeras, M., & Obermajer, N. (2010). *Cellular Signalling,* **22** (10), 1397–405.

[29] Tabarani, G., Reina, J. J., Ebel, C., Vivès, Lortat-Jacob, H., Rojo, J., & Fieschi, F. (2006). *FEBS Lett,* **580** (10), 2402–8.

[30] Tabarani, G., Thépaut, M., Stroebel, D., Ebel, C., Vivès, C., Vachette, P., Durand, D., & Fieschi, F. (2009). *J Biol Chem,* **284** (32), 21229–40.

[31] Thepaut, M., Guzzi, C., Sutkeviciute, I., Sattin, S., Ribeiro-Viana, R., Varga, N., Chabrol, E., Rojo, J., Angulo, J., Bernardi, A., Nieto, P. M., & Fieschi, F. (2012). *submitted,* , 1–28.

[32] Torre, J. G. D. L., Huertas, M. L., & Carrasco, B. (2000). *Biophys J,* **78** (2), 719–30.

[33] Turnbull, W. B. & Daranas, A. H. (2003). *Journal of the American Chemical Society,* **125** (48), 14859–66.

[34] van Kooyk, Y. & Geijtenbeek, T. B. H. (2003). *Nature Reviews Immunology,* **3** (9), 697–709.

[35] Zhao, H., Balbo, A., Brown, P. H., & Schuck, P. (2011). *Methods,* **54** (1), 16–30.

#### 9.1.4.5. *The bis-benzylamide derivatives of psDi to achieve better affinity and selectivity to DC-SIGN*

The mannose-based glycomimics development focused on the improvement of psDi molecule. The solved structure of psDi in complex with DC-SIGN CRD (fig. 9.16) showed that the methyl groups at cyclohexane ring of the compound are in a favorable position to be changed to more complicated moieties that could extend over the protein surface by making additional contacts and thus increasing the affinity to the protein.

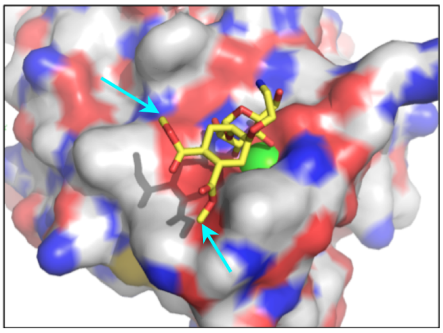


**Figure 9.16.:** psDi binding within DC-SIGN CRD.
The D-mannose moiety of the molecule is oriented to $Ca^{2+}$ (a green sphere) in the CRD through 3-OH and 4-OH groups, while the cyclohexane derivative is solution exposed, but still it makes Van der Waals contacts with Val351 residue of the protein. The cyan arrows point to the methyl groups that can be substituted to more complex moieties in order to improve the affinity of the molecule to DC-SIGN CRD. Color scheme: carbons of the protein and psDi are white and yellow, respectively; oxygens and nitrogens are red and blue; sulfur of the protein is yellow.

A group of psDi-based compounds, bis-benzylamides, with a general formula shown in figure 9.17A have been synthesized by a PhD student Norbert Varga in the team of Pr. A. Bernardi. The efficiency of these compounds to inhibit DC-SIGN and their selectivity to DC-SIGN versus langerin was assessed, and the results are presented in the following article by Varga et al.

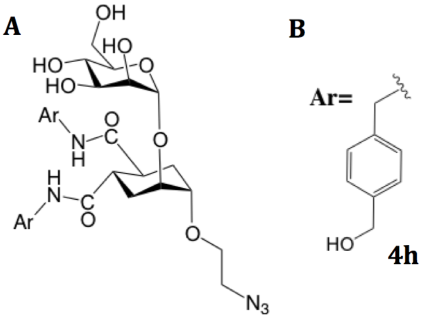


**Figure 9.17.:** The structure of new psDi-based compound series.
**A**, The general formula of bis-benzylamides; “Ar” stands for aromatic substituent. **B**, The “Ar” substituent of the new lead compound **4h**.

One of the compounds in this series, the bis-amide **4h** (fig. 9.17B), showed a proper combination of such features as good affinity and selectivity to DC-SIGN together with good water solubility and synthetic accessibility. It was chosen for further affinity improvement by mounting it on the multivalent scaffolds. To get a better understanding of its interaction with DC-SIGN, the structural analysis by NMR spectroscopy was performed. These NMR studies included the determination of the binding epitope of **4h** and the estimation of the compound's conformation in free and in the bound states. I have been directly involved in these NMR experiments and data analysis during the long-term secondment in the group of our collaborator Dr. P. Nieto, where I had an opportunity for the first time to learn the basics of NMR and try to apply the knowledge to study DC-SIGN/**4h** interaction. The stay supervised by Dr. Jesus Angulo and Dr. Pedro Nieto in the NMR laboratory in Seville (Instituto de Investigaciones Químicas, CSIC) lasted for 2 months and a half starting 26th September, 2011.

The saturation transfer difference (STD) NMR experiment was used to determine the **4h** binding epitope. Figure 9.18 summarizes the basic principle and the steps of STD NMR experiment, which is based on Nuclear Overhauser Effect (NOE) [171].

The first step is the so-called off-resonance experiment, which is done to record a standard 1D $^1$H spectrum, or a reference spectrum. Then the on-resonance experiment follows: some of the protein protons are selectively irradiated, or saturated, (ligand protons remain unaffected) using a low power radiofrequency for a specific period, i.e. saturation time. Subsequently, the intermolecular NOE arise between protein and the bound ligand protons, that is the "magnetization" (perturbation, or saturation, of protein proton polarization) from protein protons is transferred to the bound ligand protons, which are in close proximity with the protein (<6 Å).

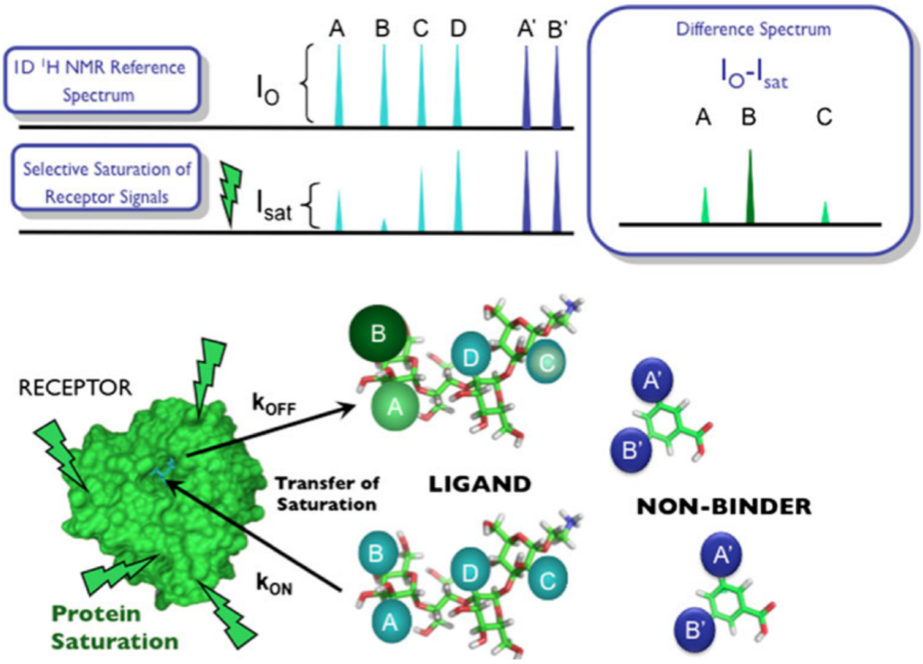


**Figure 9.18.:** Schematic explanation of 1D STD NMR experiment.
The top part illustrates the experiment itself, while the bottom part shows the sample composed of a protein and a pool of small test molecules in a molar excess with respect to the protein. (From [172])

The information about the binding appears in the saturated spectrum where the intensities of the protons that were closer to the protein are reduced, while the intensities of the protons that did not participate in binding are not affected. The subtraction of the saturated spectrum from reference spectrum yields a difference spectrum, where the STD intensities only of those protons that were close to protein are visible, and the stronger intensities indicate the closer distance to the protein. Normalizing the intensities with respect to the most intense, to which 100% value is assigned, the ligand epitope can be mapped.

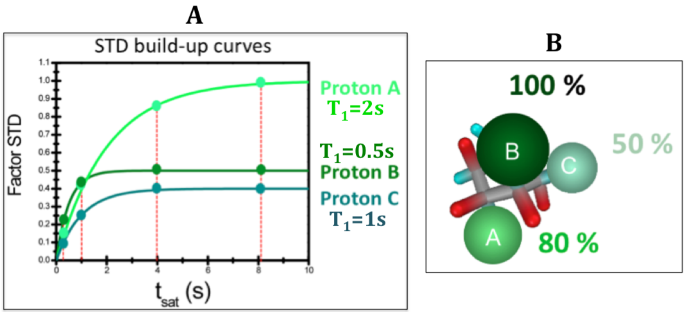


**Figure 9.19.:** An illustration of the determination of a binding epitope of a small molecule. **A**, An example of STD build-up curves with the relaxation times given for the corresponding protons $A$, $B$ and $C$; without estimation of initial slopes and taking STD intensities at tsat of 4 s or 8 s, the epitope would be misinterpreted putting the highest STD intensity possessing proton $A$ closest to the protein (100% STD), then protons $B$ and $C$. **B**, A corresponding correct binding epitope deduced from initial slopes of shown STD build-up curves: when the molecule is bound to the protein, proton $B$ (highest value of initial slope, 100%) is closest to the protein and proton $C$ is further, while proton $A$ takes the intermediate position with respect to protons $B$ and $C$.

In order to avoid artifacts arising from the different relaxation $R_1$ (i.e. loss of magnetization) rates for different protons (fig. 9.19A), a proper ligand epitope map (fig. 9.19B) is obtained from the STD intensity build-up curves (fig. 9.19A) [172]. For this reason, 1D STD NMR experiments are performed at several different saturation times, and after plotting the resulting STD intensities against saturation time, the initial slopes are calculated by fitting mono-exponential function:

$$STD(t_{sat}) = STD_{max} \cdot (1 - \exp(-k_{sat} \cdot t_{sat})) \tag{9.4}$$

where $STD(t_{sat})$ is the observed STD intensity, $STD_{max}$ is the asymptotic maximum of the build-up curve, $t_{sat}$ is the saturation time, and $k_{sat}$ is a rate constant related to the relaxation properties of a given proton that measures the speed of STD build-up. When $k_{sat}$ and $STD_{max}$ are derived by least-squares fitting, the initial slope may be calculated:

$$\frac{dSTD}{dt(0)} = STD_{max} \cdot k_{sat} \tag{9.5}$$

The calculated initial slopes for each interacting proton are normalized with respect to the highest value, and thus the proper ligand epitope is determined.

The main outcome of these studies:

- ✧ Almost all of the bis-benzylamide derivatives of psDi yielded a better affinity and selectivity to DC-SIGN than langerin.
- ✧ A new lead compound, **4h**, selected for further improvement by tethering on multivalent scaffolds.
- ✧ **4h** binds in the same binding mode as psDi validating the rational optimization approach.

Contributions:

All psDi derivatives were synthesized by a PhD student Norbert Varga in the team of Pr. A. Bernardi. The docking studies of **4h** were performed by postdoctoral researcher John McGeagh in the team of Dr. J. Weiser in order to envision the binding mode of **4h** in DC-SIGN CRD.

My contribution to this study:

- ✧ Prepared DC-SIGN ECD required for SPR and NMR studies.
- ✧ Performed all of the described SPR assays, and analyzed the data for all psDi derivatives.
- ✧ Participated in STD NMR experiments and contributed to data analysis with the help of Dr. Jesus Angulo and Cinzia Guzzi.
- ✧ Selection of a plausible **4h** docking pose out of 60 calculated poses by checking it against STD NMR data.
- ✧ Participated in article writing.

# 6

# 6

***Paper n$^o$6:*** **Selective targeting of DC-SIGN with mannose-based glycomimetics. Synthesis and interaction studies of bis-benzylamide derivatives of a pseudomannobioside**

CHEMISTRY
A EUROPEAN JOURNAL

FULL PAPER

DOI: 10.1002/chem.201202764

# 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

**Norbert Varga,[a] Ieva Sutkeviciute,[b, c, d] Cinzia Guzzi,[e] John McGeagh,[f] Isabelle Petit-Haertlein,[b, c, d] Serena Gugliotta,[a] Jörg Weiser,[f] Jesùs Angulo,[e] Franck Fieschi,[b, c, g] and Anna Bernardi*[a]**

**Abstract:** Dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN) and Langerin are C-type lectins of dendritic cells (DCs) that share a specificity for mannose and are involved in pathogen recognition. HIV is known to use DC-SIGN on DCs to facilitate transinfection of T-cells. Langerin, on the contrary, contributes to virus elimination; therefore, the inhibition of this latter receptor is undesired. Glycomimetic molecules targeting DC-SIGN have been reported as promising agents for the inhibition of viral infections and for the modulation of immune responses mediated by DC-SIGN. We show here for the first time that glycomimetics based on a mannose anchor can be tuned to selectively inhibit DC-SIGN over Langerin. Based on structural and binding studies of a mannobioside mimic previously described by us (**2**), a focused library of derivatives was designed. The optimized synthesis gave fast and efficient access to a group of bis(amides), decorated with an azide-terminated tether allowing further conjugation. SPR inhibition tests showed improvements over the parent pseudomannobioside by a factor of 3–4. A dimeric, macrocyclic structure (**11**) was also serendipitously obtained, which afforded a 30-fold gain over the starting compound (**2**). The same ligands were tested against Langerin and found to exhibit high selectivity towards DC-SIGN. Structural studies using saturation transfer difference NMR spectroscopy (STD-NMR) were performed to analyze the binding mode of one representative library member with DC-SIGN. Despite the overlap of some signals, it was established that the new ligand interacts with the protein in the same fashion as the parent pseudodisaccharide. The two aromatic amide moieties showed relatively high saturation in the STD spectrum, which suggests that the improved potency of the bis(amides) over the parent dimethyl ester can be attributed to lipophilic interactions between the aromatic groups of the ligand and the binding site of DC-SIGN.

**Keywords:** DC-SIGN · glycomimetics · HIV · NMR spectroscopy · proteins

## Introduction

The dendritic cell (DC) membrane C-type lectin receptor dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN) has been implicated in the early stages of HIV infection and is, therefore, considered an interesting target for the design of antiviral agents.[1,2] Immature dendritic cells in mucosal tissue use DC-SIGN to recognize high-mannose glycans present on the viral envelope glycoprotein gp120 of HIV. This recognition event appears to contribute to infection by promoting viral transmission in a manner dependent on the composition of gp120

---

[a] N. Varga,[+] S. Gugliotta, A. Bernardi
Universita' di Milano, Dipartimento di Chimica
via Golgi 19, 20133 Milano (Italy)

[b] I. Sutkeviciute,[+] I. Petit-Haertlein, F. Fieschi
Institut de Biologie Structurale, Université Grenoble I
41 rue Jules Horowitz, Grenoble, 38027 (France)

[c] I. Sutkeviciute,[+] I. Petit-Haertlein, F. Fieschi
CNRS, UMR 5075, Grenoble, 38000 (France)

[d] I. Sutkeviciute,[+] I. Petit-Haertlein
CEA, DSV, Grenoble, 38000 (France)

[e] C. Guzzi, J. Angulo
Glycosystems Laboratory, Instituto de Investigaciones
Químicas (IIQ), CSIC - Universidad de Sevilla
Américo Vespucio 49, 41092 Sevilla (Spain)

[f] J. McGeagh, J. Weiser
Anterio Consult&Research GmbH, Augustaanlage 23
68165 Mannheim (Germany)

[g] F. Fieschi
Institut Universitaire de France
103 boulevard Saint-Michel 75005 Paris (France)

[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201202764.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

These are not the final page numbers! ↗↗

WILEY-VCH Galley proofs

glycans.[3] Ligands that can be used to explore the multiple functions of DC-SIGN and to inhibit DC-SIGN-mediated pathogen binding, are actively pursued by different groups.[4,5,6] Selective inhibition of DC-SIGN, however, has rarely been addressed, despite the presence in the immune system of many other C-type lectins that share DC-SIGNs ability to bind high-mannose oligosaccharides. Among these, Langerin, which is expressed on the surface of a different subset of antigens presenting cells known as Langerhans cells (LCs), has been suggested to have protective effects against HIV infection.[7] Selective DC-SIGN ligands, which interact only weakly with Langerin, are potentially useful therapeutic tools against sexually transmitted HIV infection.[6d,8]

Glycomimetic structures designed to inhibit DC-SIGN have been reported.[4a–c,6] In our previous work, we focused on the design and synthesis of monovalent *N*-fucosylamides[6b,d] or on mannose conjugates, such as **1**–**3** (Figure 1).[6a,c,g] These molecules have been conceived to be metabolically more stable than native oligosaccharides and can be equipped with reactive moieties that enable a multimeric presentation on different scaffolds. Indeed, the pseudotrisaccharide **1**, mimicking the linear Man-α-1,2-Man-α-1,6-trimannoside, the D3 arm of the high mannose structure Man$_9$, was found to inhibit DC-SIGN binding to mannosylated BSA■■please define■■ (Man-BSA) with an IC$_{50}$ of 130 μM by surface-plasmon resonance (SPR).[6c] Moreover, when presented on a tetravalent dendron, **1** was also found to inhibit HIV transinfection of CD4+T lymphocytes[6c] and of cervical tissue explants[6f] at micromolar concentrations. We have also recently reported a small library of mannose conjugates of general formula **3**, which were tested by using a dendritic cell adhesion assay to mannan-coated plates.[6 g] The activity of these compounds was tuned by the nature of the amide R group, but most of them efficiently inhibited DC-SIGN-mediated adhesion. The data showed that the activity increased if the R group was lipophilic. However, most compounds in this series possess a low solubility in aqueous media, which limits their applicability. Preliminary NMR spectroscopic studies of their interaction with DC-SIGN also indicated that some nonspecific association occurred through the allyl linker.

Recently, structural studies combining X-ray crystallography and NMR spectroscopic analysis allowed us to characterize the binding mode of **2** within the DC-SIGN binding site and to demonstrate that it corresponds to a unique well-defined orientation, making **2** a good lead compound for chemical improvement of the binding affinity and specificity towards DC-SIGN.[8] Analysis of the X-ray structure showed that the methyl ester groups of **2** in the complex extend towards the DC-SIGN surface in an area in which larger substituents can be accommodated. Larger groups in this position can be expected to reach the protein surface and make additional contacts. These observations validated the initial strategy that had led to the synthesis of bis(amides) **3** and suggested that their structures should be reconsidered with the goal of increasing water solubility and removing the allyl group, as the source of undesired nonspecific interactions with the protein. In the group of compounds previously examined, the best balance of synthetic accessibility, selectivity, and solubility was achieved for benzylamides **3a**–**b**, which showed an IC$_{50}$ in the μM range in the adhesion assay. These molecules provided interesting leads, which could be improved by optimizing the nature of the benzylamide substituents and by replacing their allylether appendage with a functional tether, which would also allow multimeric presentation of the ligands.

With these goals in mind, a group of functionalized bis(benzylamides) of general formula **4** (Scheme 1) was synthesized. They contain an azidoethanol linker, which can be exploited for conjugation, and different polar groups on the benzylamide ring, to increase water solubility and possibly improve affinity for the receptor. These molecules were tested for their activity as DC-SIGN ligands by using an SPR competition assay. Selectivity against Langerin was also addressed. NMR spectroscopic studies were performed to analyze the interaction of the ligands with DC-SIGN and establish the epitope region. Altogether these results showed that bis(benzylamido) pseudodimannosides bind selectively to DC-SIGN and allowed the selection

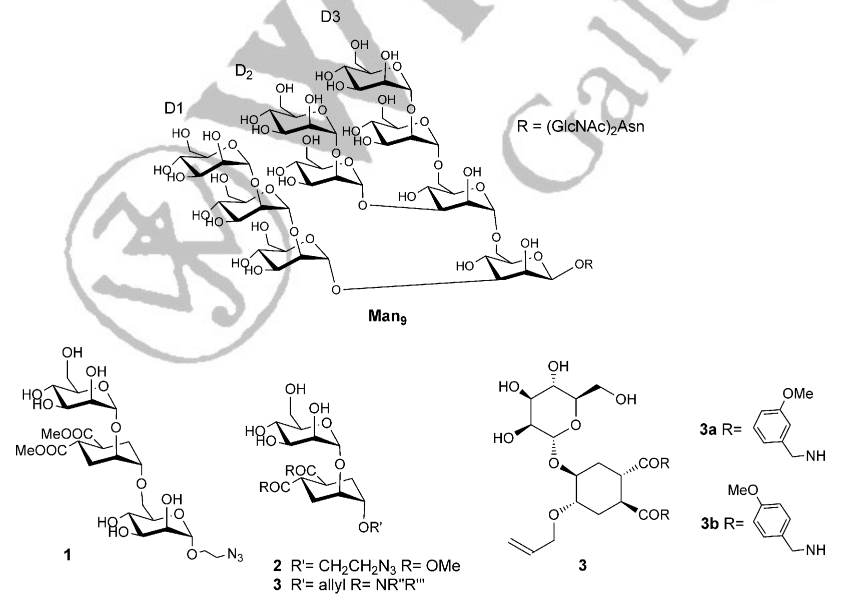
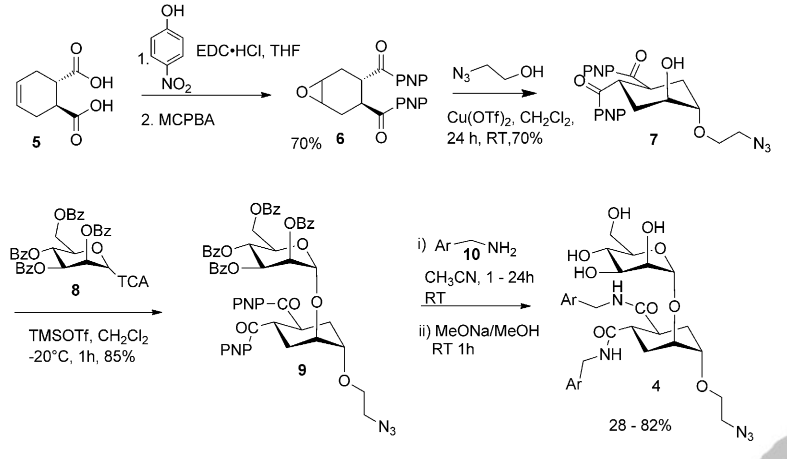


Figure 1. Structure of **Man9** and of previously reported mannose-based DC-SIGN inhibitors **1**–**3**.

WILEY-VCH Galley Proofs



Scheme 1. Synthesis of bis(benzylamides) **4**.

of **4h** as the optimal ligand for the future construction of multivalent systems. In the course of this study a divalent macrocyclic ligand (**11**) was also serendipitously obtained, which showed a low μM affinity for DC-SIGN in the SPR assay used.

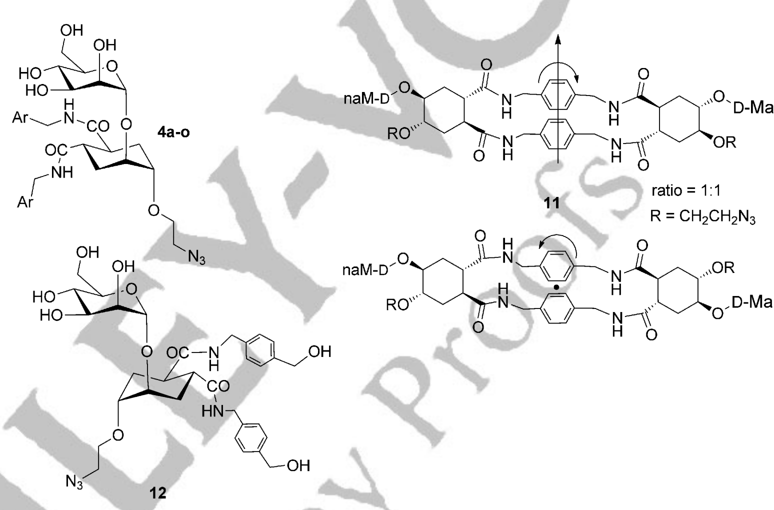


Figure 2. The bis-benzylamides **4a–o, 11** and **12** synthesized and tested in this work.

## Results and Discussion

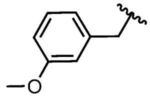
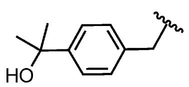
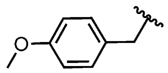
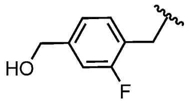
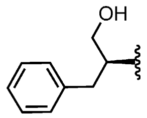
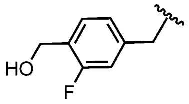
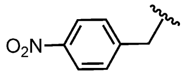
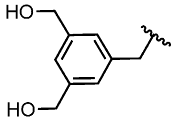
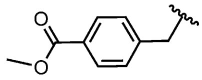
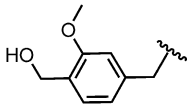
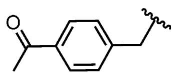
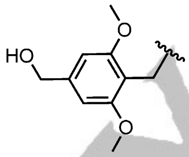
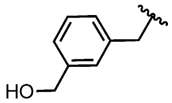
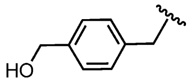
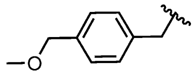
A high-yielding and flexible synthesis of the entire library was achieved by small modifications of the previously reported procedure.[6g] By starting from enantiomerically pure (1*S*,2*S*)-diacid **5**,[9] the bis(*p*-nitrophenylester) was synthesized and the double bond oxidized by using *m*CPBA (*meta*-chloroperbenzoic acid), to afford epoxide **6**. Copper-catalyzed epoxide opening[10] with azidoethanol,[11] followed by mannosylation of alcohol **7** with trichloroacetemidate **8**[12] afforded the pseudodisaccharide **9** in only four steps and with 42 % overall yield. The two activated *p*-nitrophenyl esters, used as acid-protecting groups throughout the sequence so far, were transformed into amides by reaction with an excess of the appropriate benzylamine (**10**), yielding bis(amides) **4** after Zemplén deprotection of the sugar.

Docking studies[6g] had suggested that H-bond donors and acceptors out of the plane of the aromatic ring could be beneficial for interaction with the protein surface. This consideration, together with water solubility and synthetic availability, drove the selection of benzylamines **10** and of the group of 16 compounds **4a–o** (Figure 2, Table 1) to be synthesized, all bearing polar, oxygen-containing groups on the aromatic ring. Bis(benzylamines) were also considered; however, their reaction with **9** led mostly to complex mixtures. By only using *p*-xylylendiamine (**10p**, see Scheme S1 in the Supporting Information) one major product could be chromatographically isolated in modest yields (18 %). MS and $^{13}$C NMR spectroscopic analysis revealed that the dimeric macrocyclic structure **11** was assembled, as a 1:1 mixture of inseparable diastereoisomers that was tested as such in the following assays. Finally, to analyze the effect of the stereochemistry of the cyclohexane scaffold, compound **12**, a stereoisomer of **4h**, was synthesized by starting from the (1*R*,2*R*)-enantiomer of **5** (see Scheme S8 in the Supporting Information). Analysis of the $^1$H NMR spectra confirmed that all compounds share the same chair conformation of the cyclohexane ring observed in **2**[13] (see Figure S1 in the Supporting Information).

The activity of these compounds as inhibitors of DC-SIGN binding to mannosylated BSA (Man-BSA) was tested by SPR (Figure 3). In the assay, Man-BSA was immobilized on a CM4 chip and the extracellular domain (ECD) of DC-SIGN was injected over the Man-BSA surface in the absence or in the presence of increasing concentrations of the inhibitors. The obtained IC$_{50}$ values are shown in Table 1■ ■ok?■■. The parent pseudodisaccharide bis(methyl ester) **2** and the pseudotrisaccharide **1** are included in the Figure 3 data as reference compounds.■■ok?■■ Many of the diamides prepared showed a remarkable increase in inhibitory activity compared to diester **2** and some approached the affinity observed for **1**, a molecule of significantly higher

WILEY-VCH
Galley Proofs

Table 1. Structure and activity of DC-SIGN ligands **4a–o**, **11**, and **12**.

| | Structure (R=) | Yield[a] [%] | IC$_{50}$[b] [μM] | | Structure (R=) | Yield[a] [%] | IC$_{50}$[b] [μM] |
|---|---|---|---|---|---|---|---|
| **4a** |  | 67 | 296 | **4j** |  | 65 | 367 |
| **4b** |  | 73 | 441[c] | **4k** |  | 76 | 398 |
| **4c** |  | 75 | 814 | **4l** |  | 64 | 335 |
| **4d** |  | 72 | 810[c] | **4m** |  | 63 | 317 |
| **4e** |  | 60 | 324 | **4n** |  | 80 | 405 |
| **4f** |  | 59 | 310 | **4o** |  | 28 | 1461 |
| **4g** |  | 65 | 356 | **11** | dimer (Figure 2) | – | 31 |
| **4h** |  | 82 | 325 | **12** | diastereoisomer of **4h** (Figure 2) | 65[d] | 1421 |
| **4i** |  | 78 | 290 | **2** | see Figure 1 | – | 986 |

[a] Unoptimized overall yield from **9**. [b] In SPR competition test with immobilized Man-BSA. [c] Low solubility in water. [d] Isolated as a byproduct from glycosylation of a batch of nonenantiomerically pure **7** (see the Supporting Information for details).

structural and synthetic complexity. Low solubility remains a problem for some of the structures (**4a** and **4d**, shown in light grey in Figure 3)■■ok?■■, but most compounds displayed good solubility in water at the concentration required for the assay.

The group of diamides **4e–n**, built upon the modeling suggestion mentioned above, showed a remarkable activity, as all the compounds were found to be only 2.5-fold less potent than the pseudotrisaccharide **1** (Figure 1). Among them, **4h**, featuring a hydroxymethylene group in the *para* position, was selected for structural modifications. The hydroxy group appears to play a role as a H-bond acceptor, since the corresponding methyl ether shows the same inhibitory potency (cf. **4h** and **4i**). Addition of fluorine atoms on the ring (**4k–l**) or additional lipophylic groups in the proximity of the acceptor (**4j**) did not significantly improve the affinity. On the contrary, two methoxy substituents *meta* to the hydroxymethylene group, as in compound **4o**, had a negative effect, possibly as a result of a different orientation of the aromatic residue around the *N*-benzylic bond. Similarly, compound **12**, a diastereoisomer of **4h** with the opposite configuration (1*R*,2*R*) of the cyclohexane ring, binds poorly to DC-SIGN, which indicates that the design of the pseudodisaccharide scaffold is important to optimize interaction with the protein as suggested by the structure of the **2**/DC-SIGN complex.[8] Remarkably, dimer **11** (1:1 mixture of isomers) with an IC$_{50}$ of 31 μM turned out to be the most potent inhibitor of the series, and one of the most effective reported so far. The dimension of the macrocycle in **11** (OH-4–OH-4 distance: 22 Å) is too short to span two adjacent binding sites of the DC-SIGN tetramer (which are separated by no less than 35 Å),[14] thus the potency of this compound must derive from proximity effects (statistical rebinding) or by aggregation of the protein,[15] which, in turn, could block the availability of DC-SIGN binding sites and thus effectively inhibit binding to immobilized Man-BSA. Since the effect is rather large, protein oligomerization under the assay conditions appears to be the most likely explanation. Protein aggregation by polyvalent carbohydrates is very well documented;[16] however, given the heterogeneity of the compound and the low yield with which it was synthesized, proving the mechanism underlying the activity of **11** would be challenging. Furthermore, even if an oligomerization mechanism could be demonstrated under the conditions of the current assay (by using soluble DC-SIGN extra-cellular domain (ECD)), it is very unlikely that it will be operative in the biologically relevant conditions, in which the protein would be membrane-bound. Thus the issue was not pursued further.

In parallel, most of the library members were further tested for inhibition of Langerin by using an SPR assay that we have recently described.[6d] Selectivity for DC-SIGN versus Langerin is particularly important to develop inhibitors that would block sexually transmitted HIV infection. As we have already remarked, while interaction with DC-

WILEY-VCH Gallery Proofs

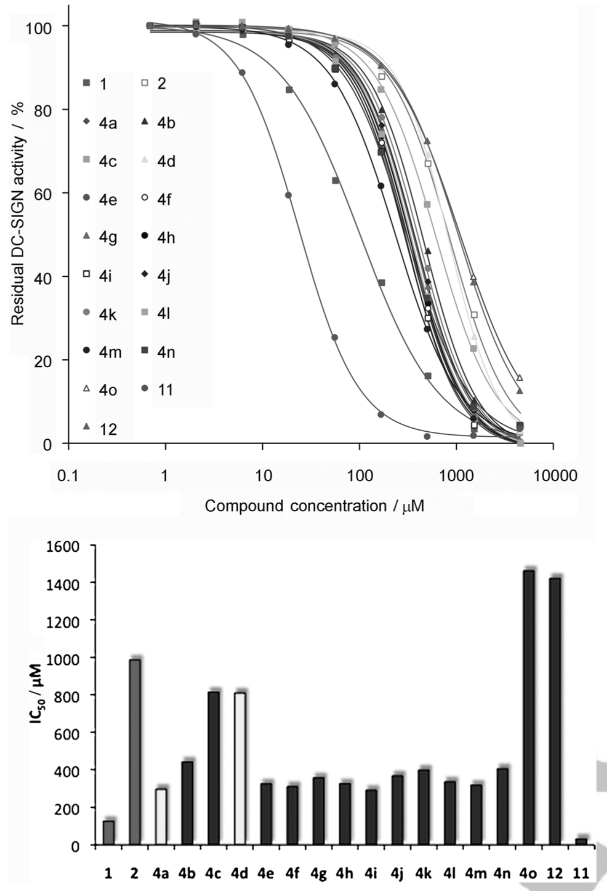


Figure 3. SPR results representing inhibition of DC-SIGN ECD binding to Man-BSA surface by mannose-based compounds: inhibition curves (top) and extracted corresponding IC$_{50}$ values of mannose-based ligands (bottom). In dark grey: IC$_{50}$ bars of the reference compounds **1** and **2**. In light grey: compounds that displayed low solubility under the assay conditions.

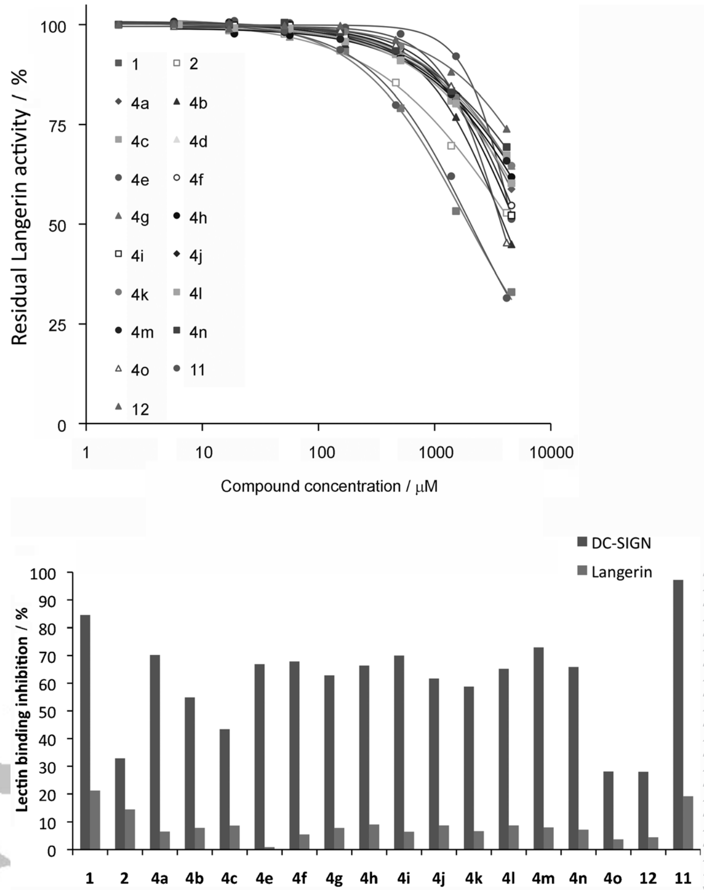
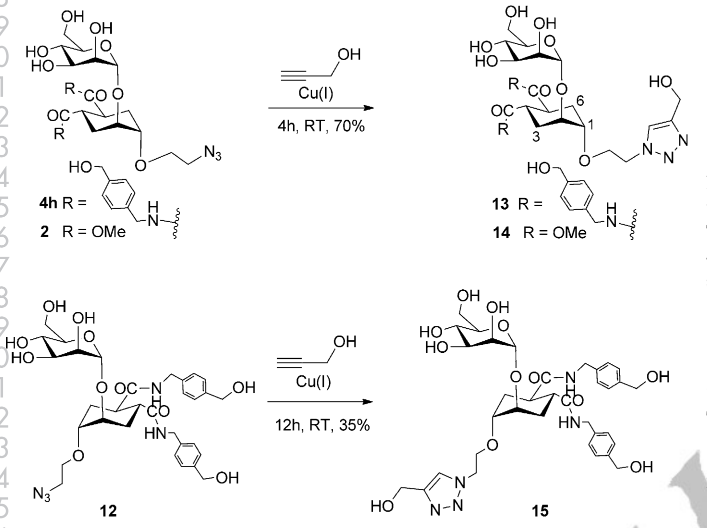


Figure 4. Inhibition of Langerin ECD binding to the Man-BSA/dextran SPR surface by mannose-based compounds: inhibition curves (a), DC-SIGN and langerin Inhibition level obtained for 0.5 mM of the compounds (b).

SIGN on mucosal DCs is used by the virus to invade the host immune system, Langerin was shown to have protective effects against HIV infection.[7] To get an insight into the selectivity of ligands **1**–**4**, they were tested in the same concentration range as inhibitors of DC-SIGN (20 µM) and of Langerin (15.5 µM) binding to dextran/Man-BSA surface. As can be seen from the inhibition curves (Figure 4a), none of the tested compounds were able to fully inhibit Langerin binding to the surface, despite the lower concentration of Langerin used in the assay relative to DC-SIGN ECD concentration (due to limited availability of the protein). The most potent inhibitors were found to be dimer **11** and the references **1** and **2**. Because most of the compounds at the highest concentration used did not reach 50% inhibition of Langerin, it was not possible to calculate their IC$_{50}$ values. Nevertheless, to get a comparison of DC-SIGN versus Langerin inhibition, the inhibition levels (%) achieved at three compound concentrations, 0.5, 1.0, and/or 1.5 mM, were determined. Data obtained at 0.5 mM are plotted in Figure 4b (see Figure S2 in the Supporting Information for DC-SIGN and Langerin inhibition levels at 1 and 1.5 mM ligand concentration). The data show that the pseudo-disaccharide bis(methylester) **2** is endowed with a modest selectivity for DC-SIGN as previously reported.[8] The corresponding pseudotrisaccharide **1** and many of the bis(amides) **4**, that is, **4a**, **4f**, **4h**, and **4i**, were found to inhibit DC-SIGN with an improved selectivity relative to Langerin, a very desirable feature for further development as antiviral agents. Fucose-based ligands, both natural[17] and unnatural[6d] have been reported to recognize DC-SIGN selectively. However, mannose-containing oligosaccharides appear to bind both lectins[17,18] and it has long been believed that selectivity could not be achieved with mannose-based ligands. Our current results show that this goal can be reached.

With this information in hand, bis(amide) **4h** was selected for future development in multivalent constructs, since this ligand appears to combine good affinity and selectivity for DC-SIGN with good solubility, synthetic accessibility, and atom efficiency. This ligand was therefore chosen for structural analysis by NMR spectroscopy to eventually identify ligand-receptor contacts that are essential for complex formation and compare its binding mode to that observed for **2** in the X-ray structure of its complex with DC-SIGN.[8] Initial attempts were frustrated because the high-field (500 and 600 MHz) $^1$H NMR spectrum of **4h** showed some overlapping of signals in areas crucial for our analysis. Previous NMR work on the parent bis(methylester) **2** had shown that its triazole derivative **14** (Scheme 2) is a better substrate for spectroscopic analysis. The triazole residue in the linker of **14** did not affect binding to DC-SIGN ECD relative to **2**, neither in terms of affinity nor in terms of protein-ligand contacts, but it significantly improved the dispersion of signals in the $^1$H NMR spectrum and allowed a full characterization of the binding epitope.[8] Thus, based on this experience, **4h** was transformed into the corresponding triazole **13** (Scheme 2). For comparison, and to examine the role played by the scaffold configuration in pseudomannobioside mimics, the diastereomeric bis(amide) **12** was also transformed into the corresponding triazole **15** and its interaction with DC-SIGN was studied by NMR spectroscopy.



Scheme 2. Triazole derivatives used for NMR spectroscopic interaction studies.

The spectrum of **13** is indeed more resolved than the spectrum of **4h**, but not fully resolved. In particular the 6ax proton of the cyclohexane ring (H6ax(C)) and the 3 proton of mannose (H3(M)), both belonging to the epitope of **14**, overlap with protons 6eq (H6eq(C)) and 3eq (H3eq(C)) (m, at $\delta = 1.89$–$1.82$ ppm) and with proton 2 (H2(C)) (m, at $\delta = 3.84$ ppm) of the cyclohexane ring, respectively (Figure 5a, see Scheme 2 for numbering of the cyclohexane ring). Nonetheless saturation transfer difference (STD) NMR spectroscopic studies could be performed and the data were compared with those obtained for the reference ligand **14**. STD-NMR spectra is a useful and robust technique to obtain structural information on molecular recognition processes within the range of fast exchange kinetics ($K_D$ from the 0.1 to 10 mM range).[19] Attempts to resolve signal overlapping by varying the temperature were ineffective and some of the most interesting signals could not be quantified. The STD spectrum of **13** in the presence of DC-SIGN ECD (50:1 ligand/protein ratio) at room temperature is reported in Figure 5b.

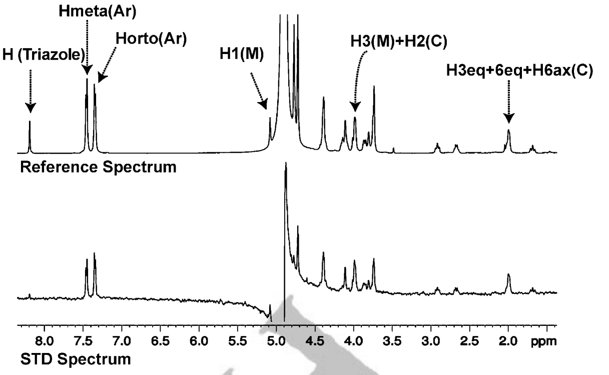


Figure 5. STD NMR spectroscopic study of the interaction of ligand **13** with DC-SIGN in buffered $D_2O$ (150 mM NaCl, 4 mM $CaCl_2$, 25 mM d-Tris, pD 8). a) $^1$H NMR reference spectrum (off-resonance frequency 40 ppm) and b) STD spectrum (on-resonance frequency $\delta = 0$ ppm) of a sample containing **13** (1 mM) and DC-SIGN ECD (19 μM) at 25 °C (500 MHz) and a saturation time of 0.5 s. Key proton signals are labeled.

To obtain information on the binding epitope, the initial slope approximation ($STD_0$)[20] was used, to reduce interferences from differences in relaxation properties of each proton of the ligand and from rebinding processes.[21] The experimental STD growth curves are collected in Figure 6.

Despite the signal overlap problems, a detailed comparison of the STD data obtained for **13** and **14** allows us to qualitatively establish that the two molecules share the same binding mode. For **14**, quantitative analysis of the STD results had shown a very good match with the three-dimensional structure of the DC-SIGN/**2** complex determined by X-ray crystallography (Figure 7a): binding occurs through coordination of the calcium ion by OH-3 and OH-4 of the mannose residue in an orientation which allows the cyclohexane ring to establish a van der Waals contact with the V351 side chain of the receptor; this results in important

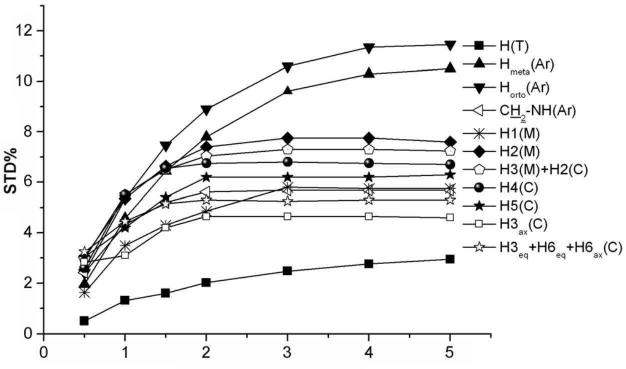


Figure 6. Experimental STD growth curves of **13** as a function of saturation time.

© WILEY-VCH Galley Proofs

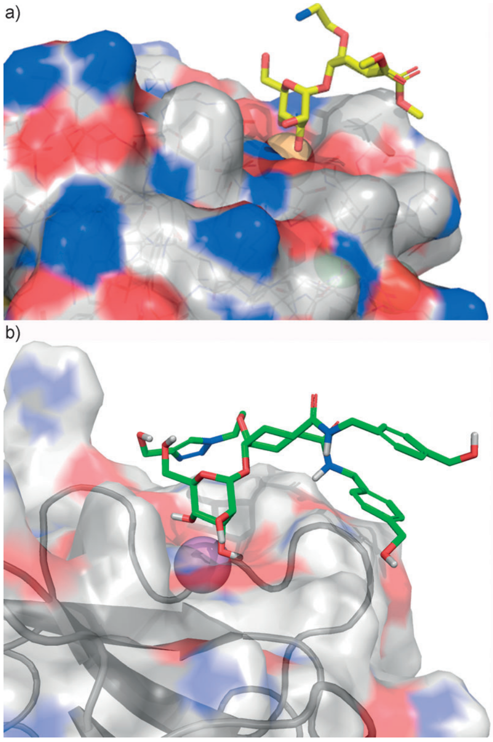


Figure 7. a) X-ray structure of the DC-SIGN/**2** complex (from ref. [8]). b) Docked pose of **13** that qualitatively accounts for the STD data.

STD signals for the mannose moiety and for all cyclohexane protons, particularly H6ax(C) (see Figure S3 in the Supporting Information). Although no accurate integration of some of the key protons was possible for **13**, overall the same pattern of STD intensities emerged (for a comparison of the $STD_0$ values, see the Table S1 in the Supporting Information). For a comparison of the full STD spectra see Figure S3 in the Supporting Information) leaving little doubt that **13** has the same binding mode as **14**. Additionally, the $STD_0$ NMR spectroscopic signals obtained for **13** are higher than those for the corresponding protons of **14** under the same conditions (see Table S1 in the Supporting Information). If both binding modes are the same, this is also indicative of a higher affinity of **13** relative to the parent compound. Indeed, the bis(benzylamido) moieties of **13** appear to establish further contacts with the protein (H(Ar), Figures 5 and 6), which can explain the significant increase in affinity for **13** compared to **14**. Docking of mimic **13** in the 1L4 crystal structure[22] was carried out by using the quantum polarized ligand docking (QPLD) workflow[23] based on the docking algorithm Glide.[24] This procedure identified poses consistent with the crystal structure of the DC-SIGN/**2** complex,[8] and also displayed additional interactions between the bis(benzylamido) moieties and the protein surface. A docking pose that qualitatively accounts for the STD data is shown in Figure 7b.

The STD spectra of the diastereomeric ligand **15** obtained under the same conditions suggested a very different binding mode involving the mannose moiety (H3(M), H4(M), and H2(M)), but showing only a single contact between the protein and the cyclohexane ring, through the H4 proton (H4(C); see Table S1 and Figure S3 in the Supporting Information). These data are in agreement with the lower affinity measured for **15** in the SPR experiments and further support the notion that appropriate design of the cyclohexane ring configuration is crucial to the activity of the mannobioside mimics in this series.

Analysis of the NOESY spectra of **13** revealed the clear existence of the same diagnostic cross-peaks critically defining inter-residue contacts (highlighted in Figure 8) in both NOESY (Figure 8a, free-ligand state) and TR-NOESY (Figure 8b, bound ligand state). This suggests that there is no major variation in the conformation of **13** upon binding to DC-SIGN, or, in other words, that the lectin recognizes the main conformation(s) existing in solution.

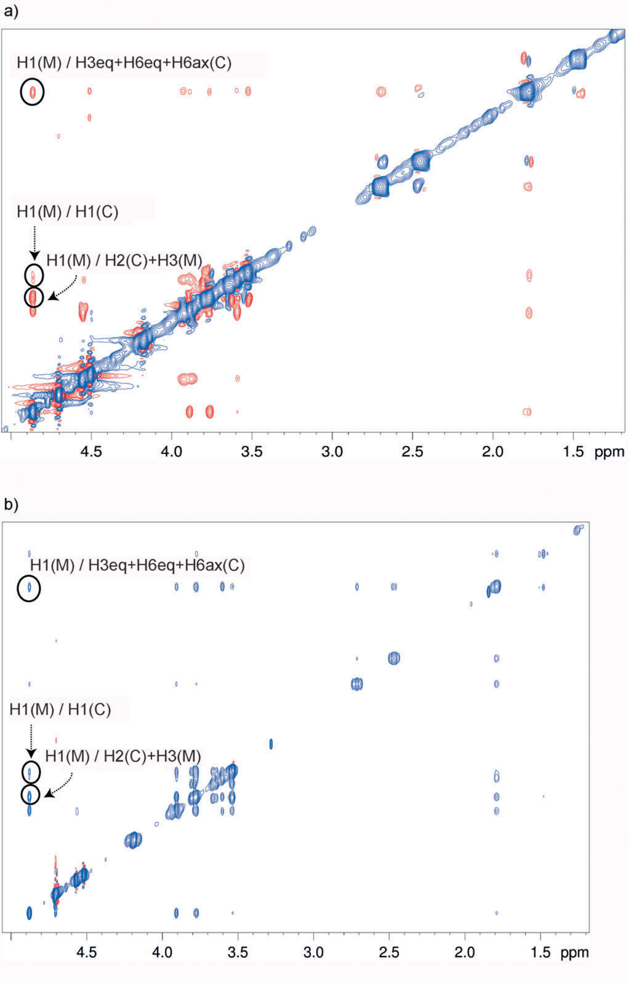


Figure 8. Expansions of NOESY experiments at 25 °C (500 MHz) of **13**. a) Free state, mixing time 500 ms. b) In the presence of DC-SIGN ECD, mixing time 500 ms. Labels indicate some key NOE peaks.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

## Conclusion

We have shown here that benzylic bis(amides) of general formula **4** bind to DC-SIGN with affinities approaching that of the synthetically more demanding pseudotrisaccharide **1**. Replacement of an allylic linker and inclusion of substituents designed to increase water solubility afforded a class of molecules that outperform the group of amides of general formula **3**[6g] previously found to bind somewhat nonspecifically to the protein and to display low solubility. Indeed, NMR interaction studies showed no sign of nonspecific association between DC-SIGN and compound **13**, derived from **4h**. These studies also allowed us to determine that **13** binds to DC-SIGN with the same binding mode of the parent pseudodisaccharide bis(methylester) **2**, that was recently determined by high-resolution X-ray analysis.[8] Among the group of molecules synthesized, the dimeric structure **11** was found to display the best affinity for DC-SIGN (31 μM). However, this ligand was not pursued further, since its activity is likely to depend on its ability to cross-link the DC-SIGN ECD tetramer, a feature that can contribute to increased selectivity in the format of the assay performed (inhibition of DC-SIGN ECD binding to immobilized Man-BSA), but might be not relevant in the real physiological situation (DC-SIGN "immobilized" on DC surface). Among the molecules prepared here, ligand **4h** was selected for further studies, due to a combination of good solubility, synthetic accessibility, activity, and selectivity. Polyvalent presentations of this ligand on dendrimeric scaffolds are presently being examined and will be reported in due course.

Probably the most important characteristic that we could determine for compounds **4a–o** is their selectivity for DC-SIGN against a second mannose-binding lectin, Langerin, also expressed by a (different) subset of dendritic cells and endowed with protective properties against HIV infection. Selective inhibition of DC-SIGN versus Langerin is considered an important requirement for the development of anti-infective agents able to block the first stages of infection during sexual transmission of the virus. The current observation that amides **4** display enhanced selectivity that remains significant up to 100 % DC-SIGN inhibition shows that the goal of selective mannose-based DC-SIGN ligands can be achieved. More generally, these results strongly suggest that the activity and selectivity of glycomimetic molecules based on a monosaccharide anchor can be tuned by structural modifications of secondary elements in the molecule and bode well for the development of selective low molecular weight lectin ligands.

## Experimental Section

DC-SIGN ECD protein (residues 66–404) were overexpressed and purified as described previously.[14] Langerin ECD construct, comprising residues 68–328, were overexpressed as inclusion body, then refolded and purified to homogeneity in a functional form as already described.[25] Surface plasmon resonance experiments were performed on a Biacore 3000 by using a CM4 chip, functionalized at 5 μL min$^{-1}$. Flow cells (Fc) 1 and 2 were activated with 50 μL of an 0.2 M EDC■■definition?■■/0.05 M NHS■■*N*-hydroxysuccinimide?■■ mixture, then Fc2 was functionalized with mannosylated bovine serum albumin (Man-BSA) by injecting protein (60 μg mL$^{-1}$) prepared in sodium acetate (10 mM, pH 4). Finally, the remaining activated groups of both flow cells were blocked with ethanolamine (30 μL, 1 M). After blocking, the three flow cells were treated with HCl (10 μL, 10 mM) to remove unspecific bound protein and then EDTA (20 μL, 50 mM) to expose surface to regeneration protocol. After these steps, BSA-Man (2000 RU, on average) was immobilized.■■ok?■■

For inhibition studies, mixtures of each lectin (DC-SIGN ECD (20 μM) or Langerin (15.5 μM)) with increasing concentrations of inhibiting compounds were prepared in a running buffer composed of Tris (25 mM, pH 8), NaCl (150 mM), $CaCl_2$ (4 mM), and P20 surfactant (0.005 %), and each sample (13 μL) was injected onto the surfaces at a 5 μL min$^{-1}$ flow rate. Concentrations of inhibiting compounds ranged from 0.7 to 4583 μM. The resulting sensorgrams were reference surface corrected, except for the case of Langerin; because this lectin displayed high affinity to the dextran matrix, a dextran/Man-BSA surface was considered as a combined ligand of Langerin, as previously described.[6d] The lectin-binding responses were extracted from the sensorgrams, converted to percent residual activity values ($y$) with respect to lectin-alone binding, and then plotted against corresponding compound concentrations. The 4-parameter logistic model [Eq. (1)] was fitted to the plots, and the $IC_{50}$ values were calculated by using the values of fitted parameters ($R_{hi}$, $R_{lo}$, $A_1$, and $A_2$) and Equation (2).

$$y = R_{int} - \frac{R_{hi} - R_{lo}}{1 + \left(\frac{c}{A_1}\right)^{A_2}} \tag{1}$$

$$IC_{50} = A_1 \left(\frac{R_{hi} - R_{lo}}{R_{hi} - 50} - 1\right)^{\frac{1}{A_2}} \tag{2}$$

**NMR spectroscopy experiments**: NMR spectroscopic experiments were recorded on a Bruker Avance DRX 500 MHz instrument equipped with a 5 mm inverse triple-resonance probe head. NMR samples were prepared in $D_2O$ (550 μL, 99.9 %) and, for the experiments in the presence of the receptor, in the same amount of buffer $D_2O$ (NaCl (150 mM), $CaCl_2$ (4 mM), d-Tris (25 mM, pD■■pH?■■ 8) and DC-SIGN ECD (19 μM). The concentration of the ligand was 5 mM for assignment experiments ($^1H$, COSY, TOCSY, NOESY, HSQC) and 1 mM for both STD NMR and transferred NOESY spectroscopic experiments.

STD NMR spectroscopic experiments were carried out at 10, 25, and 35 °C by using a train of Gaussian shaped pulses of 49 ms (field strength of ca. 80 Hz), an interpulse delay of 1 ms[19] and 15 ms spin-lock pulse (field strength of 3.7 kHz) prior to acquisition. The on-resonance frequency was set to $\delta = 0$ ppm■■not Hz for frequency?■■ and the off-resonance frequency was 40 ppm. Appropriate blank experiments were performed to assure the lack of direct saturation of the ligand protons. Saturation times of 0.5, 1, 1.5, 2, 3, 4, and 5 s were used to obtain the STD build-up curves.

The binding epitope was characterized by the analysis of initial slopes of the STD intensities[26]: the experimental ($I_0 - I_{sat}/I_0$) curves were fitted to an exponential function described by the equation: STD ($t_{sat}$) = $STD_{max}$ ($1 - e^{-ksat \cdot tsat}$), which allows us to calculate STD at zero saturation time (initial slopes) by the product of resulting parameters $STD_{max}$ and $k_{sat}$.[20]

NOESY experiments were carried out at 0.1, 0.2, 0.3, and 0.5 s, by using a phase-sensitive pulse program with gradient pulses in the mixing time and with presaturation.[27,28]

**Docking studies**

Protein setup: By starting from the crystal structure (resolution = 1.55 Å) of human DC-SIGN in a complex with Man4 (PDB code 1L4),[22] a molecular model of the protein was constructed. By using the 'Protein Preparation Wizard' within the Maestro graphical interface,[29] the crystal structure was modified by deleting all crystal waters, assigning bond orders and adding hydrogen atoms. Protonation states of basic and acidic residues were assigned by optimization of the hydrogen-bonding net-

© WILEY-VCH
Galley Proofs

work. An Impref minimization was carried out on the final protein structure by using the OPLS2005 force field,[30] converging all heavy atoms to a root-mean-squared deviation (RMSD) of 0.30 Å of the crystallographic positions.

**Ligand conformational search**: A conformational search was carried out on each ligand by using a mixed torsional/low-mode sampling method.[31] The conformational sampling was performed by using the OPLS2005 force field,[30] an implicit water model[32] with constant dielectric, and with the van der Waals, electrostatic and hydrogen bond cutoffs set to 8, 20, and 4 Å, respectively. Extended torsion sampling was used with the maximum number of steps set to 1000. The energy window for saving structures was set to 42 kJ mol$^{-1}$ (10.04 kcal mol$^{-1}$). The resulting conformations were clustered by using an average linkage clustering technique, with representatives closest to the centroid of each cluster selected for docking.

**Docking**: The protein structure described above was used to generate a grid to be employed in the docking study. An 'enclosing box' with dimensions of 36 Å, centered on the ligand in the crystal structure, was set; nonplanar conformations of amide bonds were penalized, and the partial charge cutoff was set to 0.25. No constraints were used so as to allow ligands to freely explore the grid.

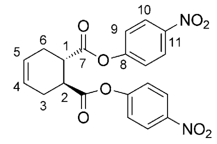
Docking studies were carried out by using the QPLD[23] workflow based on Glide.[24] This technique allows for charge polarization of the ligands induced by the protein environment, which is essential within highly charged active sites, such as DC-SIGN. QPLD improves the partial charges on the ligand atoms in Glide docking run by replacing them with charges generated from a quantum-mechanical (QM) calculation on the ligand in the field of the receptor. By using the previously described grid, along with the input structures obtained from conformational clustering, QPLD was performed. Initial charges for each ligand were generated with a semi-empirical method and were rigidly docked in Glide Extra Precision (XP) mode[33] with van der Waals radii scaled by 0.8. The QM charges were calculated by using density functional theory at the B3LYP/3-21g level of theory by using the Jaguar program.[34] The generated poses from the initial docking stage were subsequently re-docked rigidly in XP mode by using the calculated QM charges for the ligands.

Redundant binding modes, in which no calcium coordination was observed, were removed by filtering docked poses by the OH-3–Ca$^{2+}$ distance. Distances greater than 4 Å were discarded and the remaining poses were minimized, and binding energies estimated, with the Prime MM-GBSA program.[35]

**Synthesis**

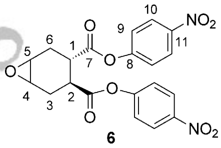
General: Dichloromethane, methanol, *N*,*N*-diisopropylethylamine (DIPEA), and triethylamine (TEA) were dried over calcium hydride; tetrahydrofuran (THF) was distilled over sodium, and *N*,*N*-dimethylacetamide (DMA) was dried over activated molecular sieves. Reactions requiring anhydrous conditions were performed under nitrogen. $^1$H and $^{13}$C spectra were recorded at 400 MHz on a Bruker AVANCE-400 instrument. Chemical shifts ($\delta$) for $^1$H and $^{13}$C spectra are expressed in ppm relative to the internal standard (CDCl$_3$: $\delta = 7.24$ ppm for $^1$H and $\delta = 77.23$ ppm for $^{13}$C; CD$_3$OD: $\delta = 3.31$ ppm for $^1$H and $\delta = 49.15$ ppm for $^{13}$C). Signals were abbreviated as s, singlet; br s, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet. Sugar signals were numbered as customary; cyclohexane protons are indicated with the letter C followed by numbers. The unusual numbering of the pseudodisaccharide derivatives was adopted to facilitate the comparison with the native disaccharide. Mass spectra were obtained with a ThermoFisherLCQapparatus (ESI ionization), or iontrap ESI Esquire 6000 from Bruker, or a Microflex apparatus (MALDI ionization) from Bruker, or Apex II ICR FTMS (ESI ionization–HRMS). Specific optical rotation values were measured at 589 nm by using a Perkin–Elmer 241 with a 1 dm cell. TLC analysis was carried out with pre-coated Merck F254 silica gel plates. Flash chromatography (FC) was carried out with Macherey–Nagel silica gel 60 (230–400 mesh). Amines **10 a–e** and **10 p** are commercially available. The synthesis of amine **10 f–o** is reported in the Supporting Information file. The synthesis of **4 h**, **11**, **13**, and **15** is described below. The synthesis of all other amides (**4 a–g**, **4 i–o**, and **12**) are reported in the Supporting Information.

**4,5-Epoxy-cyclohexan-1,2-dicarboxylic acid bis(4-nitrophenyl)ester (1*S*,2*S*) (6)**: EDC·HCl (394 mg, 2.05 mmol, 3.5 equiv) was added to a solution of the diacid **5**[10] (100 mg, 0.59 mmol, 1 equiv) in dry THF (5.8 mL) with stirring and under a nitrogen atmosphere. After 10 min, *p*-nitrophenol (245 mg, 1.76 mmol, 3 equiv) was added. The solution was stirred at room temperature for 2 h. After completion of the reaction, the solvent was evaporated under reduced pressure, the residue was taken up in Et$_2$O, washed with 1 M HCl, saturated Na$_2$CO$_3$ (3×), and water; the mixture was then dried over sodium sulphate. The solvent was evaporated under reduced pressure to yield 74 % of pure (1*S*,2*S*)-4-cyclohexen-1,2-dicarboxylic acid bis(*p*-nitrophenyl) (PNP) ester as a pale-yellow solid.



$[\alpha]_D^{20} = +129.6$ ($c = 1$ in CHCl$_3$); $^1$H NMR (400 MHz, CDCl$_3$): $\delta = 8.28$–8.22 (m, 4 H; H$_{10}$), 7.27–7.22 (m, 4 H; H$_9$), 5.83 (app d, $J = 2.8$ Hz, 2 H; H$_4$, H$_5$), 3.27–3.19 (m, 2 H; H$_1$, H$_2$), 2.78–2.68 (m, 2 H; H$_{3ps-eq}$, H$_{6ps-eq}$), 2.48–2.37 ppm (m, 2 H; H$_{3ps-ax}$, H$_{6ps-ax}$); $^{13}$C NMR (100 MHz, CDCl$_3$): $\delta = 172.8$ (C$_7$); 155.4 (C$_{11}$); 145.7 (C$_8$); 125.5 (C$_{10}$); 124.9 (C$_5$, C$_4$); 122.5 (C$_9$); 41.5 (C$_1$, C$_2$); 28.0 ppm (C$_3$, C$_6$).

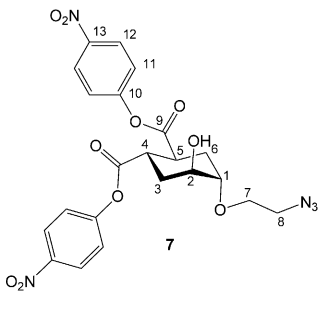
*m*CPBA (77 %, 891 mg, 3.98 mmol, 1.2 equiv) was added to a solution of the PNP ester prepared in the previous step (1367 mg, 3.32 mmol, 1 equiv) in dry CH$_2$Cl$_2$ (11 mL) under stirring. The reaction was stirred under nitrogen at room temperature for 16 h. The solvent was evaporated at reduced pressure and then the reaction mixture was diluted with EtOAc and washed with saturated NaHCO$_3$ (3×) and with water. The organic phase was dried over sodium sulphate and the solvent evaporated under reduced pressure to yield 91 % of pure product **6** as a white solid.



**6**

$[\alpha]_D^{20} = +82.2$ ($c = 1.1$ in CHCl$_3$); $^1$H NMR (400 MHz, CDCl$_3$): $\delta = 8.28$–8.22 (m, 4 H; H$_{10}$), 7.27–7.21 (m, 4 H; H$_9$), 3.43–3.36 (m, 1 H; H$_4$ or H$_5$), 3.35–3.31 (m, 1 H; H$_4$ or H$_5$), 3.32–3.24 (m, 1 H; H$_1$ or H$_2$), 3.08–3.00 (m, 1 H; H$_1$ or H$_2$), 2.79–2.71 (m, 1 H, H$_{3eq}$ or H$_{6eq}$), 2.63–2.54 (m, 1 H; H$_{3eq}$ or H$_{6eq}$), 2.42–2.33 (m, 1 H; H$_{3ax}$ or H$_{6ax}$), 2.21–2.12 ppm (m, 1 H; H$_{3ax}$ or H$_{6ax}$); $^{13}$C NMR (100 MHz, CDCl$_3$): $\delta = 172.6$, 171.5 (C$_7$); 155.4, 155.3 (C$_8$); 149.9 (C$_{11}$); 125.5 (C$_{10}$); 122.5, 122.6 (C$_9$); 51.8, 50.3 (C$_4$, C$_5$); 40.1, 38.1 (C$_1$, C$_2$); 26.8, 26.3 ppm (C$_3$, C$_6$); MS (ESI): *m*/*z* (%): for [C$_{20}$H$_{16}$N$_2$O$_9$Na]$^+$: 451.3; found: 452.0; HRMS: *m*/*z* calcd for [C$_{20}$H$_{16}$N$_2$O$_9$Na]$^+$: 451.07480; found: 451.07525

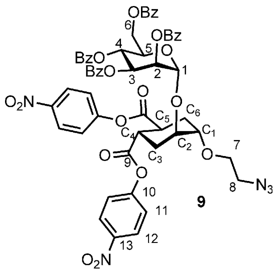
**1,2-Cyclohexanedicarboxylic acid, (1*S*,2*S*,4*S*,5*S*)-5-(2-azidoethoxy)-4-hydroxy-1,2-bis-*p*-nitrophenyl ester (7)**:■■ Epoxide **6** (50 mg, 0.12 mmol, 1 equiv) and Cu(OTf)$_2$ (4 mg, 0.01 mmol, 0.1 equiv) were added to a solution of 2-azidoethanol[11] (ca. 600 mg, 7 mmol, 58 equiv) in dichloromethane (3 mL) and the mixture stirred under nitrogen at room temperature. After completion (16 h), the solvents were removed under reduced pressure and the crude was purified by flash chromatography (hexane with gradient of ethyl acetate from 20 to 50 %) to yield 70 % of pure product **7** as a colorless wax.



**7**

$[\alpha]_D^{20} = +33.8$ (c=1.1 in $CHCl_3$); $^1H$ NMR (400 MHz, $CDCl_3$): $\delta = 8.27$–8.17 (m, 4H; $H_{12}$), 7.31–7.23 (m, 4H; $H_{11}$), 4.19–4.14 (m, 1H; $H_2$), 3.88–3.82 (m, 1H; $H_{7a}$), 3.73–3.63 (m, 2H; $H_{7b}$, $H_1$), 3.49–3.29 (m, 4H; $H_4$, $H_5$, $H_{8a,b}$), 2.40–2.10 ppm (m, 4H; $H_3$, $H_6$); $^{13}C$ NMR (100 MHz, $CDCl_3$): $\delta =$ 172.7, 172.6 ($C_9$); 155.5 ($C_{10}$); 145.8, 145.7 ($C_{13}$); 125.5 ($C_{12}$); 122.6, 122.6 ($C_{11}$); 76.5 ($C_1$); 68.7 ($C_7$); 66.3 ($C_2$); 51.1 ($C_8$); 39.4, 39.1 ($C_4$, $C_5$); 30.5, 27.0 ppm ($C_3$, $C_6$); MS (ESI): *m*/*z* calcd for $[C_{22}H_{21}N_3O_{10}Na]^+$: 512.4 $[M+Na]^+-N_2$■■; found: 512.0.

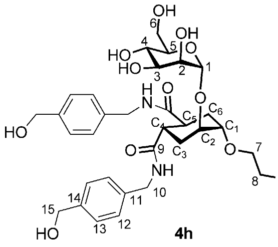
**1,2-Cyclohexanedicarboxylic acid, (1*S*,2*S*,4*S*,5*S*)-4-(2-azidoethoxy)-5-[(2,3,4,6-tetra-*O*-benzoyl-α-D-mannopyranosyl)oxy]-1,2-bis(*p*-nitrophenyl ester) (9)**: A mixture of the acceptor **7** (37 mg, 0.071 mmol, 1 equiv) and the donor **8**[12] (65 mg, 0.086 mmol, 1.2 equiv) was co-evaporated 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 with dry $CH_2Cl_2$ (1 mL). After cooling at −20 °C, TMSOTf (3 µL, 0.014 mmol, 0.2 equiv) was added to the mixture whilst stirring. The reaction was stirred at −20 °C for 1 h and was then quenched with $Et_3N$. The mixture was warmed to room temperature and filtered over a Celite pad. The filtrate was evaporated at reduced pressure and the crude product purified by flash chromatography (toluene with a gradient of ethyl acetate from 0 to 10%) to yield 85% of pure product **9** as a white foam.



**9**

$[\alpha]_D^{20} = -18.0$ (*c*=0.5 in $CHCl_3$); $^1H$ NMR (400 MHz, $CDCl_3$): $\delta = 8.30$–8.20 (m, 4H; $H_{12}$), 8.09–8.00 (m, 4H; $H_{Bz}$), 7.99–7.94 (m, 2H; $H_{Bz}$,), 7.83–7.79 (m, 3H; $H_{Bz}$), 7.63–7.48 (m, 3H; $H_{Bz}$), 7.46–7.20 (m, 12H; $H_{Bz}$, $H_{11}$), 6.07 (t, 1H; $H_4$, $J_{4-3} = J_{4-5} = 10.0$ Hz), 5.90 (dd, 1H; $H_3$, $J_{3-4} = 10.0$, $J_{3-2} =$ 3.3 Hz), 5.75 (dd, 1H; $H_2$, $J_{2-1} = 1.7$, $J_{2-3} = 3.3$ Hz), 5.34 (d, 1H; $H_1$, $J_{1-2} =$ 1.7 Hz), 4.73 (dd, 1H; $H_{6b}$, $J_{6b-5} = 2.9$, $J_{6a-6b} = 12.0$ Hz), 4.53 (dd, 1H; $H_{6a}$, $J_{6a-5} = 5.3$, $J_{6a-6b} = 12$ Hz), 4.48–4.40 (m, 1H; $H_5$), 4.23–4.18 (m, 1H; $C_2$), 3.89–3.84 (m, 1H; $C_1$), 3.74–3.67 (m, 1H; $H_{7a}$), 3.60–3.52 (m, 1H; $H_{7b}$), 3.44–3.34 (m, 2H; $C_4$, $C_5$), 3.35–3.23 (m, 2H; $H_8$), 2.52–2.41 (m, 2H; $C_{3eq}$, $C_{6eq}$), 2.25–2.09 ppm (m, 2H; $C_{3ax}$, $C_{6ax}$); $^{13}C$ NMR (100 MHz, $CDCl_3$): $\delta = 172.3$, 172.2 ($C_9$); 166.3, 165.8, 165.7 ($CO_{BZ}$); 155.5, 155.4 ($C_{10}$); 145.8 ($C_{13}$); 133.9, 133.8, 133.6, 133.5 ($CH_{BZ}$); 130.1, 130.1, 129.9, 129.9, 129.9 ($CH_{BZ}$); 129.3, 129.2, 129.1, 129.0 ($Cquat_{BZ}$); 128.9, 128.7, 128.9, 128.4 ($CH_{BZ}$); 125.5, 125.4 ($C_{12}$); 122.7, 122.6 ($C_{11}$); 96.8 ($C_1$); 74.9 ($C_{C1}$); 72.3 ($C_{C2}$); 70.9 ($C_2$); 70.1 ($C_5$); 69.9 ($C_3$); 68.7 ($C_7$); 67.4 ($C_4$); 63.5 ($C_6$); 51.0 ($C_8$); 39.4, 39.2 ($C_{C4}$,$C_{C5}$); 27.8 ($C_{C3}$); 27.0 ppm ($C_{C6}$); MS (ESI): *m*/*z* calcd for $[C_{56}H_{47}N_5O_{19}Na]^+$: 1117.0; found 1116.1; HRMS: *m*/*z* calcd for $[C_{56}H_{47}N_5O_{19}Na]^+$: 1116.27575; found: 1116.27734.

**Synthesis of *N*$^1$,*N*$^2$-bis[4-(hydroxymethylene)benzyl]amide (4h)**: The amine **10h** (38 mg, 0.275 mmol, 3 equiv) was added to a 0.1 M solution of **9** (100 mg, 0.091 mmol, 1 equiv) in dry acetonitrile (0.9 mL) whilst stirring and under a nitrogen atmosphere at room temperature. After completion (3 h; TLC, hexane/EtOAc 1:1 and 2:8) the solvent was evaporated under reduced pressure. The crude product was dissolved in dry methanol (*c*=0.1 M), under nitrogen at room temperature and a 1 M solution of sodium methoxide in MeOH (2 equiv) was added. After reaction completion (1 h), the reaction mixture was diluted with MeOH and neutralized with prewashed Amberlite IRA 120-H$^+$. The resin was filtered off and the filtrate was concentrated under reduced pressure. The crude was purified by flash chromatography ($CHCl_3$ with gradient of methanol from 0 to 20%).

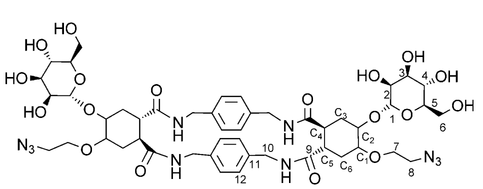


**4h**

Yield=82%; $[\alpha]_D^{20} = +12.1$ (*c*=0.81 in $CH_3OH$); $^1H$ NMR (400 MHz, $CD_3OD$): $\delta = 7.29$ (d, 4H; $H_{12}$, $J_{12-13} = 8$ Hz), 7.23 (d, 4H; $H_{13}$, $J_{13-12} =$ 8 Hz), 4.96 (d, 1H; $H_1$, $J_{1-2} = 1.6$ Hz), 4.58 (s, 4H; $H_{15a,b}$), 4.31 (s, 4H; $H_{10a,b}$), 4.08–4.03 (m, 1H; $C_2$), 3.93–3.89 (m, 1H; $H_2$), 3.89–3.84 (m, 1H; $H_{6a}$), 3.84–3.65 (m, 5H; $C_1$, $H_{6b,}$ $H_{7a,b}$, $H_{3,}$), 3.64–3.54 (m, 2H; $H_4$, $H_5$), 3.47–3.35 (m, 2H; $H_{8a,b}$), 3.02–2.85 (m, 2H; $C_4$,$C_5$), 2.06–1.86 ppm (m, 4H; $C_3$, $C_6$); $^{13}C$ NMR (100 MHz, $CD_3OD$): $\delta = 177.2$, 177.0 ($C_9$); 141.7 ($C_{14}$); 139.2 ($C_{11}$); 128.6, 128.7 ($C_{12}$); 128.3 ($C_{13}$); 100.4 ($C_1$); 76.6 ($C_3$); 75.7 ($C_5$); 72.7 ($C_{C1}$); 72.5 ($C_2$); 72.4 ( $C_{C2}$); 69.3 ($C_7$); 68.9 ($C_4$); 65.1 ($C_{15}$); 63.2 ($C_6$); 52.1 ($C_8$); 43.8 ($C_{10}$); 42.1, 41.9 ($C_{C4}$, $C_{C5}$); 29.9, 29.0 ppm ($C_{C3}$, $C_{C6}$); HRMS: *m*/*z* calcd for: $[C_{32}H_{43}N_5O_{11}Na]^+$: 696.28568; found: 696.28423.

**Synthesis of macrocycle 11**: A solution of *p*-xylylbenzylamine **10p** (3.8 mg, 0.028 mmol, 0.5 equiv) in MeCN (0.55 mL) was added dropwise under nitrogen to a flask charged with scaffold **9** (60 mg, 0.055 mmol, 1 equiv). The reaction was stirred for 5 h. TLC (dichloromethane/MeOH 9:1, dichloromethane/MeOH 9:1+1% TEA, hexane/AcOEt 6:4) indicated the presence of scaffold **9**, whereas amine **10p** was not observed. Therefore, another portion of amine **10p** (0.5 equiv) was added and the reaction mixture was stirred for an additional 16 h after which TLC analysis indicated almost no **9** and a complex mixture of products. The solvent was removed under reduced pressure and the crude was purified by flash chromatography (silica, hexane with gradient of EtOAc from 30 to 70%) to afford 18.6 mg of benzoyl-protected intermediate (TLC, hexane/EtOAc 2:8, $R_f = 0.33$); MS (ESI): *m*/*z* calcd for $[C_{104}H_{98}N_{10}O_{26}Na]^+$: 1926.9; found: 1926.6.

A solution of sodium methoxide in MeOH (1 M, 100 µL, 0.037 mmol, 4 equiv) was added to the solution of the product obtained in the previous reaction (18.6 mg, 0.0097 mmol, 0.18 equiv) in dry MeOH (0.7 mL). After 45 min, the reaction mixture was diluted with methanol and neutralized with prewashed Amberlite IRA 120-H$^+$. The resin was filtered off and the filtrate was concentrated under reduced pressure. The crude product was purified by flash chromatography ($CHCl_3$ with a gradient of MeOH from 0 to 20% with 10% water in methanol) to afford 8.7 mg of product.



**11** (only one isomer shown)

Yield=15%; $^1H$ NMR (400 MHz, $CD_3OD$): $\delta = 7.11$ (s, 8H; $H_{12}$), 4.96 (d, 2H; $H_1$, $J_{2-1} = 1.7$ Hz), 4.72–4.65 (m, 4H; $H_{10a}$), 4.05–4.02 (m, 2H; $C_2$), 3.91 (dd, 2H; $H_2$, $J_{2-1} = 1.7$, $J_{2-3} = 3.2$ Hz), 3.88–3.81 (m, 2H; $H_{6b}$), 3.80–3.62 (m, 14H; $H_{10b}$, $H_3$, $H_7$, $C_1$, $H_{6b}$), 3.61–3.52 (m, 2H: $H_{4,}$ $H_5$), 3.45–3.33 (m, 2H; $H_8$), 2.98–2.81 (m, 2H; $C_4$, $C_5$), 2.00–1.86 ppm (m, 4H; $C_3$, $C_6$); $^{13}C$ NMR (100 MHz, $CD_3OD$): $\delta = 177.1$, 177.0, 176.8, 176.7 ($C_9$); 138.9, 138.9 ($C_{11}$); 128.9 ($C_{12}$); 100.3 ($C_1$); 76.7 ($C_{C1}$); 75.7 ($C_5$); 72.7 ($C_3$); 72.6 ($C_2$); 72.5 ($C_{C2}$); 69.3 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 52.2 ($C_8$); 43.8 ($C_{10}$);

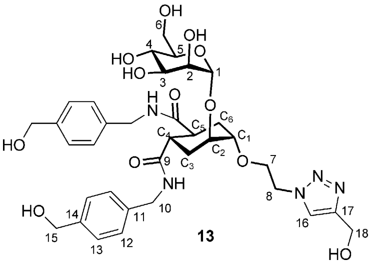
WILEY-VCH Galley Proof

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

These are not the final page numbers!

41.9, 41.8 ($C_{C4}$, $C_{C5}$); 30.0, 29.1 ppm ($C_{C3}$, $C_{C6}$); HRMS: *m*/*z* calcd for $[C_{48}H_{66}N_{10}O_{18}Na]^+$: 1093.44543; found: 1093.44341.

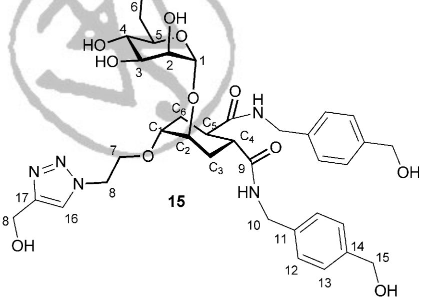
**Synthesis of 13 (triazol derivative of 4h)**: Bisamide **4h** (20 mg, 0.030 mmol, 1 equiv), propargyl alcohol (5 mg, 0.090 mmol, 3 equiv), copper(II) sulphate pentahydrate (0.7 mg, 0.003 mmol, 0.1 equiv), sodium ascorbate (2.3 mg, 0.012 mmol, 0.4 equiv), and tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA, 3.1 mg, 0.006 mmol, 0.2 equiv) were dissolved in THF/$H_2O$ (1 mL, 1:1). After reaction completion (4 h; TLC, $CHCl_3$/MeOH 8:2) the solvent was removed under reduced pressure and the resulting crude was purified by flash chromatography (silica, chloroform with gradient of MeOH from 10 to 30 %) to afford 14.8 mg of pure product.



**13**

Yield = 70 %; $[\alpha]_D^{20} = +11.7$ ($c = 0.30$ in $CH_3OH$); $^1H$ NMR (400 MHz, $CD_3OD$): $\delta = 8.00$ (s, 1 H; $H_{16}$), 7.31–7.17 (m, 8 H; $H_{12}$, $H_{13}$), 4.88 (br s, 1 H; $H_1$), 4.65 (s, 2 H; $H_{18}$), 4.59 (t, 2 H; $H_{8,}$ $J_{7-8} = 5.0$ Hz), 4.56 (s, 4 H; $H_{15}$), 4.32–4.23 (m, 4 H; $H_{10a,b}$), 3.98–3.86 (m, 3 H; $H_7$, $C_2$), 3.88–3.81 (m, 2 H; $H_2$, $H_{6a}$), 3.72–3.61 (m, 3 H; $C_1$, $H_{6b,}$ $H_{3,}$), 3.58–3.48 (m, 2 H; $H_4$, $H_5$), 2.88–2.77 (m, 1 H; $C_4$), 2.72–2.60 (m, 1 H; $C_5$), 1.94–1.64 ppm (m, 4 H; $C_3$, $C_6$); $^{13}C$ NMR (100 MHz, $CD_3OD$): $\delta = 177.0$, 176.9 ($C_9$); 141.7 ($C_{14}$); 139.2 ($C_{11}$); 128.5 ($C_{12}$); 128.3 ($C_{13}$); 125.5 ($C_{16}$); 100.4 ($C_1$); 76.1 ($C_3$); 75.7 ($C_5$); 72.7 ($C_{C1}$); 72.5 ($C_2$); 72.2 ($D_2$); 69.0 ($C_4$); 68.3 ($C_7$); 65.1 ($C_{15}$); 63.2 ($C_6$); 56.7 ($C_{18}$); 51.8 ($C_8$); 43.8 ($C_{10}$); 41.9, 41.8 ($C_{C4}$, $C_{C5}$); 29.6, 28.9 ppm ($C_{C3}$, $C_{C6}$); HRMS: *m*/*z* calcd for $[C_{35}H_{47}N_5O_{12}Na]^+$: 752.31189; found: 752.31099.

**Synthesis of 15 (triazol derivative of 12)**: Bisamide **12** (see Scheme S8 in the Supporting Information, 11 mg, 0.016 mmol, 1 equiv), propargyl alcohol (4.6 mg, 0.081 mmol, 5 equiv), copper(II) sulphate pentahydrate (0.4 mg, 0.001 mmol, 0.1 equiv), sodium ascorbate (1.3 mg, 0.006 mmol, 0.4 equiv), and TBTA (1.72 mg, 0.003 mmol, 0.2 equiv) were dissolved in THF/$H_2O$ (1 mL, 1:1). After reaction completion (12 h; TLC, $CHCl_3$/MeOH 8:2), the solvent was removed under reduced pressure and the resulting crude was purified by flash chromatography (silica, chloroform with gradient of methanol from 10 to 30 %) to afford 4.2 mg of pure product.



**15**

Yield = 35 %; $[\alpha]_D^{20} = +30.1$ ($c = 0.20$ in $CH_3OH$); $^1H$ NMR (400 MHz, $CD_3OD$): $\delta = 8.00$ (s, 1 H; $H_{16}$), 7.31–7.14 (m, 8 H; $H_{12}$, $H_{13}$), 4.84 (br s, 1 H; $H_1$), 4.88 (br s, 1 H; $H_1$), 4.65 (s, 2 H; $H_{18}$), 4.59 (t, 2 H; $H_{8,}$ $J_{7-8} = 5.0$ Hz), 4.56 (s, 2 H; $H_{15}$), 4.56 (s, 2 H; $H_{15}$), 4.29 (s, 2 H; $H_{10}$), 4.27 (s, 2 H; $H_{10}$), 3.99–3.92 (m, 1 H; $H_{7a}$), 3.90–3.84 (m, 2 H; $H_{7b}$, $C_2$), 3.84–3.80 (m, 1 H; $H_{6a}$), 3.78–3.71 (m, 2 H; $H_2$, $H_4$), 3.71–3.59 (m, 2 H; $H_5$, $H_{6b}$), 3.59–3.49 (m, 2 H; $C_1$, $H_3$), 2.94–2.84 (m, 1 H; $C_4$), 2.70–2.59 (m, 1 H; $C_5$), 1.98–1.71 ppm (m, 4 H; $C_3$, $C_6$); $^{13}C$ NMR (100 MHz, $CD_3OD$): $\delta = 177.0$ ($C_9$); 141.7, 141.6 ($C_{14}$); 139.2, 139.2 ($C_{11}$); 128.6 128.5 ($C_{12}$); 128.3, 128.3 ($C_{13}$); 125.7 ($C_{16}$); 101.8 ($C_1$); 75.8 ($C_5$); 75.2 ($D_1$); 73.9 ($C_{C2}$); 72.6, 72.5 ($C_2$, $C_4$); 69.0 ($C_3$); 68.3 ($C_7$); 65.1 ($C_{15}$); 63.3 ($C_6$); 56.7 ($C_{18}$); 51.9 ($C_8$); 43.8, 43.7 ($C_{10}$); 42.3, 41.7 ($C_{C4}$, $C_{C5}$); 31.0, 29.2 ppm ($C_{C3}$, $C_{C6}$); HRMS: *m*/*z* calcd for $[C_{35}H_{47}N_5O_{12}Na]^+$: 752.31189; found: 752.31123

## Acknowledgements

The project was supported under the EU ITN Marie-Curie program (CARMUSYS Grant number 213592).

---

[1] Y. van Kooyk, T. B. H. Geijtenbeek, *Nat. Rev. Immunol.* **2003**, *3*, 697–709.

[2] M. Anderluh, G. Jug, U. Svajger, N. Obermajer, *Curr Med Chem.* **2012**, *19*, 992–1007.

[3] T. van Montfort, D. Eggink, M. Boot, M. Tuen, C. E. Hioe, B. Berkhout, R. W. Sanders, *J. Immunol.* **2011**, *187*, 4676–4685.

[4] a) M. J. Borrok, L. L. Kiessling, *J. Am. Chem. Soc.* **2007**, *129*, 12780–12785; b) K. C. A. Garber, K. Wangkanont, E. E. Carlson, L. L. Kiessling, *Chem. Commun.* **2010**, *46*, 6747–6749; c) S. L. Mangold, L. R. Prost, L. L. Kiessling, *Chem. Sci.* **2012**, *3*, 772–777; d) T. H. Tran, R. El Baz, A. Cuconati, J. Arthos, P. Jain, Z. K. Khan, *J. Antivir. Antiretrovir.* **2011**, *3*, 49–54.

[5] a) F. Lasala, E. Arce, J. R. Otero, J. Rojo, R. Delgado, *Antimicrob. Agents Chemother.* **2003**, *47*, 3970–3972; b) J. Rojo, R. Delgado, *J. Antimicrob. Chemother.* **2004**, *54*, 579–581; c) J. J. Reina, O. S. Maldonado, G. Tabarani, F. Fieschi, J. Rojo, *Bioconjugate Chem.* **2007**, *18*, 963–969.

[6] a) 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; b) G. Timpano, G. Tabarani, M. Anderluh, D. Invernizzi, F. Vasile, D. Potenza, P. M. Nieto, J. Rojo, F. Fieschi, A. Bernardi, *Chembiochem* **2008**, *9*, 1921–1930; c) S. Sattin, A. Daghetti, M. Thepaut, A. Berzi, M. Sanchez-Navarro, G. Tabarani, J. Rojo, F. Fieschi, M. Clerici, A. Bernardi, *ACS Chem. Biol.* **2010**, *5*, 301–312; d) M. Andreini, D. Doknic, I. Sutkeviciute, J. J. Reina, J. Duan, E. Chabrol, M. Thepaut, E. Moroni, F. Doro, L. Belvisi, J. Weiser, J. Rojo, F. Fieschi, A. Bernardi, *Org. Biomol. Chem.* **2011**, *9*, 5778–5786; e) J. Luczkowiak, S. Sattin, I. Sutkeviciute, J. J. Reina, M. Sanchez-Navarro, M. Thepaut, L. Martinez-Prats, A. Daghetti, F. Fieschi, R. Delgado, A. Bernardi, J. Rojo, *Bioconjugate Chem.* **2011**, *22*, 1354–1365; f) 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; g) N. Obermajer, S. Sattin, C. Colombo, M. Bruno, U. Svajger, M. Anderluh, A. Bernardi, *Molecules Molecular Diversity* **2011**, *15*, 347–360.

[7] 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.

[8] C. G. M. Thépaut, I. Sutkeviciute, S. Sattin, R. Ribeiro-Viana, N. Varga, J. R. Eric Chabrol, J. Angulo, A. Bernardi, P. M. Nieto, F. Fieschi, **2012**, unpublished results■■any news?■■.

[9] A. Bernardi, D. Arosio, D. Dellavecchia, F. Micheli, *Tetrahedron: Asymmetry* **1999**, *10*, 3403–3407.

[10] A. Bernardi, D. Arosio, L. Manzoni, F. Micheli, A. Pasquarello, P. Seneci, *J. Org. Chem.* **2001**, *66*, 6209–6216.

[11] Y. H. Dong, X. M. Liang, H. Z. Yuan, S. H. Qi, F. H. Chen, D. Q. Wang, *Green Chem.* **2008**, *10*, 990–994.

WILEY-VCH
Galley Proof

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

These are not the final page numbers! ↗↗

[12] D. J. Lee, R. Kowalczyk, V. J. Muir, P. M. Rendle, M. A. Brimble, *Carbohydr. Res.* **2007**, *342*, 2628–2634.

[13] S. Mari, H. Posteri, G. Marcou, D. Potenza, F. Micheli, F. Canada, J. Jimenez-Barbero, A. Bernardi, *Eur. J. Org. Chem.* **2004**, 5119–5125.

[14] G. Tabarani, M. Thepaut, D. Stroebel, C. Ebel, C. Vives, P. Vachette, D. Durand, F. Fieschi, *J. Biol. Chem.* **2009**, *284*, 21229–21240.

[15] R. J. Pieters, *Org. Biomol. Chem.* **2009**, *7*, 2013–2025.

[16] K. H. Schlick, C. K. Lange, G. D. Gillispie, M. J. Cloninger, *J. Am. Chem. Soc.* **2009**, *131*, 16608–16609, and references therein; C. Sisu, A. J. Baron, H. M. Baron, Branderhorst, S. D. Connell, C. A. Weijers, R. de Vries, E. D. Hayes, A. V. Pukin, M. Gilbert, R. J. Pieters, H. Zuilhof, G. M. Visser, W. B. Turnbull, *Chembiochem* **2009**, *10*, 329–337, and references therein.

[17] C. Galustian, C. G. Park, W. Chai, M. Kiso, S. A. Bruening, Y. S. Kang, R. M. Steinman, T. Feizi, *International Immunology* **2004**, *16*, 853–866.

[18] 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.

[19] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784–1788.

[20] M. Mayer, T. L. James, *J. Am. Chem. Soc.* **2004**, *126*, 4453–4460.

[21] J. Angulo, P. M. Enriquez-Navas, P. M. Nieto, *Chem. Eur. J.* **2010**, *16*, 7803–7812.

[22] 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.

[23] A. E. Cho, V. Guallar, B. J. Berne, R. Friesner, *J. Comput. Chem.* **2005**, *26*, 915–931.

[24] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749.

[25] M. Thépaut, J. Valladeau, A. Nurisso, R. Kahn, B. Arnou, C. Vives, S. Saeland, C. Ebel, C. Monnier, C. Dezutter-Dambuyant, A. Imberty, F. Fieschi, *Biochemistry* **2009**, *48*, 2684–2698.

[26] M. Mayer, B. Meyer, *J. Am. Chem. Soc.* **2001**, *123*, 6108–6117.

[27] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546.■■author added and page changed, ok?■■

[28] R. Wagner, S. Berger, *J. Magn. Reson., Ser. A* **1996**, *123*, 119–121.

[29] Schrödinger Suite 2011 Protein Preparation Wizard: Epik version 2.2, Schrödinger, LLC, New York, NY, **2011**; Impact version 5.7, Schrödinger, LLC, New York, NY, **2011**; Prime version 3.0, Schrödinger, New York, NY, **2011**.

[30] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

[31] I. Kolossváry, W. C. Guida, *J. Am. Chem. Soc.* **1996**, *118*, 5011–5019.

[32] a) W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129; b) J. Weiser, P. S. Shenkin, W. C. Still, *J. Comput. Chem.* **1999**, *20*, 217–230.

[33] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177–6196.

[34] Jaguar, version 7.8, Schrödinger, LLC, New York, NY, **2011**.

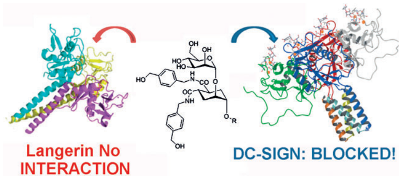
[35] Prime, version 3.0, Schrödinger, LLC, New York, NY, **2011**.

Received: August 1, 2012
Revised: December 17, 2012
Published online: ■■ ▮▮, 0000

WILEY-VCH Galley Proofs

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

**Receptor targeting**: We show here for the first time that glycomimetics based on a mannose anchor can be tuned to selectively inhibit DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin) over Langerin. Based on structural and binding studies of a mannobioside mimic previously described by us, a focused library of derivatives was designed (see figure).



**NMR Spectroscopy**

*N. Varga, I. Sutkeviciute, C. Guzzi, J. McGeagh, I. Petit-Haertlein, S. Gugliotta, J. Weiser, J. Angulo, F. Fieschi, A. Bernardi** ........ ▮▮▮–▮▮▮

**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**

WILEY-VCH Galley Proofs

**The additional data not presented in the article.**

***1) Evaluation of selectivity gain for bis-benzylamide compound series.*** The selectivity gain values achieved by modifying psDi scaffold to obtain the compounds of bis-benzylamide series were calculated ($sel.g_{psDi \rightarrow psDi}$ derivative). For comparison psTri was also included. After extrapolating the IC$_{50}$ values for langerin using eq. 6.2 on page 114, the $sel.g_{psDi \rightarrow psDi}$ derivative values were calculated as described in sub-subsection 9.1.1.4 on page 150. and the results are shown in figure 9.20.

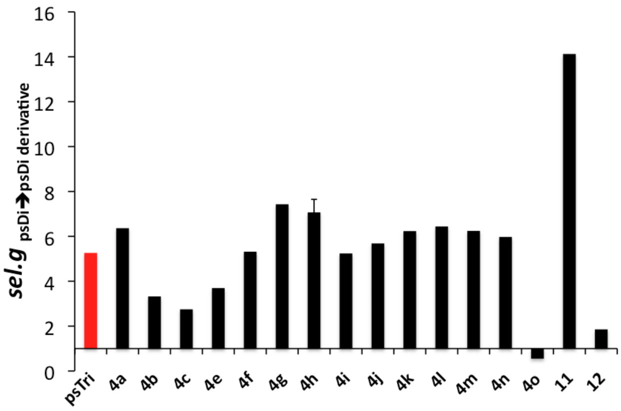


**Figure 9.20.:** The comparison of the selectivity gains of bis-benzylamides and psTri with respect to psDi as a reference compound.

The error bar for **4h** is the SD value from two measurements in this experiment.

With an exception for compound **4o**, all other compounds yielded improved selectivity to DC-SIGN versus langerin. The outstanding selectivity gain for compound **11**, a dimeric macrocyclic structure of two diastereoisomers (fig. 9.21), was observed.

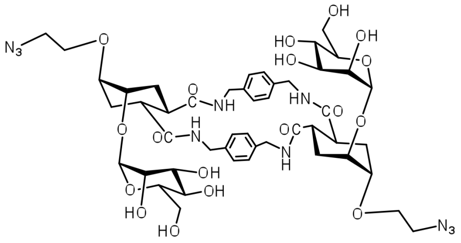


**Figure 9.21.:** The structure of compound **11**.

Taking into account a relatively very high affinity to DC-SIGN of this compound (in terms of IC$_{50}$ , 31 µM, which is even better than for psTri) that might suggest a psTri-like ability to bridge DC-SIGN tetramers, such high selectivity gain could indicate that **11** is able to bridge only DC-SIGN tetramers and not langerin trimers. Whereas a comparison with the result obtained for psTri may suggest that psTri bridges both DC-SIGN and langerin, but in addition has intrinsically higher affinity to DC-SIGN than langerin (thereof the selectivity gain, which is very similar to the gain for bis-benzylamide compound series). It is also possible that **11** bridges langerin, but in that case it has much higher intrinsic affinity to DC-SIGN than langerin.

One of the highest selectivity gains (apart from **11**) belonged to compound **4h**.

***2) Efforts to resolve the overlap of the key protons in 4h 1D $^1$H NMR spectrum.*** The compound **4h** exhibited an overlap of the key protons, shown in figure 9.22, necessary for the comparison of its binding to DC-SIGN properties with those of psDi, i.e. the protons that had highest initial slope values in STD build-up curves.

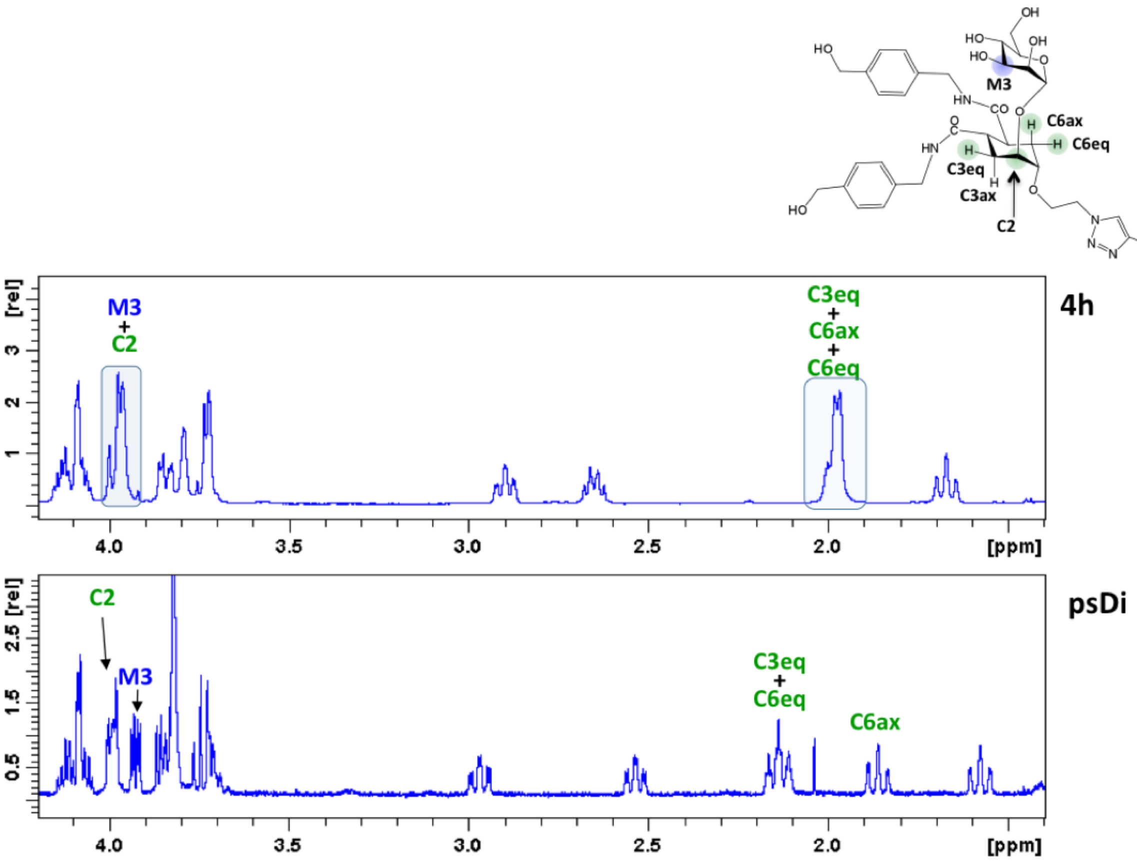


**Figure 9.22.:** 1D $^1$H NMR spectra of **4h** and psDi comparing the signals of the key protons in the two molecules.

The protons are indicated on **4h** structure on top.

It was attempted to resolve these overlaps (mainly the C3eq-C6ax-C6eq overlap since C6ax in psDi had the highest initial slope value) by lowering the temperature of the experiment from 25°C to 10°C and by adding DMSO to the sample. I was involved in these experiments during the long-term secondment in Seville.

Lowering the temperature not only didn't improve the resolution of C3eq-C6ax-C6eq overlap, but also turned to be even more disadvantageous because at 10°C anomeric proton of mannose (M1) fused with water signal and mannose protons M6 and C1 were worse resolved (fig. 9.23).

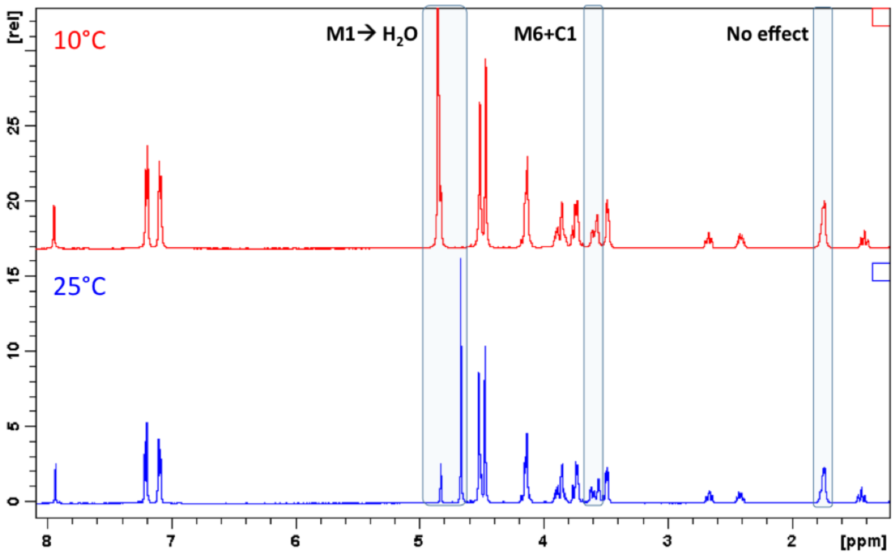
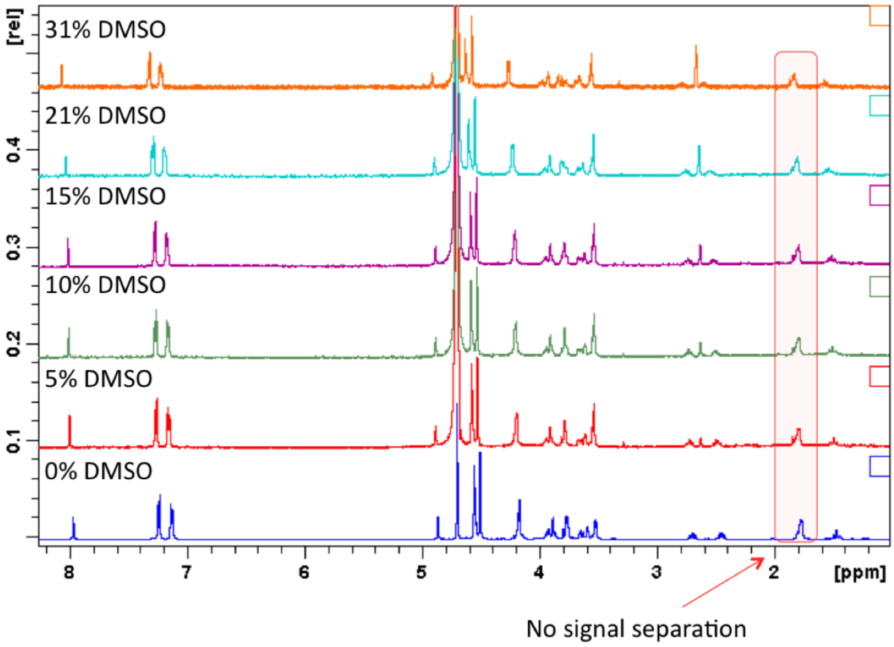


**Figure 9.23.:** 1D $^1$H NMR spectra of **4h** at 10°C and 25°C.

The changes of proton signals are indicated in the figure by blue the frames: at 10°C anomeric proton of mannose (M1) fused with water signal, mannose protons M6 and C1 were worse resolved and no change for C3eq-C6ax-C6eq overlap was observed.

Likewise, the second approach of adding DMSO up to 31% did not give any favorable results – there were virtually no changes to the spectra (fig. 9.24).



No signal separation

**Figure 9.24.:** 1D 1H NMR spectra of **4h** in presence of different DMSO concentrations, at 25°C.

Hence STD experiments were performed at 25°C with no DMSO in the sample, and consequently **4h** binding epitope was qualitatively compared with that of psDi as described in the article and the following supplementary information.

### 9.1.4.6. *The second type of psDi optimization to achieve better selectivity to DC-SIGN*

The new series of psDi-based compounds were generated with a purpose to reduce their recognition by langerin and thus improve the selectivity to DC-SIGN. The rationale of this series design was the structural knowledge about the ability of langerin CRD to recognize galactose residues sulphated at position 6 (fig. 2.19 on page 78B), where the negatively charged sulphate group is in a favorable position to make electrostatic contacts with positively charged side chain of the protein (Lys313, see DC-SIGN and langerin binding site comparison in fig. 2.23 on page 81). Therefore, different substituents at position 6 of non-reducing mannose of psDi were introduced and the corresponding structures are shown in figure 9.25. These studies were performed by master-1 student Vanessa Porkolab under my supervision.

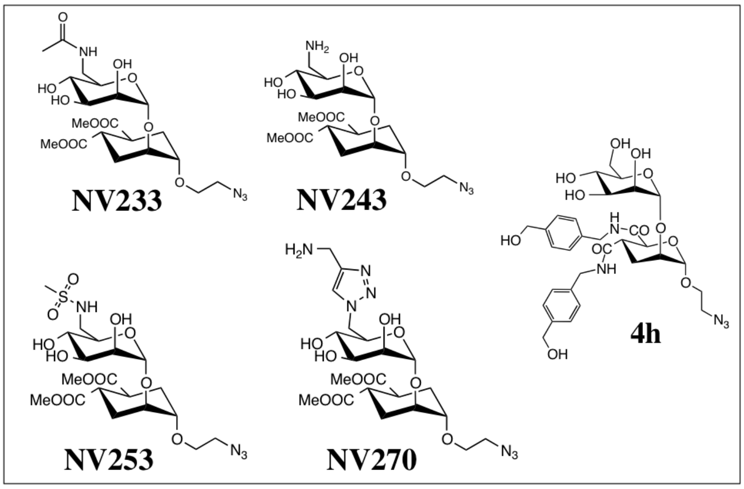


**Figure 9.25.:** The structures of psDi derivatives with different substituents at position 6 of non-reducing mannose compared to **4h** structure.

The synthesis of these compounds was carried out in the team of Pr. A. Bernardi by Norbert Varga. In these series the compound with the best inhibitory capacity towards DC-SIGN was NV243, although previously selected new lead **4h** remained the best among all the tested compounds (fig. 9.26 left panel).

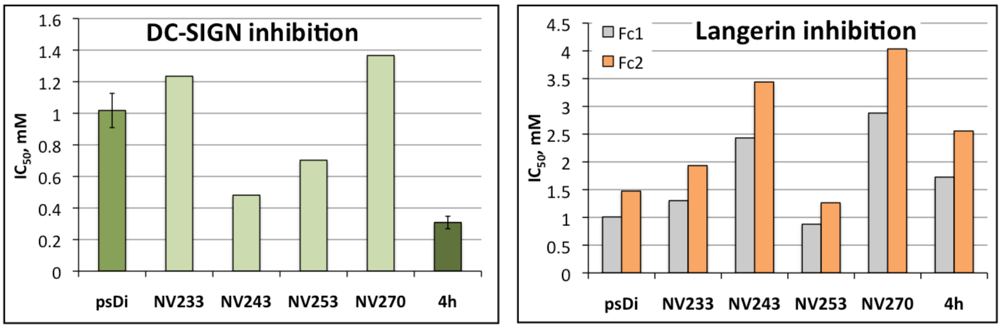


**Figure 9.26.:** The activities of new psDi derivatives to inhibit DC-SIGN and langerin.

DC-SIGN inhibition was performed using surface with 1900 RU of ManBSA immobilized; inhibition of langerin binding to reference (Fc1, activated/deactivated CM-dextran) and ManBSA/dextran (Fc2, ManBSA immobilized to 1600 RU) surfaces was measured.

The same compounds were tested for their capability to inhibit langerin ECD binding to ManBSA/dextran surface. Since langerin exhibits measurable affinity for dextran surface, the $IC_{50}$ values were determined for the inhibition of both, binding to reference (activated/deactivated CM-dextran) and ManBSA/dextran surfaces (fig. 9.26 right panel). As expected, the absolute $IC_{50}$ values were different for these two surfaces, but the general pattern of the activities of the compounds was the same. Acetyl (NV233) and methyl sulfone (NV253) groups at position 6 did not change compounds' relative affinity to langerin as compared to psDi. However, NV243 and NV270 with amino and triazole groups, respectively, appeared to have lower activity to inhibit lagerin.

To have a more quantitative evaluation of the selectivity of these compounds, the *sel.g* values were calculated as described in sub-subsection 9.1.1.4 on page 150. and compared (fig. 9.27).

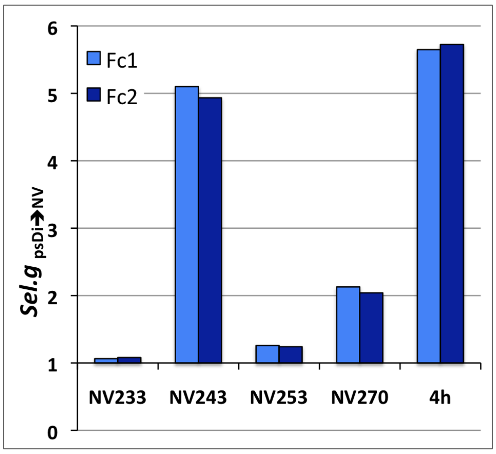


**Figure 9.27.:** The comparison of the selectivity gains with respect to psDi as a reference compound. Fc1 and Fc2 surfaces are the same as described in fig. 9.26.

As can be seen in figure 9.27, the psDi modifications leading to compounds NV233 and NV253 did not yield selectivity gain to DC-SIGN, and very slight *sel.g* was observed for compound NV270. The highest improvement was reached for compound NV243, which possesses $NH_2$ group at position 6 of non-reducing mannose, which confirms the assumption that adding there a positive charge would impair compound's interaction with langerin. The previously chosen lead **4h** had the best selectivity in the group of the tested compounds.

These results suggest that the combination of compounds **4h** and NV243 might lead to even more stronger and more selective inhibitor of DC-SIGN than the current lead **4h**. The synthesis of this compound, combining **4h** and NV243 features, is a direct perspective of this collaborative work.

## 9.2. The development of multivalent glycomimetic DC-SIGN antagonists

### 9.2.1. *The structures of all tested multivalent compounds*

The first attempts to design multivalent DC-SIGN inhibitors were made before my PhD project by our group in collaboration with Dr. J. Rojo group. The second (G2) and third generation (G3) Boltorn type dendrimers functionalized with D-mannose residues were generated and assessed for the capacity to inhibit DC-SIGN binding to gp120 surface [158]. The resulting activities of these dendrimers were encouraging, thus G3 Boltorn type dendrimers were functionalized with the selected mannose-based glycomimics psDi and psTri (fig. 9.15 on page 191). The evaluation of these dendrimers to inhibit DC-SIGN in SPR competiton assay as well as DC-SIGN-mediated Ebola infection assay is described in article **n°3** (see sub-subsection 9.1.4.2 on page 191).

The strategies of multivalent presentation of DC-SIGN ligands were further explored by the group of Dr. J. Rojo, and various glycodendrimers were generated by a PhD student Renato Ribeiro-Viana using an efficient convergent strategy, which employs a click chemistry reaction based on the Cu(I) catalyzed azide—alkyne cycloaddition (CuAAC) (fig. 9.28) [173, 174, 175], a very versatile and popular reaction allowing preparation of high diversity of compounds (small molecules, polymers, hydrogels, etc.) and bioconjugates.

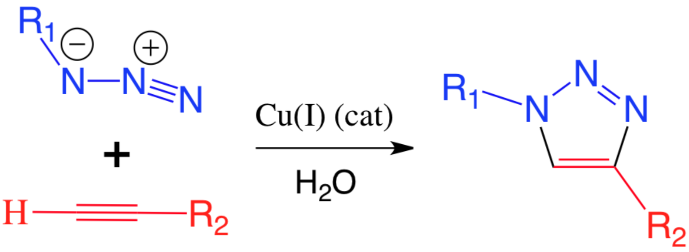


**Figure 9.28.:** The principal scheme of CuAAC reaction.

(Adapted from [176])

These dendrimers, shown in figure 9.29, allowed the ligand presentation at different valences as well as different spatial conformations.

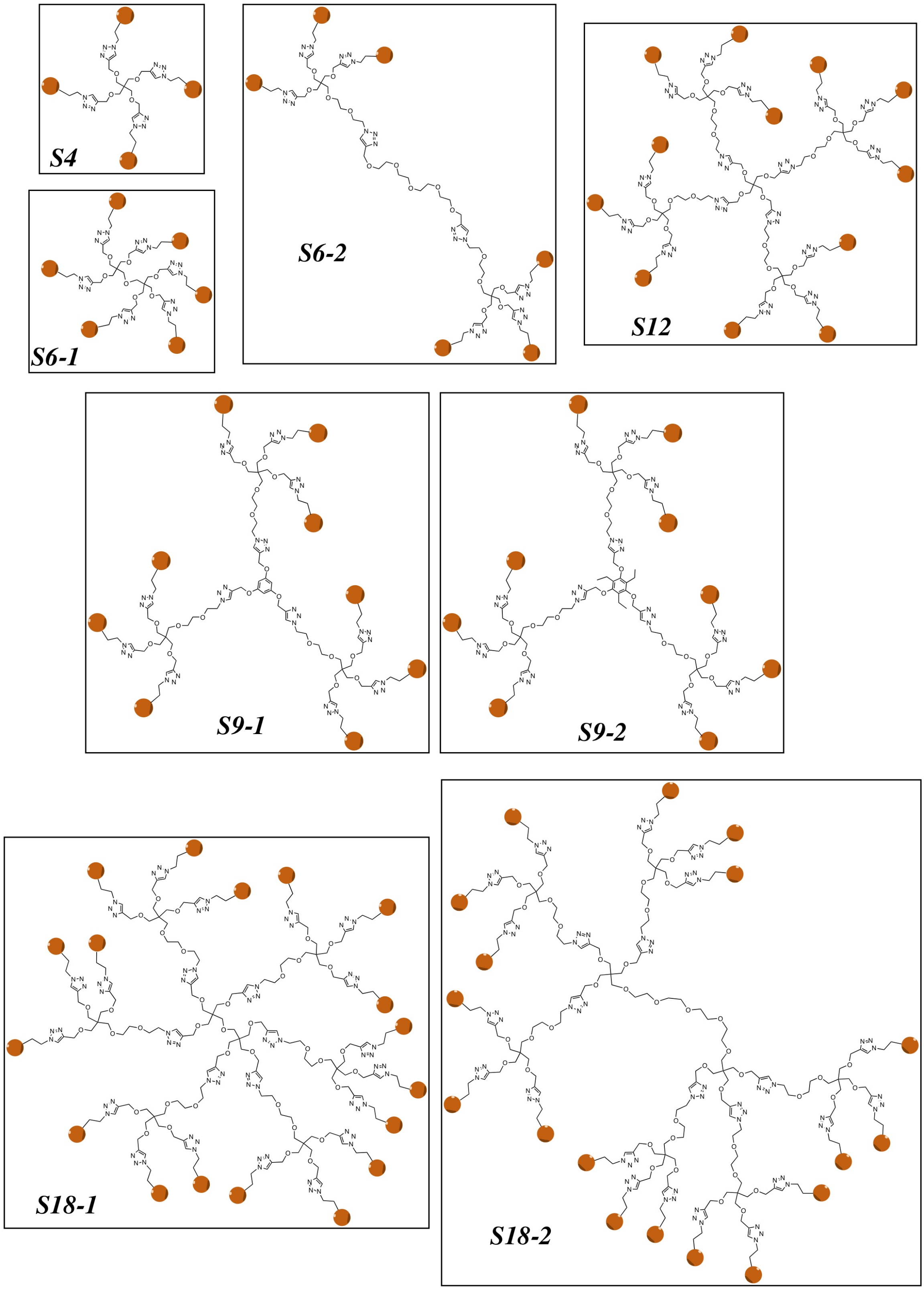


**Figure 9.29.:** The structures of different multivalent flexible scaffolds generated by CuAAC reactions.

The scaffolds are named according to their valence (e.g. ***S4*** means 4-valent Scaffold) and additional number marks the alternative scaffold with the same valence. The orange spheres represent different ligands conjugated to the scaffolds, as shown in table 9.2.

Whereas the dendrimers shown in figure 9.29 have flexible backbones, the multivalent platform of a different type with a rigid (“rod-like”) backbone was also synthesized (fig. 9.30, ***R3***) by Norbert Varga. The ***R3*** scaffold was generated in order to obtain the presentation of conjugated ligands at fixed distance and in this way to target the compound for binding to vicinal CRDs within the same DC-SIGN tetramer. The other two compounds, ***R1*** and ***R2***, are the control compounds to test the intermediate stages of ***R3*** development.

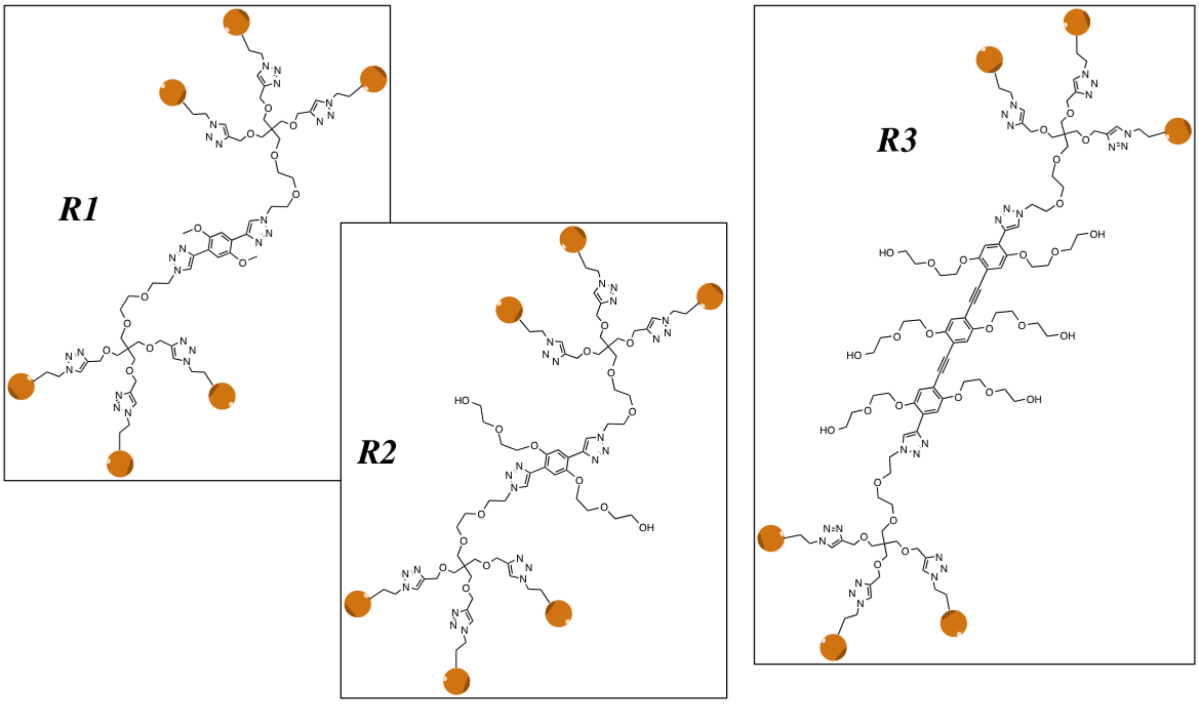


**Figure 9.30.:** The structure of rod-like scaffold and the control compounds ***R1*** and ***R2***.

The predicted distance between the two branches of the ligands in scaffold ***R3*** corresponded to a 40 Å spacing to reach the binding sites within a single DC-SIGN tetramer, as suggested by molecular modeling performed in the group of Pr. Anna Bernardi (fig. 2.11 on page 70).

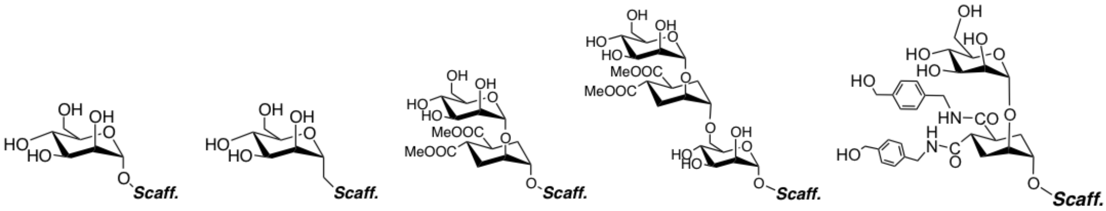
Different mannose-based ligands, including the new lead compound **4h**, were conjugated to these newly generated multivalent scaffolds (fig. 9.29 and 9.30). Table 9.2 lists all mannose-based compounds that were attached to the dendrimers and the corresponding multivalent compound names used further in this thesis.

**Table 9.2.:** Mannose-based ligands conjugated to the multivalent scaffolds and the corresponding names of the compounds.

| Ligand | Scaffold | Valence | Compound name |
|---|---|---|---|
| **D-mannose** | ***S4*** | *4* | ***S4-Oman*** |
| | ***S6-1*** | *6* | ***S6-1-Oman*** |
| | ***S9-1*** | *9* | ***S9-1-Oman*** |
| | ***S9-2*** | *9* | ***S9-2-Oman*** |
| | ***S12*** | *12* | ***S12-Oman*** |
| | ***S18-1*** | *18* | ***S18-1-Oman*** |
| | ***S18-2*** | *18* | ***S18-2-Oman*** |
| **D-mannose C-glycoside** | ***S9-1*** | *9* | ***S9-1-Cman*** |
| | ***S12*** | *12* | ***S12-Cman*** |
| **psDi** | ***S4*** | *4* | ***S4-psDi*** |
| | ***S6-1*** | *6* | ***S6-1- psDi*** |
| | ***S6-2*** | *6* | ***S6-2- psDi*** |
| | ***S9-1*** | *9* | ***S9-1- psDi*** |
| **psTri** | ***S4*** | *4* | ***S4-psTri*** |
| | ***S6-1*** | *6* | ***S6-1- psTri*** |
| **4h** | ***S4*** | *4* | ***S4-4h*** |
| | ***S6-1*** | *6* | ***S6-1-4h*** |
| | ***S9-1*** | *9* | ***S9-1-4h*** |
| | ***R1*** | *6* | ***R1*** |
| | ***R2*** | *6* | ***R2*** |
| | ***R3*** | *6* | ***R3*** |

*Scaff.* stands for "scaffold"

*Ligand structures:*



**D-mannose** **D-mannose C-glycoside** **psDi** **psTri** **4h**

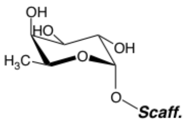
On the side of polyvalent fucose-based glycomimics development, several multivalent compounds were also synthesized (table 9.3). The compounds listed in table 9.3 were synthesized by PhD students Benedetta Bertolotti (L-fucose C-glycosides), Renato Ribeiro-Viana (***S12-Ofuc***, ***S18-1-Ofuc***), Daniela Doknic (***S4-fucβAla***, ***S4-10b-azide1***) and a post-doctoral researcher Martino Ambrosini (***S4-Ofuc***, ***S6-1-Ofuc*** compounds).

**Table 9.3.:** The list of fucose-based ligands conjugated to multivalent scaffolds and the corresponding names of the compounds.

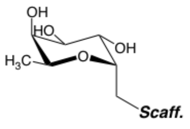
| Ligand | Scaffold | Valence | Compound name |
|---|---|---|---|
| **L-fucose** | ***S4*** | *4* | ***S4-Ofuc*** |
| | ***S6-1*** | *6* | ***S6-1-Ofuc*** |
| | ***S12*** | *12* | ***S12-Ofuc*** |
| | ***S18-1*** | *18* | ***S18-1-Ofuc*** |
| **L-fucose C-glycoside** | ***S4*** | *4* | ***S4-Cfuc*** |
| | ***S6-1*** | *6* | ***S6-1-Cfuc*** |
| | ***S9-1*** | *9* | ***S9-1-Cfuc*** |
| | ***S12-1*** | *12* | ***S12-1-Cfuc*** |
| **L-fucose β-Ala+** | ***S4*** | *4* | ***S4-fucβAla*** |
| **10b-azide1** | ***S4*** | *4* | ***S4-10b-azide1*** |

*Scaff.* stands for "scaffold"

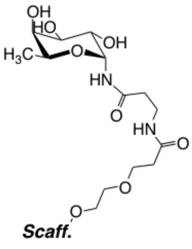
*Ligand structures:*



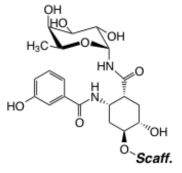
**L-fucose**



**L-fucose C-glycoside**



**L-fucose β-alanine+**



**10b-azide1**

All of these compounds were analyzed in SPR competition assay, and some of them were tested for binding to DC-SIGN in SPR direct interaction assay.

## 9.2.2. *The results of multivalent compound analysis by SPR competition assay*

The multivalent compounds were tested in SPR competition assay in the same manner as described earlier for monovalent compounds (subsection 9.1.1 on page 147). The below presented results are divided to three parts according to compound, i.e. mannose or fucose-based, and scaffold type, i.e. flexible or rod-like.

### 9.2.2.1. *The evaluation of mannose-based compounds tethered to flexible multivalent scaffolds*

The SPR competition assay results for the most of multivalent compounds on flexible scaffolds showed a gradual increase of their activity as a function of the scaffold valence increment (fig. 9.31). The only exception was psTri as monovalent and tetravalent presentation yielded same activity, the reason for this is previously described ability of "monovalent" psTri to bridge DC-SIGN tetramers.

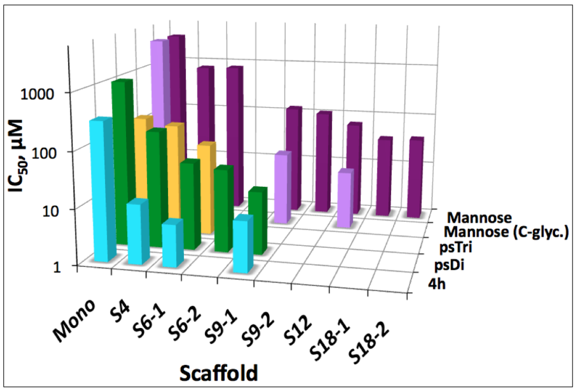


**Figure 9.31.:** DC-SIGN inhibition activities of multivalent manoose-based compounds on flexible scaffolds.

"*Mono*" stands for monovalent and "Mannose (C-glyc.)" means C-glycoside of mannose. IC$_{50}$ values for *Mono* compounds are from table 9.1 on page 149, except mannose C-glycoside, for which the value is the same as in fig. 9.7 on page 154). Compounds with **4h** on all scaffold were assayed in presence of 4% DMSO.

Comparing the activities of the compounds with different ligands but same scaffolds, it is obvious that tethering selected lead **4h** on each of the tested scaffold gives the best activity in the group. It is also noticeable that **4h** activity dramatically increases switching from monovalent to tetravalent compound and a slight improvement is still achieved when valence increases to 6, however, the nonavalent presentation gives no further improvement. It is possible that the spatial presentation of the ligands on a nonavalent scaffold does not favor the accessibility of individual ligands to DC-SIGN. The other, more plausible reason is the overestimation of the concentration of this compound due to the lower solubility (the compounds are sent lyophilized, and it was not possible to fully solubilize this particular compound).

Significantly higher activity of mannose C-glycosides as compared to mannose on the same scaffolds, despite that the corresponding monovalent ligands have the same activities, is a strange outcome with no rational explanation. All of the obtained IC$_{50}$ values are listed in table 9.4.

**Table 9.4.:** The IC$_{50}$ ± SD values (µM) obtained for multivalent mannose-based compounds on flexible scaffolds.

| | | **Ligand** | | | | |
|---|---|---|---|---|---|---|
| | | **Mannose** | **Mannose (C-glyc.)** | **psDi** | **psTri** | **4h** |
| **Scaffold** | ***Mono**** | 3292 ± 337 | 3324 ± 74 | 1018 ± 109 | 145 ± 83 | 308 ± 40 |
| | ***S4*** | 767 ± 20 | - | 136 ± 23 | 112 | 12 ± 3 |
| | ***S6-1*** | 800 | - | 39 | 51 | 5.7 ± 1.6 |
| | ***S6-2*** | - | - | 32 | - | - |
| | ***S9-1*** | 128 | 24 ± 0.2 | 14 | - | 8.2 ± 0.2 |
| | ***S9-2*** | 107 ± 10 | - | - | - | - |
| | ***S12*** | 67 | 12 ± 0.1 | - | - | - |
| | ***S18-1*** | 36 | - | - | - | - |
| | ***S18-2*** | 38 | - | - | - | - |

*values for monovalent ligands, except mannose C-glycoside (value same as in fig. 9.7 on page 154), were taken from table 9.1 on page 149. For other compounds SD values obtained due to the analysis using two or three different flow cells with ManBSA, and some of the compounds were analyzed 2 or 3 times.

To better visualize the effect of compound clustering relative to the activity of monovalent compound, the activity/affinity improvement factor β was calculated using the following equation:

$$\beta = \frac{IC_{50 monovalent}}{valence \cdot IC_{50 multivalent}} \tag{9.6}$$

The values of $\beta > 1$ mean the improvement of the multivalent compound activity with respect to monovalent ligand, and $\beta < 1$ – the decline of the activity. Figure 9.32 represents the calculated β factors for the compounds, except mannose C-glycosides, which were likely overestimated.

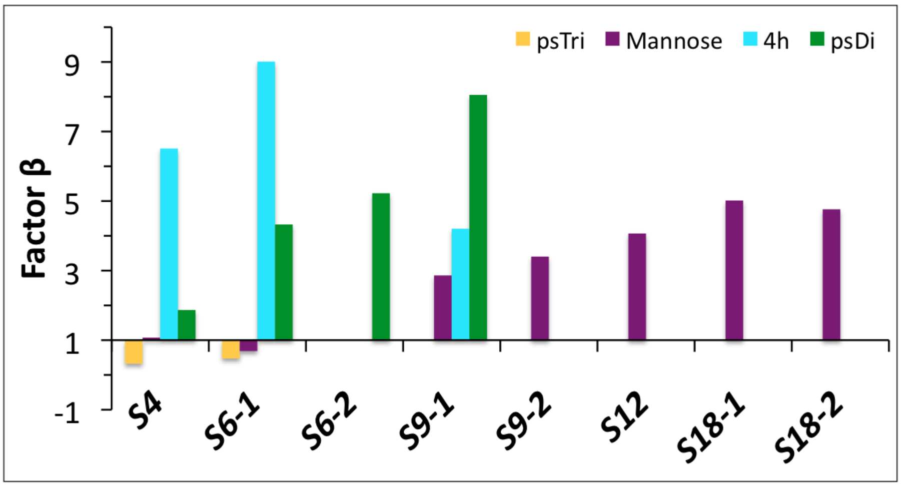


**Figure 9.32.:** DC-SIGN inhibition activity improvement factors for the multivalent compounds on flexible scaffolds with respect to corresponding monovalent ligands.

Tethering mannose on tetravalent and hexavalent scaffolds did not yield any activity improvement, but mannose on the scaffolds with the higher valence show the gradual improvement β. Such gradual improvement effect is even more profound for psDi. The best improvement of factor β was obtained for **4h** with hexavalent scaffold. The overestimation of the concentration of the nonavalent compound is most likely the reason why no further increase of β factor is observed for this compound.

However, psTri clustering on both tetra- and hexavalent scaffolds resulted in the decrease of compound's activity relative to the "monovalent" ligand because of its unique property to cluster DC-SIGN tetramers.

It must be kept in mind that in the SPR competition assay the soluble tetramers of DC-SIGN are used, therefore the $\beta > 1$ values most likely show a combined effect: 1) concentrating the monovalent ligands on the scaffold mimics high concentration of ligand; 2) possibly increased intrinsic affinity of the multivalent compound to DC-SIGN tetramer due to binding to several CRDs within the same tetramer; 3) the ability of multivalent compounds to cluster DC-SIGN tetramers [158]. Considering the small size of ***S4*** and ***S6-1*** scaffolds from the proposed affinity improvement mechanisms 2) and 3), only the third case may be expected.

#### 9.2.2.2. *The evaluation of fucose-based compounds tethered to flexible multivalent scaffolds*

The multivalent forms of L-fucose O- and C-glycosides, and a new fucose-based lead compound **10b-azide1** (table 9.3) were analyzed in SPR competition assay. In addition to these compounds, in order to evaluate the effect of the linker in the multivalent compound, the L-fucose with a linker of β-alanine with or without additional ethylenglycol (fig. 9.33) were also tested and appeared to have no effect when comparing both monovalent and tetravalent derivatives of L-fucose (fig. 9.34 and table 9.5).

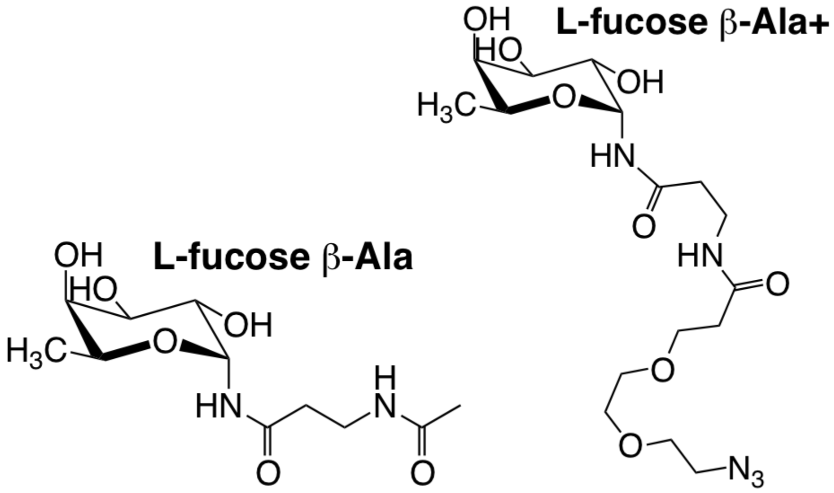


**Figure 9.33.:** The structures of L-fucose with β-alanine/ethylene glycol linkers.

As in case of mannose-based multivalent compounds, the results for multivalent fucose-based compounds showed generally the same tendency of activity increase as a function of the valence (fig. 9.34A and table 9.5).

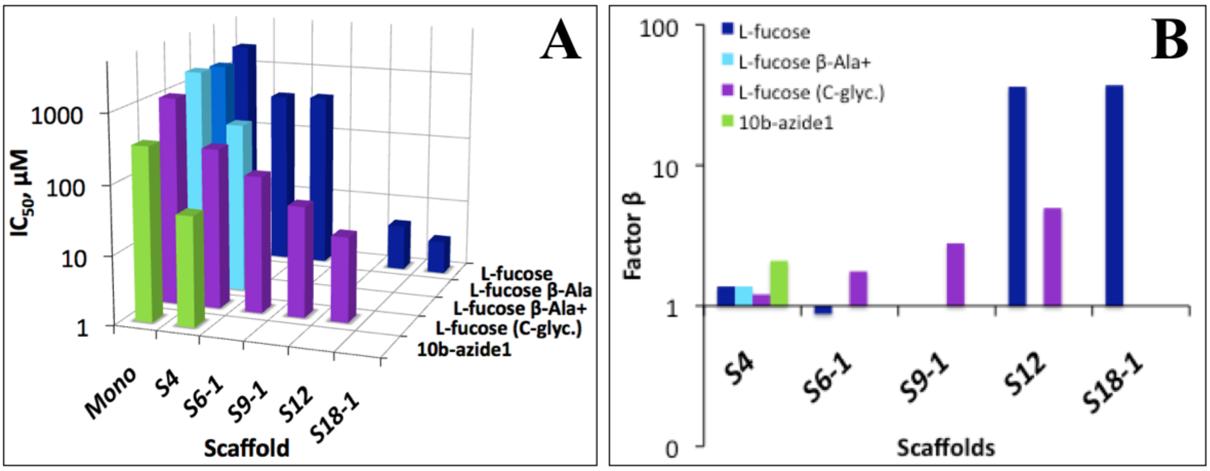


**Figure 9.34.:** DC-SIGN inhibition activities of multivalent fucose-based compounds on flexible scaffolds.

**A**, The $IC_{50}$ values. **B**, The inhibition activity improvement factors β. "L-fucose (C-glyc.)" means C-glycoside of L-fucose. $IC_{50}$ values for *Mono* compounds are from table 9.1 on page 149.

**Table 9.5.:** The IC$_{50}$ ± SD values (µM) obtained for multivalent fucose-based compounds on flexible scaffolds.

| | | **Ligand** | | | | |
|---|---|---|---|---|---|---|
| | | **L-fucose** | **L-fucose β-Ala** | **L-fucose β-Ala+** | **L-fucose (C-glyc.)** | **10b-azide1** |
| **Scaffold** | ***Mono**** | 2061 ± 182 | 1445 ± 52 | 1682 ± 61 | 1020 ± 225 | 328 ± 43 |
| | ***S4*** | 375 ± 3.2 | - | 306 ± 19 | 212 ± 4 | 39 ± 1.3 |
| | ***S6-1*** | 392 ± 1.3 | - | - | 97 ± 0.3 | - |
| | ***S9-1*** | - | - | - | 41 ± 0.6 | - |
| | ***S12*** | 4.7 ± 0.04 | - | - | 17 ± 0.1 | - |
| | ***S18-1*** | 3.1 ± 0.1 | - | - | - | - |

*values for monovalent L-fucose and **10b-azide** were taken from table 9.1 on page 149; for monovalent L-fucose C-glycoside value is the same as in fig. 9.7 on page 154. For other compounds SD values obtained due to the analysis using two or three different flow cells with ManBSA, and some of the compounds were analyzed 2 or 3 times.

The relationship between valence and compound activity is best illustrated for L-fucose C-glycoside derivatives: a proportional increase of both affinity (in terms of IC$_{50}$ value, fig. 9.34A) and β factor (fig. 9.34B) is observed. Strangely, L-fucose O-glycosides did not show such a proportional evolution: while the 4-valent and 6-valent compounds had virtually the same affinity and no activity improvement, these terms dramatically increased for the 12-valent and 18-valent forms. The latter multivalent forms had the same, relatively high, affinity and β factors, probably suggesting the maximal valence for L-fucose derivative achieved, with which the affinity is maximal under the used experimental conditions. The experimental errors cannot be excluded as well (compounds of only a single batch were tested using two ManBSA surfaces in the same run).

Additionally, comparing L-fucose O- and C- glycosides, their activity was alternating: the monovalent C-fucose has twice better affinity than O-fucose, which becomes the same when ligands are presented on 4-valent scaffold. Comparing hexavalent forms, C-fucose has again better affinity (4 times) than O-fucose, whereas 12-valent O-fucose becomes regains the better affinity than C-fucose (3 times). It is not clear what was the origin of such an alternation.

The best compound in these series, in terms of both IC$_{50}$ value and the activity improvement factor β (fig. 9.34B), appeared to be the selected lead compound **10b-azide1**. Considering that 10b-azide1 was tethered only to a tetravalent scaffold, even better affinity and factor β values can be expected if the compound was tethered to the higher valence scaffolds.

#### 9.2.2.3. *The evaluation of mannose-based compounds tethered to rod-like multivalent scaffolds*

The third type of multivalent presentation, i.e. the rod-like hexavalent scaffold, which was designed in an attempt to target binding to two CRDs within the same DC-SIGN tetramer, was used to tether the mannose-based lead compound **4h**. Figure 9.35 and table 9.6 show the results obtained for the rod-like scaffold presenting **4h**, compared to the monovalent ligand and other related hexavalent compounds with **4h**, i.e. ***R1*** and ***R2***, and the previously presented ***S6-1*** scaffold with **4h**.

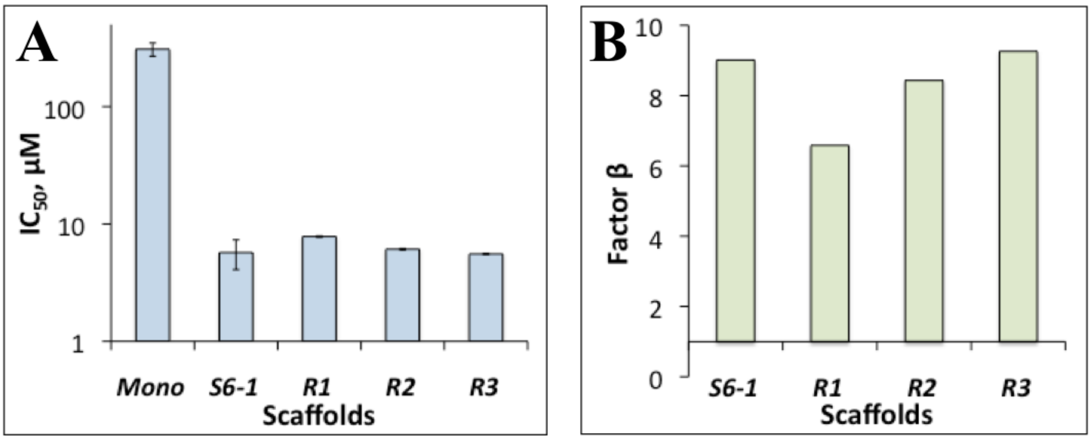


**Figure 9.35.:** DC-SIGN inhibition activities of different hexavalent scaffolds with **4h**. **A**, The $IC_{50}$ values. **B**, The inhibition activity improvement factors β. $IC_{50}$ value for monovalent **4h** is from table 9.1 on page 149, and for ***S6-1*** with **4h** from table 9.4. Compounds ***R1*** and ***R2*** were assayed in presence of 4% DMSO, while ***R3*** was assayed in presence of 4.2% DMSO. The error bars show SD values from 3 measurements.

**Table 9.6.:** The $IC_{50}$ obtained for rod-like and other hexavalent scaffolds.

| Scaffold | $IC_{50}$ ± SD, µM |
|---|---|
| ***R1*** | 7.81 ± 0.09 |
| ***R2*** | 6.10 ± 0.09 |
| ***R3*** | 5.55 ± 0.04 |

SD values calculated from three measurement in the same test performed using 3 different ManBSA surfaces on sensor chip flow cells 2, 3, and 4.

It turned out that all of the hexavalent **4h** variants have the same the affinity (fig. 9.35A) as well as the same β factors (fig. 9.35B). However, the latter three compounds, and especially ***R***3, had significant solubility problems, which led to the overestimation of their concentration. Hence, the real affinity of these compounds might be stronger.

It is not possible to answer the question, if ***R***3 is able to bind several CRDs within the same DC-SIGN tetramer from the SPR competition assay results, because the higher observed affinity may result from two phenomena (as already pointed above): the intrinsically increased affinity of the compound due to binding to several CRDs within the same tetramer, and due to increased capability to cluster DC-SIGN tetramers. It is not possible to discriminate, which of the two phenomena occurs, or do both of them occur simultaneously.

### 9.2.3. *The results of multivalent compound analysis by SPR direct interaction assay*

The SPR competition assay is suitable for efficient compound activity screening, but does not reflect the real situation since DC-SIGN tetramers are in solution and not surface-attached. This point was clearly illustrated by psTri interaction studies with DC-SIGN, where its interaction in solution induced bridging between the lectin tetramers that is not awaited if the lectin receptor is membrane-bound. Therefore the SPR direct interaction assay was designed, where DC-SIGN tetramers are immobilized to the sensor chip surface.

Additionally to screening by SPR competition assay, some of the multivalent compounds were tested in SPR direct interaction assay in three separate experiments:

1. The first tested compounds were Boltorn type dendrimers and dendrons with psDi and psTri (fig. 9.36).
2. In the second experiment mannose-based compounds from table 9.2 were analyzed, except mannose C-glycosides, ***S9-1-4h***, and ***R1***, ***R2***, ***R3*** compounds.
3. Compounds ***S4-4h***, ***S6-1-4h***, ***S9-1-4h***, and ***R1***, ***R2***, ***R3*** compounds from table 9.2 together with fucose based compounds ***S4-fucβAla*** and ***S4-10b-azide1*** (table 9.3) were tested in the third direct interaction SPR experiment.

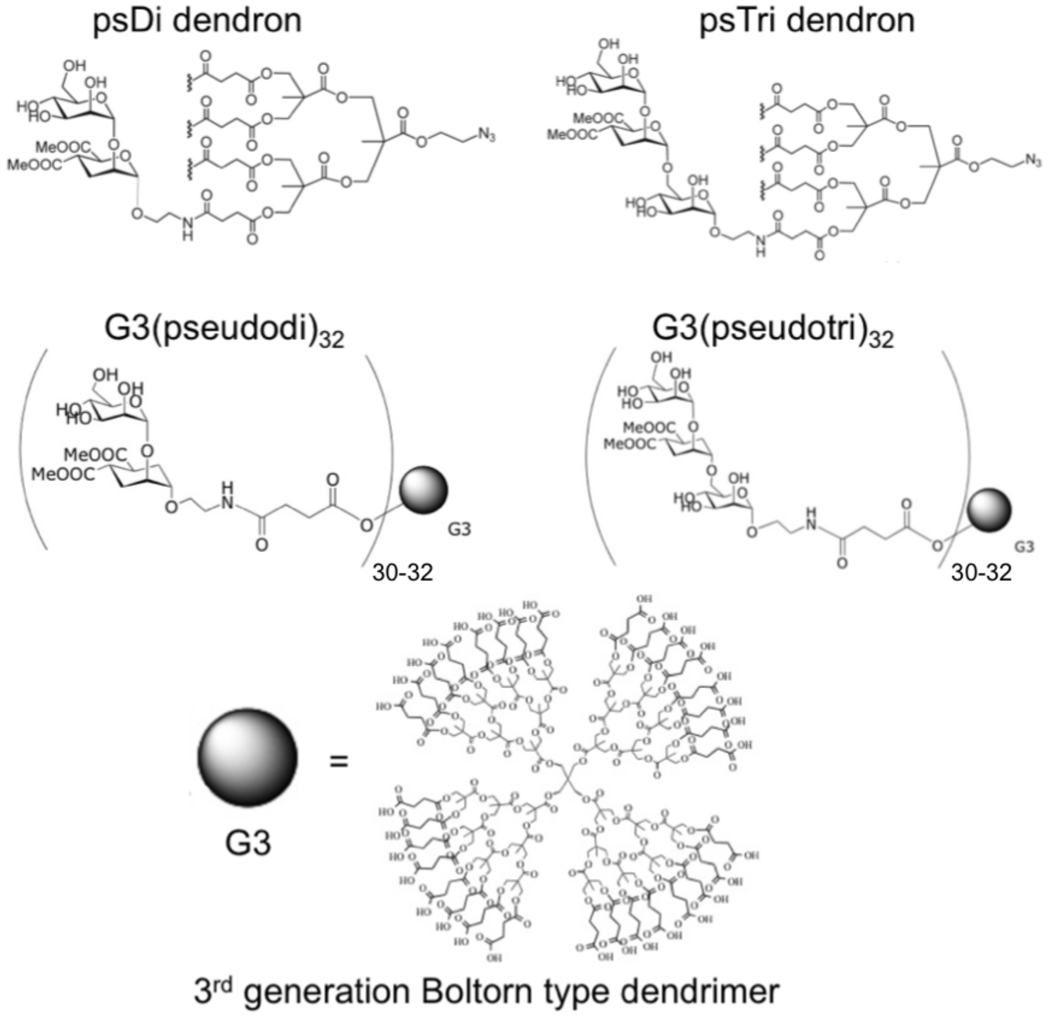


**Figure 9.36.:** The structures of psDi and psTri functionalized Boltorn type dendrons and dendrimers that were analyzed in SPR direct interaction assay.

Unfortunately, the monovalent ligands could not be assayed in this set-up due to their low affinity and low molecular weight: the injected high concentrations cause too strong bulk responses, which do not allow discriminating the real binding response from the SPR signal error.

#### 9.2.3.1. *The set-up of the direct interaction assay*

**DC-SIGN surface preparation.**

Due to the following reasons, the standard amine coupling procedure could not be used to generate surfaces with "oriented" functional DC-SIGN tetramers, meaning that all 4 CRDs of the tetramer are available for interaction. The protein to be immobilized by amine coupling should have a positive net charge; since DC-SIGN ECD has a pI of 5.16, the positive net charge can be achieved only by preparing the protein sample in a buffer with an acidic pH. However, the previous studies in our group [80] indicated that acidic pH induces irreversible dissociation of DC-SIGN tetramers, therefore preparing DC-SIGN ECD in acidic pH for amine coupling would lead to immobilization of unfolded monomeric ECDs.

To overcome this problem, DC-SIGN ECD construct with a StrepTag II in its N-terminus was generated previously (see subsection 8.1.2 on page 126). The latter tag allows the biospecific capture by StrepTactin at pH compatible with the stable structure of tetrameric DC-SIGN ECD. StrepTactin can be covalently immobilized to CM-dextran matrix on sensor chip by amine coupling. Figure 9.37 schematically illustrates the principal set-up of the direct interaction assay.

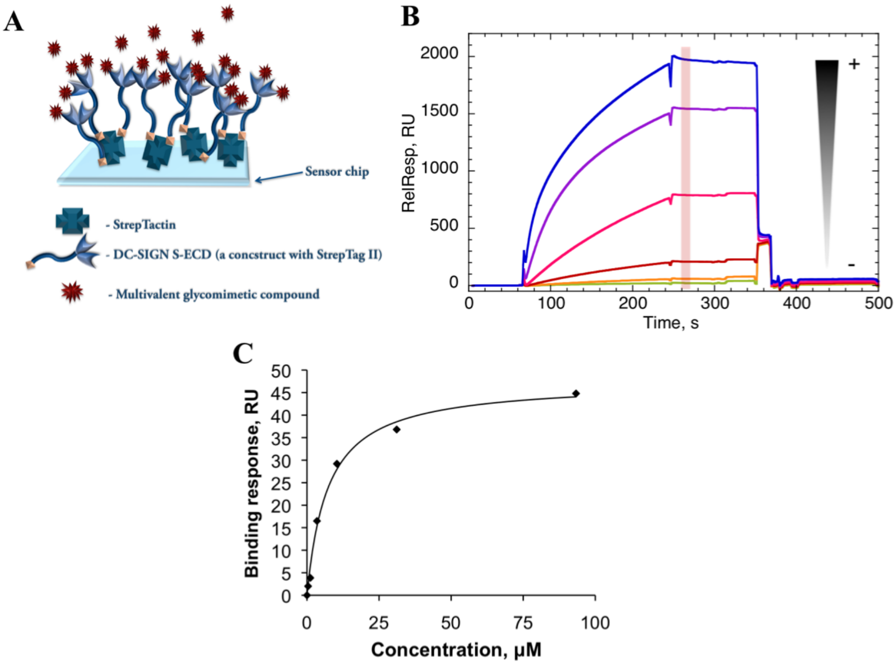


**Figure 9.37.:** A schematic illustration of the SPR direct interaction assay.

**A**, DC-SIGN S-ECD is captured by a covalently immobilized StrepTactin, then the multivalent glycomimetic compounds are flown over the prepared surface. **B**, The binding responses of the compounds, injected in increasing concentration order, can be measured at the indicated interval. **C**, The obtained binding responses can be plotted against compound concentration and the $K_\mathrm{D}$ values can be calculated using eq. 9.1 on page 148 (the shown example plot in panel C does not correspond to the sensorgrams shown in panel B). Alternatively (if possible), the interaction affinity could be determined by evaluating kinetic parameters, i.e. association and dissociation rate constants $k_\mathrm{on}$ and $k_\mathrm{off}$.

In order to determine the suitability of DC-SIGN S-ECD capturing to immobilized StrepTactin surface, the affinity of this DC-SIGN construct to StrepTactin was estimated (fig. 9.38), and resulted in a reasonable apparent $K_\mathrm{D}$ value of 57 nM. This is a higher value than the $K_\mathrm{D}$ of 1 μM indicated by the manufacturer for StrepTag II/StrepTactin interaction, which likely resulted due to the avidity effect: because DC-SIGN ECD is tetrameric, StrepTag II is concentrated, and on the other side, StrepTactin is also tetrameric protein.

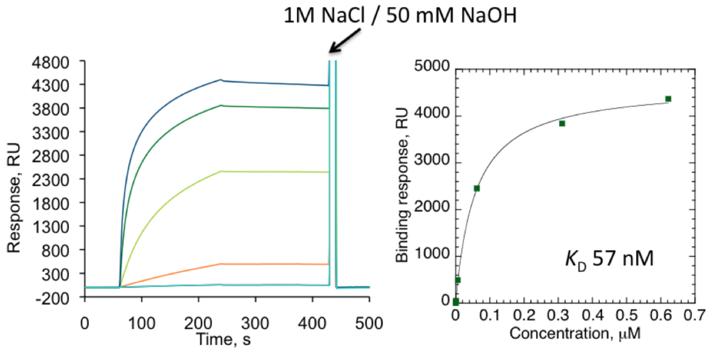


**Figure 9.38.:** DC-SIGN S-ECD affinity to StrepTactin.
The left panel shows the sensorgrams for the injections of a range of DC-SIGN S-ECD concentrations (0.62 nM – 622 nM of DC-SIGN S-ECD tetramers prepared in HBS-P buffer) over covalently immobilized StrepTactin surface (2700 RU), and the right panel represents the corresponding binding responses as a function of DC-SIGN concentration. The injection point of the indicated regeneration solution (20 µL at 100 µL/min) is indicated above. The flow rate of 5 µL/min of HBS-P buffer was maintained.

The next step was to test the activity of the prepared surface. For this reason, the dendrimer G3(pseudotri)$_{32}$ (see fig. 9.36) was injected over DC-SIGN surface (fig. 9.39).

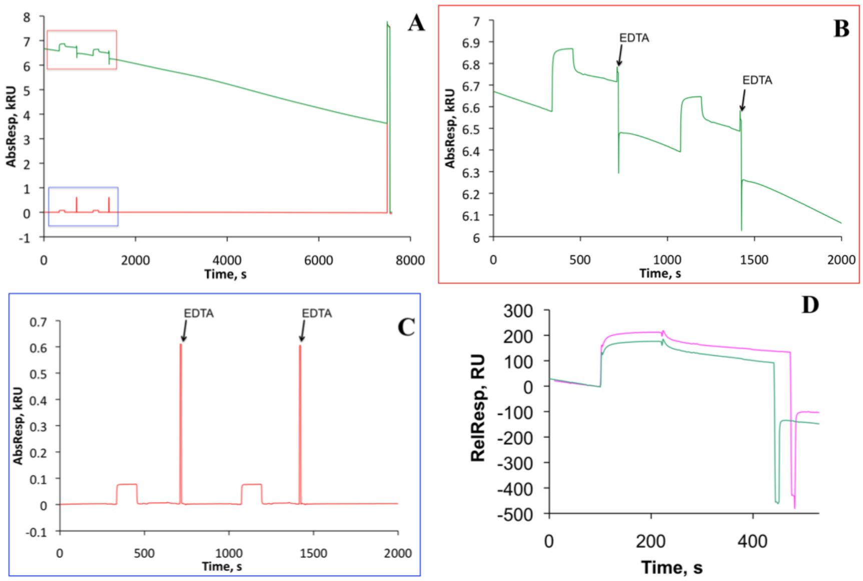


**Figure 9.39.:** The activity of oriented DC-SIGN surface.
**A**, The sensorgrams showing two subsequent injections of G3(pseudotri)$_{32}$ (10 µL of 20 µM at 5 µL/min) over the reference (red) and DC-SIGN S-ECD surface (green) followed by the running buffer flow (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl$_2$, 5 µL/min). **B** and **C**, The insets, highlighted in panel A by a black and blue frames, showing the two injections of G3(pseudotri)$_{32}$ over DC-SIGN (panel B) and StrepTactin (panel C) surfaces; the injection of regeneration solution (50 mM EDTA pH 8, 5 µL at 5 µL/min) followed each injection of the compound, as indicated. **D**, Reference surface corrected overlaid sensorgrams corresponding to the first (pink) and the second (green) injections of G3(pseudotri)$_{32}$. 25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM CaCl$_2$ was used as the running buffer.

The prepared DC-SIGN surface was indeed active as the dendrimer bound (about 200 RU) in a $Ca^{2+}$-dependent manner (interaction could be abolished by EDTA). Moreover, the compound did not bind to the reference surface, which showed that the observed interaction is specific to DC-SIGN.

However, a rapid dissociation of DC-SIGN S-ECD from StrepTactin surface was observed, as indicated by the drifting baseline, despite the reasonable apparent affinity determined for DC-SIGN S-ECD / StrepTactin interaction. Such an unstable surface could not be used for the DC-SIGN/compound interaction studies: the compounds should be injected multiple times, making the comparison between injections impossible due to continuously decreasing amount of DC-SIGN. As illustrated in fig. 9.39D, the binding responses of two subsequent injections of the same sample are markedly different.

Therefore, the stabilization of DC-SIGN S-ECD surface had to be explored, and the approach of covalent stabilization of captured DC-SIGN S-ECD was chosen. For this reason, after covalently immobilizing StrepTactin, an additional injection of EDC/NHS mixture was done in order to activate carboxyl groups on StrepTactin and/or residual carboxyl groups on CM-dextran. Then DC-SIGN S-ECD, prepared in HBS-P buffer, was injected over such a reactivated surface (fig. 9.40A). To control the binding specificity, DC-SIGN ECD, i.e. without a construct without Strep-Tag II, prepared at the same concentration as S-ECD construct, was also injected over reactivated dextran/StrepTactin surface (fig. 9.40B).

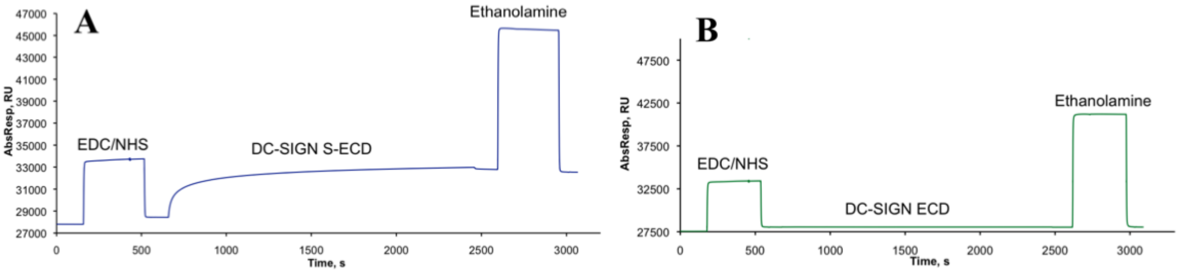


**Figure 9.40.:** The sensorgrams showing the immobilization of DC-SIGN constructs on reactivated dextran/StrepTactin surface.

The dextran/StrepTactin surface was reactivated by 50 µL injection EDC/NHS mixture. Then DC-SIGN S-ECD (panel A) or DC-SIGN ECD (panel B) constructs prepared at 60 µg/mL concentration in HBS-P buffer were injected (150 µL/min) over reactivated surface at a flow rate of 5 µL/min of HBS-P running buffer. The remaining activated –COOH groups were blocked by 30 µL injection of 1 M ethanolamine pH 8.

The latter experiment confirmed that DC-SIGN S-ECD is indeed covalently immobilized through StrepTactin/StrepTag II interaction, thus in a well-oriented way, since DC-SIGN ECD construct without StrepTag II did not bind to the reactivated surface. The stability of the surface as shown by later experiments was indeed much better (see figure 9.43 on page 308).

### DC-SIGN surface regeneration optimization and stability.

The first compounds tested by the direct interaction assay were the Boltorn type dendrimers and dendrons functionalized with psDi and psTri (see fig. 9.36 for the structures). This initial test indicated the inability to fully remove DC-SIGN-bound dendrons by EDTA injection (fig. 9.41).

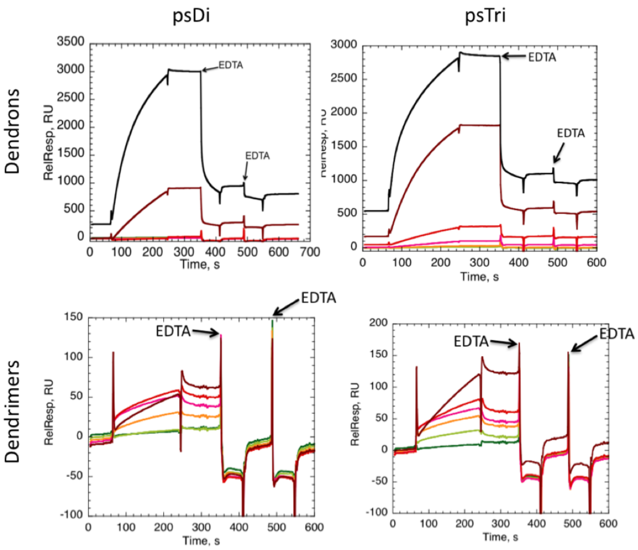


**Figure 9.41.:** Reference surface corrected overlaid sensorgrams showing the inability to regenerate DC-SIGN surface with EDTA.

The sensorgrams showing injections of psDi and psTri functionalized compounds are on left and right, respectively, and the tetravalent dendrons and 32-valent dendrimers are on top and bottom, respectively. The compounds were injected (15 µL at 5 µL/min flow rate) at a concentration range of 3 µM – 2400 µM for dendrons and 0.38 µM – 93.3 µM for dendrimers. The injection points of 50 mM EDTA pH 8 (5 µL at 5 µL/min each) are marked.

Such an inability to regenerate DC-SIGN surface with EDTA was not only a technical problem of the assay: it indicates that these compounds, for which the interaction is not abolished by EDTA, bind to DC-SIGN not only through a conventional sugar recognition site, but have $Ca^{2+}$-independent interaction with other unknown parts of the protein (no binding to the reference surface was observed, data not shown). However, better regeneration conditions were explored.

After screening eight different regeneration solutions (as listed in sub-subsection 6.2.5.2 on page 115; the sensorgrams of the screening are shown in Appendix), a mixture consisting of 50 mM Gly-NaOH pH 11.9 / 0.15% TritonX100 / 25 mM EDTA pH 8, was found to efficiently remove the bound compounds from DC-SIGN surface (fig.9.42).

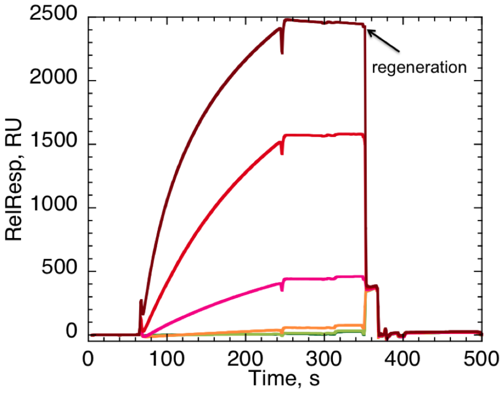


**Figure 9.42.:** An example of reference surface corrected overlaid sensorgrams showing improved compound detachment from DC-SIGN surface.

Tetravalent dendron functionalized with psDi was injected (15 µL at 5 µL/min flow rate) at a concentration range of 22 µM – 2400 µM over covalently stabilized DC-SIGN S-ECD surface. The injection point of regeneration mix (8 µL of 50 mM Gly-NaOH pH 11.9 / 0.15% TritonX100 / 25 mM EDTA pH 8, at 100 µL/min flow rate) is marked.

This solution was further tested in order to find out if using it does not lead to the loss of DC-SIGN surface activity. For this reason, ManBSA (0.56 µM) was used to monitor DC-SIGN surface activity, and 10 cycles[1] of ManBSA injections followed by regeneration (8 µL injection at 100 µL/min) were performed (fig. 9.43).

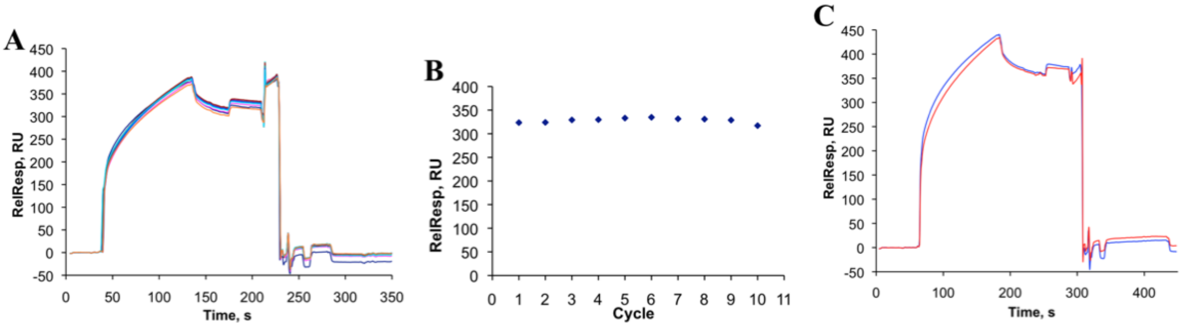


**Figure 9.43.:** The evaluation of DC-SIGN surface stability.
**A**, Reference surface corrected overlaid sensorgrams showing 10 subsequent injections of ManBSA (10 µL of 0.56 µM) over DC-SIGN surface. Each injection was followed by regeneration with a mix consisting of 50 mM Gly-NaOH pH 11.9 / 0.15% TritonX100 / 25 mM EDTA pH 8, injected (8 µL) at 100 µL/min flow rate. **B**, The corresponding binding responses measured after each injection and plotted against cycle number (i.e. injection number). **C**, Reference surface corrected overlaid sensorgrams showing two injections of ManBSA (0.48 µM), the first (blue) and the last (red) in the series of 44 cycles in total.

Indeed, using the latter regeneration conditions, a reasonable DC-SIGN surface stability was maintained as ManBSA binding responses remained very similar (fig. 9.43B). Even more, the later tests where 44 injections were performed and ManBSA was injected first and last in the run, indicated a remarkable stability of the surface as the binding responses of ManBSA remained the same (fig. 9.43C).

#### 9.2.3.2. *The studies of multivalent compound interaction with oriented DC-SIGN surface*

##### The results of the first SPR direct interaction experiment.

When the optimal surface regeneration conditions were found, the psDi and psTri functionalized dendrons and dendrimers were analyzed again. Figures 9.44 and 9.45 shows the sensorgrams and corresponding plots of binding responses as a function of compound concentration.

---

[1] The term “cycle” here means the injection of an analyte followed by surface regeneration and corresponds to one sensorgram

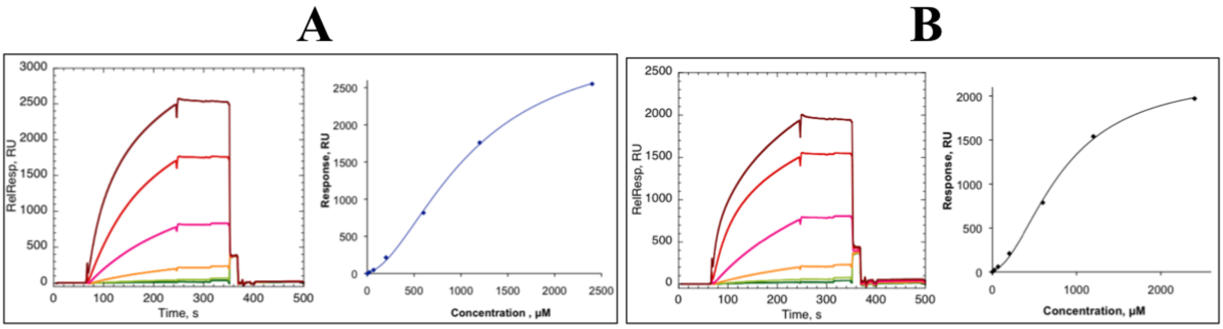


**Figure 9.44.:** The interaction of tetravalent dendrons with oriented DC-SIGN surf ace.

The left and the right plots show respectively the reference surface corrected sensorgrams and corresponding binding responses of psDi (panel A) and psTri (panel B) functionalized dendrons. Both dendrons were injected at the concentration range of 22 µM - 2400 µM. The solid lines in the left panels represent the fitted curve of 4-parameter model (eq. 6.1 on page 114).

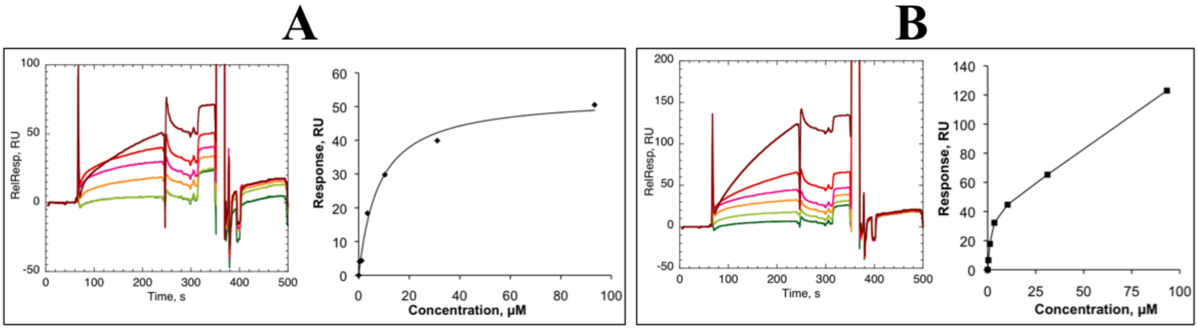


**Figure 9.45.:** The interaction of G3(pseudodi)$_{32}$ and G3(pseudotri)$_{32}$ dendrimers with oriented DC-SIGN surface.

The left and the right plots show respectively the reference surface corrected sensorgrams and corresponding binding responses of G3(pseudodi)$_{32}$ (panel A) and G3(pseudotri)$_{32}$ (panel B) dendrimers. Both compounds were injected at the concentration range of 0.38 µM - 93.3 µM. The solid line in the left of panel A represent the fitted curve of steady state affinity model (eq. 9.1 on page 148), and in panel B the line simply connects the data points.

Surprisingly, the tetravalent dendrons appeared to have more stable binding to DC-SIGN, as the dissociation phases were apparently slower than those for 32-valent dendrimers. Moreover, the binding responses of lower molecular weight dendrons were much higher than for the dendrimers that have higher molecular weights. Therefore it was interesting to check the observed maximal binding responses with the theoretically calculated $R_{\max}$ values that could be expected for the corresponding compounds.

$$R_{max} = \frac{MW_{analyte}}{MW_{ligand}} \cdot R_{ligand} \cdot n \tag{9.7}$$

where $MW_{\text{analyte}}$ and $MW_{\text{ligand}}$ are the molecular weight of the injected analyte (i.e. compound) and the immobilized ligand (i.e. DC-SIGN S-ECD, MW=161.6 kDa of a tetramer), $R_{\text{ligand}}$ is the quantity (in RU) of the immobilized ligand, and $n$ is ligand:analyte interaction stoichiometry. The results of the experiments shown in figures 9.44 and 9.45 were obtained with the surface where 3000 RU of DC-SIGN S-ECD tetramers were immobilized. Table 9.7 shows the comparison of theoretical and experimental $R_{\max}$ values calculated assuming two limiting cases: *1)* one DC-SIGN tetramer binds only one molecule of the compound, i.e. $n$=1; *2)* one DC-SIGN tetramer binds four molecules of the compound, i.e. $n$=4.

**Table 9.7.:** The comparison of calculated $R_{\max}$ values with the experimentally observed maximal binding responses of Boltorn type dendrimers and dendrons functionalized with psDi or psTri.

| | | **Calculated $R_{\max}$, RU** | | **Maximal observed response, RU** |
|---|---|---|---|---|
| | **MW, kDa** | $n$=1 | $n$=4 | |
| **psDi dendron** | 2.513 | 47 | 187 | 2550 |
| **psTri dendron** | 3.162 | 59 | 235 | 1970 |
| **G3(pseudodi)$_{32}$** | 19.592 | 364 | 1455 | 50 |
| **G3(pseudotri)$_{32}$** | 24.618 | 457 | 1828 | 123 |

The observed maximal binding responses for both dendrimers were lower than the calculated values even for the case when one tetramer binds one molecule. It might suggest that only a part of the CRDs in DC-SIGN tetramers is active due to the following reasons: *1)* a part of the CRDs is not well-folded and therefore not active, or *2)* considering that DC-SIGN surface density is high, one molecule of the dendrimer may bind several tetramers and thus block binding site availability for other dendrimer molecules. However, it must be noticed that the sensorgrams representing the highest compound concentration injected did not reach the equilibrium, therefore the real experimental Rmax is higher than the observed maximal binding response.

On the other hand, the maximal binding responses for both dendrons were significantly higher than the calculated $R_{\max}$, and considering that the binding did not reach equilibrium, the real experimental $R_{\max}$ would be even higher. In the first tests of direct interaction with compounds shown in fig. 9.36, only the dendrons could not be removed from DC-SIGN surface by EDTA (fig. 9.41), which indicated that they additionally interact with DC-SIGN in a $Ca^{2+}$-independent manner. Combining this observation and much higher experimental $R_{\max}$ values for the dendrons, it is possible to envision that several dendrons bind one DC-SIGN tetramer at multiple sites, in addition to the conventional $Ca^{2+}$-dependent sugar binding site. Since the major part of the bound dendrons could be removed by EDTA (fig. 9.41), the reason of the much higher experimental $R_{\max}$ values could also be the formation of compound aggregates, which would cause the higher binding responses due to increased molecular weight of the binding entity. It is likely that both of these phenomena occur simultaneously.

The compound aggregation, although in a much lesser extent than for dendrons, could be tracked also for G3(pseudotri)$_{32}$ dendrimer, as at the higher injected concentrations the response increased disproportionally[2] (fig. 9.45B).

---

[2] The term “disproportional(ly)” will be used to define a situation where binding occurs without and even above the saturation awaited for a standard ligand-receptor interaction, whereas in this context the term “proportional(ly)” will refer to the standard case of receptor saturation by the increased ligand concentration.

Because of the high DC-SIGN surface density, which creates conditions for mass transfer, the interaction of the tested compounds could not be evaluated in kinetic terms. Nevertheless, the measured binding responses as a function of compound concentration could be used to calculate the apparent equilibrium affinity values ($K_D$) using eq. 9.1 on page 148. Because of the unexpected above described behavior of the dendrons, this latter data analysis approach could be applied only for the dendrimers, and the calculated $K_D$ values are listed in table 9.8.

**Table 9.8.:** The calculated apparent $K_D$ values for psDi and psTri functionalized 3rd generation Boltorn type dendrimers.

| | **Apparent $K_D$ ± SD*, µM** |
|---|---|
| G3(pseudodi)$_{32}$ | 7.6 ± 0.5 |
| G3(pseudotri)$_{32}$** | 2.4 ± 0.1 |

*SD is calculated from two measurements on the same surface and same samples. ** The two highest concentrations were excluded for the $K_D$ calculation (fig. 9.45B).

Both dendrimers have very similar apparent affinities to immobilized DC-SIGN. In order to get a more quantitative evaluation of dendron interaction with DC-SIGN, the EC$_{50}$ (concentration, with which half of maximal binding response is reached) were calculated using equation 6.2 on page 114, where *50* was exchanged to the half of the observed corresponding $R_{\max}$ value. The EC$_{50}$ values of 925 µM and 730 µM for psDi and psTri dendrons, respectively, were found. These values do not correlate with and are 4 and 6 times (for psDi and psTri dendrons, respectively) higher than the IC$_{50}$ values determined in SPR competition assay. Such a difference arose most likely due to the strange behavior of the dendrons, i.e. a possible compound aggregation, which led to the artificially high concentration for which the maximal binding is observed, thus the overestimation of EC$_{50}$ value.

### The results of the second SPR direct interaction experiment.

The new type dendrimers, built by the click chemistry reactions, with mannose-based compounds were tested in the second SPR direct interaction experiment. Firstly, all of the compounds listed in table 9.2 were analyzed, except mannose C-glycosides, ***S9-1-4h***, and ***R1***, ***R*2**, ***R*3** compounds. These dendrimers were prepared at concentration ranges of 4 µM – 2.5 mM in running buffer without DMSO, and injected (15 µL, 5 µL/min) over the surface with DC-SIGN S-ECD (3000 RU) covalently captured to StrepTactin. Sensor chip CM5 was used. The majority of the compounds bound to DC-SIGN surface as can be seen in the represented sensorgrams (fig. 9.46).

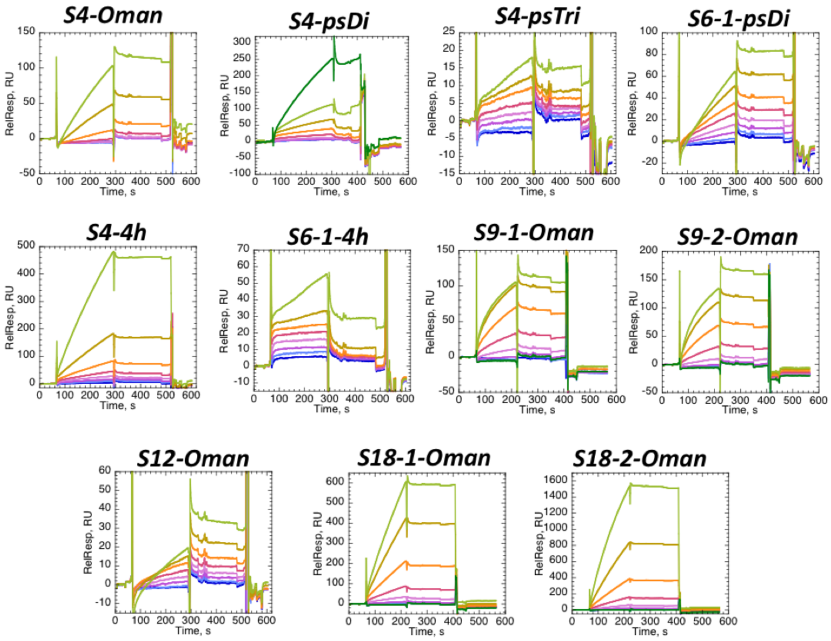


**Figure 9.46.:** Reference surface corrected sensorgrams showing the binding of flexible dendrimers with mannose-based compounds to DC-SIGN surface.

As it can be noticed from the sensorgrams, the binding equilibrium was far from being reached for the majority of the tested compounds, probably with the exception for compounds ***S6-1-4h***, ***S9-1-Oman*** and ***S9-1-Oman***. Regarding the observable kinetic aspect of the compound interaction with DC-SIGN, nearly all of them had a slow association phase followed by very slow dissociation, which indicated a stable binding.

The hexavalent dendrimer with the selected lead **4h**, ***S6-1-4h***, showed an exceptional behavior: the association was rather fast and the binding equilibrium was reached rapidly, while the dissociation had two phases, the rapid one followed by a very slow dissociation. Also the sensorgrams corresponding to the last two (highest) concentrations injected deviated from the general pattern for lower concentrations, as the binding appeared to continue further without reaching equilibrium.

Unfortunately, during this set of experiments DC-SIGN surface was not as stable as observed before: only 10% of its prior activity remained as judged by ManBSA binding responses in the beginning and the end of the experiments. Measuring absolute responses of the baseline at the beginning of each cycle, showed a significant loss of immobilized DC-SIGN, i.e. about 8 RU per cycle. Therefore the loss of absolute surface activity most likely resulted from the dissociation of DC-SIGN than the deactivation of binding sites. It is not possible to tell the reason why DC-SIGN S-ECD tetramers, captured by StrepTactin and covalently stabilized, were still dissociating, especially because in the previous experiment the surface displayed a remarkable stability. Therefore, the interaction of the compounds could be evaluated only qualitatively.

The binding responses of the compounds were plotted against corresponding concentration as shown in figure 9.47.

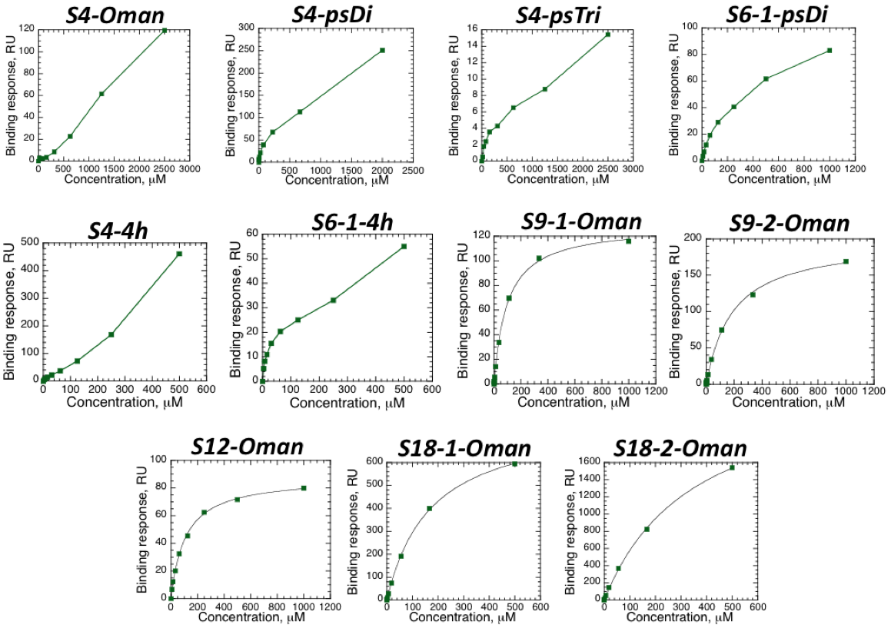


**Figure 9.47.:** DC-SIGN surface titration with flexible dendrimers.

The binding responses were measured after injections for the most of the compounds, except ***S6-1-4h***, for which binding responses were collected just before the end of injections.

All of the compounds with the scaffold valence lower than 9 displayed a disproportional increase of the binding responses with the higher concentrations injected.

Even though DC-SIGN density was decreasing during the course of the experiment, the observed maximal binding responses for all of the compounds were roughly compared with the calculated $R_{\text{max}}$ values using the value of 3000 RU of immobilized DC-SIGN (table 9.9, for a better comparison, the compounds are listed in the order as they were analyzed).

**Table 9.9.:** The comparison of calculated $R_{\text{max}}$ values with the experimentally observed maximal binding responses of the flexible dendrimers.

| | **MW, kDa** | **Calculated $R_{\text{max}}$, RU** | | **Maximal observed response, RU** |
|---|---|---|---|---|
| | | **$R_{\text{max}}$, $n=1$** | **$R_{\text{max}}$, $n=4$** | |
| ***S4-psDi*** | 2.156 | 40 | 160 | 251 |
| ***S9-1-Oman*** | 3.573 | 66 | 265 | 125 |
| ***S9-2-Oman*** | 3.7 | 69 | 275 | 170 |
| ***S18-1-Oman*** | 7.149 | 133 | 531 | 600 |
| ***S18-2-Oman*** | 7.325 | 136 | 544 | 1550 |
| ***S4-Oman*** | 1.285 | 24 | 95 | 120 |
| ***S6-1-psDi*** | 3.263 | 61 | 242 | 90 |
| ***S4-psTri*** | 2.791 | 52 | 207 | 15 |
| ***S4-4h*** | 2.983 | 55 | 222 | 480 |
| ***S6-1-4h*** | 4.525 | 84 | 336 | 60 |
| ***S12-Oman*** | 4.733 | 88 | 351 | 40 |

Compounds, for which the observed binding responsesare higher than calculated Rmax, are highlighted in red.

In fact, those compounds, for which disproportional binding response increase was observed, had also higher maximal binding responses than could be expected. Although ***S4-psTri***, ***S6-1-***

***psDi*** and ***S6-1-4h*** did not demonstrate higher experimental binding than the expected one, it must be kept in mind that they were analyzed using DC-SIGN surface with a decreased activity, hence the higher binding responses could be expected with the fully active freshly prepared DC-SIGN surface. As discussed earlier, the higher experimental binding responses than calculated $R_{\max}$ values may indicate the aggregation of the compounds. Combining the latter result with the observation of the disproportional increase of the signal when increasing concentration, it can be speculated that these compounds start to aggregate (or aggregation becomes more favored) only at the higher compound concentrations.

Although for the compounds ***S18-1-Oman*** and ***S18-2-Oman*** the binding responses increased proportionally[3] to the change of concentration, they also had significantly higher experimentally observed binding than the calculated one (table 9.9). This case could be explained by envisioning that these two compounds at all concentrations in water-based solution exist as larger aggregates and not monomeric molecules, for which the MW values were determined, and/or they bind to other sites on DC-SIGN, additionally to conventional sugar binding site.

The compounds ***S9-1-Oman*** and ***S9-2-Oman*** were the only ones, for which the binding responses increased proportionally to the concentration change, they had lower than expected maximal binding responses when analyzed with freshly prepared DC-SIGN surface, and in the sensorgrams observed the binding seemed to approach equilibrium. For those two compounds the apparent $K_{\mathrm{D}}$ values could be determined using steady state affinity model (eq. 9.1 on page 148), and resulted in 97 µM and 189 µM for ***S9-1-Oman*** and ***S9-2-Oman***, respectively. However, the 2-fold difference in affinity between these two nonavalent dendrimers with D-mannose is strange, and might be a result of measuring binding responses prior to reaching equilibrium (it was not possible to reach the binding equilibrium because of the high required compound volumes to be injected, thus the high amount of the compound, which was very limited).

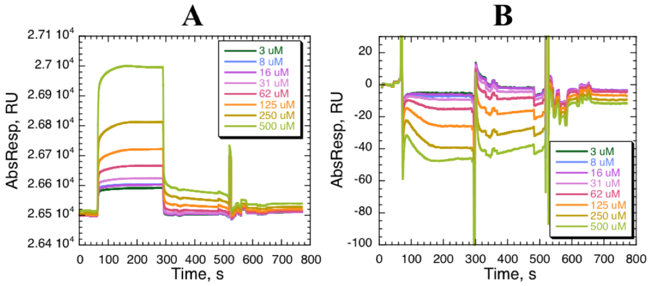


**Figure 9.48.:** The example sensorgrams showing a stronger interaction of the flexible dendrimer with StrepTactin than DC-SIGN surface.

**A**, Raw sensorgrams showing the binding of ***S9-1-psDi*** to immobilized StrepTactin (reference) surface. **B**, Reference surface corrected sensorgrams showing ***S9-1-psDi*** injections over DC-SIGN surface. The concentration range of ***S9-1-psDi*** was used as indicated.

The other compounds, analyzed in this first set included ***S6-1-Oman***, ***S6-1-psTri***, ***S6-2-psDi***, and ***S9-1-psDi***. All these four compounds had stronger, concentration-dependent interaction with the reference (StrepTactin) surface than DC-SIGN (fig. 9.48), which led to negative binding responses observed in reference surface corrected sensorgrams. No further analysis of these compounds could be performed.

---

[3] The term "proportional(ly)" refers to the standard case of receptor saturation by the increased ligand concentration

**The results of the third SPR direct interaction experiment.**

In this experiment the multivalent compounds functionalized with **4h** were analyzed. They included newly synthesized ***S4-4h*** and ***S6-1-4h*** dendrimers, and compounds ***S9-1-4h***, ***R1***, ***R2*** and ***R3*** from table 9.2 on page 296. The fucose-based dendrimers ***S4-fucβAla*** and ***S4-10b-azide1*** (table 9.3 on page 297) were also analyzed in the same experiment.

The priority of this run was to determine if the rod-like dendrimer ***R3*** is capable to bind to two CRDs within the same DC-SIGN tetramer. The underlying idea for the experiment (illustrated in figure 9.49) was the following: the interaction of ***R3*** with DC-SIGN should be investigated using the surfaces with DC-SIGN tetramers covalently captured at low and high densities. If the compound is indeed capable to bind to two CRDs within the same DC-SIGN tetramer, the surface density should not have an effect to the compound's affinity, thus the determined apparent affinities for all the three surfaces should be the same. Whereas, if one molecule of the compound binds to one CRDs of the tetramer, the affinity for the densely populated DC-SIGN surface would be higher: the close proximity of DC-SIGN tetramers would allow the bridging to occur, which would not be possible on the low density surface.

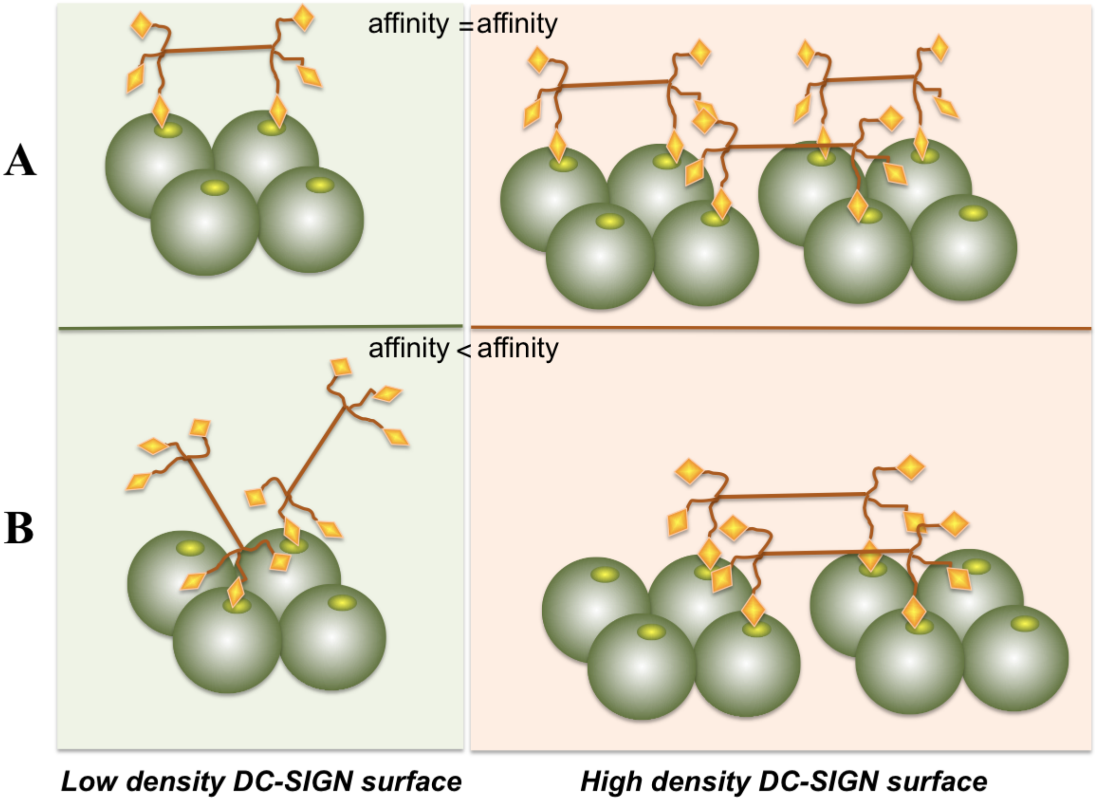


**Figure 9.49.:** The experimental set-up of SPR direct interaction assay to determine if the compound binds to 2 CRDs within the same DC-SIGN tetramer.

**A**, The case when compound is capable to bind to 2 CRDs within the same tetramer. **B**, The case when one molecule of the compound binds only to one CRD of the tetramer.

DC-SIGN S-ECD was covalently captured to three StrepTactin surfaces, immobilized to about 2800 RU by amine coupling procedure using sensor chip CM3 (procedure: 80 µL EDC/NHS, 100 µL of 100 µg/mL StrepTactin in sodium acetate pH 4, 80 µL ethanolamine, running buffer HBS-P at 5 µL/min). After reactivating CM-dextran/StrepTactin surfaces with 50 µL EDC/NHS injection, three different DC-SIGN densities were achieved (the approximate required DC-SIGN S-ECD concentrations were estimated from StrepTactin surface titration with DC-SIGN S-ECD, shown in figure 9.38 on page 305):

- DCSIGN-LD: 230 RU captured after injecting 15 µL of 0.35 µg/mL DC-SIGN S-ECD;
- DCSIGN-MD: 2030 RU captured after injecting 7 µL of 50.27 µg/mL DC-SIGN S-ECD;
- DCSIGN-HD: 3820 RU captured after injecting 50 µL of 696 µg/mL DC-SIGN S-ECD.

The non-reacted carboxyl groups were blocked by ethanolamine (50 µL injection). In order to remove non-covalently bound DC-SIGN S-ECD, which might improve DC-SIGN surface stability, the prepared surfaces were rinsed with 1 M NaCl / 50 mM NaOH solution (20 µL at 100 µL/min). The example sensorgram of this DC-SIGN S-ECD covalent capture is shown in figure 9.50.

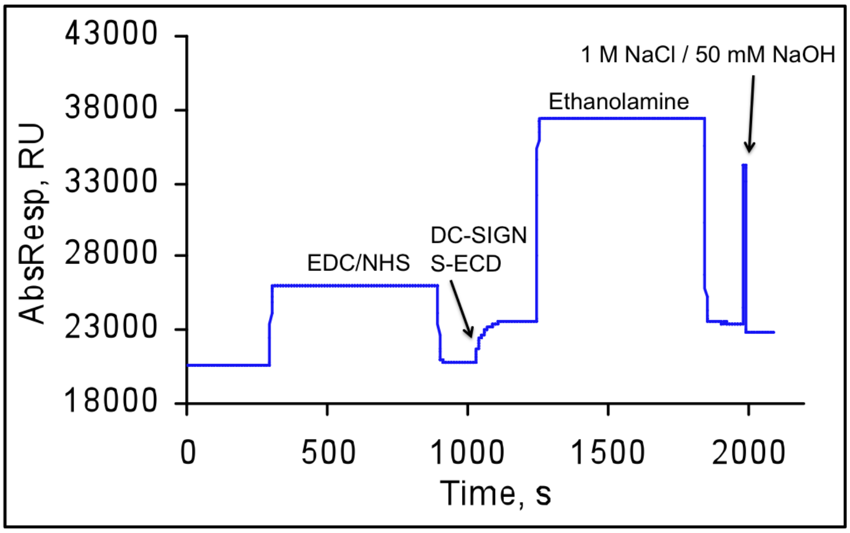


**Figure 9.50.:** The example sensorgram DC-SIGN S-ECD covalent capture followed by surface rinsing to remove non-covalently bound protein.

As usually, flow cell 1 was used as a reference surface and contained immobilized StrepTactin (2800 RU).

*Mass transfer effect evaluation for DC-SIGN surfaces.* The kinetic parameters of ***R3*** interaction with different density DC-SIGN surfaces could be also helpful: on a low density surface, faster dissociation could be predicted in the case where ***R3*** molecule binds one CRD of the tetramer (fig. 9.49B), and it would become slower when the surface density increases. Therefore, the prepared surfaces were tested for the mass transfer effect [177].

The recommended procedure for testing mass transfer effect consists of several injections of the same analyte at different flow rates and comparison of the shape of association and dissociation phases. If these shapes superimpose at all flow rates, the mass transfer is not significant.

In order to test this, ManBSA (27 µM) prepared in the running buffer consisting of 24 mM Tris-HCl pH 8, 144 mM NaCl, 3.8 mM $CaCl_2$, 0.005% P20 and 4% DMSO (also used for all of the below described tests), was injected over the surfaces at flow rates of 5, 30 and 80 µL/min.

The injected volumes corresponded to the association for 1 min, i.e. 5, 30 and 80 µL were injected, respectively. The obtained sensorgrams are compared in figure 9.51.

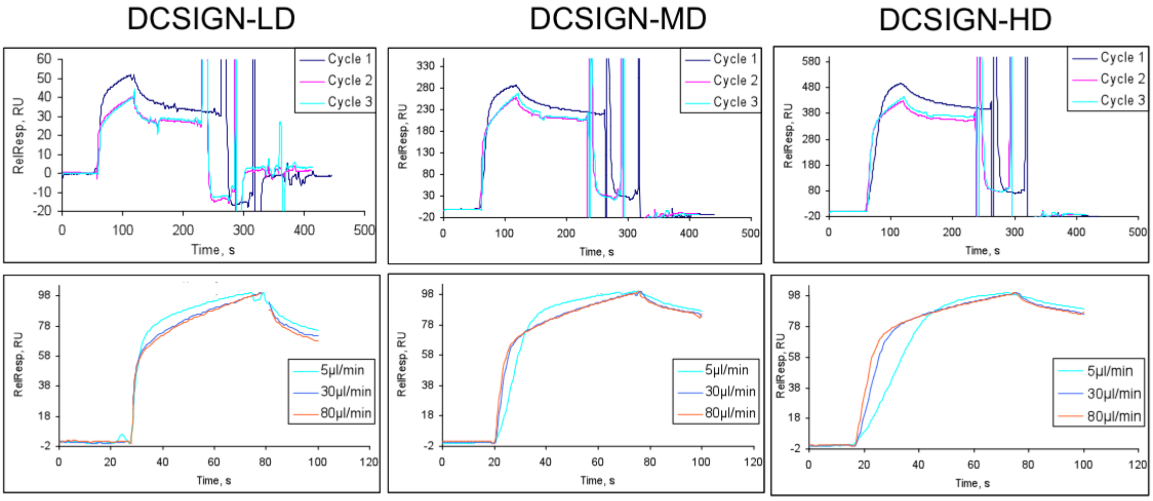


**Figure 9.51.:** The results of mass transfer test for different density DC-SIGN S-ECD surfaces. ManBSA (27 µM) was injected over the above indicated surfaces at flow rates of 5 µL/min (cycle 1), 30 µL/min (cycle 2) and 80 µL/min (cycle 3). All the upper panels show reference surface corrected overlaid sensorgrams obtained for the above-indicated surfaces, and all the lower panels represent the corresponding normalized sensorgrams.

For some reason ManBSA injection at 5 µL/min flow rate resulted in higher binding response than the other two injections, which gave very comparable ManBSA binding responses. In order to compare the association and dissociation phases obtained at all three flow rates, the sensorgrams were normalized (fig. 9.51, lower panels). Even if the first injection (at 5 µL/min) is not taken into account for comparison, it is obvious that the mass transfer effect is significant for the high DC-SIGN density surface (DCSIGN-HD), as association phases for injections at 30 µL/min and 80 µL/min do not superimpose. Even higher mass transfer effect could be expected for the glycomimetic dendrimers, since they have much lower MW than ManBSA, thus their diffusion to the surface would be much slower than that of ManBSA.

*Evaluation of ManBSA interaction with DC-SIGN surfaces.* In order to better characterize DC-SIGN surfaces prior the compound analysis, ManBSA titration was performed. Despite the observed mass transfer effect, the experiment was done in a kinetic regime, i.e. at a flow rate of 30 µL/min. ManBSA samples were prepared in the running buffer with the concentrations ranging from 2.7 nM to 266 nM. The injections of 30 µL of each sample were made and followed by 200 s dissociation phase. The binding was abolished with 50 mM EDTA pH 8 (5 µL, 5 µL/min). The resulting sensorgrams are shown in the upper panels of figure 9.52.

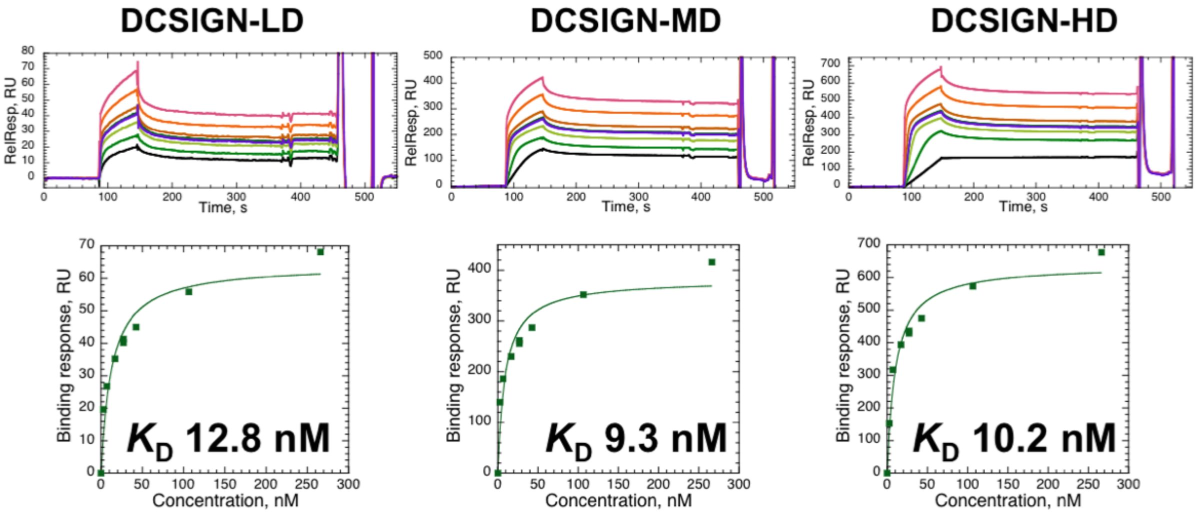


**Figure 9.52.:** The interaction of ManBSA with DC-SIGN surfaces of different densities. Upper panels show reference surface corrected sensorgrams obtained by ManBSA titration over the indicated density DC-SIGN surfaces. Lower panels represent the titration curves obtained by plotting binding responses at the respective density surfaces against ManBSA concentration.

None of the available interaction kinetic models fitted the data. However, the interaction kinetics could be qualitatively evaluated by visually inspecting the dissociation phases observed for different density surfaces. Clearly, the dissociation slows down with an increase of DC-SIGN surface density, which is particularly visible for the lowest concentration ManBSA sample (fig. 9.52, upper panels).

The interaction was evaluated quantitatively by fitting steady state affinity model (eq. 9.1 on page 148) to the plots of ManBSA binding responses against concentration, and the obtained apparent affinity values are displayed on the plots of lower panels of figure 9.52. The latter model obviously did not fit well the data indicating that the system is more complicated than assumed by this simple 1:1 binding model.

The apparent affinity values of $K_\mathrm{D}$ of about 10 nM were similar to all DC-SIGN surfaces with slightly lower affinity for low density DC-SIGN surface. It is also noteworthy that the affinity measured with immobilized DC-SIGN is much higher (2 orders of magnitude) than when ManBSA is immobilized. This could be expected because the absolute affinity of this interaction is determined by the avidity, which is more favorable when the binding partner with lower number of binding sites (i.e. DC-SIGN, 4 CRDs) is immobilized on the surface while the partner with more binding sites (ManBSA, 12 site of glycosylation) is free in the solution.

*Multivalent compound analysis using different density DC-SIGN surfaces.* Hence, it was chosen to perform the further experiment with the multivalent compounds using steady state interaction set-up, since the presence of mass transfer would prevent the reliable evaluation of kinetic parameters. The compounds were prepared at the concentration ranges similar to those used in SPR competition assay: 0.38 µM – 2.5 mM for compounds ***S4-4h*** and ***S6-1-4h***, 0.15 µM – 1 mM for compounds ***S9-1-4h***, ***R1*** and ***R2***, 0.3 µM – 2 mM for compounds ***S4-fucβAla*** and ***S4-10b-azide1***, and compound ***R3*** was prepared at 0.03 µM – 185 µM range. The prepared samples were injected (120 µL, 5 µL/min) over the surfaces and the binding was abolished by the injection (8 µL, 100 µL/min) of the previously optimized surface regeneration mix (50 mM Gly-NaOH pH 12 / 0.15% TritonX100 / 25 mM EDTA pH 8). The activity of DC-SIGN surfaces during the run was controlled by the injections (30 µL, 5 µL/min) of ManBSA (10 nM). The obtained sensorgrams are presented in the Appendix. The binding responses measured at the end of compound injections were plotted against compound concentration and are represented in figures 9.53, 9.54 and 9.55.

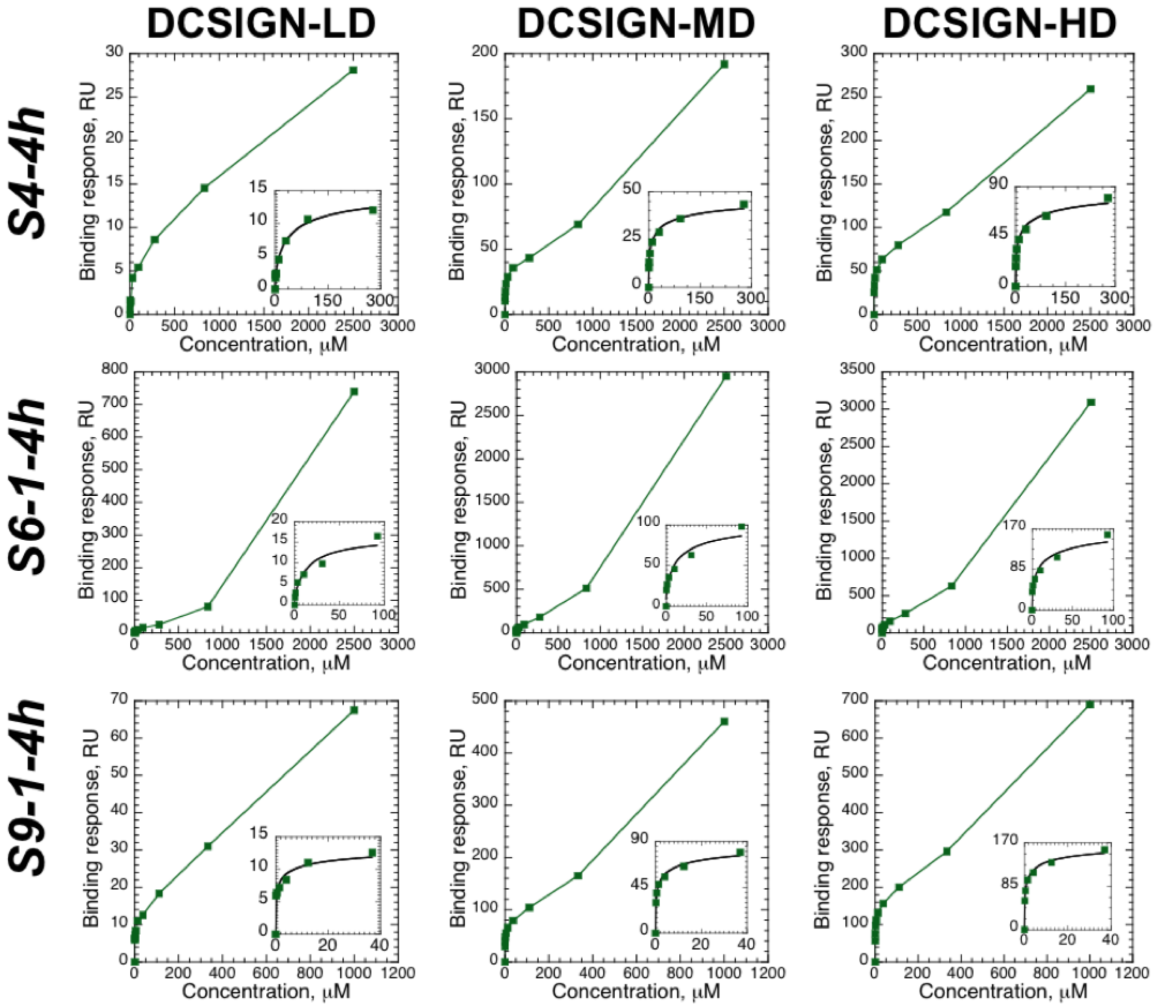


**Figure 9.53.:** The results of the interaction of **4h** functionalized dendrimers with DC-SIGN surfaces of different density.

The plots in each row correspond to the indicated compound, and the plots in each column correspond to the indicated different DC-SIGN surface densities. The insets show the first portion of the curves with 4 parameter logistic model (eq. 6.1 on page 114) fits.

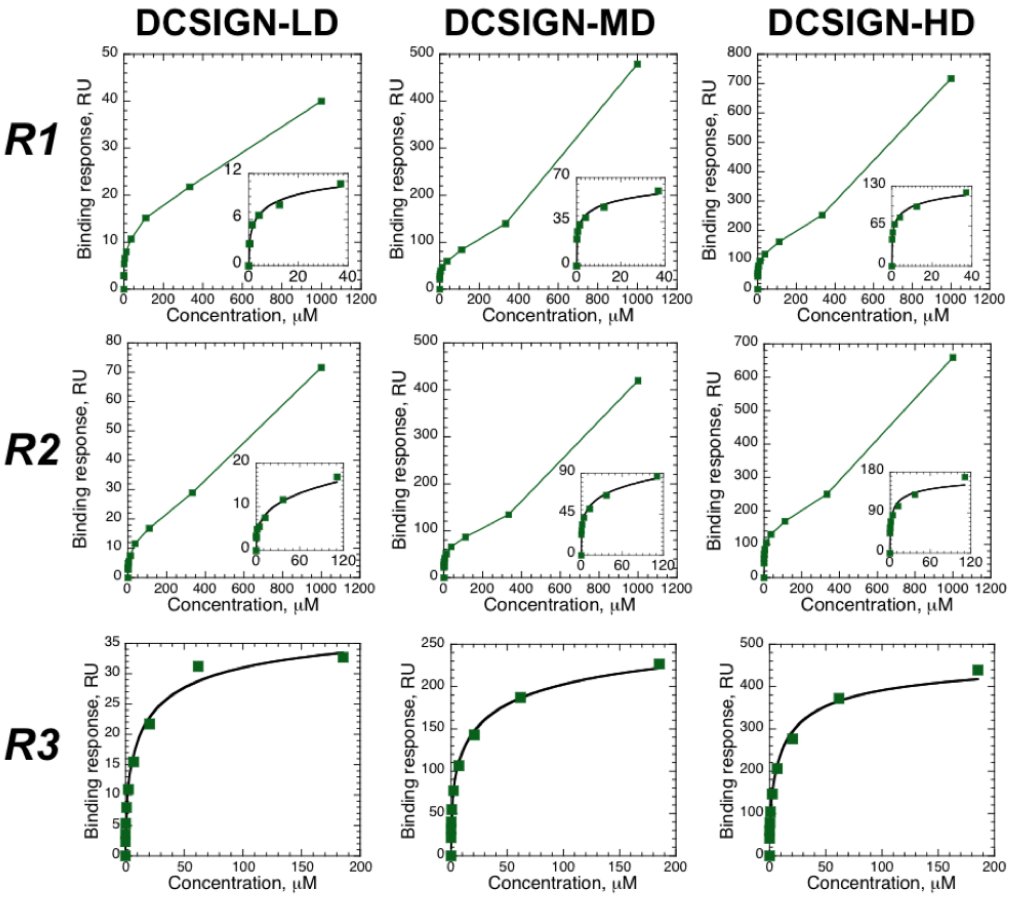


**Figure 9.54.:** The results of the interaction of **4h** functionalized rod-like and related dendrimers with DC-SIGN surfaces of different density.

The plots in each row corresponds to the indicated compound, and the plots in each column correspond to the indicated different DC-SIGN surface densities. The insets show the first portion of the curves with 4 parameter logistic model (eq. 6.1 on page 114) fits; for compound *R3*, the fitting was performed all of the data points obtained for all concentrations used.

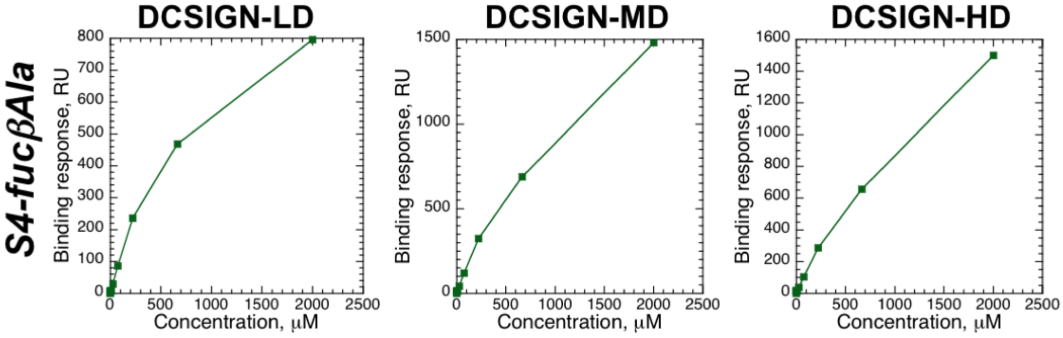


**Figure 9.55.:** The results of the interaction of tetravalent fucosylated dendrimer with DC-SIGN surfaces of different density.

Apparently, the results were very similar to the previously observed behavior of the flexible dendrimers: most of the compounds displayed the disproportional binding response increase at the higher injected concentrations. As previously, the theoretical $R_{\max}$ values were calculated for all of the compounds and compared with the maximal binding responses observed (table 9.10).

**Table 9.10.:** The comparison of calculated $R_{\mathrm{max}}$ values with the experimentally observed maximal binding responses for the tested dendrimers.

| Compound | MW, kDa | Calculated $R_{\mathrm{max}}$, $n=1$ | | | Calculated $R_{\mathrm{max}}$, $n=4$ | | | Observed maximal responses, RU | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | DCSIGN-LD | DCSIGN-MD | DCSIGN-HD | DCSIGN-LD | DCSIGN-MD | DCSIGN-HD | DCSIGN-LD | DCSIGN-MD | DCSIGN-HD |
| ***S4-4h*** | 2.983 | 4 | 37 | 71 | 17 | 150 | 282 | 21 | 192 | 259 |
| ***S6-1-4h*** | 4.525 | 6 | 57 | 107 | 26 | 227 | 428 | 740 | 2949 | 3093 |
| ***S9-1-4h*** | 7.394 | 11 | 93 | 175 | 42 | 372 | 699 | 68 | 460 | 690 |
| ***R1*** | 4.955 | 7 | 62 | 117 | 28 | 249 | 469 | 40 | 479 | 717 |
| ***R2*** | 5.103 | 7 | 64 | 121 | 29 | 256 | 483 | 72 | 420 | 659 |
| ***R3*** | 5.72 | 8 | 72 | 135 | 33 | 287 | 541 | 33 | 227 | 439 |
| ***S4-fucβAla*** | 1.966 | 3 | 25 | 46 | 11 | 99 | 186 | 796 | 1480 | 1500 |

Compounds, for which the observed binding responses are higher than calculated $R_{\mathrm{max}}$, are highlighted in red.

With the only exception for the rod-like dendrimer ***R3***, all other compounds had higher binding responses than the expected Rmax, with the most significant increase in maximal binding response for compounds ***S6-1-4h*** for ***S4-fucβAla***. Again, such a disproportional binding response suggests that these dendrimers form aggregates at higher concentrations, and the proportional evolution of too high binding responses observed for ***S4-fucβAla*** may mean that this dendrimer binds to additional sites on DC-SIGN other than $Ca^{2+}$-dependent sugar recognition site, or that it exists as higher molecular weight aggregates at all concentrations, or both.

Although the binding responses for ***R3*** did not exceed the calculated $R_{\mathrm{max}}$, they are very close to the theoretical value, calculated with an assumption that four molecules of the compound bind to the same DC-SIGN tetramer. In practice, the experimental $R_{\mathrm{max}}$ values are usually lower than the calculated ones because of the surface activity loss during experiment and/or ligand immobilization in an orientation not favorable for the interaction. In this experiment the control injections of ManBSA indicated a loss of surface activity (about 66%, 86% and 89% of primary activity of DCSIGN-LD, DCSIGN-MD and DCSIGN-HD surfaces remained after all tests). Therefore it is more likely that the rod-like dendrimer ***R3*** also forms aggregates (or binds DC-SIGN to additional sites), especially considering its low water solubility, than that all of the CRDs of immobilized DC-SIGN tetramers are active and each is occupied by one ***R3*** molecule. Even though all of the studied dendrimers had this unexpected behavior, a rough estimation of the affinities for their interaction with different density DC-SIGN surfaces was attempted. For this reason, the 4 parameter logistic model (eq. 6.1 on page 114) was fit to the selected first portion of the curves until the concentration, with which the rapid increase of the signal was obtained (the insets in figures 9.53 and 9.54). The apparent $EC_{50}$ values were calculated as described above and compared as shown in figure 9.56.

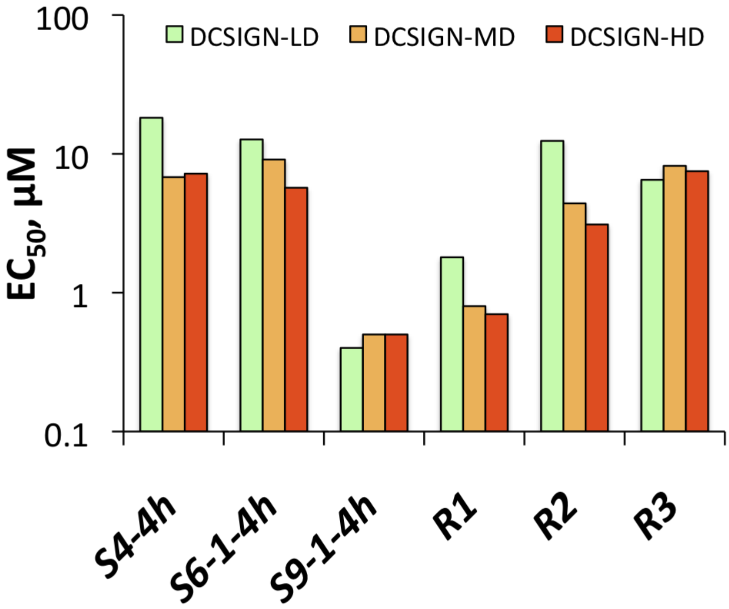


**Figure 9.56.:** A relative comparison of $EC_{50}$ values obtained for **4h** functionalized series of dendrimers.

Although the reliable affinity of these compounds to DC-SIGN surfaces could not be determined, the relative comparison of the calculated $EC_{50}$ values showed a general tendency: for four compounds the apparent affinity increased with the DC-SIGN surface density increase. The other two compounds, namely ***S9-1-4h*** and ***R3***, appeared to have similar affinities to the surfaces of all three different densities. This might indicate that the two latter compounds bind to several CRDs within the same DC-SIGN tetramer, a result initially anticipated for compound ***R3***, and the probable ability of ***S9-1-4h*** to bind to several CRDs of same tetramer may suggest that this compound in solution adopts a conformation favorable for such a binding. However, this is just a qualitative comparison of the compounds, which may only give a clue of the real situation.

The estimation of an apparent $EC_{50}$ value for the dendrimer ***S4-fucβAla*** was not possible because of a nearly linear concentration-binding relationship.

Regarding fucosed-based dendrimer ***S4-10b-azide1***, it was analyzed after ***S4-fucβAla*** dendrimer, which had very strong binding to the surfaces and could not be fully regenerated by the used regeneration solution. Thus the compound ***S4-10b-azide1*** was injected to the surfaces that still contained ***S4-fucβAla***, which made the comparison of ***S4-10b-azide1*** binding responses impossible.

The failure to fully remove ***S4-fucβAla*** dendrimer from DC-SIGN surfaces with 50 mM Gly-NaOH pH 12 / 0.15% TritonX100 / 25 mM EDTA pH 8 solution, indicated that the binding of this compound is in part $Ca^{2+}$-independent. Therefore the test of DC-SIGN surface regeneration was performed with all of the compounds in order to determine if the binding could be abolished by EDTA, i.e. the compounds interacted in a $Ca^{2+}$-dependent manner. For this reason, each compound was injected (5 µL, 5 µL/min) over the DC-SIGN surfaces. The tested compounds and their concentrations are shown in table 9.11.

**Table 9.11.:** The list of multivalent compounds, which were tested for ability to be removed from DC-SIGN surface by EDTA.

| Cycle | Compound | Concentration |
|---|---|---|
| 1 | ManBSA | 17 nM |
| 2 | ***S4-4h*** | 1 mM |
| 3 | ***S6-1-4h*** | 1 mM |
| 4 | ***S9-1-4h*** | 1 mM |
| 5 | ***R1*** | 1 mM |
| 6 | ***R2*** | 1 mM |
| 7 | ***R3*** | 185 µM |
| 8 | ***S4-psDi*** | 1 mM |
| 9 | ***S4-fucβAla*** | 2 mM |
| 10 | ***S4-10b-azide1*** | 2 mM |
| 11 | ManBSA | 17 nM |

After the compound injection, two subsequent regeneration solutions were injected: at first 50 mM EDTA pH 8 was injected (5 µL, 5 µL/min), then the previously optimized surface regeneration mix was injected (8 µL, 100 µL/min). The baseline responses (AbsResp) were measured in the beginning of each cycle (i.e. before the compound injection), after the first regeneration with EDTA, after the second regeneration and were plotted against cycle number (fig. 9.57).

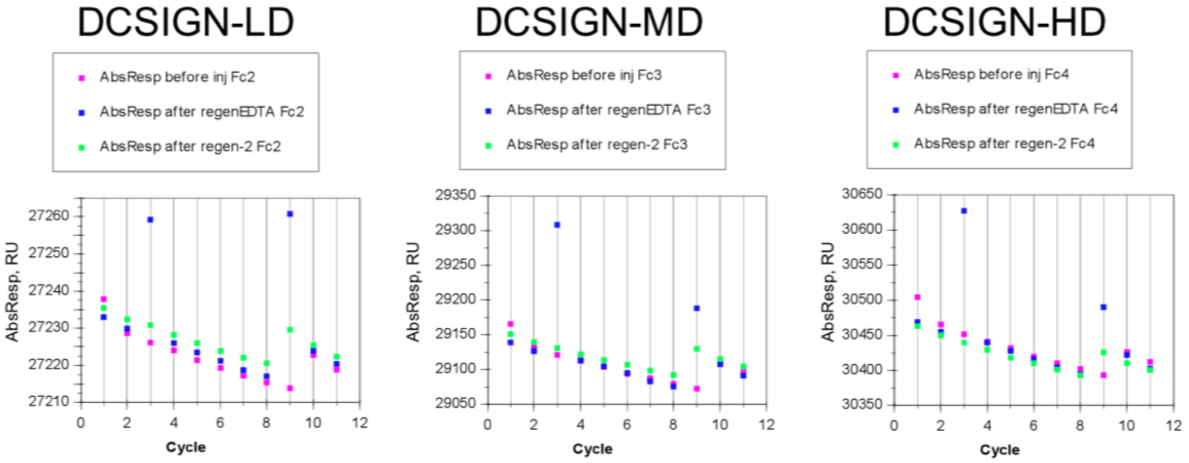


**Figure 9.57.:** The results of the test of DC-SIGN S-ECD surface regeneration after multivalent compound binding.

The three panels correspond to the results obtained with different DC-SIGN density surfaces, as indicated. Pink, blue and green data points correspond to baseline responses before compound injection, after EDTA injection and after the injection of the second regeneration mix (50 mM Gly-NaOH pH 12 / 0.15% TritonX100 / 25 mM EDTA pH 8), respectively.

The results of this test indicated two compounds (cycles 3 and 9), which failed to be removed from DC-SIGN surface by EDTA. These compounds were ***S6-1-4h***, and as expected, ***S4-fucβAla***. Recalling that for both of those compounds the experimentally observed maximal binding responses exceeded theoretical ones most significantly, it can be assumed that they interact with DC-SIGN at multiple sites in addition to $Ca^{2+}$-dependent sugar binding site.

In summary, the majority of the tested multivalent compounds display a complicated behavior:

aggregation, interaction with other sites than $Ca^{2+}$-dependent sugar binding site. Unfortunately, these phenomena make the quantitative evaluation of the compound interaction with DC-SIGN and an accurate comparison of the individual compounds barely feasible, and only the general tendencies can be tracked. Due to the significant mass transfer effect, even the qualitative comparison of the association and dissociation phases of the compound interaction with different density DC-SIGN surfaces would not be reliable.

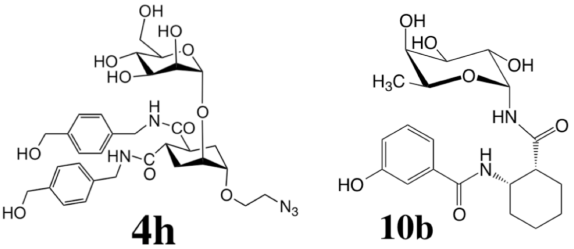
On the other hand, this SPR direct interaction assay showed that most of the compounds interact with DC-SIGN in a $Ca^{2+}$-dependent manner, with only two above indicated exceptions.

# Part IV.

# Discussion and future perspectives

# 10. The newly identified monovalent lead compounds as DC-SIGN antagonists.

Three different types of monovalent DC-SIGN ligands were developed including mannose-based glycomimics, fucose-based Lewis$^X$ trisaccharide mimics and various C-glycosidic compounds. In the pool of these monovalent compounds, two new leads, namely **4h** and **10b** (fig. 10.1), were identified by SPR competition assay among the *bis-benzylamide derivatives* of previously designed pseudodisaccharide psDi (fig. 9.14 on page 173) and the *Lewis$^X$ mimicking compounds* (Lewis$^X$ structure may be found in fig. 2.9 on page 68).


**4h**

**10b**

***IC*$_{50}$** 0.31 ± 0.04 mM
(DC-SIGN inhibition)

***IC*$_{50}$** 0.395 ± 0.043 mM
(DC-SIGN inhibition)

**Figure 10.1.:** The new identified glycomimetic leads.

Although both of these compounds have similar activities to inhibit DC-SIGN, the activity improvement comparing to their parental compounds is slightly better for mannose-based glycomimics **4h**. While fucose-based compound **10b** has an improvement of relative affinity compared to Lewis$^X$ trisaccharide by a factor of 1.9, the comparison of **4h** with its predecessors psDi and natural Manα1-2Man disaccharide shows slightly more than a 3-fold affinity improvement. Furthermore, the Lewis$^X$-type compounds displayed a remarkable selectivity to DC-SIGN compared with langerin, as could be expected knowing that langerin does not bind Lewis$^X$ trisaccharide [57, 113]. On the other hand, langerin is known to bind mannosylated carbohydrates [60, 113, 115], therefore the selectivity improvement for mannose-based compounds is a more challenging task in the course of the development of mannose-based DC-SIGN antagonists. Indeed, the selectivity gain of a factor of $\approx 6$ was reached for compound **4h** as compared to its predecessor psDi. Knowing that psDi already had the selectivity improvement of $sel.g = 3.5$ compared to natural disaccharide Manα1-2Man, even much higher $sel.g$ value (compared to Manα1-2Man) for **4h** may be predicted. Finally, the slightly better activity of **4h** than **10b** could be also observed for the multivalent forms of these two compounds: the tetravalent dendrimer with **4h** had about 3 times better activity than **10b** ($IC_{50}$ values 39 µM and 12 µM, respectively).

The design of **4h** and its identification as a new lead compound became possible only after the previous lead compound pseudotrisaccharide psTri (fig. 9.14 on page 173) was discovered to possesses the ability to cross-link two DC-SIGN tetramers (figure 7 in paper **n°5**). This unique feature of psTri ligand led to artificial overestimation of its activity in SPR competition assay where soluble DC-SIGN tetramers are used allowing the cross-linking to occur, which has an important effect on $IC_{50}$ value estimation. Along the latter discovery, a shorter analogue of psTri, the above-mentioned compound psDi, was found to have the required qualities for further development as a drug candidate (paper **n°4**). Therefore, the further development focused on psDi scaffold and yielded the lead compound **4h**. However, the detection of the cross-linking phenomenon characteristic to psTri, highlights the importance of testing the selected leads in multiple biophysical techniques in order to avoid the false positives identified due to the particular set-up of the competition experiment.

The activity and selectivity of **4h** may be further improved. The second type of psDi derivatives with substituents introduced at carbon 6 of the non-reducing mannose moiety of psDi molecule were designed in order to decrease the recognition by langerin. Among these derivatives, one compound, with an amino group instead of hydroxyl at carbon 6 (NV243, fig. 9.25 on page 290), exhibited an improved selectivity to DC-SIGN as compared to langerin (*sel.g* of ≈ 5 comparing to psDi). This compound also had a better activity of DC-SIGN inhibition than psDi. Therefore, combining the structural features of both **4h** and NV243 may lead to the design of a compound (fig. 10.2) with a much better activity and much higher selectivity.

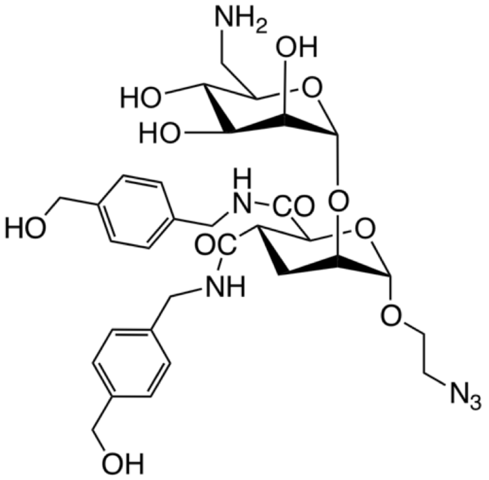


**Figure 10.2.:** A perspective mannose-based compound with combined features of the lead **4h** and compound NV243.

The structural basis of the improvement of the selectivity to DC-SIGN *vs* langerin with this new compound (fig. 10.2) could be also anticipated. In addition to the previous discovery that langerin is able to bind galactose with a sulphate group at position 6 [115], our group found that GlcNAc with an analogous sulphation at position 6 is also recognized by langerin, and as in the case of $6SO_4$-Gal, the sulphate group makes electrostatic contacts with Lys313 side chain (fig. 10.3A). If this sulphate was replaced by a group carrying a positive charge at a respective position, an electrostatic repulsion of the resulting compound could be expected.

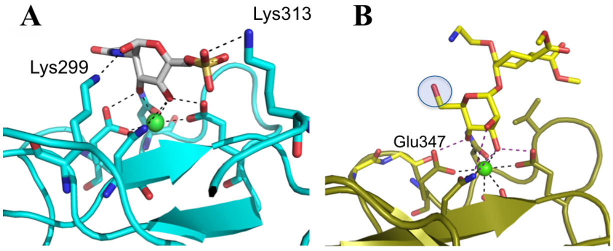


**Figure 10.3.:** The anticipated structural basis for the improvement of the selectivity to DC-SIGN *vs* langerin by the perspective mannose-based compound.

**A**, The structure of langerin in complex with $6SO_4$-GlcNAc and important lysines indicated (non-published data: the complex was obtained and the structure was solved in our group by a PhD student Eric Chabrol, and is described in his PhD thesis). **B**, DC-SIGN binding site in complex with psDi; 6-OH group on mannose moiety, which is replaced by amino group in NV243, is highlighted by a blue circle. Langerin (A) and DC-SIGN (B) binding sites are arranged in similar respective positions, $Ca^{2+}$ ions are shown as green spheres and polar contacts are represented by dashed lines.

The selectivity may be gained not only by decreasing the affinity to langerin but also by increasing it to DC-SIGN. The crystal structure of psDi binding mode in DC-SIGN CRD (fig. 10.3B) allows to predict that amino group instead of 6-OH of mannose may interact through electrostatic contact and/or H-bond with the neighboring surface of the lectin and notably carbonyl of Glu347. Indeed, the psDi derivative with $NH_2$ instead of 6-OH (NV243) showed improved affinity compared to psDi, and the additional aromatic groups of **4h** would improve the affinity even further.

In order to further improve *Lewis$^X$ mimics*, the structural knowledge of the lead compound interaction with DC-SIGN would be helpful. Therefore, co-crystallizing this compound with DC-SIGN CRD and solving the X-ray structure of the complex could be a future perspective for this project.

Regarding the *C-glycosidic compounds*, an interesting result was found: a simple L-fucose with a short alkyl-azide appendage linked to an anomeric carbon by C-glycosidic bond (compound BB133, fig. 9.6 on page 153) has twice higher activity than the natural L-fucose, whereas the analogous D-mannose C-glycoside (compound BB112, fig. 9.6 on page 153) had the same activity as natural D-mannose. Comparing the latter two C-glycosidic monosaccharides, L-fucose has 3 times better activity than D-mannose. The underlying reason for this affinity improvement observed for L-fucose C-glycoside could be investigated further by other biophysical techniques, such as ITC combined with structural studies of the interaction by NMR or X-ray crystallography. However, it may be speculated that the alkyl-azide linker of L-fucose is in a favorable position to make additional contacts in the binding site of DC-SIGN, while this is not possible in case of D-mannose with the same linker. Indeed, when the C-glycosidic L-fucose was tethered to the dendrimeric scaffolds (fig. 9.29 on page 294 and table 9.3 on page 297), the activity of the resulting compounds was comparable or even lower than of the corresponding dendrimers with O-glycosidic L-fucose (fig. 9.34 on page 300), suggesting that the contribution of the alkyl-azide linker to affinity improvement may have been lost in the multivalent presentation. On the contrary, the multivalent platforms with the C-glycosidic D-mannose for some reason appeared to have a better affinity than corresponding dendrimers with O-glycosidic D-mannose (fig. 9.31 on page 298). Finally, when the corresponding dendrimers with C-glycosidic L-fucose and C-glycosidic D-mannose are compared, a switch in activities can be noticed: while C-glycosidic L-fucose had higher affinity than C-glycosidic D-mannose as a monosaccharide, the multivalent presentation results in favor for C-glycosidic D-mannose and a slightly higher affinity is observed for the dendrimers with mannose than corresponding compounds with fucose (fig. 10.4).

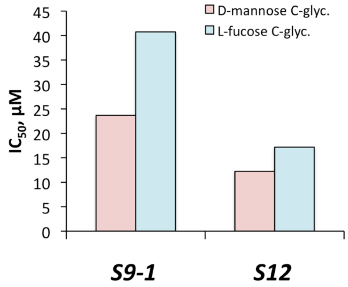


**Figure 10.4.:** The comparison of DC-SIGN inhibition activities of the corresponding multivalent compounds with C-glycosidic L-fucose and C-glycosidic D-mannose.

On the other side, when Manα1-3Man-mimicking C-glycosidic disaccharides were built (compounds ZL1(D), ZL2(L), BB207, BB213, fig. 9.6 on page 153), it appeared that their activities were even worse than the natural Manα1-3Man sugar, and one of these compounds (ZL2(L)) possessed a bizarre ability to induce DC-SIGN affinity to the dextran. Considering that even the natural Manα1-3Man disaccharide has an activity twice lower than the Manα1-2Man disaccharide, it can be concluded that it is not a good idea to design Manα1-3Man mimics as DC-SIGN antagonists.

The disaccharide of the other type, i.e. the artificial dimannoside with mannose moieties linked through C1α carbons by an alkyl chain (compound BB203, fig. 9.6 on page 153), were found to have 2 times better activity as compared to D-mannose monosaccharide. The similar result was found also for the analogous disaccharide with C-glycosidic L-fucose (compound BB178, fig. 9.6 on page 153): it had twice better activity than the corresponding C-glycosidic L-fucose monosaccharide. Thus such a 2-fold improvement of the activity for the disaccharides may result from the ligand concentration effect. The other three L-fucose C-glycosidic disaccharides (compounds BO, BO2 and BO3, fig. 9.6 on page 153) had the same activity as the corresponding C-glycosidic L-fucose monosaccharide.

Although these results give a first insight to the affinities of the C-glycosides to DC-SIGN, more importantly these compounds should be evaluated in the enzymatic digestion assay to test their resistance to glycosidases and compare it with corresponding O-glycosides.

# 11. The development of multivalent scaffolds.

Several series of multivalent scaffolds with different valences and ligand presentation in space were produced. Functionalizing them with simple monosaccharides such as D-mannose or L-fucose allowed a preliminary evaluation of the efficiencies of the scaffolds themselves (fig. 11.1).

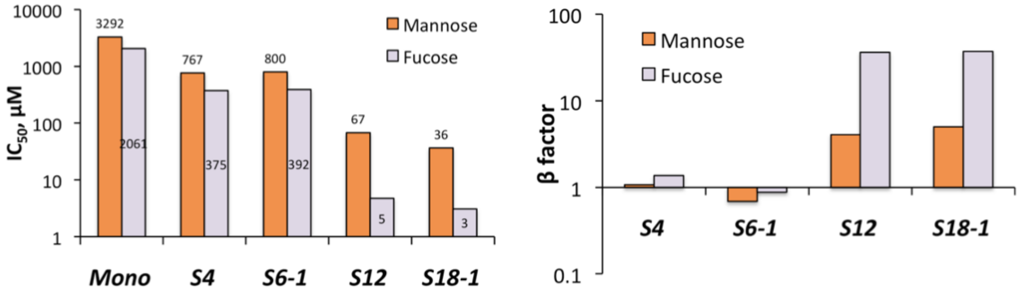


**Figure 11.1.:** The comparison of DC-SIGN inhibition activities of the corresponding multivalent compounds with O-glycosidic L-fucose and D-mannose.

Left and right panels show IC$_{50}$ and β factor values, respectively. The IC$_{50}$ values (µM) are displayed on each bar in the left panel.

Generally, the activity of the compounds increased with the increase of the scaffold valence. Fucosylated compounds had better activity at all valences than the mannose-bearing compounds, and the activity improvement factor β was much stronger for fucosylated compounds. It is interesting to notice that, for both mannose and fucose, there was no activity improvement achieved when the valence is raised from 4 to 6, and the same absence of activity increase also found changing valence from 12 to 18. This might suggest that the configuration of tetravalent scaffold is more favorable in terms of ligand presentation than that of the hexavalent one in case of these two simple O-glycosidic monosaccharide ligands, while for 12-valent and 18-valent scaffolds either the maximal efficiency with these ligands has been already reached (compared to IC$_{50}$ value of 50 µM for 32-valent 3$^{rd}$ generation Boltorn type dendrimer functionalized with mannose was found previously [158]) or also the configuration (folding) of 18-valent scaffold is not favorable for further affinity improvement as compared to 12-valent scaffold. However, no definite conclusion can be drawn from this observation because a very smooth efficiency increase with valence increase was found for dendrimers with C-glycosidic L-fucose (fig. 9.34 on page 300).

The dendrimers with glycomimetic ligands, which have higher affinity compared to natural sugars, had higher activity than the same scaffolds with natural monosaccharides. This indicates that tethering to multivalent platforms conserved their initially better affinity to DC-SIGN. Even more, this better initial affinity was further amplified by the multivalent presentation, as could be seen from $\beta > 1$ values for all of the dendrimers with the glycomimics, except psTri.

Among the glycomimetic compounds, comparing the same scaffolds the best activity and β factor values were achieved for the new lead compound **4h**, even though the Lewis$^X$ mimics **10b-azide1** displayed the same activity as **4h** at the monovalent presentation (fig. 11.2).

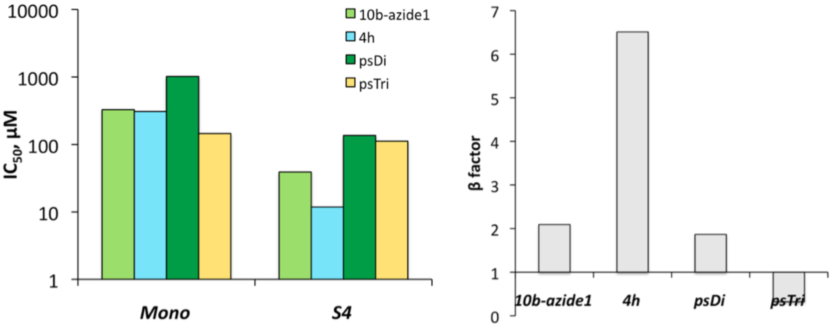


**Figure 11.2.:** The comparison of DC-SIGN inhibition activities of all glycomimetic compounds on tetravalent dendrimer scaffolds.
Left and right panels show $IC_{50}$ and β factor values, respectively.

As discussed above, for psTri multivalent presentation resulted in unfavorable β factor < 1 due to actually bivalent nature of this compound when not on the multivalent scaffold. The comparison of these tetravalent dendrimer scaffolds with previously designed tetravalent dendrons (fig. 11.3), indicated that a new type of tetravalent presentation was beneficial for the activity of psDi, while it had no effect for psTri. Again this could be related to scaffold configuration/folding, which might also be ligand-dependent.

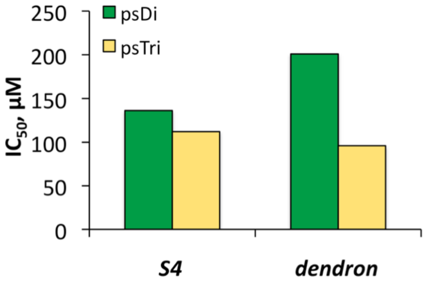


**Figure 11.3.:** The comparison of DC-SIGN inhibition activities of tetravalent psDi and psTri dendrimers and dendrons.

The most potent DC-SIGN antagonists identified up to now are the 32-valent Boltorn type dendrimers with psDi and psTri with the $IC_{50}$ values of 2 µM and 1.25 µM, respectively. They were also showed to be good inhibitors of HIV *trans* infection in a cell-based assay (paper **n°5**). However, a new lead compound **4h** had a rather close activity when presented only on a hexavalent dendrimer scaffold, giving a promising perspective to acquire even higher potencies with higher valence scaffolds. Morever, the Boltorn type dendrimers are built based on ester bonds, which are less stable than ether and triazole bonds in new type dendrimers, which is one more advantage to further develop this new type multivalent platforms.

The only drawback observed for some of these dendrimers was their low solubility, namely for compounds ***S4-Oman***, ***S4-psTri***, ***S4-fucβAla***, ***S6-1-Oman***, ***S9-1-Cfuc***, and for all **4h** functionalized compounds, in particular rod-like dendrimer ***R3***. This limited solubility first of all affects the accurate determination of the $IC_{50}$ values due to overestimated compound concentration, which is why the further increase of **4h** activity on a 9-valent scaffold was not observed (fig. 9.31 on page 298). The limited solubility may also be related with the anticipated aggregation phenomena observed in SPR direct interaction assay (it must be noted that the compounds before SPR experiments were centrifuged, and only soluble part of the samples was analyzed). However, the aggregation was observed also for those compounds (***S4-psDi*** and ***S6-1-psDi***), which did not have any apparent water solubility problems. All these presumably aggregating compounds comprised all of the tetra- and hexavalent dendrimers, in addition to all **4h**-functionalized dendrimers. For the lower valence compounds it could be that the aggregation is driven by the scaffold folding in a way, where hydrophobic parts become exposed and available for formation of the aggregates. Whereas for **4h** multivalent compounds the nature of the ligand, i.e. probably the presence of two aromatic rings on each ligand molecule, also plays a role in aggregation process, and for this reason the solubility of the latter compounds is particularly lower. Such aggregation phenomenon most likely has also an effect on compound analysis in SPR competition assay, as described below.

The second presumable phenomenon observed in SPR direct interaction assay was the ability of some compounds to interact with DC-SIGN in $Ca^{2+}$-independent manner. These compounds included:

- dendrons with psDi and psTri since EDTA failed to fully remove them from DC-SIGN surface (fig. 9.41 on page 307)
- ***S6-1-4h*** and ***S4-fucβAla*** as seen in surface double regeneration assay (fig. 9.57 on page 323, EDTA also didn't fully remove these compounds from DC-SIGN)
- compounds ***S18-1-Oman***, ***S18-2-Oman***, ***S4-Oman*** and ***S4-4h*** can be predicted to bind DC-SIGN in other than $Ca^{2+}$-dependent sugar recognition site from the observation that after using the regeneration solution (composed of 50 mM Gly-NaOH pH 11.9 / 0.15% TritonX100 / 25 mM EDTA pH 8) still part of the compounds stayed bound on DC-SIGN surface (data not shown).

Possibly, the initial binding of these compounds occurs through $Ca^{2+}$-dependent interaction in the conventional sugar binding site of DC-SIGN, but then is stabilized by additional, probably hydrophobic, compound-protein interactions. Since the double regeneration assay was not performed with all of the tested dendrimers, it is possible that there were more compounds binding to DC-SIGN in $Ca^{2+}$-independent manner.

Finally, the third unexpected behavior of some of the dendrimers (i.e. compounds ***S6-1-Oman***, ***S6-1-psTri***, ***S6-2-psDi***, and ***S9-1-psDi***) was an apparently higher affinity to StrepTactin than to DC-SIGN. The latter effect, together with ability of some other compounds to interact in $Ca^{2+}$-independent manner, should be taken with caution considering cellular assays since such unspecific compound-protein interaction might have an effect to the course and the results of these assays, where many different protein receptors are available for interaction.

All of these observed phenomena for the compounds are summarized in table 12.2.

# 12. Does the shape of the inhibition curve indicate “the mode of action” of the compound?

After performing multiple SPR competition experiments, I have noticed that the shapes of the inhibition curves, in terms of the slope, for some compounds vary significantly. It is obvious that the same $IC_{50}$ value may be achieved from the curves with different slopes as shown in figure 12.1.

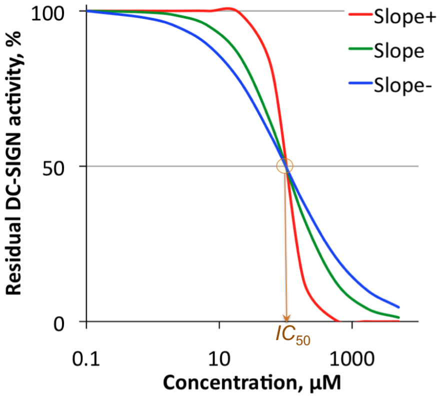


**Figure 12.1.:** The comparison of three inhibition curves with different slopes, for which $IC_{50}$ values are the same.

It raised the question whether the slope of the curve could indicate a particular mechanism of how the compound interacts with DC-SIGN. One of the first observations was that for psTri the slope values were always lower than for the majority of other compounds. When it was found that psTri is capable to cross-link DC-SIGN tetramers, it gave an idea how to interpret such a lower slope value, which is illustrated in the middle panel of figure 12.2 and compared to a “normal” non-bridging compound (figure 12.2, left panel). At low psTri concentrations (the top of the curve), DC-SIGN tetramers are in excess, which makes bridging favorable. At this point, psTri displays an activity characteristic so higher affinity compounds, which in turn forces the inhibition curve to start to incline at low psTri concentrations. Increasing psTri concentration further, the binding sites of DC-SIGN gradually become saturated, which makes bridging less and less favored, and this way affects the appearance of the curve to have a milder slope than for the “normal” compound.

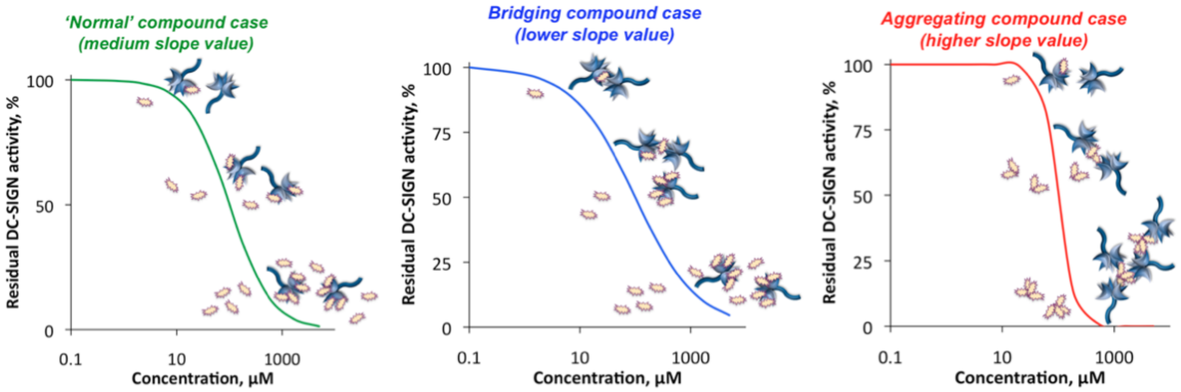


**Figure 12.2.:** Possible mechanisms of how compounds inhibit DC-SIGN in the SPR competition assay.

The curves with the evolutions of compound and DC-SIGN states for the three indicated cases are shown. The compound evolution is illustrated on the left, and the change of DC-SIGN state is shown on the right of each curve. A compound, which neither cross-links DC-SIGN nor aggregates, is referred as "normal".

However, for some of the compounds the slope value was obviously higher than for the "normal" compounds. Thinking of this case, I considered the following situation (fig. 12.2, right panel): since the slope is steep, it might suggest that the compound exhibits higher affinity at higher concentrations. This could be possible only if the compound at higher concentrations forms aggregates, which display more ligands on their surface in a favorable manner for DC-SIGN binding. This way the clustering of even more DC-SIGN tetramers becomes possible. In competition assay the binding of DC-SIGN is measured, so such an aggregation of the compounds at higher concentrations would reduce the amount of free DC-SIGN more efficiently forcing the steepness of the curve.

I have selected six monovalent compounds (L-fucose, Lewis$^{X}$, Manα1-2Man, Manα1-3Man, psDi and D-mannose), which are expected to behave "normally", i.e. they don't bridge DC-SIGN nor they aggregate. I calculated the mean slope (parameter $A_2$ in eq. 6.1 on page 114) and SD values from multiple SPR competition experiments (table 12.1).

**Table 12.1.:** The determination of upper and lower A2 limits for the "normal" compound.

| | **Mean $A_2$** | **SD** | | |
|---|---|---|---|---|
| D-mannose | 1.24 | 0.17 | | |
| Manα1-2Man | 1.17 | 0.11 | | |
| Manα1-3Man | 1.03 | 0.06 | **0.97** | **←lower limit** |
| psDi | 1.16 | 0.15 | | |
| L-fucose | 1.25 | 0.18 | **1.43** | **←upper limit** |
| Lewis$^{X}$ | 1.21 | 0.15 | | |

By subtracting the SD value from the lowest observed $A_2$ (for Manα1-3Man) and by adding SD value to the highest observed $A_2$ (for L-fucose), I defined the approximate lower and upper slope limits for the "normal" mode of inhibition: the compounds that lie below the lower limit would be likely to cross-link DC-SIGN tetramers, while the compounds that go above the upper limit are probably aggregating, and all others that are in-between the two limits are "normal".

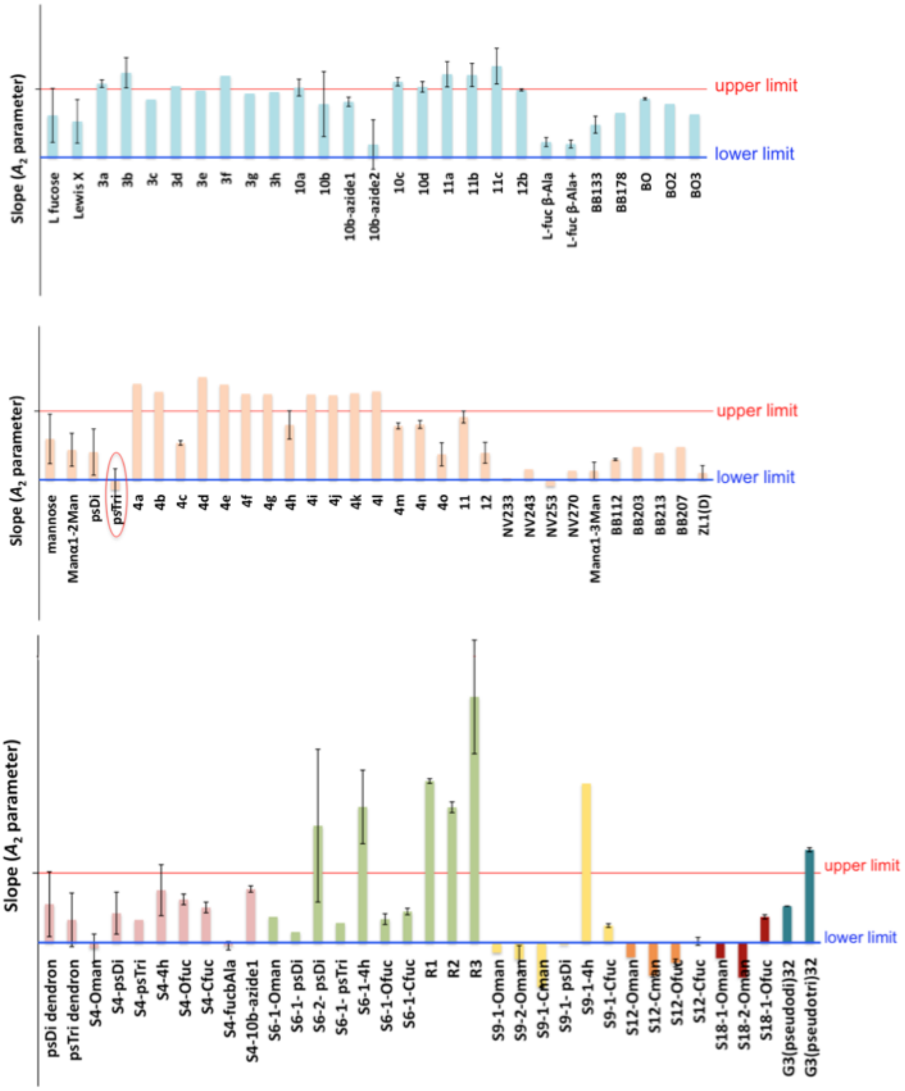


**Figure 12.3.:** The analysis of all the tested compounds in terms of the slope of the inhibition curve. Fucose-based, mannose-based and all multivalent compounds are shown in top, middle and bottom panels, respectively.

Such an analysis of all the tested compounds in terms of the inhibition curve slope value showed that all fucose-based (fig. 12.3 top panel) and most of the mannose-based compounds (fig. 12.3 middle panel) behave in a “normal” mode. Among mannose-based compound series, psTri was found to be below the lower limit indicating its ability to cross-link DC-SIGN. The second compound slightly below the lower limit was NV253. However, this compound was tested only once and considering relatively high variation of slope values, it behaves most likely like a “normal” compound. Similarly, there were several mannose-based compounds above the upper limit, which were measured also only once, thus possibly behaving as “normal” ones as well (although the lower solubility displayed by **4a** and **4d** might be related with aggregation).

A very different result is seen for multivalent compounds (fig. 12.3 bottom panel): part of them (***S6-2-psDi***, ***S6-1-4h***, ***S9-1-4h***, ***R1***, ***R***2, ***R3*** and G3(pseudotri)32) demonstrated a significant sign of aggregation, which correlates well with SPR direct interaction assay. The others (***S4-Oman***, ***S9-1-Oman***, ***S9-2-Oman***, ***S9-1-Cman***, ***S9-1-psDi***, ***S12-Oman***, ***S12-Cman***,

***S12-Ofuc***, ***S4-fucβAla***, ***S18-1-Oman*** and ***S18-2-Oman***) displayed the ability to cross-link DC-SIGN, an expected behavior for the multivalent compounds. The rest of the compounds were in-between the two limits. It may suggest the "normal" mode of their interaction with DC-SIGN, however, considering that multivalent compounds are likely to bridge DC-SIGN, these observed slope values between the limits might likely result from the cross-linking effect compensated by aggregation effect, and thus the slopes do not cross either limit.

**Table 12.2.:** The summary of all observed and assumed characteristics of all multivalent compounds.

| Compound | Solubility problem | Aggregation-1* | Aggregation-2** | $Ca^{2+}$-independent interaction | Interaction with StrepTactin |
|---|---|---|---|---|---|
| ***S4-Oman*** | yes | yes | no | yes | no |
| ***S6-1-Oman*** | yes | - | ? | *NA* | yes |
| ***S9-1-Oman*** | no | no | no | *NA* | no |
| ***S9-2-Oman*** | no | no | no | *NA* | no |
| ***S12-Oman*** | no | no | no | *NA* | no |
| ***S18-1-Oman*** | no | no | no | yes | no |
| ***S18-2-Oman*** | no | no | no | yes | no |
| ***S9-1-Cman*** | no | *NA* | no | *NA* | *NA* |
| ***S12-Cman*** | no | *NA* | no | *NA* | *NA* |
| ***S4-psDi*** | no | yes | ? | no | no |
| ***S6-1- psDi*** | no | yes | ? | *NA* | no |
| ***S6-2- psDi*** | no | - | yes | *NA* | yes |
| ***S9-1- psDi*** | no | - | no | *NA* | yes |
| ***S4-psTri*** | yes | yes | ? | *NA* | no |
| ***S6-1- psTri*** | no | - | ? | *NA* | yes |
| ***S4-4h*** | yes | yes | ? | yes | no |
| ***S6-1-4h*** | yes | yes | yes | yes | no |
| ***S9-1-4h*** | yes | yes | yes | no | no |
| ***R1*** | yes | yes | yes | no | no |
| ***R2*** | yes | yes | yes | no | no |
| ***R3*** | yes | yes | yes | no | no |
| ***S4-Ofuc*** | no | *NA* | ? | *NA* | *NA* |
| ***S6-1-Ofuc*** | no | *NA* | ? | *NA* | *NA* |
| ***S12-Ofuc*** | no | *NA* | no | *NA* | *NA* |
| ***S18-1-Ofuc*** | no | *NA* | ? | *NA* | *NA* |
| ***S4-Cfuc*** | no | *NA* | ? | *NA* | *NA* |
| ***S6-1-Cfuc*** | no | *NA* | ? | *NA* | *NA* |
| ***S9-1-Cfuc*** | yes | *NA* | ? | *NA* | *NA* |
| ***S12-Cfuc*** | no | *NA* | ? | *NA* | *NA* |
| ***S4-fucβAla*** | yes | no | no | yes | no |
| ***S4-10b-azide1*** | no | - | ? | no | no |
| ***G3(pseudodi)***$_{32}$ | no | no | ? | no | no |
| ***G3(pseudotri)***$_{32}$ | no | yes | yes | no | no |
| ***psDi dendron*** | no | no | ? | yes | no |
| ***psTri dendron*** | no | no | ? | yes | no |

* assumed from disproportional signal increase with concentration change in SPR direct interaction assay;

** assumed from slope evaluation; the "-" means the result was not available for the corresponding compound, the "?" is an uncertainty of the results, and *NA* stands for "not analyzed".

Although this type of analysis may give a clue, it does not provide a definite answer about what is really happening in solution. The other biophysical assays, such as DLS and/or AUC, should be performed on the compounds alone in order to track and characterize their aggregation. Performing the same experiment in the presence of DC-SIGN would indicate which of the compounds are able to cross-link DC-SIGN. These experiments would also answer the question if the higher maximal binding responses observed in SPR direct interaction assay indicate compound aggregation or binding to other sites on DC-SIGN.

In summary, the low solubility of the compounds is problematic for both performing the assay and interpreting the results. It could be a good idea to determine water solubility of the compounds and provide as a guideline in order to avoid the misevaluation of their activities. Moreover and most importantly, the solubility problem is a drawback for development of the compounds as the drug substances. When tethered to multivalent scaffolds, a new perspective ligand with combined features of **4h** and NV243 (fig. 10.2) could be predicted to induce the electrostatic repulsion between the molecules due to the presence of positively charged amino group, and lead to both increased solubility and reduced aggregation.

In general, the strategy of using multivalent glycomimetic compounds for protection from HIV *trans* infection has been proved promising (article **n°2**), and the further improvement of the scaffolds may possibly lead to a successful drug substance discovery for the development of vaginal microbicides.

The closest way to achieve that may be envisioned: after synthesizing and tethering on multivalent scaffolds a new ligand with combined **4h** and NV243 features, it should be analyzed in proposed biophysical assays (DLS, AUC and SPR) to control its behavior in solution and activity. In case of positive results, it should be tested for HIV *trans* infection inhibition in cell-based or *ex vivo* assays together with toxicity determination. The good results of the latter tests would pave a way for such compound to be tested in preclinical animal studies. Only the confirmed positive tendency of the resistance to infection in monkeys may approve the further investigation of this molecule in clinical trials.

Finally, even if the use of such molecule for protection from HIV infection wasn't approved, it may find other applications in inhibiting infections and/or blocking the hijack of DCs by other pathogens. This type of molecule may also be used for the development of synthetic vaccines to target DCs specifically through the interaction with DC-SIGN.

# Part V.

# Bibliography

# Bibliography

[1] He, J. & Deem, M. W. (2010) Heterogeneous diversity of spacers within CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). *Physical Review Letters* **105**, 128102.

[2] Rowley, A. F. & Powell, A. (2007) Invertebrate immune systems specific, quasi-specific, or nonspecific? *Journal of Immunology* **179**, 7209–7214.

[3] Goldsby, R.; Kindt, T.; Osborne, B.; & Kuby, J. Book: *Immunology. 5th edition. W. H. Freeman*. 2002.

[4] Parham, P. Elements of the immune system and their roles in defense. In book: *The Immune System. 3rd edition. Garland Science*. 2009.

[5] Miller, L. S & Cho, J. S. (2011) Immunity against *Staphylococcus aureus* cutaneous infections. *Nature Reviews Immunology* **11**, 505–518.

[6] Murphy, K.; Geha, R. & Notarangelo, L. Innate immunity: The first lines of defense. In book: *Janeway's Immunobiology. 8th edition. Garland Science.* 2011.

[7] Pathak, S. & Palan, U. Book: *Immunology: Essential And Fundamental. 2nd edition. Science Publishers.* 2005.

[8] DeFranco, A.; Locksley, R. & Robertson, M. Book: *Immunity: The Immune Response in Infectious and Inflammatory Disease. New Science Press*. 2007.

[9] Paul, W. E. Book: *Fundamental Immunology. 6th edition*. *Lippincott Williams & Wilkins.* 2008.

[10] Bedard, K. & Krause, K.-H. (2007) The NOX family of ROS-generating NADPH oxidases: physiology and pathophysiology. *Physiological Reviews* **87**, 245–313.

[11] Janeway, C. A. & Medzhitov, R. (2002) Innate immune recognition. *Annual Review of Immunology* **20**, 197–216.

[12] Kishore, U. Book: *Target pattern recognition in innate immunity. Springer.* 2009.

[13] Peiser, L.; Mukhopadhyay, S. & Gordon, S. (2002) Scavenger receptors in innate immunity. *Current Opinion in Immunology* **14**, 123–128.

[14] Lombardi, G. & Riffo-Vasquez, Y. Book: *Dendritic Cells: Handbook of Experimental Pharmacology. Springer*. 2008.

[15] Elgert, K. D. Book: *Immunology: Understanding the immune system. 2nd edition. John Wiley & Sons.* 2009.

[16] Chen, G.; Shaw, M. H.; Kim, Y.-G. & Nuñez, G. (2009) NOD-like receptors: role in innate immunity and inflammatory disease. *Annual Review of Pathology: Mechanisms of Disease* **4**, 365–398.

[17] Zelensky, A. N. & Gready, J. E. (2005) The C-type lectin-like domain superfamily. *The FEBS journal* **272**, 6179–6217.

[18] van den Berg, L. M.; Gringhuis, S. I. & Geijtenbeek, T. B. H. (2012) An evolutionary per- spective on C-type lectins in infection and immunity. *Annals of the New York Academy of Sciences*. doi: 10.1111/j.1749-6632.2011.06392.x

[19] Geijtenbeek, T. B. H. & Gringhuis, S. I. (2009) Signalling through C-type lectin receptors: shaping immune responses. *Nature Reviews Immunology* **9**, 465–479.

[20] Redelinghuys, P. & Brown, G. D. (2011) Inhibitory C-type lectin receptors in myeloid cells. *Immunology Letters* **136**, 1–12.

[21] de Jong, M. A. W. P.; Vriend, L. E. M.; Theelen, B.; Taylor, M. E.; Fluitsma, D.; Boekhout, T. & Geijtenbeek, T. B. H. (2010) C-type lectin langerin is a β-glucan receptor on human Langerhans cells that recognizes opportunistic and pathogenic fungi. *Molecular Immunology* **47**, 1216– 1225.

[22] Stoitzner, P. & Romani, N. (2011) Langerin, the "catcher in the rye": an important receptor for pathogens on Langerhans cells. *European Journal of Immunology* **41**, 2526–2529.

[23] Bashirova, A. A.; Geijtenbeek, T. B.; van Duijnhoven, G. C.; van Vliet, S. J.; Eilering, J. B.; Martin, M. P.; Wu, L.; Martin, T. D.; Viebig, N.; Knolle, P. A.; KewalRamani, V. N.; van Kooyk, Y. & Carrington, M. (2001) 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. *Journal of Experimental Medicine* **193**, 671–678.

[24] Snyder, G. A.; Ford, J.; Torabi-Parizi, P.; Arthos, J. A.; Schuck, P.; Colonna, M. & Sun, P. D. (2005) Characterization of DC-SIGN/R interaction with human immunodeficiency virus type 1 gp120 and ICAM molecules favors the receptor's role as an antigen-capturing rather than an adhesion receptor. *Journal of Virology* **79**, 4589–4598.

[25] Yabe, R.; Tateno, H. & Hirabayashi, J. (2010) Frontal affinity chromatography analysis of constructs of DC-SIGN, DC-SIGNR and LSECtin extend evidence for affinity to agalactosylated N-glycans. *FEBS J* **277**, 4010–4026.

[26] Kolatkar, A. R.; Leung, A. K.; Isecke, R.; Brossmer, R.; Drickamer, K. & Weis, W. I. (1998) Mechanism of N-acetylgalactosamine binding to a C-type animal lectin carbohydrate-recognition domain. *Journal of Biological Chemistry* **273**, 19502–19508.

[27] Williams, A. Book: *Immunology: Mucosal and Body Surface Defences. 2nd edition. John Wiley & Sons*. 2011.

[28] Gessani, S. & Belardelli, F. Book: *The Biology of Dendritic Cells and HIV Infection. Springer*. 2007.

[29] Geijtenbeek, T. B. H.; den Dunnen, J. & Gringhuis, S. I. (2009) Pathogen recognition by DC-SIGN shapes adaptive immunity. *Future Microbiology* **4**, 879–890.

[30] Bleijs, D. A.; Geijtenbeek, T. B.; Figdor, C. G. & van Kooyk, Y. (2001) DC-SIGN and LFA-1: a battle for ligand. *Trends in Immunology* **22**, 457–463.

[31] Dustin, M. L.; Tseng, S.-Y.; Varma, R. & Campi, G. (2006) T cell-dendritic cell immunological synapses. *Current Opinion in Immunology* **18**, 512–516.

[32] Wilen, C. B.; Tilton, J. C. & Doms, R. W. (2012) HIV: Cell binding and entry. *Cold Spring Harbor Perspectives in Medicine* **2**. doi: 10.1101/cshperspect.a006866

[33] Saphire, A. C.; Bobardt, M. D.; Zhang, Z.; David, G. & Gallay, P. A. (2001) Syndecans serve as attachment receptors for human immunodeficiency virus type 1 on macrophages. *Journal of Virology* **75**, 9187–9200.

[34] Arthos, J.; Cicala, C.; Martinelli, E.; Macleod, K.; Ryk, D. V.; Wei, D.; Xiao, Z.; Veenstra, T. D.; Conrad, T. P.; Lempicki, R. A.; McLaughlin, S.; Pascuccio, M.; Gopaul, R.; Mcnally, J.; Cruz, C. C.; Censoplano, N.; Chung, E.; Reitano, K. N.; Kottilil, S.; Goode, D. J. & Fauci, A. S. (2008) HIV-1 envelope protein binds to and signals through integrin $\alpha_4\beta_7$, the gut mucosal homing receptor for peripheral T cells. *Nature Immunology* **9**, 301–309.

[35] Cicala, C.; Martinelli, E.; McNally, J. P.; Goode, D. J.; Gopaul, R.; Hiatt, J.; Jelicic, K.; Kottilil, S.; Macleod, K.; O'Shea, A.; Patel, N.; Ryk, D. V.; Wei, D.; Pascuccio, M.; Yi, L.; McKinnon, L.; Izulla, P.; Kimani, J.; Kaul, R.; Fauci, A. S. & Arthos, J. (2009) The integrin $\alpha_4\beta_7$ forms a complex with cell-surface CD4 and defines a T-cell subset that is highly susceptible to infection by HIV-1. *PNAS* **106**, 20877–20882.

[36] 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. & 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.

[37] Orloff, G. M.; Orloff, S. L.; Kennedy, M. S.; Maddon, P. J. & McDougal, J. S. (1991) Penetration of CD4 T cells by HIV-1. The CD4 receptor does not internalize with HIV, and CD4-related signal transduction events are not required for entry. *Journal of Immunology* **146**, 2578–2587.

[38] Sherer, N. M.; Jin, J. & Mothes, W. (2010) Directional spread of surface-associated retroviruses regulated by differential virus-cell interactions. *Journal of Virology* **84**, 3248–3258.

[39] Miyauchi, K.; Kim, Y.; Latinovic, O.; Morozov, V. & Melikyan, G. B. (2009) HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. *Cell* **137**, 433–444.

[40] Melikyan, G. B. (2008) Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. *Retrovirology* **5**. doi:10.1186/1742-4690-5-111

[41] Cameron, P. U.; Freudenthal, P. S.; Barker, J. M.; Gezelter, S.; Inaba, K. & Steinman, R. M. (1992) Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to $CD4^+$ T cells. *Science* **257**, 383–387.

[42] Pope, M.; Betjes, M. G.; Romani, N.; Hirmand, H.; Cameron, P. U.; Hoffman, L.; Gezelter, S.; Schuler, G. & Steinman, R. M. (1994) Conjugates of dendritic cells and memory T lymphocytes from skin facilitate productive infection with HIV-1. *Cell* **78**, 389–398.

[43] Hladik, F. & McElrath, M. J. (2008) Setting the stage: host invasion by HIV. *Nature Reviews Immunology* **8**, 447–457.

[44] Hladik, F. & Hope, T. J. (2009) HIV infection of the genital mucosa in women. *Current HIV/AIDS Reports* **6**, 20–28.

[45] Felts, R. L.; Narayan, K.; Estes, J. D.; Shi, D.; Trubey, C. M.; Fu, J.; Hartnell, L. M.; Ruthel, G. T.; Schneider, D. K.; Nagashima, K.; Bess, J. W.; Bavari, S.; Lowekamp, B. C.; Bliss, D.; Lifson, J. D. & Subramaniam, S. (2010) 3D visualization of HIV transfer at the virological synapse between dendritic cells and T cells. *PNAS* **107**, 13336–13341.

[46] Kwon, D. S.; Gregorio, G.; Bitton, N.; Hendrickson, W. A. & Littman, D. R. (2002) DC-SIGN-mediated internalization of HIV is required for *trans*-enhancement of T cell infection. *Immunity* **16**, 135–144.

[47] Sol-Foulon, N.; Moris, A.; Nobile, C.; Boccaccio, C.; Engering, A.; Abastado, J.-P.; Heard, J.-M.; van Kooyk, Y. & Schwartz, O. (2002) HIV-1 Nef-induced upregulation of DC-SIGN in dendritic cells promotes lymphocyte clustering and viral spread. *Immunity* **16**, 145–155.

[48] Gurney, K. B.; Elliott, J.; Nassanian, H.; Song, C.; Soilleux, E.; McGowan, I.; Anton, P. A. & Lee, B. (2005) Binding and transfer of human immunodeficiency virus by DC-SIGN$^{+}$ cells in human rectal mucosa. *Journal of Virology* **79**, 5762–5773.

[49] Sattin, S.; Daghetti, A.; Thépaut, M.; Berzi, A.; Sánchez-Navarro, M.; Tabarani, G.; Rojo, J.; Fieschi, F.; Clerici, M. & Bernardi, A. (2010) Inhibition of DC-SIGN-mediated HIV infection by a linear trimannoside mimic in a tetravalent presentation. *ACS Chemical Biology* **5**, 301–312.

[50] Berzi, A.; Reina, J. J.; Ottria, R.; Sutkeviciute, I.; Antonazzo, P.; Sanchez-Navarro, M.; Chabrol, E.; Biasin, M.; Trabattoni, D.; Cetin, I.; Rojo, J.; Fieschi, F.; Bernardi, A. & Clerici, M. (2012) A glycomimetic compound inhibits DC-SIGN-mediated HIV infection in cellular and cervical explant models. *AIDS (London, England)* **26**, 127–137.

[51] de Witte, L.; Nabatov, A. & Geijtenbeek, T. B. H. (2008) Distinct roles for DC-SIGN$^{+}$-dendritic cells and Langerhans cells in HIV-1 transmission. *Trends in Molecular Medicine* **14**, 12–19.

[52] Burleigh, L.; Lozach, P.-Y.; Schiffer, C.; Staropoli, I.; Pezo, V.; Porrot, F.; Canque, B.; Virelizier, J.-L.; Arenzana-Seisdedos, F. & Amara, A. (2006) Infection of dendritic cells (DCs), not DC-SIGN-mediated internalization of human immunodeficiency virus, is required for long-term transfer of virus to T cells. *Journal of Virology* **80**, 2949–2957.

[53] Gringhuis, S. I.; van der Vlist, M.; van den Berg, L. M.; den Dunnen, J.; Litjens, M. & Geijtenbeek, T. B. H. (2010) HIV-1 exploits innate signaling by TLR8 and DC-SIGN for productive infection of dendritic cells. *Nature Immunology* **11**, 419–426.

[54] Hladik, F.; Sakchalathorn, P.; Ballweber, L.; Lentz, G.; Fialkow, M.; Eschenbach, D. & McElrath, M. J. (2007) Initial events in establishing vaginal entry and infection by human immunodeficiency virus type-1. *Immunity* **26**, 257–270.

[55] Turville, S. G.; Cameron, P. U.; Handley, A.; Lin, G.; Pöhlmann, S.; Doms, R. W. & Cunningham, A. L. (2002) Diversity of receptors binding HIV on dendritic cell subsets. *Nature Immunology* **3**, 975– 983.

[56] Hussain, L. A. & Lehner, T. (1995) Comparative investigation of Langerhans' cells and potential receptors for HIV in oral, genitourinary and rectal epithelia. *Immunology* **85**, 475–484.

[57] de Witte, L.; Nabatov, A.; Pion, M.; Fluitsma, D.; de Jong, M. A. W. P.; de Gruijl, T.; Piguet, V.; van Kooyk, Y. & Geijtenbeek, T. B. H. (2007) Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nature Medicine* **13**, 367–371.

[58] de Jong, M. A. W. P.; de Witte, L.; Taylor, M. E. & Geijtenbeek, T. B. H. (2010) Herpes simplex virus type 2 enhances HIV-1 susceptibility by affecting Langerhans cell function. *Journal of Immunology* **185**, 1633–1641.

[59] Ballweber, L.; Robinson, B.; Kreger, A.; Fialkow, M.; Lentz, G.; McElrath, M. J. & Hladik, F. (2011) Vaginal Langerhans cells nonproductively transporting HIV-1 mediate infection of T cells. *Journal of Virology* **85**, 13443–13447.

[60] Chatwell, L.; Holla, A.; Kaufer, B. B. & Skerra, A. (2008) The carbohydrate recognition domain of langerin reveals high structural similarity with the one of DC-SIGN but an additional, calcium-independent sugar-binding site. *Molecular Immunology* **45**, 1981–1994.

[61] Curtis, B. M.; Scharnowske, S. & Watson, A. J. (1992) Sequence and expression of a membrane- associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120. *PNAS* **89**, 8356–8360.

[62] Geijtenbeek, T. B.; Torensma, R.; van Vliet, S. J.; van Duijnhoven, G. C.; Adema, G. J.; van Kooyk, Y. & Figdor, C. G. (2000) Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell* **100**, 575–585.

[63] Granelli-Piperno, A.; Pritsker, A.; Pack, M.; Shimeliovich, I.; Arrighi, J.-F.; Park, C. G.; Trumpfheller, C.; Piguet, V.; Moran, T. M. & Steinman, R. M. (2005) Dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin/CD209 is abundant on macrophages in the normal human lymph node and is not required for dendritic cell stimulation of the mixed leukocyte reaction. *Journal of Immunology* **175**, 4265–4273.

[64] Lee, B.; Leslie, G.; Soilleux, E.; O'doherty, U.; Baik, S.; Levroney, E.; Flummerfelt, K.; Swiggard, W.; Coleman, N.; Malim, M. & Doms, R. W. (2001) *cis* expression of DC-SIGN allows for more efficient entry of human and simian immunodeficiency viruses via CD4 and a coreceptor. *Journal of Virology* **75**, 12028–12038.

[65] Soilleux, E. J.; Morris, L. S.; Leslie, G.; Chehimi, J.; Luo, Q.; Levroney, E.; Trowsdale, J.; Montaner, L. J.; Doms, R. W.; Weissman, D.; Coleman, N. & Lee, B. (2002) Constitutive and

induced expression of DC-SIGN on dendritic cell and macrophage subpopulations *in situ* and *in vitro*. *Journal of Leukocyte Biology* **71**, 445–457.

[66] Krutzik, S. R.; Tan, B.; Li, H.; Ochoa, M. T.; Liu, P. T.; Sharfstein, S. E.; Graeber, T. G.; Sieling, P. A.; Liu, Y.-J.; Rea, T. H.; Bloom, B. R. & Modlin, R. L. (2005) TLR activation triggers the rapid differentiation of monocytes into macrophages and dendritic cells. *Nature Medicine* **11**, 653–660.

[67] Rappocciolo, G.; Piazza, P.; Fuller, C. L.; Reinhart, T. A.; Watkins, S. C.; Rowe, D. T.; Jais, M.; Gupta, P. & Rinaldo, C. R. (2006) DC-SIGN on B lymphocytes is required for transmission of HIV-1 to T lymphocytes. *PLoS Pathogenes* **2**, 691-704.

[68] Mummidi, S.; Catano, G.; Lam, L.; Hoefle, A.; Telles, V.; Begum, K.; Jimenez, F.; Ahuja, S. S. & Ahuja, S. K. (2001) Extensive repertoire of membrane-bound and soluble dendritic cell-specific ICAM-3-grabbing nonintegrin 1 (DC-SIGN1) and DC-SIGN2 isoforms. Inter-individual variation in expression of DC-SIGN transcripts. *Journal of Biological Chemistry* **276**, 33196–33212.

[69] Martinez, O.; Brackenridge, S.; El-Idrissi, M. E.-A. & Prabhakar, B. S. (2005) DC-SIGN, but not sDC-SIGN, can modulate IL-2 production from PMA- and anti-CD3-stimulated primary human CD4 T cells. *International Immunology* **17**, 769–778.

[70] Plazolles, N.; Humbert, J.-M.; Vachot, L.; Verrier, B.; Hocke, C. & Halary, F. (2011) Pivotal advance: The promotion of soluble DC-SIGN release by inflammatory signals and its enhancement of cytomegalovirus-mediated *cis*-infection of myeloid dendritic cells. *Journal of Leukocyte Biology* **89**, 329–342.

[71] Soilleux, E. J.; Barten, R. & Trowsdale, J. (2000) DC-SIGN; a related gene, DC-SIGNR; and CD23 form a cluster on 19p13. *Journal of Immunology* **165**, 2937–2942.

[72] Svajger, U.; Anderluh, M.; Jeras, M. & Obermajer, N. (2010) C-type lectin DC-SIGN: an adhesion, signalling and antigen-uptake molecule that guides dendritic cells in immunity. *Cellular Signalling* **22**, 1397–1405.

[73] Feinberg, H.; Mitchell, D. A.; Drickamer, K. & Weis, W. I. (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* **294**, 2163–2166.

[74] Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I. & Drickamer, K. (2004) Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nature Structural & Molecular Biology* **11**, 591–598.

[75] Mitchell, D. A.; Fadden, A. J. & Drickamer, K. (2001) A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. *Journal of Biological Chemistry* **276**, 28939–28945.

[76] Frison, N.; Taylor, M. E.; Soilleux, E.; Bousser, M.-T.; Mayer, R.; Monsigny, M.; Drickamer, K. & Roche, A.-C. (2003) Oligolysine-based oligosaccharide clusters: selective recognition and endocytosis by the mannose receptor and dendritic cell-specific intercellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin. *Journal of Biological Chemistry* **278**, 23922–23929.

[77] Dam, T. K. & Brewer, C. F. (2010) Lectins as pattern recognition molecules: the effects of epitope density in innate immunity. *Glycobiology* **20**, 270–279.

[78] Lupas, A. (1996) Coiled coils: new structures and new functions. *Trends in Biochemical Sciences* **21**, 375–382.

[79] Feinberg, H.; Guo, Y.; Mitchell, D. A.; Drickamer, K. & Weis, W. I. (2005) Extended neck regions stabilize tetramers of the receptors DC-SIGN and DC-SIGNR. *Journal of Biological Chemistry* **280**, 1327–1335.

[80] Tabarani, G.; Thépaut, M.; Stroebel, D.; Ebel, C.; Vivès, C.; Vachette, P.; Durand, D. & Fieschi, F. (2009) DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain. *Journal of Biological Chemistry* **284**, 21229–21240.

[81] Menon, S.; Rosenberg, K.; Graham, S. A.; Ward, E. M.; Taylor, M. E.; Drickamer, K. & Leckband, D. E. (2009) Binding-site geometry and flexibility in DC-SIGN demonstrated with surface force measurements. *PNAS* **106**, 11524–11529.

[82] Cambi, A.; de Lange, F.; van Maarseveen, N. M.; Nijhuis, M.; Joosten, B.; van Dijk, E. M. H. P.; de Bakker, B. I.; Fransen, J. A. M.; Bovee-Geurts, P. H. M.; van Leeuwen, F. N.; Hulst, N. F. V. & Figdor, C. G. (2004) Microdomains of the C-type lectin DC-SIGN are portals for virus entry into dendritic cells. *Journal of Cell Biology* **164**, 145–155.

[83] de Bakker, B. I.; de Lange, F.; Cambi, A.; Korterik, J. P.; van Dijk, E. M. H. P.; Hulst, N. F. V.; Figdor, C. G. & Garcia-Parajo, M. F. (2007) Nanoscale organization of the pathogen receptor DC-SIGN mapped by single-molecule high-resolution fluorescence microscopy. *European Journal of Chemical Physics and Physical Chemistry* **8**, 1473–1480.

[84] Koopman, M.; Cambi, A.; de Bakker, B. I.; Joosten, B.; Figdor, C. G.; van Hulst, N. F. & Garcia-Parajo, M. F. (2004) Near-field scanning optical microscopy in liquid for high resolution single molecule detection on dendritic cells. *FEBS Letters* **573**, 6–10.

[85] Neumann, A. K.; Thompson, N. L. & Jacobson, K. (2008) Distribution and lateral mobility of DC-SIGN on immature dendritic cells–implications for pathogen uptake. *Journal of Cell Science* **121**, 634–643.

[86] Itano, M. S.; Neumann, A. K.; Liu, P.; Zhang, F.; Gratton, E.; Parak, W. J.; Thompson, N. L. & Jacobson, K. (2011) DC-SIGN and influenza hemagglutinin dynamics in plasma membrane microdomains are markedly different. *Biophysical Journal* **100**, 2662–2670.

[87] Itano, M. S.; Steinhauer, C.; Schmied, J. J.; Forthmann, C.; Liu, P.; Neumann, A. K.; Thompson, N. L.; Tinnefeld, P. & Jacobson, K. (2012) Super-resolution imaging of C-type lectin and influenza hemagglutinin nanodomains on plasma membranes using blink microscopy. *Biophysical Journal* **102**, 1534–1542.

[88] Liu, P.; Wang, X.; Itano, M. S.; Neumann, A. K.; Jacobson, K. & Thompson, N. L. (2012) The formation and stability of DC-SIGN microdomains require its extracellular moiety. *Traffic* **13**, 715-726.

[89] Gringhuis, S. I.; den Dunnen, J.; Litjens, M.; van der Vlist, M. & Geijtenbeek, T. B. H. (2009) Carbohydrate-specific signaling through the DC-SIGN signalosome tailors immunity to *Mycobacterium tuberculosis*, HIV-1 and *Helicobacter pylori*. *Nature Immunology* **10**, 1081–1088.

[90] Cambi, A.; Koopman, M. & Figdor, C. G. (2005) How C-type lectins detect pathogens. *Cellular microbiology* 7, 481–488.

[91] Bogoevska, V.; Nollau, P.; Lucka, L.; Grunow, D.; Klampe, B.; Uotila, L. M.; Samsen, A.; Gahmberg, C. G. & Wagener, C. (2007) DC-SIGN binds ICAM-3 isolated from peripheral human leukocytes through Lewis X residues. *Glycobiology* **17**, 324–333.

[92] Geijtenbeek, T. B.; Krooshoop, D. J.; Bleijs, D. A.; van Vliet, S. J.; van Duijnhoven, G. C.; Grabovsky, V.; Alon, R.; Figdor, C. G. & van Kooyk, Y. (2000) DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. *Nature Immunology* **1**, 353–357.

[93] García-Vallejo, J. J.; van Liempt, E.; da Costa Martins, P.; Beckers, C.; van het Hof, B.; Gringhuis, S. I.; Zwaginga, J.-J.; van Dijk, W.; Geijtenbeek, T. B. H.; van Kooyk, Y. & van Die, I. (2008) DC-SIGN mediates adhesion and rolling of dendritic cells on primary human umbilical vein endothelial cells through LewisY antigen expressed on ICAM-2. *Molecular Immunology* **45**, 2359–2369.

[94] van Gisbergen, K. P. J. M.; Sanchez-Hernandez, M.; Geijtenbeek, T. B. H. & van Kooyk, Y. (2005) Neutrophils mediate immune modulation of dendritic cells through glycosylation-dependent interactions between MAC-1 and DC-SIGN. *Journal of Experimental Medicine* **201**, 1281–1292.

[95] van Gisbergen, K. P. J. M.; Ludwig, I. S.; Geijtenbeek, T. B. H. & van Kooyk, Y. (2005) Interactions of DC-SIGN with MAC-1 and CEACAM1 regulate contact between dendritic cells and neutrophils. *FEBS Letters* **579**, 6159–6168.

[96] van Gisbergen, K. P. J. M.; Aarnoudse, C. A.; Meijer, G. A.; Geijtenbeek, T. B. H. & van Kooyk, Y. (2005) Dendritic cells recognize tumor-specific glycosylation of carcinoembryonic antigen on colorectal cancer cells through dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin. *Cancer Research* **65**, 5935–5944.

[97] Sabatte, J.; Faigle, W.; Ceballos, A.; Morelle, W.; Rodrígues, C. R.; Lenicov, F. R.; Thépaut, M.; Fieschi, F.; Malchiodi, E.; Fernández, M.; Arenzana-Seisdedos, F.; Lortat-Jacob, H.; Michalski, J.-C.; Geffner, J. & Amigorena, S. (2011) Semen clusterin is a novel DC-SIGN ligand. *Journal of Immunology* **187**, 5299–5309.

[98] Pöhlmann, S.; Baribaud, F. & Doms, R. W. (2001) DC-SIGN and DC-SIGNR: helping hands for HIV. *Trends in Immunology* **22**, 643–646.

[99] Ludwig, I. S.; Geijtenbeek, T. B. H. & van Kooyk, Y. (2006) Two way communication between neutrophils and dendritic cells. *Current Opinion in Pharmacology* **6**, 408–413.

[100] Engering, A.; Geijtenbeek, T. B. H.; van Vliet, S. J.; Wijers, M.; van Liempt, E.; Demaurex, N.; Lanzavecchia, A.; Fransen, J.; Figdor, C. G.; Piguet, V. & van Kooyk, Y. (2002) The dendritic cell-specific adhesion receptor DC-SIGN internalizes antigen for presentation to T cells. *Journal of Immunology* **168**, 2118–2126.

[101] van Kooyk, Y. & Geijtenbeek, T. B. H. (2003) DC-SIGN: escape mechanism for pathogens. *Nature Reviews Immunology* **3**, 697–709.

[102] Cambi, A.; Beeren, I.; Joosten, B.; Fransen, J. A. & Figdor, C. G. (2009) The C-type lectin DC-SIGN internalizes soluble antigens and HIV-1 virions via a clathrin-dependent mechanism. *Journal of Immunology* **39**, 1923–1928.

[103] Tacken, P. J.; Ginter, W.; Berod, L.; Cruz, L. J.; Joosten, B.; Sparwasser, T.; Figdor, C. G. & Cambi, A. (2011) Targeting DC-SIGN via its neck region leads to prolonged antigen residence in early endosomes, delayed lysosomal degradation and cross-presentation. *Blood* **118**, 4111-4119.

[104] Engering, A.; Geijtenbeek, T. B. H. & van Kooyk, Y. (2002) Immune escape through C-type lectins on dendritic cells. *Trends in Immunology* **23**, 480–485.

[105] Valladeau, J.; Ravel, O.; Dezutter-Dambuyant, C.; Moore, K.; Kleijmeer, M.; Liu, Y.; Duvert-Frances, V.; Vincent, C.; Schmitt, D.; Davoust, J.; Caux, C.; Lebecque, S. & Saeland, S. (2000) Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. *Immunity* **12**, 71–81.

[106] Romani, N.; Clausen, B. E. & Stoitzner, P. (2010) Langerhans cells and more: langerin-expressing dendritic cell subsets in the skin. *Immunological Reviews* **234**, 120–141.

[107] Idoyaga, J.; Suda, N.; Suda, K.; Park, C. G. & Steinman, R. M. (2009) Antibody to langerin/CD207 localizes large numbers of CD8$\alpha^+$ dendritic cells to the marginal zone of mouse spleen. *PNAS* **106**, 1524–1529.

[108] Hunger, R. E.; Sieling, P. A.; Ochoa, M. T.; Sugaya, M.; Burdick, A. E.; Rea, T. H.; Brennan, P. J.; Belisle, J. T.; Blauvelt, A.; Porcelli, S. A. & Modlin, R. L. (2004) Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *Journal of Clinical Investigation* **113**, 701–708.

[109] Girolomoni, G.; Caux, C.; Lebecque, S.; Dezutter-Dambuyant, C. & Ricciardi-Castagnoli, P. (2002) Langerhans cells: still a fundamental paradigm for studying the immunobiology of dendritic cells. *Trends in Immunology* **23**, 6–8.

[110] Cunningham, A. L.; Carbone, F. & Geijtenbeek, T. B. H. (2008) Langerhans cells and viral immunity. *European Journal of Immunology* **38**, 2377–2385.

[111] Stambach, N. S. & Taylor, M. E. (2003) Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. *Glycobiology* **13**, 401–410.

[112] Thépaut, M.; Valladeau, J.; Nurisso, A.; Kahn, R.; Arnou, B.; Vivès, C.; Saeland, S.; Ebel, C.; Monnier, C.; Dezutter-Dambuyant, C.; Imberty, A. & Fieschi, F. (2009) Structural studies of langerin and Birbeck granule: a macromolecular organization model. *Biochemistry* **48**, 2684–2698.

[113] Feinberg, H.; Powlesland, A. S.; Taylor, M. E. & Weis, W. I. (2010) Trimeric structure of langerin. *Journal of Biological Chemistry* **285**, 13285–13293.

[114] Tateno, H.; Ohnishi, K.; Yabe, R.; Hayatsu, N.; Sato, T.; Takeya, M.; Narimatsu, H. & Hirabayashi, J. (2010) Dual specificity of langerin to sulfated and mannosylated glycans via a single C-type carbohydrate recognition domain. *Journal of Biological Chemistry* **285**, 6390–6400.

[115] Feinberg, H.; Taylor, M. E.; Razi, N.; Mcbride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K. & Weis, W. I. (2011) Structural basis for langerin recognition of diverse pathogen and mammalian glycans through a single binding site. *Journal of Molecular Biology* **405**, 1027– 1039.

[116] Ward, E. M.; Stambach, N. S.; Drickamer, K. & Taylor, M. E. (2006) Polymorphisms in human langerin affect stability and sugar binding activity. *Journal of Biological Chemistry* **281**, 15450–15456.

[117] Drickamer, K. (1999) C-type lectin-like domains. *Current Opinion in Structural Biology* **9**, 585–590.

[118] UNAIDS. (2011) UNAIDS world AIDS day report 2011. 1–52.

[119] Hester, E. K. (2012) HIV medications: an update and review of metabolic complications. *Nutrition in Clinical Practice* **27**, 51–64.

[120] Subbaraman, R.; Chaguturu, S. K.; Mayer, K. H.; Flanigan, T. P. & Kumarasamy, N. (2007) Adverse effects of highly active antiretroviral therapy in developing countries. *Clinical Infectious Diseases* **45**, 1093–1101.

[121] d'Arminio Monforte, A.; Lepri, A. C.; Rezza, G.; Pezzotti, P.; Antinori, A.; Phillips, A. N.; Angarano, G.; Colangeli, V.; Luca, A. D.; Ippolito, G.; Caggese, L.; Soscia, F.; Filice, G.; Gritti, F.; Narciso, P.; Tirelli, U. & Moroni, M. (2000) Insights into the reasons for discontinuation of the first highly active antiretroviral therapy (HAART) regimen in a cohort of antiretroviral naïve patients. I.CO.N.A. study group. Italian cohort of antiretroviral-naïve patients. *AIDS (London, England)* **14**, 499–507.

[122] Rerks-Ngarm, S.; Pitisuttithum, P.; Nitayaphan, S.; Kaewkungwal, J.; Chiu, J.; Paris, R.; Premsri, N.; Namwat, C.; de Souza, M.; Adams, E.; Benenson, M.; Gurunathan, S.; Tartaglia, J.; McNeil, J. G.; Francis, D. P.; Stablein, D.; Birx, D. L.; Chunsuttiwat, S.; Khamboonruang, C.; Thongcharoen, P.; Robb, M. L.; Michael, N. L.; Kunasol, P. & Kim, J. H. (2009) Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *New England Journal of Medicine* **361**, 2209–2220.

[123] Moscicki, A. B. (2008) Vaginal microbicides: where are we and where are we going? *Journal of Infection and Chemotherapy* **14**, 337–341.

[124] Stein, Z. A. (1990) HIV prevention: the need for methods women can use. *American Journal of Public Health* **80**, 460–462.

[125] Cutler, B. & Justman, J. (2008) Vaginal microbicides and the prevention of HIV transmission. *The Lancet Infectious Diseases* **8**, 685–697.

[126] Mâsse, B. R.; Boily, M.-C.; Dimitrov, D. & Desai, K. (2009) Efficacy dilution in randomized placebo-controlled vaginal microbicide trials. *Emerging Themes in Epidemiology* **6**, doi:10.1186/1742-7622-6-5.

[127] Huskens, D. & Schols, D. (2012) Algal lectins as potential HIV microbicide candidates. *Marine Drugs* **10**, 1476–1497.

[128] Reina, J. J.; Bernardi, A.; Clerici, M. & Rojo, J. (2010) HIV microbicides: state-of-the-art and new perspectives on the development of entry inhibitors. *Future Medicinal Chemistry* **2**, 1141–1159.

[129] Obiero, J.; Mwethera, P. G. & Wiysonge, C. S. (2012) Topical microbicides for prevention of sexually transmitted infections. *Cochrane Database of Systematic Reviews* **6**, doi: 10.1002/14651858.CD007961.pub2.

[130] Karim, Q. A.; Karim, S. S. A.; Frohlich, J. A.; Grobler, A. C.; Baxter, C.; Mansoor, L. E.; Kharsany, A. B. M.; Sibeko, S.; Mlisana, K. P.; Omar, Z.; Gengiah, T. N.; Maarschalk, S.; Arulappan, N.; Mlotshwa, M.; Morris, L. & Taylor, D. (2010) Effectiveness and safety of Tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science* **329**, 1168–1174.

[131] Kelly, C. G. & Shattock, R. J. (2011) Specific microbicides in the prevention of HIV infection. *Journal of Internal Medicine* **270**, 509–519.

[132] Anderluh, M.; Jug, G.; Svajger, U. & Obermajer, N. (2012) DC-SIGN antagonists, a potential new class of anti-infectives. *Current Medicinal Chemistry* **19**, 992–1007.

[133] Ernst, B. & Magnani, J. L. (2009) From carbohydrate leads to glycomimetic drugs. *Nature Reviews Drug Discovery* **8**, 661–677.

[134] Magnani, J. L. & Ernst, B. (2009) Glycomimetic drugs–a new source of therapeutic opportunities. *Discovery Medicine* **8**, 247–252.

[135] Rojo, J. & Delgado, R. (2004) Glycodendritic structures: promising new antiviral drugs. *Journal of Antimicrobial Chemotherapy* **54**, 579–581.

[136] Marradi, M.; García, I. & Penadés, S. (2011) Carbohydrate-based nanoparticles for potential applications in medicine. *Progress in Molecular Biology and Translational Science* **104**, 141–173.

[137] Nierengarten, J.-F.; Iehl, J.; Oerthel, V.; Holler, M.; Illescas, B. M.; Muñoz, A.; Martín, N.; Rojo, J.; Sánchez-Navarro, M.; Cecioni, S.; Vidal, S.; Buffet, K.; Durka, M. & Vincent, S. P. (2010) Fullerene sugar balls. *Chemical Communications* **46**, 3860–3862.

[138] Gómez-García, M.; Benito, J. M.; Butera, A. P.; Mellet, C. O.; Fernández, J. M. G. & Blanco, J. L. J. (2012) Probing carbohydrate-lectin recognition in heterogeneous environments with monodisperse cyclodextrin-based glycoclusters. *Journal of Organic Chemistry* **77**, 1273–1288.

[139] André, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H.-J. & Ungaro, R. (2008) Calix[*n*]arene-based glycoclusters: bioactivity of thiourea-linked galactose/lactose moieties as inhibitors of binding of medically relevant lectins to a glycoprotein and cell-surface glycoconjugates and selectivity among human adhesion/growth-regulatory galectins. *ChemBioChem* **9**, 1649–1661.

[140] Bossu, I.; Šulc, M.; Křenek, K.; Dufour, E.; Garcia, J.; Berthet, N.; Dumy, P.; Křen, V. & Renaudet, O. (2011) Dendri-rafts: a second generation of cyclopeptide-based glycoclusters. *Organic & Biomolecular Chemistry* **9**, 1948–1959.

[141] Blanco, J. L. J.; Mellet, C. O. & Fernández, J. M. G. (2013) Multivalency in heterogeneous glycoenvironments: hetero-glycoclusters, -glycopolymers and -glycoassemblies. *Chemical Society Reviews*, doi: 10.1039/C2CS35219B.

[142] Mitchell, D. A.; Jones, N. A.; Hunter, S. J.; Cook, J.; Jenkinson, S. F.; Wormald, M. R.; Dwek, R. A. & Fleet, G. W. J. (2007) Synthesis of 2-*C*-branched derivatives of D-mannose: 2-*C*- aminomethyl-D-mannose binds to the human C-type lectin DC-SIGN with affinity greater than an order of magnitude compared to that of D-mannose. *Tetrahedron: Asymmetry* **18**, 1502– 1510.

[143] Reina, J. J.; Sattin, S.; Invernizzi, D.; Mari, S.; Martínez-Prats, L.; Tabarani, G.; Fieschi, F.; Delgado, R.; Nieto, P. M.; Rojo, J. & Bernardi, A. (2007) *1,2*-mannobioside mimic: synthesis, DC-SIGN interaction by NMR and docking, and antiviral activity. *ChemMedChem* **2**, 1030–1036.

[144] Obermajer, N.; Sattin, S.; Colombo, C.; Bruno, M.; Svajger, U.; Anderluh, M. & Bernardi, A. (2010) Design, synthesis and activity evaluation of mannose-based DC-SIGN antagonists. *Molecular diversity* **15**, 347-360.

[145] Timpano, G.; Tabarani, G.; Anderluh, M.; Invernizzi, D.; Vasile, F.; Potenza, D.; Nieto, P. M.; Rojo, J.; Fieschi, F. & Bernardi, A. (2008) Synthesis of novel DC-SIGN ligands with an α-fucosylamide anchor. *ChemBioChem* **9**, 1921–1930.

[146] Garber, K. C. A.; Wangkanont, K.; Carlson, E. E. & Kiessling, L. L. (2010) A general glycomimetic strategy yields non-carbohydrate inhibitors of DC-SIGN. *Chemical Communications* **46**, 6747–6749.

[147] Prost, L. R.; Grim, J. C.; Tonelli, M. & Kiessling, L. L. (2012) Noncarbohydrate glycomimetics and glycoprotein surrogates as DC-SIGN antagonists and agonists. *ACS Chemical Biology* **7**, 1603–1608.

[148] Borrok, M. J. & Kiessling, L. L. (2007) Non-carbohydrate inhibitors of the lectin DC-SIGN. *Journal of the American Chemical Society* **129**, 12780–12785.

[149] Mangold, S. L.; Prost, L. R. & Kiessling, L. L. (2012) Quinoxalinone inhibitors of the lectin DC-SIGN. *Chemical Science* **3**, 772–777.

[150] Scanlan, C. N.; Offer, J.; Zitzmann, N. & Dwek, R. A. (2007) Exploiting the defensive sugars of HIV-1 for drug and vaccine design. *Nature* **446**, 1038–1045.

[151] Martínez-Avila, O.; Hijazi, K.; Marradi, M.; Clavel, C.; Campion, C.; Kelly, C. & Penadés, S. (2009) Gold manno-glyconanoparticles: multivalent systems to block HIV-1 gp120 binding to the lectin DC-SIGN. *Chemistry* **15**, 9874–9888.

[152] Gianvincenzo, P. D.; Chiodo, F.; Marradi, M. & Penadés, S. (2012) Gold manno- glyconanoparticles for intervening in HIV gp120 carbohydrate-mediated processes. *Methods in Enzymology* **509**, 21–40.

[153] Arnáiz, B.; Martínez-Ávila, O.; Falcon-Perez, J. M. & Penadés, S. (2012) Cellular uptake of gold nanoparticles bearing HIV gp120 oligomannosides. *Bioconjugate Chemistry* **23**, 814-825.

[154] Martínez-Avila, O.; Bedoya, L. M.; Marradi, M.; Clavel, C.; Alcamí, J. & Penadés, S. (2009) Multivalent manno-glyconanoparticles inhibit DC-SIGN-mediated HIV-1 *trans*-infection of human T cells. *ChemBioChem* **10**, 1806–1809.

[155] Wang, S.-K.; Liang, P.-H.; Astronomo, R. D.; Hsu, T.-L.; Hsieh, S.-L.; Burton, D. R. & Wong, C.-H. (2008) Targeting the carbohydrates on HIV-1: Interaction of oligomannose dendrons with human monoclonal antibody 2G12 and DC-SIGN. *PNAS* **105**, 3690–3695.

[156] Becer, C. R.; Gibson, M. I.; Geng, J.; Ilyas, R.; Wallis, R.; Mitchell, D. A. & Haddleton, D. M. (2010) High-affinity glycopolymer binding to human DC-SIGN and disruption of DC-SIGN interactions with HIV envelope glycoprotein. *Journal of the American Chemical Society* **132**, 15130–15132.

[157] Lasala, F.; Arce, E.; Otero, J. R.; Rojo, J. & Delgado, R. (2003) Mannosyl glycodendritic structure inhibits DC-SIGN-mediated Ebola virus infection *in cis* and *in trans*. *Antimicrobial Agents and Chemotherapy* **47**, 3970–3972.

[158] Tabarani, G.; Reina, J. J.; Ebel, C.; Vivès, C.; Lortat-Jacob, H.; Rojo, J. & Fieschi, F. (2006) Mannose hyperbranched dendritic polymers interact with clustered organization of DC-SIGN and inhibit gp120 binding. *FEBS Letters* **580**, 2402–2408.

[159] Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

[160] Powlesland, A. S.; Fisch, T.; Taylor, M. E.; Smith, D. F.; Tissot, B.; Dell, A.; Pöhlmann, S. & Drickamer, K. (2008) A novel mechanism for LSECtin binding to Ebola virus surface glycoprotein through truncated glycans. *Journal of Biological Chemistry* **283**, 593–602.

[161] Dominguez-Soto, A.; Aragoneses-Fenoll, L.; Martin-Gayo, E.; Martinez-Prats, L.; Colmenares, M.; Naranjo-Gomez, M.; Borras, F. E.; Munoz, P.; Zubiaur, M.; Toribio, M. L.; Delgado, R. & Corbi, A. L. (2007) The DC-SIGN-related lectin LSECtin mediates antigen capture and pathogen binding by human myeloid cells. *Blood* **109**, 5337–5345.

[162] Domínguez-Soto, A.; Aragoneses-Fenoll, L.; Gómez-Aguado, F.; Corcuera, M. T.; Clária, J.; García-Monzón, C.; Bustos, M. & Corbí, A. L. (2009) The pathogen receptor liver and lymph node sinusoidal endotelial cell C-type lectin is expressed in human Kupffer cells and regulated by PU.1. *Hepatology* **49**, 287–296.

[163] 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. & Pöhlmann, S. (2005) LSECtin interacts with filovirus glycoproteins and the spike protein of SARS coronavirus. *Virology* **340**, 224–236.

[164] Gramberg, T.; Soilleux, E.; Fisch, T.; Lalor, P. F.; Hofmann, H.; Wheeldon, S.; Cotterill, A.; Wegele, A.; Winkler, T.; Adams, D. H. & 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.

[165] Suzuki, N.; Yamamoto, K.; Toyoshima, S.; Osawa, T. & Irimura, T. (1996) Molecular cloning and expression of cDNA encoding human macrophage C-type lectin. Its unique carbohydrate binding specificity for Tn antigen. *Journal of Immunology* **156**, 128–135.

[166] Denda-Nagai, K.; Aida, S.; Saba, K.; Suzuki, K.; Moriyama, S.; Oo-Puthinan, S.; Tsuiji, M.; Morikawa, A.; Kumamoto, Y.; Sugiura, D.; Kudo, A.; Akimoto, Y.; Kawakami, H.; Bovin, N. V. & Irimura, T. (2010) Distribution and function of macrophage galactose-type C-type lectin 2 (MGL2/CD301b): efficient uptake and presentation of glycosylated antigens by dendritic cells. *Journal of Biological Chemistry* **285**, 19193–19204.

[167] Iida, S.; Yamamoto, K. & Irimura, T. (1999) Interaction of human macrophage C-type lectin with O-linked N-acetylgalactosamine residues on mucin glycopeptides. *Journal of Biological Chemistry* **274**, 10697–10705.

[168] van Vliet, S. J.; van Liempt, E.; Saeland, E.; Aarnoudse, C. A.; Appelmelk, B.; Irimura, T.; Geijtenbeek, T. B. H.; Blixt, O.; Alvarez, R.; van Die, I. & van Kooyk, Y. (2005) Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *International Immunology* **17**, 661–669.

[169] Takada, A.; Fujioka, K.; Tsuiji, M.; Morikawa, A.; Higashi, N.; Ebihara, H.; Kobasa, D.; Feldmann, H.; Irimura, T. & Kawaoka, Y. (2004) Human macrophage C-type lectin specific for galactose and N-acetylgalactosamine promotes filovirus entry. *Journal of Virology* **78**, 2943–2947.

[170] Upham, J. P.; Pickett, D.; Irimura, T.; Anders, E. M. & Reading, P. C. (2010) Macrophage receptors for influenza a virus: role of the macrophage galactose-type lectin and mannose receptor in viral entry. *Journal of Virology* **84**, 3730–3737.

[171] Mayer, M. & Meyer, B. (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie International Edition* **38**, 1784–1788.

[172] Angulo, J. & Nieto, P. M. (2011) STD-NMR: application to transient interactions between biomolecules-a quantitative approach. *European Biophysoical Journal* **40**, 1357–1369.

[173] Kolb, H. C.; Finn, M. G. & Sharpless, K. B. (2001) Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie International Edition* **40**, 2004–2021.

[174] Rostovtsev, V. V.; Green, L. G.; Fokin, V. V. & Sharpless, K. B. (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie International Edition* **41**, 2596–2599.

[175] Tornøe, C. W.; Christensen, C. & Meldal, M. (2002) Peptidotriazoles on solid phase: [1,2,3]- triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *Journal of Organic Chemistry* **67**, 3057–3064.

[176] Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B. & Fokin, V. V. (2005) Copper(I)-catalyzed synthesis of azoles. DFT study predicts unprecedented reactivity and intermediates. . *Journal of the American Chemical Society* **127**, 210–216.

[177] Kortt, A. A.; Gruen, L. C. & Oddie, G. W. (1997) Influence of mass transfer and surface ligand heterogeneity on quantitative biacore binding data. Analysis of the interaction of NC10 Fab with an anti-idiotype Fab'. *Journal of Molecular Recognition* **10**, 148–158.

# Part VI.

# Appendices

# 13. Supplementary information to paper n°1

Supplementary Information

# Second generation of Fucose-based DC-SIGN ligands : affinity improvement and specificity *versus* Langerin

Manuel Andreini[a] §, Daniela Doknic[a] §, Ieva Sutkeviciute[b], [e] §, José J. Reina[a], Janxin Duan[c], Eric Chabrol[b], [e], Michel Thepaut[b], [e], [f], Elisabetta Moroni[a], Fabio Doro[a], Laura Belvisi[a], Joerg Weiser[c], Javier Rojo[d], Franck Fieschi*[b], [e], [g], Anna Bernardi*[a]

§ these authors, listed in alphabetical order, contributed equally to this work

*a) Universita' degli Studi di Milano, Dipartimento di Chimica Organica e Industriale and CISI, via Venezian 21, 20133 Milano, Italy, anna.bernardi@unimi.it*
*b) Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, 38027 Grenoble, France, fieschi@ibs.fr*
*c) Anterio consult&research, Augustaanlage 23, D-68165 Mannheim, Germany*
*d) Glycosystems Laboratory, Instituto de Investigaciones Químicas, CSIC – Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla (Spain)*
*e) CNRS, UMR 5075, Grenoble, France*
*f) CEA, Grenoble, France*
*g) Institut Universitaire de France, 103 boulevard Saint-Michel 75005 Paris, France*

**1. Supplementary Information – Synthetic procedures** p 1 - 30

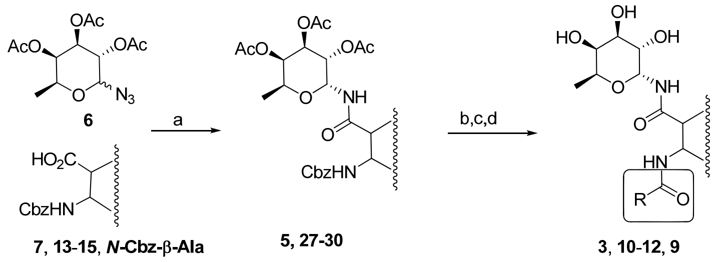
**2. Supplementary Information – Surface Plasmon Resonance Experiments** p 31 - 35

---

Supplementary Information

## 1. Supplementary Information – Synthetic procedures

The libraries were synthesized through the common approach shown in Scheme 1.[1]



a) De Shong reaction (ref 2); b) $H_2$/Pd; c) $RCO_2H$ coupling; d) MeONa

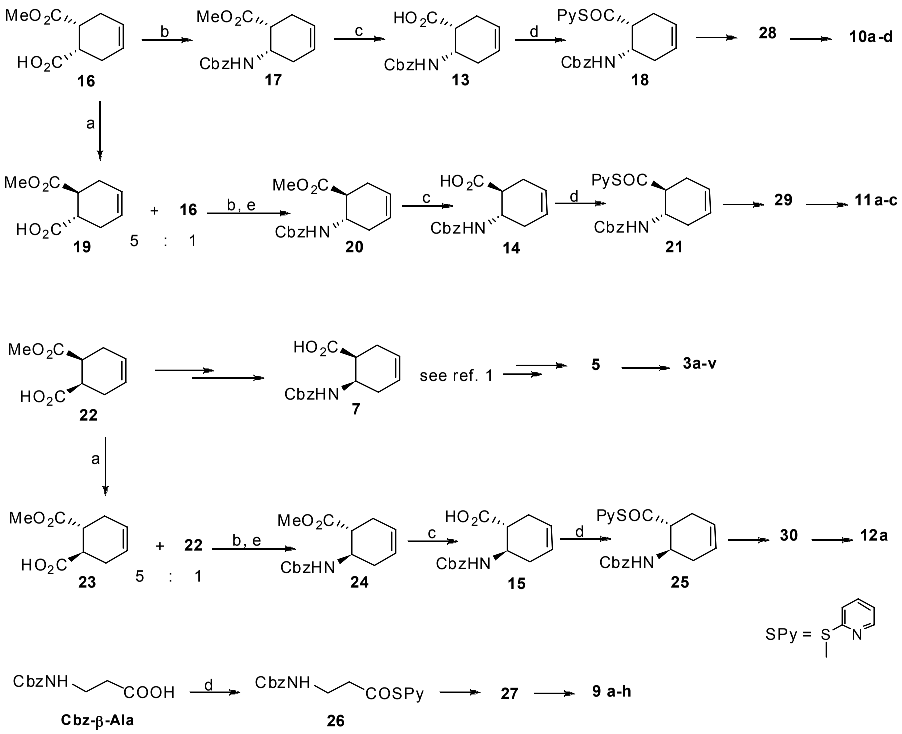
**SI-Scheme 1.** General approach for the synthesis of the ligands

DeShong reaction[2] of tri-*O*-acetyl-fucosylazide **6** with the appropriate *N*-Cbz-protected β-aminoacid (**7**, **13-15**, and *N*-Cbz-Ala) yielded α-*N*-fucosylamides. After Cbz removal, the free amines were coupled with the acid partners ($RCO_2H$) without previous purification. The synthesis of **3a**,[1] (through intermediates **7** and **5** ) and **9a**[3] (through **27**) have been reported elsewhere**.**

### Synthesis and activation of the β-aminoacids.

Carbobenzyloxy-β-alanine is commercially available. The synthesis of (1*S*,2*R*)- *N*-Carbobenzyloxy-2-amino-cyclohexanecarboxylic acid **7** has been described.[1] The synthesis of the isomeric *N*-carbobenzyloxy-2-amino-cyclohexanecarboxylic acids **13-15** was performed as shown in Scheme 2, starting from the *cis* monoacids **16** and **22.**[4] They were either transformed directly into the *cis* *N*-Cbz-protected β-aminoacids **13** and **7**, respectively, using a Curtius rearrangement followed by LiOH hydrolysis, or were transformed in the *trans* isomers **19** and **23** by base-promoted equilibration[4b] (Scheme

2). Curtius rearrangement of the *trans* monoacids **19** and **23** afforded the corresponding *trans* *N*-Cbz-protected β-aminoacids **14** and **15** after LiOH hydrolysis of esters **20** and **24.**

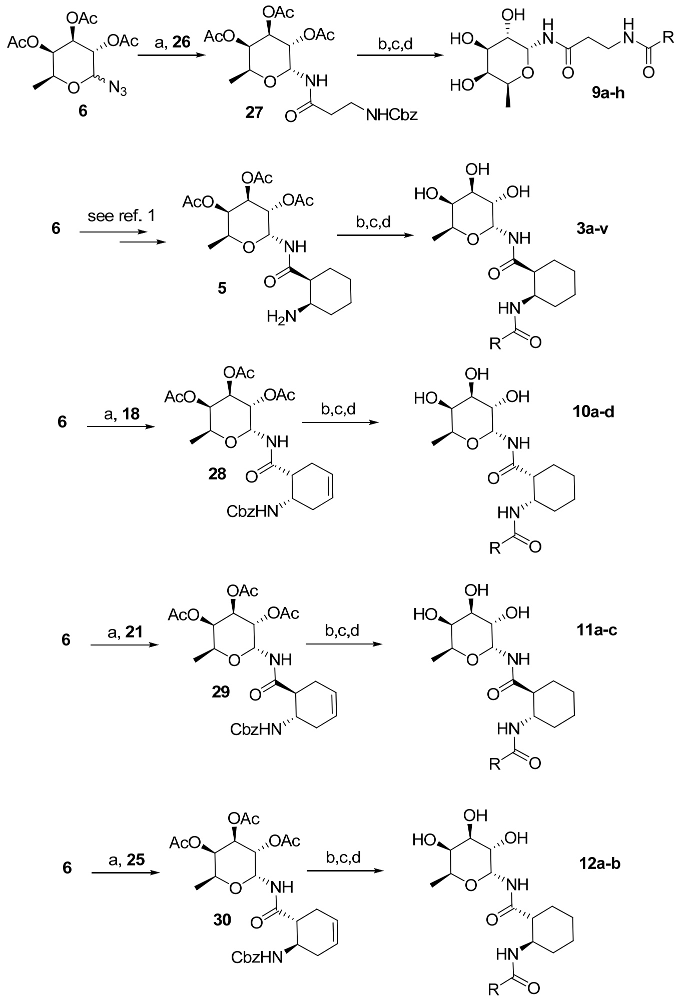

MeO2C
HO2C
16
b
MeO2C
CbzHN
17
c
HO2C
CbzHN
13
d
PySOC
CbzHN
18
28
10a-d
a
MeO2C
HO2C
19
5 : 1
+ 16
b, e
MeO2C
CbzHN
20
c
HO2C
CbzHN
14
d
PySOC
CbzHN
21
29
11a-c
MeO2C
HO2C
22
HO2C
CbzHN
7
see ref. 1
5
3a-v
a
MeO2C
HO2C
23
5 : 1
+ 22
b, e
MeO2C
CbzHN
24
c
HO2C
CbzHN
15
d
PySOC
CbzHN
25
30
12a
SPy =
S
N
CbzNH
COOH
Cbz-β-Ala
d
CbzNH
COSPy
26
27
9 a-h

a) tAmOK, THF; b) Curtius rearrangement; c) LiOH; d) Ph$_3$P, PySSPy; e) flash chromatography

**SI-Scheme 2.** Synthesis of the β-aminoacids **7, 13-15** and activation as pyridyl thiolesters

All Cbz-protected aminoacids, including commercially available Cbz-β-Ala, were activated as pyridyl thiolesters (**18, 21, 25** and **26**, Scheme 2 and 3) and used in the DeShong reactions with **6** (Scheme 3). The pyridyl thiolesters of Cbz-β-Ala **26** is a known compound.[5]

---


AcO
OAc
OAc
N3
6
a, 26
27
NH
NHCbz
b,c,d
HO
OH
HO
9a-h
R
6
see ref. 1
5
H2N
b,c,d
3a-v
6
a, 18
28
CbzHN
10a-d
6
a, 21
29
11a-c
6
a, 25
30
12a-b
HN
R
O

a) Ph$_3$P, then pyridyl thiolester; b) H$_2$/Pd; c) RCO$_2$H, coupling; d) MeONa, MeOH

**SI-Scheme 3**. DeShong reactions and synthesis of the libraries

This step produced the fucosylamides **5**,[1] **27**,[3] and **28-30**, which were transformed in the final ligands after Cbz removal (H$_2$/Pd) and coupling with the acid partner.

## Experimental

Solvents were dried by standard procedures: dichloromethane, methanol, *N*,*N*-diisopropylethylamine and triethylamine were dried over calcium hydride, chloroform and pyridine were dried over activated molecular sieves. Reactions requiring anhydrous conditions were performed under nitrogen. $^1$H, $^{13}$C and $^{31}$P-NMR spectra were recorded at 400 MHz on a Bruker AVANCE-400 instrument. Chemical shifts (*δ)* for $^1$H and $^{13}$C spectra are expressed in ppm relative to internal $Me_4Si$ as standard. Signals were abbreviated as s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet. Mass spectra were obtained with a Bruker ion-trap Esquire 3000 apparatus (ESI ionization) or an Autospec Fission Instrument (FAB ionization). Thin layer chromatography (TLC) was carried out with pre-coated Merck $F_{254}$ silica gel plates. Flash chromatography (FC) was carried out with Macherey-Nagel silica gel 60 (230-400 mesh). The *cis* monoacids **16** and **22** are known and were synthesised as described in the literature.[4] The synthesis of **3a** from **22** through **5** and **7** has already been reported.[1] The pyridyl thiolesters of Cbz-β-Ala **26**[5] and the corresponding fucosylamide **27**,[3] are known compounds.

## Synthesis of the *trans* monoacids 19 and 23

29 mL of Potassium *tert*-Amylate (1.7M in Toluene) were diluted in 40 mL of toluene under $N_2$. The solution was cooled to –15 °C. Then a solution of the *cis* monoacid **16** (or **22**) (3.5 g, 19 mmol) in toluene (20 mL) was added dropwise. The reaction mixture was stirred for 4 hours at - 15 °C. The reaction was quenched with 60 mL HCl (1M), diluted with ethyl acetate and washed with water. The organic phase was dried over $Na_2SO_4$ and the solvent was evaporated to afford a mixture of product **19** and **16** (or **23** and **22**) as a colourless oil.

Yield (**16**/**19** mixture): 3.3 g (94 %) in a 1:5 ratio.

Yield (**22**/**23** mixture): 2.9 g (84 %) in a 1:4 ratio.

The mixtures **16**/**19** (or **23**/**22**) were used in the general procedure for the *Curtius* rearrangement and the isomers were separated at that level.

## General procedure for the *Curtius* rearrangement – Synthesis of 17, 20 and 24

DPPA (1.0 mol equivalent) and $Et_3N$ (1 mol equivalent) were added to a 0.3 M solution of the acid in toluene. After the addition of benzyl alcohol (1.8 mol equivalent) the reaction mixture was refluxed for 5 h, then it was diluted with ethyl acetate and extracted with HCl (5 %), saturated $NaHCO_3$ solution and brine. The solvent was evaporated and the *N*-Cbz-protected β-aminoacids were purified by flash-chromatography or automated chromatography on silica gel to give a white solid.

**17**. Obtained from Curtius rearrangement of **16**[4] and purified by flash chromatography (toluene/ethyl acetate 15:1), $R_f = 0.31$; Yield (**17**): 4.7 g (75 %)

**$^1$H-NMR (400 MHz, $CDCl_3$):** 2.14 – 2.22 (m, 1H, H3), 2.28 –2.41 (m, 2H, H3 and H6), 2.46 – 2.54 (m, 1H, H6), 2.79 – 2.83 (m, 1H, H1), 3.67 (s, 3H, COO*Me*), 4.21 – 4.27 (m, 1H, H2), 5.05 – 5.08 (m, 2H, *CH*$_2$Ph), 5.38 (brd, 1H, NH), 5.56-5.67 (m, 2H, H4 and H5), 7.28 – 7.36 (m, 5H, *Ar*). **$^{13}$C-NMR (100 MHz, $CDCl_3$):** 25.7 (C6), 30.8 (C3), 42.3 (C1), 47.0 (C2), 52.1 (COO*Me*), 66.9 (*CH*$_2$Ph), 124.9 (C4 or C5), 125.2 (CH, C4 or C5), 128.3 (CH, Ar), 128.7 (CH, Ar), 136.7 (Cquart., Ar). ESI-MS *m/z* = 312 [(M + Na)$^+$, 100 %].

**20.** Obtained from Curtius rearrangement of a 5:1 **19:16** mixture and separated from the *cis* isomer by automated chromatography (toluene/ethyl acetate with a gradient from 0 % to 10 % ethyl acetate), $R_f =$ 0.31 (toluene/EtOAc 19:1). Yield (**20**): 1.7 g (38 %)

**$^1$H-NMR (400 MHz, $CDCl_3$)**: 1.92 –2.00 (m, 1H, H3), 2.24 – 2.31 (m, 1H, H6), 2.45 – 2.52 (m, 2H, H3 and H6), 2.67 – 2.73 (m, 1H, H1), 3.63 (s, 3H, COO*Me*), 4.04 – 4.12 (m, 1H, H2), 4.86 (brs, 1H, NH), 5.08 (s, 2H, *CH*$_2$Ph), 5.56 – 5.59 (m, 1H, H4), 5.62 – 5.66 (m, 1H, H5), 7.27 – 7.35 (m, 5H, *Ar*). **$^{13}$C-NMR (100 MHz, $CDCl_3$)**: 26.8 (C6), 31.3 (C3), 44.6 (C1), 48.1 (C2), 52.1 (COO*Me*), 66.9 (*CH*$_2$Ph), 124.3 (C4), 125.2 (C5), 128.3 (Ar), 128.7 (Ar), 128.7 (Ar), 136.7 (Cquart., Ar), 155.7 (NHCO), 174.1 (COOMe). ESI-MS *m/z* = 312 [(M + Na)$^+$, 100 %]

**24.** Obtained from Curtius rearrangement of a 5:1 **23:22** mixture and separated from the *cis* isomer by automated chromatography (toluene/ethyl acetate with a gradient from 0 % to 10 % ethyl acetate), $R_f = 0.31$ (toluene/EtOAc 19:1); Yield (**24**): 1.1 g (32 %).

**$^1$H-NMR (400 MHz, CDCl$_3$)**: 1.92 – 2.00 (m, 1H, H3), 2.24 – 2.32 (m, 1H, H6), 2.46 – 2.51 (m, 2H, H3 and H6), 2.67 – 2.73 (m, 1H, H1), 3.62 (s, 3H, COO*Me*), 4.06 – 4.13 (m, 1H, H2), 4.84 (brs, 1H, NH), 5.08 (s, 2H, *CH$_2$*Ph), 5.56 – 5.59 (m, 1H, H4), 5.63 – 5.66 (m, 1H, H5), 7.26 – 7.36 (m, 5H, Ar). **$^{13}$C-NMR (100 MHz, CDCl$_3$)**: 26.8 (C6), 31.2 (C3), 44.6 (C1), 48.1 (C2), 52.2 (COO*Me*), 66.9 (*CH$_2$*Ph), 124.3 (C4), 125.2 (C5), 128.4 (Ar), 128.7 (Ar), 128.8 (Ar), 136.6 (Cquart., Ar), 174.1 (COOMe). ESI-MS *m/z* = 312 [(M + Na)$^+$, 100 %]

**Ester hydrolysis – Synthesis of *N*-Cbz-β-aminoacids 13 - 15**

A 0.17 M solution of β-amino ester **17** (or **20** or **24**) in MeOH/H$_2$O (4:1) was prepared. LiOH*H$_2$O (2.4 mol equivalents) was added to the solution at 0 °C and the mixture was stirred for 12 h at room temperature. After the reaction was completed, ca. 2/3 of the solvent was evaporated and, if necessary, the pH of the mixture was adjusted to pH 9 with NaHCO$_3$. To remove benzyl alcohol that was occasionally still present after the previous reaction step, the mixture was extracted with diethyl ether. Then the inorganic phase was acidified to pH 1 with HCl (6M) and extracted with ethyl acetate. The organic phases were combined and the solvent was evaporated to afford the acid (**13** or **14** or **15**) as a white solid.

**13** Yield (**13**): 4.29 g (93 %)

**$^1$H-NMR (400 MHz, CDCl$_3$)**: 2.16 – 2.23 (m, 1H, H3), 2.34 – 2.39 (m, 2H, H3 and H6), 2.48 – 2.55 (m, 1H, H6), 2.85 – 2.90 (m, 1H, H1), 4.17 – 4.28 (m, 1H, H2), 5.03 – 5.11 (m, 2H, *CH$_2$*Ph), 5.46 (brd, 1H, NH), 5.59 - 5.67 (m, 2H, H4 and H5), 7.27 – 7.35 (m, 5H, Ar). **$^{13}$C-NMR (100 MHz, CDCl$_3$)**: 26.1 (C6), 30.6 (C3), 42.1 (C1), 46.9 (C2), 67.1 (*CH$_2$*Ph), 125.0 (C4 or C5), 125.1 (C4 or C5), 128.4 (CH, Ar), 128.8 (CH, Ar), 178.6 (COOH). ESI-MS *m/z* = 298 [(M + Na)$^+$, 59 %].

**14** Yield (**14**): 0.64 g (83 %)

**$^1$H-NMR (400 MHz, CDCl$_3$)**: 1.96 - 2.05 (m, 1H, H3), 2.31 – 2.40 (m, 1H, H6), 2.46 – 2.56 (m, 2H, H3 and H6), 2.76 – 2.81 (m, 1H, H1), 4.10 – 4.15 (m, 1H, H2), 4.97 (brd, 1H, NH), 5.04 – 5.14 (m, 2H, *CH$_2$*Ph), 5.59 – 5.69 (m, 2H, H4 and H5), 7.29 – 7.38 (m, 5H, Ar). **$^{13}$C-NMR (100 MHz, CDCl$_3$)**: 26.4 (C6), 30.9 (C3), 43.9 (C1), 47.7 (C2), 67.0 (*CH$_2$*Ph), 124.4 (C4 or C5), 125.1 (C4 or *C5*), 128.38 (Ar), 128.75 (Ar), 136.5 (Cquart., Ar), 163.5 (NHCO), 178.5 (COOH). ESI-MS *m/z* = 298 [(M + Na)$^+$, 11 %]

**15** Yield (**15**): 643 mg (43%)

**$^1$H-NMR (400 MHz, CDCl$_3$)**: 1.93 – 2.03 (m, 1H, H3), 2.28 – 2.41 (m, 1H, H6), 2.42 – 2.57 (m, 2H, H3 and H6), 2.73 – 2.81 (m, 1H, H1), 4.06 – 4.16 (m, 1H, H2), 4.92 (brd, 1H, NH), 5.03 – 5.13 (m, 2H, *CH$_2$*Ph), 5.55 – 5.60 (m, 1H, H4), 5.62 – 5.67 (m, 1H, H5), 7.28 – 7.38 (m, 5H, Ar). **$^{13}$C-NMR (100 MHz, CDCl$_3$)**: 26.3 (C6), 30.7 (C3), 43.7 (C1), 47.5 (C2), 66.9 (*CH$_2$*Ph), 124.2 (C4 or C5), 124.9 (C4 or *C5*), 128.2 (Ar), 128.56 (Ar), 128.6 (Ar). ESI-MS *m/z* = 298 [(M + Na)$^+$, 11 %]

**Synthesis of pyridyl thiolesters 18, 21, 25**

PPh$_3$ (1.3 mol equivalent) and dipyridyl disulfide (1.3 mol equivalent) were added to a 0.1 M solution of **13** (or **14** or **15**) in CH$_3$CN. The reaction mixture was refluxed for 2 h, then the solvent was evaporated and the product (**18** or **21** or **25**) was isolate by flash chromatography on a short path of SiO$_2$ (*n*-hexane/ethyl acetate 6:4 or 7:3).

**18** Yield (**18**): 830 mg (79 %), R$_f$ = 0.25 (n-hexane/ethyl acetate 7:3)

**$^1$H-NMR (400 MHz, CDCl$_3$)**: 2.21 – 2.38 (m, 1H, H3), 2.38 – 2.48 (m, 2H, H3 and H6), 2.56 – 2.65 (m, 1H, H6), 3.17 – 3.21 (m, 1H, H1), 4.33 – 4.39 (m, 1H, H2), 5.07 (s, 2H, *CH$_2$*Ph), 5.26 (brd, 1H, NH), 5.61-5.72 (m, 2H, H4 and H5), 7.25 – 7.34 (m, 6H, Ar and Pyr), 7.55 – 7.58 (m, 1H, Pyr), 7.67 – 7.71 (m, 1H, Pyr), 8.58 – 8.61 (m, 1H, Pyr)

**21** Yield (**21**): 528 mg (79 %), R$_f$ = 0.25 (n-hexane/ethyl acetate 7:3)

**$^1$H-NMR (100 MHz, CDCl$_3$)**: 2.04 – 2.13 (m, 1H, H6), 2.38 – 2.54 (m, 3H, H3 and H6), 3.06 – 3.14 (m, 1H, H1), 4.12 – 4.19 (m, 1H, H2), 4.99 (brd, 1H, NH), 5.09 (s, 2H, *CH$_2$*Ph), 5.58 - 5.69 (m, 2H, H4 and H5), 7.26 – 7.35 (m, 6H, Ar and pyr ), 7.51 - .53 (m, 1H, pyr), 7.66 – 7.70 (m, 1H, pyr), 8.57 – 8.60 (m, 1H, pyr).

**25** Yield (**25**): 334 mg (83 %), $R_f$ = 0.33 (n-hexane/ethyl acetate 6:4)

**$^1$H-NMR (400 MHz, $CDCl_3$):** 2.03 – 2.14 (m, 1H, H6), 2.39 – 2.54 (m, 3H, H3 and H6), 3.07 – 3.14 (m, 1H, H1), 4.12 – 4.19 (m, 1H, H2), 4.96 – 5.03 (brd, 1H, NH), 5.08 (s, 2H, *CH$_2$*Ph), 5.58 – 5.69 (m, 2H, H4 and H5), 7.24 – 7.34 (m, 6H, Ar and pyr ), 7.51 – 7.53 (m, 1H, pyr), 7.66 – 7.71 (m, 1H, pyr), 8.58 – 8.61 (m, 1H, pyr).

**DeShong reaction. Synthesis of the α-fucosylamides 28-30**

To a 0.07 M solution of fucosyl azide **6** in $EtNO_2$ grounded molecular sieves and a 0.08 M solution of $PPh_3$ (1.1 mol equivalent) in $EtNO_2$ were added. The reaction mixture was refluxed for 14 h. After cooling to room temperature, a 0.5 M solution of pyridyl thiolester **18** (or **21** or **25**) (1.3 mol equivalent) in $EtNO_2$ and $CuCl_2$*$H_2O$(1.3 mol equivalent) were added to the reaction mixture. The mixture was stirred at 40 °C for 24 h monitoring by TLC. After completion, molecular sieves and catalyst were filtered on a celite pad washing with ethyl acetate, the organic phase was extracted with $NH_3/NH_4Cl$ (1:1) and washed with water. The organic phase was dried over $Na_2SO_4$ and the solvent was evaporated. The obtained crude was purified by flash-chromatography on $SiO_2$ (*n*-hexane/ethyl acetate 6:4).

**28** Yield (**28**): 320 mg (31 %), $R_f$ = 0.22 (*n*-hexane/ethyl acetate 6:4)

**$^1$H-NMR (400 MHz, $CDCl_3$)**: 1.15 (d, $J_{5\text{-}6}$ = 6.4 Hz, $H_{F6}$), 1.94 (s, 3H, Fuc-Ac), 1.96 (s, 3H, Fuc-Ac), 2.05 – 2.14 (m, 1H, $H_{Cy3ax}$), 2.15 (s, 3H, Fuc-*Ac*), 2.19 – 2.25 (m, 1H, $H_{Cy6ax}$), 2.43 – 2.51 (m, 1H, $H_{Cy3eq}$), 2.54 – 2.63 (m, 1H, $H_{Cy6eq}$), 2.70 – 2.75 (m, 1H, $H_{Cy1}$), 4.02 – 4.08 (m, 1H, $H_{F5}$), 4.32 – 4.38 (m, 1H, $H_{Cy2}$), 5.06 (d, *J* = 12.2 Hz, 1H, *CH$_2$*Ph), 5.16 (d, *J* = 12.2 Hz, 1H, *CH$_2$*Ph), 5.29 - 5.36 (m, 4H, $H_{F4}$, $H_{F3}$, Cy-*NH*, $H_{F2}$), 5.62 – 5.80 (m, 2H, $H_{Cy4}$ and $H_{Cy5}$), 5.94 (dd, *J* = 4.3 Hz, *J* = 8.1 Hz, 1H, $H_{F1}$), 7.28 – 7.36 (m, 5H, Ar), 7.55 (brs, 1H, Fuc-*NH*). **$^{13}$C-NMR (100 MHz, $CDCl_3$)**: 16.5 ($C_{F6}$), 20.8 (Fuc-Ac-Me), 20.9 (Fuc-Ac-Me), 25.6 ($C_{Cy6}$), 32.4 ($C_{Cy3}$), 44.1 ($C_{Cy1}$), 46.2 ($C_{Cy2}$), 66.3 ($H_{F5}$), 66.4 ($C_{F2}$), 67.6 ($CH_2$, *CH$_2$*Ph), 68.3 ($C_{F3}$), 70.9 ($C_{F4}$), 74.6 ($C_{F1}$), 124.4 ($C_{Cy4}$ or $C_{Cy5}$), 126.4 ($C_{Cy4}$ or $C_{Cy5}$), 128.4 (Ar), 128.5 (Ar), 128.8 (Ar), 136.3 (Cquart., Ar), 157.4 (Cquart.), 169.7 (Cquart.), 170.2 (Cquart), 170.9 (Cquart.), 173.4 (Cquart.). ESI-MS m/z = 569 [(M + Na)$^+$, 100 %]. $[\alpha]_D$ = -17.8 (c = 1.50, EtOH)

**29** Yield (**29**): 210 mg (46 %), $R_f$ = 0.15 (*n*-hexane/ethyl acetate 6:4)

**$^1$H-NMR (400 MHz, $CDCl_3$)**: 1.02 (d, 3H, $J_{5\text{-}6}$ = 6.3 Hz, $H_{F6}$), 1.98 (s, 6H, Fuc-Ac), 2.12 (s, 3H, Fuc-Ac), 2.14 – 2.20 (m, 1H, $H_{Cy3ax}$), 2.21 – 2.45 (m, 1H, $H_{Cy6ax}$, $H_{Cy3eq}$, $H_{Cy6eq}$), 2.83 – 2.89 (m, 1H, $H_{Cy1}$), 3.82 – 3.89 (m, 1H, $H_{Cy2}$), 3.89 – 3.98 (m, 1H, $H_{F5}$), 5.00 – 5.09 (m, 2H, *CH$_2$*Ph), 5.19 – 5.22 (m, 1H, $H_{F4}$), 5.27 – 5.35 (m, 1H, Cy-*NH*, $H_{F3}$, $H_{F2}$), 5.55 – 5.60 (m, 1H, $H_{Cy4}$), 5.63 – 5.68 (m, 1H, $H_{Cy5}$), 5.91 (dd, *J* = 4.6 Hz, *J* = 8.1 Hz 1H, $H_{F1}$) 7.26 – 7.34 (m, 5H, Ar), 7.42 (brs, 1H, Fuc-*NH*). **$^{13}$C-NMR (100 MHz, $CDCl_3$)**: 16.3 ($C_{F6}$), 20.8 (Fuc-Ac-*Me*), 20.9 (Fuc-Ac-*Me*), 21.0 (Fuc-Ac-*Me*), 28.7 ($C_{Cy6}$), 31.3 ($C_{Cy3}$), 46.3 ($C_{Cy1}$), 49.1 ($C_{Cy2}$), 65.8 ($C_{F5}$), 66.3 ($C_{F2}$), 67.2 (*CH$_2$*Ph), 68.4 ($C_{F3}$), 70.9 ($C_{F4}$), 74.6 ($C_{F1}$), 124.6 ( $C_{Cy4}$), 125.6 ($C_{Cy5}$), 128.2 (Ar), 128.5 (Ar), 128.8 (Ar), 136.3 (Cquart, Ar ), 156.6 (Cquart.), 169.7 (Cquart.), 170.6 (Cquart.), 171.0 (Cquart.), 174,7 (Cquart.) ESI-MS m/z = 569 [(M + Na)$^+$, 100 %]. $[\alpha]_D$ = -93.7 (c = 0.50, MeOH)

**30** Yield (**30**): 200 mg (52%), $R_f$ = 0.15 (*n*-hexane/ethyl acetate 6:4)

**$^1$H-NMR (400 MHz, $CDCl_3$):** δ (ppm) = 1.04 (d, 3H, $J_{5-6}$ = 6.4 Hz, $H_{F6}$), 1.96 (s, 3H, OAc), 1.98 (s, 3H, OAc), 2.05 – 2.12 (m, 1H, $H_{C3ax}$), 2.15 (s, 3H, OAc), 2.24 – 2.48 (m, 3H, $H_{C3eq}$ and $H_{C6}$), 2.74 – 2.87 (m, 1H, $H_{Cy1}$), 3.88 – 3.99 (m, 2H, $H_{Cy2}$ and $H_{F5}$), 5.03 – 5.12 (m, 3H, *CH$_2$*Ph and Cyc-*NH*), 5.23 – 5.25 (m, 1H, $H_{F2}$), 5.32 – 5.38 (m, 2H, $H_{F3}$ and $H_{F4}$), 5.57 –5.62 (m, 1H, $H_{Cy4}$ or $H_{Cy5}$), 5.64 – 5-70 (m, 1H, $H_{Cy4}$ or $H_{Cy5}$), 5.85 – 5.91 (m, 1H, $H_{F1}$), 7.27 – 7.36 (m, 6H, Ar and Fuc-NH). **$^{13}$C-NMR (100 MHz, $CDCl_3$):** δ (ppm) = 16.3 ($C_{F6}$), 20.8 (Fuc-OAc-Me), 20.9 (Fuc-OAc-Me), 20.9 (Fuc-OAc-Me), 29.5 ($C_{Cy6}$), 32.0 ($C_{Cy3}$), 47.4 ($C_{Cy1}$), 48.7 ($C_{Cy2}$), 66.02 ($C_{F5}$), 66.4 ($C_{F3}$ or $C_{F4}$), 67.4 (*CH$_2$*-Ph), 68.3 ($C_{F3}$ or $C_{F4}$), 70.9 ($C_{F2}$), 74.8 ($C_{F1}$), 124.4 ($C_{Cy4}$ or $C_{Cy5}$), 125.8 ($C_{Cy4}$ or $C_{Cy5}$), 128.3 (Ar), 128.6 (Ar), 128.9 (Ar), 136.1 (Cquart., Ar), 156.8 (Cquart.), 169.7 (Cquart.), 170.4 (Cquart.), 170.9 (Cquart.), 174.4 (Cquart.). ESI-MS *m/z* = 569 [(M + Na)$^+$, 100 %].

**Synthesis of the ligands 3, 9-12**

The fucosylamides isolated after the DeShong reaction were treated with $H_2$ and Pd/C in MeOH (0.05 M) for 2h at room temperature. The crude was filtered through celite and used for the coupling reactions, which were run using the general procedures for coupling reported below for amine **5.** The final amides were deacetylated using the general procedure for Zemplen's deprotection described below for compound **8**.

***General procedure of acetylation with $Ac_2O$***

To a solution of **5** (0.072 mmol, 1 equiv) in $CH_2Cl_2$ (2 mL), pyridine (0.01 mL, 1.3 equiv) and Ac2O (0.01 mL, 1.3 equiv) were added and the solution was stirred overnight. The mixture was diluted with $CH_2Cl_2$, extracted with water, diluted HCl, water. The organic phase was dried over $Na_2SO_4$ and evaporated.

***General procedure of coupling using HBTU***

To a solution of **5** (0.02 mmol) in 0.2 mL of $CH_2Cl_2$, $Et_3N$ (0.072 mmol, 3 eq.) and carboxylic partner (0.036 mmol, 1.5 eq.) in 0.2 mL of $CH_2Cl_2$ were added. Subsequently, HBTU (0.036 mmol, 1.5 eq.) was added and reaction mixture stirred at room temperature. After 18 h, 10 mL of $CH_2Cl_2$ were added to the reaction mixture and organic phase was washed with 0.5 M NaOH (10 mL), 1M $KHSO_4$ (10 mL), water (10 mL) and brine (10 mL). The organic phase was dried over $Na_2SO_4$, filtered and the solvent evaporated under vacuum. The residue was purified via Biotage using the appropriate solvent gradient.

***General procedure of coupling using acid chloride***

The carboxylic partner (1.2 mmol, 1.2 eq.) was refluxed in 1 mL of toluene in presence of oxalyl chloride (3 mmol, 3.3 eq.) during 3 h. The solution was then evaporated and added to a solution of **5** (1 mmol) and $Et_3N$ (2 mmol, 2 eq.) in 1 mL of THF. The solution was stirred at room temperature for 18 h. After completion, the solvent was evaporated. The crude was taken up with $CH_2Cl_2$, was washed with a saturated solution of $NaHCO_3$, 10% citric acid, and an aqueous saturated NaCl solution. Separations have been done using isolute column separative phase. The solvent was evaporated under vacuum and the residue was purified via Biotage using the appropriate solvent gradient.

---

***General procedure of coupling using EDC/HOBt***

To the mixture of the acidic partner (0.07 mmol, 1.1 eq.) in 1.5 mL of $CH_2Cl_2$ were added HOBt (0.10 mmol, 1.4 eq.) and compound **5** (0.07 mmol, 1 eq.). The mixture was stirred at 0°C for 5 min then a solution of EDC.HCl (0.09 mmol, 1.3 eq.) and triethylamine (0.09 mmol, 1.2 eq.) in 0.5 mL of $CH_2Cl_2$ was added. The resulting mixture was allowed to warm at r.t and stirred at r.t overnight. The reaction was then washed with $NaHCO_3$, 10% citric acid, and an aqueous saturated NaCl solution. Separations were performed using isolute column separative phase. The solvent was evaporated under vacuum and the residue was purified via Biotage using the appropriate solvent gradient.

***General procedure of Zemplen's deprotection***

To a solution of protected amide **8** (0.07 mmol) in 1.5 mL of dry MeOH, NaOMe (0.01 mmol) was added and the solution was stirred at room temperature. After completion, amberlite IRA $120^+$ was added until pH 7 and the beads were filtered off. The solvent was evaporated under vacuum and the residue was purified by reverse phase automated chromatography using the appropriate solvent gradient.

**(1*S*,2*R*)-2-Aminocyclohexanecarboxylic acid series: Ligands 3**

***N*-[(1*S*,2*R*)-2-(3-Hydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3b)**

Compound **5**[13] was coupled with 3-acetoxybenzoic acid using the general HBTU procedure. The crude was purified by flash chromatography (1:4 hexane:AcOEt , $R_f = 0.63$, 72% yield). Zemplen deprotection and flash chromatography (85:15 $CHCl_3$:MeOH, $R_f = 0.21$) afforded **3b** (quant). **$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 0.88 (d, 3H, $J_{5\text{-}6} = 6.4$ Hz, $H_{F6}$), 1.31-2.17 (m, 8H, $CH_{2Cy}$), 2.94 (m, 1H, $H_{Cy1}$), 3.54 (m, 1H, $H_{F4}$), 3.59 (m, 1H, $H_{F5}$), 3.75 (dd, 1H, $J_{2\text{-}3} = 10.3$ Hz, $J_{3\text{-}4} = 3.3$ Hz, $H_{F3}$), 3.93 (dd, 1H, $J_{1\text{-}2} = 5.6$ Hz, $H_{2F}$), 4.37 (m, 1H, $H_{Cy2}$), 5.50 (d, 1H, $H_{F1}$), 6.94 (m, 1H, $H_{Ar}$), 7.19-7.27 (m, 3H, $H_{Ar}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.8 ($C_{F6}$), 23.3, 23.5, 26.2, 30.7 (4×$CH_{2Cy}$), 46.0 ($C_{Cy1}$), 50.3 ($C_{Cy2}$), 68.0 ($C_{F2}$), 68.4 ($C_{F5}$), 71.5 ($C_{F3}$), 73.4 ($C_{F4}$), 78.5 ($C_{F1}$), 115.4 ($CH_{Ar}$), 119.4 ($CH_{Ar}$), 119.8 ($CH_{Ar}$), 130.8 ($CH_{Ar}$), 137.4 ($C_{Ar}$), 159.0 ($C_{Ar}$), 170.2 (C=O), 177.8 (C=O). $R_f = 0.21$ (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z*

= 391 [(M-OH)$^+$, 100%]. HR-MS (ESI): calculated for $C_{20}H_{28}N_2O_7$ [M+Na]$^+$: 431.17887; found [M+Na]$^+$: 486.17851.

### *N*-[(1*S*,2*R*)-2-(3,5-Dihydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3c)

Compound **5**[13] was coupled with 3,5-diacetoxybenzoic acid using the general HBTU procedure. The crude was purified by flash chromatography (1:1.5 hexane-AcOEt, $R_f$ =0.33, 55 % yield) Zemplen deprotection and flash chromatography (85:15 $CHCl_3$ : MeOH, $R_f$ = 0.18) afforded **3c** (quant). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 0.89 (d, 1H, $J_{5\text{-}6}$ = 6.4 Hz, $H_{F6}$), 1.78-1.29 (m, 6H, $H_{Cy3}$, $2H_{Cy4}$, $2H_{Cy5}$, $H_{Cy6}$), 1.96-1.85 (m, 1H, $H_{Cy6}$), 2.16-2.07 (m, 1H, $H_{Cy3}$), 2.89 (td, 1H, $J$ = 8.4 & 4.2 Hz, $H_{Cy1}$), 3.53 (d, 1H, $J_{3\text{-}4}$ = 3.2 Hz, $H_{F4}$), 3.57 (q, 1H, $J_{5\text{-}6}$ = 6.6 Hz, $H_{F5}$), 3.73 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_{F3}$), 3.91 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.32-4.25 (m, 1H, $H_{Cy2}$), 5.45 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 6.37 (t, 1H, $J$ = 2.2 Hz, $CH_{Ar}$), 6.64 (d, 2H, $J$ = 2.2 Hz, $2CH_{Ar}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.7 ($C_{F6}$), 23.2 ($C_{Cy4}$), 24.3 ($C_{Cy5}$), 26.1 ($C_{Cy6}$), 30.6 ($C_{Cy3}$), 45.8 ($C_{Cy1}$), 50.2 ($C_{Cy2}$), 67.9 ($C_{F2}$), 68.4 ($C_{F5}$), 71.4 ($C_{F3}$), 73.3 ($C_{F4}$), 78.4 ($C_{F1}$), 106.7 ($CH_{Ar}$), 106.8 ($CH_{Ar}$), 127.7 ($CH_{ArTrp}$), 138.0 ($C_{Ar}$), 160.0 ($2xC_{Ar}$), 170.2 (C=O), 177.7 (C=O). $R_f$ = 0.17 (4:1 $CHCl_3$:MeOH). ESI-MS: $m/z$ = 447.4 [(M+Na)$^+$, 100%]. HR-MS (ESI): calculated for $C_{20}H_{28}N_2O_8$ [M+Na]$^+$: 447.17397; found [M+Na]$^+$: 486.17400.

### *N*-[(1*S*,2*R*)-2-(3-Pyridinecarboxamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3d)

Compound **5**[13] was coupled with nicotinic acid using the general HBTU procedure. The crude was purified by flash chromatography (AcOEt, $R_f$ = 0.30, 77% yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$:MeOH, $R_f$ = 0.33) afforded **3d** (95% yield). **$^1$H-NMR (400 MHz, $D_2O$) :** δ (ppm) = 0.70 (3H, d, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.35-1.91 (m, 8H, $CH_{2Cy}$), 2.91 (m, 1H, $H_{Cy1}$), 3.55 (m, 1H, $H_{F5}$), 3.36 (pseudo-d,1H, $J_{4\text{-}5}$ = 3.4 Hz, $H_{F4}$), 3.77 (dd, 1H, $J_{2\text{-}3}$ = 10.6 Hz, $H_{F3}$), 3.92 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $H_{F2}$), 4.47 (m, 1H, $H_{Cy2}$), 5.43 (d, 1H, $H_{F1}$), 7.52 (dd, 1H, $J$ = 5.0 Hz, $J$ = 7.7 Hz, $H_{Ar}$), 8.05 (d, 1H, $J$ = 8.0 Hz, $H_{Ar}$), 8.69 (m, 2H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $D_2O$) :** δ (ppm) = 16.2 ($C_{F6}$), 20.6, 23.6, 30.3 ($4xCH_{2Cy}$), 45.6 ($C_{Cy1}$), 49.7 ($C_{Cy2}$), 66.7 ($C_{F2}$), 67.9 ($C_{F5}$), 70.3 ($C_{F3}$), 72.1 ($C_{F4}$), 77.5 ($C_{F1}$), 137.1, 148.2,

152.4 ($4xC_{Ar}$), 169.1 (C=O), 178.6 (C=O). $R_f$ = 0.33 ($CHCl_3$/MeOH = 4/1). ESI-MS: $m/z$ = 394.4 [(M+H)$^+$, 43%]. HR-MS (ESI): calculated for $C_{19}H_{27}N_3O_6$ [M+Na]$^+$: 416.17921; found [M+Na]$^+$: 416.17906.

### *N*-[(1*S*,2*R*)-2-(3-Methoxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3e)

Compound **5**[13] was coupled with 3-methoxybenzoic acid using the general HBTU procedure. The crude was purified by flash chromatography (1:1.5 hexane-AcOEt, $R_f$ =0.31, 64 % yield) Zemplen deprotection and flash chromatography (85:15 $CHCl_3$ : MeOH, $R_f$ = 0.30) afforded **3e** (quant). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 0.83 (d, 1H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.79-1.34 (m, 6H, $H_{Cy3}$, $2H_{Cy4}$, $2H_{Cy5}$, $H_{Cy6}$), 2.02-1.89 (m, 1H, $H_{Cy6}$), 2.18-2.10 (m, 1H, $H_{Cy3}$), 2.92 (td, 1H, $J$ = 8.6 & 4.3 Hz, $H_{Cy1}$), 3.51 (d, 1H, $J_{3\text{-}4}$ = 2.5 Hz, $H_{F4}$), 3.56 (q, 1H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F5}$), 3.73 (dd, 1H, $J_{2\text{-}3}$ = 10.4 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_{F3}$), 3.83 (s, 3H, OMe), 3.91 (dd, 1H, $J_{2\text{-}3}$ = 10.4 Hz, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.40-4.33 (m, 1H, $H_{Cy2}$), 5.46 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 7.07 (td, 1H, $J$ = 6.0 & 2.7 Hz, $H_{Ar}$), 7.35 (m, 3H, $3H_{Ar}$);**$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.8 ($C_{F6}$), 23.2 ($C_{Cy5}$), 24.3 ($C_{Cy4}$), 26.0 ($C_{Cy6}$), 30.7 ($C_{Cy3}$), 45.9 ($C_{Cy1}$), 50.3 ($C_{Cy2}$), 55.9 (OMe), 67.9 ($C_{F2}$), 68.3 ($C_{F5}$), 71.4 ($C_{F3}$), 73.3 ($C_{F4}$), 78.4 ($C_{F1}$), 113.9 ($CH_{Ar}$), 118.5 ($CH_{Ar}$), 120.6 ($CH_{Ar}$), 130.7 ($CH_{Ar}$), 137.4 ($C_{Ar}$), 161.3 ($C_{ArOMe}$), 169.9 (C=O), 177.7 (C=O). $R_f$ = 0.40 ($CHCl_3$/MeOH = 4/1). ESI-MS: $m/z$ = 445.3 [(M+Na)$^+$, 100%]. HR-MS (ESI): calculated for $C_{21}H_{30}N_2O_7$ [M+Na]$^+$: 445.19452; found [M+Na]$^+$: 445.19401.

### *N*-[(1*S*,2*R*)-2-((2*R*)-2-Hydroxy-phenylacetamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3f)

Compound **5**[13] was coupled with (*R*)-(−)-α-acetoxyphenylacetic acid using the general acid chloride procedure. The crude was purified by flash chromatography (1:4 hexane:AcOEt, $R_f$ = 0.6, 60% yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.25) afforded **3f** (quant). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 0.96 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_{F6}$), 1.40-2.08 (m, 8H, $CH_{2Cy}$), 2.76 (m, 1H, $H_{Cy1}$), 3.57 (m, 1H, $H_{F4}$), 3.65 (dq, 1H, $J_{4\text{-}5}$ < 1Hz, $H_{F5}$), 3.73 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.6 Hz, $H_{F3}$), 3.95 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.16 (m, 1H, $H_{Cy2}$), 4.98 (s, 1H, -C*H*(OH)Ph), 5.51 (d, 1H, $H_{F1}$), 7.26-7.42 (m, 5H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 17.1 ($C_{F6}$), 23.3, 24.3, 26.7, 30.9 (4x$CH_{2Cy}$), 46.4 ($C_{Cy1}$, $C_{Cy2}$), 68.2 ($C_{F2}$), 68.5 ($C_{F5}$), 71.6 ($C_{F3}$), 73.5 ($C_{F4}$), 75.7 (-*C*H(OH)Ph), 78.4 ($C_{F1}$), 128.6 (2x$CH_{Ar}$), 129.2 ($CH_{Ar}$), 129.6 (2×$CH_{Ar}$), 141.8 ($C_{Ar}$), 174.6, 177.5 (2×C=O). $R_f$ = 0.25 (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z* = 445 [$(M+Na)^+$, 25%]. HR-MS (ESI): calculated for $C_{21}H_{30}N_2O_7$ $[M+Na]^+$: 445.19452; found $[M+Na]^+$: 445.19406.

### *N*-[(1*S*,2*R*)-2-(D-*N'*-Acetyl-triptophanoyl)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3g)

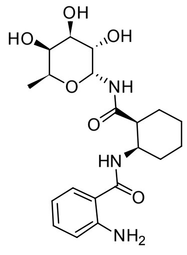
Compound **5**[13] was coupled with D-N-acetyl-triptophane using the general HBTU procedure. The crude was purified by flash chromatography (100:2 $CHCl_3$: MeOH, $R_f$ = 0.28, 85% yield). Zemplen deprotection and flash chromatography (85:15 $CHCl_3$: MeOH, $R_f$ = 0.23) afforded **3g** (quant). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.17 (d, 1H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.36-1.12(m 4H, $CH_{2Cy}$), 1.73-1.54 (m, 4H, $CH_{2Cy}$), 1.95 (s, 3H, $CH_3CONH$-), 2.70-2.63 (m, 1H, $H_{Cy1}$), 3.09 (dd, 1H, *J* = 14.4 & 7.4 Hz, $CH_{2Trp}$), 3.23 (dd, 1H, *J* = 14.4 & 7.2 Hz, $CH_{2Trp}$), 3.64 (d, 1H, $J_{3\text{-}4}$ = 3.0 Hz, $H_{4F}$), 3.80-3.74 (m, 2H, $H_{F3}$, $H_{F5}$), 3.95 (dd, 1H, $J_{2\text{-}3}$ = 10.3, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.16-4.10 (m, 1H, $H_{Cy2}$), 4.66 (t, 1H, *J* = 7.3 Hz, $CH_{Trp}$), 5.48 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 7.02 (t, 1H, *J* = 7.0 Hz, $CH_{ArTrp}$), 7.12-7.06 (m, 2H, 2$CH_{ArTrp}$), 7.33 (d, 2H, *J* = 8.1 Hz, $CH_{ArTrp}$), 7.60 (d, 1H, *J* = 7.8 Hz, $CH_{ArTrp}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 17.1 ($C_{F6}$), 22.6 ($C_{Cy4}$ or $C_{Cy5}$), 22.7 ($CH_3CONH$-), 24.5 ($C_{Cy4}$ or $C_{Cy5}$), 25.6 ($C_{Cy6}$), 29.3 ($CH_{2Trp}$), 30.7 ($C_{Cy3}$), 46.4 ($C_{Cy1}$), 49.1 ($C_{Cy2}$), 56.1 ($CH_{Trp}$), 68.0 ($C_{F2}$), 68.6 ($C_{F5}$), 71.4 ($C_{F3}$), 73.3 ($C_{F4}$), 78.3 ($C_{F1}$), 111.1 ($C_{ArTrp}$), 112.3 ($CH_{ArTrp}$), 119.4 ($CH_{ArTrp}$), 119.8 ($CH_{ArTrp}$), 120.4 ($CH_{ArTrp}$), 122.5 ($CH_{ArTrp}$), 124.5 ($CH_{ArTrp}$), 128.9 ($C_{ArTrp}$), 138.1 ($C_{ArTrp}$), 173.1 (C=O), 173.2 (C=O), 177.3 (C=O). $R_f$ = 0.30 (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z* = 539.4 [$(M+Na)^+$, 100%]. HR-MS (ESI): calculated for $C_{26}H_{36}N_4O_7$ $[M+Na]^+$: 539.24726; found $[M+Na]^+$: 539.24754.

### *N*-[(1*S*,2*R*)-2-( D-*N'*-Acetyl-phenylalanyl)cyclohexanecarboxyl]-α-L-fucopyranosylamine (3h)

Compound **5**[13] was coupled with D-N-acetyl-phenylalanine using the general HBTU procedure. The crude was purified by flash chromatography (1:9 hexane: AcOEt, $R_f$ = 0.33, 88% yield). Zemplen deprotection and flash chromatography (85:15 $CHCl_3$: MeOH, $R_f$ = 0.32) afforded **3h** (quant). **$^1$H-NMR (400 MHz,

**$CD_3OD$) :** δ (ppm) = 1.16 (d, 1H, $J_{5\text{-}6}$ = 6.4 Hz, $H_{F6}$), 1.48-1.27 (m, 4H, $CH_{2Cy}$), 1.70-1.54 (m, 2H, $CH_{2Cy}$), 1.83-1.71 (m, 2H, $CH_{2Cy}$), 1.91 (s, 3H, $CH_3CONH$-), 2.72-2.65 (m, 1H, $H_{Cy1}$), 2.85 (dd, 1H, *J* = 13.7 & 8.6 Hz, $CH_{2Phe}$), 3.08 (dd, 1H, *J* = 13.7 & 6.4 Hz, $CH_{2Phe}$), 3.62 (d, 1H, $J_{3\text{-}4}$ = 2.3 Hz, $H_{F4}$), 3.79-3.71 (m, 2H, $H_{F3}$, $H_{F5}$), 3.93 (dd, 1H, $J_{2\text{-}3}$ = 10.3, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.22-4.14 (m, 1H, $H_{Cy2}$), 4.60 (dd, 1H, *J* = 8.5 & 6.5 Hz, $CH_{Phe}$), 5.47 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 7.29-7.16 (m, 5H, 5$H_{ArPhe}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 17.1 ($C_{F6}$), 22.6 ($CH_3CONH$-), 22.7 ($C_{Cy5}$), 24.5 ($C_{Cy4}$), 25.6 ($C_{Cy6}$), 30.9 ($C_{Cy3}$), 39.3 ($CH_{2Phe}$), 46.4 ($C_{Cy1}$), 49.1 ($C_{Cy2}$), 56.4 ($CH_{Phe}$), 68.0 ($C_{F2}$), 68.6 ($C_{F5}$), 71.4 ($C_{F3}$), 73.3 ($C_{F4}$), 78.3 ($C_{F1}$), 127.8 ($CH_{ArPhe}$), 129.5 (2×$CH_{ArPhe}$), 130.3 (2×$CH_{ArPhe}$), 138.6 ($C_{ArPhe}$), 172.7 (C=O), 173.1 (C=O), 177.3 (C=O). $R_f$ = 0.41 (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z* = 500.4 [$(M+Na)^+$, 100%]. HR-MS (ESI): calculated for $C_{24}H_{35}N_3O_7$ $[M+Na]^+$: 500.23672; found $[M+Na]^+$: 500.23646.

### *N*-[(1*S*,2*R*)-2-(2-Aminobenzamido)cyclohexanecarboxyl]- α-L-fucopyranosylamine 3i

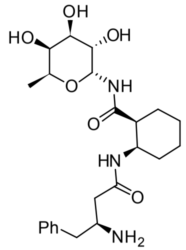


Compound **5** was coupled with *N*-2-Boc-aminobenzoic acid using the general HBTU procedure (see Supplementary Information) and the product (1:1 hexane: AcOEt, $R_f$ = 0.42) was purified by automated chromatography (55 % yield). The amide was dissolved in a mixture of $CH_2Cl_2$/TFA (1.5 mL, 5/1). The reaction mixture was stirred at r.t. during 2 h and was concentrated under vacuum. Zemplen deprotection of this crude and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.25) afforded the title product (97%).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 1.02 (d, 3H, $J_{5\text{-}6}$ = 6.3 Hz, $H_6$), 1.22-2.06 (m, 8H, $CH_{2cycl}$), 2.84 (m, 1H, $H_9$), 3.61 (pseudo-d, 1H, $J_{3\text{-}4}$ = 2.6 Hz, $H_4$), 3.75-3.81 (m, 2H, $H_3$, $H_5$), 3.96 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $J_{2\text{-}3}$ = 10.2 Hz, $H_2$), 4.52 (m, 1H, $H_{10}$), 5.50 (m, 1H, $H_1$), 5.99 (s, 1H, -CH(OAc)), 7.03 (d, 1H, $J_{10\text{-}NH}$ = 8.6 Hz, $NH_{11}$), 7.08 (d, 1H, $NH_7$), 7.10-7.82 (m, 4H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 16.06 ($C_6$), 22.22, 24.86, 25.15, 30.90 (4x$CH_{2cyl}$), 46.96 ($C_9$), 49.37 ($C_{10}$), 67.88 ($C_5$), 68.56 ($C_2$), 71.46

($C_3$), 73.36 ($C_4$), 78.28 ($C_1$), 177.97 (2xC=O). ESI-MS: *m/z* = 407 $[(M+Na)^+$, 100%]. $R_f$ = 0.25 ($CHCl_3$:MeOH 4:1).

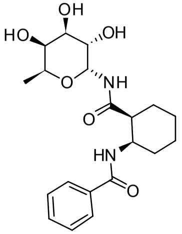
**N-[(1S,2R)-2-(L-Homophenylalanyl)cyclohexancarboxyl]- α-L-fucopyranosylamine 3j**



Compound **5** was coupled with *N*-Boc-homophenylalanine using the general HBTU procedure (see Supplementary Information) and the product ($R_f$ = 0.75, AcOEt) was isolated by automated chromatography (72 % yield). The amide was dissolved in a mixture of $CH_2Cl_2$/TFA (1.5 mL, 5/1). The reaction mixture was stirred at r.t. during 2 h and was concentrated under vacuum. Zemplen deprotection of this crude and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.2) afforded the title product (75%).

**$^1$H-NMR (400 MHz, D$_2$O):** δ (ppm) = 1.16 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_6$), 1.44-1.76 (m, 8H, $CH_{2cycl}$), 2.68 (dd, 1H, $J_{gem}$ = 16.0 Hz, $J_{CH2\text{-}CH}$ = 7.5 Hz, -$C\mathit{H}H'CHNH_2$-), 2.76 (dd, 1H, $J_{CH2\text{-}CH}$ = 7.5 Hz, -$CH\mathit{H}'CHNH_2$-), 2.85 (m, 1H, $H_9$), 3.10 (d, 2H, $J_{CH\text{-}CH2}$ = 7.4 Hz, -$CHCH_2Ph$), 3.83-3.87 (m, 2H, $H_4$, $H_5$), 3.92 (dd, 1H, $J_{2\text{-}3}$ = 10.6 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_3$), 3.97 (m, 1H, -$C\mathit{H}CH_2Ph$), 4.07 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $H_2$), 4.46 (m, 1H, $H_{10}$), 5.57 (m, 1H, $H_1$), 7.40-7.55 (m, 5H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, D$_2$O):** δ (ppm) = 16.48 ($C_6$), 20.80, 23.11, 23.43, 30.47 (4x$CH_{2cyl}$), 36.98 (-$CH\mathit{C}H_2Ph$), 38.46 (-$\mathit{C}H_2CHNH_2$-), 46.18 ($C_9$), 48.20 ($C_{10}$), 50.98 (-$CH_2\mathit{C}HNH_2$-), 66.66 ($C_2$), 68.01 ($C_5$), 70.19 ($C_3$), 72.11 ($C_4$), 77.12 ($C_1$), 128.38, 129.85, 130.10 (5x$C_{Ar}$), 135.79 ($C_{ipso}$), 171.46, 178.42 (2xC=O). ESI-MS: *m/z* = 450.5 $[(M+1)^+$, 100%]. $R_f$ = 0.20 ($CHCl_3$:MeOH 4:1).

**N-[(1S,2R)-2-(Benzamido)cyclohexanecarboxyl]- α-L-fucopyranosylamine 3k**

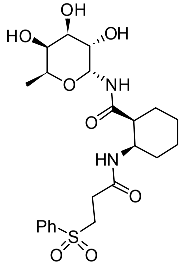


Compound **5** was coupled with benzoic acid using the general acid chloride procedure (see Supplementary Information) and the product ($R_f$ = 0.87, AcOEt) was isolated by automated chromatography (80 % yield). Zemplen deprotection

and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.17) afforded the title product (quant).

**$^1$H-NMR (400 MHz, CD$_3$OD):** δ (ppm) = 0.73 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_6$), 1.13-2.06 (m, 8H, $CH_{2cycl}$), 2.83 (m, 1H, $H_9$), 3.42 (pseudo-d, 1H, $J_{3\text{-}4}$ = 2.6 Hz, $H_4$), 3.47 (m, 1H, $H_5$), 3.64 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.3 Hz, $H_3$), 3.82 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_2$), 4.30 (m, 1H, $H_{10}$), 5.37 (m, 1H, $H_1$), 7.33-7.43 (m, 3H, $H_{Ar}$), 7.65-7.72 (m, 2H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, CD$_3$OD):** δ (ppm) = 16.86 ($C_6$), 23.22, 24.44, 26.04, 30.84 (4x$CH_{2cyl}$), 46.06 ($C_9$), 50.32 ($C_{10}$), 68.00 ($C_2$), 68.36 ($C_5$), 71.54 ($C_3$), 73.36 ($C_4$), 78.50 ($C_1$), 170.14, 177.97 (2xC=O). ESI-MS: *m/z* = 415 $[(M+Na)^+$, 20%]. $R_f$ = 0.17 ($CHCl_3$:MeOH 4:1).

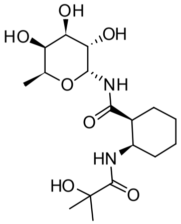
**N-[(1S,2R)-2-(3-(Phenylsulfonyl)propanamido)cyclohexancarboxyl)- α-L-fucopyranosylamine 3l**



Compound **5** was coupled with (phenylsulfonyl)propionic acid using the general acid chloride procedure (see Supplementary Information) and the product ($R_f$ = 0.59, 1:4 hexane:AcOEt) was isolated by automated chromatography (47 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.15) afforded the title product (quant).

**$^1$H-NMR (400 MHz, D$_2$O):** δ (ppm) = 1.12 (3H, d, $J_{5\text{-}6}$ = 6.5 Hz, $H_6$), 1.30-1.82 (m, 8H, $CH_{2CYCL}$), 2.68-2.82 (m, 3H, $H_9$, -$CH_2SO_2Ph$), 3.65 (m, 1H, -$C\mathit{H}H'CH_2SO_2Ph$), 3.76-3.82 (m, 3H, $H_4$, $H_5$, -$CH\mathit{H}'CH_2SO_2Ph$), 3.91 (dd, 1H, $J_{2\text{-}3}$ = 10.5 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_3$), 4.08 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $H_2$), 5.56 (d, 1H, $H_1$), 7.75-8.01 (m, 5H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, D$_2$O):** δ (ppm) = 16.40 ($C_6$), 21.11, 23.9 (3x$CH_{2CYCL}$), 29.34 (-$CH_2SO_2Ph$), 30.27 ($CH_{2CYCL}$), 45.63 ($C_9$), 48.68 ($C_{10}$), 51.75 (-$\mathit{C}H_2CH_2SO_2Ph$), 66.64 ($C_2$), 67.58 ($C_5$), 69.86 ($C_3$), 72.09 ($C_4$), 77.24 ($C_1$), 128.57, 130.45, 135.65 (3x$C_{Ar}$), 137.10 ($C_{ipso}$), 171.00 ($C_{12}$), 178.33 ($C_8$). ESI-MS: *m/z* = 485 $[(M+1)^+$, 25%]. $R_f$ = 0.15 ($CHCl_3$:MeOH 4:1).

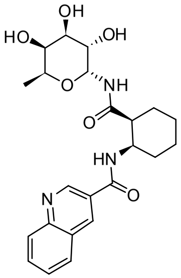
### *N*-[(1*S*,2*R*)-2-(2-Hydroxy-2-methylpropanamido)cyclohexancarboxyl]- α-L-fucopyranosylamine 3m



Compound **5** was coupled with 2-acetoxy2-methylpropanoic acid using the general acid chloride procedure (see Supplementary Information) and the product ($R_f$ = 0.59, 1:4 hexane:AcOEt) was isolated by automated chromatography (46 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.15) afforded the title product (95%).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 1.17 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_6$), 1.36 (s, 6H, -$C(CH_3)_2$), 1.31-2.03 (m, 8H, $CH_{2cycl}$), 2.80 (m, 1H, $H_9$), 3.64 (m, 1H, $H_4$), 3.76 (m, 1H, $H_5$), 3.79 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_3$), 3.96 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_2$), 4.14 (m, 1H, $H_{10}$), 5.50 (d, 1H, $H_1$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 15.61 ($C_6$), 21.46, 23.00, 24.96, 29.46 (4x$CH_{2cyl}$), 26.35, 26.46 (2x-$C(CH_3)_2$), 44.84 ($C_9$, $C_{10}$), 66.55 ($C_2$), 67.00 ($C_5$), 70.03 ($C_3$), 71.91 ($C_4$), 73.2 (-$C(CH_3)_2$), 76.88 ($C_1$), 177.39, 178.89 (2xC=O). ESI-MS: *m/z* = 397 [$(M+Na)^+$, 38%]. $R_f$ = 0.15 ($CHCl_3$:MeOH 4:1).

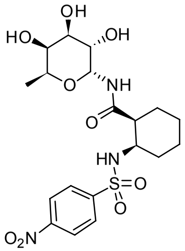
### *N*-[(1*S*,2*R*)-2-(Quinoline-3-carboxamido)cyclohexanecarboxyl]- α-L-fucopyranosylamine 3n



Compound **5** was coupled with quinoline-3-carboxylic acid using the general acid chloride procedure and the product ($R_f$ = 0.45, 1:4 hexane:AcOEt) was isolated by automated chromatography (48 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.22) afforded the title product (quant).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 0.74 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_6$), 11.46-2.15 (m, 8H, $CH_{2cycl}$), 2.96 (m, 1H, $H_9$), 3.53 (m, 1H, $H_4$), 3.61 (m, 1H, $H_5$), 3.77 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.3 Hz, $H_3$), 3.93 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_2$), 4.57 (m, 1H, $H_{10}$), 5.50 (d, 1H, $H_1$), 7.72 (m, 1H, $H_{19}$), 7.91 (m, 1H, $H_{18}$), 8.08, 8.10 (2xm, 2H, $H_{17}$, $H_{20}$), 8.78 (m, 1H, $H_{22}$), 9.20 (m, 1H, $H_{14}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 16.89 ($C_6$), 22.91, 24.65, 25.55, 31.12 (4x$CH_{2cyl}$), 46.30 ($C_9$), 50.34 ($C_{10}$), 68.00 ($C_2$), 68.38 ($C_5$), 71.55 ($C_3$), 73.30 ($C_4$), 78.46 ($C_1$), 128.58 ($C_{21}$), 129.15 ($C_{13}$), 129.06, 129.32, 130.45, 132.96, 138.02 ($C_{17}$, $C_{18}$, $C_{19}$, $C_{20}$, $C_{22}$), 149.82 ($C_{16}$), 150.22 ($C_{14}$), 168.02 ($C_{12}$), 177.73 ($C_8$). ESI-MS: *m/z* = 444 [$(M+H)^+$, 100%]. $R_f$ = 0.22 ($CHCl_3$:MeOH 4:1).

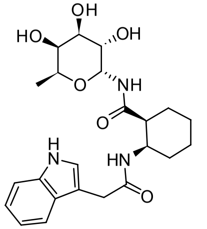
### *N*-[(1*S*,2*R*)-2-(4-Nitrobenzenesulfonamido)cyclohexanecarboxyl)-α-L-fucopyranosylamine 3o



Compound **5** was coupled with 4-nitrobenzenesulfonyl chloride and the product ($R_f$ = 0.68, 1:4 hexane:AcOEt) was isolated by automated chromatography (69 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.21) afforded the title product (63%).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 1.19 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_6$), 1.31-1.86 (m, 8H, $CH_{2cycl}$), 2.73 (m, 1H, $H_9$), 3.68-3.72 (m, 2H, $H_4$, $H_{10}$), 3.79 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_3$), 3.93 (m, 1H, $H_5$), 3.96 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_2$), 5.50 (d, 1H, $H_1$), 8.14, 8.41 (2xd 4H, *J* = 8.4 Hz, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 16.88 ($C_6$), 22.62, 23.97, 26.11, 30.85 (4x$CH_{2cyl}$), 47.24 ($C_9$), 53.91 ($C_{10}$), 68.01 ($C_2$), 68.50 ($C_5$), 71.48 ($C_3$), 73.29 ($C_4$), 78.23 ($C_1$), 125.35, 129.41 (4x$C_{Ar}$), 148.74, 151.33 (2x$C_{ipso}$), 177.00 (2xC=O). ESI-MS: *m/z* = 474 [$(M+H)^+$, 78%]. $R_f$ = 0.21 ($CHCl_3$:MeOH 4:1).

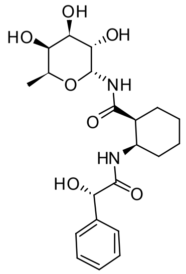
### *N*-[(1*S*,2*R*)-2-(3-Indolacetamido)cyclohexanecarboxyl)- α-L-fucopyranosylamine 3p



Compound **5** was coupled with 3-indolacetic acid acid using the general HBTU procedure and the product ($R_f$ = 0.68, AcOEt) was isolated by automated chromatography (48 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.11) afforded the title product (quant).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 1.14 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_6$), 1.30-1.84 (m, 8H, $CH_{2cycl}$), 2.84 (m, 1H, $H_9$), 3.74-3.91 (m, 5H, $CH_{2INDOL}$, $H_3$, $H_4$, $H_5$), 4.07 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $J_{2\text{-}3}$ = 10.6 Hz, $H_2$), 4.39 (m, 1H, $H_{10}$), 5.56 (d, 1H, $H_1$), 7.24-7.39 (m, 3H, $H_{Ar}$), 7.61 (d, 1H, *J* = 8.2 Hz, $H_{Ar}$), 7.70 (d, 1H, *J* = 7.9 Hz, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 16.50 ($C_6$), 21.01, 23.62, 23.81, 30.55 (4x$CH_{2cyl}$), 33.22 ($CH_{2INDOL}$), 45.73 ($C_9$), 48.81 ($C_{10}$), 66.80 ($C_2$), 67.96 ($C_5$), 70.34 ($C_3$), 72.23 ($C_4$), 77.46 ($C_1$), 108.38 ($C_{ipso\text{-}14}$), 112.67, 119.25, 120.28, 122.88, 125.60 (5x$C_{Ar}$), 128.56, 137.07 (2x$C_{ipso}$), 174.84 ($C_{12}$), 178.39 ($C_8$). ESI-MS: *m/z* = 446.3 [$(M+H)^+$, 100%].. $R_f$ = 0.11 ($CHCl_3$:MeOH 4:1).

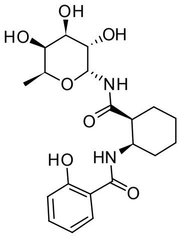
### *N*-[(1*S*,2*R*)-2-((*S*)-2-Hydroxyphenylacetamido)cyclohexanecarboxyl]- α-L-fucopyranosylamine 3q



Compound **5** was coupled with (*S*)-(+)-α-acetoxyphenylacetic acid using the general acid chloride procedure and the product ($R_f$ = 0.38, 1:1 hexane:AcOEt) was isolated by automated chromatography (47 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.72) afforded the title product (quant).

**$^1$H-NMR (400 MHz, $D_2O$):** δ (ppm) = 1.18 (d, 3H, $J_{5\text{-}6}$ = 6.4 Hz, $H_6$), 1.40-1.93 (m, 8H, $CH_{2cycl}$), 2.94 (m, 1H, $H_9$), 3.78-3.81 (m, 2H, $H_4$, $H_5$), 3.93 (dd, 1H, $J_{2\text{-}3}$ = 10.6 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_3$), 4.06 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $H_2$), 4.38 (m, 1H, $H_{10}$), 5.21 (s, 1H, -*CH*(OH)Ph), 5.58 (d, 1H, $H_1$), 7.45-7.52 (m, 5H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $D_2O$):** δ (ppm) = 16.29 ($C_6$), 20.71, 23.36, 23.67, 30.25 (4x$CH_{2cyl}$), 45.44 ($C_9$), 48.41 ($C_{10}$), 66.47 ($C_2$), 67.82 ($C_5$), 70.07 ($C_3$), 71.97 ($C_4$), 74.52 (-*C*H(OH)Ph), 77.14 ($C_1$), 127.61, 129.58 (5x$C_{Ar}$), 139.16 ($C_{ipso}$), 174.42, 178.25 (2xC=O). ESI-MS: *m/z* = 405.5 [$(M-OH)^+$, 25%]. $R_f$ = 0.72 ($CHCl_3$:MeOH 4:1).

### *N*-((1*S*,2*R*)-2-(2-Hydroxybenzamido)cyclohexanecarboxyl)- α-L-fucopyranosylamine 3r



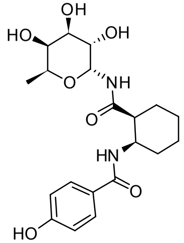
Compound **5** was coupled with 2-acetoxybenzoic acid using the general acid chloride procedure and the product ($R_f$ = 0.3, 1:1 hexane:AcOEt) was isolated by automated chromatography (50 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.75) afforded the title product (quant).

**$^1$H-NMR (400 MHz, $D_2O$):** δ (ppm) = 0.78 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_6$), 1.48-2.05 (m, 8H, $CH_{2cycl}$), 3.00 (m, 1H, $H_9$), 3.61 (dq, 1H, $J_{4\text{-}5}$ < 1 Hz, $H_5$), 3.67 (dd, 1H, $J_{3\text{-}4}$ = 3.4 Hz, $H_4$), 3.83 (dd, 1H, $J_{2\text{-}3}$ = 10.6 Hz, $H_3$), 4.00 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_2$), 4.56 (m, 1H, $H_1$), 5.48 (d, 1H, $H_1$), 7.08, 7.49, 7.84 (3xm, 4H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $D_2O$):** δ (ppm) = 16.10 ($C_6$), 21.62,

23.90, 24.21, 30.39 (4x$CH_{2cyl}$), 45.59 ($C_9$), 49.58 ($C_{10}$), 66.82 ($C_2$), 67.89 ($C_5$), 70.43 ($C_3$), 72.27 ($C_4$), 77.66 ($C_1$), 117.96 ($C_{Ar}$), 118.20 ($C_{ipso}$), 121.31, 130.36, 134.99 (3x$C_{Ar}$), 157.28 ($C_{ipso}$), 169.07, 178.63 (2xC=O). ESI-MS: *m/z* = 431.5 [$(M+Na)^+$, 30%]. $R_f$ = 0.75 ($CHCl_3$:MeOH 4:1).

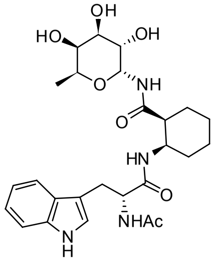
### *N*-[(1 *S*,2*R*)-2-(4-Hydroxybenzamido)cyclohexanecarboxyl]- α-L -fucopyranosylamine 3s

Compound **5** was coupled with 4-acetoxybenzoic acid using the general acid chloride procedure and the product ($R_f$ = 0.30, 1:1 hexane:AcOEt) was isolated by automated chromatography (55 % yield). Zemplen deprotection and flash chromatography (8:2 $CHCl_3$: MeOH, $R_f$ = 0.46) afforded the title product (quant).



**$^1$H-NMR (400 MHz, $D_2O$):** δ (ppm) = 0.66 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_6$), 1.31-1.95 (m, 8H, $CH_{2cycl}$), 2.94 (m, 1H, $H_9$), 3.46 (m, 1H, $H_5$), 3.59 (pseudo-d, 1H, $J_{3\text{-}4}$ = 3.4 Hz, $H_4$), 3.76 (dd, 1H, $J_{2\text{-}3}$ = 10.6 Hz, $H_3$), 3.92 (dd, 1H, $J_{1\text{-}2}$ = 5.7 Hz, $H_2$), 4.40 (m, 1H, $H_{10}$), 5.42 (d, 1H, $H_1$), 6.91, 7.60 (2xd, 4H, $H_{Ar}$). **$^{13}$C-NMR (100 MHz, $D_2O$):** δ (ppm) = 16.04 ($C_6$), 21.58, 23.57, 23.92, 30.25 (4x$CH_{2cyl}$), 45.37 ($C_9$), 49.758 ($C_{10}$), 66.67 ($C_5$), 67.82 ($C_2$), 70.27 ($C_3$), 72.15 ($C_4$), 77.52 ($C_1$), 116.09 (2x$C_{Ar}$), 126.46 ($C_{ipso}$), 130.38 (2x$C_{Ar}$), 159.96 ($C_{ipso}$), 170.88, 178.75 (2xC=O). ESI-MS: *m/z* = 431.5 [$(M+Na)^+$, 30%]. $R_f$ = 0.46 ($CHCl_3$:MeOH 4:1).

### *N*-[(1*S*,2*R*)-2-(D-*N*-Acetyl-Tryptophanoyl)cyclohexanecarboxyl]- α-L-fucopyranosylamine 3t



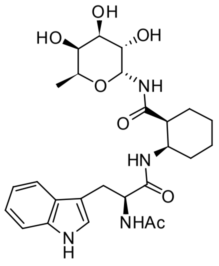
Compound **5** was coupled with N-acetyl-D-tryptophan using the general HBTU procedure and the product ($R_f$ = 0.48, 1:9 hexane:AcOEt) was isolated by automated chromatography (85 % yield). Zemplen deprotection and flash chromatography (4:1 $CHCl_3$: MeOH, $R_f$ = 0.22) afforded the title product (quant.).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 1.17 (d, 1H, *J*=6.5 Hz, $3H_{6F}$), 1.36-

1.12(m 4H, $CH_{2cycl}$), 1.73-1.54 (m, 4H, $CH_{2cycl}$), 1.95 (s, 3H, $CH_3CONH$-), 2.70-2.63 (m, 1H, $H_1$), 3.09 (dd, 1H, *J*=14.4 & 7.4 Hz, $CH_{2Trp}$), 3.23 (dd, 1H, *J*=14.4 & 7.2 Hz, $CH_{2Trp}$), 3.64 (d, 1H, *J*=3.0 Hz, $H_{4F}$), 3.80-3.74 (m, 2H, *J*=2.8 Hz, $H_{3F} + H_{5F}$), 3.95 (dd, 1H, *J*=10.3 & 5.6 Hz, $H_{2F}$), 4.16-4.10 (m, 1H, $H_2$), 4.66 (t, 1H, *J*=7.3 Hz, $CH_{Trp}$), 5.48 (d, 1H, *J*=5.6 Hz, $H_{1F}$), 7.02 (t, 1H, *J*=7.0 Hz, $CH_{ArTrp}$), 7.12-7.06 (m, 2H, $2CH_{ArTrp}$), 7.33 (d, 2H, *J*=8.1 Hz, $CH_{ArTrp}$), 7.60 (d, 1H, *J*=7.8 Hz, $CH_{ArTrp}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 17.1 ($C_{6F}$), 22.6 ($C_4$ or $C_5$), 22.7 ($CH_3CONH$-), 24.5 ($C_4$ or $C_5$), 25.6 ($C_6$), 29.3 ($CH_{2Trp}$), 30.7 ($C_3$), 46.4 ($C_1$), 49.1 ($C_2$), 56.1 ($CH_{Trp}$), 68.0 ($C_{2F}$), 68.6 ($C_{5F}$), 71.4 ($C_{3F}$), 73.3 ($C_{4F}$), 78.3 ($C_{1F}$), 111.1 ($C_{ArTrp}$), 112.3 ($CH_{ArTrp}$), 119.4 ($CH_{ArTrp}$), 119.8 ($CH_{ArTrp}$), 120.4 ($CH_{ArTrp}$), 122.5 ($CH_{ArTrp}$), 124.5 ($CH_{ArTrp}$), 128.9 ($C_{ArTrp}$), 138.1 ($C_{ArTrp}$), 173.1 (C=O), 173.2 (C=O), 177.3 (C=O). ESI-MS: *m/z* = 539.4 [$(M+Na)^+$, 100%].

**N-[(1 *S*,2*R*)-2-(L-*N*-Acetyl-tryptophanoyl)cyclohexanecarboxyl]- α-L-fucopyranosylamine 3u**

Compound **5** was coupled with *N*-acetyl-L-tryptophan using the general HBTU procedure and the product



($R_f$ = 0.26, 1:9 hexane:AcOEt) was isolated by automated chromatography (71% yield). Zemplen deprotection and flash chromatography (4:1 $CHCl_3$: MeOH, $R_f$ = 0.32) afforded the title product (quant.).

**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 1.10 (d, 1H, *J*=6.4 Hz, $3H_{6F}$), 1.73-1.23 (m, 7H, $H_3 + 2H_4 + 2H_5 + 2H_6$), 1.87 (s, 3H, $CH_3CONH$-), 2.03-1.81 (m, 1H, $H_3$), 2.75-2.69 (m, 1H, $H_1$), 3.06 (dd, 1H, *J*=14.9 & 9.1 Hz, $CH_{2Trp}$), 3.22 (dd, 1H, *J*=14.9 & 4.8 Hz, $CH_{2Trp}$), 3.59 (d, 1H, *J*=3.0 Hz, $H_{4F}$), 3.79-3.70 (m, 2H, $H_{3F} + H_{5F}$), 3.93 (dd, 1H, *J*=10.3 & 5.6 Hz, $H_{2F}$), 4.24-4.16 (m, 1H, $H_2$), 4.67 (dd, 1H, *J*=9.0 & 4.9 Hz, $CH_{Trp}$), 5.50 (d, 1H, *J*=5.6 Hz, $H_{1F}$), 6.99 (t, 1H, *J*=7.0 Hz, $CH_{ArTrp}$), 7.11-7.04 (m, 2H, $2CH_{ArTrp}$), 7.31 (d, 2H, *J*=8.1 Hz, $CH_{ArTrp}$), 7.61 (d, 1H, *J*=7.8 Hz, $CH_{ArTrp}$).

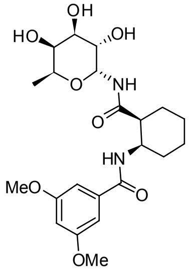
**$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 17.1 ($C_{6F}$), 22.5 ($C_4$ or $C_5$), 23.0 ($CH_3CONH$-), 24.4 ($C_4$ or $C_5$), 26.3 ($C_6$), 29.0 ($CH_{2Trp}$), 30.8 ($C_3$), 46.4 ($C_1$), 49.3 ($C_2$), 55.7 ($CH_{Trp}$), 68.1 ($C_{2F}$), 68.6 ($C_{5F}$), 71.5 ($C_{3F}$), 73.3 ($C_{4F}$), 78.3 ($C_{1F}$), 111.3 ($CH_{ArTrp}$), 112.2 ($CH_{ArTrp}$), 119.5 ($CH_{ArTrp}$), 119.8 ($CH_{ArTrp}$), 122.4 ($CH_{ArTrp}$), 124.4

---

($CH_{ArTrp}$), 128.9 ($C_{ArTrp}$), 138.1 ($C_{ArTrp}$), 173.3 (C=O), 173.7 (C=O), 177.4 (C=O). ESI-MS: *m/z* = 539.4 [$(M+Na)^+$, 100%].

***N*-[(1*S*,2*R*)-2-(3,5-Dimethoxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine 3v**

Compound **5** was coupled with 3,5-dimethoxybenzoic acid using the general HBTU procedure and the product ($R_f$ = 0.32, 1:1.5 hexane:AcOEt) was isolated by automated chromatography (87 % yield). Zemplen deprotection and flash chromatography (4:1 $CHCl_3$: MeOH, $R_f$ = 0.61) afforded the title product (quant.).



**$^1$H-NMR (400 MHz, $CD_3OD$):** δ (ppm) = 0.85 (d, 1H, *J*=6.4 Hz, $3H_{6F}$), 1.81-1.34 (m, 6H, $H_3 + 2H_4 + 2H_5 + H_6$), 2.00-1.90 (m, 1H, $H_6$), 2.19-2.10 (m, 1H, $H_3$), 2.93 (td, 1H, *J*=8.5 & 4.4 Hz, $H_1$), 3.51 (d, 1H, *J*=2.6 Hz, $H_{4F}$), 3.55 (q, 1H, *J*=6.6 Hz, $H_{5F}$), 3.73 (dd, 1H, *J*=10.3 & 3.4 Hz, $H_{3F}$), 3.81 (s, 6H, 2 x OMe), 3.91 (dd, 1H, *J*=10.3 & 5.6 Hz, $H_{2F}$), 4.37-4.31 (m, 1H, $H_2$), 5.47 (d, 1H, *J*=5.6 Hz, $H_{1F}$), 6.61 (t, 1H, *J*=2.2 Hz, $H_{Ar}$), 6.91 (d, 2H, *J*=2.2 Hz, $2H_{Ar}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$):** δ (ppm) = 16.7 ($C_{6F}$), 23.2 ($C_5$), 24.2 ($C_4$), 26.0 ($C_6$), 30.6 ($C_3$), 45.8 ($C_1$), 50.4 ($C_2$), 56.1 (2 x OMe), 67.9 ($C_{2F}$), 68.3 ($C_{5F}$), 71.4 ($C_{3F}$), 73.3 ($C_{4F}$), 78.4 ($C_{1F}$), 104.5 ($1CH_{Ar}$), 106.5 ($2CH_{Ar}$), 162.4 ($2C_{ArOMe}$), 169.9 (C=O), 177.7 (C=O). ESI-MS: *m/z* = 475.3 [$(M+Na)^+$, 100%].

**(1*S*,2*S*)-2-Aminocyclohexanecarboxylic acid series: Ligands 11**

***N*-((1*S*,2*S*)-2-Acetamido-cyclohexanecarboxyl)-α-L-fucopyranosylamine (11a)**

The crude hydrogenation product of **29** (see Supplementary Information – SI-Scheme 3) was used in the general acetylation method (see Supplementary Information). The product was purified by flash chromatography on silica gel (AcOEt, $R_f$ = 0.21). Yield: 16 mg (73 %). The Zemplen deprotection was performed and the product was not further purified. ($R_f$ = 0.10, $CHCl_3$/MeOH 4:1) Yield (**11a**): 6 mg (55 %)

**$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.17(d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.21 – 1.35 (m, 2H, $H_{Cy3ax}$ and

$H_{Cy5ax}$), 1.35 – 1.45 (m, 1H, $H_{Cy4ax}$), 1.47 – 1.57 (m, 1H, $H_{Cy6ax}$), 1.71 – 1.79 (m, 2H, $H_{Cy4eq}$ and $H_{Cy5eq}$), 1.88 (s, 3H, Ac-Me), 1.89 – 1.96 (m, 2H, $H_{Cy3eq}$ and $H_{Cy6eq}$), 2.36 (td, 1H, $J_{1\text{-}6a}$ = 3.6 Hz, $J_{1\text{-}6b}$ = 11.7 Hz, $J_{1\text{-}2}$ = 11.7 Hz , $H_{Cy1}$ ), 3.63 – 3.66 (m, 1H, $H_{F4}$), 3.73 (dd, $J_{3\text{-}4}$ = 3.4 Hz, $J_{2\text{-}3}$ = 10.3 Hz, 1H, $H_{F3}$), 3.77 (q, $J_{5\text{-}6}$ = 6.5 Hz ,1H, $H_{F5}$), 3.87 – 3.96 (m, 2H, $H_{C2}$, $H_{F2}$), 5.47 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $HF_1$) **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 17.1 ($C_{F6}$), 23.0 (Ac-$CH_3$), 26.1 ($C_{Cy4}$ or $C_{Cy5}$), 26.2 ($C_{Cy4}$ or $C_{Cy5}$), 30.9 ($C_{Cy6}$), 33.8 ($C_{Cy3}$), 51.2 ($C_{Cy2}$), 51.7 ($C_{Cy1}$), 68.1 ($C_{F2}$), 68.6 ($C_{F5}$), 71.7 ($C_{F3}$), 73.3 ($C_{F4}$), 78.6 ($C_{F1}$), 172.4 (NHCO), 177.8 (NHCO). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{15}H_{26}N_2O_6$: 353.16831 $[M + Na]^+$; found: 353.16832. $[\alpha]_D$ = -140.8 (c 0.05, MeOH)-

**N-[(1*S*,2*S*)-2-(3-Hydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (11b)**

The crude hydrogenation product of **29** (see Supplementary Information – SI-Scheme 3) was coupled with 3-hydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (AcOEt, $R_f$ = 0.15, 71% yield). Zemplen deprotection and flash chromatography (4:1 chloroform:methanol, $R_f$ = 0.15) afforded **11b**. **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 0.59 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.25 – 1.30 (m, 1H, $H_{Cy4ax}$ or $H_{Cy5ax}$), 1.36 – 1.42 (m, 1H, $H_{Cy3ax}$ and $H_{Cy4ax}$ or $H_{Cy5ax}$), 1.51 – 1.64 (m, 1H, $H_{Cy6ax}$), 1.69 – 1.79 (m, 2H, $H_{Cy4eq}$ and $H_{Cy5eq}$), 1.81 – 1.89 (m, 1H, $H_{Cy6eq}$), 1.91 – 1.98 (m, 1H, $H_{Cy3eq}$), 2.48 (td, 1H, $J_{1\text{-}6a}$ = 3.5 Hz, $J_{1\text{-}6b}$ = 11.7 Hz, $J_{1\text{-}2}$ = 11.7 Hz, $H_{Cy1}$), 3.32 – 3.34 (m, 1H, $H_{F4}$), 3.40 (q, $J_{5\text{-}6}$ = 6.3 Hz, 1H, $H_{F5}$), 3.62 (dd, 1H, $J_{3\text{-}4}$ = 3.4 Hz, $J_{2\text{-}3}$ = 10.4 Hz, $H_{F3}$), 3.81 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $J_{2\text{-}3}$ = 10.3 Hz, $H_{F2}$), 4.01 – 4.09 (m, 1H, $H_{C2}$), 5.31 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 6.74 – 6.78 (m, 1H, $H_{Ar}$), 7.01 – 7.08 (m, 3H, Ar). **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.5 ($C_{F6}$), 25.9 ($C_{Cy4}$ or $C_{Cy5}$), 26.3 ($C_{Cy4}$ or $C_{Cy5}$), 30.2 ($C_{Cy6}$), 33.9 ($C_{Cy3}$), 51.3 ($C_{Cy1}$), 52.0 ($C_{Cy2}$), 67.9 ($C_{F2}$), 68.6 ($C_{F5}$), 71.5 ($C_{F3}$), 73.2 ($C_{F4}$), 78.8 ($C_{F1}$), 116.5 (Ar), 116.7 (Ar), 122.3 (Ar), 130.4 (Ar), 136.37 (Cquart., Ar.), 164.59 (Cquart., Ar.), 169.81 (Cquart.), 178.13 (Cquart.). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{20}H_{28}N_2O_7$: 431.17887 $[M + Na]^+$; found: 431.17951. $[\alpha]_D$ = -21.0 (c = 0.25, MeOH).

**N-[(1*S*,2*S*)-2-(3,5-Dihydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (11c)**

The crude hydrogenation product of **29** (see Supplementary Information – SI-Scheme 3) was coupled with 3,5-dihydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (7:3 ethyl acetate:*n*-hexane, $R_f$ = 0.23). Yield: 19 mg (26 %). Zemplen deprotection and flash chromatography (4:1 chloroform:methanol, $R_f$ = 0.14) afforded **11c**. Yield: 9 mg (69 %)- **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 0.73 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.28 – 1.46 (m, 3H, $H_{Cy5ax}$, $H_{Cy4ax}$, $H_{Cy3ax}$), 1.51 – 1.61 (m, 1H, $H_{Cy6ax}$), 1.69 – 1.78 (m, 1H, $H_{Cy5eq}$, $H_{Cy4eq}$), 1.84 – 1.90 (m, 1H, $H_{Cy6eq}$), 1.90 – 1.97 (m, 1H, $H_{Cy3eq}$), 2.55 (td, $J_{1\text{-}6a}$ = 3.5 Hz, $J_{1\text{-}2}$ = 11.6 Hz, $J_{1\text{-}6b}$ = 11.6 Hz, 1H, $H_{Cy1}$), 3.40 – 3.42 (m, 1H, $H_{F4}$), 3.49 (q, $J_{5\text{-}6}$ = 6.5 Hz, 1H, $H_{F5}$), 3.64 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{3\text{-}4}$ = 3.4 Hz, $H_{F3}$), 3.79 (dd, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $J_{2\text{-}3}$ = 10.3 Hz, $H_{F2}$), 3.99 (td, $J_{2\text{-}3a}$ = 4.0 Hz, $J_{2\text{-}3b}$ = 11.1 Hz, $J_{1\text{-}2}$ = 11.1 Hz, 1H, $H_{Cy2}$), 5.35 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 6.35 (t, 1H, $J^4$ = 2.2 Hz, $H_{Ar4}$), 6.67 (d, 2H, $J^4$ = 2.2 Hz, $H_{Ar2}$ and $H_{Ar6}$ ). **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.6 ($C_{F6}$), 25.9 ($C_{Cy5}$), 26.3 ($C_{Cy4}$), 30.4 ($C_{Cy6}$), 33.5 ($C_{Cy3}$), 51.1 ($C_{Cy1}$), 52.3 ($C_{Cy2}$), 68.0 ($C_{F2}$), 68.6 ($C_{F5}$), 71.6 ($C_{F3}$), 73.3 ($C_{F4}$), 78.7 ($C_{F1}$), 106.7 (Ar), 106.9 (Ar), 137.8 (Cquart.) 160.0 (Cquart.), 169.3 (CONH), 178.0 (CONH). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{20}H_{28}N_2O_8$: 447.17379 $[M + Na]^+$; found: 447.17369. $[\alpha]_D$ = -76.2 (c = 0.20, MeOH).

**(1*S*,2*S*)-2-Aminocyclohexanecarboxylic acid series: Ligands 12**

**N-[(1*R*,2*R*)-2-(3-Hydroxybenzamido)cyclohexanecarboxyl]-α-L-fucopyranosylamine (12b)**

The crude hydrogenation product of **30** (see Supplementary Information – SI-Scheme 3) was coupled with 3-hydroxybenzoic acid using the HBTU general procedure (see Supplementary Information) and the product was purified by flash chromatography (6:4 ethyl acetate:petroleum ether $R_f$ = 0.17). Yield: 28 mg (53 %) Zemplen deprotection and flash chromatography (85:15 chloroform:methanol, $R_f$ = 0.10) afforded **12b**. Yield: 7 mg (50 %). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 0.93 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{F6}$), 1.24 – 1.63 (m, 4H, $H_{Cy3ax}$, $H_{Cy6ax}$ and $H_{Cy4}$ and/or $H_{Cy5}$ ), 1.76 – 1.78 (m, 2H, $H_{Cy4}$ and/or $H_{Cy5}$), 1.93 – 2.02 (m, 2H, $H_{Cy3eq}$ and $H_{Cy6eq}$), 2.47 (td, 1H, $J_{1\text{-}6b}$ = 3.6 Hz, $J_{1\text{-}2}$ = 11.9 Hz, $J_{1\text{-}6b}$ = 11.9 Hz, $H_{Cy1}$), 3.54 – 3.55

(m, 1H, $H_{F4}$), 3.60 (q, $J_{5-6}$ = 6.4 Hz, 1H, $H_{F5}$), 3.78 (dd, 1H, $J_{3-4}$ = 3.4 Hz, $J_{2-3}$ = 10.4 Hz, $H_{F3}$), 3.89 (dd, 1H, $J_{1-2}$ = 5.6 Hz, $J_{2-3}$ = 10.4 Hz, $H_{F2}$), 4.17 (td, $J_{2-3}$ = 4.0 Hz, $J_{1-2}$ = 11.3 Hz, $J_{2-3}$ = 11.3 Hz, 1H, $H_{Cy2}$), 5.41 (d, 1H, $J_{1-2}$ = 5.7 Hz, $H_{F1}$), 6.85 – 6.90 (m, 1H, $H_{Ar2}$), 7.12 – 7.20 (m, 3H, $H_{Ar4,5,6}$). **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.84 ($C_{F6}$), 26.22 ($C_{Cy4}$ or $C_{Cy5}$), 26.42 ($C_{Cy4}$ or $C_{Cy5}$), 301.70 ($C_{Cy6}$), 33.78 ($C_{Cy3}$), 51.00 ($C_{Cy2}$), 52.25 ($C_{Cy1}$), 67.99 ($C_{F2}$), 68.44 ($C_{F5}$), 71.47 ($C_{F3}$), 73.32 ($C_{F4}$), 78.63 ($C_{F1}$), 115.55 (Ar-CH), 119.52 (Ar-CH), 119.61 (Ar), 130.66 (Ar), 137.31 (Cquart.,), 158.88 (Cquart.), 169.95 (Cquart.), 178.07 (Cquart.). HRMS (FT-ICR, ESI): *m/z* calcd for $C_{20}H_{28}N_2O_7$: 431.17887 $[M + Na]^+$; found: 431.17963. $[\alpha]_D$ = -65.0 (c = 0.20, MeOH) for a sample contained 8% of the beta anomer.

## β-Alanine series: Ligands 9

### *N*-(*N'*-Acetyl-β-alanyl)-α-L-fucopyranosylamine (9a)

The crude hydrogenation product of **27** (see Supplementary Information – SI-Scheme 3) was used in the general acetylation method (see Supplementary Information). The product was purified by flash chromatography on silica gel (AcOEt, $R_f$ = 0.17, quant.). Zemplen deprotection and flash chromatography (4:1 $CHCl_3$:MeOH, $R_f$ = 0.2) afforded **9a** (quant.). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.90 (d, 3H, $J_{6-5}$ = 5.6 Hz, $H_{F6}$), 1.93 (s, 3H, NHC(O)$CH_3$), 2.50 (t, 2H, *J* = 6.0 Hz, C(O)$CH_2CH_2$NHAc), 3.43 (t, 2H, *J* = 6.0 Hz, C(O)$CH_2CH_2$NHAc), 3.65 (br s, 1H, $H_{F4}$), 3.73–3.80 (br s, 2H, $H_{F5}$, $H_{F3}$), 3.93–4.10 (m, 1H, $H_{F2}$), 5.55 (dd, 1H, $J_{1-2}$ = 5.6 Hz, $J_{1\text{-}NH}$ = 7.6 Hz, $H_{F1}$), 8.05 (br s, 1H, C(O)$CH_2CH_2$NHAc), 8.35 (d, 1H, $J_{1\text{-}NH}$ = 7.6 Hz, α-NHC(O)); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 15.5 ($C_{F6}$), 21.1 (NHC(O)$CH_3$), 35.1 (C(O)$CH_2CH_2$NHAc), 35.4 (C(O)$CH_2CH_2$NHAc), 66.5 ($C_{F2}$), 77.2, 70.0, ($C_{F3,}$ $C_{F5}$), 72.3 ($C_{F4}$), 77.1 ($C_{F1}$), 172.0 (C=O), 173.6 (C=O); $R_f$ = 0.2 (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z* = 277.3 $[(M+H)^+$, 100%].

### *N*-[*N'*-(3-Hydroxybenzoyl)-β-alanyl]-α-L-fucopyranosylamine (9b)

The crude hydrogenation product of **27** (see Supplementary Information – SI-Scheme 3) was coupled with 3-acetoxybenzoic acid using the HBTU general method (see Supplementary Information). The product was

purified by flash chromatography on silica gel (1:4 hexane: AcOEt, $R_f$ = 0.25, 50% yield). Zemplen deprotection and flash chromatography (9:1 $CHCl_3$: MeOH, $R_f$ = 0.17) afforded **9b** (quant.).

**$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.12 (d, 3H, $J_{6-5}$ = 6.5 Hz, $H_{F6}$), 2.70-2.54 (m, 2H, NHOC-$CH_2$-), 3.68-3.59 (m, 3H, $H_{F5}$ + -$CH_2$NHCO-), 3.76-3.70 (m, 1H, $H_{F4}$), 3.76 (dd, 1H, $J_{2-3}$ = 10.4 Hz, $J_{3-4}$ = 3.5 Hz, $H_{F3}$), 3.96 (dd, 1H, $J_{2-3}$ =10.4 Hz, $J_{1-2}$ = 5.6 Hz, $H_{F2}$), 5.57 (d, 1H, $J_{1-2}$ = 5.6 Hz, $H_{F1}$), 6.96-6.91 (m, 1H, $H_{Ar}$), 7.27-7.20 (m, 3H, 3x$H_{Ar}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.9 ($C_{F6}$), 36.7 (NHOC-$\underline{CH_2}$-), 37.4 (-$\underline{CH_2}$NHCO-), 68.0 ($C_{F2}$), 68.6, 71.5 ($C_{3F}$, $C_{4F}$), 73.3 ($C_{F5}$), 78.3 ($C_{F1}$), 115.3 ($CH_{Ar}$), 119.1 ($CH_{Ar}$), 120.8 ($CH_{Ar}$), 130.6 ($CH_{Ar}$), 137.0 ($C_{Ar}$), 159.0 ($C_{Ar}$), 170.4 (C=O), 175.3 (C=O); $R_f$ = 0.24 (4:1 $CHCl_3$:MeOH)-ESI-MS: *m/z* = 377.5 $[(M+Na)^+$, 100%]. HR-MS (ESI): calculated for $C_{16}H_{21}N_2O_7$ $[M-1]^-$: 353.13542; found $[M-1]^-$: 353.13565.

### *N*-[*N'*-(3,5-Dihydroxybenzoyl)-β-alanyl]-α-L-fucopyranosylamine (9c)

The crude hydrogenation product of **27** (see Supplementary Information – SI-Scheme 3) was coupled with 3,5-diacetoxybenzoic acid using the HBTU general procedure (see Supplementary Information). The product was purified by flash chromatography on silica gel (1:4 hexane: AcOEt, $R_f$ = 0.3, 55% yield). Zemplen deprotection and flash chromatography (85:15 $CHCl_3$: MeOH, $R_f$ = 0.23) afforded **9c** (quant.).

**$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.12 (d, 3H, $J_{6-5}$ = 6.5 Hz, $H_{F6}$), 2.69-2.53 (m, 2H, -NHOC*$CH_2$*-), 3.65-3.59 (m, 3H, $H_{F4}$ + -*$CH_2$*NHCO-), 3.80-3.70 (m, 2H, $H_{F3}$, $H_{F5}$), 3.96 (dd, 1H, $J_{2-3}$ =10.3 Hz, $J_{1-2}$ = 5.6 Hz, $H_{F2}$), 5.56 (d, 1H, $J_{1-2}$ =5.6 Hz, $H_{1F}$), 6.40 (dt, 1H, *J* = 2.1 & 0.9 Hz, $H_{Ar}$), 6.71-6.68 (m, 2H, 2$H_{Ar}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.9 ($C_{F6}$), 36.6 (NHOC-$\underline{CH_2}$-), 37.4 (-$\underline{CH_2}$NHCO-), 68.0 ($C_{F2}$), 68.7, 71.5 ($C_{F3}$, $C_{F5}$), 73.3 ($C_{F4}$), 78.4 ($C_{F1}$), 116.8 (3×$CH_{Ar}$), 137.7 ($C_{Ar}$), 159.9 (2×$C_{Ar}$), 170.6 (C=O), 175.3 (C=O); $R_f$ = 0.3 (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z* = 393.5 $[(M+Na)^+$, 100%]. HR-MS (ESI): calculated for $C_{16}H_{21}N_2O_8$ $[M-1]^-$: 369.13034; found $[M-1]^-$: 369.13089.

**N-[N'-(D-Acetyl-triptophanyl)-β-alanyl]- α -L-fucopyranosylamine (9g)**

The crude hydrogenation product of **27** (see Supplementary Information – SI-Scheme 3) was coupled with D-*N*-acetyl-triptophane using the HBTU general procedure (see Supplementary Information). The product was purified by flash chromatography on silica gel (AcOEt, $R_f$ = 0.67, 47% yield). Zemplen deprotection and flash chromatography (85:15 $CHCl_3$: MeOH, $R_f$ = 0.15) afforded **9g** (quant.). **$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.16 (d, 3H, $J_{6\text{-}5}$ = 6.4 Hz, $H_{F6}$), 1.92 (s, 3H, $CH_3CONH$-), 2.45-2.27 (m, 2H, NHOC-$CH_2$-), 3.09 (dd, 1H, *J* = 14.5 & 7.6 Hz, $CH_{2Trp}$), 3.22 (dd, 1H, *J* = 14.4 & 6.5 Hz, $CH_{2Trp}$), 3.44-3.30 (m, 1H, -$CH_2$NHCO-), 3.65-3.62 (m, 1H, $H_{F4}$), 3.81-3.73 (m, 2H, $H_{F3}$, $H_{F5}$), 3.95 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.56 (t, 1H, *J* = 7.1 Hz, $CH_{Trp}$), 5.54 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 7.01 (t, 1H, *J* = 7.4 Hz, $H_{ArTrp}$), 7.11-7.05 (m, 2H, $2H_{ArPhe}$), 7.32 (d, 1H, *J* = 8.0 Hz, $H_{ArTrp}$), 7.58 (d, 1H, *J* = 7.8 Hz, $H_{ArTrp}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 16.9 ($C_{F6}$), 22.6 ($CH_3CONH$-), 29.0 ($CH_{2Trp}$), 36.5 (NHOC-$\underline{CH_2}$-), 36.8 (-$\underline{CH_2}$NHCO-), 56.1 ($CH_{Trp}$), 68.1 ($C_{F2}$), 68.7, 71.5 ($C_{F3}$, $C_{F5}$), 73.3 ($C_{F4}$), 78.3 ($C_{F1}$), 101.4 ($C_{Ar}$), 111.0 (2x$C_{Ar}$), 112.3 ($CH_{Ar}$), 119.4 ($CH_{Ar}$), 119.8 ($CH_{Ar}$), 122.5 ($CH_{Ar}$), 124.5 ($CH_{Ar}$), 128.9 ($C_{Ar}$), 138.1 ($C_{Ar}$), 173.2 (C=O), 174.3 (C=O), 175.1 (C=O); $R_f$ = 0.25 ($CHCl_3$:MeOH 4:1). ESI-MS: *m/z* = 485.6 [$(M+Na)^+$, 100%]. HR-MS (ESI): calculated for $C_{22}H_{30}N_4O_7Na$ $[M+Na]^+$: 485.20067; found $[M+Na]^+$: 485.20057.

**N-[N'-(D-N-Acetyl-phenylalanyl)-β-alanyl]- α-L-fucopyranosylamine (9h)**

The crude hydrogenation product of **27** (see Supplementary Information – SI-Scheme 3) was coupled with D-*N*-acetyl-phenylalanine using the HBTU general procedure (see Supplementary Information). The product was purified by flash chromatography on silica gel (100:2, AcOEt:MeOH, $R_f$ = 0.12, 60% yield). Zemplen deprotection and flash chromatography (9:1 $CHCl_3$: MeOH, $R_f$ = 0.13) afforded **9h** (quant.).

**$^1$H-NMR (400 MHz, $CD_3OD$) :** δ (ppm) = 1.17 (d, 3H, $J_{5\text{-}6}$ = 6.5 Hz, $H_{6F}$), 1.90 (s, 3H, $CH_3CONH$-), 2.51-2.35 (m, 2H, NHOC-*$CH_2$*-), 2.86 (dd, 1H, *J* = 13.8 & 8.9 Hz, $CH_{2Phe}$), 3.08 (dd, 1H, *J* = 13.8 & 6.2 Hz, $CH_{2Phe}$), 3.50-3.30 (m, 2H, -*$CH_2$*NHCO-), 3.66-3-63 (m, 1H, $H_{F4}$), 3.81-3.74 (m, 2H, $H_{F3}$, $H_{F5}$), 3.96 (dd, 1H, $J_{2\text{-}3}$ = 10.3 Hz, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F2}$), 4.42 (dd, 1H, *J* = 8.8 & 6.3 Hz, $CH_{Phe}$), 5.56 (d, 1H, $J_{1\text{-}2}$ = 5.6 Hz, $H_{F1}$), 7.31-7.17 (m, 5H, $5H_{ArPhe}$); **$^{13}$C-NMR (100 MHz, $CD_3OD$) :** δ (ppm) = 17.6 ($C_{F6}$), 22.4 ($CH_3CONH$-), 36.7 (NHOC-$\underline{CH_2}$-), 36.8 (-$\underline{CH_2}$NHCO-), 39.0 ($CH_{2Phe}$), 58.5 ($CH_{Phe}$), 68.1 ($C_{F2}$), 68.7, 71.5 ($C_{F3}$, $C_{F5}$), 73.3 ($C_{F4}$), 78.3 ($C_{F1}$), 127.8 ($CH_{Ar}$), 129.5 ($CH_{Ar}$), 130.3 ($CH_{Ar}$), 138.5 ($C_{Ar}$), 173.2 (C=O), 173.8 (C=O), 175.1 (C=O); $R_f$ = 0.20 (4:1 $CHCl_3$:MeOH). ESI-MS: *m/z* = 446.5 [$(M+Na)^+$, 100%]. HR-MS (ESI): calculated for $C_{20}H_{29}N_3O_7Na$ $[M+Na]^+$: 446.18977; found $[M+Na]^+$: 446.19005.

---

[1] G. Timpano, G. Tabarani, M. Anderluh, D. Invernizzi, F. Vasile, D. Potenza, P. M. Nieto, J. Rojo, F. Fieschi, A. Bernardi, *ChemBioChem* **2008**, *9*, 1921-1930

[2] F. Damkaci, P. DeShong, *J. Am. Chem. Soc.* **2003**, *125*, 4408-4409

[3] M. Andreini , M. Anderluh , A. Audfray , A. Bernardi , A. Imberty, *Carb. Res.* **2010**, *345*, 1400-1407

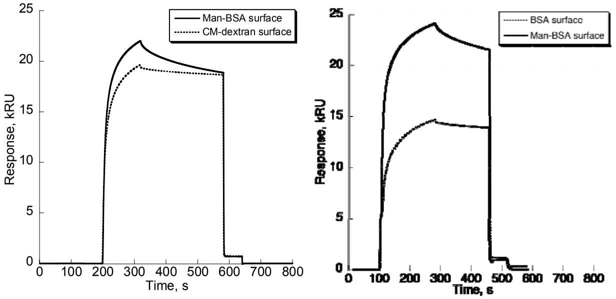
[4] a) C. Bolm, I. Schiffers, C. L. Dinter, A. Gerlach, *J. Org. Chem.* **2000**, *65*, 6984-6991; b) A. Bernardi, D. Arosio, D. Dellavecchia, F. Micheli, *Tetrahedron: Asymm.* **1999**, *10*, 3403 – 3407

[5] M. Adamczyk, J. R. Fishpaugh, *Tetrahedron Lett.* **1996**, *37*, 4305–4308

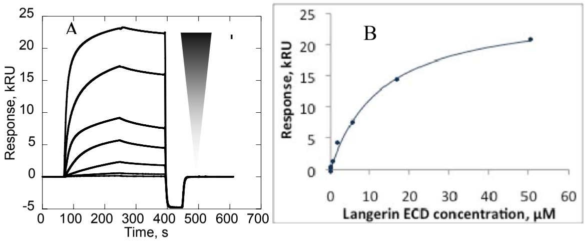
## 2. Supplementary Information – Surface Plasmon Resonance Experiments

### *1. Properties of Langerin binding to BSA-Man SPR surfaces and set up of the inhibition test.*

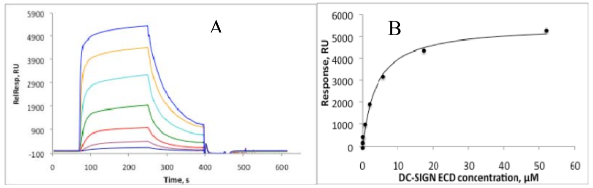
As for DC-SIGN, the ability of Langerin to bind to surface functionalized with Man-BSA was tested. In the case of Langerin, strong binding to the reference surface was observed, due to the presence of the dextran polymer (SI-Fig.1). This is limited but not suppressed by using a reference surface functionalized with non-glycosylated BSA (SI-Fig.1). Therefore, the dextran/Man-BSA surface was considered as a combined ligand of Langerin ECD and thus binding responses were not reference surface corrected.



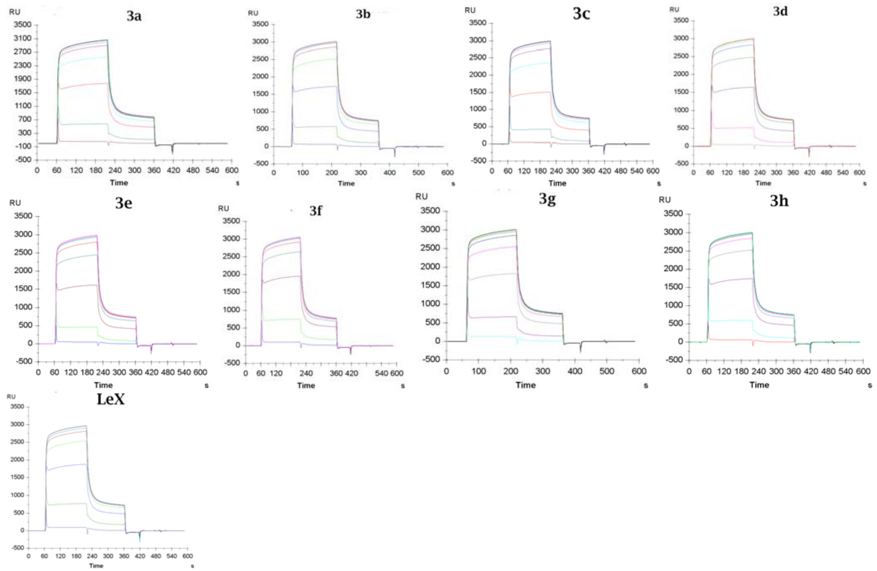
**SI-Figure 1.** Langerin interaction with reference surface. Data were obtained with 5 µl/min flow rate of running buffer (25 mM Tris-HCl pH 8, 150 mM NaCl, 4 mM $CaCl_2$, 0.005% surfactant P20). Left panel A), reference flow cell corresponds to carboxymethylated dextran surface. Right panel B), reference surface functionalized with unglycosylated BSA.



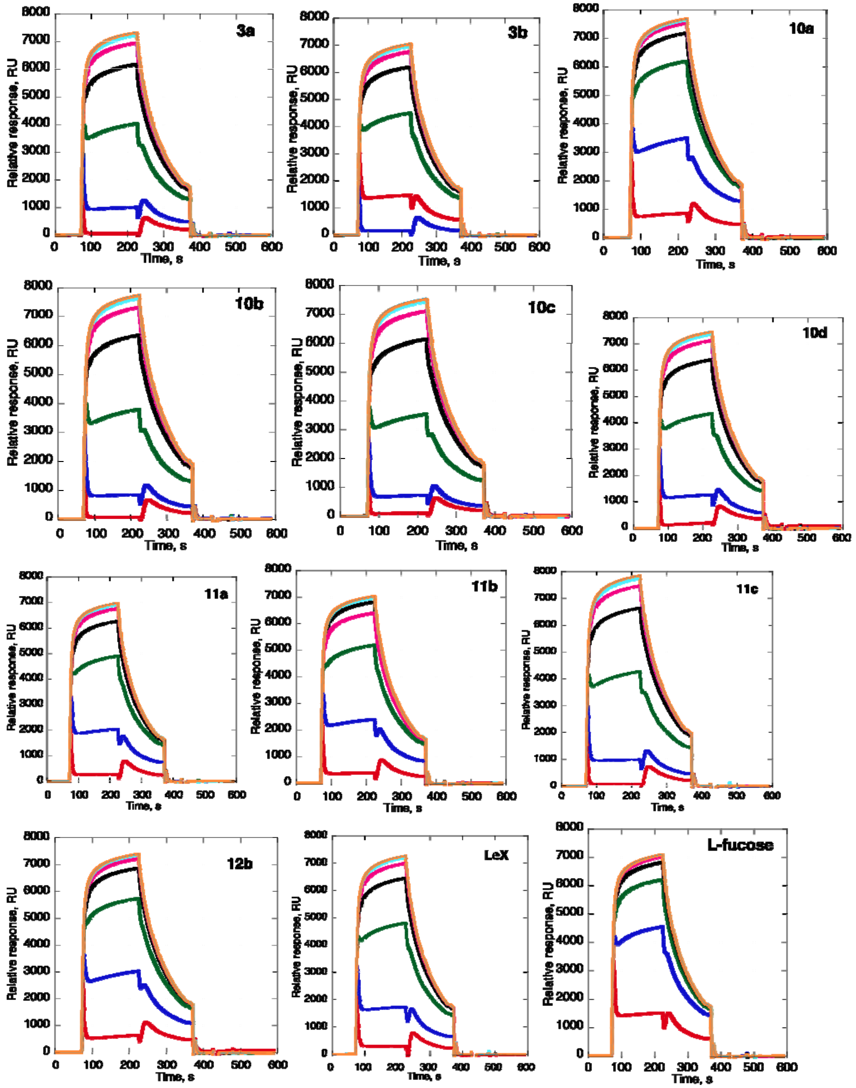
**SI-Figure 2:** Langerin titration on a Man-BSA surface. A) Langerin concentrations range from 50.35 µM to 0.07 µM by 3-fold dilutions. Regeneration is performed using 1-minute pulse injection of 50 mM EDTA pH 8, at 400 sec. B) Titration curve derived from sensorgrams A).



**SI-Figure 3:** DC-SIGN ECD titration on a Man-BSA surface. A) DC-SIGN ECD concentrations range from 50.35 µM to 0.07 µM by 3-fold dilutions. Regeneration is performed using 1 minute pulse injection of 50 mM EDTA pH 8, at 400 sec. B) Titration curve derived from sensorgrams A).

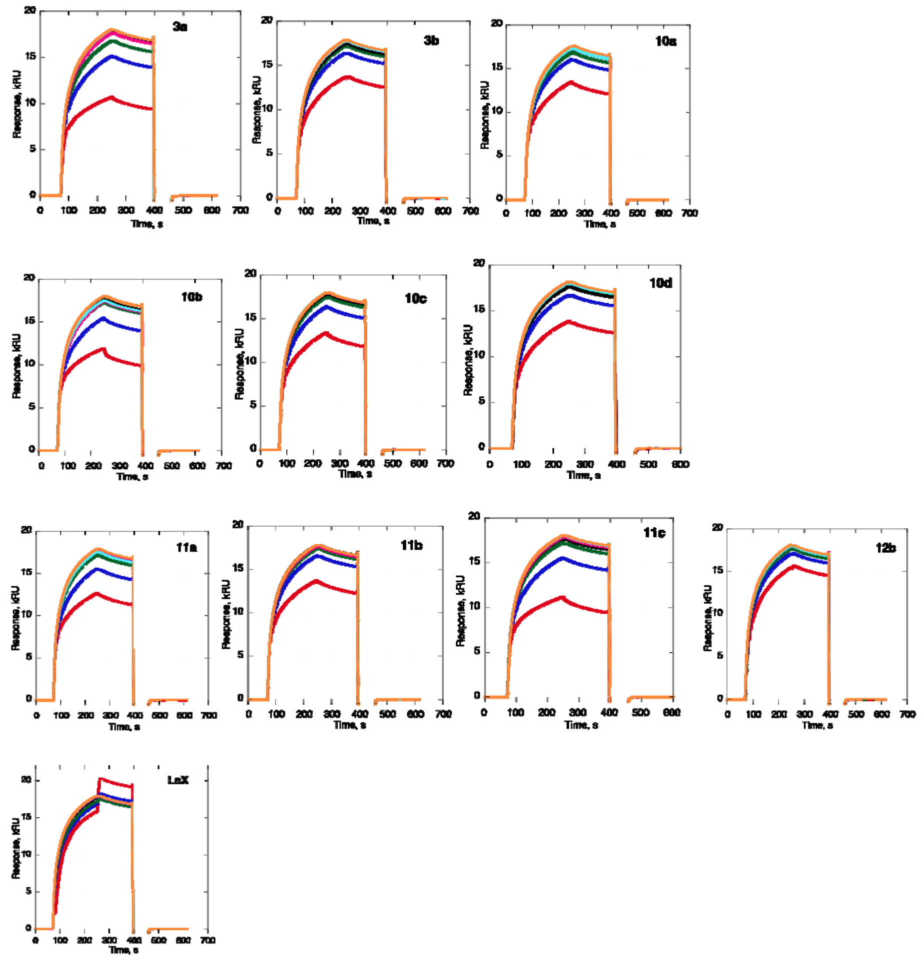


**SI-Figure 4:** Reference surface corrected overlay sensorgrams representing inhibition of DC-SIGN ECD binding to Man-BSA surface by compounds **3** and LeX.



**SI-Figure 5:** Reference surface corrected overlay sensorgrams representing inhibition of DC-SIGN ECD binding to Man-BSA surface by compounds **3**, **10**, **11**, **12**, L-fucose, and LeX.

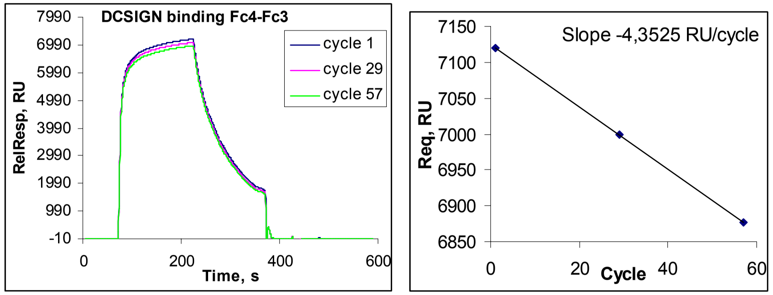
---



**SI-Figure 6:** Overlay raw sensorgrams representing inhibition of Langerin ECD binding to Man-BSA/dextran surface by compounds **3**, **10**, **11**, **12**, and LeX.

## *2. Evaluation of BSA-Man SPR surface stability.*

The stability of the surface was evaluated in two aspects: affinity and binding capacity. SI- Figure 7a shows the evolution of DC-SIGN binding level onto one of the chips used, plotted as a function of the number of cycles performed onto the chip. The binding level is very reproducible with a deviation of about 4 RU per cycle (Figure 7b) to compare to a binding level close to 7000 RU (so a loss of around 0.06 %).



**a.** **b.**

**SI- Figure 7. a.** DC-SIGN binding to BSA-Man chip, as a function of different cycles performed on the chip, b. RU variation as a function of different cycles

Finally, a complete titration of Man-BSA surface with DC-SIGN was performed, to determine apparent Kd, at the beginning and end of each SPR campaign (new surface each time). Very similar values were obtained each time (4.2 µM compared to 3.3 µM; or 6.1 µM compared to 5.5 µM).

# 14. Supplementary information to paper n°3

Supplementary Information

# Pseudosaccharide functionalized dendrimers as potent inhibitors of DC-SIGN dependent Ebola pseudotyped viral infection

Joanna Luczkowiak,[c] Sara Sattin,[b] Ieva Sutkevičiūtė,[d,f] José Juan Reina,[b] Macarena Sánchez-Navarro,[a] Michel Thépaut,[d,e] Lorena Martínez-Prats,[c] Anna Daghetti,[b] Franck Fieschi,[d,e,f,*] Rafael Delgado[c,*] Anna Bernardi,[b,*] Javier Rojo,[a,*]

[a] *Grupo Carbohidratos, Instituto de Investigaciones Químicas, CSIC – Universidad de Sevilla, Américo Vespucio 49, 41092 Seville, Spain. Tel: 954489568; Fax: 954460565*

[b] *Università degli Studi di Milano, Dipartimento di Chimica Organica e Industriale, Milano, Italy*

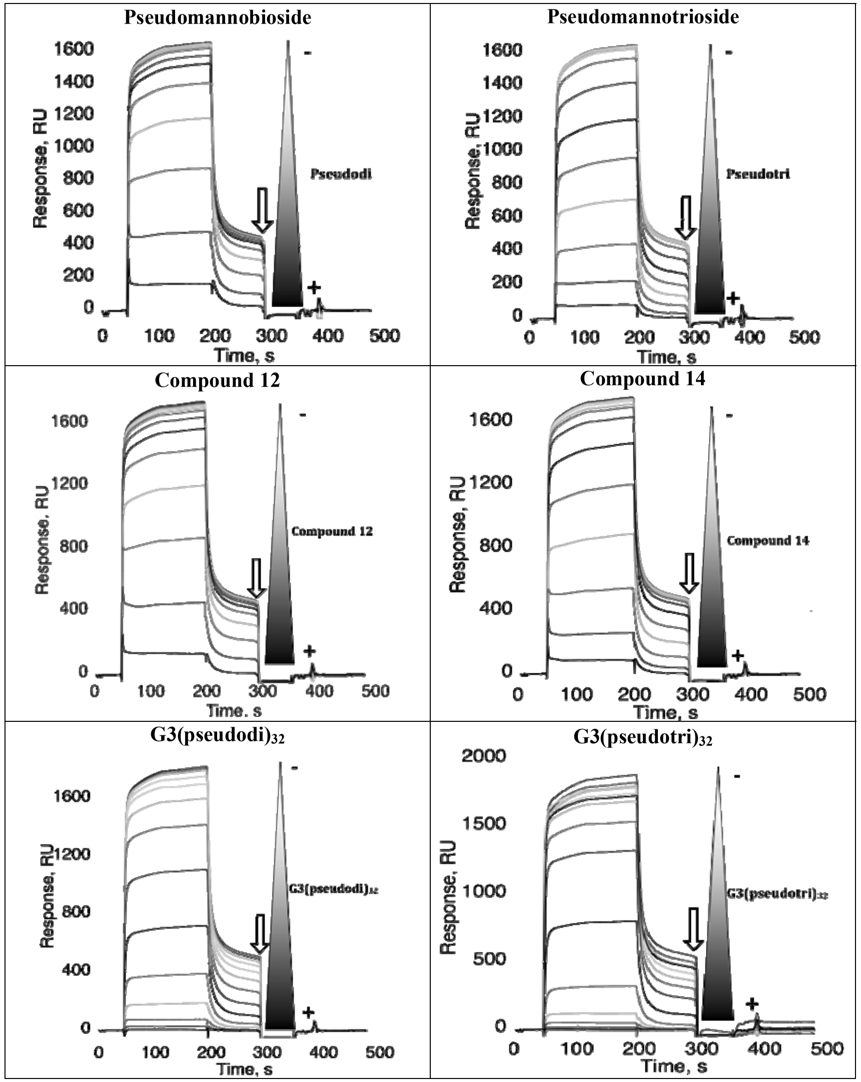
[c] *Laboratorio de Microbiología Molecular- Hospital Universitario 12 de Octubre, Madrid, Spain*

[d] *Institut de Biologie Structurale, CNRS, UMR 5075, 41 rue Jules Horowitz 38027 Grenoble France*

[e] *CEA, DSV, 38027 Grenoble France*

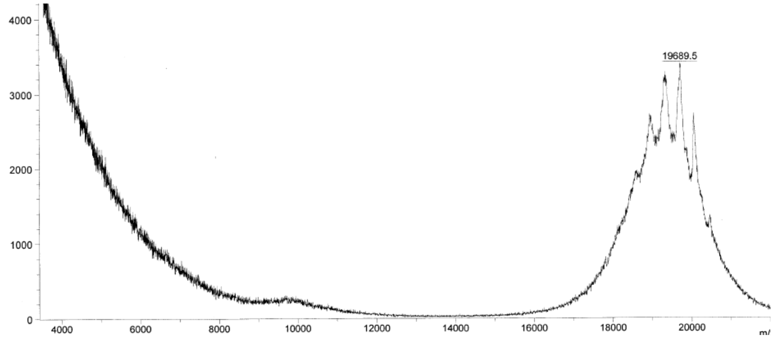
[f] *Université Joseph Fourier, 38000 Grenoble, France*

1. SPR sensorgrams for the inhibition curves reported in Figure 3. The arrows denote EDTA injections for BSA-Mannotriose surface regeneration between each cycle.

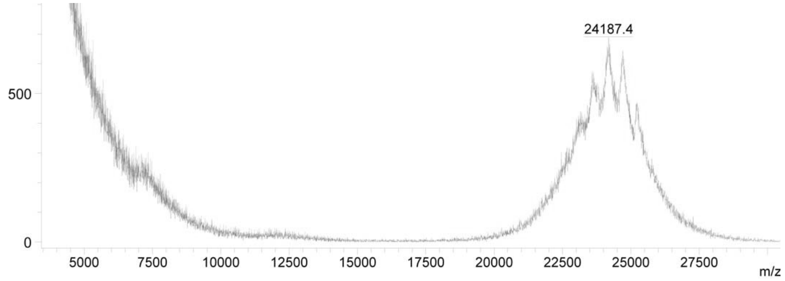


**Pseudomannobioside** | **Pseudomannotrioside** | **Compound 12** | **Compound 14** | **G3(pseudodi)$_{32}$** | **G3(pseudotri)$_{32}$**

2. Mass spectra for multivalent compounds with glycomimetics **13** and **15**.



**MALDI mass spectrum of 3G(pseudodi)$_{32}$. Matrix: SA.**



**MALDI mass spectrum of 3G(pseudotri)$_{32}$. Matrix: SA.**

**Equation 1 (**determination of relative IC$_{50}$ values)

$$y = bot + \frac{top - bot}{1 + \left(\frac{x}{IC_{50}}\right)^{slope}}$$

where *y* is the percent activity, *x* is the corresponding concentration, *bot* and *top* are the lowest and the highest values of percent activity, respectively.

# 15. Supplementary information to paper n°4

## Supporting informations

# Structure of a glycomimetic ligand in the Carbohydrate Recognition Domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design.

Michel Thépaut[1,2,3#], Cinzia Guzzi[5#], Ieva Sutkeviciute[1,2,3], Sara Sattin[4], Renato Ribeiro-Viana[5], Norbert Varga[4], Eric Chabrol[1,2,3], Javier Rojo[5], Anna Bernardi[4], Jesus Angulo[5], Pedro M. Nieto[5], and Franck Fieschi[1,2,6*]

*[1]Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, Grenoble, F-38027, France; [2]CNRS, UMR 5075, Grenoble, F-38000, France; [3]CEA, DSV, Grenoble, F-38000, France; [4]Universita' di Milano, Dipartimento di Chimica Organica e Industriale, via Venezian 21, 20133 Milano - Italy; [5]Glycosystems Laboratory, Instituto de Investigaciones Químicas (IIQ), CSIC - Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla-Spain; [6]Institut Universitaire de France, 103 boulevard Saint-Michel 75005 Paris-France.*

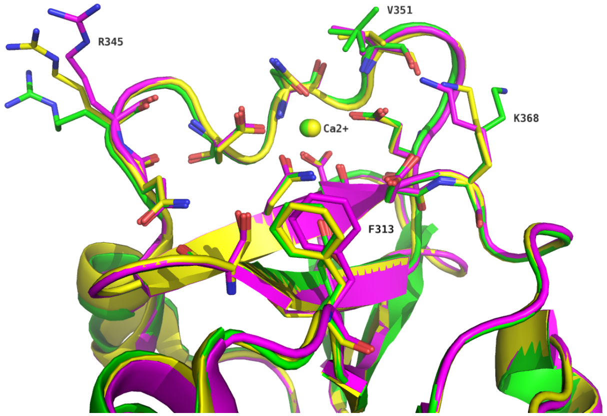


Figure S1. Structural alignment of DC-SIGN CRD from 2XR5, 2IT6 and 1SL5 structure corresponding respectively to the pdb code of DC-SIGN complexed with pseudo-1,2-mannobioside (this study, green), Manα1,2Man (magenta), and Lacto-N-fucopentaose III (yellow). Ligands are not represented for clarity (see Figure 6a to see their binding mode).

Comparison of the global DC-SIGN CRD backbones in each complex reveals a global identical structure. When examining more precisely the side chain of residues (see above) involved in the binding site ($Ca^{2+}$ and/or ligand binding), again, most of them are perfectly superimposed. Only 4 residues present some slight differences and are labeled in the figure above: F313, R345, V351 and K368.

For F313, the difference is with 2IT6 structure (complex with Manα1-2Man) where the F313 side chain is doing a slight rotation caused by hydrophobic interaction with mannose ring. No movement of this residue with pseudo-1,2-mannoside and Lewis X which are not interacting with F313.

For R345, the side chain is in different orientation in all three structures but is not involved in ligand binding. It is oriented towards solved and different orientation are more related to the various crystal structures packing (not exactly same space groups). Only, the CO of the backbone of R345, which is approximately in the same orientation in all structures, can be indirectly involved in interaction with ligands through a bridging water molecule (not shown here).

For K368, differences are more significant and implicate movements of the side-chain. Depending of the case, the lateral chains are facing the ligand or solvent. In 2IT6 terminal amine of side chain is involved in hydrogen bond with hydroxyl function of C6 of the non-reducing mannose of minor orientation. In 1SL5 an interaction with the hydroxyl group on C2 of fucose might also be possible. In 2XR5 structure there is no group from the ligand available for interaction and K368 side chain is oriented to solvent.

For V351, the differences are small and limited to various orientations of side chain around the CA-CB bond in order to fit to the corresponding ligand. In 2XR5 structure (this study) we identified two alternate conformations, one that is the same conformation as in 2IT6 structure and a second which is a 90° rotation. These two conformations establish contacts with hexane ring of pseudomannobioside. In 1SL5 structural orientation of the side chain is in between the two alternate conformations we identified in 2XR5 structure and establish contacts with hydroxyl group and C2 of the fucose ring of Lewis X.

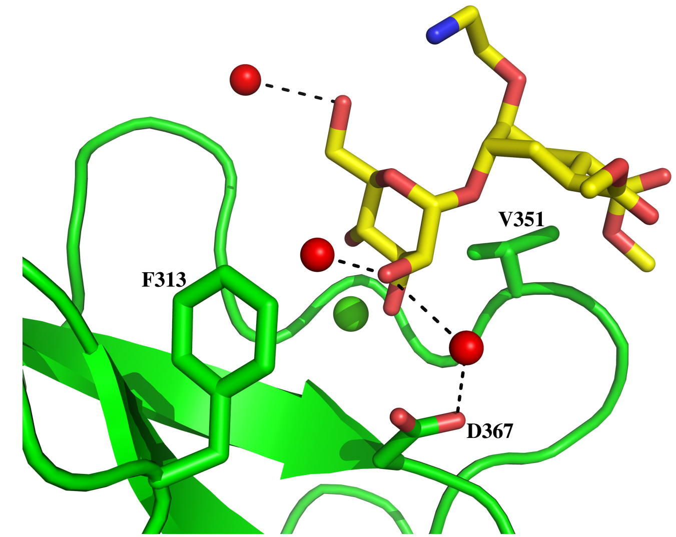


Figure S2. Crystallographic water molecule in direct contact with pseudo-1,2-mannobioside. Only one of these three mediates an interaction with a protein residue, D367. Water are represented as red sphere, $Ca^{2+}$ as green sphere. Hydrogen bonds involving water molecules are represented as dotted line.

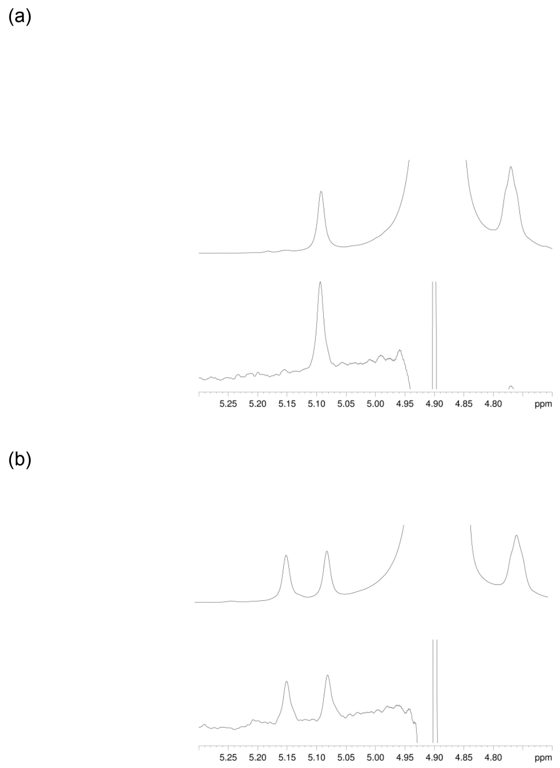


Figure S3. Zoom of anomeric region of STD spectra registered at 25 ºC (500 MHz) and saturation time of 2.0 s. (off-resonance frequency 40 ppm, on-resonance frequency 0 ppm). (a) 1H NMR reference spectrum (above) and STD spectrum (bottom) of a sample containing 1.7 mM of **1b** and 19 μM of DC-SIGN ECD. (b) 1H NMR reference spectrum (above) and STD spectrum (bottom) of the same sample after addition of 1.7 mM of **1a**. The decrease in STD around 50 % indicates that the two ligands have the same affinity and then the triazol moiety does not affect binding to DC-SIGN.

Table S1. Chemical shifts of **1b** (500 MHz,D2O, 25 ºC, HDO residual at 4.7 ppm)

| Protons | δ $^1$H (ppm) |
|---|---|
| H-1 M | 4.89 |
| H-2 M | 3.88 |
| H-3 M | 3.72 |
| H-4 M | 3.52 |
| H-5 M | 3.51 |
| H-6a M | 3.78 |
| H-6b M | 3.65 |
| Me | 3.62 |
| H-1$_{ax}$ C | 3.65 |
| H-2$_{eq}$ C | 3.79 |
| H-3$_{ax}$ C | 1.38 |
| H-3$_{eq}$ C | 1.95 |
| H-4$_{ax}$ C | 2.77 |
| H-5$_{ax}$ C | 2.33 |
| H-6$_{ax}$ C | 1.66 |
| H-6$_{eq}$ C | 1.92 |
| O-C$\underline{H}_2$-CH$_2$ | 3.86 |
| O-C$\underline{H}_2$-CH$_2$ | 3.92 |
| O-CH$_2$-C$\underline{H}_2$ | 4.56 |
| C$\underline{H}_2$-OH | 4.65 |
| H Triazol | 7.98 |

Table S2. Comparison of interproton distances of **1b** in bound state obtained from NMR data in solution, and from measurement on the crystal structure (pdb code 2xr5).

| Proton pair | Distances from TR-NOESY (Å) [a] | X-Ray (Å) |
|---|---|---|
| H-3$_{ax}$ C / H-3$_{eq}$ C | 1.7 | 1.8 |
| H-1 M / H-3$_{eq}$ C | 2.6 | 2.3 |
| H-1 M / H-4 C | 3.3 | 3.5 |
| H-2 M / H-6$_{ax}$ C | 4.2 | 4.3 |
| H-2 M /H-3$_{eq}$ C | 3.2 | 3.9 |
| H-2 M / H-4 C | 3.1 | 3.3 |
| H-3 M / H-6$_{ax}$ C | 3.4 | 3.4 |
| H-3 M / H-4 C | 4.4 | 4.0 |
| H-5 C / H-1 C | 3.5 | 3.7 |
| H-6$_{eq}$ C / H-1 C | 2.5 | 2.5 |
| H-6$_{ax}$ C / H-1 C | 2.4 | 2.4 |

[a] Distances derived using the isolated spin-pair approximation (ISPA) by comparing relative NOE intensities with that of the reference (H3ax-H3eq of cyclohexyl ring of **1b**)

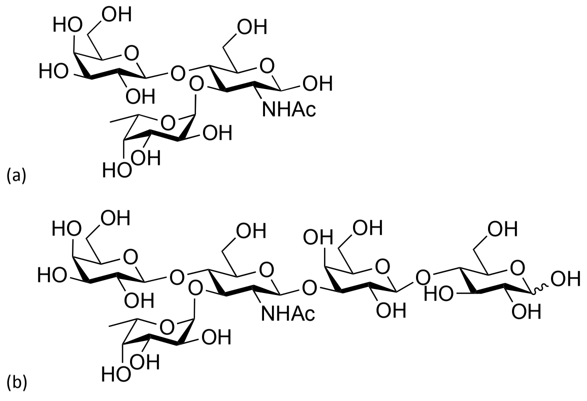


Figure S4: Structure of Lewis X (a) and Lacto-*N*-fucopentaose III (b)

Table S3. Contact distances between pseudo-1,2-mannobioside and V351 side chain in both conformation A and B.

| Contact | Distances (Å)<br>conf A [a] | conf B |
|---|---|---|
| Val 351 CG2 - 07B OAT | 3.8 | 3.0 |
| Val 351 CG2 - 07B CAR | 3.9 | 3.7 |
| Val 351 CG1 - 07B CAR | - | 4.2 |
| Val 351 CG1 - MAN C3 | - | 3.9 |
| Val 351 CG2 - MAN C3 | 3.9 | - |

Atom names are those from the pdb file 2XR5 of the structure described in this study.

MAN C3: C3 of the non reducing mannose pseudo-1,2-mannobioside
07B CAR: Carbon of the the cyclohexan ring pseudo-1,2-mannobioside
07B OAT : Oxygen of methylester bond substituent on the cyclohexane ring of pseudo-1,2-mannobioside

### Proposed formalism to quantify, in relative terms, the selectivity effect.

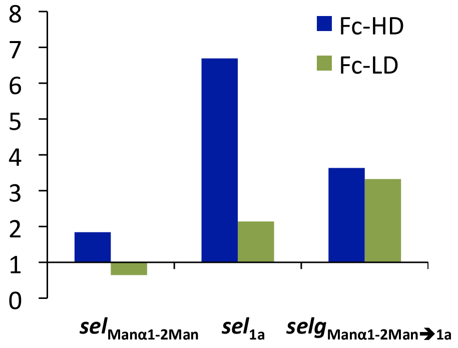


Figure S5. Histograms showing the comparison of Manα1-2Man and **1a** selectivities, and the selectivity gain obtained for both flow cells (Fc-HD – 5000 RU, Fc-LD – 1350 RU of immobilized BSA-Man).

In order to quantify the selectivity effect, a relative evaluation had to be employed. The absolute selectivity of a particular compound ($sel_{comp.}$) for DC-SIGN *vs* langerin can be defined as a ratio of the $IC_{50}$ values for langerin and DC-SIGN ($sel_{comp.} = IC_{50\ langerin} / IC_{50\ DC\text{-}SIGN}$), so that $sel_{comp.} > 1$ means a selectivity of the compound in favour of DC-SIGN. To exclude the effect of the surface density (Figure S5.), a selectivity gain (*selg*) achieved by switching from one compound to another can be considered, $selg_{comp.1 \rightarrow comp.2} = sel_{comp.2} / sel_{comp.1}$. This term provides a relative comparison of the improvement of selectivity between the two lectins when switching from one compound to another. Indeed, $selg_{comp.1 \rightarrow comp.2} > 1$ means a gain of DC-SIGN selectivity *vs*. langerin for compound 2 compared to compound 1 by a corresponding factor, while values <1 means a loss of selectivity for DC-SIGN. This approach of a relative comparison revealed that the selectivity gain is indeed surface density-independent (Figure S5). Interestingly, this formalism allows quantifying the improvement of selectivity that was qualitatively observed in Figure 7B. While Manα1-2Man has no clear selectivity between the two lectins, **1a** presents a specificity improvement towards DC-SIGN with a selectivity gain by a factor of about 3.5 (between 3.63 and 3.32) over Manα1-2Man.

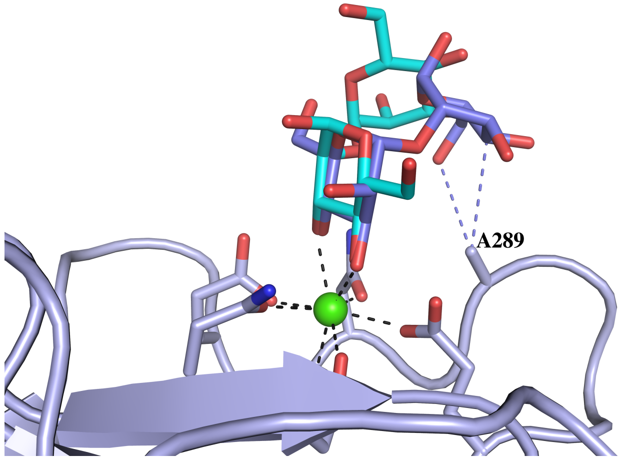


Figure S6. Manα1-2Man binding modes in langerin. Minor and major binding modes are represented in cyan and dark blue respectively. Van der Waals interactions with A289 are represented as dashed blue line.

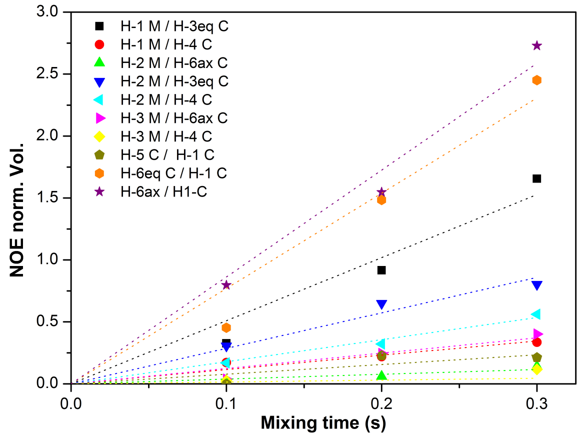


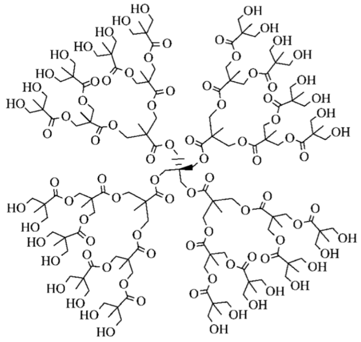
Figure S7. Transferred NOE build-up curves of selected cross peaks of **1b**. The growth of NOE normalized volumes (cross peak / diagonal peak) was well described by a linear behavior for mixing times below 0.3 s.

# 16. Supplementary information to paper n°5

# **SUPPLEMENTARY INFORMATION TO**: Lectin clustering by a glycomimetic without any multivalent presentation: a case study in DC-SIGN antagonist development

Ieva Sutkevičiūtė[1,2,3], Michel Thépaut[1,2,3], Sara Sattin[4], Angela Berzi[5], John McGeagh[6], Aline Le Roy[1,2,3], Macarena Sanchez Navaro[7], Javier Rojo[7], Mario Clerici[8], Anna Bernardi[4], Christine Ebel[1,2,3], Franck Fieschi[1,2,3*]

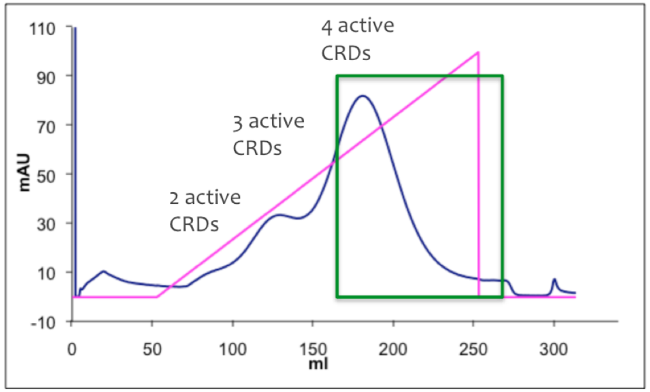
*[1]Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, Grenoble, F-38027, France; [2]CNRS, UMR 5075, Grenoble, F-38000, France; [3]CEA, DSV, Grenoble, F-38000, France; [4]Universita' di Milano, Dipartimento di Chimica, via Golgi 19, 20133 Milano - Italy; [5]Chair of Immunology, Department of Clinical Sciences L. Sacco, University of Milan, Segrate; [6]Anterio Consult&Research GmbH, Augustaanlage 23, D-68165 Mannheim-Germany; [7]Glycosystems Laboratory, Instituto de Investigaciones Químicas (IIQ), CSIC - Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla-Spain; [8]Chair of Immunology, Department of Biomedical Sciences and Technologies, University of Milan, Segrate.*



**Supplemental figure S1.** The structure of the G3 Boltorn type dendrimer.

*DC-SIGN ECD extra-purification*

DC-SIGN ECD purified by affinity chromatography using mannan-agarose column, which allows to select only those tetramers that have at least two active CRDs, and subsequent gel-filtration (as described in Tabarani *et al.* [1], was re-injected to mannan-agarose column maintaining the same conditions as already described Tabarani *et al.* [1], except that elution was performed by mannose gradient (0-30 mM, 200 mL length, 1 mL/min flow rate), and at the end the flow of mannose containing buffer was switched to 1 mM EDTA to elute any residual protein that was attached to the column.



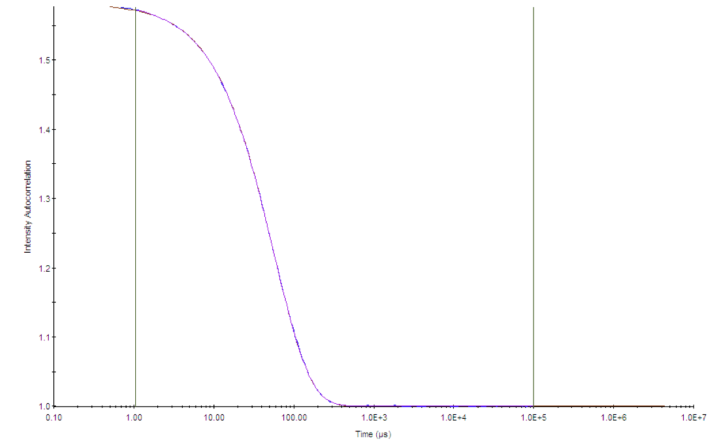
**Supplemental figure S2.** Extra-purification of DC-SIGN ECD. A chromatogram of DC-SIGN ECD injection to mannan-agarose and elution by mannose gradient (pink line). Absorption changes at $\lambda = 280$ nm are shown by a blue line. Green frame depicts collected fractions for further use in ITC experiments.

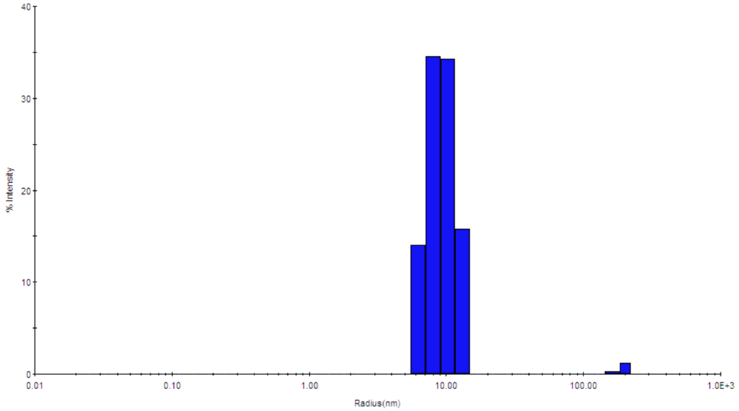
Since DC-SIGN ECD has been already purified, each tetramer must have at least 2 CRDs functional (CRD monomer has not enough of affinity to bind to the column, thus it is delayed and elution is not required (described in Thepaut *et al.* [2]). The eluted peak in Fig. S1 is composed of 3 sub-peaks, in which the earliest eluted have lowest affinity and must correspond to tetramers with 2 functional CRDs, subsequent sub-peaks should contain 3 and 4 functional CRDs possessing tetramers, respectively. Overall, the approximately estimated percentage of active CRDs per injected sample is 95%.

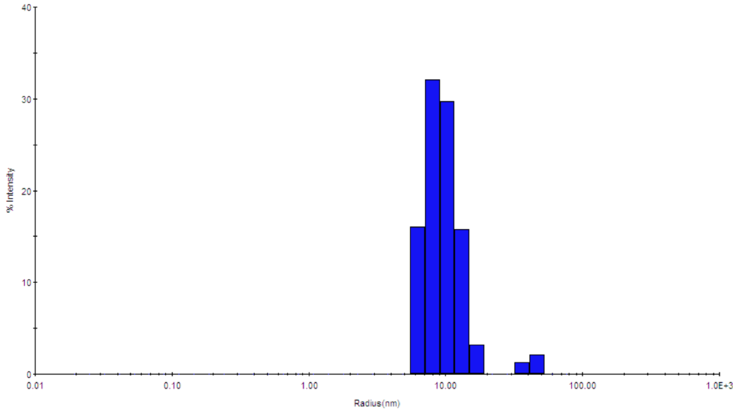
The collected 3rd sub-peak with 4 functional CRDs was further treated to eliminate the excess of mannose: EDTA was added to the concentrated sample of the protein (8 mg/mL) to final concentration of 16 mM in order to disrupt mannose/DC-SIGN complexes, centrifuged and subjected to gel-filtration (Superose12) to eliminate EDTA and mannose, and recharge the protein with $Ca^{2+}$. The eluted protein was concentrated and used in ITC experiment.

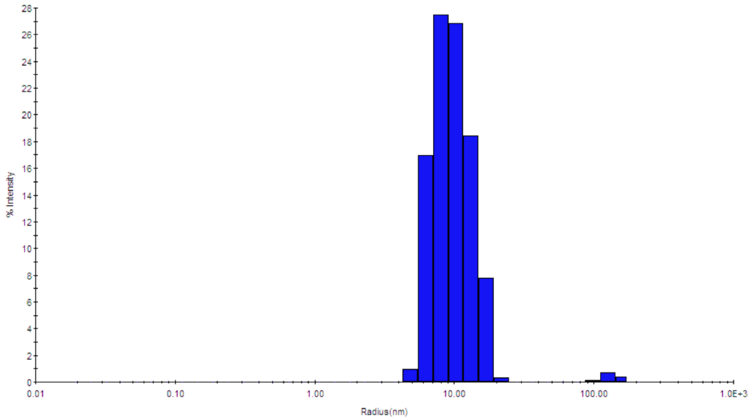
*An example of psTri/DC-SIGN ECD sample polydispersity in DLS measurements*

The correlation curves and the corresponding $R_\mathrm{H}$ distributions from three measurements of a sample containing 102 µM DC-SIGN ECD and 109 µM psTri are presented below.









## References

[1] Tabarani, G., Thépaut, M., Stroebel, D., Ebel, C., Vivès, C., Vachette, P., Durand, D., & Fieschi, F. (2009). *J Biol Chem,* **284** (32), 21229–40.

[2] Thepaut, M., Guzzi, C., Sutkeviciute, I., Sattin, S., Ribeiro-Viana, R., Varga, N., Chabrol, E., Rojo, J., Angulo, J., Bernardi, A., Nieto, P. M., & Fieschi, F. (2012). *submitted,* , 1–28.

# 17. Supplementary information to paper n°6

## Supplementary Information to :

# Selective targeting of DC-SIGN with mannose-based glycomimetics. Synthesis and interaction studies of bis-benzylamide derivatives of a pseudo-mannobioside.

Norbert Varga,[1,§] Ieva Sutkeviciute,[2,3,4,§] Cinzia Guzzi,[5] John McGeagh,[6] Isabelle Petit-Haertlein, [2,3,4]
Serena Gugliotta,[1] Jörg Weiser,[6] Jesùs Angulo,[5] Franck Fieschi,[2,3,7] Anna Bernardi[1*]

*[1]Universita' di Milano, Dipartimento di Chimica , via Golgi 19, 20133 Milano - Italy ; [2]Institut de Biologie Structurale, Université Grenoble I, 41 rue Jules Horowitz, Grenoble, F-38027, France ; [3]CNRS, UMR 5075, Grenoble, F-38000, France; [4]CEA, DSV, Grenoble, F-38000, France; [5]Glycosystems Laboratory, Instituto de Investigaciones Químicas (IIQ), CSIC - Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla-Spain; [6]Anterio Consult&Research GmbH, Augustaanlage 23, D-68165 Mannheim-Germany; [7]Institut Universitaire de France, 103 boulevard Saint-Michel 75005 Paris-France.*

§ These two authors contributed equally to this paper

## Supplementary Information

$^1$H- and $^{13}$C- NMR spectra of **13** and **15** (500 MHz, $D_2O$)

**Table SI-1.** $STD_0$ values for the protons of **13-15**

Synthesis of benzylamines **10**

Synthesis of amides **4a-o**, **11** and **12**

Synthesis of N$^1$,N$^2$-bis(4-(hydroxymethylene)benzyl)amide **12**

**Figure SI-1** – Conformation of the pseudo-sugar ring: stacked 1H NMR spectra

**Figure SI-2** - SPR additional data on DC-SIGN and Langerin inhibition comparison.

**Figure SI-3** - STD NMR spectra of the interaction of ligands **13** (a), **15** (b) and **14** (c) with DC-SIGN

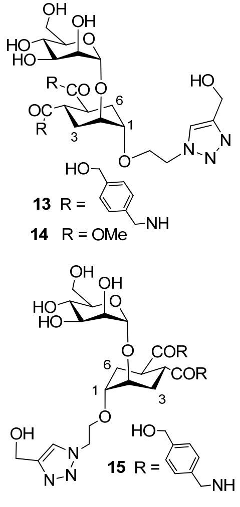
## NMR spectra of 13 and 15 (500 MHz, $D_2O$)

Compound **13.** **$^1$H NMR** (500 MHz, $D_2O$, 25 °C, HDO residual at 4.77 ppm ): δ = 8.04 (s, 1H, H Triazole); 7.31 (m, 4H, $H_{meta}$ (Ar)); 7.2 (m, 4H, $H_{orto}$ (Ar)); 4.93 (d, 1H, H1 (M); 4.64 – 4.59 (m, 4H, O-$CH_2$-$C\underline{H}_2$, $C\underline{H}_2$-OH (Triazole)); 4.57, (bs, 4H, $C\underline{H}_2$-OH (Ar));4.29 – 4.19 (m, 4H, $C\underline{H}_2$-NH (Ar)); 4.02 – 3.91 (m, 2H, O-$C\underline{H}_{2a,b}$-$CH_2$ (Triazole)); 3.95 (m, 1H, H2 (M)); 3.86 (bs, 1H, $H6_a$ (M)); 3.84 (m, 2H, H3 (M), H2 (C)); 3.71 (m, 1H, $H6_b$ (M)); 3.66 (m, 1H, H1 (C)); 3.65 – 3.56 (m, 2H, H4, H5 (M)); 2.76 (m, 1H, H4 (C)); 2.52 (m, 1H, H5 (C)); 1.89 – 1.82 (m, 3H, $H3_{eq}$, $H6_{eq}$, $H6_{ax}$ (C)); 1.54 (m, 1H, $H3_{ax}$ (C)). **$^{13}$C (HSQC) NMR** (500 MHz, $D_2O$, 25 °C): 127.6 ($C_{meta}$ (Ar)); 127.2 ($C_{orto}$ (Ar)); 124.8 (C (Triazole)); 98.7 (C1 (M)); 73.4 (C5 (M)); 70.8 (C3 (M)), C2 (C)); 70.2 (C1 (C)); 66.8 (O-$\underline{C}H_2$-$CH_2$ (Triazole)); 66.7 (C4 (M)); 63.5 ($\underline{C}H_2$-OH (Ar)); 60.9 (C6 (M)); 54.6 ($\underline{C}H_2$-OH (Triazole)); 50.3 (O-$CH_2$-$\underline{C}H_2$ (Triazole)); 42.4 $\underline{C}H_2$-NH (Ar)); 40.6 (C5 (C)); 40.5 (C4 (C)); 27.5 (C6 (C)); 27.1 (C3(C)).

Compound **15.** **$^1$H NMR** (500 MHz, $D_2O$, 25 °C, HDO residual at 4.77 ppm ): 8.04 (s, 1H, H Triazole); 7.31 (m, 4H, $H_{meta}$ (Ar)); 7.2 (m, 4H, $H_{orto}$ (Ar)); 4.89 (d, 1H, H1 (M); 4.64 – 4.60 (m, 4H, O-$CH_2$-$C\underline{H}_2$, $C\underline{H}_2$-OH (Triazole)); 4.57 (d, 4H, $C\underline{H}2$-OH (Ar)); 4.30 – 4.18 (m, 4H, $C\underline{H}_2$-NH (Ar)); 4.03 – 3.91 (m, 2H, O-$C\underline{H}_{2a,b}$-$CH_2$ (Triazole)); 3.90 (m, 1H, H2 (M)); 3.86 (m, 1H, H3 (M)); 3.80 (m, 2H, $H6_a$ (M), H2 (C)); 3.68 (m, 1H, $H6_b$ (M)); 3.67 – 3.58 (m, 3H, H1(C), H4, H5 (M)); 2.78 (m, 1H, H4 (C)); 2.49 (m, 1H, H5 (C)); 1.89 – 1.75 (m, 3H, $H3_{eq}$, $H6_{eq}$, $H6_{ax}$ (C)); 1.63 (m, 1H, $H3_{ax}$ (C)). **$^{13}$C (HSQC) NMR** (500 MHz, $D_2O$, 25 °C): 127.5 ($C_{meta}$ (Ar)); 127.2 ($C_{orto}$ (Ar)); 124.8 (C (Triazole)); 99.5 (C1 (M)); 73.1 (C5 (M)); 73.0 (C1 (C)); 71.4 (C2 (C)); 70.3 (C3 (M)); 70.2 (C2 (M); 66.9 (O-$\underline{C}H_2$-$CH_2$ (Triazole)); 66.7 (C4 (M)); 63.5 ($\underline{C}H_2$-OH (Ar)); 60.7 (C6 (M)); 54.5 ($\underline{C}H_2$-OH (Triazole)); 50.5 (O-$CH_2$-$\underline{C}H_2$ (Triazole)); 42.4 $\underline{C}H_2$-NH (Ar)); 41.0 (C4 (C)); 40.5 (C5 (C)); 29.0 (C3 (C)); 27.3 (C6 (C)).

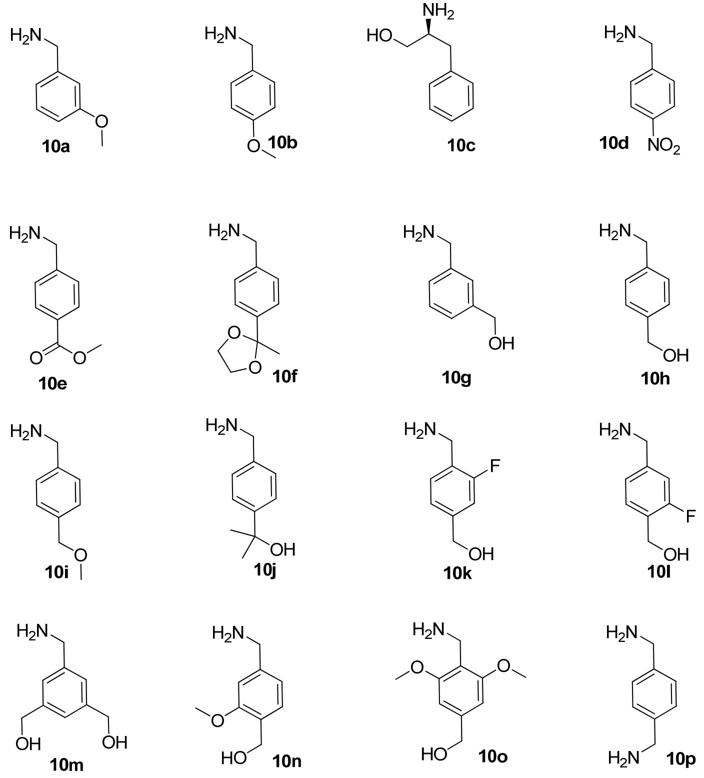
**Table SI-1.** STD$_0$ values for the protons of **14** compared to with the STD$_0$ values of the protons of **13** and its diastereoisomer **15** (500 MHz, D$_2$O, non-overlapping signals)

| Protons | STD$_0$ 13 | STD$_0$ 14 | STD$_0$ 15 |
|---|---|---|---|
| H (Triazole) | 1.5 | 0.4 | 0.76 |
| H$_{meta}$ (Ar) | 6.1 | - | 1.97 |
| H$_{orto}$ (Ar) | 7.3 | - | 1.79 |
| C$\underline{H_2}$-NH (Ar) | 7.5 | - | - |
| H1 (M) | 4.9 | 1.2 | - |
| H2 (M) | 9.0 | 2.6 | 5.2 |
| H3 (M) + H2 (C) | 8.9 | 4.9 (H3(M)) | 7.8 (H3 (M)) |
| H4 (M) | - | - | 8.4 |
| H4 (C) | 9.1 | 2.5 | 4.7 |
| H5 (C) | 7.7 | 2.0 | - |
| H3ax (C) | 6.9 | 2.6 | - |
| H3eq+H6eq+H6ax (C) | 10.1 | 5 (H6ax (C)) | - |


OH
OH
HO
HO
O
R
CO
O
OC
R
6
1
3
HO
N
N
N
O
HO
13 R =
NH
14 R = OMe
OH
OH
HO
HO
O
O
COR
COR
6
1
3
O
OH
N
N-N
HO
15 R =
NH

## Synthesis of benzylamines 10

Amines **10a-e,p** were commercially available. Benzylamines **10g-h,k-m** were prepared according to the general procedure 1 starting from commercially available materials, while compounds **10f,I,j,n,o** were prepared in multistep synthesis.

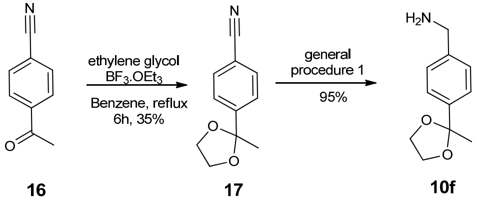

10a
10b
10c
10d
NO2
10e
10f
10g
10h
10i
10j
10k
10l
10m
10n
10o
10p
H2N
NH2
HO
OH
F
O

*Scheme S1*:*Amines* ***10a-p*** *used for the synthesis of pseudo-dimannoside based DC-SIGN ligands* ***4a-o***, ***11*** *and* ***12***.

***General procedure 1: reduction using LiAlH$_4$***

To a 0.5 M solution of LiAlH$_4$ (2 eq per functional group) in dry THF a 0.5 M solution of starting material (1 eq) in dry THF was added at 0°C dropwise under nitrogen atmosphere. After the complete addition the reaction mixture was heated up to reflux for 3 h. The reaction was cooled down to 0°C and worked up by addition of water (1 eq per $LiAlH_4$), 15% NaOH solution (1 eq per $LiAlH_4$) and water (3 eq per $LiAlH_4$) again. The precipitate was filtered off and washed with THF. The filtrate was concentrated under reduced pressure. The product was purified by flash chromatography (DCM with gradient of methanol from 0 to 20%) or used without purification if the purity was sufficient.

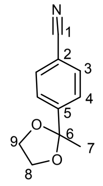
### *(4-(2-methyl-1,3-dioxolan-2-yl)phenyl)methanamine (10f)*



***Scheme S2:*** *synthesis of **10f***

### *4-(2-methyl-1,3-dioxolan-2-yl)benzonitrile (16)*[1]*:*

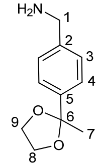
To the solution of 4-acetylbenzonitrile **16** (300 mg, 2.06 mmol, 1 eq) in benzene (5 ml) under nitrogen atmosphere ethyleneglycol (0.35 ml, 6.20 mmol, 3 eq) was added in one portion and $BF_3.Et_3$ (0.08 ml, 0.62 mmol, 0.3 eq) dropwise. The solution was stirred at room temperature for 1 h then heated up to reflux for 5 h. The reaction was quenched by addition of triethylamine (few drops) and then diluted with diethyl ether. The mixture was transferred to a separatory funnel, washed with water and brine, dried over sodium sulphate and concentrated under reduced pressure. The crude was purified by flash chromatography (silica, hexane:diethyl ether = 7:3) to yield 41 % of pure product **17**.



**17**

$^1$**H NMR** (400 MHz, $CDCl_3$): δ = 7.64 – 7.56 (m, 4H, $H_3$, $H_4$), 4.09 – 3.99 (m, 2H, $H_8$, $H_9$), 3.78 – 3.68 (m, 2H, $H_8$, $H_9$), 1.61 (s, 3H, $H_7$). $^{13}$**C NMR** (100 MHz, $CDCl_3$): δ = 148.9 ($C_5$); 132.4 ($C_3$); 126.4 ($C_4$); 119.0 ($C_1$); 112.0 ($C_2$); 108.4 ($C_6$); 64.9 ($C_8$, $C_9$); 27.6 ($C_7$)

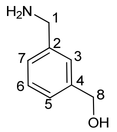
***(4-(2-methyl-1,3-dioxolan-2-yl)phenyl)methanamine (10f)*[2]*:*** Using general procedure 1 starting from nitrile **17**; yield: 95 %.



**10f**

$^1$**H NMR** (400 MHz, $CDCl_3$): δ = 7.45 – 7.21 (m, 4 H, $H_3$, $H_4$), 4.06 – 3.96 (m, 2H, $H_8$, $H_9$), 3.85 (s, 2H, $H_1$), 3.80 – 3.70 (m, 2H, $H_8$, $H_9$), 1.63 (s, 3H, $H_7$); $^{13}$**C NMR** (100 MHz, $CDCl_3$): δ = 142.6, 142.0 ($C_2$, $C_5$); 127.2 ($C_3$); 125.7 ($C_4$); 109.0 ($C_6$); 64.6 ($C_8$, $C_9$); 46.3 ($C_1$); 27.8 ($C_7$).

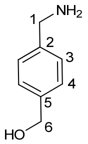
***(3-(aminomethyl)phenyl)methanol 10g*[3]*:*** Using general procedure 1 starting from 3-formylbenzonitrile; yield: 92%.



**10g**

$^1$**H NMR** (400 MHz, $CD_3OD$): δ = 7.41 – 7.28 (m, 4 H, $H_3$, $H_5$, $H_6$, $H_7$), 4.68 (s, 2H, $H_8$), 3.85 (s, 2H, $H_1$). $^{13}$**C NMR** (100 MHz, $CDCl_3$): δ = 143.9, 143.2 ($C_2$, $C_4$); 129.7($C_3$); 127.4, 127.1, 126.7 ($C_5$, $C_6$, $C_7$);, 65.3 ($C_8$); 46.8 ($C_1$).

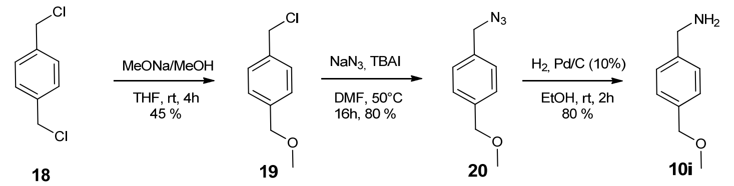
***(4-(aminomethyl)phenyl)methanol (10h)*[3]*:*** Using general procedure 1 starting from methyl 4-(aminomethyl)benzoate hydrochloride; yield: 95 %.



**10h**

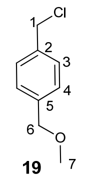
$^1$**H NMR** (400 MHz, $CD_3OD$): δ = 7.32 (s, 4H, $H_3$, $H_4$), 4.58 (s, 2H, $H_6$), 3.78 (s, 2H, $H_1$).

### *(4-(methoxymethyl)phenyl)methanamine (10i)*



***Scheme S3:*** *synthesis of* ***10i***

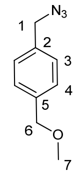
*1-(chloromethyl)-4-(methoxymethyl)benzene* ***(19)***[4]***:*** To a solution of 1,4bis(chloromethyl)benzene **18** (200mg, 1.14 mmol, 1.1 eq) in dry THF (1.3 ml) a 1 M solution of MeONa in methanol (0.9 ml, 1.01mmol, 1 eq) was added dropwise at 0°C under nitrogen atmosphere. After the complete addition the reaction was stirred at room temperature for 16 h. The reaction was diluted with diethylether washed with water and brine, dried over sodium sulphate and concentrated under reduced pressure. The crude product was purified by flash chromatography (hexane with gradient of DCM from 20% to 50%) to yield 45% of pure product **19**.



**19**

$^1$**H NMR** (400 MHz, $CDCl_3$): δ = 7.39 (d, 2H, $H_3$, $J$ = 8.2 Hz), 7.34 (d, 2H, $H_4$, $J$ = 8.2 Hz), 4.60 (s, 2H, $H_1$), 4.48 (s, 2H, $H_6$), 3.41 (s, 3H, $H_7$).

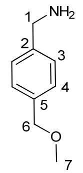
*1-(azidomethyl)-4-(methoxymethyl)benzene* ***(20):*** To a solution of **19** (65 mg, 0.38 mmol, 1 eq) in DMF (1.3 ml) NaN3 (198 mg, 3.05 mmol, 8 eq) and TBAI (14 mg, 0.04 mmol, 0.1 eq) was added. The reaction was heated up to 50°C and stirred for 16 h. The solvent was removed under reduced pressure and the residue was taken up with ether, washed with water and brine, dried over sodium sulphate and concentrated under reduced pressure. The crude product was purified by flash chromatography (hexane with gradient of DCM from 30 % to 70 %) to yield 80 % of pure product **20**.



**20**

$^1$**H NMR** (400 MHz, $CDCl_3$): δ = 7.35 (d, 2H, $H_3$, $J$ = 8.1 Hz), 7.29 (d, 2H, $H_4$, $J$ = 8.1 Hz), 4.45 (s, 2H, $H_6$), 4.31 (s, 2H, $H_1$), 3.38 (s, 3H, $H_7$) $^{13}$**C NMR** (100 MHz, $CDCl_3$): δ = 138.6 ($C_2$); 134.9 ($C_5$); 128.4, 128.2 ($C_3$, $C_4$); 74.4 ($C_6$); 58.3 ($C_7$); 54.7 ($C_1$).

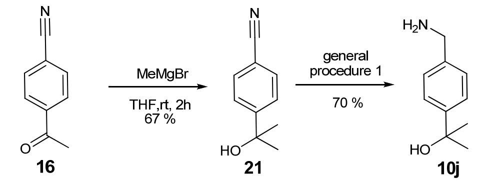
*(4-(methoxymethyl)phenyl)methanamine* ***(10i)***[5]***:*** To a solution of **20** (50 mg, 0.28 mmol, 1 eq) in ethanol (9 ml) 10 % Pd/C was added in catalytic amount. The reaction was stirred under $H_2$ (1 atm) at room temperature for 2 h. The catalyst was filtered off through a celite pad. The filtrate was concentrated under reduced pressure to yield 80 % of pure product **10i**



**10i**

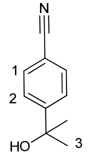
$^1$**H NMR** (400 MHz, $CDCl_3$): δ = 7.30 - 7.24 (m, 4H, $H_3$, $H_4$), 4.41 (s, 2H, $H_6$), 3.83 (s, 2H, $H_1$), 3.35 (s, 3H, $H_7$). $^{13}$**C NMR** (100 MHz, $CDCl_3$): δ = 142.9 ($C_2$); 136.9 ($C_5$); 128.2, 127.3 ($C_3$, $C_4$); 74.6 ($C_6$); 58.2 ($C_7$); 46.4 ($C_1$).

### *2-(4-(aminomethyl)phenyl)propan-2-ol (10j)*



***Scheme S4:*** *Synthesis of* ***10j***

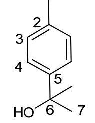
*4-(2-hydroxypropan-2-yl)benzonitrile* ***(21)[6]:*** To a solution of MeMgBr (3 M in $Et_2O$, 2.3 ml, 6.88 mmol, 5 eq) in THF (5 ml) a solution of 4-acetylbenzonitrile **16** (200mg, 1.38 mmol, 1 eq) was added dropwise at 0°C under nitrogen atmosphere. The reaction was stirred for 2 h at room temperature, then cooled down to 0°C and quenched by addition of water. The reaction was diluted with ether, washed with water and brine. The organic phase was dried over sodium sulphate and concentrated under reduced pressure. The crude product was purified by flash chromatography (hexane with gradient of ethyl acetate from 10% - 30%) to yield 67 % of pure product **21**.



**21**

**$^1$H NMR** (400 MHz, $CDCl_3$): δ = 7.65 – 7.55 (m, 4 H, $H_1$, $H_2$), 1.57 (s, 6H, $H_3$).

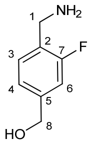
*2-(4-(aminomethyl)phenyl)propan-2-ol* ***(10j)[9]:*** Using general procedure 1 starting from **21**; Yield: 70 %.



**10j**

**$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.46 (d, 2H, $H_4$, $J$ = 8.4 Hz), 7.29 (d, 2H, $H_3$, $J$ = 8.4 Hz), 3.79 (s, 2H, $H_1$), 1.52 (s, 6H, $H_7$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 149.2 ($C_2$, $C_5$); 128.4 ($C_3$);, 126.0 ($C_4$); 73.0 ($C_6$); 47.3 ($C_1$); 32.1 ($C_7$).

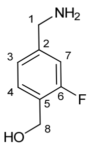
***(4-(aminomethyl)-3-fluorophenyl)methanol (10k):*** Using the general procedure 1 starting from 2-fluoro-4-formylbenzonitrile, yield: 53 %.



**10k**

**$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.40 (t, 1H, $H_3$, $J_{7\text{-}6} = J_{3\text{-}F}$ = 7.8 Hz), 7.19 - 7.09 (m, 2H, $H_6$, $H_4$), 4.63 (s, 2H, $H_8$), 3.86 (s, 2H, $H_1$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 162.4 (d, $C_3$, $J_{C3\text{-}F}$ = 244 Hz); 144.7 (d, $C_5$, $J_{C5\text{-}F}$ = 7.2 Hz); 130.7 (d, $C_7$, $J_{C7\text{-}F}$ = 5.0 Hz); 129.17 (d, $C_2$, $J_{C2\text{-}F}$ = 15.5 Hz); 123.7 (d, $C_4$, $J_{C4\text{-}F}$ = 3.22 Hz); 114.5 (d, $C_6$, $J_{C6\text{-}F}$ = 22.4 Hz); 64.4 (d, $C_8$, $J_{C8\text{-}F}$ = 1.6 Hz); 40.3 (d, $C_1$, $J_{C1\text{-}F}$ = 3.9 Hz).

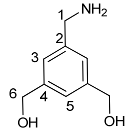
***(4-(aminomethyl)-2-fluorophenyl)methanol (10l):*** Using the general procedure 1 starting from 4-cyano-2-fluorobenzoic acid; yield: 70 %.



**10l**

**$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.44 (t, 1H, $H_4$, $J_{4\text{-}3} = J_{4\text{-}F}$ = 7.8 Hz), 7.19 - 7.09 (m, 2H, $H_3$, $H_7$), 4.68 (s, 2H, $H_8$), 3.81 (s, 2H, $H_1$) **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 162.1 (d, $C_6$, $J_{C6\text{-}F}$ = 245 Hz); 145.8 (d, $C_2$, $J_{C2\text{-}F}$ = 7.1 Hz); 130.7 (d, $C_4$, $J_{C4\text{-}F}$ = 4.9 Hz); 128.1 (d, $C_5$, $J_{C5\text{-}F}$ = 15.2 Hz); 124.2 (d, $C_3$, $J_{C3\text{-}F}$ = 3.16 Hz); 115.0 (d, $C_7$, $J_{C7\text{-}F}$ = 22.0 Hz); 58.8 (d, $C_8$, $J_{C8\text{-}F}$ = 4.4 Hz); 46.2 (d, $C_1$, $J_{C1\text{-}F}$ = 1.5 Hz).

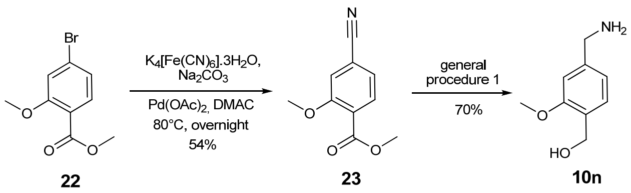
***(5-(aminomethyl)-1,3-phenylene)dimethanol (10m):*** Using general procedure 1 starting from 5-cyanoisophthalic acid, yield: 35 %.



**10m**

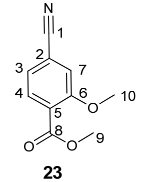
**$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.26 (s, 1H, $H_5$), 7.23 (s, 2H, $H_3$), 4.62 (s, 4H, $H_6$), 3.88 (s, 2H, $H_1$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 143.6 ($C_4$); 141.5 ($C_2$); 126.4 ($C_3$); 125.9 ($C_5$); 65.2 ($C_6$); 46.1 ($C_1$).

***(4-(aminomethyl)-2-methoxyphenyl)methanol (10n)***


K$_4$[Fe(CN)$_6$].3H$_2$O, Na$_2$CO$_3$
Pd(OAc)$_2$, DMAC
80°C, overnight
54%
general procedure 1
70%
22
23
10n

***Scheme S5:*** *Synthesis of* ***10n***

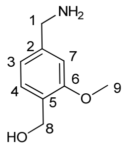
*Methyl 4-cyano-2-methoxybenzoate* ***(23):*** Flask charged with methyl 4-bromo-2-methoxybenzoate **22** (200 mg, 0.82 mmol, 1 eq), $K_4[Fe(CN)_6].3H_2O$ (76 mg, 0.18 mmol, 0.22 eq), $Na_2CO_3$ (87 mg, 0.82 mmol, 1 eq), $Pd(OAc)_2$ (1 mg, 0.04 mmol, 0.05 eq) was degassed and nitrogen atmosphere was introduced (3 cycle), then the reagents were dissolved by addition of dry DMAC (1.4 ml). The reaction mixture was heated up to 80°C and stirred under nitrogen atmosphere for 24 h. The TLC (hexane : AcOEt = 8:2), indicated presence of starting material so another portion of $K_4[Fe(CN)_6].3H_2O$ (172 mg, 0.408 mmol, 0.5 eq), $Na_2CO_3$ (87 eq, 0.82 mmol, 1 eq), $Pd(OAc)_2$ (1 mg, 0.04 mmol, 0.05 eq), and 1mL of DMAC was added. The reaction mixture was stirred for additional 24 h at 70°C then cooled down to room temperature, diluted with ethyl acetate and filtered through a short silica pad. The filtrate was concentrated under reduced pressure and the crude residue was purified by flash chromatography (silica, hexane:EtOAc = 9:1) to yield 54 % of pure product **23**.


1
2
3
4
5
6
7
8
9
10

**23**

$^1$**H NMR** (400 MHz, $CDCl_3$): $\delta = 7.83$ (d, 1H, $H_4$, $J_{4\text{-}3} = 7.9$ Hz), 7.29 (dd, 1H, $H_3$, $J_{3\text{-}4} = 7.9$ Hz, $J_{3\text{-}7} = 1.3$ Hz ), 7.26 (d, 1H, $H_7$, $J_{7\text{-}3} = 1.3$ Hz), 3.95 (s, 3H, $H_{10}$) 3.93 (s, 1H, $H_9$). $^{13}$**C NMR** (100 MHz, $CDCl_3$): $\delta = 165.3$ ($C_8$); 159.0 ($C_2$); 132.2 ($C_4$); 124.8 ($C_5$); 124.0 ($C_3$); 118.0 ($C_1$); 116.6 ($C_6$); 115.4 ($C_7$); 56.3 ($C_9$); 52.7 ($C_{10}$).

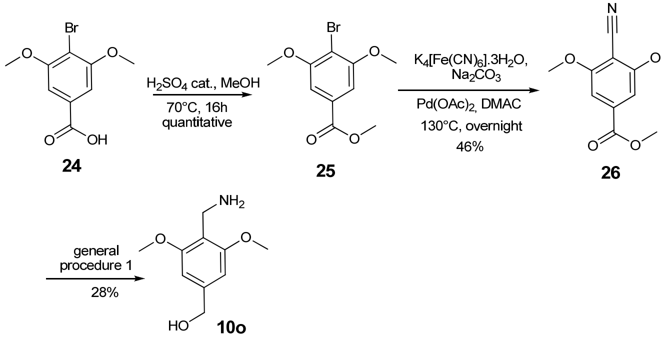
*(4-(aminomethyl)-2-methoxyphenyl)methanol* ***(10n):*** Using general procedure 1 starting from **23**; yield: 70 %.



**10n**

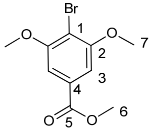
$^1$**H NMR** (400 MHz, $CD_3OD$): $\delta = 7.30$ (d, 1H, $H_4$, $J_{4\text{-}3} = 7.6$ Hz), 6.96 (s, 1H, $H_7$), 6.91 (d, 1H, $H_3$, $J_{3\text{-}4} = 7.6$ Hz), 4.60 (s, 2H, $H_8$) 3.88 (s, 3H, $H_9$), 3.82 (s, 3H, $H_1$). $^{13}$**C NMR** (100 MHz, $CD_3OD$): $\delta = 158.8$ ($C_6$); 143.8 ($C_2$); 129.6 ($C_5$); 129.4 ($C_4$); 120.5 ($C_3$); 110.6 ($C_7$); 60.5 ($C_8$); 56.1 ($C_9$); 46.8 ($C_1$).

***(4-(aminomethyl)-3,5-dimethoxyphenyl)methanol (10o)***


H$_2$SO$_4$ cat., MeOH
70°C, 16h
quantitative
K$_4$[Fe(CN)$_6$].3H$_2$O, Na$_2$CO$_3$
Pd(OAc)$_2$, DMAC
130°C, overnight
46%
general procedure 1
28%
24
25
26
10o

***Scheme S6:*** *synthesis of* ***10o***

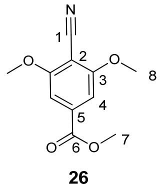
*Methyl 4-bromo-3,5-dimethoxybenzoate* ***(25):*** To a solution of methyl 4-bromo-3,5-dimethoxybenzoate **24** (1,00 g, 3.83 mmol, 1 eq) in methanol (4 ml) $H_2SO_4$ (25 µL, 0.46 mmol, 0.12 eq).was added. The reaction was stirred under reflux for 16 h. The solvent was removed under reduced pressure and the crude residue was dissolved in DCM, washes with saturated $NaHCO_3$ solution and water. The organic phase was dried over sodium sulphate and concentrated under reduced pressure to afford the pure product **25** in quantitative yield.



**25**

**$^1$H NMR** (400 MHz, $CDCl_3$): δ = 7.25 (s, 2H, $H_3$), 3.96 (s, 6H, $H_7$), 3.94 (s, 3H, $H_6$). **$^{13}$C NMR** (100 MHz, $CDCl_3$): δ 166.8 ($C_5$); 157.3 ($C_2$); 130.2 ($C_4$); 107.0 ($C_1$); 105.6 ($C_3$); 56.9 ($C_7$); 52.6 ($C_6$).

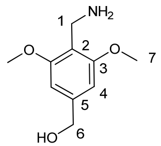
*Methyl 4-cyano-3,5-dimethoxybenzoate **(26):*** Flask charged with **25** (200 mg, 0.73 mmol, 1 eq), $K_4[Fe(CN)_6].3H_2O$ (307 mg, 0.73 mmol, 1 eq), $Na_2CO_3$ (77 mg, 0.727 mmol, 1 eq) and $Pd(OAc)_2$ (3-5%) was degassed and nitrogen atmosphere was introduced (3 cycle), then the reagents were dissolved by addition of dry DMAC (1 ml). The reaction mixture was heated up to 130°C and stirred under nitrogen atmosphere for 24 h. The reaction was cooled down to room temperature, diluted with ethyl acetate and filtered through a short silica pad. The filtrate was concentrated under reduced pressure and the crude residue was purified by flash chromatography (hexane : ethylacetate = 9 : 1) to yield 46 % of pure product **26**.



**26**

**$^1$H NMR** (400 MHz, $CDCl_3$): δ = 7.22 (s, 2H, $H_4$), 3.98 (s, 6H, $H_8$), 3.96 (s, 3H, $H_7$). **$^{13}$C NMR** (100 MHz, $CDCl_3$): δ = 166.4 ($C_6$); 162.6 ($C_3$); 136.1 ($C_5$); 113.6 ($C_1$); 105.0 ($C_4$); 56.9 ($C_8$); 53.2 ($C_7$).

*(4-(aminomethyl)-3,5-dimethoxyphenyl)methanol **(10o):*** Using general procedure 1 starting from **26**; yield: 28 %.



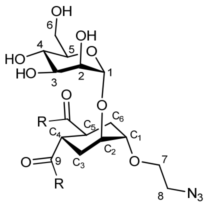
**10o**

**$^1$H NMR** (400 MHz, $CDCl_3$): δ = 6.66 (s, 2H, $H_4$), 4.61 (s, 2H, $H_6$), 3.93 (s, 2H, $H_1$), 3.85 (s, 6H, $H_7$). **$^{13}$C NMR** (100 MHz, $CDCl_3$): δ = 159.8 ($C_3$); 145.4 ($C_5$); 113.8 ($C_2$); 103.0 ($C_4$); 65.4 ($C_6$); 56.6 ($C_7$) 34.1 ($C_1$).

**Synthesis of amides 4a-o, 11 and 12**

***General procedure 2: Synthesis of 1,2-Cyclohexanedicarboxamide, 4-(2-azidoethoxy)-5-(α-D-mannopyranosyloxy),- (1S,2S,4S,5S),4a-e,g-o, 11 and 12***

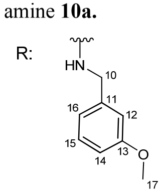
The amine **10** (3 eq) was added to a 0.1 M PNP-scaffold **9** (1 eq) in dry MeCN under stirring and under nitrogen atmosphere at room temperature. After completion (1-24 h; TLC, hex:EtOAc = 1:1 or 2:8) the solvent was evaporated under reduced pressure. The crude product was dissolved in dry methanol (c = 0.1 M), under nitrogen at room temperature, and 1M solution of sodium methoxide in MeOH (2 eq) was added. After reaction (1 h; TLC, $CHCl_3$:MeOH = 9:1 or 8:2) the reaction mixture was diluted with methanol and neutralized with prewashed Amberlite IRA 120-$H^+$. The resin was filtered off and the filtrate was concentrated under reduced pressure. The crude was purified by flash chromatography ($CHCl_3$ with gradient of MeOH from 0 to 20%).



*General structure and numbering of pseudo-dimannoside bis-amides **4a-o***

**N$^1$,N$^2$*-bis(3-methoxybenzyl)amide (4a)***

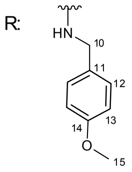
According to general procedure 2 using amine **10a.**



**Yield** = 67 %; **$[\alpha]_D$20** = + 16.2 (c 0.78, MeOH); **$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.20 (t, 2H, $H_{15}$, $J$ = 8.0 Hz), 6.85 – 6.75 (m, 6H, $H_{12}$, $H_{14}$, $H_{16}$), 4.96 (br s, 1H, $H_1$), 4.28 (s, 4H, $H_{10a,b}$), 4.09 - 4.03 (m, 1H, $C_2$), 3.94 - 3.90 (m, 1H, $H_2$), 3.90 – 3.84 (m, 1H, $H_{6a}$), 3.83 – 3.65 (m, 11H, $C_1$, $H_{17}$, $H_{6b}$, $H_{7a,b}$, $H_3$), 3.64 - 3.54 (m, 2H, $H_4$, $H_5$), 3.47 – 3.31 (m, 2H, $H_8$), 3.03 - 2.85 (m, 2H, $C_4$, $C_5$), 2.08 - 1.90 (m, 4H, $C_3$, $C_6$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 177.2, 177.0 ($C_9$); 161.5 ($C_{13}$); 141.8 ($C_{11}$); 130.6 ($C_{15}$); 120.7, 120.7 ($C_{16}$); 114.0, 113.9, 113.7, 113,7 ($C_{14}$, $C_{12}$); 100.4 ($C_1$); 76.6 ($C_3$); 75.7 ($C_5$); 72.7 ($C_{C1}$); 72.5 ($C_2$); 72.4 ($C_{C2}$); 69.3 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 55.8 ($C_{17}$); 52.1

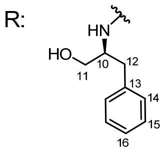
($C_8$); 44.0 ($C_{10}$); 42.1, 41.9 ($C_{C4}$, $C_{C5}$); 29.9, 29.0 ($C_{C3}$, $C_{C6}$). **MS** (HRMS): calculated for $[C_{32}H_{43}N_5O_{11}Na]^+$: 696.28568; found: 696.28599

**N$^1$,N$^2$*-bis(4-methoxybenzyl)amide (4b)***
According to general procedure 2 using amine **10b.**



**Yield** = 73 %; **[α]$_D$20** = + 1.81 (c 0.15, EtOH); **$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.20 (d, 4H, $H_{12}$, $J_{12\text{-}13}$ = 8 Hz), 6.88 (d, 4H, $H_{13}$, $J_{13\text{-}12}$ = 8 Hz), 4.99 (d, 1H, $H_1$,), 4.26 (d, 4H, $H_{10a,b}$, $J_{10a\text{-}10b}$ = 2.4 Hz), 4.10 - 4.06 (m, 1H, $C_2$), 3.95 (dd, 1H, $H_2$, $J_{2\text{-}1}$ = 1.6 Hz, $J_{2\text{-}3}$ = 3.2 Hz), 3.93 – 3.87 (m, 1H, $H_{6a}$), 3.86 – 3.69 (m, 11H, $C_1$, $H_{6b}$, $H_{7a,b}$, $H_3$, $H_{15}$ ), 3.51 - 3.7 (m, 2H, $H_4$, $H_5$), 3.49 – 3.35 (m, 2H, $H_{8a,b}$), 3.01 - 2.86 (m, 2H, $C_4$,$C_5$), 2.08 - 1.90 (m, 4H, $C_3$, $C_6$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 177.0, 176.8 ($C_9$); 160.4 ($C_{14}$); 132.2 ($C_{11}$); 129.9, 129.8 ($C_{12}$); 115.0 ($C_{13}$); 100.4 ($C_1$); 76.6 ($C_3$); 75.6 ($C_5$); 72.2 ($C_{C1}$); 72.5 ($C_2$); 72.5 ( $C_{C2}$); 69.3 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 55.8 ($C_{15}$); 52.2 ($C_8$); 43.6 ($C_{10}$); 42.1, 42.0 ($C_{C4}$, $C_{C5}$); 29.8, 29.0 ($C_{C3}$, $C_{C6}$). **MS** (HRMS): calculated for: $[C_{32}H_{43}N_5O_{11}Na]^+$: 696.28568; found: 696.28465

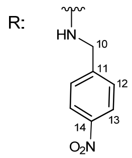
**N$^1$,N$^2$*-bis((L)-phenylalaninol)amide (4c)***
According to general procedure 2 using amine **10c.**



**Yield** = 75 %; **[α]$_D$20** = + 17.8 (c0.5, MeOH); **$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.31 - 7.10 (m, 10H, $H_{14}$, $H_{15}$, $H_{16}$), 4.96 (s, 1H, $H_1$), 4.12 - 4.00 (m, 2H, $H_{10}$), 3.88 (dd, 1H, $H_2$, $J_{2\text{-}1}$ = 1.6 Hz, $J_{2\text{-}3}$ = 3.1 Hz), 3.86 (dd, 2H, $H_5$, $H_{6b}$, $J_{6\text{-}5}$ = 2.1 Hz, $J_{6\text{-}6}$ = 11.7 Hz), 3.72. - 3.63 (m, 4H, $C_2$, $H_{6a}$, $H_7$),

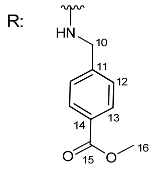
3.63 - 3.59 (m, 2H, $H_4$, $C_1$), 3.58 - 3.54 (m, 1H, $H_3$), 3.47 (d, 4H, $H_{11}$, $J_{11\text{-}10}$= 5.4), 3.39 - 3.33 (m, 2H, $H_8$), 3.00 - 2.88 (m, 2H, $H_{12}$), 2.71 - 2.57 (m, 4H, $H_{11}$, $C_4$, $C_5$), 1.63 - 1.40 (m, 4H, $C_3$, $C_6$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 177.2, 177.1 ($C_9$); 140.1 ($C_{13}$); 130.6, 129.5, 129.4, 127.4 ($C_{14}$, $C_{15}$, $C_{16}$); 100.2 ($C_1$); 76.5 ($C_3$); 75.6 ($C_{C1}$); 72.8 ($C_{C2}$); 72.6 ($C_5$); 72.1 ($C_2$); 69.3 ($C_7$); 68.8 ($C_4$); 64.8 ($C_{11}$); 63.2 ($C_6$); 53.9 ($C_{10}$); 52.1 ($C_8$); 41.9, 41.7 ($C_{C4}$, $C_{C5}$); 38.1 ($C_{12}$); 29.9.5, 28.7 ($C_{C3}$, $C_{C6}$). **MS** (HRMS): calculated for: $[C_{34}H_{47}N_5O_{11}Na]^+$: 724.31698; found: 724.31698

**N$^1$,N$^2$*-bis(4-nitrobenzyl)amide (4d)***
According to general procedure 2 using amine **10d.**



**Yield** = 72 %; **[α]$_D$20** = - 21.5 (c 0.2, MeOH); **$^1$H NMR** (400 MHz, $CD_3OD$): δ = 7.99 (d, 4H, $H_{13}$, $J_{13\text{-}12}$ = 7.4 Hz), 7.40 (d, 4H, $H_{12}$, $J_{12\text{-}13}$ = 7.4 Hz), 4.99 (s, 1H, $H_1$), 4.60 (d, 2H, $H_{10a}$, $J_{10a\text{-}10b}$ = 16.4 Hz), 4.30 (d, 2H, $H_{10b}$, $J_{10b\text{-}10a}$ = 16.4 Hz), 4.10 - 4.06 (m, 1H, $C_2$), 3.94 - 3.91 (m, 1H, $H_2$), 3.89 – 3.84 (m, 1H, $H_{6a}$), 3.84 – 3.66 (m, 5H, $C_1$, $H_{6b}$, $H_{7a,b}$, $H_3$, ), 3.65 - 3.56 (m, 2H, $H_4$, $H_5$), 3.49 - 3.35 (m, 2H, $H_{8a,b}$), 3.10 - 2.95 (m, 2H, $C_4$,$C_5$), 2.10 - 1.92 (m, 4H, $C_3$, $C_6$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): δ = 177.8, 177.6 ($C_9$); 148.3 ($C_{14}$); 148.1 ($C_{11}$); 128.8 ($C_{12}$); 124.5 ($C_{13}$); 100.6 ($C_1$); 76.6 ($C_3$); 75.7 ($C_5$); 72.8 ($C_{C1}$); 72.6 ($C_2$); 72.6 ($C_{C2}$); 69.4 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 52.2 ($C_8$); 43.3 ($C_{10}$); 41.9, 41.8 ($C_{C4}$, $C_{C5}$); 30.1, 29.3 ($C_{C3}$, $C_{C6}$). **MS** (HRMS) calculated for: $[C_{30}H_{37}N_7O_{13}Na]^+$: 726.23470; found: 726.23526

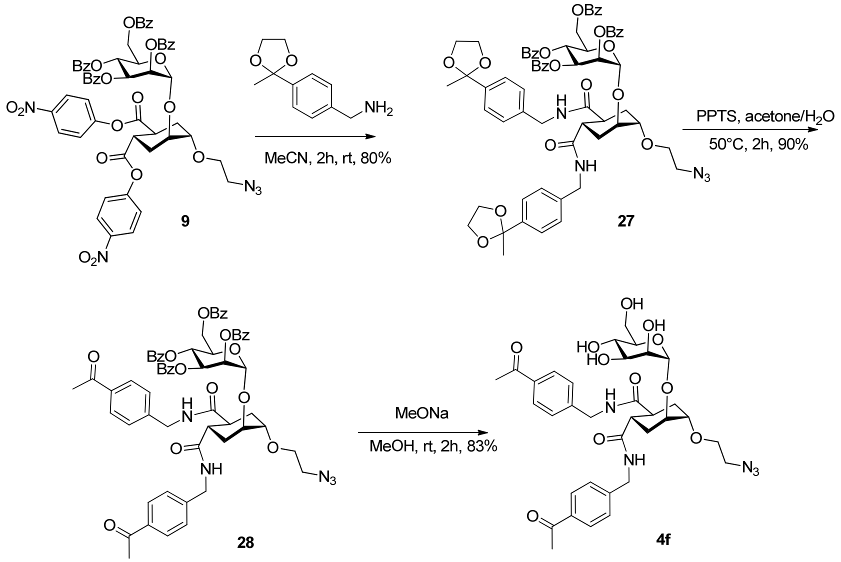
**N$^1$,N$^2$*-bis(4-carbomethoxybenzyl)amide (4e)***
According to general procedure 2 using amine **10e.**



**Yield** = 60 %; **[α]$_D$20** = +54.3 (c = 0.55 in methanol); **$^1$H NMR** (400 MHz, $CD_3OD$): 7.81 (d, 4H, $H_{13}$, $J_{13\text{-}12}$ = 8 Hz), 7.28 (d, 4H, $H_{12}$, $J_{12\text{-}13}$ = 8 Hz), 4.96 (s, 1H, $H_1$), 4.51 - 4.21 (m, 4H, $H_{10a,b}$), 4.08 - 4.03 (m, 1H, $C_2$), 3.92 - 3.89 (m, 1H, $H_2$), 3.88 – 3.83 (m, 7H, $H_{6a}$, $H_{16}$), 3.83 – 3.64 (m, 5H, $C_1$, $H_{6b,}$ $H_{7a,b}$, $H_{3,}$ ), 3.62 - 3.52 (m, 2H, $H_{4,}$ $H_5$), 3.45 – 3.33 (m, 2H, $H_{8a,b}$), 3.05 - 2.90 (m, 2H, $C_4$,$C_5$), 2.52 (s, 3H, $H_{16}$), 2.06 - 1.91 (m, 4H, $C_3$, $C_6$). **$^{13}$C NMR** (100 MHz, $CD_3OD$): 177.5, 177.3 (C9); 168.5 ($C_{15}$); 145.9 ($C_{11}$); 130.8 ($C_{13}$); 130.0 ($C_{14}$); 128.2, 128.1 ($C_{12}$); 100.5 ($C_1$); 76.6 ($C_3$); 75.7 ($C_5$); 72.7 ($C_{C1}$); 72.6 ($C_2$); 72.5 ( $C_{C2}$); 69.3 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 52.7 ($C_{16}$); 52.2 ($C_8$); 43.6 ($C_{10}$); 42.0, 41.9 ($C_{C4}$, $C_{C5}$); 30.0, 29.2 ($C_{C3}$, $C_{C6}$). **MS** (HRMS) calculated for: $[C_{34}H_{43}N_5O_{13}Na]^+$: 752.27551; found: 752.27418

### N$^1$,N$^2$-*bis(4-acetylbenzyl)amide (4f)*

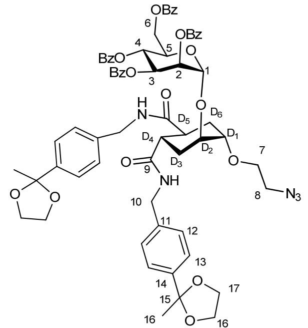
This compound was prepared by the following procedure:



**Scheme S7**: synthesis of **4f**

N$^1$,N$^2$-*bis(4-(2-methyl-1,3-dioxolan-2-yl)benzyl)amide (27):* The amine **10f** (39 mg, 2 mmol, 2.5 eq) was added to a the PNP-scaffold **9** (87 mg, 0.08 mmol, 1eq) in dry MeCN under stirring and under nitrogen atmosphere at room temperature. After completion (2 h) the solvent was evaporated

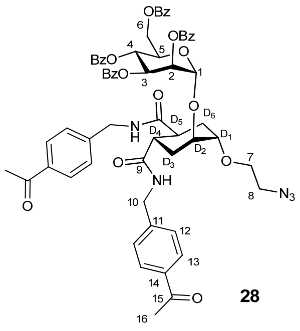
under reduced pressure. The crude was purified by flash chromatography (hexane with gradient of EtOAc from 20 % to 60 %) to yield 80 % of product **27**.



**27**

**1H NMR** (400 MHz, $CDCl_3$): δ = 8.06 – 8.02 (m, 2H, $H_{Bz}$), 8.01 – 7.96 (m, 4H, $H_{Bz}$,), 7.75 – 7.70 (m, 2H, $H_{Bz}$), 7.60 – 7.46 (m, 3H, $H_{Bz}$), 7.43–7.32 (m, 11H, $H_{Bz}$, $H_{12}$), 7.26 – 7.19 (m, 6H, $H_{Bz}$, $H_{13}$), 6.81 (t, 1H, $H_{NH}$, $J_{NH\text{-}10}$ = 5.8 Hz), 6.21 (t, 1H, $H_{NH}$, $J_{NH\text{-}10}$ = 5.8 Hz), 6.10 (t, 1H, $H_4$, $J_{4\text{-}3}$ = $J_{4\text{-}5}$= 10.0 Hz), 5.86 (dd, 1H, $H_3$, $J_{3\text{-}4}$ = 10.0 Hz, $J_{3\text{-}2}$= 3.3), 5.66 (dd, 1H, $H_2$, $J_{2\text{-}1}$ = 1.7 Hz, $J_{2\text{-}3}$= 3.3 Hz), 5.24 (d, 1H, $H_1$, $J_{1\text{-}2}$= 1.7 Hz), 4.70 – 4.64 (m, 1H, $H_{6a}$), 4.53 – 4.20 (m, 6H, $H_{6b}$, $H_{10a,b}$, $H_5$), 4.09 - 4.04 (m, 1H, $C_2$), 4.03 - 3.91 (m, 4H, $H_{16a,\ 17a}$), 3.78 – 3.74 (m, 1H, $C_1$), 3.74 - 3.63 (m, 4H, $H_{16b,\ 17b}$), 3.61 – 3.52 (m, 2H, $H_7$), 3.34 - 3.17 (m, 2H, $H_8$), 3.03 – 2.87 (m, 2H, $C_4$, $C_5$), 2.30 - 2.10 (m, 2H, $C_{3eq}$, $C_{6eq}$), 2.10 - 2.95 (m, 2H, $C_{3ax}$, $C_{6ax}$), 1.60 (s, 3H, $H_{16}$), 1.55 (s, 3H, $H_{16}$). **$^{13}$C NMR** (100 MHz, $CDCl_3$): δ = 174.2, 172.2 ($C_9$); 166.3, 165.9, 165.9, 165.8 ($CO_{BZ}$); 142.7, 142.6 ($C_{11}$); 138.1 ($C_{14}$); 133.9, 133.8, 133.6, 133.4 ($CH_{BZ}$); 130.1, 130.0, 129.9 ($CH_{BZ}$); 129.2, 129.0, 129.0 ($Cquat_{BZ}$); 128.8, 128.7, 128.7, 128.5 ($CH_{BZ}$); 127.7 ($C_{13}$); 125.8, 125.8 ($C_{12}$); 108,9 ($C_{15}$); 97.3 ($C_1$); 75.5 ($C_{C1}$); 74.2 ($C_{C2}$); 71.5 ($C_2$); 70.3 ($C_3$); 70.1 ($C_5$); 68.7 ($C_7$); 66.8 ($C_4$); 64.6, 64.6 ($C_{16}$, $C_{17}$); 63.1 ($C_6$); 50.9 ($C_8$); 43.3 ($C_{10}$); 41.8, 41.1 ($C_{C4}$, $C_{C5}$); 28.8 ($C_{C3}$); 28.6 ($C_{C6}$); 27.8, 27.8 ($C_{16}$); **MS** (ESI) calculated for $[C_{66}H_{67}N_5O_{17}Na]^+$: 1225.2; found: 1224.9.

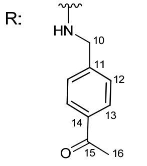
N$^1$,N$^2$-*bis(4-acetylbenzyl)amide (28):* The acetal **27** (60 mg, 0,05mmol, 1 eq) was dissolved in a mixture of acetone and water (10/1, 0.5ml) and to this solution pyridinium 4-toluenesulfonate (PPTS, 1.2 mg, 0.005 mmol, 0.1 eq) was added. The reaction was stirred at 50°C for 4 h. The solvent was evaporated under reduced pressure and the crude was purified by flash chromatography (hexane: EtOAc = 3:7) to yield 90 % of product **28**.



**28**

**1H NMR** (400 MHz, $CDCl_3$): δ = 8.04 (d, 2H, $H_{Bz}$, $J$ = 8.3 Hz), 7.98 (d, 2H, $H_{Bz}$, $J$ = 8.3 Hz), 7.93 (d, 2H, $H_{Bz}$, $J$ = 8.3 Hz), 7.85 (d, 2H, $H_{13}$, $J_{13\text{-}12}$ = 8.1 Hz), 7.80 (d, 2H, $H_{13}$, $J_{13\text{-}12}$ = 8.1 Hz), 7.65 (d, 2H, $H_{Bz}$, $J$ = 8.3 Hz), 7.90 – 7.45 (m, 3H, $H_{Bz}$), 7.41 – 7.28 (m, 11H, $H_{12}$, $H_{Bz}$), 7.22 – 7.16 (m, 2H, $H_{Bz}$), 7.00 (t, 1H, $H_{NH}$, $J_{NH\text{-}10}$ = 5.8 Hz), 6.32 (t, 1H, $H_{NH}$, $J_{NH\text{-}10}$ = 5.8 Hz), 6.12 (t, 1H, $H_4$, $J_{4\text{-}3}$ = $J_{4\text{-}5}$ = 10.0 Hz), 5.86 (dd, 1H, $H_3$, $J_{3\text{-}4}$ = 10.0 Hz, $J_{3\text{-}2}$ = 3.2), 5.66 (m, 1H, $H_2$), 5.24 (br s, 1H, $H_1$), 4.73 – 4.63 (m, 1H, $H_{6a}$), 4.53 – 4.32 (m, 6H, $H_{6b}$, $H_{10a,b}$, $H_5$), 4.13 - 4.06 (m, 1H, $C_2$), 3.79 – 3.74 (m, 1H, $C_1$), 3.66 – 3.54 (m, 2H, $H_7$), 3.34 - 3.17 (m, 2H, $H_8$), 3.09 – 2.88 (m, 2H, $C_4$, $C_5$), 2.53 (s, 3H, $H_{16}$), 2.45 (s, 3H, $H_{16}$), 2.26 - 2.03 (m, 4H, $C_3$, $C_6$). $^{13}$**C NMR** (100 MHz, $CDCl_3$): δ = 197.8, 197.8 ($C_{15}$); 174.5, 174.5 ($C_9$); 166.3, 166.1, 166.1, 165.7 ($CO_{BZ}$); 144.0, 144.9 ($C_{11}$); 136.4, 136.3 ($C_{14}$); 134.0, 133.8, 133.6, 133.4 ($CH_{BZ}$); 130.1, 130.0, 129.9, 129.9 ($CH_{BZ}$); 129.1, 128.9, 129.0 ($Cquat_{BZ}$); 128.9, 128.9, 128.7, 128.5 ($CH_{BZ}$, $C_{13}$); 127.7, 127,6 ($C_{12}$); 97.2 ($C_1$); 75.3 ($C_{C1}$); 74.3 ($C_{C2}$); 71.5 ($C_2$); 70.4 ($C_3$); 70.2 ($C_5$); 68.8 ($C_7$); 66.7 ($C_4$); 63.0 ($C_6$); 50.9 ($C_8$); 43.3, 43.2 ($C_{10}$); 41.7, 41.0 ($C_{C4}$,$C_{C5}$); 29.0 ($C_{C3}$); 28.9 ($C_{C6}$); 26.8, 26.7 ($C_{16}$). **MS** (ESI) calculated for $[C_{62}H_{59}N_5O_{15}Na]^+$: 1137.1; found: 1136.3.

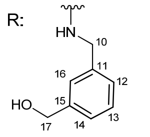
N$^1$,N$^2$-*bis(4-acetylbenzyl)amide* ***(4f):*** Compound **27** (30 mg, 0.026 mmol, 1eq.) was dissolved in dry methanol (0.3 ml), under nitrogen at room temperature, and 1M solution of sodium methoxide in MeOH (0.05 ml, 2 eq.) was added. After reaction completion the reaction mixture was diluted with methanol and neutralized with prewashed Amberlite IRA 120-H$^+$. The resin was filtered off and the solvent evaporated under reduced pressure. The crude was purified by flash chromatography ($CHCl_3$ with gradient of methanol from 5% to 20%) to yield 83 % of product **4g**.



**[α]$_D$20** = - 42.8 (c = 0.1 in methanol); $^1$**H NMR** (400 MHz, $CD_3OD$): δ = 7.82 (d, 4H, $H_{13}$, $J_{13\text{-}12}$ = 7.3 Hz), 7.33 (d, 4H, $H_{12}$, $J_{12\text{-}13}$ = 8.1 Hz), 4.98 (d, 1H, $H_1$, $J_{1\text{-}2}$ = 1.6 Hz), 4.54 - 4.26 (m, 4H, $H_{10a,b}$), 4.10 - 4.05 (m, 1H, $C_2$), 3.94 - 3.90 (m, 1H, $H_2$), 3.90 – 3.84 (m, 1H, $H_{6a}$), 3.84 – 3.66 (m, 5H, $C_1$, $H_{6b}$, $H_{7a,b}$, $H_3$), 3.64 - 3.57 (m, 2H, $H_5$, $H_4$), 3.47 – 3.35 (m, 2H, $H_{8a,b}$), 3.08 - 2.92 (m, 2H, $C_4$,$C_5$), 2.52 (s, 3H, $H_{16}$), 2.52 (s, 3H, $H_{16}$), 2.08 - 1.92 (m, 4H, $C_3$, $C_6$). $^{13}$**C NMR** (100 MHz, $CD_3OD$): δ = 200.2 ($C_{15}$); 177.5, 177.3 ($C_9$); 146.2 ($C_{14}$); 137.1 ($C_{11}$); 129.8 ($C_{13}$); 128.3 ($C_{12}$); 100.5 ($C_1$); 76.6 ($C_3$); 75.7 ($C_5$); 72.8 ($C_{C1}$); 72.6 ($C_2$); 72.6 ( $C_{C2}$); 69.4 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 52.2 ($C_8$); 43.6 ($C_{10}$); 42.0, 41.9 ($C_{C4}$, $C_{C5}$); 30.0, 29.2 ($C_{C3}$, $C_{C6}$); 26.8 ($C_{16}$). **MS** (HRMS) calculated for: $[C_{34}H_{43}N_5O_{11}Na]^+$: 720.28568; found: 720.28552

**N$^1$,N$^2$-*bis(3-(hydroxymethylene)benzyl)amide (4g)***

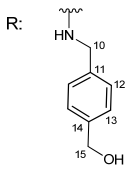
According to general procedure 2 using amine **10g.**



**Yield** = 65 %; **[α]$_D$20** = + 21.5 (c = 0.33 in methanol); $^1$**H NMR** (400 MHz, $CD_3OD$): δ = 7.35 – 7.18 (m, 8H, $H_{12}$, $H_{13}$, $H_{14}$, $H_{16}$), 5.01 (d, 1H, $H_1$, $J_{1\text{-}2}$ = 1,6 Hz), 4,61 (s, 4H, $H_{17a,b}$) 4.41 – 4.31 (m, 4H, $H_{10a,b}$), 4.13 - 4.08 (m, 1H, $C_2$), 3.96 (dd, 1H, $H_2$, $J_{2\text{-}1}$ = 1.6 Hz, $J_{2\text{-}3}$ = 3.2 Hz), 3.94 – 3.89 (m, 1H, $H_{6a}$), 3.88 – 3.71 (m, 5H, $C_1$, $H_{6b}$, $H_{7a,b}$, $H_3$, ), 3.71 - 3.59 (m, 2H, $H_4$, $H_5$), 3.51 – 3.38 (m, 2H, $H_{8a,b}$), 3.06 – 2.92 (m, 2H, $C_4$,$C_5$), 2.09 - 1.95 (m, 4H, $C_3$, $C_6$). $^{13}$**C NMR** (100 MHz, $CD_3OD$): δ = 177.2, 177.0 ($C_9$); 143.1 ($C_{15}$); 140.3 ($C_{11}$); 129.7 ($C_{16}$); 127.7, 127.1, 126.8 ($C_{12}$, $C_{13}$, $C_{14}$); 100.5 ($C_1$); 76.7 ($C_3$); 75.7 ($C_5$); 72.8 ($C_{C1}$); 72.6 ($C_2$, $C_{C2}$); 69.3 ($C_7$); 69.0 ($C_4$); 65.3 ($C_{17}$); 63.2 ($C_6$); 52.2 ($C_8$); 44.0 ($C_{10}$); 42.1, 41.9 ($C_{C4}$, $C_{C5}$); 29.9, 29.1 ($C_{C3}$, $C_{C6}$). **MS** (HRMS) calculated for: $[C_{32}H_{43}N_5O_{11}Na]^+$: 696.28568; found: 696.28578

### N$^1$,N$^2$-*bis(4-(hydroxymethylene)benzyl)amide (4h)*

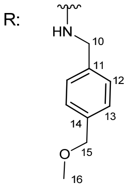
According to general procedure 2 using amine **10h.**



**Yield** = 82 %; **[α]$_D$20** = + 12.1 (c = 0.81 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 7.29 (d, 4H, H$_{12}$, $J_{12-13}$ = 8 Hz), 7.23 (d, 4H, H$_{13}$, $J_{13-12}$ = 8 Hz), 4.96 (d, 1H, H$_1$, $J_{1-2}$ = 1,6 Hz), 4,58 (s, 4H, H$_{15a,b}$), 4.31 (s, 4H, H$_{10a,b}$), 4.08 - 4.03 (m, 1H, C$_2$), 3.93 - 3.89 (m, 1H, H$_2$), 3.89 – 3.84 (m, 1H, H$_{6a}$), 3.84 – 3.65 (m, 5H, C$_1$, H$_{6b,}$ H$_{7a,b}$, H$_{3,}$), 3.64 - 3.54 (m, 2H, H$_4$, H$_5$), 3.47 – 3.35 (m, 2H, H$_{8a,b}$), 3.02 - 2.85 (m, 2H, C$_4$,C$_5$), 2.06 - 1.86 (m, 4H, C$_3$, C$_6$). **$^{13}$C NMR** (100 MHz, CD$_3$OD): δ = 177.2, 177.0 (C$_9$); 141.7 (C$_{14}$); 139.2 (C$_{11}$); 128.6, 128.7 (C$_{12}$); 128.3 (C$_{13}$); 100.4 (C$_1$); 76.6 (C$_3$); 75.7 (C$_5$); 72.7 (C$_{C1}$); 72.5 (C$_2$); 72.4 ( C$_{C2}$); 69.3 (C$_7$); 68.9 (C$_4$); 65.1 (C$_{15}$); 63.2 (C$_6$); 52.1 (C$_8$); 43.8 (C$_{10}$); 42.1, 41.9 (C$_{C4}$, C$_{C5}$); 29.9, 29.0 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: [C$_{32}$H$_{43}$N$_5$O$_{11}$Na]$^+$: 696.28568; found: 696.28423.

### N$^1$,N$^2$-*bis(4-(methoxymethylene)benzyll)amide (4i)*

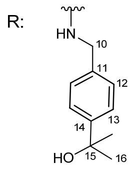
According to general procedure 2 using amine **10i.**



**Yield** = 78 %; **[α]$_D$20** = + 5.2 (c = 0.22 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 7.28 – 7.18 (m, 8H, H$_{12}$, H$_{13}$), 4.94 (br s, 1H, H$_1$), 4.40 (s, 4H, H$_{15}$), 4.25 (br s, 4H, H$_{10}$), 4.06-4.01 (m, 1H, C$_2$), 3.91 - 3.87 (m, 1H, H$_2$), 3.87 – 3.82 (m, 1H, H$_{6a}$), 3.80 – 3.64 (m, 5H, C$_1$, H$_{6b,}$ H$_{7a,b}$, H$_{3,}$), 3.61 - 3.54 (m, 2H, H$_{4,}$ H$_5$), 3.44 – 3.32 (m, 8H, H$_{8a,b}$, H$_{16}$), 3.98 - 2.83 (m, 2H, C$_4$,C$_5$), 2.02 - 1.87 (m, 4H, C$_3$, C$_6$). **$^{13}$C NMR** (100 MHz, CD$_3$OD): δ = 177.2, 177.0 (C$_9$); 139.8 (C$_{14}$); 138.4, 138.3 (C$_{11}$); 129.3 (C$_{13}$); 128.6, 128.5 (C$_{12}$); 100.4 (C$_1$); 76.6 (C$_3$); 75.7 (C$_5$); 75.5 (C$_{15}$); 72.7 (C$_{C1}$); 72.5 (C$_2$); 72.4 ( C$_{C2}$); 69.3 (C$_7$); 68.9 (C$_4$); 63.2 (C$_6$); 58.4 (C$_{16}$); 52.1 (C$_8$); 43.8 (C$_{10}$); 42.1, 41.9 (C$_{C4}$, C$_{C5}$); 29.9, 29.0 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: [C$_{34}$H$_{47}$N$_5$O$_{11}$Na]$^+$: 724.31698; found: 724.31565

### N$^1$,N$^2$-*bis(4-(hydroxy(α,α-dimethyl)methylen)benzyl)amide (4j)*

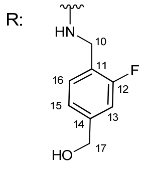
According to general procedure 2 using amine **10j.**



**Yield** = 65 %; **[α]$_D$20** = + 5.3 (c = 0.48 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 7.40 (d, 4H, H$_{12}$, $J_{12-13}$ = 8.1 Hz), 7.18 (d, 4H, H$_{13}$, $J_{13-12}$ = 8.1 Hz), 4.93 (br s, 1H, H$_1$), 4.26 (s, 4H, H$_{10a,b}$), 4.04 - 3.00 (m, 1H, C$_2$), 3.88 - 3.86 (m, 1H, H$_2$), 3.86 – 3.79 (m, 1H, H$_{6a}$), 3.78 – 3.61 (m, 5H, C$_1$, H$_{6b,}$ H$_{7a,b}$, H$_{3,}$), 3.59 - 3.50 (m, 2H, H$_5$, H$_{4,}$), 3.43 – 3.32 (m, 2H, H$_{8a,b}$), 2.98 - 2.82 (m, 2H, C$_4$,C$_5$), 2.00 - 1.84 (m, 4H, C$_3$, C$_6$), 1.47 (s, 12H, H$_{16}$). **$^{13}$C NMR** (100 MHz, CD$_3$OD): δ = 177.1, 176.9 (C$_9$); 149.9 (C$_{14}$); 138.2 (C$_{11}$); 128.3, 128.2 (C$_{12}$); 125.9 (C$_{13}$); 100.4 (C$_1$); 76.6 (C$_3$); 75.7 (C$_5$); 73.0 (C$_{15}$); 72.7 (C$_{C1}$); 72.5 (C$_2$); 72.4 ( C$_{C2}$); 69.3 (C$_7$); 68.9 (C$_4$); 63.2 (C$_6$); 52.1 (C$_8$); 43.8 (C$_{10}$); 42.1, 42.0 (C$_{C4}$, C$_{C5}$); 32.1 (C$_{16}$); 29.9, 29.0 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: [C$_{36}$H$_{51}$N$_5$O$_{11}$Na]$^+$: 752.34828; found: 752. 34702.

### N$^1$,N$^2$-*bis(2-fluoro-4-(hydroxymethylene)benzyl)amide (4k)*

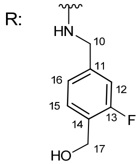
According to general procedure 2 using amine **10k.**



**Yield** = 76 %; **$[\alpha]_D$20** = + 9.3 (c = 0.44 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 7.22 (t, 2H, H$_{16}$, $J_{16\text{-}15}$ = 7.7 Hz, $J_{16\text{-}F}$ = 7.7 Hz), 7.05 (d, 2H, H$_{13}$, $J_{13\text{-}F}$ = 11.5 Hz), 7.04 (d, 2H, H$_{15}$, $J_{15\text{-}16}$ = 7.7 Hz), 4.94 (br s, 1H, H$_1$), 4,55 (s, 4H, H$_{17a,b}$), 4.37 – 4.26 (m, 4H, H$_{10a,b}$), 4.04 - 4.00 (m, 1H, C$_2$), 3.89 (dd, 1H, H$_2$, $J_{2\text{-}1}$ = 1.6 Hz, $J_{2\text{-}3}$ = 3.1 Hz ), 3.87 – 3.81 (m, 1H, H$_{6a}$), 3.81 – 3.63 (m, 5H, C$_1$, H$_{6b,}$ H$_{7a,b}$, H$_{3,}$ ), 3.62 - 3.51 (m, 2H, H$_4$, H$_5$), 3.44 – 3.31 (m, 2H, H$_{8a,b}$), 2.99 - 2.83 (m, 2H, C$_4$,C$_5$), 2.01 - 1.86 (m, 4H, C$_3$, C$_6$). **$^{13}$C NMR** (100 MHz, CD$_3$OD): δ = 177.3, 177.1 (C$_9$); 162.1 (d, C$_{12}$, $J_{12\text{-}F}$ = 250 Hz); 144.8, 144.7 (d, C$_{14}$, $J_{14\text{-}F}$ = 4.3 Hz); 130.6, 130.5 (d, C$_{16}$, $J_{16\text{-}F}$ = 4.5 Hz); 125.5 (d, C$_{11}$, $J_{11\text{-}F}$ = 13.8 Hz); 123.5 (d, C$_{15}$, $J_{15\text{-}F}$ = 3.1 Hz); 114.3 (d, C$_{13}$, $J_{13\text{-}F}$ = 22.1 Hz); 100.4 (C$_1$); 76.6 (C$_3$); 75.7 (C$_5$); 72.7 (C$_{C1}$); 72.5 (C$_2$); 72.5 ( C$_{C2}$); 69.3 (C$_7$); 68.9 (C$_4$); 64.3 (C$_{17}$); 63.1 (C$_6$); 52.1 (C$_8$); 42.0, 41.8 (C$_{C4}$, C$_{C5}$); 37.7 (d, C$_{10}$, $J_{10\text{-}F}$ = 4.6 Hz); 29.8, 29.0 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: $[C_{32}H_{41}F_2N_5O_{11}Na]^+$: 732.26683; found: 732.26691.

**N$^1$,N$^2$-*bis(3-fluoro-4-(hydroxymethylene)benzyl)amide, (4l)***

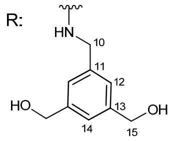
According to general procedure 2 using amine **10l.**



**Yield** = 64 %; **$[\alpha]_D$20** = + 6.9 (c = 0.33 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 77.22 (t, 2H, H$_{15}$, $J_{15\text{-}16}$ = 7.7 Hz, $J_{15\text{-}F}$ = 7.7 Hz), 7.12 (d, 2H, H$_{16}$, $J_{16\text{-}15}$ = 7.7 Hz), 7.04 (d, 2H, H$_{12}$, $J_{12\text{-}F}$ = 11.5 Hz), 5.04 (d, 1H, H$_1$, $J_{1\text{-}2}$ = 1.6 Hz), 4.69 (s, 4H, H$_{17a,b}$) 4.44 – 4.33 (m, 4H, H$_{10a,b}$), 4.15 - 4.11 (m, 1H, C$_2$), 3.99 (dd, 1H, H$_2$, $J_{2\text{-}1}$ = 1.6 Hz, $J_{2\text{-}3}$ = 3.1 Hz ), 3.96 – 3.91 (m, 1H, H$_{6a}$), 3.90 - 3.73 (m, 5H, C$_1$, H$_{6b,}$ H$_{7a,b}$, H$_{3,}$ ), 3.73 - 3.62 (m, 2H, H$_4$, H$_5$), 3.53 – 3.41 (m, 2H, H$_{8a,b}$), 3.08 - 2.93 (m, 2H, C$_4$,C$_5$), 2.13 - 1.96 (m, 4H, C$_3$, C$_6$). **$^{13}$C NMR** (100 MHz, CD$_3$OD): δ = 177.3, 177.1 (C$_9$); 162.0 (d, C$_{13}$, $J_{13\text{-}F}$ = 245.0 Hz); 142.1 (d, C$_{11}$, $J_{11\text{-}F}$ = 7.3 Hz); 130.6 (d, C$_{15}$, $J_{15\text{-}F}$ = 5.0 Hz); 128.4, 128.3 (d, C$_{14}$, $J_{14\text{-}F}$ = 15.2 Hz); 124.2, 124.1 (d, C$_{16}$, $J_{16\text{-}F}$ = 4.7 Hz); 115.1, 115.0 (d, C$_{12}$, $J_{13\text{-}F}$ = 22.5 Hz); 100.5 (C$_1$); 76.6 (C$_3$); 75.7 (C$_5$); 72.8 (C$_{C1}$); 72.6 (C$_2$); 72.5 ( C$_{C2}$); 69.3 (C$_7$); 69.0 (C$_4$); 63.2 (C$_6$); 58.8 (d, C$_{17}$, $J_{17\text{-}F}$ = 4.3 Hz); 52.2 (C$_8$); 43.4 (C$_{10}$); 42.1, 42.0 (C$_{C4}$, C$_{C5}$); 29.9, 29.0 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: $[C_{32}H_{41}F_2N_5O_{11}Na]^+$: 732.26683; found: 732.26529.

**N$^1$,N$^2$-*bis(4,5-di-(hydroxymethylene)benzyl)amide (4m)***

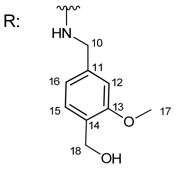
According to general procedure 2, using amine **10m.**



**Yield** = 63 %; **$[\alpha]_D$20** = + 9.4 (c = 0.2 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 7.20 (s, 2H, H$_{14}$), 7.13 (s, 2H, H$_{12}$), 4.95 (s, 1H, H$_1$), 4.67 - 4.46 (m, 8H, H$_{15}$), 4.36 - 4.24 (m, 4H, H$_{10}$), 4.01 (d, 1H, H$_5$, $J$ = 2.7 Hz), 3.90 (dd, 1H, H$_2$, $J_{2\text{-}1}$ = 1.7 Hz, $J_{2\text{-}3}$ = 3.2 Hz), 3.88 (d, 1H, H$_{6b,}$ $J_{6\text{-}6}$ = 11.3 Hz), 3.83 - 3.76(m, 1H, H$_3$), 3.76 - 3.63 (m, 4H, C$_{2,}$ H$_{6a,}$ H$_7$), 3.62 - 3.52 (m, 2H, H$_4$, C$_1$), 3.45 - 3.33 (m, 2H, H$_8$), 3.04 - 2.84 (m, 2H, C$_4$, C$_5$), 2.07 - 1.90 (m, 4H, C$_3$, C$_6$). **$^{13}$C NMR** (100 MHz, CD$_3$OD): δ = 177.9, 177,7 (C$_9$); 143.8, (C$_{13}$); 140.9 (C$_{11}$); 126.4, 126.3 (C$_{12}$); 125.8 (C$_{14}$); 100.0 (C$_1$); 77.2 (C$_3$); 76.2 (C$_{C1}$); 73.2 (C$_{C2}$); 73.1 (C$_5$); 72.9 (C$_2$); 69.8 (C$_7$); 69.4 (C$_4$); 65.2(C$_{15}$); 63.6 (C$_6$); 52.2 (C$_8$); 44.4 (C$_{10}$); 42.5, 42.3 (C$_{C4}$, C$_{C5}$); 30.5, 29.6 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: $[C_{34}H_{47}N_5O_{13}Na]^+$: 756.30681; found: 756.30576.

**N$^1$,N$^2$-*bis(3-methoxy-4-(hydroxymethylene)benzyl)amide (4n)***

According to general procedure 2, using amine **10n.**

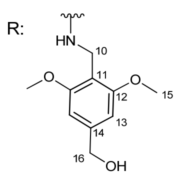


**Yield** = 80 %; **$[\alpha]_D$20** = - 24.3 (c = 0.15 in methanol); **$^1$H NMR** (400 MHz, CD$_3$OD): δ = 7.23 (d, 2H, H$_{15}$, $J_{15\text{-}16}$ = 7.6 Hz), 6.81 (s, 2H, H$_{12}$), 6.80 (d, 2H, H$_{16}$, $J_{15\text{-}16}$ = 7.6 Hz), 4.94 (s, 1H, H$_1$), 4.55 (s, 4H, H$_{18}$), 4.27 (m, 4H, H$_{10}$), 4.04 (d, 1H, H$_5$, $J_{6\text{-}5}$ = 2.7 Hz), 3.91 (dd, 1H, H$_2$, $J_{2\text{-}1}$ = 1.7 Hz, $J_{2\text{-}3}$ =

3.2 Hz), 3.87 - 3.81 (m, 1H, $H_{6b}$), 3.79 - 3.76 (m, 6H, $H_{17}$), 3.76 - 3.72 (m, 2H, $H_3$, $C_2$), 3.72. - 3.64 (m, 3H, $H_{6a,}$ $H_7$), 3.59 - 3.54 (m, 2H, $H_4$, $C_1$), 3.44 - 3.33 (m, 2H, $H_8$), 3.08 - 2.77 (m, 4H, $C_4$, $C_5$), 2.05 - 1.82 (m, 4H, $C_3$, $C_6$). $^{13}$**C NMR** (100 MHz, $CD_3OD$): δ = 177.3, 177.0 ($C_9$); 158.6 ($C_{13}$); 140.9, 140.8 ($C_{11}$); 129.5 ($C_{14}$); 129.1 ($C_{15}$); 120.2, 120.1 ($C_{16}$); 110.4, 110.3 ($C_{12}$); 100.4 ($C_1$); 76.7 ($C_3$); 75.7 ($C_{C1}$); 72.7 ($C_{C2}$); 72.5 ($C_5$); 72.4 ($C_2$); 69.3 ($C_7$); 68.9 ($C_4$); 63.2 ($C_6$); 60.4 ($C_{18}$); 56.0 ($C_{17}$); 52.1 ($C_8$); 43.9 ($C_{10}$); 42.0, 41.9 ($C_{C4}$, $C_{C5}$); 30.0, 29.1 ($C_{C3}$, $C_{C6}$). **MS** (HRMS) calculated for: $[C_{34}H_{47}N_5O_{13}Na]^+$: 756.30681; found: 756.30567

### N$^1$,N$^2$*-bis(2,6-dimethoxy-4-(hydroxymethylene)benzyl)amide (4o)*

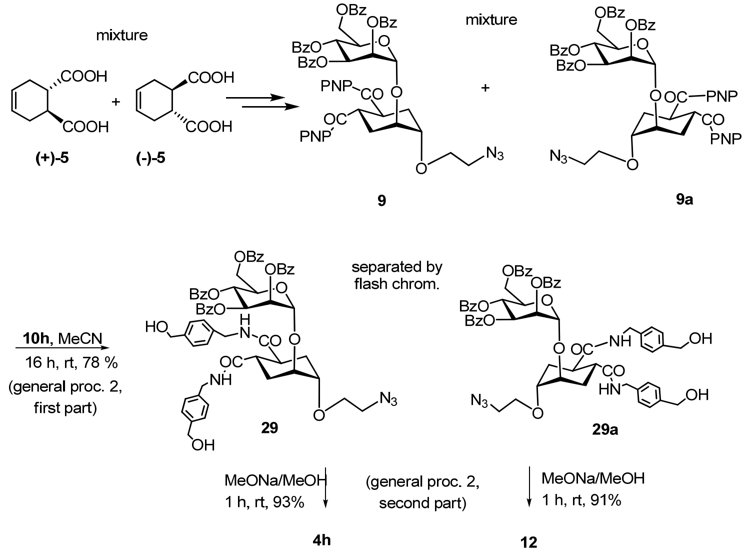
According to general procedure 2, using amine **10o.**


R:
HN
10
11
12
15
13
14
16
OH

**Yield** = 28 %; **[α]$_D$20** = + 36.5 (c = 0.1 in methanol); $^1$**H NMR** (400 MHz, $CD_3OD$): δ = 6.63 (s, 4H, $H_{13}$), 4.97 (s, 1H, $H_1$), 4.63 (s, 4H, $H_{16}$), 4.46 (dd, 2H, $H_{10a,}$ $J_{10a\text{-}NH}$ = 7.6 Hz, $J_{10a\text{-}10b}$ = 13.4 Hz), 4.17 (dd, 2H, $H_{10b}$, $J_{10b\text{-}NH}$ = 5.8 Hz), 3.97 (d, 1H, $H_5$, $J_{6\text{-}5}$ = 2.6 Hz), 3.90 - 3.76 (m, 14H, $H_{15'}$, $H_2$, $H_{6b}$), 3.74 - 3.59 (m, 5H, $H_7$, $H_{6a,}$ $H_3$, $C_2$), 3.59 - 3.48 (m, 2H, $H_{4,}$ $C_1$), 3.35 - 3.32 (m, 4H, $H_8$), 2.81 - 2.62 (m, 2H, $C_4$, $C_5$), 1.98 - 1.72 (m, 4H, $C_3$, $C_6$). $^{13}$**C NMR** (100 MHz, $CD_3OD$): δ = 176.4, 176.1 ($C_9$); 160.2, 160.1 ($C_{12}$); 144.7 ($C_{11}$); 129.1 ($C_{14}$); 103.3 ($C_{13}$); 100.4 ($C_1$); 76.8 ($C_3$); 75.6 ($C_{C1}$); 72.7 ($C_{C2}$); 72.5 ($C_5$); 72.2 ($C_2$); 69.3 ($C_7$); 68.9 ($C_4$); 65.4 ($C_{16}$); 63.2 ($C_6$); 56.4 ($C_{15}$); 52.1 ($C_8$); 42.1, 42.0 ($C_{C4}$, $C_{C5}$); 33.5, 33.3 ($C_{10}$); 29.5, 28.6 ($C_{C3}$, $C_{C6}$). **MS** (HRMS) calculated for: $[C_{36}H_{51}N_5O_{15}Na]^+$: 816.32794; found: 816.32600

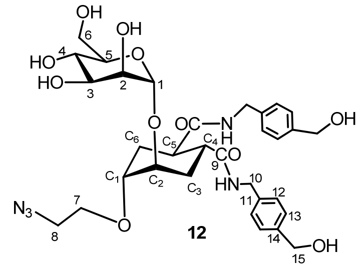
### *Synthesis of N$^1$,N$^2$-bis(4-(hydroxymethylene)benzyl)amide 12 (diastereoisomer of 4h)*

During the large scale synthesis of **4h**, the diacid **5** was used as the commercially available 4:1 mixture of (+)-**5** and (-)-**5**. Following the reaction path described in scheme 1 (paper), a mixture of diastereoisomers **9** and **9a** was obtained, which was treated with amine **10h** according to the first part of general procedure 2. This led to two diastereoisomers **29** and **29a** which were partially separable by flash chromatography ($Et_2O$ with gradient of EtOAc from 30% to 70%).


mixture
COOH
COOH
+
COOH
COOH
(+)-5
(-)-5
OBz
OBz
BzO
BzO
O
PNP
CO
O
OC
PNP
O
N3
9
mixture
+
OBz
OBz
BzO
BzO
O
O
OC—PNP
CO
PNP
N3
O
9a
10h, MeCN
16 h, rt, 78 %
(general proc. 2, first part)
OBz
OBz
BzO
BzO
HO
H
N
CO
O
OC
NH
OH
O
N3
29
separated by flash chrom.
OBz
OBz
BzO
BzO
O
OC
NH
OH
CO
HN
OH
N3
O
29a
MeONa/MeOH
1 h, rt, 93%
(general proc. 2, second part)
MeONa/MeOH
1 h, rt, 91%
4h
12

**Scheme S8**: synthesis of **12**

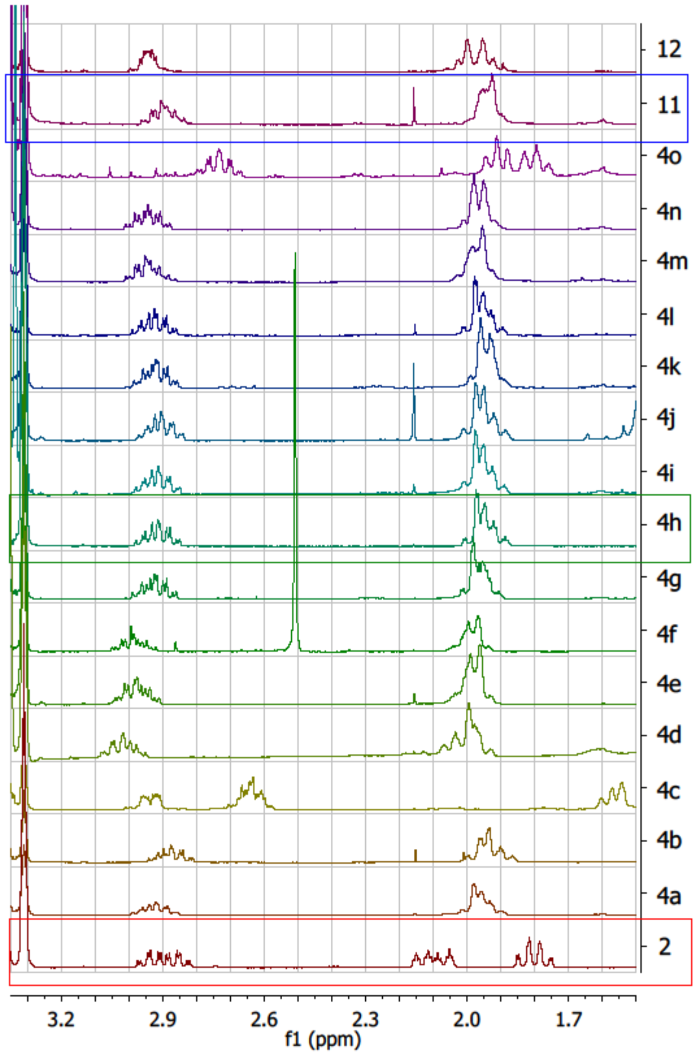
Deprotection of compounds **29** and **29a** with MeONa following the second part of the general procedure 2, gave the final DC-SIGN ligands **4h** and **12,** respectively.


OH
6
OH
HO
4
5
O
HO
3
2
1
O
OC
N
OH
C6
C5
C4
H
CO
9
C1
C2
C3
HN
10
11
12
N3
7
8
O
12
13
14
OH
15

**12: [α]$_D$20** = + 36.9 (c 0.23, MeOH); $^1$**H NMR** (400 MHz, CD$_3$OD): δ = 7.27 (d, 4H, H$_{12}$, $J_{12\text{-}13}$ = 8.0 Hz), 7.22 (d, 4H, H$_{13}$, $J_{13\text{-}12}$ = 8 Hz), 4.91 (d, 1H, H$_1$, $J_{1\text{-}2}$ = 1,5 Hz), 4.56 (s, 4H, H$_{15a,b}$) 4.30 – 4.27 (m, 4H, H$_{10a,b}$), 4.01 – 3.97 (m, 1H, C$_2$), 3.88 – 3.82 (m, 1H, H$_{6a}$), 3.82 – 3.74 (m, 3H, H$_2$, H$_{7a,b}$, ), 3.74 - 3.69 (m, 1H, H$_3$), 3.69 - 3.61 (m, 3H, C$_1$, H$_5$, H$_{6b}$), .3.60 – 3.54 (m, 1H, H$_4$), 3.45 – 3.33 (m, 2H, H$_{8a,b}$), 3.01 - 2.84 (m, 2H, C$_4$,C$_5$), 2.07 - 1.88 (m, 4H, C$_3$, C$_6$). $^{13}$**C NMR** (100 MHz, CD$_3$OD): δ = 177.2, 177.0 (C9); 141.7, 141.6 (C$_{14}$); 139.3, 139.2 (C$_{11}$); 128.5 128.5 (C$_{12}$); 128.3, 128.3 (C$_{13}$); 101.9 (C$_1$); 75.8, 75.7 (C$_3$, C$_5$); 74.4 (C$_{C2}$); 72.7 (C$_2$); 72.4 (C$_1$); 69.2 (C$_7$); 69.0 (C$_4$); 65.1, 65.1 (C$_{15}$); 63.3 (C$_6$); 52.2 (C$_8$); 43.8, 43.7 (C$_{10}$); 42.5, 41.8 (C$_{C4}$, C$_{C5}$); 31.1, 29.5 (C$_{C3}$, C$_{C6}$). **MS** (HRMS) calculated for: $[C_{32}H_{43}N_5O_{11}Na]^+$ = 696.8568, found: 696.8560

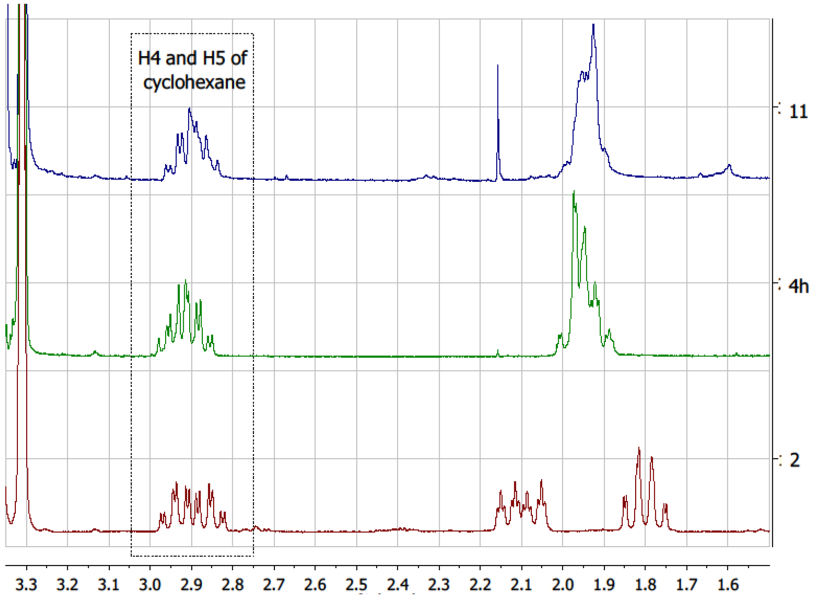
**Figure SI-1** – Conformation of the pseudo-sugar ring: stacked 1H NMR spectra

(400 MHz, CD$_3$OD, methanol reference peak at 3.31 ppm)

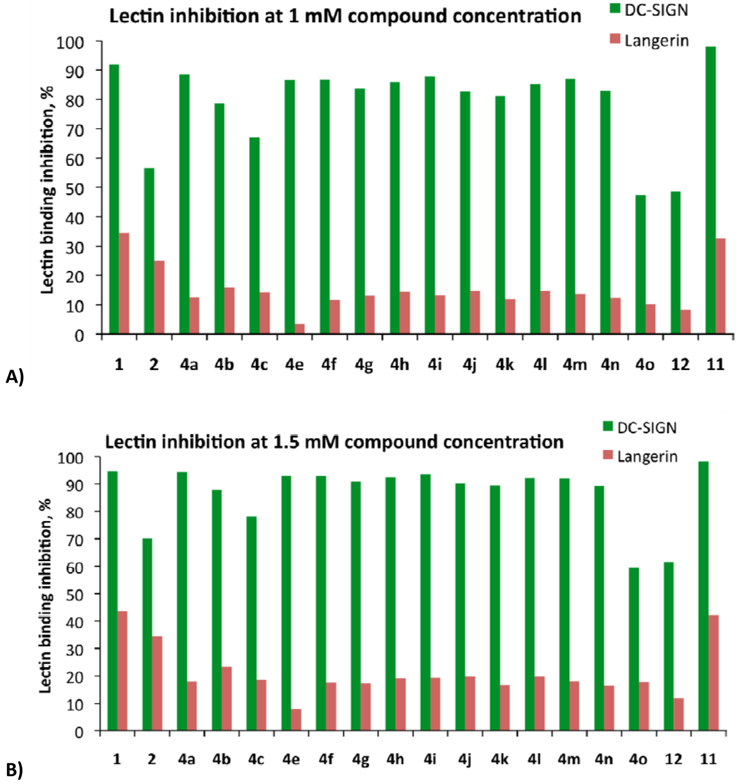


Stacked 1H NMR spectra of **2**, **4a-4o**, **11** and **12**. The diagnostic peaks for (axial) protons H4(C) and H5(C) are shown. Depending on the actual structure of the ligand, signals can be shifted upfield or downfield relative to the reference protons in **2** (2.8-3.0 ppm), but the shape of the multiplets remains the same (also see the expansion below).

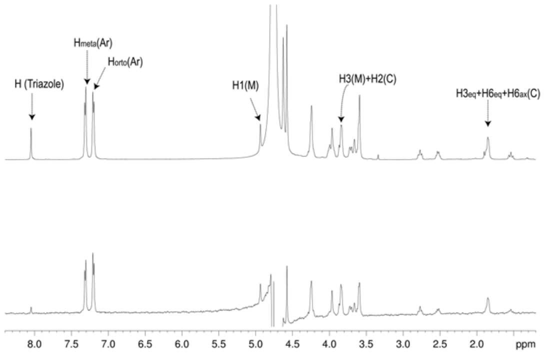
The H4(C) and H5(C) cyclohexane signals of the diastereisomer **12** are isochronous (2.95 ppm), but analysis of the spectrum of its derivative **15** (see Figure SI-3) clearly shows the two protons are axial.



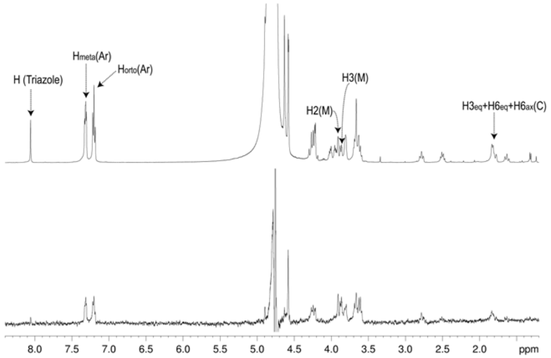
**SPR additional data on DC-SIGN and Langerin inhibition comparison.**



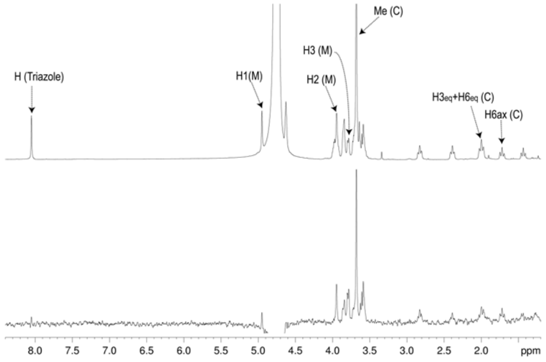
**Supplemental Figure SI-2.** SPR results representing inhibition of Langerin ECD binding to Man-BSA/dextran surface by mannose-based compounds: DC-SIGN and langerin Inhibition level obtained for 1 mM (A) and 1.5 MM (B) concentration of the compounds.



**a) 13**



**b) 15**



**c) 14**

**Figure SI-3.** STD NMR spectra of the interaction of ligands **13** (a), **15** (b) and **14** (c, from *ref 8*) with DC-SIGN in buffered $D_2O$ (150 mM NaCl, 4 mM $CaCl_2$, 25 mM d-Tris, pD 8). Upper trace: $^1H$ NMR reference spectrum

(off-resonance frequency 40 ppm); Lower trace: STD spectrum (on-resonance frequency 0 ppm). 1 mM lignad, DC-SIGN ECD (19 µM), at 25 ºC (500 MHz), saturation time of 1 s. Key proton signals are labeled.

[1] Z. Hyder, J. W. Ruan and J. L. Xiao, *Chemistry-a European Journal* **2008**, *14*, 5555-5566.
[2] H. Langhals, A. Obermeier, Y. Floredo, A. Zanelli and L. Flamigni, *Chemistry-a European Journal* **2009**, *15*, 12733-12744.
[3] J. R. Martinelli, D. A. Watson, D. M. M. Freckmann, T. E. Barder and S. L. Buchwald, *Journal of Organic Chemistry* **2008**, *73*, 7102-7107.
[4] P. Y. Keng, I. Shim, B. D. Korth, J. F. Douglas and J. Pyun, *Acs Nano* **2007**, *1*, 279-292.
[5] C. Salome, E. Salome-Grosjean, K. D. Park, P. Morieux, R. Swendiman, E. DeMarco, J. P. Stables and H. Kohn, *Journal of Medicinal Chemistry* **2010**, *53*, 1288-1305.
[6] J. L. Tucker, M. Couturier, M. J. Castaldi, C. K. F. Chiu and D. Gestmann, *Synthetic Communications* **2006**, *36*, 2145-2150.

# 18. The list of compounds and their names used in this thesis

# The compound names used in thesis and corresponding initial names of the same compounds

| Compound name in thesis/papers | Compound name as the received samples |
|---|---|
| 3a | Fuc 1-02 |
| 3f | Fuc 1-09 |
| 3b | Fuc 1-14 |
| 3d | Fuc 1-19 |
| 3g | Fuc 1-28 |
| 3h | Fuc 1-30 |
| 3c | Fuc 1-31 |
| 3e | Fuc 1-32 |
| 10a | Fuc4-2 |
| 10c | Fuc4-3 |
| 10d | Fuc4-4 |
| 11b | Fuc5-1 |
| 11c | Fuc5-2 |
| 11a | Fuc5-3 |
| - | Fuc5-4 |
| 12b | Fuc6-1 |
| 10b | Fuc 4-1 |
| 10b-azide1 | Fuc4-1 azide |
| 10b-azide2 | Fuc4-1 azide-2 |
| L fucose | L fucose |
| Lewis X | Lewis X |
| BB178 | BB178 |
| BB203 | BB203 |
| BB213 | BB213 |
| BB207 | BB207 |
| BO2 | BO2 |
| BO3 | BO3 |
| BB112 | BB112 |
| BB133 | BB133 |
| ZL1(D) | ZL1(D) |
| Manα1-3Man | Manα1-3Man |
| Manα1-2Man | Manα1-2Man |
| NV233 | NV233 |
| NV243 | NV243 |
| NV253 | NV253 |
| NV270 | NV270 |
| - | Man-029 |
| 4j | Man-031 |
| 4f | Man-032 |
| 4i | Man-033 |
| 4e | Man-035 |
| 4b | Man-036 |
| 4d | Man-037 |
| 4a | Man-038 |
| 4g | Man-039 |
| 4k | Man-040 |
| 4l | Man-041 |
| - | Man-042 |
| - | Man043 |

| Compound name in thesis/papers | Compound name as the received samples |
|---|---|
| 4m | Man044 |
| 4n | Man045 |
| 4o | Man046 |
| 4c | Man047 |
| 12 | Man048 |
| 11 | Man049 |
| 4h | Man-030 |
| psDi | psDi/Man002/Man001 |
| Mannose | Mannose |
| psTri | psTri/Man005 |
| psDi dendron | poly001 (Polymann001)/SARA127 |
| psTri dendron | poly002 (Polymann002)/SARA133 |
| ***S6-1-Oman*** | poly006 (Polymann006) |
| ***S4-Oman*** | poly005 (Polymann005) |
| ***S4-psDi*** | Poly008 (Polymann008) |
| ***S6-1- psDi*** | poly009 (Polymann009) |
| ***S9-1-Oman*** | poly012 (Polymann012) |
| ***S9-2-Oman*** | poly013 (Polymann013) |
| ***S12-Oman*** | poly014 (Polymann014) |
| ***S18-1-Oman*** | poly015 (Polymann015) |
| ***S18-2-Oman*** | poly016 (Polymann016) |
| ***S4-psTri*** | poly017 (Polymann017) |
| ***S4-4h*** | poly018 (Polymann018) |
| ***S6-1-4h*** | poly019 (Polymann019) |
| ***S9-1- psDi*** | poly-020 (Polymann020) |
| ***S6-1- psTri*** | poly021 (Polymann021) |
| ***S6-2- psDi*** | poly022 (Polymann022) |
| ***S9-1-4h*** | PM023 |
| ***R1*** | PM024 |
| ***R2*** | PM025 |
| ***R3*** | PM026 |
| ***L-fucose β-Ala*** | 9a |
| ***L-fucose β-Ala+*** | L-fuc b-Ala (Fuc β-Ala azide) |
| ***S4-fucβAla*** | Dendri b-Ala fuc (Dendrimer Fuc β-Ala) |
| ***S4-10b-azide1*** | Dendri Fuc4-1 (Dendrimer Fuc4-1) |
| ***S6-1-Cfuc*** | BB02 |
| ***S4-Ofuc*** | MA05 |
| ***S6-1-Ofuc*** | MA07 |
| ***S18-1-Ofuc*** | RCF-26 |
| ***S12-Ofuc*** | RR390b |
| ***S4-Cfuc*** | BB01 |
| ***S9-1-Cfuc*** | BB013 |
| ***S12-Cfuc*** | BB014 |
| ***S9-1-Cman*** | BB015 |
| ***S12-Cman*** | BB016 |
| BO | BO |
| G3(pseudodi)$_{32}$ | Dendri-Man-003 |
| G3(pseudotri)$_{32}$ | Dendri-Man-004 |

# 19. The list of $IC_{50}$ values for all compounds

# All acquired $IC_{50}$ values for all compounds during the thesis

| Date | Compound | $IC_{50}$, μM | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| November 2009 | psDi Fc2 | 1252 | 1.3467 | -0.79138 | 99.208 | 1282.1 |
| | psTri Fc2 | 166 | 1.0517 | 0.6077 | 100.61 | 162.65 |
| | psDi Fc3 | 1193 | 1.3491 | -0.75415 | 99.245 | 1219.6 |
| | psTri Fc3 | 168 | 1.0597 | 0.60731 | 100.61 | 164.31 |
| | Dendri-Nan004 fc2 | 1 | 1.573 | -0.7239 | 99.275 | 1.3158 |
| | Dendri-Man004 fc3 | 1 | 1.5926 | -1.0721 | 98.927 | 1.2556 |
| | Dendri-Man003 fc2 | 2 | 1.2108 | -0.7727 | 99.226 | 2.2628 |
| | Dendri-Man003 fc3 | 2 | 1.2099 | -0.87759 | 99.121 | 2.1015 |
| | dendri-G3 fc2 | 20 | 1.1132 | -1.2221 | 98.777 | 20.471 |
| | dendri-G3 fc3 | 18 | 1.0387 | -0.85847 | 99.141 | 18.858 |
| | SARA133 fc2 | 118 | 1.2282 | -0.5601 | 99.439 | 120.33 |
| | SARA133 fc3 | 115 | 1.2163 | -0.55428 | 99.445 | 116.76 |
| | SARA127 fc2 | 225 | 1.3654 | -1.5647 | 98.434 | 235.13 |
| | SARA127 fc3 | 214 | 1.3287 | -1.9026 | 98.096 | 226.95 |
| | MS517 fc2 | 898 | 1.0951 | -1.4972 | 98.502 | 948.72 |
| | MS517 fc3 | 824 | 1.0765 | -2.1316 | 97.867 | 892.07 |
| | mannose fc2 | 2850 | 1.2995 | -1.2833 | 98.716 | 2965 |
| | mannose fc3 | 2728 | 1.3987 | -1.2966 | 98.702 | 2831.2 |
| 13 September 2010 | Fuc5-2 Fc2-Fc3 | 444 | 1.5 | 0.00525 | 99 | 450 |
| | Fuc4-1 Fc2-Fc3 | 394 | 1.46 | 0.00284 | 99.5 | 397 |
| | Fuc4-2 Fc2-Fc3 | 1080 | 1.4 | 0.00996 | 100 | 1080 |
| | Fuc4-3 Fc2-Fc3 | 376 | 1.46 | 0.00184 | 99.5 | 379 |
| | Fuc4-4 Fc2-Fc3 | 515 | 1.42 | 0.0000817 | 99.6 | 518 |
| | Fuc6-1 Fc2-Fc3 | 966 | 1.42 | 0.998 | 100 | 952 |
| | Fuc1-2 Fc2-Fc3 | 480 | 1.44 | 0.428 | 99.3 | 482 |
| | LewisX Fc2-Fc3 | 684 | 1.37 | 0.0951 | 99.1 | 692 |
| | Man005 Fc2-Fc3 | 109 | 0.881 | 0.1 | 104 | 99.7 |
| | Fuc5-2 Fc4-Fc3 | 524 | 1.67 | 7.66E-04 | 98.2 | 536 |
| | Fuc4-1 Fc4-Fc3 | 468 | 1.58 | 0.379 | 98.2 | 477 |
| | Fuc4-2 Fc4-Fc3 | 1493 | 1.48 | 0.0636 | 97.7 | 1540 |
| | Fuc4-3 Fc4-Fc3 | 459 | 1.5 | 0.241 | 98.2 | 469 |
| | Fuc4-4 Fc4-Fc3 | 625 | 1.47 | 0.88 | 98.1 | 634 |
| | Fuc6-1 Fc4-Fc3 | 1286 | 1.43 | 0.0596 | 98.1 | 1320 |
| | Fuc1-2 Fc4-Fc3 | 596 | 1.47 | 1.82 | 98.3 | 595 |
| | LewisX Fc4-Fc3 | 833 | 1.47 | 4.03 | 98.8 | 800 |
| | Man005 Fc4-Fc3 | 155 | 0.946 | 5.55 | 102 | 131 |
| 14 September 2010 | Fuc4-1 Fc2-Fc3 | 382 | 1.5 | 1.23E-01 | 98.5 | 389 |
| | L-fucose Fc2-Fc3 | 1814 | 1.43 | 0.293 | 99.1 | 1830 |
| | Fuc1-14 Fc2-Fc3 | 595 | 1.52 | 0.0718 | 98.3 | 608 |
| | Fuc5-1 Fc2-Fc3 | 845 | 1.47 | 2.11E-01 | 98.6 | 859 |
| | Fuc5-3 Fc2-Fc3 | 742 | 1.47 | 0.182 | 98.2 | 759 |
| | Fuc5-4 Fc2-Fc3 | 776 | 1.5 | 0.184 | 97.9 | 797 |
| | Fuc4-1 Fc4-Fc3 | 438 | 1.62 | 4.81E-04 | 97.8 | 450 |
| | L-fucose Fc4-Fc3 | 2043 | 1.56 | 0.0118 | 97.9 | 2100 |
| | Fuc1-14 Fc4-Fc3 | 672 | 1.65 | 7.35E-03 | 96.9 | 698 |
| | Fuc5-1 Fc4-Fc3 | 954 | 1.58 | 0.0198 | 96.6 | 997 |
| | Fuc5-3 Fc4-Fc3 | 828 | 1.59 | 0.0208 | 96 | 872 |
| | Fuc5-4 Fc4-Fc3 | 869 | 1.61 | 1.69E-02 | 95.4 | 922 |

| Date | Compound | IC$_{50}$, µM | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| 10 March 2010 | Fuc 1-02 | 596 | 1.49 | 0.0138 | 99.3 | 602 |
| | Fuc 1-09 | 757 | 1.52 | 0.565 | 98.9 | 762 |
| | Fuc 1-14 | 592 | 1.45 | 0.0139 | 98.6 | 604 |
| | Fuc 1-19 | 554 | 1.45 | 0.0178 | 98.4 | 566 |
| | Fuc 1-28 | 657 | 1.4 | 0.0072 | 99.1 | 666 |
| | Fuc 1-30 | 609 | 1.41 | 0.042 | 99.4 | 614 |
| | Fuc 1-31 | 490 | 1.36 | 0.00616 | 99 | 497 |
| | Fuc 1-32 | 545 | 1.42 | 0.0138 | 99 | 553 |
| | Lewis X | 750 | 1.35 | 0.00901 | 99.2 | 759 |
| | Man 005 | 114 | 1 | 0.276 | 97.9 | 118 |
| | Man 002 | 774 | 1.37 | 0.0471 | 95.4 | 830 |
| | Fuc 4-1 | 347 | 1.53 | 0.00481 | 97.6 | 358 |
| | Man-030 | 291 | 1.54 | 0.000161 | 97.8 | 300 |
| 19 July 2010 | Man-035 | 324 | 1.6085 | -1.2626 | 98.736 | 334.51 |
| | Man-036 | 441 | 1.5596 | -1.5425 | 98.456 | 459.07 |
| | Man-002 | 830 | 1.4248 | -1.2203 | 98.779 | 858.69 |
| | Man-005 | 101 | 1.0629 | 0.013663 | 100.01 | 101.09 |
| | mannose | 79 | 1.62 | 12 | 94.9 | 71.7 |
| | Man-030 | 325 | 1.561 | -1.3005 | 98.699 | 335.6 |
| | Man-031 | 367 | 1.5355 | -1.1181 | 98.881 | 378.07 |
| | Man-032 | 310 | 1.545 | -1.432 | 98.567 | 321.44 |
| | Man-033 | 290 | 1.5416 | -1.4121 | 98.587 | 301.23 |
| 20 July 2010 | man-040 | 398 | 1.5495 | -1.2961 | 98.703 | 411.54 |
| | man-041 | 335 | 1.5618 | -1.3352 | 98.664 | 347.16 |
| | man-042 | 5283 | 2.24 | 43.8 | 92.8 | 2230 |
| | man-001 | 1054 | 1.4834 | -0.93711 | 99.062 | 1081.5 |
| | RR122 | 159 | 1.5 | 0.179 | 90.5 | 183 |
| | man-029 | 797 | 1.34 | 0.00537 | 96.9 | 836 |
| | man-037 | 810 | 1.6584 | -1.2391 | 98.76 | 834.48 |
| | man-038 | 296 | 1.6134 | -1.3499 | 98.649 | 305.77 |
| | man-039 | 356 | 1.5434 | -1.2866 | 98.712 | 368.49 |
| 15 March 2011 | ZL2(1) | 3275 | 1.21 | 20.1 | 106 | 1950 |
| | psDi | 1056 | 1.1 | -16.8 | 105 | 1260 |
| | mannose | 2982 | 1.11 | -14.1 | 104 | 3480 |
| | BB112 | 3272 | 1.11 | -13.8 | 105 | 3740 |
| | BB133 | 1179 | 1.23 | -5.46 | 106 | 1170 |
| | ZL1(D) | 4540 | 1.05 | 7.26 | 106 | 3510 |
| 16 March 2011 | BB112 | 3376 | 1.1 | -14 | 101 | 4150 |
| | ZL1(D) | 4587 | 0.982 | -2.67 | 101 | 4740 |
| | ZL2(L) | #NOMBRE! | 1.49 | 57 | 99.8 | 1010 |
| | Man1-2Man | 915 | 1.09 | -14.4 | 99.7 | 1160 |
| | Man1-3Man | 2341 | 1.07 | -15.8 | 101 | 2970 |
| | psDi | 1012 | 1.07 | -16.1 | 101 | 1290 |
| 17 March 2011 | Man1-3Man-DC | 2245 | 0.989 | -28.9 | 101 | 3490 |
| | ZL2(L)-DC | 8840 | 1.53 | 48.2 | 99 | 1020 |
| | Fucose-DC | 3157 | 1.06 | -14 | 98.9 | 4070 |
| | BB133-DC | 862 | 1.15 | -4.64 | 98.6 | 954 |

| Date | Compound | IC$_{50}$, μM | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| 15 June 2011 | dcsign-mannose-FC2-1 | 3050 | 1.28 | -7.46 | 100 | 3400 |
| | dcsign-man12man1-FC2-1 | 901 | 1.23 | -7.4 | 99.9 | 1010 |
| | dcsign-psDi-FC2-1 | 1020 | 1.19 | -9.46 | 100 | 1180 |
| | dcsign-psTri-FC2-1 | 246 | 1.01 | -3.31 | 102 | 252 |
| | dcsign-man12man2-FC2-1 | 869 | 1.22 | -7.35 | 99.9 | 974 |
| | dcsign-mannose-FC3-1 | 3722 | 1.29 | -7.01 | 100 | 4120 |
| | dcsign-man12man1-FC3-1 | 1119 | 1.28 | -6.89 | 99.9 | 1240 |
| | dcsign-psDi-FC3-1 | 1289 | 1.22 | -9.16 | 100 | 1480 |
| | dcsign-psTri-FC3-1 | 319 | 1.01 | -3.74 | 101 | 336 |
| | dcsign-man12man2-FC3-1 | 1096 | 1.29 | -6.78 | 100 | 1210 |
| | Lg mannose Fc2 | 10300 | 0.916 | 0 | 100 | 10300 |
| | Lg man12man Fc2 | 1630 | 0.988 | 0 | 99.7 | 1640 |
| | Lg psDi Fc2 | 6827 | 1.03E+00 | 0 | 99.6 | 6880 |
| | Lg psTri Fc2 | 5050 | 0.881 | 0 | 100 | 5050 |
| | Lg mannose Fc3 | 4107 | 1.1 | 0 | 99.7 | 4130 |
| | Lg man12man Fc3 | 713 | 1.19 | 0 | 98.8 | 728 |
| | Lg psDi Fc3 | 2760 | 1.21 | 0 | 98.1 | 2850 |
| | Lg psTri Fc3 | 1669 | 1.09 | 0 | 99 | 1700 |
| 19 July 2011 | Man030 Fc2-1 1 | 261 | 1.28 | -3.82 | 101 | 272 |
| | Man030 Fc2-1 2 | 258 | 1.3 | -3.7 | 98.9 | 277 |
| | Man043 Fc2-1 | 4288 | 1.05 | -50.1 | 101 | 8150 |
| | Man044 Fc2-1 | 234 | 1.33 | 0.443 | 99.6 | 234 |
| | Man045 Fc2-1 | 300 | 1.36 | 0.539 | 98.9 | 303 |
| | Man046 Fc2-1 | 1080 | 1.23 | 1.43 | 99.7 | 1060 |
| | Man047 Fc2-1 | 610 | 1.22 | -7.37 | 99.2 | 692 |
| | Man048 Fc2-1 | 1047 | 1.23 | -3.07 | 99.4 | 1110 |
| | Man049 Fc2-1 | 24 | 1.43 | 1.75 | 98.9 | 23.8 |
| | Man030 Fc3-1 1 | 352 | 1.38 | -4.04 | 99.8 | 374 |
| | Man030 Fc3-1 2 | 351 | 1.36 | -4.17 | 98.5 | 381 |
| | Man043 Fc3-1 | 5850 | 1.03 | -71.6 | 101 | 13600 |
| | Man044 Fc3-1 | 326 | 1.35 | -3.25 | 99.9 | 342 |
| | Man045 Fc3-1 | 418 | 1.35 | -4.4 | 99.7 | 447 |
| | Man046 Fc3-1 | 1527 | 1.1 | -31.1 | 100 | 2370 |
| | Man047 Fc3-1 | 844 | 1.23 | -11.1 | 99.6 | 1000 |
| | Man048 Fc3-1 | 1491 | 1.12 | -23.3 | 99.7 | 2110 |
| | Man049 Fc3-1 | 32 | 1.39 | 0.31 | 98.4 | 32.5 |
| | Man030 Fc4-1 1 | 340 | 1.33 | -5.11 | 102 | 355 |
| | Man030 Fc4-1 2 | 322 | 1.32 | -5.03 | 97.2 | 362 |
| | Man043 Fc4-1 | 5490 | 0.999 | -80.6 | 102 | 13800 |
| | Man044 Fc4-1 | 308 | 1.31 | -4.88 | 101 | 326 |
| | Man045 Fc4-1 | 392 | 1.31 | -5.81 | 100 | 426 |
| | Man046 Fc4-1 | 1395 | 1.09 | -30.5 | 100 | 2160 |
| | Man047 Fc4-1 | 784 | 1.2 | -12.1 | 99.3 | 950 |
| | Man048 Fc4-1 | 1350 | 1.1 | -25.5 | 99 | 2000 |
| | Man049 Fc4-1 | 30 | 1.35 | -0.691 | 97.6 | 31.7 |
| 21 July 2011 | Fuc4-1 Fc2-1 1 | 367 | 1.13 | -25.2 | 99.6 | 530 |
| | Fuc4-1 Fc2-1 2 | 372 | 1.13 | -25.5 | 100 | 535 |
| | Fuc4-1 azide Fc2-1 | 227 | 1.28 | -3.05 | 99.3 | 240 |
| | Lewis X Fc2-1 | 628 | 1.15 | -11.7 | 100 | 754 |
| | Fuc4-1 Fc3-1 1 | 527 | 1.29 | -6.29 | 99.6 | 581 |
| | Fuc4-1 Fc3-1 2 | 529 | 1.29 | -6.2 | 99.4 | 585 |
| | Fuc4-1 azide Fc3-1 | 311 | 1.36 | -3.59 | 99.3 | 331 |
| | Lewis X Fc3-1 | 886 | 1.18 | -13 | 99.9 | 1080 |
| | Fuc4-1 Fc4-1 1 | 529 | 1.05 | -37 | 101 | 879 |
| | Fuc4-1 Fc4-1 2 | 524 | 1.04 | -38.3 | 99.8 | 908 |
| | Fuc4-1 azide Fc4-1 | 304 | 1.32 | -4.41 | 99.5 | 327 |
| | Lewis X Fc4-1 | 856 | 1.15 | -14.5 | 99.9 | 1070 |

| **Date** | **Compound** | **IC$_{50}$, μM** | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| 1 February 2011 | poly-020 | 14 | 0.956 | 6.98 | 100 | 12 |
| | psDi | 943 | 1.11 | -10.5 | 101 | 1100 |
| | Man030 | 226 | 1.22 | -1.22 | 100 | 231 |
| | poly008 | 113 | 1.11 | -1.75 | 100 | 117 |
| | poly009 | 39 | 1.04 | 2.12 | 100 | 37.6 |
| | poly015 | 36 | 0.867 | 3.54 | 100 | 33.5 |
| | poly016 | 38 | 0.74 | 0.774 | 101 | 36.6 |
| | poly018 | 8 | 1.5 | 6.53 | 99.4 | 6.89 |
| | poly019 | 4 | 2.12 | 7.45 | 97.3 | 3.37 |
| 2 February 2011 | psTri | 81 | 0.637 | -7.87 | 120 | 60.4 |
| | poly005 | 753 | 0.999 | 5.37 | 101 | 659 |
| | poly006 | 800 | 1.14 | -2.53 | 100 | 835 |
| | poly012 | 128 | 0.903 | 6.2 | 101 | 108 |
| | poly013 | 114 | 0.924 | 3.83 | 100 | 105 |
| | poly014 | 67 | 0.873 | 4.76 | 101 | 58.8 |
| | poly017 | 112 | 1.12 | -0.267 | 99.8 | 113 |
| | poly021 | 51 | 1.1 | 3.24 | 99.6 | 48.5 |
| | poly022 | 32 | 0.986 | 2.83 | 100 | 30.6 |
| 3 February 2011 | poly002 | 55 | 0.916 | -3.19 | 102 | 56.5 |
| | man a1-2man | 869 | 1.04 | -15.8 | 101 | 1110 |
| | psDi | 1023 | 1 | -18.8 | 101 | 1380 |
| | psTri | 113 | 0.776 | -3.68 | 105 | 110 |
| | Man030 | 254 | 1.19 | -2.47 | 100 | 265 |
| | poly018 | 10 | 1.5 | 4.45 | 99.9 | 9.19 |
| | poly019 | 4 | 2.14 | 1.37 | 98.1 | 4.41 |
| | poly005 | 781 | 0.867 | -9.04 | 102 | 904 |
| 7 February 2011 | Poly001 | 164 | 0.977 | -13 | 99.9 | 208 |
| | Poly008 | 136 | 0.987 | -12.2 | 99.9 | 170 |
| | psDi | 896 | 0.994 | -17.7 | 98.6 | 1250 |
| | Man1-2Man | 905 | 1.04 | -15.6 | 98.5 | 1210 |
| | psTri | 68 | 0.802 | -4.7 | 102 | 72.2 |
| | Poly013 | 101 | 0.797 | -2.44 | 98.8 | 110 |
| | Mannose | 2825 | 1.01 | -18.8 | 97.7 | 4060 |
| 24 January 2012 | psDi 1 | 994 | 1.06 | -18.2 | 99.3 | 1350 |
| | psDi 2 | 918 | 1.07 | -17.1 | 94.4 | 1350 |
| | NV233 | 1235 | 0.981 | -29.3 | 98 | 2060 |
| | NV243 | 481 | 1.04 | -8.96 | 96.8 | 601 |
| | NV253 | 703 | 0.927 | -21.2 | 96.6 | 1110 |
| | NV270 | 1366 | 1.03 | -27.8 | 96.1 | 2270 |
| | Man030 | 298 | 1.31 | -4.37 | 94.3 | 348 |

**All acquired IC$_{50}$ values for all compounds during the thesis**

| Date | Compound | IC$_{50}$, µM | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| 20-22 February 2012 | L-fucose Fc2 | 11760 | 1.09 | -22.8 | 100 | 16600 |
| | BB01 | 211 | 1.21 | -7.01 | 100 | 235 |
| | MA05 Fc2 | 378 | 1.28 | -7.36 | 101 | 414 |
| | BB02 Fc2 | 96 | 1.19 | -3.29 | 101 | 100 |
| | MA07 Fc2-1 | 391 | 1.15 | -15.3 | 101 | 485 |
| | BB014 Fc2-1 | 17 | 0.997 | -1.34 | 102 | 17 |
| | RR390b Fc2-1 | 5 | 0.854 | -0.152 | 109 | 3.94 |
| | BB01 Fc2-1 2 | 215 | 1.21 | -7.05 | 100 | 240 |
| | L fucose Fc3-1 | 11743 | 1.09 | -9.15 | 100 | 13700 |
| | BB01 Fc3-1 1 | 211 | 1.15 | -8.38 | 100 | 241 |
| | MA05 Fc3-1 | 373 | 1.23 | -8.16 | 100 | 422 |
| | BB02 Fc3-1 | 97 | 1.16 | -4.17 | 101 | 102 |
| | MA07 Fc3-1 | 393 | 1.1 | -16.6 | 101 | 501 |
| | BB014 Fc3-1 | 17 | 0.96 | -1.98 | 102 | 17.1 |
| | RR390b Fc3-1 | 5 | 0.818 | -0.468 | 110 | 3.81 |
| | BB01 Fc3-1 2 | 217 | 1.17 | -8.33 | 100 | 248 |
| | L-fucose Fc2-1 | 7348 | 1.06 | -29.9 | 99.7 | 11500 |
| | BB013 Fc2-1 1 | 40 | 1.08 | -1.93 | 100 | 41.4 |
| | RCF-26 Fc2-1 | 3 | 1.15 | 1.61 | 102 | 2.89 |
| | BB01 Fc2-1 | 206 | 1.24 | -5.93 | 100 | 225 |
| | D-mannose Fc2-1 | 3220 | 1.22 | -11.2 | 100 | 3800 |
| | BB015 Fc2-1 | 24 | 0.679 | -1.15 | 106 | 20.6 |
| | BB016 Fc2-1 | 12 | 0.758 | -1.12 | 106 | 10.8 |
| | BB013 Fc2-1 2 | 41 | 1.1 | -1.86 | 99.9 | 42.3 |
| | L-fucose Fc3-1 | 7465 | 1.05 | -30.70 | 99.90 | 11800 |
| | BB013 Fc3-1 1 | 41 | 1.07 | -2.73 | 100.00 | 42.8 |
| | RCF-26 Fc3-1 | 3 | 1.13 | 1.50 | 102.00 | 2.9 |
| | BB01 Fc3-1 | 210 | 1.23 | -6.65 | 100.00 | 232 |
| | D-mannose Fc3-1 | 3273 | 1.20 | -11.90 | 100.00 | 3910 |
| | BB015 Fc3-1 | 24 | 0.67 | -2.39 | 106.00 | 21.6 |
| | BB016 Fc3-1 | 12 | 0.75 | -1.95 | 106.00 | 11.1 |
| | BB013 Fc3-1 2 | 41 | 1.08 | -2.58 | 99.90 | 43.5 |
| | BO Fc2-1 | 985 | 1.36 | -2.41 | 99.3 | 1030 |
| | BO Fc3-1 | 998 | 1.37 | -3.66 | 99.4 | 1060 |
| 24 April 2012 | PM008 FC2-1 | 219 | 1.16 | -15 | 99.6 | 277 |
| | PM018 FC2-1 | 14 | 1.21 | -0.755 | 101 | 14.2 |
| | PM019 FC2-1 | 7 | 1.72 | 0.444 | 98.9 | 6.99 |
| | PM023 FC2-1 | 8 | 2.02 | -0.629 | 98.6 | 8.43 |
| | PM024 FC2-1 | 8 | 2.04 | 0.107 | 98.8 | 7.94 |
| | PM025 FC2-1 | 6 | 1.86 | 0.0434 | 97.7 | 6.29 |
| | PM026 FC2-1 | 26 | 2.13 | -1.14 | 98.5 | 26.7 |
| | Man30 FC2-1 | 344 | 1.26 | -4.9 | 99.1 | 376 |
| | PM008 FC3-1 | 219 | 1.15 | -15.5 | 99.7 | 279 |
| | PM018 FC3-1 | 14 | 1.2 | -0.811 | 101 | 14.2 |
| | PM019 FC3-1 | 7 | 1.71 | 0.356 | 99 | 7 |
| | PM023 FC3-1 | 8 | 2.01 | -0.564 | 98.6 | 8.41 |
| | PM024 FC3-1 | 8 | 2.02 | 0.0122 | 98.9 | 7.95 |
| | PM025 FC3-1 | 6 | 1.83 | -0.065 | 97.8 | 6.31 |
| | PM026 FC3-1 | 26 | 2.13 | -0.985 | 98.5 | 26.7 |
| | man30 FC3-1 | 344 | 1.25 | -5.15 | 99.1 | 378 |
| | PM008 FC4-1 | 213 | 1.11 | -14.4 | 100 | 267 |
| | PM018 FC4-1 | 13 | 1.17 | -0.453 | 101 | 13.3 |
| | PM019 FC4-1 | 7 | 1.64 | 0.491 | 99.4 | 6.71 |
| | PM023 FC4-1 | 8 | 1.95 | -0.51 | 98.4 | 8.11 |
| | PM024 FC4-1 | 8 | 2.05 | 0.249 | 96.8 | 7.93 |
| | PM025 FC4-1 | 6 | 1.9 | 0.307 | 93.4 | 6.44 |
| | PM026 FC4-1 | 25 | 2.17 | -0.823 | 96.1 | 26.6 |
| | man30 FC4-1 | 324 | 1.2 | -5.26 | 97.9 | 365 |

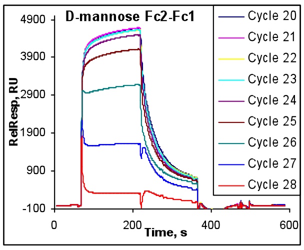
| Date | Compound | IC$_{50}$, µM | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| 25 April 2012 | PM008 Fc2-1 | 217 | 1.18 | -16.1 | 97 | 290 |
| | MAN30 Fc2-1 | 259 | 1.41 | -3.08 | 99.6 | 272 |
| | PSDi Fc2-1 | 1100 | 1.05 | -22.6 | 101 | 1540 |
| | PM008 Fc3-1 | 218 | 1.17 | -15 | 96.8 | 289 |
| | MAN30 Fc3-1 | 258 | 1.4 | -2.89 | 99.4 | 271 |
| | PSDi Fc3-1 | 1076 | 1.05 | -21.4 | 100 | 1510 |
| | PM008 Fc4-1 | 194 | 1.11 | -14.1 | 94.5 | 270 |
| | man30 fc 4-1 | 256 | 1.45 | -2.32 | 99.4 | 266 |
| | PSDi Fc4-1 | 1037 | 1.14 | -14.7 | 100 | 1300 |
| 27 April 2012 | FC4-1 man30 | 134 | 1.28 | -2.29 | 99.6 | 140 |
| | FC4-1 PM026 | 8 | 2.54 | -1.33 | 97.3 | 7.83 |
| | FC3-1 man30 | 137 | 1.25 | -2.41 | 99.1 | 144 |
| | FC3-1 PM026 | 8 | 2.64 | 0.0646 | 96.5 | 7.99 |
| | FC2-1 Man030 | 139 | 1.26 | -2.43 | 99.6 | 145 |
| | FC2-1 PM026 | 8 | 2.72 | 0.641 | 96.7 | 7.98 |
| 2 May 2012 | 9a Fc4-1 | 1504 | 1.04 | -25.4 | 101 | 2190 |
| | Dendri b-Ala fuc Fc4-1 | 327 | 0.918 | -9.28 | 101 | 385 |
| | Dendri Fuc4-1 Fc4-1 | 41 | 1.3 | 0.0829 | 99.3 | 41 |
| | L-fucose Fc4-1 | 2281 | 1.12 | -18.2 | 100 | 3010 |
| | LewisX Fc4-1 | 1490 | 1.06 | -28.5 | 100 | 2280 |
| | Fuc4-1 azide old Fc4-1 2 | 364 | 1.33 | -4.48 | 99.2 | 393 |
| | LewisX Fc2-1 | 1420 | 1.1 | -25.7 | 100 | 2070 |
| | Fuc4-1 azide old Fc2-1 2 | 344 | 1.38 | -4.43 | 99.1 | 371 |
| | Fuc4-1 azide old Fc3-1 1 | 348 | 1.36 | -3.97 | 99.3 | 372 |
| | Fuc4-1 azide-2 Fc3-1 | 2755 | 1.08 | -41.5 | 99.9 | 4830 |
| | L-fuc b-Ala Fc3-1 | 1660 | 1.07 | -33.5 | 100 | 2680 |
| | 9a Fc3-1 | 1427 | 1.08 | -22.8 | 100 | 2020 |
| | Dendri b-Ala fuc Fc3-1 | 300 | 0.96 | -8.41 | 99.6 | 356 |
| | Dendri Fuc4-1 Fc3-1 | 39 | 1.34 | 0.0995 | 99 | 39.5 |
| | L-fucose Fc3-1 | 2196 | 1.15 | -15.7 | 99.9 | 2790 |
| | LewisX Fc3-1 | 1427 | 1.08 | -29 | 100 | 2180 |
| | Fuc4-1 azide old Fc3-1 2 | 352 | 1.36 | -3.92 | 99.2 | 376 |
| | Fuc4-1 azide old Fc4-1 1 | 362 | 1.34 | -4.02 | 99.4 | 387 |
| | Fuc4-1 azide-2 Fc4-1 | 2910 | 0.88 | -131 | 99.5 | 12700 |
| | L-fuc b-Ala Fc4-1 | 1750 | 1.03 | -36.4 | 101 | 2920 |
| | Fuc4-1 azide old Fc2-1 1 | 342 | 1.37 | -4.57 | 99.5 | 367 |
| | Fuc4-1 azide-2 Fc2-1 | 2656 | 1.21 | -20.8 | 100 | 3540 |
| | L-fuc b-Ala Fc2-1 | 1635 | 1.08 | -34.9 | 100 | 2670 |
| | 9a Fc2-1 | 1404 | 1.1 | -23.4 | 100 | 1990 |
| | Dendri b-Ala fuc Fc2-1 | 291 | 0.97 | -9.24 | 99.5 | 350 |
| | Dendri Fuc4-1 Fc2-1 | 38 | 1.33 | -0.842 | 99 | 39.1 |
| | L-fucose Fc2-1 | 2143 | 1.17 | -16.4 | 99.8 | 2740 |
| 30 April 2012 | Man030 DCSIGN 1 | 319 | 1.4 | -3.94 | 99.1 | 341 |
| | Man030 DCSIGN 2 | 326 | 1.39 | -4.09 | 99.3 | 349 |
| | psDi Lg Fc1 | 1008 | 1.68 | 9.21 | 99 | 904 |
| | Man030 Lg Fc1 | 1725 | 1.37 | -0.516 | 98.4 | 1780 |
| | NV233 Lg Fc1 | 1302 | 1.29 | 1.87 | 99.7 | 1270 |
| | NV243 Lg Fc1 | 2432 | 1.25 | -18.4 | 99.3 | 3160 |
| | NV253 Lg Fc1 | 877 | 1.5 | 3.55 | 101 | 824 |
| | NV270 Lg Fc1 | 2880 | 1.26 | -18.9 | 99.4 | 3750 |
| | psDi Lg Fc2 | 1474 | 1.51 | 20.4 | 99.4 | 1050 |
| | Man030 Lg Fc2 | 2556 | 1.28 | 14.2 | 98.7 | 2010 |
| | NV233 Lg Fc2 | 1933 | 1.22 | 15.3 | 99.7 | 1440 |
| | NV243 Lg Fc2 | 3440 | 1.16 | -4.94 | 99.7 | 3750 |
| | NV253 Lg Fc2 | 1263 | 1.29 | 10.8 | 101 | 1030 |
| | NV270 Lg Fc2 | 4035 | 1.26 | 3.5 | 99.5 | 3840 |

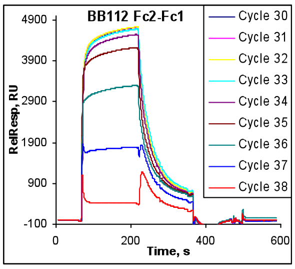
| Date | Compound | IC$_{50}$, μM | $A_2$ | $R_{hi}$ | $R_{lo}$ | $A_1$ |
|---|---|---|---|---|---|---|
| 3 May 2012 | PM026 fc2-1 | 6 | 3.03 | 1.18 | 93.6 | 5.79 |
| | man030 fc2-1 | 348 | 1.35 | -2.68 | 99.5 | 364 |
| | PM026 FC3-1 | 6 | 3 | 0.21 | 93.6 | 5.82 |
| | MAN030 FC3-1 | 348 | 1.34 | -3.55 | 99.9 | 367 |
| | PM026 FC4-1 | 6 | 2.97 | -0.689 | 94.3 | 5.76 |
| | MAN030 FC4-1 | 348 | 1.39 | -3.63 | 99.7 | 368 |
| 30 May 2012 | D-mannose 1 | 3597 | 1.23 | -8.73 | 99.7 | 4120 |
| | L-fucose | 1886 | 1.28 | -4.51 | 99.6 | 2030 |
| | BB178 | 435 | 1.27 | -5.93 | 99.7 | 477 |
| | BB203 | 2000 | 1.19 | -19 | 99.8 | 2630 |
| | BB213 | 3430 | 1.15 | -27.3 | 100 | 5010 |
| | BB207 | 3154 | 1.19 | -22.5 | 100 | 4310 |
| | BO2 | 1108 | 1.33 | -10.1 | 99.6 | 1280 |
| | BO3 | 1170 | 1.26 | -10.9 | 99.9 | 1370 |
| | D-mannose 2 | 3667 | 1.23 | -9.13 | 99.9 | 4210 |
| 8 October 2010 | Man038 Lg | 5788 | 1.0167 | -4.7854 | 95.213 | 6990.9 |
| | Man039 Lg | 8128 | 0.85874 | -5.0693 | 94.93 | 10301 |
| | Man040 Lg | 7625 | 0.90671 | -4.8698 | 95.129 | 9459.6 |
| | Man041 Lg | 6629 | 0.85703 | -4.5188 | 95.48 | 8190.9 |
| 7 October 2010 | Man 002 Lg | 2915 | 1.0322 | 2.9381 | 102.94 | 2600.6 |
| | Man 005 Lg | 2037 | 0.96396 | 3.6266 | 103.63 | 1751.8 |
| | Man 029 Lg | 3083 | 1.1661 | -2.933 | 97.066 | 3410.3 |
| | Man 030 Lg | 7467 | 0.8808 | -3.0726 | 96.927 | 8586.7 |
| | Man 031 Lg | 6410 | 1.0516 | -4.0032 | 95.996 | 7466.5 |
| | Man 032 Lg | 5061 | 1.2031 | -3.4695 | 96.529 | 5681 |
| | Man 033 Lg | 4668 | 1.1664 | -4.5742 | 95.425 | 5463.1 |
| | Man 036 Lg | 4504 | 2.117 | -4.4524 | 95.547 | 4900.7 |
| | Man 035 Lg | 3680 | 1.2217 | -5.1281 | 94.871 | 4354.9 |
| | Man-038 Lg | 5788 | 1.0167 | -4.7854 | 95.213 | 6990.9 |
| | Man-039 Lg | 8128 | 0.85874 | -5.0693 | 94.93 | 10301 |
| | Man-040 Lg | 7625 | 0.90671 | -4.8698 | 95.129 | 9459.6 |
| | Man-041 Lg | 6629 | 0.85703 | -4.5188 | 95.48 | 8190.9 |
| 26 January 2012 | psDi Lg | 4849 | 0.70 | 0.73 | 100.73 | 4651.20 |
| | Man044 Lg | 9244 | 0.82 | 0.42 | 100.42 | 9056.60 |
| | Man045 Lg | 11298 | 0.79 | 0.82 | 100.82 | 10839.00 |
| | Man046 Lg | 3747 | 1.68 | -0.45 | 99.55 | 3786.70 |
| | Man047 Lg | 10457 | 0.75 | 0.78 | 100.78 | 10032.00 |
| | Man048 Lg | 12323 | 0.95 | 0.14 | 100.14 | 12251.00 |
| | Man049 Lg | 2045 | 1.02 | 0.31 | 100.31 | 2020.40 |
| | Man030 Lg | 10111 | 0.75 | 0.79 | 100.79 | 9692.70 |
| | psDi 2 Lg | 4534 | 0.77 | 0.71 | 100.71 | 4369.70 |

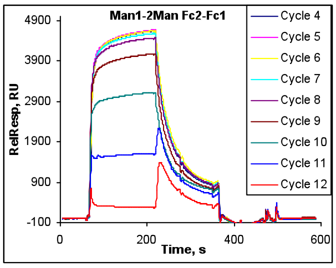
# 20. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by monovalent C-glycosides

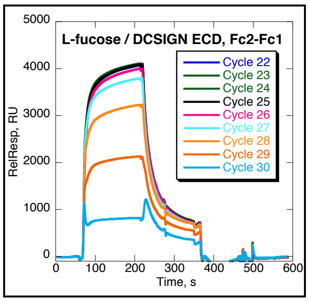
# Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by monovalent C-glycosides

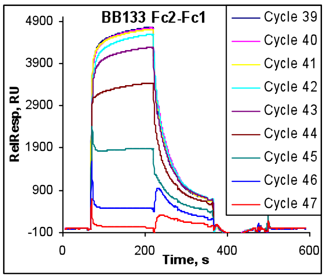
## Sensorgrams:

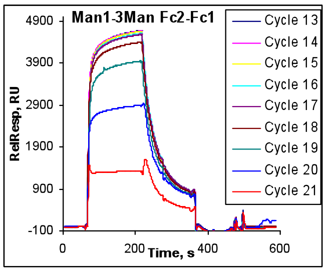


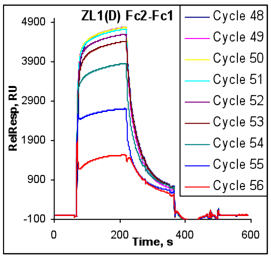


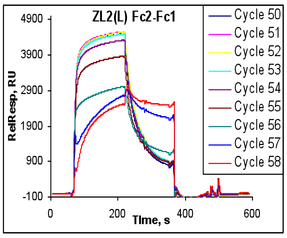


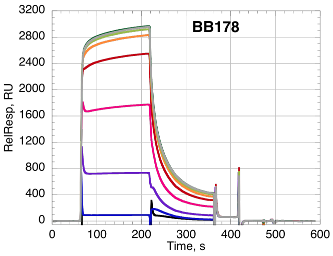


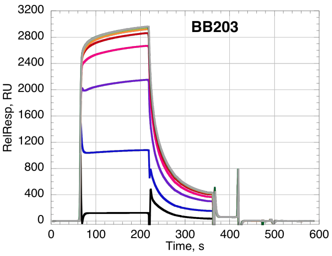


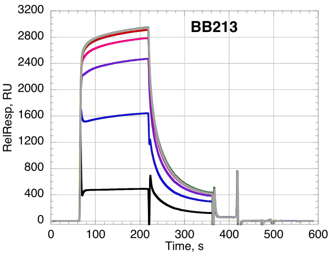


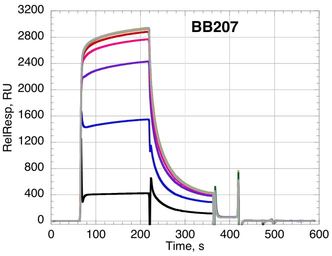


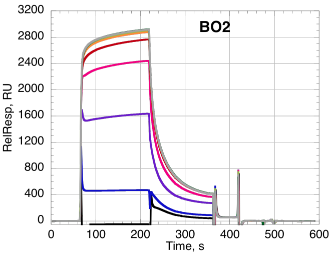


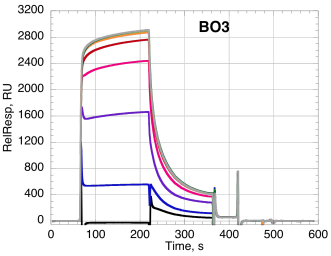


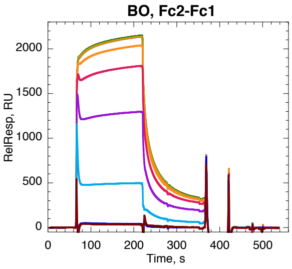




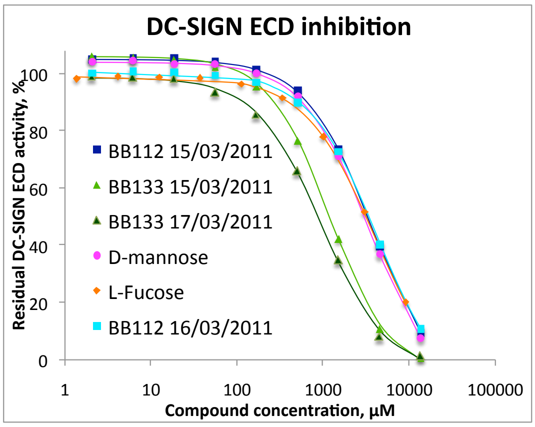


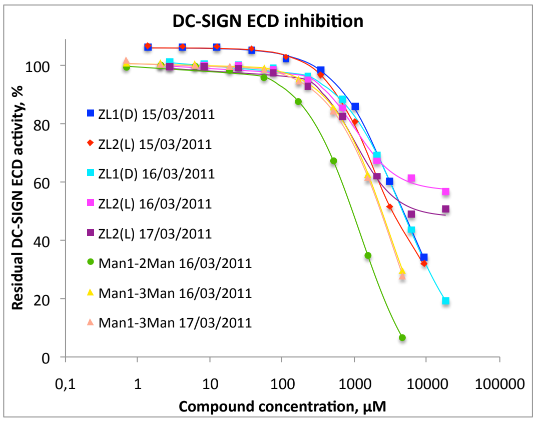


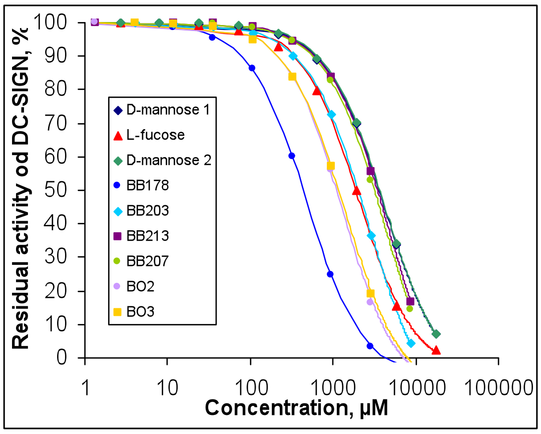


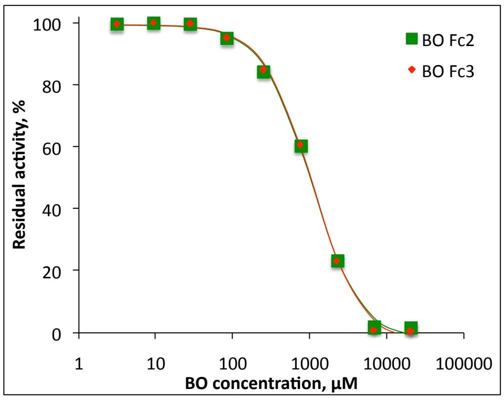


Inhibition curves:





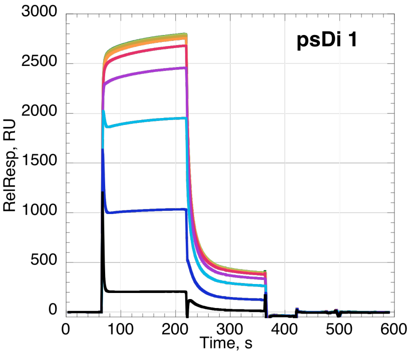


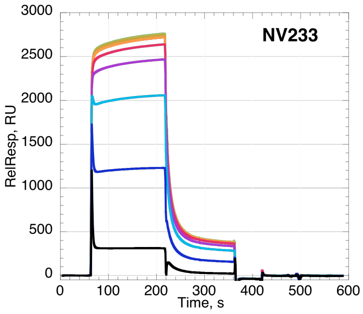


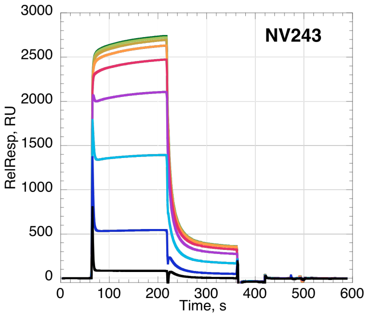
# 21. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by *NV*-type compounds

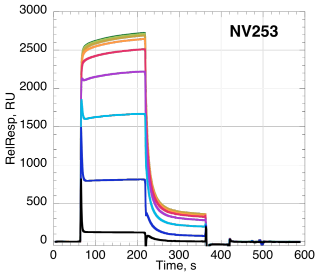
# Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by NV compounds

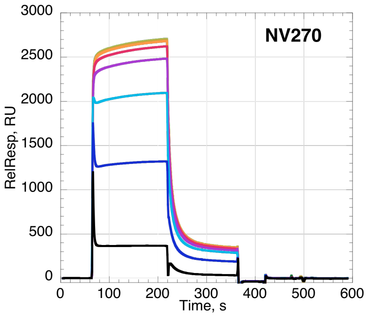
Sensorgrams:

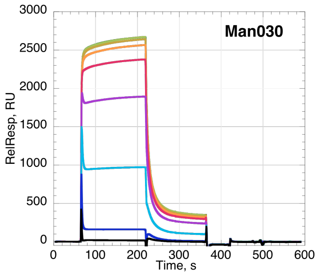


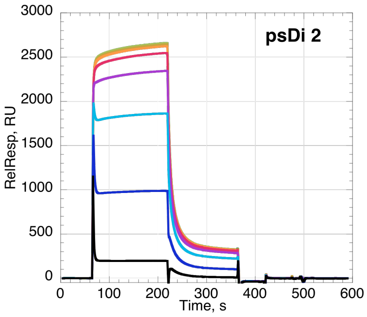




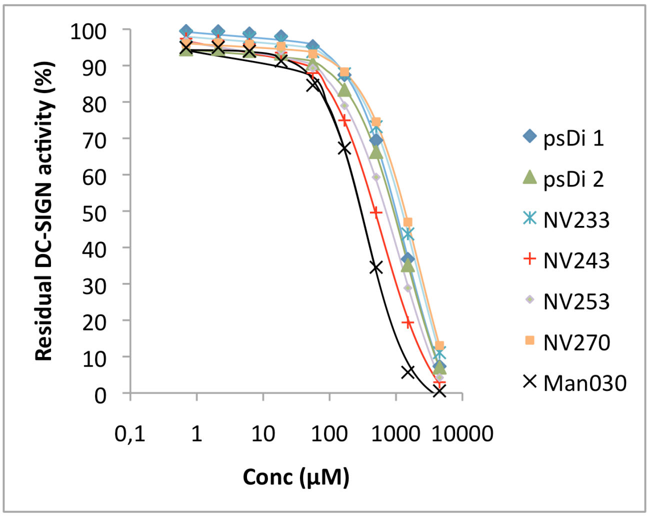








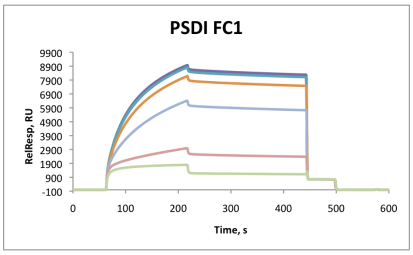
Inhibition curves:

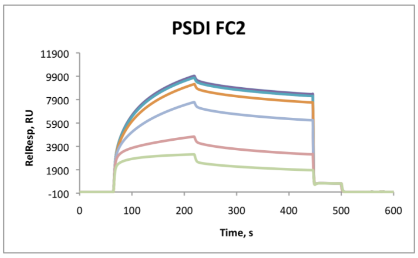


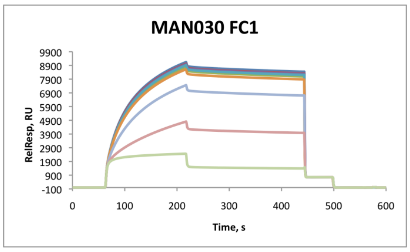
# 22. All of the sensorgrams and inhibition curves for langerin inhibition by *NV*-type compounds

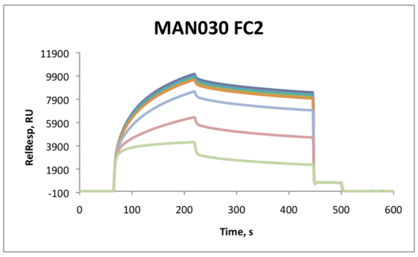
# Sensorgrams and inhibition curves of langerin inhibition by NV-type compounds

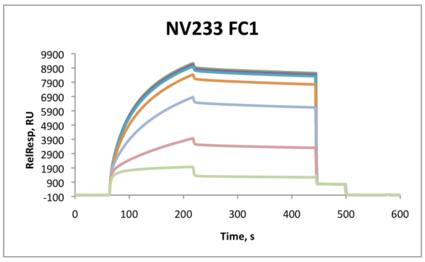
## Sensorgrams:

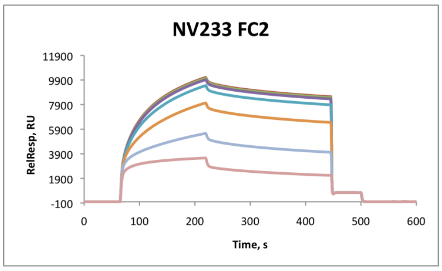


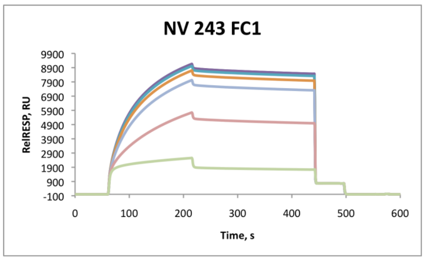


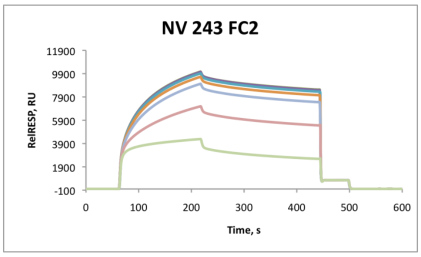


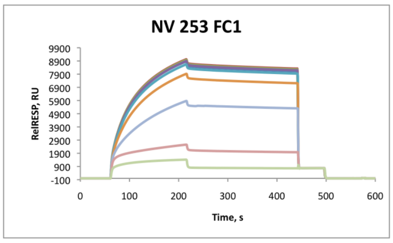


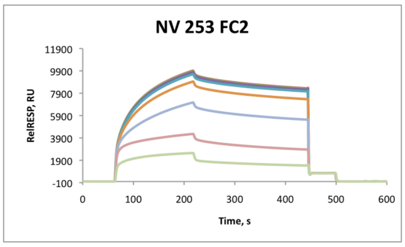


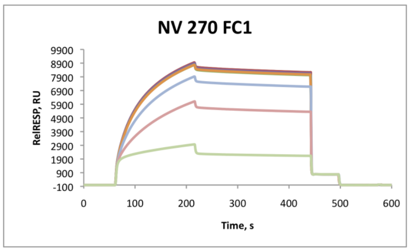


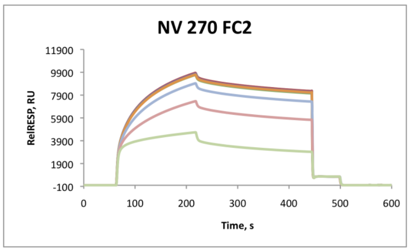




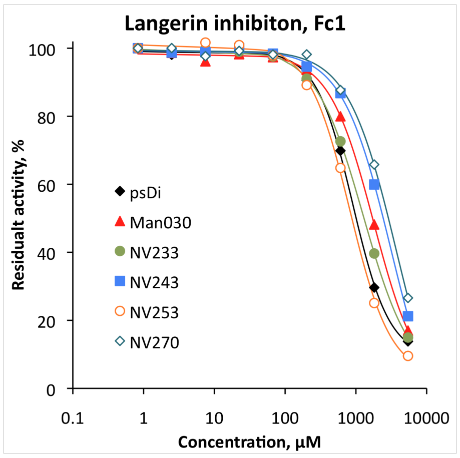


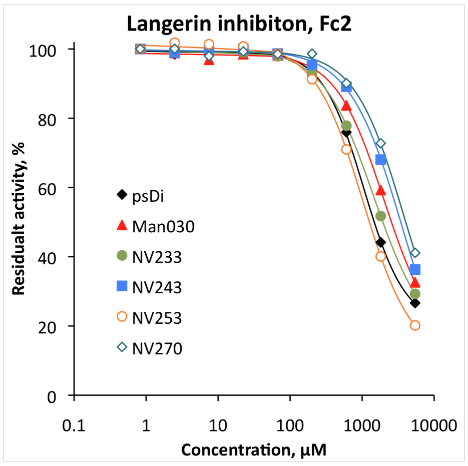






# Inhibition curves:

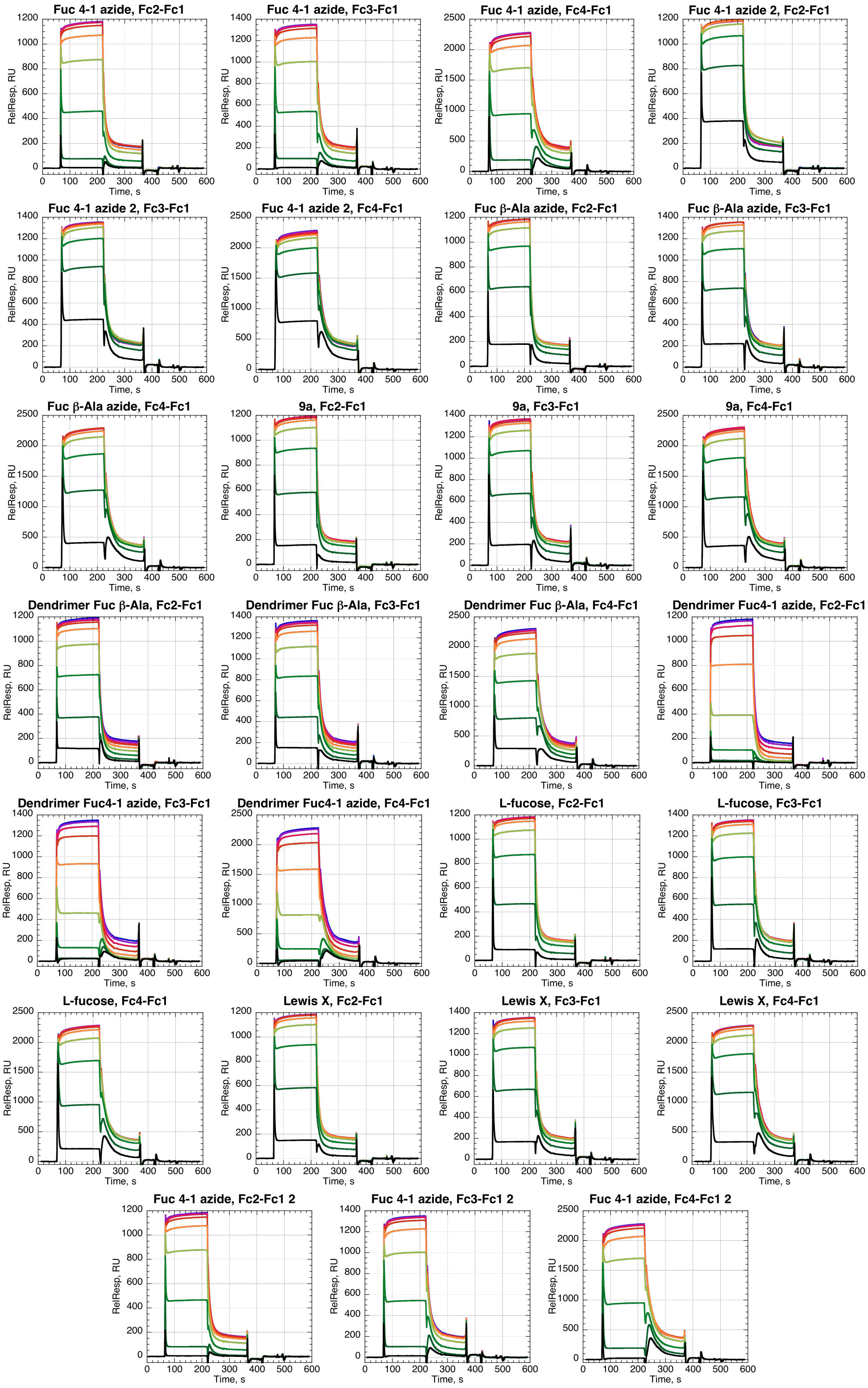




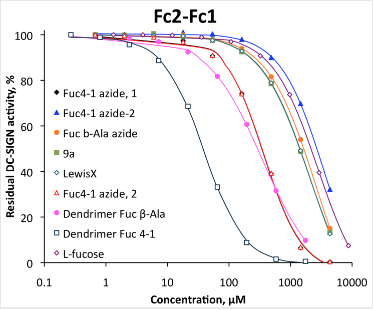
# 23. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by fucose-based monovalent and multivalent compounds

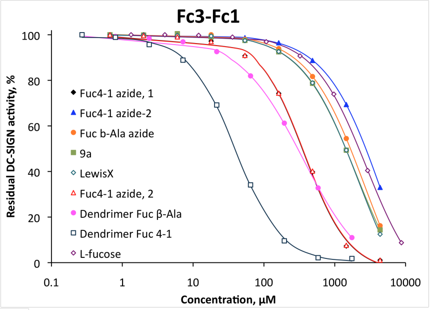
# Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by fucose-based monovalent and multivalent compounds

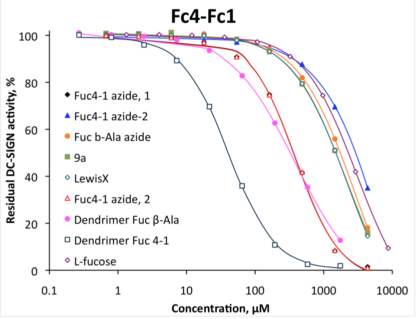
## Sensorgrams:



# Inhibition curves:

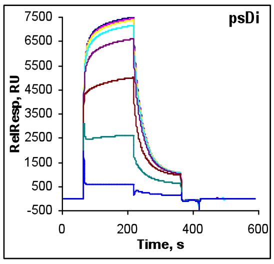


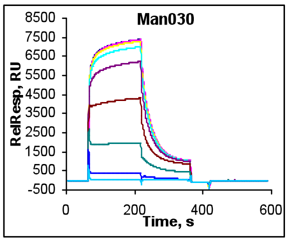


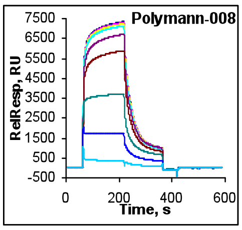


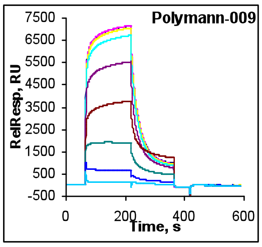
# 24. All of the sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h

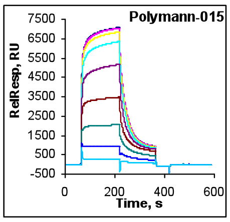
# Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h

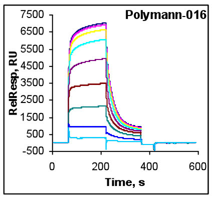


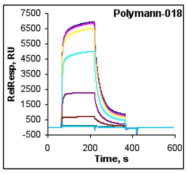


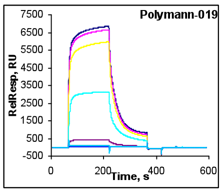


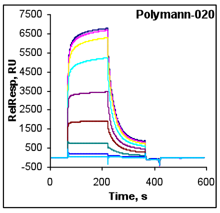


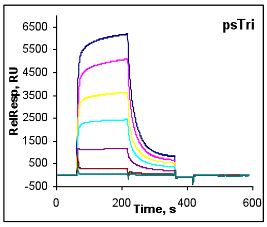


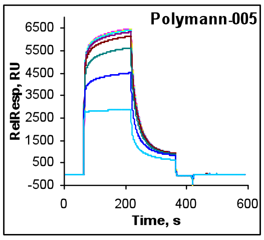


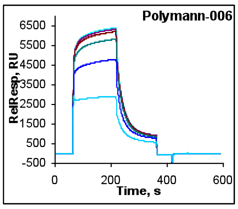


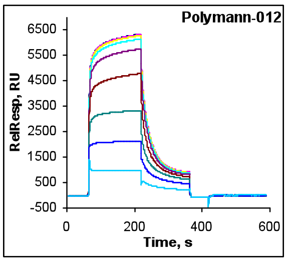


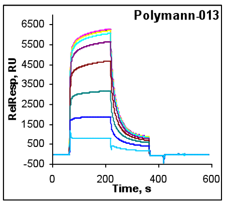


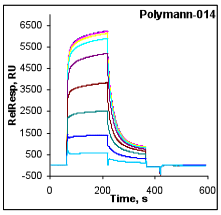


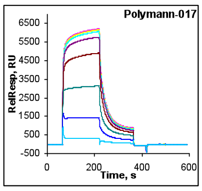


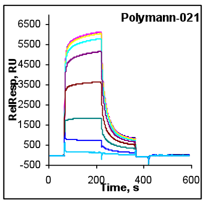


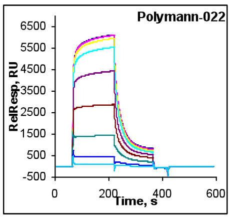


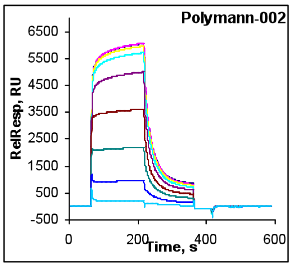


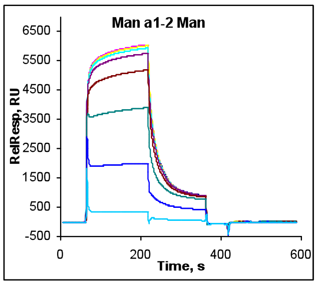


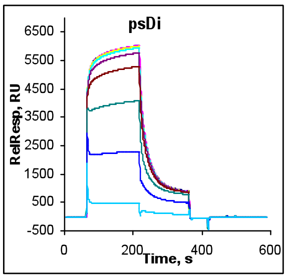


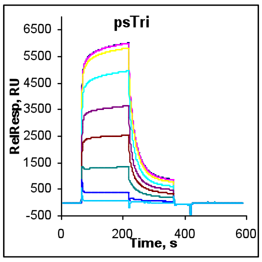


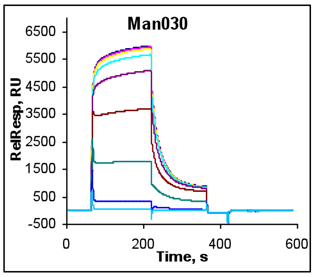


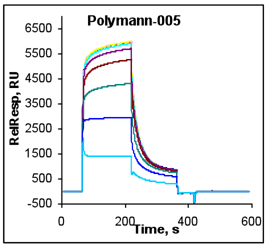


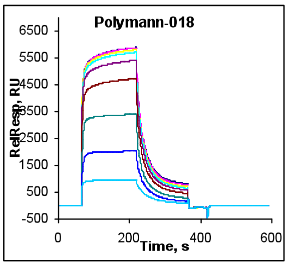


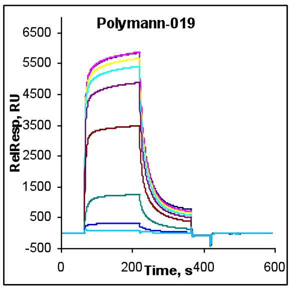


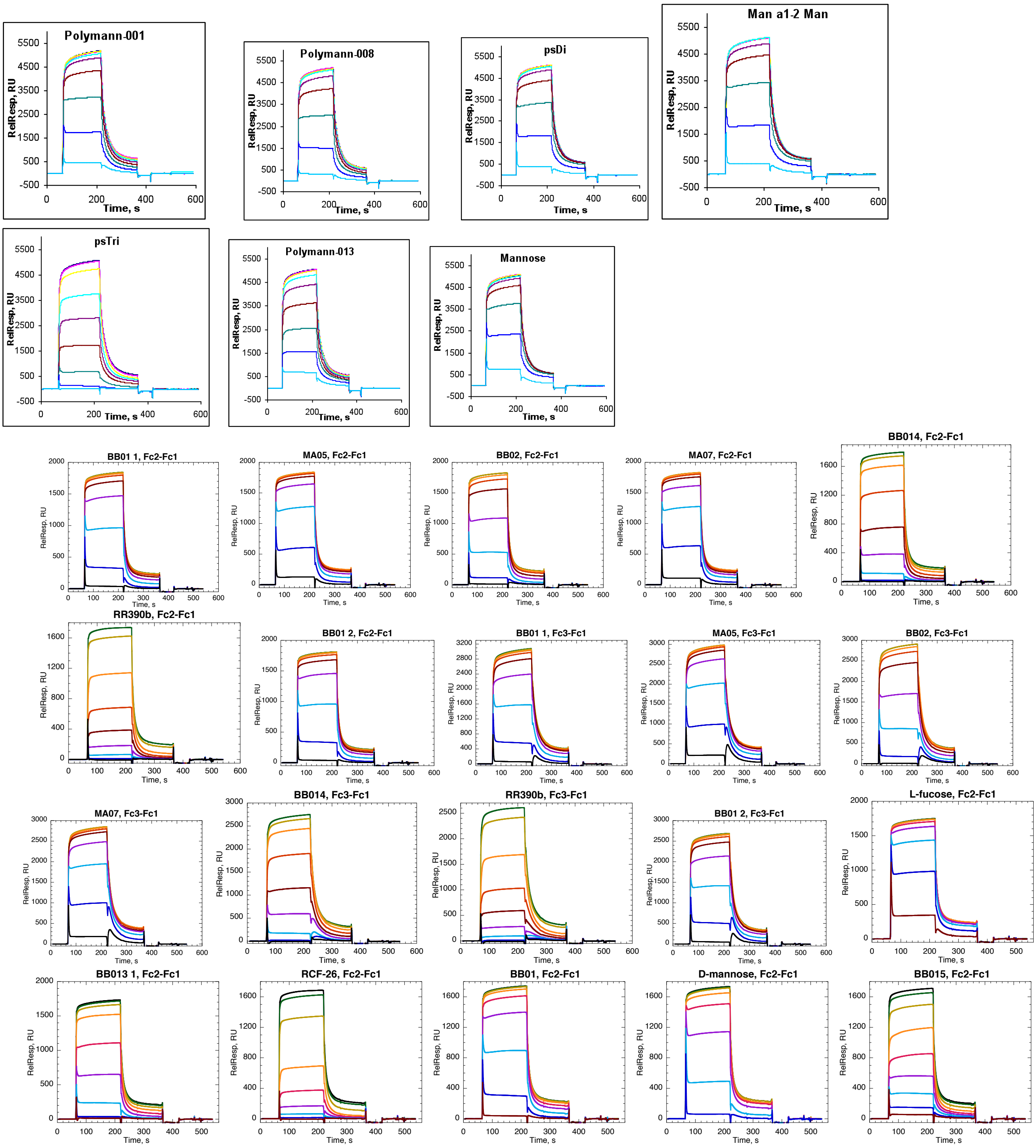




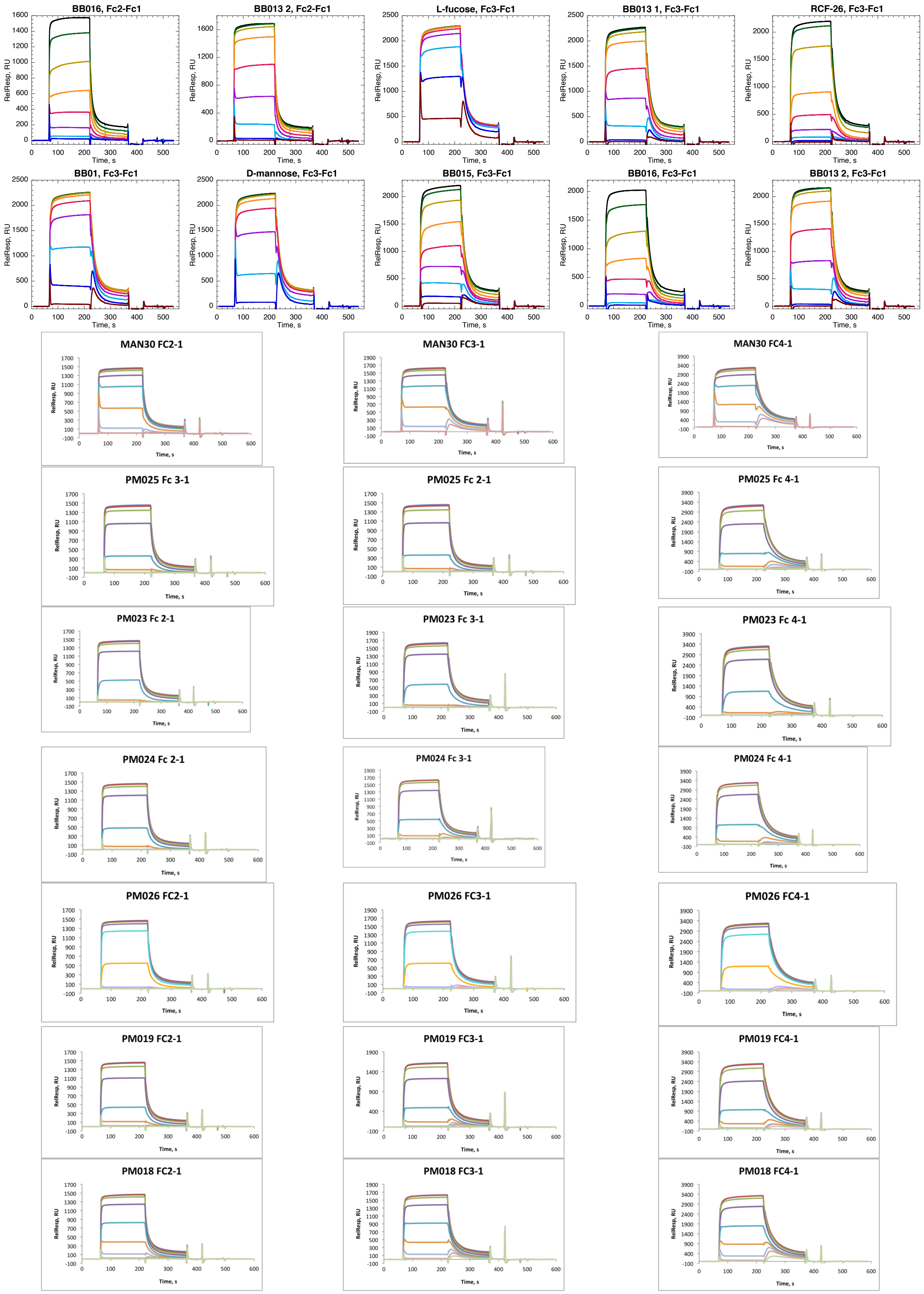




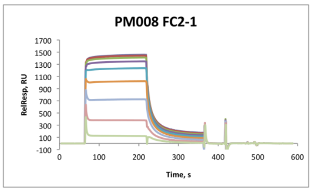


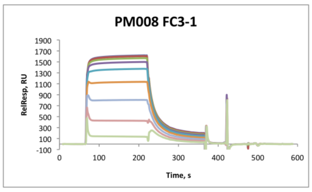


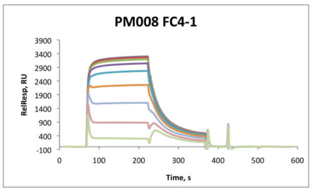
**Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h**

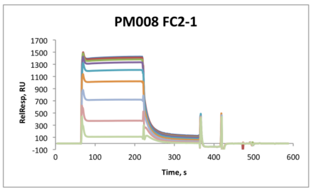


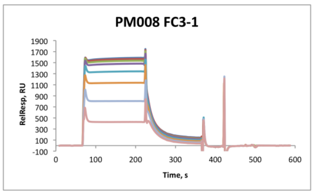
**Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h**

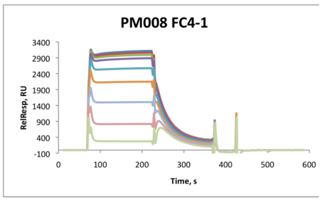


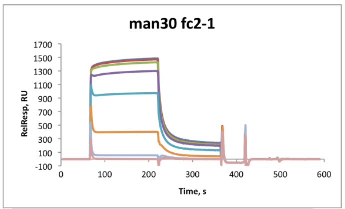


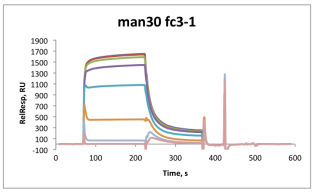


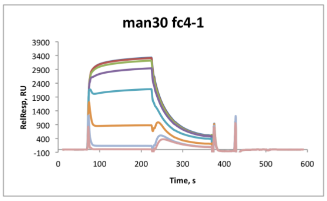


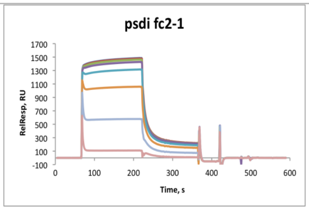


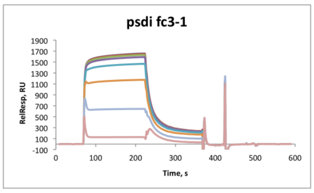


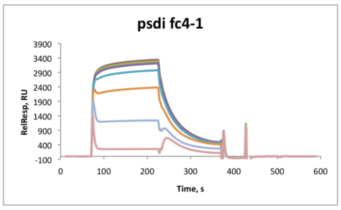


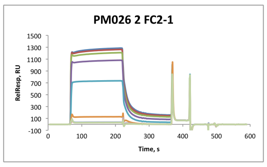


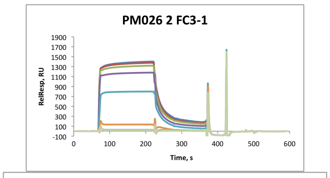


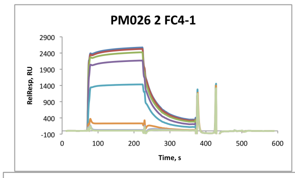


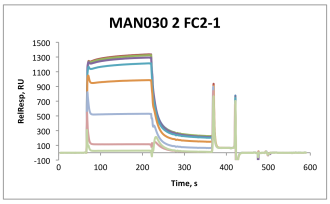


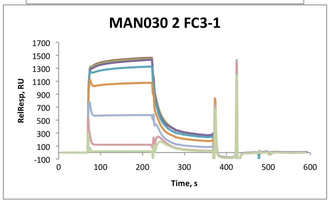


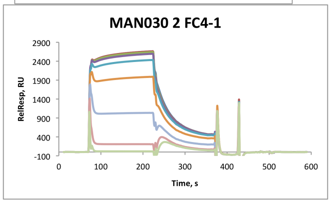






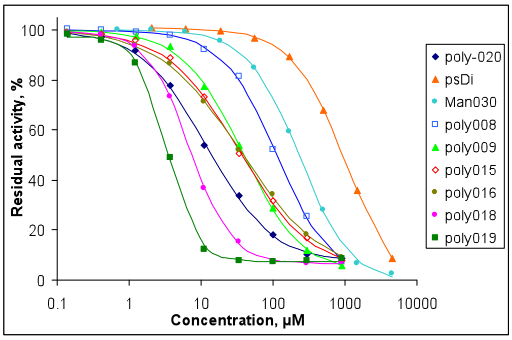


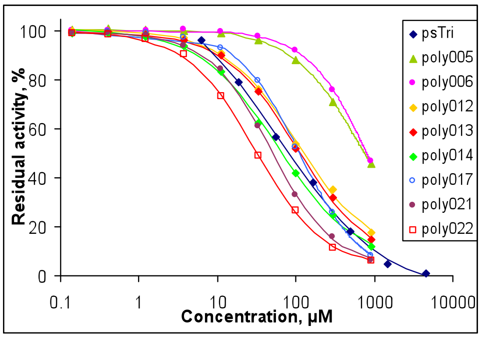


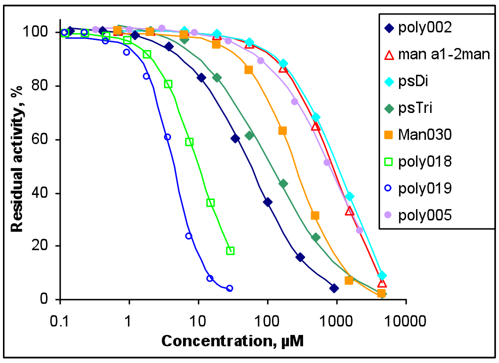


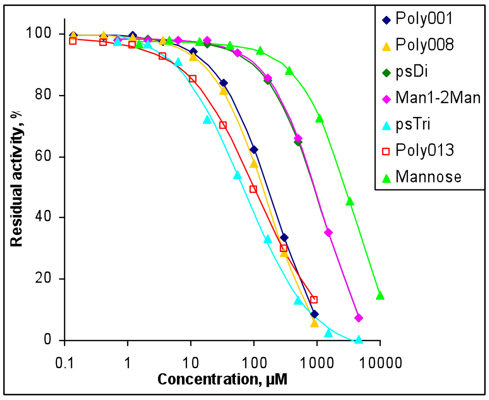
**Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h**

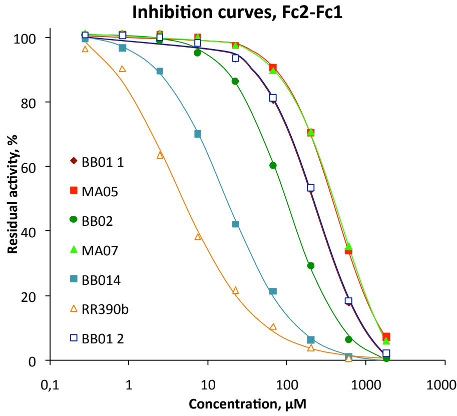
Inhibition curves :

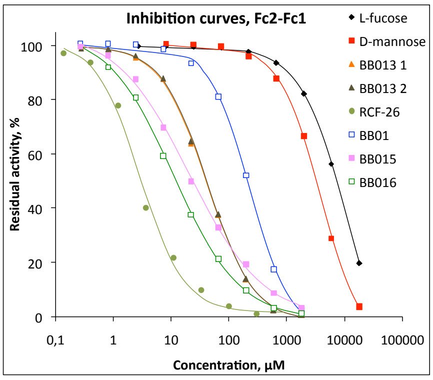


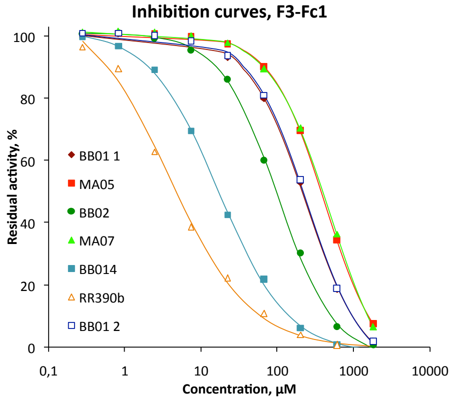


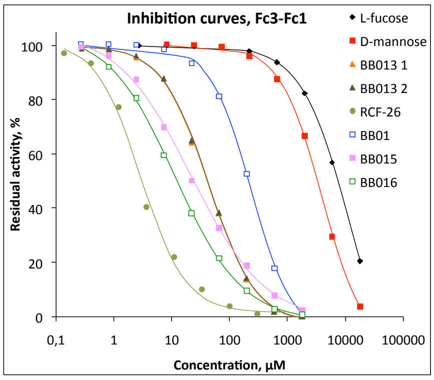




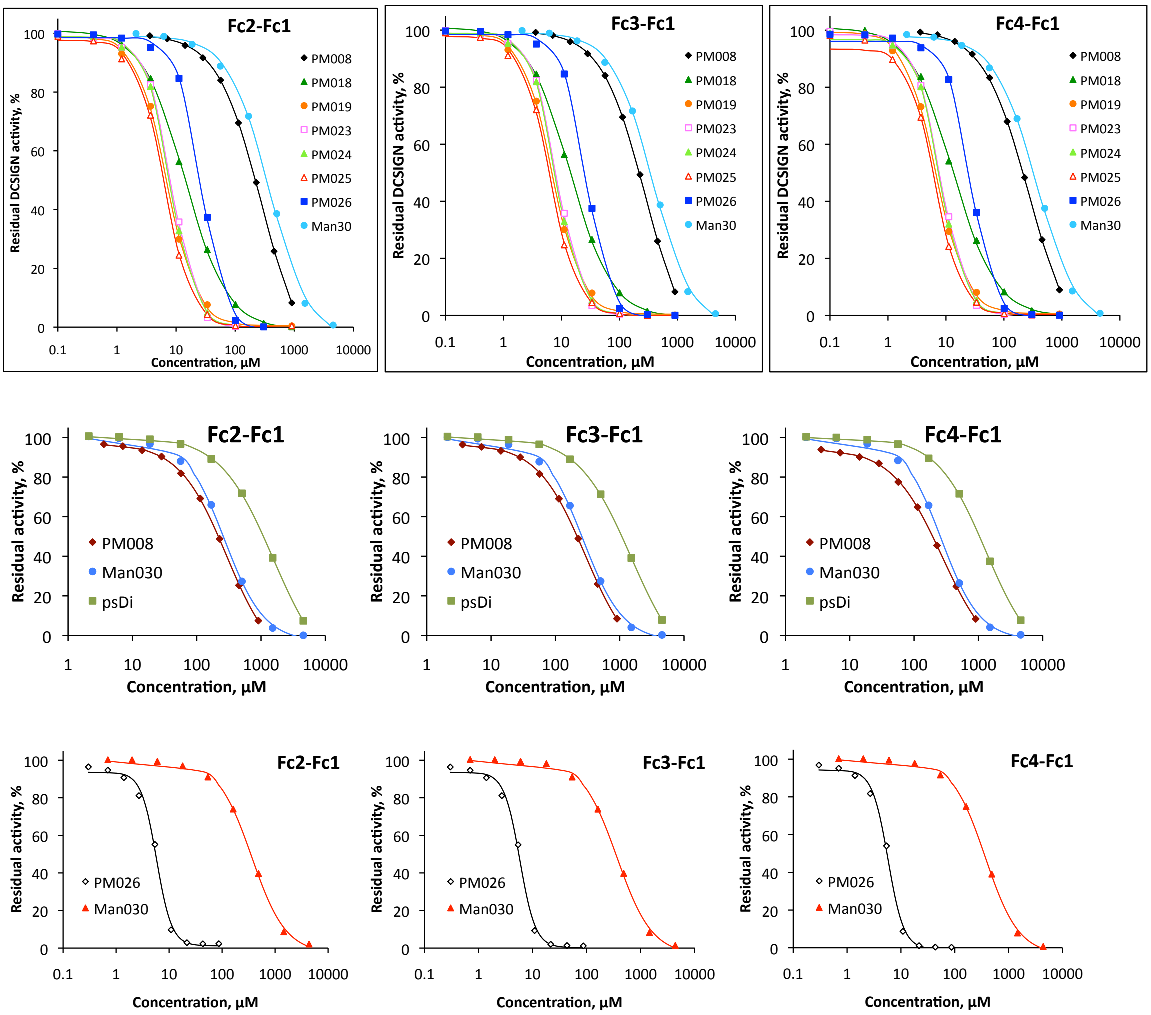








**Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h**

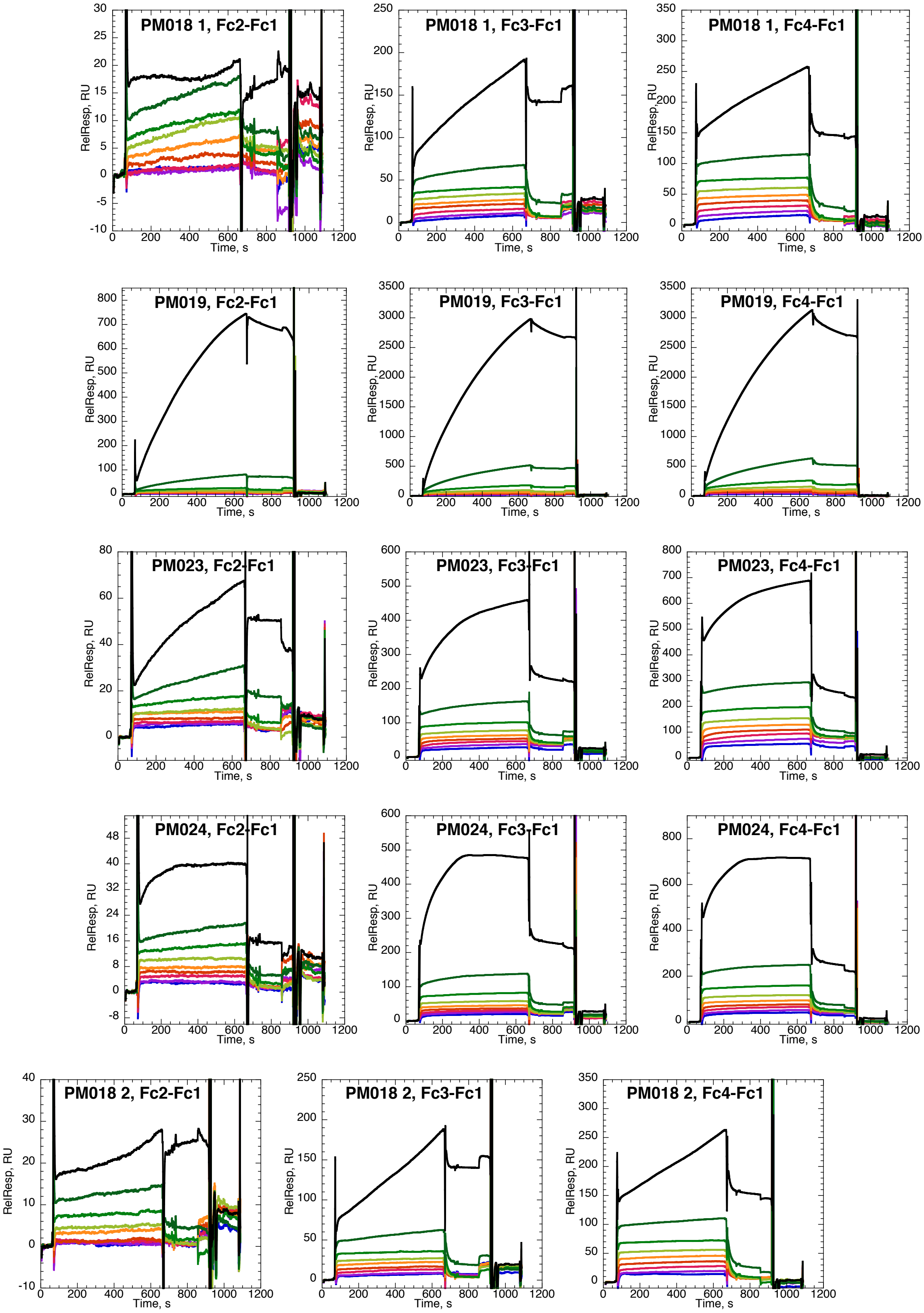


Reference surface corrected sensorgrams and inhibition curves for DC-SIGN inhibition by various dendrimers with mannose, psDi, psTri and 4h

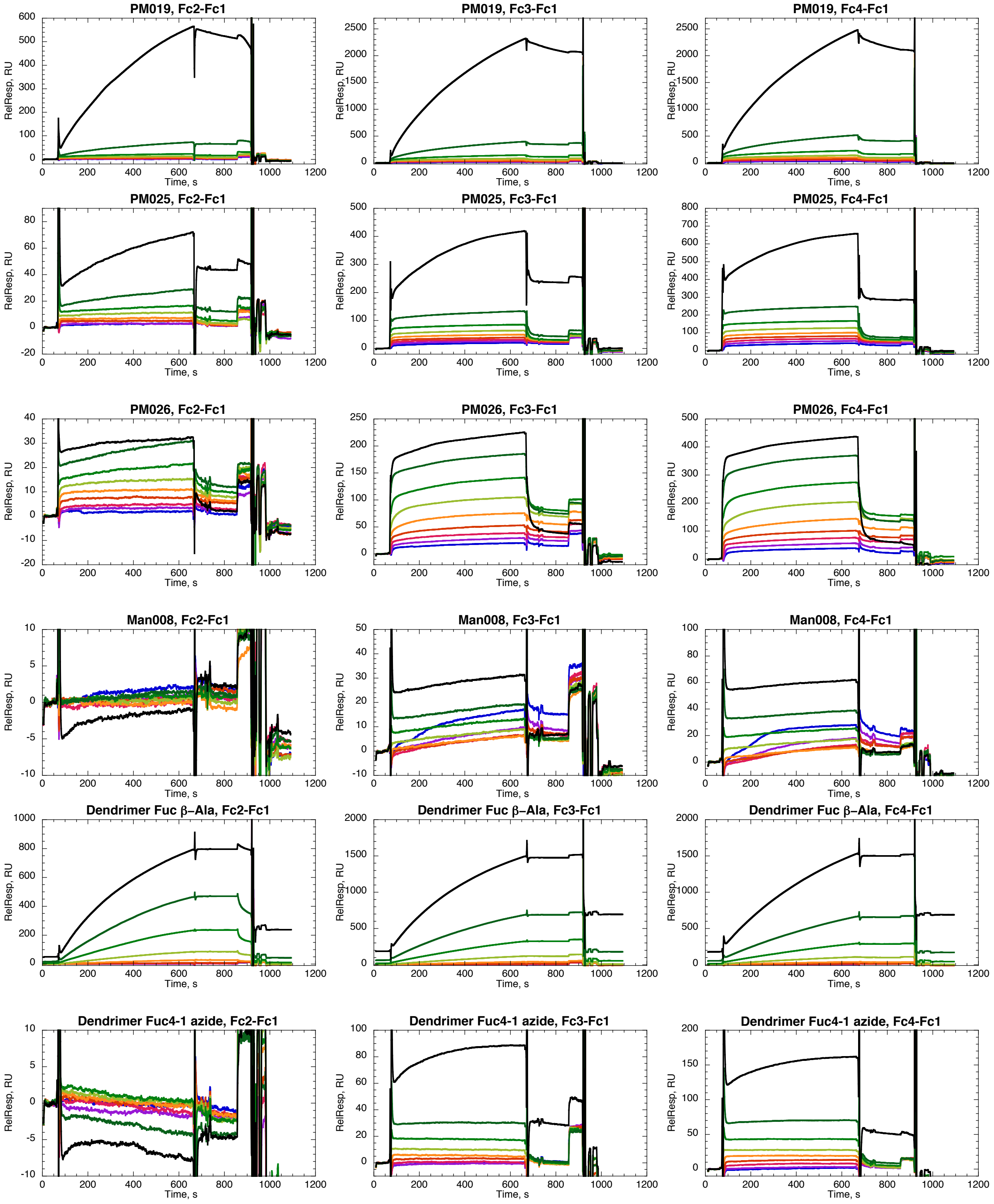
# 25. All of the sensorgrams showing the interaction of 4h-functionalized dendrimers with DC-SIGN surfaces

# Reference surface corrected sensorgrams of 4h-functionalized dendrimers interaction with DC-SIGN surfaces

Fc1 = StrepTactin; Fc2 = DC-SIGN-LD; Fc3 = DC-SIGN-MD; Fc4 = DC-SIGN-HD



Reference surface corrected sensorgrams 4h-functionalized dendrimers interaction with DC-SIGN surfaces

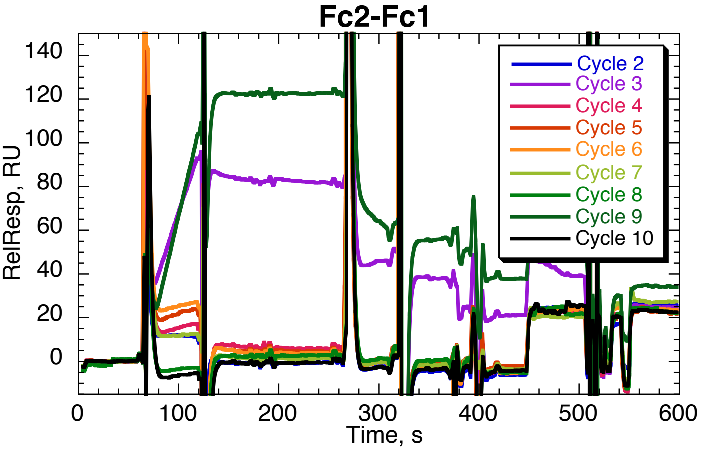


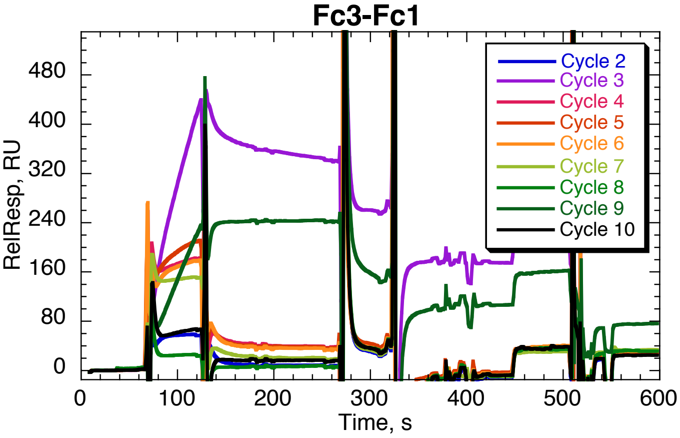
Reference surface corrected sensorgrams 4h-functionalized dendrimers interaction with DC-SIGN surfaces

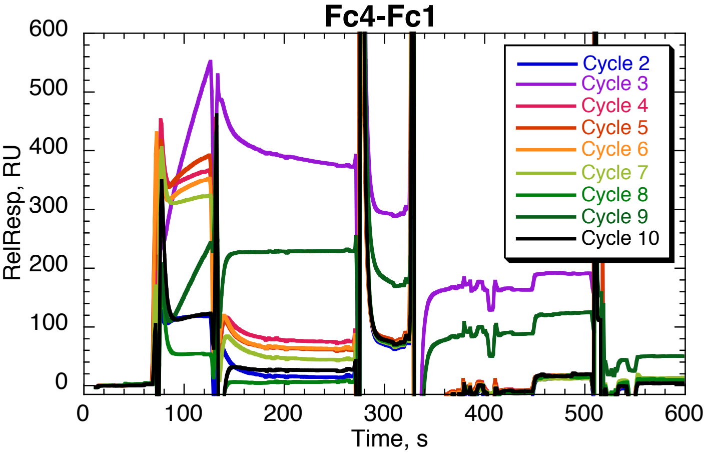
# 26. The sensorgrams of double regeneration test of DC-SIGN surfaces

# Reference surface corrected sensorgrams of double regeneration test of DC-SIGN surfaces

Fc1 = StrepTactin; Fc2 = DC-SIGN-LD; Fc3 = DC-SIGN-MD; Fc4 = DC-SIGN-HD


Fc2-Fc1


Fc3-Fc1


Fc4-Fc1

## *Abstract*

An amazing defense system, the immunity, protects most of the living organisms from harmful pathogens. The innate and acquired immunity components work together to provide efficient protection of humans as well as all other jawed vertebrates.

Dendritic cells, the component of the innate immunity, routinely survey the peripheral tissues, capture and process the invading pathogens, and finally, present the antigens to the T cells to boost the pathogen specific adaptive immune responses. These cells recognize the foreign organisms with the help of multiple pattern recognition receptors (PRRs), which specifically bind molecules on the pathogen surfaces, so-called pathogen associated molecular patterns (PAMPs).

Among PRRs, the C-type lectin receptors (CLRs) have an important role in pathogen recognition and capturing. However, one of these CLRs, DC-SIGN, has been shown to be hijacked by many dangerous pathogens, including HIV, to promote their infection.

This work aims to develop the antagonists of the C-type lectin receptor DC-SIGN, which would be able to block the use of this receptor by pathogens. To achieve that, the strategy of the development of glycomimetic ligands of DC-SIGN and the multivalent presentation of the selected monovalent glycomimics has been employed.

The presented studies were accomplished in collaboration with several chemists groups, who have designed and synthesized different glycomimetic compounds as well as different multivalent platforms. The mannose-based, fucose-based and the C-glycosidic compounds were explored as monovalent DC-SIGN ligands, and multivalent platforms with different valence as well as ligand presentation in space were investigated.

Using surface plasmon resonance (SPR), the activity of the compounds to inhibit DC-SIGN was estimated. Moreover, the compounds were evaluated for their selectivity to DC-SIGN *vs* langerin, another CLR with a protective role from HIV infection. Some of the compounds were structurally characterized by X-ray crystallography and NMR spectrometry studies. The SPR studies of multivalent compounds confirmed the improved activity, but also revealed possible complications.

Overall, these studies allowed to identify two new monovalent leads and to draw perspectives for their further improvement, and suggested the improvement of multivalent presentation platforms.

Keywords: dendritic cells, DC-SIGN, langerin, HIV, glycomimetics, dendrimers, inhibition, selectivity, clustering, surface Plasmon resonance, analytical ultracentrifugation, dynamic light scattering, X-ray crystallography, nuclear magnetic resonance.

## *Résumé*

L'immunité, un système de défense incroyable, protège la plupart des organismes vivants des pathogènes nuisibles. Les composantes innée et acquise de l'immunité travaillent ensemble pour assurer une protection efficace de l'homme ainsi que de tous les autres vertébrés à mâchoires.

Les cellules dendritiques, la composante de l'immunité innée, passent régulièrement en revue les tissus périphériques, capturent et traitent les agents pathogènes envahisseurs et enfin, présentent les antigènes aux lymphocytes T pour stimuler les réponses immunitaires adaptatives spécifiques. Ces cellules reconnaissent les organismes étrangers à l'aide de plusieurs "Pattern Recognition Receptors" (PRRs), qui se lient spécifiquement à des molécules à la surface des agents pathogènes, dits "Pathogen Associated Molecular Patterns (PAMPs).

Parmi les PRR, les récepteurs lectine de type C (CLRs) ont un rôle important dans la reconnaissance et dans la capture des pathogènes. Cependant, DC-SIGN, l'un de ces CLR, peut être détourné par de nombreux agents pathogènes dangereux, y compris le VIH, afin de promouvoir leur infection.

Ce travail vise à développer des antagonistes de DC-SIGN, qui serait en mesure de bloquer l'utilisation de ce récepteur par des agents pathogènes. Pour atteindre cet objectif, la stratégie du développement de ligands glycomimétique de DC-SIGN et la présentation multivalente des glycomimétiques monovalents sélectionnés a été employée.

Les études présentées ont été réalisées en collaboration avec plusieurs groupes de chimistes, qui ont conçu et synthétisé différents composés glycomimétiques ainsi que différentes plates-formes multivalentes. Des ligands monovalents de DC-SIGN basés sue le mannose, le fucose et sur les composés C-glycosidiques ont été explorés et différentes plates-formes de valence avec différents modes de présentation dans l'espace ont été étudiées.

En utilisant la résonance plasmonique de surface (SPR), l'activité des composés à inhiber DC-SIGN a été estimée. De plus, les composés ont été évalués pour leur sélectivité pour DC-SIGN par rapport à la Langerine, un autre CLR ayant un rôle protecteur contre l'infection par le VIH. Certains de ces composés ont été caractérisés par cristallographie aux rayons X et par spectrométrie RMN. Les études de SPR de composés multivalents ont confirmé l'amélioration de l'activité, mais a également révélé des complications possibles.

Dans l'ensemble, ces études ont permis d'identifier les deux meilleurs nouveaux composés et de suggérer des perspectives pour l'amélioration de ces composés et des plates-formes de présentation multivalentes.

Mots-clés: cellules dendritiques, DC-SIGN, langerin, VIH, glycomimetiques, dendrimères, inhibition, selectivité, clustering, résonance plasmonique de surface, ultracentrifugation analytique, diffusion dynamique de la lumière, cristallographie aux rayons X, résonance magnétique nucléaire.